UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in Fisica, Astrofisica e Fisica Applicata

Dipartimento di Fisica, Via Celoria 16, 20133, Milano

Corso di Dottorato in Fisica, Astrofisica e Fisica Applicata

Ciclo XXXI

Advances in quantum parameter estimation and other topics

PhD Thesis

Luigi Seveso

Supervisor:

Prof. Matteo G. A. Paris

Academic Year 2017/2018

Calvin: You can't just turn on creativity like a faucet.
You have to be in the right mood.
Hobbes: What mood is that?
Calvin: Last-minute panic.

– BILL WATTERSON

Acknowledgments

First of all, I sincerely thank my supervisor Prof. Matteo Paris, for having introduced me to the exciting research field of quantum information theory, at a time when I knew very little about about it. I would also like to thank all members, past and present, of the Quantum Technology Lab, with whom I have had the pleasure to interact, for the many valuable discussions and shared moments. I thank Prof. Karol Życzkowski and Dr. Dardo Goyeneche, as well as the members of the Jagellionian University's quantum information theory group, for the beautiful stay in Kraków and for their kind hospitality. Similarly, I thank Prof. Antonio Acín and the many members of his group, for hosting me in Barcelona – and for being probably the nicest group of people on the planet. Last but not least, I thank my family, whose contribution to the present work is difficult to estimate, but is definitely non-negligible.

Contents

Part I

Advances in quantum parameter estimation

Abstract

The first half of this thesis deals with the statistical inference of an unknown parameter from the empirical data generated by a quantum system via repeated measurements. This is the subject matter of *quantum parameter estimation theory*, a research field which has expanded considerably after the first pioneering works by Helstrom [1] and Holevo [2] in the 1970s. In particular, recent progress in quantum parameter estimation has been spurred by the prospect of a second quantum revolution, aimed at harnessing quantum phenomena to enhance the performance of several information-theoretic tasks. Parameter estimation theory is being applied to the design of ultra-precise measurements at the quantum frontier, by providing the theoretical tools for the estimation of e.g. atomic-scale magnetic fields [3], the displacement caused by gravitational waves in gravitational interferometers [4], the temperature of nanoscopic samples [5], as well as other physically important quantities that cannot be measured directly, either in principle or due to some technical obstruction [6].

It has been proved that estimation strategies exploiting quantum effects (such squeezing [7] and entanglement [8]) are able to achieve better precisions than any classical strategy employing the same amount of resources. For instance, the error in the estimation of a classical parameter via N independent measurements scales asymptotically $\sim 1/\sqrt{N}$. In contrast, quantum mechanics allows for a quadratic improvement, the error scaling $\sim 1/N$, which is often referred to as the *Heisenberg limit* [9, 10]. The multiplicative factor in the previous asymptotic relations is the inverse of the *Fisher information* [11]. The Fisher information is the most relevant metric in assessing the performance of quantum measurements (besides having rich statistical and geometrical underpinnings [12–14]). In fact, it provides an ordering criterion: measurements achieving a larger Fisher information are ranked higher by an experimentalist, since they lead to a lower estimation mean square error [15, 16]. In a seminal paper, Braunstein and Caves have maximized the Fisher information over the set of all (regular) measurements [17]. The result of such optimization is called the *quantum Fisher information*. It encodes the ultimate precision limit, beyond which no further improvement is possible, due to the inherent stochasticity of quantum experiments. An estimation strategy which achieves the quantum Fisher

information is thus usually considered optimal.

On the other hand, in this thesis we will show that, for certain kinds of estimation problems, a better precision is achievable than predicted by a quantum Fisher information analysis. This is true if some regularity assumptions about the underlying quantum parametric model are relaxed. In particular, we will investigate the following scenarios:

- The interaction Hamiltonian between the probing system and the measuring apparatus is parameter-dependent.
- The parameter to be estimated appears non-linearly in the probe's Hamiltonian and a measurement of the energy is employed to estimate the parameter.

In such cases, the quantum Fisher information does not completely capture the best possible performance of quantum measurements, and a different approach must be followed.

* * *

This first part of the thesis is based on the articles listed below. In particular, Chapter 2 and Chapter 3 contain entirely original material: the former is based on Refs. [1] and [2], while the latter on Ref. [3]. Chapter 4 is based on unpublished work, currently under review [4].

[1] L. Seveso, M. A. Rossi, and M. G. A. Paris, "Quantum metrology beyond the quantum Cramér-Rao theorem," *Phys. Rev. A*, vol. 95, no. 1, p. 012111, Jan. 2017.

[2] L. Seveso and M. G. A. Paris, "Estimation of general Hamiltonian parameters via controlled energy measurements," *Phys. Rev. A*, vol. 98, no. 3, p. 032114, 2018.

[3] L. Seveso and M. G. A. Paris, "Trade-off between information and disturbance in qubit thermometry," *Phys. Rev. A*, vol. 97, no. 3, p. 032129, Mar. 2018.

[4] L. Seveso, C. Benedetti, and M. G. A. Paris, "The walker speaks its graph: global and nearly-local probing of the tunnelling amplitude in continuous-time quantum walks," arXiv:1809.09211, Sept. 2018.

Chapter 1

Basic concepts

This chapter provides a concise overview of parameter estimation theory, by first lying its foundations in a classical setting, and then generalizing to a quantum scenario, i.e. when the data used for inference is obtained from a quantum system via quantum measurements.

1.1 Probability theory

1.1.1 Probability spaces

The outcome of a random experiment is an *event*. At this stage, an event has no numerical counterpart: it is an abstract subset of a *sample space* Ω.

Example 1.1.1. *Let us consider a random experiment that consists in drawing a single card from the deck* $\Omega = \{A\spadesuit\,,\ K\heartsuit\,,\ Q\diamondsuit\,,\ J\clubsuit\}$. *We define the following event: a face card is drawn. It corresponds to the subset* $\{K\heartsuit\,,\ Q\diamondsuit\,,\ J\clubsuit\} \subset \Omega$.

In general, not every possible subset of Ω constitutes an event. A few desirable requirements are the following: an experiment may have no outcome, so the empty set should be an event; if A is a possible event, then its complement A^c, or logical negation, should also be an event; if A and B are events, then their union $A \cup B$, or logical conjunction, should also be an event. Such requirements naturally lead to the introduction of a σ-algebra structure on the set of events.

Definition 1.1.2. *(σ-**algebra**) A σ-algebra $\mathcal{A}$ on a sample space Ω is a family of subsets of Ω having the following properties:*

(**P1**) *The empty set $\emptyset$ is an element of $\mathcal{A}$;*

(**P2**) *If A is an element of $\mathcal{A}$, then also its relative complement $A^c \in \mathcal{A}$;*

(**P3**) *If* $\{A_i\}_{i=1}^{\infty}$ *is a countable collection of elements of* $\mathcal{A}$*, then also* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

The tuple $(\Omega, \mathcal{A})$ is called a *measurable space* and the elements of $\mathcal{A}$ the *measurable sets*. Making use of properties (**P2**) – (**P3**), one may prove that, if $\{A_i\}_{i=1}^{\infty}$ is a countable collection of elements in $\mathcal{A}$, then also their countable intersection $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$. It follows that if $\mathcal{A}_1$ and $\mathcal{A}_2$ are two different σ-algebras on the same sample space Ω, then their intersection $\mathcal{A}_1 \cap \mathcal{A}_2$ is also a σ-algebra. From this, one may go on to prove that, given any family of sets τ, there is a unique smallest σ-algebra containing τ. A case of major interest is when τ is a topology on Ω, i.e. the tuple (Ω, τ) is a *topological space*.

Definition 1.1.3. ***(topological space)*** *A topological space* (Ω, τ) *is a set* Ω *provided with a topology* τ*, i.e. a family of subsets of* Ω *having the following properties:*

(**P1**) *Both* Ω *and the empty set* $\emptyset$ *are elements of* τ*;*

(**P2**) *If* $\{T_i\}_{i=1}^{\infty}$ *is a countable collection of elements of* τ*, then also* $\bigcup_{i=1}^{\infty} T_i \in \tau$.

(**P3**) *If* $\{T_i\}_{i=1}^{n}$ *is a finite collection of elements of* τ*, then also* $\bigcap_{i=1}^{n} T_i \in \tau$.

The elements of a topology τ on Ω are called the *open sets* of Ω. If Ω is endowed with a topological structure to start with, a σ-algebra structure can be introduced by taking countable unions, countable intersections and relative complements of its open sets. The resulting σ-algebra is called the *Borel* σ*-algebra* $\mathscr{B}(\Omega)$: it is the smallest σ-algebra containing the open sets of Ω.

Once a σ-algebra structure $\mathcal{A}$ has been introduced on Ω, the probability of different events is specified by a *probability measure* μ. The triple $(\Omega, \mathcal{A}, \mu)$ is called a *probability space*.

Definition 1.1.4. ***(probability space)*** *A probability space* $(\Omega, \mathcal{A}, \mu)$ *is a set* Ω *together with a* σ*-algebra structure* $\mathcal{A}$ *and a probability measure* μ*, i.e. a function* $\mu : \mathcal{A} \to [0, 1]$ *having the following properties:*

(**P1**) $\mu(\Omega) = 1$

(**P2**) *If* $\{A_i\}_{i=1}^{\infty}$ *is a countable collection of mutually disjoint elements of* $\mathcal{A}$*, then*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)\ . \tag{1.1.1}$$

With the help of property (**P2**), one may also prove that a probability measure satisfies the following intuitive properties: $\mu(\emptyset) = 0$; if $B \subset A$, then $\mu(B) \leq \mu(A)$; for any two events $A, B \in \mathcal{A}$, $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Remark 1.1.5. *If $(\Omega, \mathcal{A})$ is a measurable space and $\mu : \mathcal{A} \to \overline{\mathbb{R}}_+$ is a function from the measurable sets to the extended (nonnegative) real line, satisfying property* **(P2)**, *the triple $(\Omega, \mathcal{A}, \mu)$ is called a* measure space *and μ a* measure. *A measure μ is said to be* finite *if $\mu(\Omega)$ is a finite real number (it is said σ-*finite *if Ω is countable union of measurable sets having finite measure). A probability space is thus equivalent to a measure space with finite measure, normalized according to property* **(P1)**.

1.1.2 Random variables

While the random outcomes of an experiment are only required to have a σ-algebra structure, a *random variable* is needed in order to associate values to elements of Ω.

Definition 1.1.6. ***(random variable)*** *Given a probability space $(\Omega, \mathcal{A}, \mu)$ and a measurable space $(\mathcal{X}, \mathcal{B})$, a random variable is a function $\mathbb{X} : \Omega \to \mathcal{X}$ having the following property: if $B \in \mathcal{B}$, then the preimage of B under $\mathbb{X}$, i.e. $\mathbb{X}^{-1}(B) = \{\omega \in \Omega : \mathbb{X}(\omega) \in B\}$, is an element of $\mathcal{A}$.*

Example 1.1.7. *A random experiment consists in drawing a single card from the deck $\Omega = \{A\spadesuit, K\heartsuit, Q\diamondsuit, J\clubsuit\}$. The corresponding σ-algebra is taken to be $\mathcal{A} = \mathscr{P}(\Omega)$, the power set of Ω. The probability measure is the constant function $\mu(\omega) = 1/4$, $\forall \omega \in \Omega$. A random variable $\mathbb{X}$ is defined as follows: $\mathbb{X}(A\spadesuit) = 11$, $\mathbb{X}(K\heartsuit) = 4$, $\mathbb{X}(Q\diamondsuit) = 3$, $\mathbb{X}(J\clubsuit) = 2$. Thus, the image set is $\mathcal{X} = \{2, 3, 4, 11\}$. Let us take $\mathcal{B} = \mathscr{P}(\mathcal{X})$. The probability of obtaining, e.g., an even number is the probability of the event $\mathbb{X}^{-1}(\{2, 4\}) = \{K\heartsuit, J\clubsuit\} := E$, from which one obtains $\mu(E) = 1/2$.*

Notice that the measurable space $(\mathcal{X}, \mathcal{B})$ in Def. 1.1.6 can be naturally made into a probability space, by introducing the probability measure ν defined via the relation $\nu(B) = \mu(\mathbb{X}^{-1}(B))$, where B is any measurable set in $\mathcal{B}$. In practice, one often blurs the distinction between the two probability spaces $(\Omega, \mathcal{A}, \mu)$ and $(\mathcal{X}, \mathcal{B}, \nu)$, and says that the outcome of a random experiment is a real value $x \in \mathcal{X}$, rather than an event $A \in \mathcal{A}$. We will also make use of such abuse of terminology when the distinction can be safely ignored.

Remark 1.1.8. *By definition, a measurable function between two measurable spaces is a function such that the preimage of any measurable set is measurable. Therefore, a random variable can equivalently be defined as a measurable function between probability spaces.*

For most random variables of interest, the image set $\mathcal{X}$ is a subset of the real line $\mathbb{R}$. If the subset is finite or countably infinite, the random variable is said to be *discrete*; otherwise, it is a *continuous*. In the following, by a random variable, it will always be

meant a *real* random variable, either discrete or continuous. We will also assume that the σ-algebra $\mathcal{B}$ is fixed by defining first a topological structure on $\mathcal{X}$ (i.e., the subspace topology induced by the real line standard topology) and then a σ-algebra structure, i.e. the Borel algebra of $\mathcal{X}$.

1.1.3 Integration on probability spaces

We now sketch how to define a notion of integration of a random variable with respect to a probability measure. This is done initially only for *simple* random variables.

Definition 1.1.9. ***(simple random variable)*** *A random variable $\mathbb{X} : \Omega \to \mathcal{X}$ is simple if $\mathcal{X}$ is a finite set.*

As a consequence, a simple random variable $\mathbb{X}$ can be written as $\mathbb{X} = \sum_{i=1}^{n} x_i \mathbf{1}_{A_i}$, where $\{x_i\}_{i=1}^{n}$ are real numbers, $\{A_i\}_{i=1}^{n}$ are elements of $\mathcal{A}$ and $\mathbf{1}_{A_i}$ is the characteristic function of A_i, i.e.

$$\mathbf{1}_{A_i}(\omega) := \begin{cases} 1 & \text{if } \omega \in A_i \\ 0 & \text{if } \omega \notin A_i \end{cases} . \tag{1.1.2}$$

This representation is, in general, non-unique.

If $\mathbb{X}$ is a simple random variable, its *expectation* is defined as

$$E(\mathbb{X}) := \sum_{i=1}^{n} x_i \, \mu(A_i) \, , \tag{1.1.3}$$

which can also be denoted by $\int_\Omega \mathbb{X} \, d\mu$. It can be proven that $E(\mathbb{X})$ does not depend on the representation.

The next step is to define the expectation of nonnegative random variables. A random variable $\mathbb{X}$ is nonnegative if it takes only nonnegative values. Two random variables satisfy $\mathbb{X} \geq \mathbb{Y}$ if their difference $\mathbb{X} - \mathbb{Y}$ is nonnegative. One defines:

$$E(\mathbb{X}) := \sup \left(E(\mathbb{Y}), \ \mathbb{Y} \text{ a simple random variable with } 0 \leq \mathbb{Y} \leq \mathbb{X} \right) . \tag{1.1.4}$$

Let us remark that, by definition, $E(\mathbb{X}) \geq 0$ and that $E(\mathbb{X})$ always exists, but might be equal to $+\infty$, even if $\mathbb{X}$ is everywhere finite.

The final step is to consider an arbitrary random variable $\mathbb{X}$. Let $\mathbb{X}^{(+)} = \max(\mathbb{X}, 0)$ and $\mathbb{X}^{(-)} = -\min(\mathbb{X}, 0)$. Thus, $\mathbb{X} = \mathbb{X}^{(+)} - \mathbb{X}^{(-)}$, where $\mathbb{X}^{(+)}$ and $\mathbb{X}^{(-)}$ are positive random variables. Then, one defines

$$E(\mathbb{X}) := E(\mathbb{X}^{(+)}) - E(\mathbb{X}^{(-)}) \, . \tag{1.1.5}$$

A random variable $\mathbb{X}$ is *integrable* if both $E(\mathbb{X}^{(+)})$ and $E(\mathbb{X}^{(-)})$ are finite; then, its expectation is given by Eq. (1.1.5). It is easy to check that the set of integrable random variables on a probability space $(\Omega, \mathcal{A}, \mu)$ is a vector space, denoted by $\mathcal{L}^1$, with expectation acting as a linear map on it. Notice that, if two random variables satisfy $\mathbb{X} = \mathbb{Y}$ almost surely, i.e. $\mu(\{\omega \in \Omega : \mathbb{X}(\omega) = \mathbb{Y}(\omega)\}) = 1$, then $E(\mathbb{X}) = E(\mathbb{Y})$. Therefore, equality almost surely is an equivalence relation, denoted by $\sim$, and equivalent random variables have the same expectation. To remove this redundancy, one introduces the quotient space $L^1 := \mathcal{L}^1/\sim$, whose elements are equivalence classes of almost surely equal random variables. However, by abuse of terminology, one usually still refers to elements of L^1 as random variables. In a similar way, for $1 \leq p < \infty$, one defines $\mathcal{L}^p$ as the vector space of random variables such that $|\mathbb{X}|^p \in \mathcal{L}^1$, where $|\mathbb{X}| := \mathbb{X}^{(+)} + \mathbb{X}^{(-)}$. By taking equivalence classes with respect to $\sim$, one then obtains the spaces L^p of p-integrable random variables. In the following, we will only need the spaces L^1 and L^2.

If two random variables are square-integrable, they satisfy the following inequality.

Proposition 1.1.10. ***(Cauchy-Schwarz inequality)*** *If $\mathbb{X}, \mathbb{Y} \in L^2$, then $\mathbb{X} \cdot \mathbb{Y} \in L^1$ and*

$$|E(\mathbb{X} \cdot \mathbb{Y})| \leq \sqrt{E(\mathbb{X}^2)E(\mathbb{Y}^2)}\,. \tag{1.1.6}$$

Given square-integrable random variables $\{\mathbb{X}_i\}_{i=1}^n$ with $\mathbb{X}_i \in L^2$, one defines their covariance matrix as follows:

Definition 1.1.11. ***(covariance matrix)*** *Let $\{\mathbb{X}_i\}_{i=1}^n$ be a collection of square-integrable random variables in L^2. Their covariance matrix is the matrix with entries:*

$$\mathrm{Cov}(\mathbb{X}_i, \mathbb{X}_j) := E\left[(\mathbb{X}_i - E(\mathbb{X}_i))(\mathbb{X}_j - E(\mathbb{X}_j))\right]\,. \tag{1.1.7}$$

Remark 1.1.12. *In particular, the diagonal elements of a covariance matrix are the variances* $\mathrm{Var}(\mathbb{X}_i) := E[(\mathbb{X}_i - E(\mathbb{X}_i))^2]$.

As a concluding remark, since the product of two measurable functions is a measurable function and the characteristic function $\mathbf{1}_A$ of a set A is measurable if and only if A is measurable, the integral of a random variable on any measurable set $A \in \mathcal{A}$ is well-defined: one has to take the expectation of the product $\mathbf{1}_A \cdot \mathbb{X}$, i.e. $\int_A \mathbb{X}\, d\mu = \int_\Omega \mathbf{1}_A \cdot \mathbb{X}\, d\mu$.

1.1.4 Probability densities

We now introduce the concept of *probability density* of a random variable. As discussed before, a random variable $\mathbb{X}$ on a probability space $(\Omega, \mathcal{A}, \mu)$ gives rise to a probability space $(\mathcal{X}, \mathcal{B}, \nu)$, where $\mathcal{X} \subseteq \mathbb{R}$, $\mathcal{B}$ is the Borel algebra generated by the natural topology

of $\mathcal{X}$ and ν is a probability measure. Notice that there are already two natural notions of a measure on $\mathcal{X}$: the *Lebesgue measure* (if $\mathcal{X}$ is an uncountable subset of $\mathbb{R}$) and the *counting measure* (if $\mathcal{X}$ is a countable subset). The measure ν can always be expressed in terms of either the Lebesgue measure or the counting measure, provided it satisfies a technical assumption, which is contained in the following definition.

Definition 1.1.13. ***(absolutely continuous measures)*** *If ν and ν' are any two measures with the same σ-algebra $\mathcal{B}$ of subsets of $\mathcal{X}$, then ν is said to be absolutely continuous with respect to ν', denoted $\nu \ll \nu'$, if $\nu(B) = 0$ for any $B \in \mathcal{B}$ such that $\nu'(B) = 0$.*

We henceforth assume that, if $\mathbb{X}$ is a continuous random variable, ν is absolutely continuous with respect to the Lebesgue measure, i.e. it agrees with the Lebesgue measure on any set with Lebesgue measure zero. If instead $\mathbb{X}$ is discrete, every probability measure ν is already absolutely continuous with respect to the counting measure (since the counting measure vanishes only on the empty set and $\nu(\emptyset) = 0$ always). The following theorem applies to any two absolutely continuous measures.

Theorem 1.1.14. ***(Radon-Nikodym)*** *Let ν and ν' be two σ-finite measures on the same measurable space $(\mathcal{X}, \mathcal{B})$ such that $\nu \ll \nu'$. Then:*

(T1) *There exists a measurable function $h : \mathcal{X} \to \mathbb{R}_+$ such that, for all $B \in \mathcal{B}$,*

$$\nu(B) = \int_B h \, d\nu' \, . \tag{1.1.8}$$

(T2) *Such a function h is almost unique: any two functions satisfying Eq. (1.1.8) can differ only on sets of measure zero with respect to ν'.*

(T3) *h is integrable with respect to ν' if and only ν is a finite measure.*

The function h is called the Radon-Nikodym derivative of ν with respect to ν', denoted $h = d\nu/d\nu'$. It allows to convert between the two measures by means of the symbolic identity $d\nu = h \, d\nu'$.

Definition 1.1.15. *Let $\mathbb{X}$ be a random variable with probability space $(\mathcal{X}, \mathcal{B}, \nu)$.*

(D1) *If $\mathbb{X}$ is continuous, its probability density function (p.d.f.) is the Radon-Nikodym derivative of ν with respect to the Lebesgue measure, i.e. $p = d\nu/dx$.*

(D2) *If $\mathbb{X}$ is discrete, its probability mass function (p.m.f.) is the Radon-Nikodym derivative of ν with respect to the counting measure, i.e. $p = d\nu/d\#$.*

Knowledge of the p.d.f. p (*resp.*, the p.m.f.) fully characterizes a random experiment whose outcomes are described by a random variable $\mathbb{X}$, since the probability of any event B can be obtained by integrating p on B with respect to the Lebesgue measure (*resp.*, the counting measure) by means of Eq. (1.1.8).

1.2 Classical parameter estimation

1.2.1 Classical statistical models

Consider a random experiment whose outcomes are described by a random variable $\mathbb{X}$, with probability space $(\mathcal{X}, \mathcal{B}, \nu)$ and probability density p. The task is to reconstruct p, which is referred to as the *true* probability density, starting from N independent sample points or observations of $\mathbb{X}$ (in the following, a sample point is denoted by a lowercase letter, e.g. $x \in \mathcal{X}$, whereas a sample of N observations by a boldface letter, e.g. $\boldsymbol{x} \in \mathcal{X}^{\times N}$).

There are many ways to approach the problem of learning p but, if the functional form of p is already known, or can be guessed with reasonable accuracy, a *parametric approach* is quite natural. The true probability density p is assumed to belong to a parametric family of probability densities $\{p_\theta\}_{\theta\in\Theta}$, where $\Theta \subset \mathbb{R}^m$ is the *parameter space*. It is also assumed that there exists a suitable choice $\theta^* \in \Theta$ such that $p_{\theta^*} = p$. In this way, all lack of knowledge about p is reduced to lack of knowledge about the *true* parameter θ^* – a considerable simplification of the problem.

Definition 1.2.1. ***(classical statistical model)*** *A classical statistical model S is a family of probability densities on $\mathcal{X}$ parametrized by m real parameters $\theta \in \Theta \subset \mathbb{R}^m$:*

$$S = \{p_\theta \,:\, \theta = (\theta^1, \theta^2, \dots, \theta^m) \in \Theta\} \,, \tag{1.2.1}$$

where the parametrization map $\theta \to p_\theta$ is injective, the support $\mathcal{X}$ is parameter-independent and p_θ can be differentiated as many times as needed with respect to the parameters, i.e. all possible derivatives $\partial_1^{k_1} \dots \partial_m^{k_m} p_\theta$ (where ∂_i is short for ∂_{θ^i}) exist.

Remark 1.2.2. *If $\mathcal{X}$ is countable, then p_θ is a p.m.f. normalized such that*

$$\sum_{x\in\mathcal{X}} p_{x,\theta} = 1 \,, \qquad \forall\theta \in \Theta \,. \tag{1.2.2}$$

If $\mathcal{X}$ is uncountable, then p_θ is a p.d.f. normalized such that

$$\int_{\mathcal{X}} p_\theta(x)\, dx = 1 \,, \qquad \forall\theta \in \Theta \,. \tag{1.2.3}$$

In the following, we will employ the notation for continuous variables; for discrete variables, one should replace the Lebesgue measure dx by the counting measure $d\#$.

1.2.2 Fisher information metric

Given a statistical model $S = \{p_\theta\}_{\theta\in\Theta}$, the map $\varphi : S \to \mathbb{R}^m$ defined by $\varphi(p_\theta) = \theta$ can be considered as providing a coordinate system for S. If ψ is a smooth reparametrization

which maps $\Theta \to \Theta'$, nothing prevents using $\psi(\theta) = \theta'$ as the new parameters, so that the model is rewritten as $S = \{p_{\psi^{-1}(\theta')} : \theta' \in \Theta'\}$. This defines the structure of a differentiable manifold on S, with different parametrizations representing different coordinate systems. Moreover, a Riemannian metric can be defined on the statistical manifold S as follows.

Definition 1.2.3. ***(Fisher information)*** *Let S be a statistical model. Given a point θ, the (classical) Fisher information matrix $\mathcal{F}_C(\theta)$ at that point is the matrix having $(i,j)^{th}$ element*

$$[\mathcal{F}_C(\theta)]_{ij} = \int_{\mathcal{X}} dx\, p_\theta(x)\, \partial_i \log p_\theta(x)\, \partial_j \log p_\theta(x)\ . \tag{1.2.4}$$

When $m = 1$ and only one parameter $\theta = \theta^1$ is present, $\mathcal{F}_C(\theta)$ is referred to as the Fisher information (FI). For $m > 1$, $\mathcal{F}_C(\theta)$ is indeed a symmetric $m \times m$ real matrix. It is always positive semi-definite and, in particular, positive-definite if and only if for every $\theta \in \Theta$ the elements of the set $\{\partial_1 p_\theta, \ldots, \partial_m p_\theta\}$ are linearly independent. Moreover, $\mathcal{F}_C(\theta)$ has the correct transformation properties of a $(0,2)$ tensor under reparametrizations [12]. It follows that $\mathcal{F}_C(\theta)$ provides a Riemannian metric on S.

There is a precise sense in which the Fisher geometry, i.e. the geometry implied by the Fisher information metric, is the only possible geometry on a statistical manifold. To explain this, we introduce the notion of a *statistic*.

Definition 1.2.4. ***(statistic)*** *Given a random variable $\mathbb{X}$ and a function $T : \mathcal{X} \to \mathcal{Y}$ which maps $x \to y = T(x)$, a statistic based on T is the random variable $\mathbb{Y} = T(\mathbb{X})$.*

If $\mathbb{X}$ is associated with a statistical model $S = \{p_\theta\}_{\theta\in\Theta}$, then a statistic T gives rise to a model $S_T = \{q_\theta\}_{\theta\in\Theta}$ associated with $\mathbb{Y} = T(\mathbb{X})$. A statistic is said to be *sufficient* if the two models are related as follows: $p_\theta(x) = h(x)\, q_\theta(y(x))$, $\forall x \in \mathcal{X}$, i.e. all dependence on the parameter θ is contained in q_θ. Intuitively, a sufficient statistic leads to no loss of information about θ. Notice that a one-to-one function is always a sufficient statistic, but there exist sufficient statistics which are not one-to-one functions. We now have the following theorem.

Theorem 1.2.5. *The Fisher information matrix $\mathcal{F}_C^{(T)}$ of the statistical model S_T induced by a statistic T satisfies the monotonicity property $\mathcal{F}_C^{(T)} \leq \mathcal{F}_C$ (where $\mathcal{F}_C$ is the Fisher information matrix of the original model S). The previous inequality must be interpreted in the sense that the difference $\mathcal{F}_C^{(T)} - \mathcal{F}_C$ is a positive semi-definite matrix. Equality holds if and only if T is a sufficient statistic.*

A Riemannian metric satisfying the monotonicity property is said to be a monotone metric. Monotone metrics are the natural metrics on classical statistical models: they

reflect the fact that the points of the manifold are probability distributions and distances between points can only contract under any information processing. In this regard, the following theorem [18–20] singles out the Fisher information metric as the only natural metric on statistical manifolds.

Theorem 1.2.6. ***(Chentsov)*** *The Fisher information metric $\mathcal{F}_C$ is the essentially unique monotone Riemannian metric on a classical statistical model, in the sense that any other such metric is a scalar multiple of $\mathcal{F}_C$.*

Chentsov's theorem establishes a first link between the statistical properties of parametric models and the geometry defined by the Fisher metric. A further link comes from the (classical) Cramér-Rao theorem, which we now introduce.

1.2.3 Cramér-Rao theorem

Let us return to the problem of estimating the true parameter θ^* from a sample $\boldsymbol{x} \in \mathcal{X}^{\times N}$. To this end, we introduce the following definition.

Definition 1.2.7. ***(estimator)*** *An estimator $\hat{\theta}^{(N)} : \mathcal{X}^{\times N} \to \Theta$ is a random variable from the sample space $\mathcal{X}^{\times N}$ to the parameter space Θ. In particular:*

(D1) *An unbiased estimator is an estimator satisfying $E_\theta(\hat{\theta}^{(N)}) = \theta$, $\forall \theta \in \Theta$, where $E_\theta(\cdot)$ denotes expectation with respect to p_θ, i.e.*

$$E_\theta(\hat{\theta}^{(N)}) = \int_{\mathcal{X}^{\times N}} dx_1 \dots dx_N \, p_\theta(x_1) \dots p_\theta(x_N) \, \hat{\theta}^{(N)}(\boldsymbol{x}) \, . \tag{1.2.5}$$

(D2) *A locally unbiased estimator is an estimator which is unbiased at $\theta = \theta^*$, i.e.*

$$E_{\theta^*}(\hat{\theta}^{(N)}) = \theta^* \, , \tag{1.2.6}$$

and, moreover, satisfies

$$\partial_i E_\theta[(\hat{\theta}^{(N)})^j]\Big|_{\theta=\theta^*} = \delta_i^j \, . \tag{1.2.7}$$

(D3) *An asymptotically unbiased estimator is an estimator such that*

$$\lim_{N\to\infty} E_\theta(\hat{\theta}^{(N)}) = \theta \, . \tag{1.2.8}$$

A typical (classical) estimation protocol consists in collecting a sample $\boldsymbol{x} \in \mathcal{X}^{\times N}$ and processing it via an estimator $\hat{\theta}^{(N)}$, thus yielding an estimate $\hat{\theta}^{(N)}(\boldsymbol{x})$. If the estimator is unbiased, the estimate will fluctuate around the true value θ^* over many independent repetitions of the protocol. To quantify the performance of an estimator, it is usual to take as a figure of merit its mean square error:

$$[\mathrm{MSE}(\hat{\theta}^{(N)})]_{ij} := E_\theta([(\hat{\theta}^{(N)})^i - \theta][(\hat{\theta}^{(N)})^j - \theta]) \, . \tag{1.2.9}$$

Estimators with a smaller MSE are said to perform better than estimators with a larger one. Notice that for unbiased estimators, the MSE matrix coincides with the covariance matrix $[\mathrm{Cov}(\hat{\theta}^{(N)})]_{ij}$. The following theorem provides a lower bound to the covariance matrix of unbiased estimators [15, 16].

Theorem 1.2.8. ***(Cramér-Rao)*** *If S is a classical statistical model and $\hat{\theta}^{(N)}$ an unbiased estimator, its covariance matrix is bounded from below as follows:*

$$\mathrm{Cov}(\hat{\theta}^{(N)}) \geq \frac{1}{N} \left[\mathcal{F}_C(\theta)\right]^{-1}, \tag{1.2.10}$$

where $\mathcal{F}_C$ is the Fisher information matrix of S.

The proof of Thm. 1.2.8 amounts to an application of the Cauchy-Schwarz inequality of Prop. 1.1.10.

Remark 1.2.9. *Under the weaker assumption that $\hat{\theta}^{(N)}$ is only locally unbiased, inequality* (1.2.10) *still holds, but only at $\theta = \theta^*$.*

Notice that the Cramér-Rao theorem only provides a lower-bound: it does not guarantee that an estimator achieving the bound actually exists. If such an estimator exists, it is said to be *efficient*. An efficient estimator is the best unbiased estimator, since it minimizes the MSE among all unbiased estimators. Unfortunately, efficient estimators exist only under special circumstances (when the statistical model is of the exponential type and the parameters are its natural parameters, see e.g. Ref. [21]). Finding the best unbiased estimator becomes then a non-trivial task.

The situation improves in the asymptotic limit of a large number of samples. Let us remark that unbiasedness is a strong condition: for some models there exists no such estimator. A far more reasonable condition is that of consistency. A *consistent* estimator is such that, in the limit $N \to \infty$, its probability density becomes concentrated around θ, i.e. $\forall \epsilon > 0$ and $\forall \theta \in \Theta$, $\lim_{N\to\infty} \mathrm{Pr}_\theta(|\hat{\theta}^{(N)} - \theta| > \epsilon) = 0$, where $\mathrm{Pr}_\theta(\cdot)$ denotes the probability of an event computed with respect to p_θ. Under mild conditions (e.g. that $\mathrm{Cov}(\hat{\theta}^{(N)})$ is uniformly bounded with respect to the number of samples N), one can prove that a consistent estimator is asymptotically unbiased, i.e. $\lim_{N\to\infty} E_\theta(\hat{\theta}^{(N)}) = \theta$, and satisfies $\lim_{N\to\infty} \partial_i E_\theta[(\hat{\theta}^{(N)})^j] = \delta_i^j$. With the help of the last two properties, one can prove the following asymptotic version of the Cramér-Rao theorem:

$$\lim_{N\to\infty} N \cdot \mathrm{Cov}(\hat{\theta}^{(N)}) \geq \left[\mathcal{F}_C(\theta)\right]^{-1}. \tag{1.2.11}$$

A consistent estimator achieving equality is said to be *asymptotically efficient*. Remarkably, asymptotically efficient estimators always exist, e.g. the maximum-likelihood estimator and Bayes estimators are asymptotically efficient [21]. In conclusion, at the classical

level and in the asymptotic regime $N \gg 1$, the optimal protocol consists in collecting a sample and processing it via an asymptotically efficient estimator; the asymptotic optimal rate at which distinct values of the parameters can be distinguished is given by the inverse Fisher information.

1.3 Quantum measurement theory

The outcomes of a quantum experiment are probabilistic. This means that there must exist a suitable probability measure $\nu_\rho^{(\mathcal{M})}$ such that, if $(\mathcal{X}, \mathcal{B})$ is the measurable space of outcomes (where $\mathcal{X} \subseteq \mathbb{R}$ is the sample space and $\mathcal{B}$ the σ-algebra induced by the natural topology of $\mathcal{X}$), then the probability of any event $B \in \mathcal{B}$ is $\nu_\rho^{(\mathcal{M})}(B)$. The main difference compared with the classical case is that $\nu_\rho^{(\mathcal{M})}$ is not arbitrary, but is a specific function of both the state of the system ρ and the measurement $\mathcal{M}$. The mapping $(\rho, \mathcal{M}) \to \nu_\rho^{(\mathcal{M})}$ is given by Born's rule.

We will deal exclusively with *finite-dimensional* quantum systems, with Hilbert space $\mathcal{H} = \mathbb{C}^d$. A state is a *density matrix* $\rho \in \mathsf{Her}_d^+(\mathbb{C})$, i.e. an Hermitian positive semi-definite matrix, usually normalized such that $\mathrm{tr}(\rho) = 1$. The set $\mathcal{S}(\mathcal{H})$ of all possible density operators on $\mathcal{H}$ is a convex set. Its extremal elements are the pure states $|\psi\rangle\langle\psi|$, with $|\psi\rangle \in \mathcal{H}$ such that $\langle\psi|\psi\rangle = 1$. The *Hamiltonian* matrix $H \in \mathsf{Her}_d(\mathbb{C})$ completely determines the dynamics of the system (assuming it is isolated from any external environment). That is, if $U_t := \exp(-itH)$ is the matrix exponential of H and ρ_0 is the state at time $t = 0$, then the state of the system at any subsequent time t is $\rho_t := U_t \rho_0 U_t^\dagger$.

A measurement on a quantum system can be described at three different levels of details. We begin with the first level, which is the more coarse-grained of the three.

(L1) POVM description: At this level, a measurement $\mathcal{M}$ is a mapping that associates to any event $B \in \mathcal{B}$ a positive semi-definite operator $\mathcal{M}(B) \in \mathsf{Her}_d^+(\mathbb{C})$. A few natural requirements are that $\mathcal{M}(\emptyset) = \mathbb{0}_d$; $\mathcal{M}(\mathcal{X}) = \mathbb{1}_d$; if $\{B_i\}_{i=1}^n$ are mutually disjoint measurable sets such that $\bigcup_{i=1}^n B_i := B \in \mathcal{B}$, then $\mathcal{M}(B) = \sum_{i=1}^\infty \mathcal{M}(B_i)$. These properties imply that $\mathcal{M}$ is a positive-operator valued (probability) measure (POVM) on $(\mathcal{X}, \mathcal{B})$. In particular, they imply that if $\{B_i\}_{i=1}^n$ are mutually disjoint and $\bigcup_{i=1}^n B_i = \mathcal{X}$, then $\sum_{i=1}^n \mathcal{M}(B_i) = \mathbb{I}_d$. Apart for this normalization condition and for being non-negative, the operators $\mathcal{M}(B)$ are completely arbitrary.

The link between a measurement $\mathcal{M}$ and the probability measure $\nu_\rho^{(\mathcal{M})}$ is provided by Born's rule, i.e.

$$\nu_\rho^{(\mathcal{M})}(B) = \mathrm{tr}(\rho\, \mathcal{M}(B)) \,. \tag{1.3.1}$$

It can be proven that Born's rule is actually the unique possibility under a few

reasonable assumptions [22]. Eq. (1.3.1) completely determines the statistics of any quantum experiment.

If $\mathcal{X}$ is a countable sample space, one defines the *probability operators* $\{\Pi_x\}_{x\in\mathcal{X}}$ of a given measurement as follows: $\Pi_x := \mathcal{M}(x)$. The probability operators are sufficient to compute the probability of any other event. A special case is when each Π_x is a projector P_x, i.e. $P_x^2 = P_x$. One can then associate to the measurement an Hermitian operator $X = \sum_{x\in\mathcal{X}} xP_x$, also called an *observable*. Vice versa, every Hermitian operator gives rise to a projective measurement via its eigendecomposition. An example is the Hamiltonian: a projective measurement over its eigenstates $\{|\xi_j\rangle\}_{j=0}^{d-1}$ is called an energy measurement.

(L2) Instrument description: A POVM description assigns probabilities to measurement outcomes, but does not specify how the state of the system is modified as a result of the measurement. However, quantum measurements can have dynamical effects: if the measurement is non-destructive, the state of the system is updated depending on the outcome. This requires introducing an *instrument*.

Formally, an instrument $\mathcal{I}$ is a mapping $\mathcal{B} \to \mathcal{T}(\mathcal{H})$, where $\mathcal{T}(\mathcal{H})$ denotes the set of bona fide quantum operations on the system (i.e. completely-positive, trace preserving maps). If $B \in \mathcal{B}$ is the observed event, then the state of the system after the measurement is, by definition, $\mathcal{I}_B(\rho)$. Assuming $\mathcal{X}$ is countable, it is enough to consider the set $\{\mathcal{I}_x\}_{x\in\mathcal{X}}$. It can be proven [23] that the most general form for $\mathcal{I}_x$ is as follows,

$$\mathcal{I}_x(\rho) = \frac{\sum_{j=1}^{n} M_x^{(j)} \rho\, M_x^{(j)\dagger}}{\Pr(x)} \,, \tag{1.3.2}$$

where the operators $M_x^{(j)}$ are called *measurement operators* and $\Pr(x) := \mathrm{tr}(\rho\,\Pi_x)$. Since the post-measurement state $\mathcal{I}_x(\rho)$ must be normalized, one has the identification

$$\Pi_x = \sum_{j=1}^{n} M_x^{(j)\dagger} M_x^{(j)} \,. \tag{1.3.3}$$

In particular, if $n = 1$, $\forall x \in \mathcal{X}$, the measurement is said to be *fine-grained*. Notice that, in general, many different instruments correspond to the same positive-operator valued measure. This is true even for fine-grained measurements, since the condition $\Pi_x = M_x^\dagger M_x$ is solved by $M_x = U_x\sqrt{\Pi_x}$, where $\sqrt{\Pi_x}$ is the principal square-root of Π_x but U_x is an arbitrary unitary operator. If the measurement is fine-grained and $U_x = \mathbb{1}_d$, $\forall x \in \mathcal{X}$, the measurement is said to be *bare* and the corresponding instrument is known as the *Lüders instrument*.

(L3) Measurement model description: This is the most detailed level of description of a measurement and is obtained by explicitly modelling the interaction between the system and the measuring apparatus. It is assumed that the system is coupled to an ancillary system with Hilbert space $\mathcal{H}_A$; the ancilla is prepared in an initial state $\eta \in \mathcal{S}(\mathcal{H}_A)$; the two systems evolve together for an interaction time t_{int} via a quantum channel $\mathcal{E}^{(t_{int})} \in \mathcal{T}(\mathcal{H} \otimes \mathcal{H}_A)$; finally, an observable $X = \sum_{x \in \mathcal{X}} x P_x$ on $\mathcal{H}_A$ is measured, producing an outcome $x \in \mathcal{X}$. A *measurement model* is therefore a quadruple $(\mathcal{H}_A, \eta, \mathcal{E}^{(t_{int})}, X)$. It gives rise to a positive-operator valued measure via the relation:

$$\mathrm{tr}(\rho\, \Pi_x) = \mathrm{tr}[\mathcal{E}^{(t_{int})}(\rho \otimes \eta)\, \mathbb{I}_d \otimes P_x] \,. \tag{1.3.4}$$

Moreover, it defines an instrument via

$$\mathcal{I}_x(\rho) = \frac{\mathrm{tr}_A[\mathcal{E}^{(t_{int})}(\rho \otimes \eta)\, \mathbb{I}_d \otimes P_x]}{\Pr(x)} \,, \tag{1.3.5}$$

where $\mathrm{tr}_A(\cdot)$ denotes the partial trace over the ancilla's degrees of freedom.

Clearly, many measurement models can lead to the same instrument. In fact, Ozawa's theorem [24] states that one can recover all possible instruments just by considering measurement models $(\mathcal{H}_A, \eta, \mathcal{E}^{(t_{int})}, X)$ where η is pure, $\mathcal{E}^{(t_{int})}$ is a unitary channel and each P_x is rank-1. More precisely, let H_A be the free Hamiltonian of the ancillary system and H_I the interaction Hamiltonian between the system and the apparatus. Let $\eta = |\phi\rangle\langle\phi|$ be the initial preparation of the ancilla. Then, the unitary channel $\mathcal{U}^{(t)}$ generated by the total Hamiltonian $H_T = H + H_A + H_I$ acts as follows:

$$\rho \otimes |\phi\rangle\langle\phi| \to \mathcal{U}^{(t)}(\rho \otimes |\phi\rangle\langle\phi|) := U_t\, \rho \otimes |\phi\rangle\langle\phi|\, U_t^\dagger \,, \qquad U_t := e^{-itH_T} \,. \tag{1.3.6}$$

From conditions (1.3.4) and (1.3.5), one may prove that the measurement operators M_x and probability operators Π_x take the following form, respectively,

$$M_x = \langle x|U_{t_{int}}|\phi\rangle \;, \qquad \Pi_x = \langle\phi|U_{t_{int}}^\dagger\, \mathbb{I}_d \otimes P_x\, U_{t_{int}}|\phi\rangle \;. \tag{1.3.7}$$

1.4 Quantum parameter estimation

1.4.1 Quantum statistical models

By analogy with the classical case, a *quantum statistical model* is defined as follows.

Definition 1.4.1. ***(quantum statistical model)*** *Given a quantum system with Hilbert space $\mathcal{H}$, a quantum statistical model S is a family of density operators in $\mathcal{S}(\mathcal{H})$ parametrized*

by m real parameters $\theta \in \Theta \subset \mathbb{R}^m$:

$$S = \{\rho_\theta \,:\, \theta = (\theta^1, \theta^2, \ldots, \theta^m) \in \Theta\} \,, \tag{1.4.1}$$

where the parametrization map $\theta \to \rho_\theta$ is injective, the rank $\mathrm{rk}(\rho_\theta)$ is parameter-independent and ρ_θ can be differentiated as many times as needed with respect to the parameters.

A quantum statistical model typically arises in this way: the system is prepared at time $t = 0$ in an initial state ρ_0 and then goes through a quantum channel $\mathcal{E}_{\theta^*} \in \mathcal{T}(\mathcal{H})$, which depends on the true value θ^* of one or more parameters. The associated model is defined as $\rho_\theta := \mathcal{E}_\theta(\rho_0)$, with $\theta \in \Theta$ and Θ containing, by assumption, the true value θ^*. The mapping $\rho_0 \to \mathcal{E}_\theta(\rho_0)$ is called the *dynamical encoding*. A typical example is the unitary channel generated by the system's Hamiltonian, i.e.

$$\rho_\theta = U_t \rho_0 U_t^\dagger \,, \qquad U_t = e^{-itH_\theta} \,. \tag{1.4.2}$$

The parameter θ is referred to as a *Hamiltonian parameter*. One further distinguishes between Hamiltonian *phase* (or shift) parameters and *general* parameters. In the former case, the parameter appears linearly, as an overall multiplicative constant, i.e. $H_\theta = \theta\, G$. In the latter, the parameter enters non-linearly, so that the eigenvectors $|\xi_{j,\theta}\rangle$ of H_θ depend in general on θ.

Dynamical encoding is not, however, the only possibility. For certain models, the encoding is static. A typical example is that of a *thermal model*, describing the equilibrium state of a quantum system in contact with a thermal bath,

$$\rho_\beta = \frac{e^{-\beta H}}{\mathrm{tr}(e^{-\beta H})} \,, \tag{1.4.3}$$

where the parameter, conventionally denoted by β, is the inverse temperature of the bath and H is the Hamiltonian of the system.

In both cases, given a quantum statistical model $S = \{\rho_\theta\}_{\theta\in\Theta}$, performing a measurement with probability operators $\{\Pi_x\}_{x\in\mathcal{X}}$ gives rise to a classical statistical model, via the relation $p_\theta(x) = \mathrm{Pr}_\theta(x) = \mathrm{tr}(\rho_\theta \Pi_x)$ (where the sample space $\mathcal{X}$ is henceforth assumed to be countable). Notice that the choice of the measurement to perform is an additional degree of freedom the experimentalist is called to optimize upon, which is not present in the classical case. Furthermore, if the encoding is dynamical, one also has to optimize over the initial state of the probe ρ_0. As a consequence, the search for optimal quantum estimation protocols is considerably more complicated.

1.4.2 Quantum Riemannian metrics

A quantum statistical model can be naturally given the structure of a differentiable manifold. Whereas in the classical case there is a fundamentally unique metric, in the

quantum case non-commutativity breaks uniqueness and, in fact, leads to an infinite number of possible metrics. Notice that monotonicity now translates into the requirement that, for any completely-positive, trace-preserving map $\mathcal{E} \in \mathcal{T}(\mathcal{H})$, the difference between the metric on the original statistical model $\{\rho_\theta\}_{\theta\in\Theta}$ and on the derived model $\{\mathcal{E}(\rho_\theta)\}_{\theta\in\Theta}$ is positive semi-definite. In the quantum case, all possible monotone Riemannian metrics have been classified by Petz [25]. Each such metric is in one-to-one correspondence with an operator monotone function, which in turn is one-to-one related to an operator mean. We give the following definition:

Definition 1.4.2. ***(operator mean)*** *An operator mean* $\mathfrak{m} : \mathsf{Her}_d^+ \times \mathsf{Her}_d^+ \to \mathsf{Her}_d^+$ *is a function such that, for any positive semi-definite operators* A, B, C, D*:*

(P1) $\mathfrak{m}(A, A) = A$

(P2) $\mathfrak{m}(\alpha A, \alpha A) = \alpha A$, $\forall \alpha \in \mathbb{R}$

(P3) $A \geq C$, $B \geq D \implies \mathfrak{m}(A, B) \geq \mathfrak{m}(C, D)$

(P4) $\mathfrak{m}(UAU^\dagger, UBU^\dagger) = U\,\mathfrak{m}(A, B)\,U^\dagger$, $\forall\, U$ *unitary*

(P5) $\mathfrak{m}(A, B) = \mathfrak{m}(B, A)$

Any function aspiring to be a mean for positive semi-definite matrices should intuitively satisfy conditions **(P1)** through **(P5)**. The following proposition fully characterizes the family of operator means.

Proposition 1.4.3. *Every operator mean can be written in the form*

$$\mathfrak{m}^{(f)}(A, B) = \sqrt{A}\, f\left(\frac{1}{\sqrt{A}} B \frac{1}{\sqrt{A}}\right) \sqrt{A}\,, \tag{1.4.4}$$

where f *is an operator monotone function (i.e. a function such that,* $\forall A, B \in \mathsf{Her}_d^+$*,* $A \geq B \implies f(A) \geq f(B)$*) with the constraints* $f(1) = 1$ *and* $f(1/x) = f(x)/x$*. Vice versa, any such function gives rise to an operator mean.*

Each quantum monotone metric is now put in one-to-one correspondence with a suitable operator mean via Petz's classification theorem.

Theorem 1.4.4. ***(Petz [25])*** *If* $S = \{\rho_\theta\}_{\theta\in\Theta}$ *is a quantum statistical model such that,* $\forall\theta \in \Theta$*,* ρ_θ *is full-rank, the generic monotone Riemannian metric on* S *is of the form:*

$$[\mathcal{F}_Q^{(f)}(\theta)]_{ij} = \mathrm{tr}(\partial_i \rho_\theta\, \mathcal{J}^{-1} \partial_j \rho_\theta)\,, \tag{1.4.5}$$

where $\mathcal{J}$ *is the superoperator* $\mathcal{J} = R\, f(LR^{-1})$*,* f *is an operator-monotone function satisfying* $f(1) = 1$ *and* $f(1/x) = f(x)/x$*, and* L *(*resp. R*) is the left (*resp. *right) multiplication superoperator, which by definition acts on* $\eta \in \mathcal{S}(\mathcal{H})$ *as follows:*

$$L(\eta) = \rho_\theta \eta\,, \qquad R(\eta) = \eta \rho_\theta\,. \tag{1.4.6}$$

One may rewrite (1.4.5) more expressively by introducing the *logarithmic derivative* operators $L_{i,\theta}^{(f)}$ which satisfy the following relations:

$$\partial_i \rho_\theta = \mathcal{J} L_{i,\theta}^{(f)} , \qquad i \in \{1, \dots, m\} . \tag{1.4.7}$$

The metric $\mathcal{F}_Q^{(f)}$ can therefore be rewritten as

$$[\mathcal{F}_Q^{(f)}(\theta)]_{ij} = \mathrm{tr}[\partial_i \rho_\theta L_{j,\theta}^{(f)}] = \mathrm{tr}[\mathcal{J}(L_{i,\theta}^{(f)}) L_{j,\theta}^{(f)}] . \tag{1.4.8}$$

For each choice of an operator monotone function f, one obtains a corresponding monotone metric.

(M1) Let us consider the operator monotone function $f_{\mathrm{ari}}(x) = (1+x)/2$. The corresponding operator mean is the *arithmetic mean* since, if A, B are commuting matrices, then $\mathfrak{m}^{(f_{\mathrm{ari}})} = (A+B)/2$. The logarithimic derivative operator $L_{i,\theta}^{(f_{\mathrm{ari}})}$ satisfies, from Eq. (1.4.7),

$$\partial_i \rho_\theta = \frac{R+L}{2} L_{i,\theta}^{(f_{\mathrm{ari}})} = \frac{1}{2}\{\rho_\theta, L_{i,\theta}^{(f_{\mathrm{ari}})}\} , \tag{1.4.9}$$

so that $L_{i,\theta}^{(f_{\mathrm{ari}})}$ is also called the *symmetric* logarithmic derivative (SLD) of ρ_θ. The corresponding quantum metric is

$$[\mathcal{F}_Q^{(f_{\mathrm{ari}})}(\theta)]_{ij} = \Re\,\mathrm{tr}(\rho_\theta\, L_{i,\theta}^{(f_{\mathrm{ari}})} L_{j,\theta}^{(f_{\mathrm{ari}})}) , \tag{1.4.10}$$

which is usually referred to as the *quantum Fisher information* (QFI) metric and denoted simply by $\mathcal{F}_Q(\theta)$. It can be obtained by "quantizing" the *Bures distance* d_B^2 [26], in the sense that

$$d_B^2(\rho_\theta, \rho_{\theta+d\theta}) = \frac{1}{4}\,[\mathcal{F}_Q(\theta)]_{ij}\, d\theta^i d\theta^j , \tag{1.4.11}$$

where $d_B^2(\rho,\sigma) = 2[1 - \sqrt{F(\rho,\sigma)}]$ and $F(\rho,\sigma) = (\mathrm{tr}[\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}])^2$ is the *fidelity*.

(M2) The operator monotone function $f_{\mathrm{har}} = 2x/(1+x)$ corresponds to the harmonic mean, since for commuting matrix A, B one has $\mathfrak{m}^{(f_{\mathrm{har}})}(A,B) = 2AB/(A+B)$. From Eq. (1.4.7), one finds:

$$\partial_i \rho_\theta = \frac{2LR}{L+R} L_{i,\theta}^{(f_{\mathrm{har}})} \implies L_{i,\theta}^{(f_{\mathrm{har}})} = \frac{1}{2}\{\rho_\theta^{-1}, \partial_i \rho_\theta\} . \tag{1.4.12}$$

The corresponding metric is

$$[\mathcal{F}_Q^{(f_{\mathrm{har}})}]_{ij} = \Re\,\mathrm{tr}(\partial_i \rho_\theta \partial_j \rho_\theta \rho_\theta^{-1}) . \tag{1.4.13}$$

(M3) The logarithmic mean corresponds to $f_{\log} = (x-1)/\log x$ since, for commuting A and B, $\mathfrak{m}^{(f_{\log})}(A,B) = (B-A)/(\log B - \log A)$. From Eq. (1.4.7), one obtains the condition:

$$\partial_i \rho_\theta = \frac{L-R}{\log L - \log R} L_{i,\theta}^{(f_{\log})} \implies [\log \rho_\theta, \partial_i \rho_\theta] = [\rho_\theta,\, L_{i,\theta}^{(f_{\log})}] \ . \tag{1.4.14}$$

One can solve for $L_{i,\theta}^{(f_{\log})}$ as follows. First of all, let us recall the identity

$$\log \rho_\theta = \int_0^\infty \frac{dt}{1+t} - \int_0^\infty \frac{dt}{\rho_\theta + t} \ . \tag{1.4.15}$$

The commutator $[\log \rho_\theta, \partial_i \rho_\theta]$ can now be rewritten as follows:

$$\begin{aligned} [\log \rho_\theta, \partial_i \rho_\theta] &= \int_0^\infty dt \ \left[\partial_i \rho_\theta, \frac{1}{\rho_\theta + t}\right] \\ &= \int_0^\infty dt \ \left[\partial_i \frac{1}{\rho_\theta + t}, \rho_\theta\right] \\ &= \int_0^\infty dt \ \left[\rho_\theta, \frac{1}{\rho_\theta + t} \partial_i \rho_\theta \frac{1}{\rho_\theta + t}\right] \end{aligned} \tag{1.4.16}$$

where we made use of the fact that, for any invertible matrix M, $\partial_i M^{-1} = -M^{-1} \partial_i M M^{-1}$. From Eq. (1.4.16), $L_{i,\theta}^{(f_{\log})}$ can be read-off directly, i.e.

$$L_{i,\theta}^{(f_{\log})} = \int_0^\infty dt \ \frac{1}{\rho_\theta + t} \partial_i \rho_\theta \frac{1}{\rho_\theta + t} \ . \tag{1.4.17}$$

The corresponding metric is the *Bogoliubov-Kubo-Mori metric*:

$$[\mathcal{F}_Q^{(f_{\log})}]_{ij} = \int_0^\infty dt \ \mathrm{tr}\left(\partial_i \rho_\theta \frac{1}{\rho_\theta + t} \partial_j \rho_\theta \frac{1}{\rho_\theta + t}\right) \ . \tag{1.4.18}$$

It can be obtained by "quantizing" the *quantum relative entropy* $S(\rho||\sigma)$ [26], in the sense that

$$S(\rho_\theta || \rho_{\theta + d\theta}) = \frac{1}{2} \, [\mathcal{F}_Q(\theta)^{(f_{\log})}]_{ij} \, d\theta^i d\theta^j \ , \tag{1.4.19}$$

where $S(\rho||\sigma) = \mathrm{tr}[\rho(\log \rho - \log \sigma)]$.

It is also possible to derive a closed-form expression for $\mathcal{F}_Q^{(f)}$, with f an arbitrary operator monotone function. Notice that the superoperators L and R commute. Moreover, if $\rho_\theta = \sum_{k=1}^d p_k \, |k\rangle \langle k|$ (where $\{|k\rangle\}_{k=1}^d$ are the normalized eigenvectors of ρ_θ), then

$$L \, |k\rangle \langle l| = p_k \, |k\rangle \langle l| \ , \qquad R \, |k\rangle \langle l| = p_l \, |k\rangle \langle l| \ . \tag{1.4.20}$$

It follows that $\{|k\rangle \langle l|\}_{k,l=1}^d$ is a complete system of eigenvectors for both R and L. They are also the eigenvectors of the superoperator $\mathcal{J} = Rf(LR^{-1})$, with eigenvalues:

$$\mathcal{J} \, |k\rangle \langle l| = p_l \, f\left(\frac{p_k}{p_l}\right) |k\rangle \langle l| \ . \tag{1.4.21}$$

Let us expand the symmetric derivative operators as

$$L_{i,\theta}^{(f)} = \sum_{k,l=1}^{d} \ell_{kl}^{(i)} \, |k\rangle \langle l| \ . \tag{1.4.22}$$

Notice that since ρ_θ is full-rank, the coefficients $\ell_{kl}^{(i)}$ completely determine $L_{i,\theta}^{(f)}$. Next, one substitutes Eq. (1.4.22) into Eq. (1.4.16) and compares terms, which leads to the conditions:

$$\ell_{kl}^{(i)} = \begin{cases} \dfrac{\partial_i p_k}{p_k} & (k = l) \ , \\ \dfrac{p_l - p_k}{p_l \, f(p_k/p_l)} \, \langle k|\partial_i l\rangle & (k \neq l) \ . \end{cases} \tag{1.4.23}$$

From Eq. (1.4.8) and the previous relation, one finds:

$$[\mathcal{F}_Q^{(f)}(\theta)]_{ij} = \sum_{k=1}^{d} \frac{\partial_i p_k \, \partial_j p_k}{p_k} + \sum_{l \neq k} \frac{(p_l - p_k)^2}{p_l \, f(p_k/p_l)} \, \langle k|\partial_i l\rangle \, \langle \partial_j l|k\rangle \ , \tag{1.4.24}$$

which is our final result.

If the statistical model is not full-rank, one can still recover all possible monotone metrics by extending the metrics of Eq. (1.4.5) via a suitable fiber bundle construction (see e.g. [27]). In particular, for a pure model $S = \{|\psi_\theta\rangle\}_{\theta \in \Theta}$, the extension of the metric $\mathcal{F}_Q^{(f)}$ on S exists if and only if $f(0) \neq 0$, in which case it is always proportional to the Fubini-Study metric (which is in fact the unique unitarily invariant metric on pure states [26]). For instance, the quantum Fisher information metric evaluates to:

$$[\mathcal{F}_Q(\theta)]_{ij} = 4\Re \left[\langle \partial_i \psi_\theta | \partial_j \psi_\theta \rangle + \langle \psi_\theta | \partial_i \psi_\theta \rangle \, \langle \psi_\theta | \partial_j \psi_\theta \rangle \right] \ . \tag{1.4.25}$$

See also Ref. [28] for a closed-form expression of $\mathcal{F}_Q(\theta)$ when $1 < \mathrm{rk}(\rho_\theta) < d$.

1.4.3 Braunstein-Caves argument and the quantum Fisher information

In spite of the infinite number of possible metrics, Braunstein and Caves [17] have shown that the quantum Fisher information metric $\mathcal{F}_Q(\theta)$ is the only relevant one from an estimation viewpoint. This is true, at least, in the case of uniparametric models (i.e., when there is only one parameter $\theta = \theta^1$ to be estimated), to which from now on we restrict our attention (see however Rem. 1.4.7).

Let us recall that a typical quantum estimation protocol is specified by a triple $(\rho_0, \mathcal{M}, \hat{\theta}^{(N)})$ and can be broken down into the following steps:

(S1) Initialization: The statistical model ρ_θ is prepared by suitably encoding the parameter into an initial state ρ_0.

(S2) Measurement: A measurement $\mathcal{M}$ is performed, yielding an outcome $x \in \mathcal{X}$. When N independent measurements are taken onto identically prepared systems, one obtains a sample $\boldsymbol{x} \in \mathcal{X}^{\times N}$.

(S3) Data processing: The sample $\boldsymbol{x}$ is processed through the estimator $\hat{\theta}^{(N)}$.

The problem is to optimize over each step in order to minimize a given objective function, which is generally taken to be the mean-square-error $\mathrm{MSE}(\hat{\theta}^{(N)})$. Notice that, among the three steps, only **(S1)** and **(S2)** are properly quantum. Moreover, in the asymptotic limit of a large number of sample points, optimization over **(S3)** is trivially carried out by employing an asymptotically efficient estimator. In contrast, optimization over the measurement step **(S2)** is a non-trivial task. However, as long as $N \gg 1$, minimization of $\mathrm{MSE}(\hat{\theta}^{(N)})$ is equivalent to maximization of the Fisher information $\mathcal{F}_C(\theta)$ corresponding to the classical statistical model $p_\theta(x) = \mathrm{tr}(\rho_\theta \Pi_x)$ (with $\{\Pi_x\}_{x\in\mathcal{X}}$ the probability operators of a generic measurement $\mathcal{M}$). Therefore, the strategy usually followed is first to identify the family $\mathscr{F}$ of measurements which are available to the experimentalist, and then to maximize the Fisher information over all measurements $\mathcal{M} \in \mathscr{F}$.

We now introduce the family of *regular* measurements.

Definition 1.4.5. ***(regular measurement)*** *A measurement $\mathcal{M}$ is called regular if its probability operators are parameter-independent, i.e.*

$$\partial_\theta \Pi_x = 0 \,, \qquad \forall x \in \mathcal{X} \,; \tag{1.4.26}$$

otherwise, the measurement is non-regular.

Braunstein and Caves have maximized the Fisher information over the family $\mathscr{F}_R$ of regular measurements.

Theorem 1.4.6. ***(Braunstein-Caves [17])*** *For uniparametric model, the maximum Fisher information, optimized over the family $\mathscr{F}_R$, is the quantum Fisher information:*

$$\mathcal{F}_Q(\theta) = \max_{\mathcal{M}\in\mathscr{F}_R} \mathcal{F}_C(\theta) \,. \tag{1.4.27}$$

Proof. For a generic measurement, the Fisher information can be written as

$$\mathcal{F}_C(\theta) = \sum_{x\in\mathcal{X}^*} \frac{[\partial_\theta \,\mathrm{tr}(\rho_\theta \Pi_x)]^2}{\mathrm{tr}(\rho_\theta \Pi_x)} \,, \tag{1.4.28}$$

where $\mathcal{X}^* := \{x \in \mathcal{X} : \mathrm{tr}(\rho_\theta \Pi_x) \neq 0\}$. Notice that, in Def. (1.2.4), summation is only over those outcomes belonging to the support of p_θ. In the quantum case the role of p_θ

is taken by $\mathrm{Pr}_\theta(x) = \mathrm{tr}(\rho_\theta \Pi_x)$, so one should exclude outcomes $x \in \mathcal{X} \setminus \mathcal{X}^*$ for which $\mathrm{Pr}_\theta(x) = 0$. This clarification becomes irrelevant if ρ_θ is full-rank, since then $\mathcal{X} = \mathcal{X}^*$. Eq. (1.4.28) can be manipulated as follows:

$$\begin{aligned}
\mathcal{F}_C(\theta) &= \sum_{x\in\mathcal{X}^*} \frac{\Re^2\ \mathrm{tr}(\rho_\theta L_\theta \Pi_x)}{\mathrm{tr}(\rho_\theta \Pi_x)} && (1.4.29)\\
&\leq \sum_{x\in\mathcal{X}^*} \frac{|\,\mathrm{tr}(\rho_\theta L_\theta \Pi_x)|^2}{\mathrm{tr}(\rho_\theta \Pi_x)} && (1.4.30)\\
&= \sum_{x\in\mathcal{X}^*} \frac{|\,\mathrm{tr}(\sqrt{\Pi_x}\sqrt{\rho_\theta}\sqrt{\rho_\theta} L_\theta \sqrt{\Pi_x})|^2}{\mathrm{tr}(\rho_\theta \Pi_x)} && (1.4.31)\\
&\leq \sum_{x\in\mathcal{X}^*} \frac{\mathrm{tr}(\rho_\theta \Pi_x)\,\mathrm{tr}(L_\theta \rho_\theta L_\theta \Pi_x)}{\mathrm{tr}(\rho_\theta \Pi_x)} && (1.4.32)\\
&= \sum_{x\in\mathcal{X}^*} \mathrm{tr}(L_\theta \rho_\theta L_\theta \Pi_x) && (1.4.33)\\
&\leq \sum_{x\in\mathcal{X}} \mathrm{tr}(L_\theta \rho_\theta L_\theta \Pi_x) && (1.4.34)\\
&= \mathrm{tr}(\rho_\theta L_\theta^2) = \mathcal{F}_Q(\theta)\ . && (1.4.35)
\end{aligned}$$

In the first line, we have employed the defining relation of the symmetric logarithmic derivative $\partial_\theta \rho_\theta = \{\rho_\theta, L_\theta\}/2$; in the second line, the inequality $\Re^2 z \leq |z|^2$, $\forall z \in \mathbb{C}$; in the fourth line, the Cauchy-Schwarz inequality; in the sixth, we have extended summation over all outcomes $\mathcal{X}$, noting that $\mathrm{tr}(L_\theta \rho_\theta L_\theta \Pi_x) \geq 0$, $\forall x \in \mathcal{X}$ [1]; finally, in the last line, we have made use of the completeness relation $\sum_{x\in\mathcal{X}} \Pi_x = \mathbb{I}_d$. We have thus proved that, for any regular measurement $\mathcal{M} \in \mathscr{F}_R$, $\mathcal{F}_C(\theta) \leq \mathcal{F}_Q(\theta)$.

We will now show that there always exists a measurement saturating the previous inequality, which will establish the theorem. The above manipulations involved three separate inequalities, that to be simultaneously saturated require:

(R1) $\Im\ \mathrm{tr}(\rho_\theta L_\theta \Pi_x) = 0$, $\forall x \in \mathcal{X}^*$;

(R2) There exist complex numbers $\{\alpha_x\}_{x\in\mathcal{X}^*}$ such that $\sqrt{\rho_\theta} L_\theta \sqrt{\Pi_x} = \alpha_x \sqrt{\rho_\theta}\sqrt{\Pi_x}$;

(R3) $\sum_{x\in\mathcal{X}\setminus\mathcal{X}^*} \mathrm{tr}(L_\theta \rho_\theta L_\theta \Pi_x) = 0$.

It is easy to check that requirements **(R1)** through **(R3)** are satisfied by performing a projective measurement of the symmetric logarithmic derivative L_θ. More precisely, let us remark that the defining relation $\partial_\theta \rho_\theta = \{\rho_\theta, L_\theta\}/2$ determines L_θ only on the support of ρ_θ: outside the support $\mathrm{supp}(\rho_\theta)$, L_θ may be defined in an arbitrary way, compatible

[1] In fact, $L_\theta \rho_\theta L_\theta$ and Π_x are positive-semidefinite matrices and the trace of the product of two positive semi-definite matrices is always nonnegative.

with Hermiticity. The SLD L_θ may thus be written as follows:

$$L_\theta = \sum_{x\in\mathcal{X}} \lambda_{x,\theta} \left|\lambda_{x,\theta}\right\rangle \left\langle\lambda_{x,\theta}\right| \ , \tag{1.4.36}$$

where $\{|\lambda_{x,\theta}\rangle\}_{x\in\mathcal{X}\setminus\mathcal{X}^*}$ are chosen arbitrarily so as to give rise to an orthonormal basis. The eigenvectors and eigenvalues of L_θ are, in general, parameter-dependent. Then, if θ^* is the true value of the parameter, the optimal measurement is

$$\Pi_x^{(opt)} = \left|\lambda_{x,\theta}\right\rangle \left\langle\lambda_{x,\theta}\right| \Big|_{\theta=\theta^*} \ , \qquad \forall x \in \mathcal{X} \ , \tag{1.4.37}$$

i.e. the corresponding Fisher information satisfies $\mathcal{F}_C(\theta^*) = \mathcal{F}_Q(\theta^*)$. Notice that, for each $\theta^* \in \Theta$, there is a *different* optimal measurement: it is not required to engineer the measurement so that it satisfies Eq. (1.4.37) for any possible value of θ^*. Such a measurement would instead have probability operators $|\lambda_{x,\theta}\rangle \langle\lambda_{x,\theta}|$ and would be non-regular. However, implementing the optimal measurement *does* require to know the value of θ^* for the problem at hand, which is a priori unknown. The obstacle is overcome by employing an *adaptive procedure*, which involves constructing a sequence of estimates $\{\theta_n^*\}$ such that $\theta_n^* \to \theta^*$ and modifying the implemented measurement at each step so as to match condition (1.4.37). See e.g. Ref. [29] for more details. □

Remark 1.4.7. *One may generalize Thm. 1.4.6 to the multiparameter case. The quantum Fisher information $\mathcal{F}_Q(\theta)$ can be proven to be the least monotone metric such that $\mathcal{F}_Q(\theta) - \mathcal{F}_C(\theta)$ is positive semi-definite for any regular measurement. However, equality is not in general attainable, unless the commutativity condition $\mathrm{tr}(\rho_\theta[L_{i,\theta}, L_{j,\theta}]) = 0$ is satisfied $\forall i,j \in \{1,\dots,m\}$ [30, 31]. A widely employed solution [32] is to regularize the problem, by changing the objective function to $\mathrm{tr}[C\cdot\mathrm{Cov}(\theta)]$ (where C is a positive-definite diagonal matrix assigning different weights to different parameters). However, for this problem, the QFI metric is no longer necessarily the one providing the tightest bound [33].*

With some caveats, the quantum Fisher information therefore sets the ultimate asymptotic sensitivity bound in uniparametric problems.

Theorem 1.4.8. (quantum Cramér-Rao) *For any uniparametric estimation protocol $(\rho_0, \mathcal{M}, \hat{\theta}^{(N)})$, where $\mathcal{M} \in \mathscr{F}_R$ and the estimator $\hat{\theta}^{(N)}$ is unbiased, the following inequality holds:*

$$\mathrm{Var}(\hat{\theta}^{(N)}) \geq \frac{1}{N\cdot\mathcal{F}_Q(\theta)} \ . \tag{1.4.38}$$

Remark 1.4.9. *As in the classical case, the bound 1.4.38 is saturable only for a few special statistical models (see Ref. [13] for a precise statement). In contrast, in the asymptotic limit $N \gg 1$, one has that, for any regular measurement and any consistent estimator,*

$$\lim_{N\to\infty} N\cdot\mathrm{Var}(\hat{\theta}^{(N)}) \geq \frac{1}{\mathcal{F}_Q(\theta)} \ . \tag{1.4.39}$$

Equality can be achieved by resorting to the optimal measurement of Eq. (1.4.37) *and to an asymptotically efficient estimator.*

The last logical step is to maximize the QFI over the choice of the initial state ρ_0. To this end, the following *extended convexity* property is going to be useful.

Proposition 1.4.10. *Given a quantum statistical model $S = \{\rho_\theta\}_{\theta\in\Theta}$, where each ρ_θ is written as a convex superposition of the form $\rho_\theta = \sum_i \lambda_{i,\theta}\,\rho_{i,\theta}$, the quantum Fisher information satisfies the inequality:*

$$\mathcal{F}_Q[\rho_\theta] \leq \sum_i \lambda_{i,\theta}\mathcal{F}_Q[\rho_{i,\theta}] + \mathcal{F}_C[\{\lambda_{i,\theta}\}]\ . \tag{1.4.40}$$

The terms in square brackets specify the statistical models on which the (quantum) Fisher information is computed. From Prop. 1.4.10, assuming that the system is prepared in the parameter-independent state $\rho_0 = \sum_i \lambda_i\rho_i$ and that the parameter is encoded via a channel $\mathcal{E}_\theta$, one has

$$\mathcal{F}_Q[\rho_\theta] \leq \sum_i \lambda_i\mathcal{F}_Q[\mathcal{E}_\theta(\rho_i)]\ ; \tag{1.4.41}$$

notice that the classical term $\mathcal{F}_C[\{\lambda_i\}]$ vanishes since $\partial_\theta\lambda_i = 0$. It follows that the QFI achieves its maximum on the set of pure states. It is not possible, in general, to further determine the optimal preparation, with the significant exception of unitary models.

Let us assume that $\rho_0 = |\psi_0\rangle\langle\psi_0|$ and the encoding is provided by the unitary channel associated to $U_t = \exp(-itH_\theta)$. Then, substituting into Eq. (1.4.25), one obtains

$$\mathcal{F}_Q(\theta) = 4[\langle\psi_\theta|\mathfrak{g}_\theta^2[U_t]|\psi_\theta\rangle - (\langle\psi_\theta|\mathfrak{g}_\theta[U_t]|\psi_\theta\rangle)^2]\ , \tag{1.4.42}$$

where $|\psi_\theta\rangle = U_t\,|\psi_0\rangle$ and $\mathfrak{g}_\theta[U_t] := i\partial_\theta U_t U_t^\dagger$ is the local generator of U_t. Eq. (1.4.42) may be rewritten as

$$\mathcal{F}_Q(\theta) = 4\,\mathrm{Var}_{U_t|\psi_0\rangle}\left[\mathfrak{g}_\theta[U_t]\right] = 4\,\mathrm{Var}_{|\psi_0\rangle}[U_t^\dagger\,\mathfrak{g}_\theta[U_t]\,U_t]\ , \tag{1.4.43}$$

where $\mathrm{Var}_{|\psi\rangle}[O]$ is by definition the variance of the operator O over a state $|\psi\rangle$. Let us recall that, by Popoviciu's inequality [34], for any random variable $\mathbb{Y}$,

$$\mathrm{Var}(\mathbb{Y}) \leq \frac{(Y-y)^2}{4}\ , \tag{1.4.44}$$

where Y (*resp.* y) is the maximum (*resp.* minimum) value of $\mathbb{Y}$ and equality holds when $\mathbb{Y}$ is equally distributed over the two values Y and y. Let us also introduce the following standard notation for the eigenvalues of a matrix $M \in \mathsf{Her}_d(\mathbb{C})$: M has d real eigenvalues $\mathrm{spec}(M) = \{\lambda_1(M),\ldots,\lambda_d(M)\}$, ordered non-decreasingly, i.e. $\lambda_1(M) \geq \cdots \geq \lambda_d(M)$. It follows that

$$\mathcal{F}_Q(\theta) \leq [\lambda_1(U_t^\dagger\mathfrak{g}_\theta[U_t]U_t) - \lambda_d(U_t^\dagger\mathfrak{g}_\theta[U_t]U_t)]^2 = [\lambda_1(\mathfrak{g}_\theta[U_t]) - \lambda_d(\mathfrak{g}_\theta[U_t])]^2 = [\sigma(\mathfrak{g}_\theta[U_t])]^2\ ,$$

where the *spectral gap* of a matrix $M \in \mathsf{Her}_d(\mathbb{C})$ is defined as $\sigma[M] := \lambda_1(M) - \lambda_d(M)$. Moreover, equality is achieved by any balanced superposition of the extremal eigenvectors of the generator. We have thus established the following proposition.

Proposition 1.4.11. *Given the unitary model* $\{\rho_\theta\}_{\theta\in\Theta}$*, with* $\rho_\theta = U_t\rho_0 U_t^\dagger$ *and* $U_t = exp(-itH_\theta)$*, one has*

$$\max_{\rho_0} \mathcal{F}_Q[U_t\rho_0 U_t^\dagger] = [\sigma(\mathfrak{g}_\theta[U_t])]^2 \,. \tag{1.4.45}$$

The maximum is reached upon setting $\rho_0 = |\psi_0^{(opt)}\rangle\langle\psi_0^{(opt)}|$*, where* $|\psi_0^{(opt)}\rangle$ *is a balanced superposition of the extremal eigenvectors of the generator* $\mathfrak{g}_\theta[U_t]$*:*

$$|\psi_0^{(opt)}\rangle = \frac{1}{\sqrt{2}}\Big(|\lambda_1(\mathfrak{g}_\theta[U_t])\rangle + e^{i\phi}\,|\lambda_d(\mathfrak{g}_\theta[U_t])\rangle\Big) \,, \qquad \phi \in \mathbb{R} \,. \tag{1.4.46}$$

Chapter 2

Parameter estimation beyond the quantum Cramér-Rao theorem

In this chapter, we extend the theory of quantum parameter estimation, by enlarging the class of measurements under consideration to non-regular measurements, i.e. measurements carrying an intrinsic dependence on the unknown value of the parameter. Such measurements will be shown to lead to an improvement of the achievable precision, beyond the bound encoded by the quantum Cramér-Rao theorem [35, 38].

2.1 Non-regular measurements

A measurement $\mathcal{M}_\theta$ is said to be *non-regular* if its probability operators $\{\Pi_{x,\theta}\}_{x\in\mathcal{X}}$ are parameter-dependent. Since non-regular measurements, by definition, do not belong to the family $\mathscr{F}_R$ over which the Fisher information was optimized in Thm. 1.4.6, they might outperform the optimal Braunstein-Caves measurement. Explicitly, their Fisher information $\mathcal{F}_C(\theta)$ reads

$$\mathcal{F}_C(\theta) = \sum_{x\in\mathcal{X}^*} \frac{\Re^2 \operatorname{tr}(\rho_\theta L_\theta \Pi_{x,\theta})}{\operatorname{tr}(\rho_\theta \Pi_{x,\theta})} + \sum_{x\in\mathcal{X}^*} \frac{[\operatorname{tr}(\rho_\theta\, \partial_\theta \Pi_{x,\theta})]^2}{\operatorname{tr}(\rho_\theta \Pi_{x,\theta})} + 2 \sum_{x\in\mathcal{X}^*} \frac{\Re \operatorname{tr}(\rho_\theta L_\theta \Pi_{x,\theta}) \operatorname{tr}(\rho_\theta\, \partial_\theta \Pi_{x,\theta})}{\operatorname{tr}(\rho_\theta \Pi_{x,\theta})} \,.$$

The first term on the RHS is the same that appears on the first line of Eq. (1.4.29) and that is bounded from above by the QFI, but there are also two additional contributions. In general, they will have an important effect on the achievable sensitivity (though they are not always positive, so a precision enhancement is not guaranteed).

It is not immediately clear how to implement non-regular measurements. Seemingly, one would need to know beforehand the true value of the parameter. The same could be said of the statistical model ρ_θ but, in the latter case, the true value of the parameter is encoded into the initial state, e.g. by making use of the time-evolution of the system

as a resource. In the same way, a non-regular measurement requires the parameter to be suitably encoded into its probability operators. We now describe two scenarios where this is possible.

(S1) Measurement models with parameter-dependent interactions: Let us model a non-regular measurement as in Sect. 1.3, by specifying the interaction between the system and the apparatus. The total Hamiltonian is $H_T = H_\theta + H_A + H_{I,\theta}$, where it is assumed that the free Hamiltonian H_A of the apparatus does not depend on the parameter, but the coupling term $H_{I,\theta}$ does. We also suppose that the duration of the measurement t_{int} is so short and the interaction so strong, that is possible to neglect the free evolution of the two systems. Equivalently, the time-evolution operator during the measurement process is $U_t \sim \exp(-itH_{I,\theta})$. If the apparatus is prepared in a reference state $|\phi\rangle$ and a projective measurement $\{P_x\}_{x\in\mathcal{X}}$ is made on the ancilla after a time t_{int}, the resulting probability operators read

$$\Pi_{x,\theta} = \langle\phi|e^{itH_{I,\theta}}\, \mathbb{1}_d \otimes P_x\, e^{-itH_{I,\theta}}|\phi\rangle \tag{2.1.1}$$

and are, in general, parameter-dependent.

Example 2.1.1. *The parameter to be estimated is the frequency ω of a bosonic mode in a cavity. The system's Hamiltonian is $H_\omega = \omega(a^\dagger a + 1/2)$, the initial state is chosen as $|\psi_0\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ and the statistical model at time t is $|\psi_\omega\rangle := U_t|\psi_0\rangle = \alpha_0 e^{-i\omega t/2}|0\rangle + \alpha_1 e^{-3i\omega t/2}|1\rangle$, where $U_t := \exp(-itH_\omega)$. The QFI evaluates to $\mathcal{F}_Q(\omega) = 4t^2|\alpha_0|^2|\alpha_1|^2$, which is the maximum information extractable via regular measurements.*

A non-regular measurement can be engineered by coupling the bosonic mode to a two-level atom, which is initially in its ground state $|g\rangle$, and by measuring whether the atom has been excited or not after an interaction time t_{int}. The interaction Hamiltonian is of the Jaynes-Cummings type $H_I = \Omega(a^\dagger\sigma_- + a\sigma_+)$, where $\Omega := d\sqrt{\omega/2\epsilon_0 V}$, $d := \vec{\epsilon}\cdot\langle e|\,\vec{d}\,|g\rangle$, $\vec{\epsilon}$ is the photon polarization, ϵ_0 is the dielectric constant, V the volume of the cavity, $\vec{d}$ the dipole operator, $|g\rangle$ the atom's ground state, $|e\rangle$ the excited state, $\sigma_+ := |e\rangle\langle g|$ and $\sigma_- := |g\rangle\langle e|$. Notice that the interaction Hamiltonian is parameter-dependent. Explicitly, the evolution operator U_t during the measurement process is

$$U_t = U_{gg}|g\rangle\langle g| + U_{ge}|g\rangle\langle e| + U_{eg}|e\rangle\langle g| + U_{ee}|e\rangle\langle e|\ , \tag{2.1.2}$$

where, letting $N := a^\dagger a$ denote the number operator for the radiation field, we have

defined

$$U_{gg} := \cos(\Omega t\sqrt{N}) \,, \qquad U_{ge} := -i\frac{\sin(\Omega t\sqrt{N})}{\sqrt{N}}a^\dagger \,, \\ U_{eg} := -i\frac{\sin(\Omega t\sqrt{1+N})}{\sqrt{1+N}}a \,, \qquad U_{ee} := \cos(\Omega t\sqrt{1+N}) \,. \tag{2.1.3}$$

By convention, the outcome 0 *is obtained if the atom is measured in the ground state and the outcome* 1 *if measured in the excited state. From Eq.* (1.3.7)*, the measurement operators are*

$$M_{0,\omega} = \langle g| U_{t_{int}} |g\rangle = \cos(\Omega t_{int}\sqrt{N}) \,, \quad M_{1,\omega} = \langle e| U_{t_{int}} |g\rangle = -i\frac{\sin(\Omega t_{int}\sqrt{1+N})}{\sqrt{1+N}}a \,. \tag{2.1.4}$$

The corresponding probability operators are

$$\Pi_{0,\omega} = \cos^2(\Omega t_{int}\sqrt{N}) \,, \qquad \Pi_{1,\omega} = \sin^2(\Omega t_{int}\sqrt{N}) \,. \tag{2.1.5}$$

They depend on the parameter ω via the coupling constant Ω. The corresponding Fisher information is

$$\mathcal{F}_C(\omega) = \left(\frac{\Omega t_{int}}{\omega}\right)^2 \frac{|\alpha_1|^2 \cos^2(\Omega t_{int})}{1 - |\alpha_1|^2 \sin^2(\Omega t_{int})} \,, \tag{2.1.6}$$

which is not necessarily bounded from above by the QFI. For instance, if the system is initially prepared in an energy eigenstate, e.g. $\alpha_1 = 0$, then the QFI vanishes (there is no regular measurement that can estimate the parameter with finite precision), but $\mathcal{F}_C(\omega) = \Omega^2 t_{int}^2/\omega^2$.

(S2) Energy measurements of non-linear Hamiltonians: If the Hamiltonian H_θ depends on the parameter θ in a non-linear way (i.e. it is not of the form $H_\theta = \theta G$), its eigenstates $\{|\xi_{j,\theta}\rangle\}_{j=0}^{d-1}$ are in general parameter-dependent. An energy measurement corresponds to the projective probability operators $\Pi_{\xi_j,\theta} = |\xi_{j,\theta}\rangle\langle\xi_{j,\theta}|$, thus the measurement is non-regular.

Example 2.1.2. *The parameter to be estimated is the strength g of a uniform gravitational field. The probing system is a mechanical oscillator, with Hamiltonian $H_g = -\partial_x^2/2m + kx^2/2 + mgx$, where m is the mass of the oscillator, k its elastic constant and x denotes the vertical displacement of the oscillator from equilibrium. The energy eigenstates have the following wavefunctions:*

$$\psi_j = \left(\frac{m\omega}{\pi}\right)^{1/4} \frac{1}{\sqrt{2^j\, j!}} H_j(\varkappa + \varkappa_g)\, e^{-(\varkappa+\varkappa_g)^2/2} \,, \tag{2.1.7}$$

where $j \in \mathbb{N}_0$, H_j is the j^{th} Hermite polynomial, $\omega := \sqrt{k/m}$, $\varkappa$ is the dimensionless coordinate $\varkappa := x/\ell$, ℓ is the characteristic length of the oscillator $\ell := 1/\sqrt{m\omega}$ and $\varkappa_g := mg/k\ell$. The corresponding eigenvalues are $\xi_{j,g} = \omega(j+1/2) - mg^2/2\omega^2$. At time $t = 0$, the oscillator is cooled to its ground state ψ_0; it is henceforth mechanically displaced from its equilibrium point by a distance δx, so that the initial state is

$$\psi(x,0) = \left(\frac{m\omega}{\pi}\right)^{1/4} e^{-(\varkappa+\varkappa_\delta)^2/2} , \qquad \varkappa_\delta := \delta x/\ell . \tag{2.1.8}$$

At the generic time t, the wavefunction of the oscillator reads

$$\psi(x,t) = \left(\frac{m\omega}{\pi}\right)^{1/4} e^{-i\omega t(1-\varkappa_g^2)/2}\, e^{-(\varkappa+\varkappa_g)^2/2} \exp\left[\Phi_g\right] , \tag{2.1.9}$$

where

$$\Phi_g = -e^{-i\omega t}\left(\frac{(\varkappa_\delta - \varkappa_g)^2}{2}\cos\omega t + (\varkappa_\delta - \varkappa_g)(\varkappa + \varkappa_g)\right) . \tag{2.1.10}$$

The computation of the QFI for the statistical model of Eq. (2.1.10) *can be carried out straightforwardly (see Ref. [35] for details); the final result is*

$$\mathcal{F}_Q(g) = \frac{8m}{\omega^3}\sin^2\left(\frac{\omega t}{2}\right) . \tag{2.1.11}$$

It should be compared with the Fisher information $\mathcal{F}_C(g)$ corresponding to an energy measurement, which is $\mathcal{F}_C(g) = 2m/\omega^3$. Notice that $\mathcal{F}_C(g)$ exceeds the QFI for certain values of the interrogation time t.

2.2 Non-regular estimation of general Hamiltonian parameters

In this section, we further study non-regular estimation protocols based on energy measurements of non-linear Hamiltonians. The plan is to introduce a family of measurements that are non-regular and have a clear-cut physical interpretation; to maximize the Fisher information over such a family; to identify the best-performing measurement and, finally, to compare it with the optimal Braunstein-Caves measurement.

2.2.1 Controlled energy measurements

Let us consider a projective measurement of H_θ, with θ a general Hamiltonian parameter. It is assumed that H_θ has eigenvalues $\xi_{j,\theta} = \lambda_{d-j}(H_\theta)$. With no significant loss of generality, the spectrum is taken to be non-degenerate. The probability of each measurement outcome is

$$\Pr_\theta(\xi_{j,\theta}) = \mathrm{tr}(\rho_\theta P_{\xi_{j,\theta}}) = \langle \xi_{j,\theta}|\rho_0|\xi_{j,\theta}\rangle , \tag{2.2.1}$$

where $\rho_\theta = U_t \rho_0 U_t^\dagger$, $U_t = \exp(-itH_\theta)$ and $P_{\xi_{j,\theta}} = |\xi_{j,\theta}\rangle \langle \xi_{j,\theta}|$.

The corresponding sample space $\mathcal{X}_\theta = \{\xi_{j,\theta}\}_{j=0}^{d-1}$ is, in general, parameter-dependent. This is a significant complication, since there is no established theory for statistical models with parameter-dependent sample spaces. In fact, if the sample space is allowed to depend on θ, the proof of the classical Cramér-Rao theorem, Thm. 1.4.6, breaks down. In some cases, it is even possible to construct unbiased estimators having vanishing variance [36]. To exclude such pathological situations, we assume in the following that either the eigenstates of H_θ are parameter-dependent, but not its eigenvalues; or that the outcomes of an energy measurement are processed via a suitable statistic $Y : \mathcal{X}_\theta \to \mathcal{Y}$, where $\mathcal{Y}$ is a conventional parameter-independent sample space. Estimators having vanishing variance can no longer occur and the Fisher information is again the only relevant performance metric.

We now introduce a family of non-regular measurements, referred to as *controlled energy measurements*, obtained by first applying a unitary control $V \in U(d)$ and then performing a projective energy measurement.

Definition 2.2.1. ***(controlled energy measurement)*** *A controlled energy measurement $\mathcal{M}_\theta^{(V)}$ has sample space $\mathcal{Y} = \{\zeta_j := Y(\xi_{j,\theta})\}_{j=0}^{d-1}$ and probability operators $\{\Pi_{\zeta_j}\}_{j=0}^{d-1}$, where $\Pi_{\zeta_j} := V^\dagger P_{\xi_{j,\theta}} V$, $V \in U(d)$ is a unitary parameter-independent control and $P_{\xi_{j,\theta}}$ is the projector over the j^{th} energy eigenstate of H_θ.*

The Fisher information of a controlled energy measurement $\mathcal{M}_\theta^{(V)}$ is denoted by $\mathcal{F}_C^{(V)}(\theta)$. Let us remark that an energy measurement corresponds to the choice $V = \mathbb{I}_d$. Its probability measure (see Eq. (2.2.1)) is t-independent, which implies that also the Fisher information does not depend on t. In contrast, the QFI generically grows quadratically with t [37]. Therefore, for sufficiently long times, an energy measurement can never outperform the optimal Braunstein-Caves measurement. If, however, a control is applied before the measurement, then the Fisher information $\mathcal{F}_C^{(V)}(\theta)$ can grow again like t^2 and, in fact, can even outperform the optimal Braunstein-Caves measurement for any t, as will be discussed in the following.

If an experimentalist is allowed to implement arbitrary controlled energy measurements, the maximum Fisher information she can extract is

$$\mathcal{G}(\theta) := \max_{\rho_0} \max_{V \in U(d)} \mathcal{F}_C^{(V)}[\rho_\theta] \, . \tag{2.2.2}$$

Compared with regular measurements, an enhancement is achievable if and only if $\mathcal{G}(\theta) > [\sigma(\mathfrak{g}_\theta[U_t])]^2$. However, computing $\mathcal{G}(\theta)$ directly from its definition is a non-trivial task. In the next section, a closed-form formula for $\mathcal{G}(\theta)$ is derived under the assumption that the Hamiltonian H_θ satisfies a rather general condition.

2.2.2 Bounding $\mathcal{G}$

For a generic controlled energy measurement $\mathcal{M}_\theta^{(V)}$, the probability of the outcome ζ_j is

$$\Pr_\theta(\zeta_j) = \mathrm{tr}(\rho_\theta V^\dagger P_{\xi_{j,\theta}} V) \ . \tag{2.2.3}$$

Let us denote by $\{|j\rangle\}_{j=0}^{d-1}$ the computational basis on the Hilbert space $\mathcal{H}$ of the system. The two orthonormal basis $\{|j\rangle\}_{j=0}^{d-1}$ and $\{|\xi_{j,\theta}\rangle\}_{j=0}^{d-1}$ are connected by a unitary transformation, denoted by $S \in U(d)$, such that $|j\rangle = S\,|\xi_{j,\theta}\rangle$. Explicitly, the matrix elements of S are $\langle j|S|k\rangle = \langle \xi_{j,\theta}|k\rangle$. Notice that, for a general Hamiltonian parameter, the matrix S is θ-dependent and that S reduces H_θ to diagonal form, i.e. $SH_\theta S^\dagger = \mathrm{diag}(\xi_{0,\theta},\, \ldots,\, \xi_{d-1,\theta})$. One may thus rewrite Eq. (2.2.3) as follows,

$$\Pr_\theta(\zeta_j) = \mathrm{tr}\left[(SVU_t)\rho_0(SVU_t)^\dagger P_j\right] = \mathrm{tr}\left(\tilde{U}^{(V)}\rho_0\tilde{U}^{(V)\dagger}P_j\right) \ , \tag{2.2.4}$$

where $P_j := |j\rangle\langle j|$ and all dependence on θ has been collected into the unitary matrix $\tilde{U}^{(V)} := SVU_t$. Formally, a controlled energy measurement on the model ρ_θ is equivalent to a projective measurement in the computational basis on the model $\rho_\theta^{(V)} := \tilde{U}^{(V)}\rho_0\tilde{U}^{(V)\dagger}$. The Fisher information corresponding to $\mathcal{M}_\theta^{(V)}$ can thus be written as

$$\mathcal{F}_C^{(V)}(\theta) = \sum_{j\in\mathcal{J}^*} \frac{[\partial_\theta \,\mathrm{tr}(\rho_\theta^{(V)}P_j)]^2}{\mathrm{tr}(\rho_\theta^{(V)}P_j)} \ , \tag{2.2.5}$$

where $\mathcal{J}^*$ is the subset of $\mathcal{J} := \{0,\ldots,d-1\}$ such that $j \in \mathcal{J}^*$ if and only if $\Pr_\theta(\zeta_j) \neq 0$. The task is to maximize the RHS of Eq. (2.2.5) over the unitary group $U(d)$ of available controls V and over the initial preparation ρ_0.

Theorem 2.2.2. *The maximum Fisher information $\mathcal{G}(\theta)$ that can be extracted via controlled energy measurements satisfies the inequality*

$$\mathcal{G}(\theta) \leq \left[\sigma(\mathfrak{g}_\theta[U_t]) + \sigma(\mathfrak{g}_\theta[S])\right]^2 \ , \tag{2.2.6}$$

*where $U_t = exp(-itH_\theta)$ is the unitary encoding, S is the similarity transformation diagonalizing H_θ, $\mathfrak{g}_\theta[U_t]$ (*resp.*, $\mathfrak{g}_\theta[S]$) is the generator of U_t (*resp.*, S), i.e.*

$$\mathfrak{g}_\theta[U_t] = i\partial_\theta U_t U_t^\dagger \ , \qquad \mathfrak{g}_\theta[S] = i\partial_\theta S S^\dagger \ , \tag{2.2.7}$$

and $\sigma(M)$ denotes the spectral gap of a matrix $M \in \mathsf{Her}_d(\mathbb{C})$.

Proof. The Fisher information for $\mathcal{M}_\theta^{(V)}$ is given by Eq. (2.2.5). Introducing the symmetric logarithmic derivative $L_\theta^{(V)}$ of $\rho_\theta^{(V)}$,

$$\mathcal{F}_C^{(V)}(\theta) = \sum_{j\in\mathcal{J}^*} \frac{\Re^2\,\mathrm{tr}(\rho_\theta^{(V)}L_\theta^{(V)}P_j)}{\mathrm{tr}(\rho_\theta^{(V)}P_j)} \ . \tag{2.2.8}$$

Using the inequality $\Re z \leq |z|$, $\forall z \in \mathbb{C}$, and then the Cauchy-Schwarz inequality, the numerator can be bounded as follows,

$$\begin{aligned} \Re^2 \operatorname{tr}(\rho_\theta^{(V)} L_\theta^{(V)} P_j) &\leq |\operatorname{tr}(\rho_\theta^{(V)} L_\theta^{(V)} P_j)|^2 \\ &\leq \operatorname{tr}(L_\theta^{(V)} \rho_\theta^{(V)} L_\theta^{(V)} P_j) \operatorname{tr}(\rho_\theta^{(V)} P_j) \ . \end{aligned} \tag{2.2.9}$$

Therefore,

$$\begin{aligned} \mathcal{F}_C^{(V)}(\theta) &\leq \sum_{j \in \mathcal{J}^*} \operatorname{tr}(L_\theta^{(V)} \rho_\theta^{(V)} L_\theta^{(V)} P_j) \\ &\leq \sum_{j \in \mathcal{J}} \operatorname{tr}(L_\theta^{(V)} \rho_\theta^{(V)} L_\theta^{(V)} P_j) \\ &= \operatorname{tr}[\rho_\theta^{(V)} (L_\theta^{(V)})^2] \ . \end{aligned} \tag{2.2.10}$$

Taking the maximum over the initial preparation,

$$\max_{\rho_0} \mathcal{F}_C^{(V)}(\theta) \leq \max_{\rho_0} \ \operatorname{tr}[\rho_\theta^{(V)} (L_\theta^{(V)})^2] \ . \tag{2.2.11}$$

By convexity, the maximum of the expression on the RHS is achieved when the system is prepared in a pure state. Let us set $\rho_0 = |\psi_0\rangle \langle\psi_0|$. One can then rewrite it as

$$\operatorname{tr}[\rho_\theta^{(V)} (L_\theta^{(V)})^2]\Big|_{\rho_0 = |\psi_0\rangle\langle\psi_0|} = 4 \operatorname{Var}_{|\psi_0\rangle}(\tilde{U}^{(V)\dagger} \mathfrak{g}_\theta[\tilde{U}^{(V)}] \, \tilde{U}^{(V)}) \ , \tag{2.2.12}$$

where

$$\mathfrak{g}_\theta[\tilde{U}^{(V)}] = \mathfrak{g}_\theta[S] + (SV) \, \mathfrak{g}_\theta[U_t] \, (SV)^\dagger \tag{2.2.13}$$

is the local generator of $\tilde{U}^{(V)}$. By Popoviciu's inequality,

$$\max_{\rho_0} \mathcal{F}_C^{(V)}(\theta) \leq [\sigma(\mathfrak{g}_\theta[S] + (SV) \, \mathfrak{g}_\theta[U_t] \, (SV)^\dagger)]^2 \ . \tag{2.2.14}$$

After maximizing over the choice of the unitary control V,

$$\mathcal{G}(\theta) \leq \max_{V \in U(d)} [\sigma(\mathfrak{g}_\theta[S] + (SV) \, \mathfrak{g}_\theta[U_t] \, (SV)^\dagger)]^2 \ . \tag{2.2.15}$$

The maximization on the RHS can be carried out explicitly with the help of the following lemma: the maximum spectral gap of the sum of any two Hermitian matrices with given spectra is equal to the sum of their spectral gaps, i.e.

$$\max_{U_1, U_2 \in U(d)} \sigma(U_1 M_1 U_1^\dagger + U_2 M_2 U_2^\dagger) = \sigma(M_1) + \sigma(M_2) \ , \qquad M_1, M_2 \in \mathsf{Her}_d(\mathbb{C}) \ . \tag{2.2.16}$$

See Ref. [38] for a proof. From Eq. (2.2.16), Eq. (2.2.6) follows immediately. □

2.2.3 Saturability of inequality (2.2.6)

We now discuss tightness of inequality (2.2.6). The proof of Thm. 2.2.2 can be broken down into three main steps:

(S1) In Eq. (2.2.10), the Fisher information $\mathcal{F}_C^{(V)}(\theta)$ was bounded from above. This step actually made use of three different inequalities: the inequality $\Re z \leq |z|$ (on the first line of Eq. (2.2.9)), the Cauchy-Schwarz inequality (on the second line of Eq. (2.2.9)) and the inequality on the second line of Eq. (2.2.10), which follows from

$$\sum_{j\in\mathcal{J}\setminus\mathcal{J}^*} \mathrm{tr}(L_\theta^{(V)}\rho_\theta^{(V)}L_\theta^{(V)}P_j) \geq 0 \; . \tag{2.2.17}$$

(S2) Next, the quantity on the RHS of Eq. (2.2.11) was maximized over the initial preparation ρ_0, which led to Eq. (2.2.14).

(S3) Finally, maximization over the unitary control V was performed.

Steps **(S2)** and **(S3)** are proper maximizations, that can be made tight by implementing the optimal control $V^{(opt)}$ and the optimal initial preparation $|\psi_0^{(opt)}\rangle$. It is easy to check that the optimal control has the form

$$V^{(opt)} = S^\dagger R_1^\dagger R_2 \; , \tag{2.2.18}$$

where R_1 (*resp.*, R_2) is the similarity transformation that diagonalizes $\mathfrak{g}_\theta[S]$ (*resp.*, $\mathfrak{g}_\theta[U_t]$), with eigenvalues ordered decreasingly, i.e.

$$\begin{aligned} R_1\mathfrak{g}_\theta[S]R_1^\dagger &= \mathrm{diag}(\lambda_1(\mathfrak{g}_\theta[S]),\ldots,\lambda_d(\mathfrak{g}_\theta[S])) \; , \\ R_2\mathfrak{g}_\theta[U_t]R_2^\dagger &= \mathrm{diag}(\lambda_1(\mathfrak{g}_\theta[U_t]),\ldots,\lambda_d(\mathfrak{g}_\theta[U_t])) \; . \end{aligned} \tag{2.2.19}$$

Moreover, from Popoviciu's inequality, the optimal initial preparation is

$$|\psi_0^{(opt)}\rangle = \frac{1}{\sqrt{2}}\,\tilde{U}^{(V^{(opt)})\dagger}[|\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle + e^{i\phi}\,|\lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle] \; , \qquad \phi\in\mathbb{R} \; , \tag{2.2.20}$$

where $\tilde{U}^{(V^{(opt)})} = SV^{(opt)}U_t$. The previous expression for $|\psi_0^{(opt)}\rangle$ can be slightly simplified by noticing that the extremal eigenvalues of the generator of $\tilde{U}^{(V^{(opt)})}$ coincide with the extremal eigenvalues of the generator of S. This can be proven as follows. From Eq. (2.2.13) and Eq. (2.2.18), the generator of $\tilde{U}^{(V^{(opt)})}$ can be written as

$$\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}] = \mathfrak{g}_\theta[S] + R_1^\dagger R_2\mathfrak{g}_\theta[U_t]R_2^\dagger R_1 = R_1^\dagger D R_1 \; , \tag{2.2.21}$$

where D is the diagonal matrix

$$D = \mathrm{diag}[\lambda_1(\mathfrak{g}_\theta[S]) + \lambda_1(\mathfrak{g}_\theta[U_t]),\ldots,\lambda_d(\mathfrak{g}_\theta[S]) + \lambda_d(\mathfrak{g}_\theta[U_t])] \; . \tag{2.2.22}$$

Therefore, the extremal eigenvectors of $\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]$ are given by

$$|\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle = R_1^\dagger\,|0\rangle\ , \qquad |\lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle = R_1^\dagger\,|d-1\rangle\ . \tag{2.2.23}$$

But, by the very definition of R_1, $R_1^\dagger\,|0\rangle = |\lambda_1(\mathfrak{g}_\theta[S])\rangle$ and $R_1^\dagger\,|d-1\rangle = |\lambda_d(\mathfrak{g}_\theta[S])\rangle$, which establishes our claim. One may thus write

$$|\psi_0^{(opt)}\rangle = \frac{1}{\sqrt{2}}\,\tilde{U}^{(V^{(opt)})\dagger}[|\lambda_1(\mathfrak{g}_\theta[S])\rangle + e^{i\phi}\,|\lambda_d(\mathfrak{g}_\theta[S])\rangle]\ , \qquad \phi\in\mathbb{R}\ . \tag{2.2.24}$$

Proving tightness of inequality (2.2.6) is therefore equivalent to proving that of step **(S1)**, under the constraints that the control and the initial preparation are chosen according to Eq. (2.2.18) and Eq. (2.2.24), respectively. Let us first consider the majorization based on the Cauchy-Schwarz inequality, which is saturated if and only if, $\forall j\in\mathcal{J}^*$, there exist complex numbers $\{\alpha_j\}$ such that

$$\sqrt{\rho_\theta^{(V^{(opt)})}}P_j = \alpha_j\sqrt{\rho_\theta^{(V^{(opt)})}}L_\theta^{(V^{(opt)})}P_j\ . \tag{2.2.25}$$

When the model is pure, condition (2.2.25) is automatically satisfied since it reduces to

$$\langle\psi_\theta^{(V^{(opt)})}|j\rangle\ |\psi_\theta^{(V^{(opt)})}\rangle\,\langle j| = \alpha_j\,\langle\psi_\theta^{(V^{(opt)})}|L_\theta^{(V^{(opt)})}|j\rangle\ |\psi_\theta^{(V^{(opt)})}\rangle\,\langle j| \tag{2.2.26}$$

(where we have set $|\psi_\theta^{(V^{(opt)})}\rangle := \tilde{U}^{(V^{(opt)})}\,|\psi_0^{(opt)}\rangle$), which implies

$$\alpha_j = \frac{\langle\psi_\theta^{(V^{(opt)})}|L_\theta^{(V^{(opt)})}|j\rangle}{\langle\psi_\theta^{(V^{(opt)})}|j\rangle}\ . \tag{2.2.27}$$

The remaining two inequalities used in step **(S1)** cannot be saturated without making further assumptions about the Hamiltonian H_θ. For the inequality $\Re z\leq|z|$ to be tight, one should have, $\forall j\in\mathcal{J}^*$,

$$\Im\mathfrak{s}\left[\langle j|L_\theta^{(V^{(opt)})}|\psi_\theta^{(V^{(opt)})}\rangle\,\langle\psi_\theta^{(V^{(opt)})}|j\rangle\right] = 0\ . \tag{2.2.28}$$

The SLD $L_\theta^{(V^{(opt)})}$ can be checked to be given by

$$L_\theta^{(V^{(opt)})} = 2\left[|\partial_\theta\psi_\theta^{(V^{(opt)})}\rangle\,\langle\psi_\theta^{(V^{(opt)})}| + |\psi_\theta^{(V^{(opt)})}\rangle\,\langle\partial_\theta\psi_\theta^{(V^{(opt)})}|\right]\ , \tag{2.2.29}$$

which, introducing the generator $\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]$, leads to

$$L_\theta^{(V^{(opt)})}\,|\psi_\theta^{(V^{(opt)})}\rangle = 2i\left(\langle\psi_\theta^{(V^{(opt)})}|\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]|\psi_\theta^{(V^{(opt)})}\rangle - \mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]\right)|\psi_\theta^{(V^{(opt)})}\rangle\ . \tag{2.2.30}$$

Substituting the last relation in Eq. (2.2.28), one obtains

$$\langle\psi_\theta^{(V^{(opt)})}|\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]|\psi_\theta^{(V^{(opt)})}\rangle\ |\,\langle\psi_\theta^{(V^{(opt)})}|j\rangle\,|^2 = \Re\left[\langle j|\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]|\psi_\theta^{(V^{(opt)})}\rangle\ \langle\psi_\theta^{(V^{(opt)})}|j\rangle\right]\ . \tag{2.2.31}$$

Using the explicit form of the optimal preparation given in Eq. (2.2.20),

$$\left(|\langle j|\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle|^2 - |\langle j|\lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle|^2\right)[\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}]) - \lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])] = 0\,. \tag{2.2.32}$$

By Eq. (2.2.21), the previous relation can also be written as

$$\left(|\langle j|\lambda_1(\mathfrak{g}_\theta[S])\rangle|^2 - |\langle j|\lambda_d(\mathfrak{g}_\theta[S])\rangle|^2\right)[\sigma(\mathfrak{g}_\theta[S]) + \sigma(\mathfrak{g}_\theta[U_t])] = 0\,. \tag{2.2.33}$$

Therefore, the inequality is tight provided that

$$|\langle j|\lambda_1(\mathfrak{g}_\theta[S])\rangle| = |\langle j|\lambda_d(\mathfrak{g}_\theta[S])\rangle|\,, \qquad \forall j \in \mathcal{J}^*\,, \tag{2.2.34}$$

i.e., the extremal eigenvectors of the generator of S, written in the computational basis, are such that corresponding entries have the same complex moduli.

It remains to discuss tightness of inequality (2.2.17). Let $j \in \mathcal{J}\backslash\mathcal{J}^*$. This is equivalent to $\langle j|\psi_\theta^{(V^{(opt)})}\rangle = 0$ where, as before,

$$|\psi_\theta^{(V^{(opt)})}\rangle = \frac{1}{\sqrt{2}}\left[|\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle + e^{i\phi}\,|\lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle\right]. \tag{2.2.35}$$

For $\langle j|\psi_\theta^{(V^{(opt)})}\rangle = 0$ to hold, there are two possibilities: either both

$$\langle j|\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle = 0 \quad \text{and} \quad \langle j|\lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle = 0\,; \tag{2.2.36}$$

or they are different from zero, have the same moduli and the correct phase difference to cancel each other out. This last possibility can be excluded since the phase ϕ is arbitrary and can always be set such that no cancellation occurs. So the only possibility is for Eq. (2.2.36) to hold. Now, to prove tightness, one should show that

$$\langle j|L_\theta^{(V^{(opt)})}|\psi_\theta^{(V^{(opt)})}\rangle = 0\,, \qquad \forall j \in \mathcal{J}\setminus\mathcal{J}^*. \tag{2.2.37}$$

Using Eq. (2.2.30) and (2.2.35), one arrives at the equivalent condition

$$\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\,\langle j|\lambda_1(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle + \lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\,e^{i\phi}\,\langle j|\lambda_d(\mathfrak{g}_\theta[\tilde{U}^{(V^{(opt)})}])\rangle = 0\,, \tag{2.2.38}$$

which is trivially satisfied because of Eq. (2.2.36). Thus, no additional assumption is needed for equality to hold in Eq. (2.2.17). We summarize our results via the following proposition.

Proposition 2.2.3. *If H_θ is such that the extremal eigenvectors of the generator of its diagonalizing matrix S satisfy the condition*

$$|\langle j|\lambda_1(\mathfrak{g}_\theta[S])\rangle| = |\langle j|\lambda_d(\mathfrak{g}_\theta[S])\rangle|\,, \qquad \forall j \in \mathcal{J}^*\,, \tag{2.2.39}$$

then the maximum Fisher information extractable via controlled energy measurements is

$$\mathcal{G}(\theta) = \left[\sigma(\mathfrak{g}_\theta[U_t]) + \sigma(\mathfrak{g}_\theta[S])\right]^2 \ . \tag{2.2.40}$$

The optimal preparation is given by Eq. (2.2.20) *and the optimal control by Eq.* (2.2.18)*.*

The condition imposed by Eq. (2.2.39) on H_θ may seem quite restricting. However, it turns out to be satisfied for many Hamiltonians of practical use in quantum metrology (as discussed in Sect. (2.2.5)). Eq. (2.2.40) thus often provides a way to directly compute $\mathcal{G}(\theta)$, without the need of any optimization procedure.

2.2.4 Metrological applications

In this section, we discuss how to implement controlled energy measurements in a realistic metrological scenario. In principle, a controlled energy measurement requires to apply a unitary control V, and then to measure the energy projectively. The question is how to perform a projective measurement of the Hamiltonian when the Hamiltonian is not fully known. The problem has first been investigated in Refs. [39, 40]. In the following, we associate to each controlled energy measurement $\mathcal{M}_\theta^{(V)}$ a family of measurements, called *realistic controlled energy measurements*, denoted by $\mathcal{M}_{n,m}^{(V)}$ (with $n, m \in \mathbb{N}$), that are experimentally feasible and allow to approximate $\mathcal{M}_\theta^{(V)}$ to any desired level of accuracy (in the sense that, as $n, m \to \infty$ the probability measure of $\mathcal{M}_{n,m}^{(V)}$ converges to that of $\mathcal{M}_\theta^{(V)}$).

Our exposition can be divided into two parts. First, we describe a simplified version, denoted by $\mathcal{M}_n^{(V)}$, which is based on the *phase estimation* algorithm [41–43]. It is assumed that the experimentalist can implement the controlled time-evolution operator

$$C_{U_t} := |0\rangle\langle 0| \otimes \mathbb{I}_d + |1\rangle\langle 1| \otimes U_t \ . \tag{2.2.41}$$

This is an unrealistic assumption, since C_{U_t} still depends on the true value of the parameter via U_t. Next, we remove such assumption, which will lead to the introduction of realistic controlled energy measurements.

In order to implement $\mathcal{M}_n^{(V)}$, one introduces n control qubits, each one having Hilbert space $\mathcal{H}_c = \mathbb{C}^2$. The total Hilbert space is thus $\mathcal{H}_c^{\otimes n} \otimes \mathcal{H}$, with $\mathcal{H} = \mathbb{C}^d$ the Hilbert space of the original system. Every control qubit is prepared in its ground state $|0\rangle$. At time $t = 0$, the state of the total system is $|0\ldots0\rangle\langle 0\ldots0| \otimes \rho_0$. Next, a Hadamard gate is applied to each control qubit, i.e. $|0\rangle \to H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Meanwhile, the parameter is encoded into the model $\rho_\theta = U_t \rho_0 U_t^\dagger$ and the unitary control V is applied. Therefore, at time t, the state of the system is

$$\frac{1}{2^n} \sum_{x,y \in \{0,1\}^{\times n}} |x_1 \ldots x_n\rangle\langle y_1 \ldots y_n| \otimes V \rho_\theta V^\dagger \ , \tag{2.2.42}$$

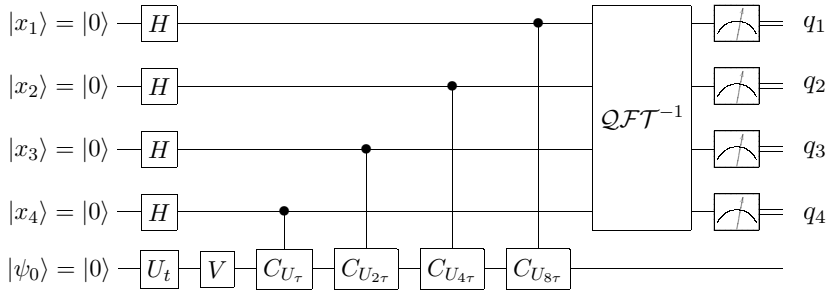


Figure 2.1. Circuit diagram of $\mathcal{M}_n^{(V)}$ with $n = 4$ control qubits. A realistic controlled energy measurement replaces each application of $\mathcal{C}_{U_\tau}$ by m repeated applications of $\Gamma_{U_{\tau/m}}$, defined in Eq. (2.2.56).

where x stands for the generic binary n-string $x_1 \dots x_n$ and y for the binary string $y_1 \dots y_n$.

Next, given an arbitrary unitary U on $\mathcal{H}$, we define the superoperator $\mathcal{C}_U$ as follows,

$$\mathcal{C}_U[\rho] := C_U \rho \, C_U^\dagger \ . \tag{2.2.43}$$

For $l = 1, \dots, n$, the n superoperators $\mathcal{C}_{U_\tau^{2^{l-1}}}$ are applied between the l^{th} control qubit and the main system (τ is a free parameter giving the timescale of the measurement process). Notice that, when $\mathcal{C}_{U_\tau^{2^{l-1}}}$ is applied to $\rho_l := |x_l\rangle \langle y_l| \otimes V \rho_\theta V^\dagger$, one obtains

$$\mathcal{C}_{U_\tau^{2^{l-1}}}[\rho_l] = |x_l\rangle \langle y_l| \otimes U_\tau^{x_l 2^{l-1}} V \rho_\theta V^\dagger \left(U_\tau^\dagger\right)^{y_l 2^{l-1}} \ . \tag{2.2.44}$$

Denoting by $X = x_1 + 2 \cdot x_2 + \ldots + 2^{n-1} \cdot x_n$ the decimal representation of the binary string x, one obtains

$$\frac{1}{2^n} \sum_{X=0}^{2^n-1} \sum_{Y=0}^{2^n-1} |x\rangle \langle y| \otimes U_\tau^X V \rho_\theta V^\dagger (U_\tau^\dagger)^Y \ . \tag{2.2.45}$$

Let us now expand $V \rho_\theta V^\dagger$ on the energy eigenbasis, i.e.

$$V \rho_\theta V^\dagger = \sum_{j=0}^{d-1} \sum_{k=0}^{d-1} c_{jk} \, |\xi_{j,\theta}\rangle \langle \xi_{k,\theta}| \ . \tag{2.2.46}$$

Eq. (2.2.45) then becomes

$$\frac{1}{2^n} \sum_{j,k=0}^{d-1} \sum_{X,Y=0}^{2^n-1} c_{jk} \, e^{-i\tau(X\xi_{j,\theta} - Y\xi_{k,\theta})} \, |x\rangle\langle y| \otimes |\xi_{j,\theta}\rangle\langle \xi_{k,\theta}| \ . \tag{2.2.47}$$

The next step is to apply an inverse quantum Fourier transform $\mathcal{QFT}^{-1}$ on the n control qubits. By definition, $\mathcal{QFT}^{-1}$ acts as follows on the computational basis of $\mathcal{H}_c^{\otimes n}$:

$$\mathcal{QFT}^{-1} |x\rangle = \frac{1}{2^{n/2}} \sum_{Q=0}^{2^n-1} e^{-\frac{2\pi i X Q}{2^n}} \, |q\rangle \ . \tag{2.2.48}$$

After application of $\mathcal{Q}\mathcal{F}\mathcal{T}^{-1}$, the total state of the system is

$$\frac{1}{2^{2n}} \sum_{j,k=0}^{d-1} \sum_{X,Y=0}^{2^n-1} \sum_{Q,P=0}^{2^n-1} \tilde{c}_{jk} \, |q\rangle \langle p| \otimes |\xi_{j,\theta}\rangle \langle \xi_{k,\theta}| \ . \tag{2.2.49}$$

where

$$\tilde{c}_{jk} = c_{jk} \, e^{-iX\left(\tau\xi_{j,\theta} + \frac{2\pi Q}{2^n}\right)} \, e^{iY\left(\tau\xi_{k,\theta} + \frac{2\pi P}{2^n}\right)} \ . \tag{2.2.50}$$

The last step is to perform a measurement of the n control qubits in the computational basis. The probability $\mathrm{Pr}_\theta(q)$ of obtaining as outcome the binary string q is

$$\mathrm{Pr}_\theta(q) = \frac{1}{2^{2n}} \sum_{j=0}^{d-1} \sum_{X,Y=0}^{2^n-1} \mathrm{Pr}_\theta(\xi_{j,\theta}) \, e^{-i(X-Y)\alpha_{j,Q}} \ , \tag{2.2.51}$$

where

$$\alpha_{j,Q} := \tau\xi_{j,\theta} + \frac{2\pi Q}{2^n} \ , \qquad \mathrm{Pr}_\theta(\xi_{j,\theta}) = \langle \xi_{j,\theta} | V \rho_\theta V^\dagger | \xi_{j,\theta} \rangle \ . \tag{2.2.52}$$

By algebraic manipulation, Eq. (2.2.51) can also be written as

$$\mathrm{Pr}_\theta(q) = \sum_{j=0}^{d-1} \mathrm{Pr}_\theta(\xi_{j,\theta}) \left(\frac{1}{2^n} \frac{\sin(2^n \alpha_{j,Q}/2)}{\sin(\alpha_{j,Q}/2)} \right)^2 \ . \tag{2.2.53}$$

In the limit $n \to \infty$, $\mathrm{Pr}_\theta(q)$ converges to the probability $\mathrm{Pr}_\theta(\xi_{j,\theta})$, which corresponds to a controlled energy measurement $\mathcal{M}_\theta^{(V)}$.

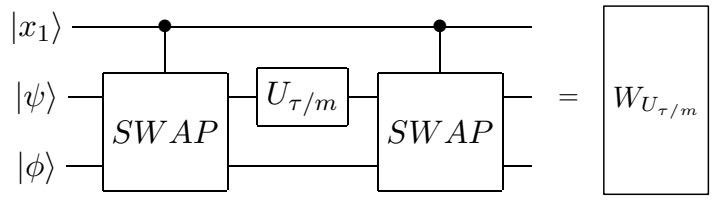


Figure 2.2. Circuit diagram of $W_{U_{\tau/m}}$.

A realistic controlled energy measurement $\mathcal{M}_{n,m}^{(V)}$ is obtained from $\mathcal{M}_n^{(V)}$ by implementing the controlled time-evolution operator C_{U_t} via a quantum subroutine, known as *universal controllization* [39]. For notational simplicity, let us consider the case $l = 1$: the problem is to approximate the action of $\mathcal{C}_{U_\tau}$ on the state $\rho_1 = |x_1\rangle \langle y_1| \otimes V \rho_\theta V^\dagger$. The operator $\mathcal{C}_{U_\tau}$ is replaced by m applications of the superoperator $\Gamma_{U_{\tau/m}}$, constructed as follows. First, an ancilla, having the same dimensionality as the main system, is introduced. The total Hilbert space is $\mathcal{H}_c^{\otimes n} \otimes \mathcal{H} \otimes \mathcal{H}_a$, with $\mathcal{H}_a = \mathbb{C}^d$. The ancilla is prepared in the maximally mixed state. Therefore, the state of the first control qubit, the main system and the ancilla before application of $\mathcal{C}_{U_\tau}$ is $\rho_1' = |x_1\rangle \langle y_1| \otimes V \rho_\theta V^\dagger \otimes \mathbb{I}_d/d$. Let us now define the following quantum operation,

$$W_{U_\tau} := C_{SWAP} (\mathbb{I}_2 \otimes U_\tau \otimes \mathbb{I}_d) \, C_{SWAP} \ , \tag{2.2.54}$$

where C_{SWAP} is the controlled-SWAP gate acting on $\mathcal{H}_c \otimes \mathcal{H} \otimes \mathcal{H}_a$ as

$$C_{SWAP}(|0\rangle \otimes |\psi\rangle \otimes |\phi\rangle) = |0\rangle \otimes |\phi\rangle \otimes |\psi\rangle \ , \qquad C_{SWAP}(|1\rangle \otimes |\psi\rangle \otimes |\phi\rangle) = |0\rangle \otimes |\psi\rangle \otimes |\phi\rangle \ . \tag{2.2.55}$$

The key remark is that implementation of W_{U_τ} does not require knowledge of the Hamiltonian, as it only makes use of the *uncontrolled* version of the time-evolution operator U_τ. We now subdivide τ into m subintervals of length τ/m. During each subinterval, $W_{U_{\tau/m}}$ is applied; then the ancilla is discarded; finally, the ancilla is refreshed to its initial state. For instance, after the first interval, one obtains $\Gamma_{U_{\tau/m}}[\rho_1] \otimes \mathbb{I}_d/d$, where

$$\Gamma_{U_{\tau/m}}[\rho_1] := \mathrm{tr}_{\mathcal{H}_a} \left(W_{U_{\tau/m}} \rho_1' \, W_{U_{\tau/m}}^\dagger \right) \ . \tag{2.2.56}$$

A simple computation reveals that

$$\Gamma_{U_{\tau/m}}[\rho_1] = \frac{1}{d} \, \mathrm{tr} \left(U_{\tau/m}^{y_1 - x_1} \right) \, \mathcal{C}_{U_{\tau/m}}[\rho_1] \ . \tag{2.2.57}$$

For future convenience, we write

$$\frac{1}{d} \, \mathrm{tr} \left(U_{\tau/m} \right) = a_{\tau/m} \, e^{i \phi_{\tau/m}} \ , \tag{2.2.58}$$

where $a_{\tau/m} \in \mathbb{R}^+$ and $\phi_{\tau/m} \in \mathbb{R}$. Note that, since $x_1 - y_1 \in \{-1, 0, 1\}$, one can write

$$\Gamma_{U_{\tau/m}}^m[\rho_1] = a_{\tau/m}^{|x_1 - y_1| m} \, e^{i (y_1 - x_1) m \phi_{\tau/m}} \, \mathcal{C}_{U_\tau}[\rho_1] \ . \tag{2.2.59}$$

Universal controllization thus replaces $\mathcal{C}_{U_\tau}$ with $\Gamma_{U_{\tau/m}}^m$. In the limit $m \to \infty$, it can be proven that the error

$$\epsilon_m := \left[\mathrm{tr} \left(U_{\tau/m} \right) / d \right]^m - 1 \tag{2.2.60}$$

tends to zero. A realistic controlled energy measurement is obtained by substituting each application of $\mathcal{C}_{U_\tau^{2^{l-1}}}$ by $2^{l-1} m$ applications of $\Gamma_{U_{\tau/m}}$. For instance, instead of Eq. (2.2.45), one would have

$$\frac{1}{2^n} \sum_{X,Y=0}^{2^n - 1} \pi_{X,Y} e^{i (Y - X) m \phi_{\tau/m}} \, |x\rangle\langle y| \otimes U_\tau^X V \rho_\theta V^\dagger (U_\tau^\dagger)^Y \ , \tag{2.2.61}$$

where

$$\pi_{X,Y} := \prod_{l=1}^{n} a_{\tau/m}^{|x_l - y_l| 2^{l-1} m} \ . \tag{2.2.62}$$

After applying the inverse quantum Fourier transform and measuring in the computational basis, the probability of obtaining the outcome $q \in \{0,1\}^{\times n}$ is

$$\mathrm{Pr}_\theta(q) = \frac{1}{2^{2n}} \sum_{j=0}^{d-1} \mathrm{Pr}_\theta(\xi_{j,\theta}) \sum_{X,Y=0}^{2^n - 1} \pi_{X,Y} \, e^{i (Y - X) \beta_{j,Q}} \ , \tag{2.2.63}$$

with

$$\beta_{j,Q} := \alpha_{j,Q} + m\phi_{\tau/m} \ . \tag{2.2.64}$$

Eq. (2.2.63) can be further expanded by rewriting it as follows,

$$\begin{aligned} \Pr_\theta(q) &= \frac{1}{2^{2n}} \sum_{j=0}^{d-1} \Pr_\theta(\xi_{j,\theta}) \prod_{l=1}^{n} \sum_{u,v=0}^{1} a_{\tau/m}^{|u-v|2^{l-1}m} e^{i(v-u)2^{l-1}\beta_{j,Q}} \\ &= \frac{1}{2^n} \sum_{j=0}^{d-1} \Pr_\theta(\xi_{j,\theta}) \prod_{l=1}^{n} \left[1 + a_{\tau/m}^{2^{l-1}m} \cos\left(2^{l-1}\beta_{j,Q}\right)\right] \ . \end{aligned} \tag{2.2.65}$$

If $m \to \infty$, then $\phi_{\tau/m} \to 0$ and $a_{\tau/m} \to 1$, so that Eq. (2.2.65) converges to Eq. (2.2.53). In conclusion, a realistic controlled energy measurement allows to approximate to any desired precision a controlled energy measurement $\mathcal{M}_\theta^{(V)}$, without requiring any a priori knowledge about the parameter θ.

2.2.5 Examples

In this section, we work out a collection of examples. For each example, we compute the QFI $\mathcal{F}_Q(\theta)$ and compare it with $\mathcal{G}(\theta)$. We will find that, in general, $\mathcal{G}(\theta)$ majorizes $\mathcal{F}_Q(\theta)$, thus controlled energy measurements lead to a precision enhancement. Moreover, we study numerically the performance of realistic controlled energy measurements $\mathcal{M}_{n,m}^{(V^{(opt)})}$. From the previous section, as $n, m \to \infty$, $\mathcal{M}_{n,m}^{(V^{(opt)})}$ converges to $\mathcal{M}_\theta^{(V^{(opt)})}$, and thus its Fisher information also converges to $\mathcal{G}_\theta$. We will show that, already for relatively small values of n and m, realistic controlled energy measurements perform very close to the the ultimate bound $\mathcal{G}_\theta$.

Example 2.2.4. *The parameter to be estimated is the polar angular direction θ of an external magnetic field of known magnitude B. The probing system is a two-level atom, with Hilbert space $\mathcal{H} = \mathbb{C}^2$. The Hamiltonian is $H_\theta = \omega(\cos\theta\,\sigma_z + \sin\theta\,\sigma_x)$, where the energy splitting ω is proportional to B and thus is assumed to be known. At time $t = 0$, the atom is initialized in its ground state: $|\psi_0\rangle = |0\rangle$. At the generic time t, the state of the probe is $|\psi_\theta\rangle = U_t\,|\psi_0\rangle$, with $U_t := \exp(-iH_\theta t)$.*

If an experimentalist is constrained to perform regular measurements, the best performance she can achieve is quantified by the QFI:

$$\mathcal{F}_Q(\theta) = 4\sin^2(\omega t) - \sin^2(2\omega t)\sin^2\theta \ . \tag{2.2.66}$$

Optimizing also over the initial preparation,

$$\max_{|\psi_0\rangle} \mathcal{F}_Q(\theta) = 4\sin^2(\omega t) \ . \tag{2.2.67}$$

If instead the experimentalist is allowed to implement only controlled energy measurements, the maximum Fisher information that she can extract is given by $\mathcal{G}(\theta)$. To compute $\mathcal{G}(\theta)$, one first computes the matrix S, built from the eigenvectors of H_θ:

$$S = \begin{pmatrix} -\mathrm{sgn}\left[\cos\left(\frac{\theta}{2}\right)\right]\sin\left(\frac{\theta}{2}\right) & \mathrm{sgn}\left[\cos\left(\frac{\theta}{2}\right)\right]\cos\left(\frac{\theta}{2}\right) \\ \mathrm{sgn}\left[\sin\left(\frac{\theta}{2}\right)\right]\cos\left(\frac{\theta}{2}\right) & \mathrm{sgn}\left[\sin\left(\frac{\theta}{2}\right)\right]\sin\left(\frac{\theta}{2}\right) \end{pmatrix} , \tag{2.2.68}$$

where $\mathrm{sgn}(x) := |x|/x$. *Its generator* $\mathfrak{g}_\theta[S]$ *reads:*

$$\mathfrak{g}_\theta[S] = \begin{pmatrix} 0 & -\frac{i}{2}\,\mathrm{sgn}(\sin\theta) \\ \frac{i}{2}\,\mathrm{sgn}(\sin\theta) & 0 \end{pmatrix} . \tag{2.2.69}$$

The extremal eigenvectors of $\mathfrak{g}_\theta[S]$ *are*

$$|\lambda_1(\mathfrak{g}_\theta[S])\rangle = \frac{1}{\sqrt{2}}\,(-i,\ 1)^t\ , \qquad |\lambda_2(\mathfrak{g}_\theta[S])\rangle = \frac{1}{\sqrt{2}}\,(i,\ 1)^t\ . \tag{2.2.70}$$

Since condition (2.2.39) *is satisfied,* $\mathcal{G}(\theta)$ *can be obtained via Prop.* (2.2.3). *The explicit expressions for* U_t *and its generator are*

$$U_t = \begin{pmatrix} A & B \\ B & A^* \end{pmatrix} , \qquad \mathfrak{g}_\theta[U_t] = \begin{pmatrix} -C & D \\ D^* & C \end{pmatrix} , \tag{2.2.71}$$

where

$$A = \cos\omega t - i\cos\theta\sin\omega t\ , \qquad C = \frac{1}{2}\sin\theta\sin 2\omega\ ,$$
$$B = -i\sin\theta\sin\omega t\ , \qquad D = (\cos\theta\cos\omega t - i\sin\omega t)\sin\omega t\ .$$

One thus obtains

$$\mathcal{G}(\theta) = (2|\sin(\omega t)| + 1)^2\ ; \tag{2.2.72}$$

see Fig. 2.3 for a comparison with the QFI.

Finally, we study numerically the performance of $\mathcal{M}_{n,m}^{(V^{(opt)})}$ *(with* $V^{(opt)}$ *the optimal control of Eq.* (2.2.18)*). Recall that* n *is the number of ancillary qubits needed to implement the phase estimation algorithm, while* m *is the number of subintervals the timescale* τ *is subdivided into. During each subinterval, the action of the controlled time-evolution operator* $\mathcal{C}_{U_\tau}$ *is approximated by applying* m *times the superoperator* $\Gamma_{U_{\tau/m}}$ *of Eq.* (2.2.56). *As* $n, m \to \infty$, *the probability measure of* $\mathcal{M}_{n,m}^{(V^{(opt)})}$ *converges to that of the optimal controlled energy measurement. The two panels of Fig. 2.4 show the performance of* $\mathcal{M}_{n,m}^{(V^{(opt)})}$ *for different values of* n *and* m. *Reasonably small values of the two parameters (i.e.* $n = 6$, $m = 3$*) are already enough to come close to the ultimate bound* $\mathcal{G}(\theta)$.

Example 2.2.5. *The parameter to be estimated is the component of a magnetic field along the* x *direction. The probing system is again a two-level atom. The Hamiltonian is*

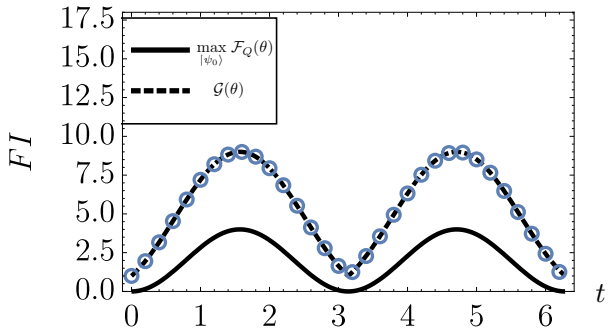


Figure 2.3. Comparison between the optimal Braunstein-Caves measurement and the optimal controlled energy measurement, for the estimation of the polar angular direction of a magnetic field via a qubit probe. The solid line is the QFI, optimized over the initial preparation, while the dashed line corresponds to $\mathcal{G}(\theta)$, computed by Eq. (2.2.72). The circular marks correspond to values of $\mathcal{G}(\theta)$ computed by numerical optimization from its definition (2.2.2).

$H_\theta = -\omega\sigma_z + \theta\sigma_x$, with eigenvalues $\pm\Omega_\theta$ and $\Omega_\theta := \sqrt{\omega^2+\theta^2}$. We report the matrices U_t and S, with their corresponding generators. For U_t and $\mathfrak{g}_\theta[U_t]$, one obtains

$$U_t = \begin{pmatrix} A & B \\ B & A^* \end{pmatrix} , \qquad \mathfrak{g}_\theta[U_t] = \begin{pmatrix} -C & D \\ D^* & C \end{pmatrix} , \tag{2.2.73}$$

where

$$A = \cos(\Omega_\theta t) + \frac{i\omega\sin(\Omega_\theta t)}{\Omega_\theta} , \qquad C = -\frac{\omega\theta\left[\sin(2\,\Omega_\theta t) - 2\,\Omega_\theta t\right]}{2\,\Omega_\theta^3} ,$$
$$B = -\frac{i\theta\sin(\Omega_\theta t)}{\Omega_\theta} , \qquad D = \frac{\sin(2\,\Omega_\theta t)\omega^2 - i\Omega_\theta\cos(2\,\Omega_\theta t)\omega + \Omega_\theta\left(2t\theta^2 + i\omega\right)}{2\,\Omega_\theta^3} .$$

For the matrix S and its generator,

$$S = \frac{1}{\sqrt{2\,\Omega_\theta}} \begin{pmatrix} -\frac{\omega+\Omega_\theta}{\sqrt{\Omega_\theta+\omega}} & \frac{\theta}{\sqrt{\Omega_\theta+\omega}} \\ \frac{\theta}{\sqrt{\Omega_\theta+\omega}} & \frac{\theta}{\sqrt{\Omega_\theta-\omega}} \end{pmatrix} , \qquad \mathfrak{g}_\theta[S] = \begin{pmatrix} 0 & \frac{i\omega}{2\theta^2} \\ -\frac{i\omega}{2\theta^2} & 0 \end{pmatrix} . \tag{2.2.74}$$

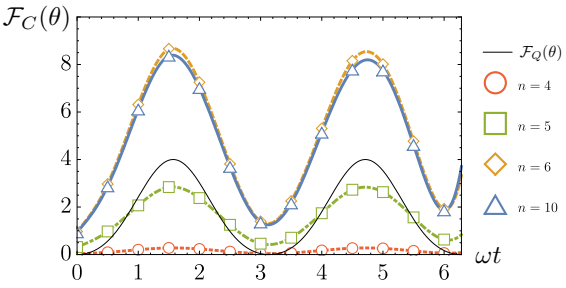
The maximum QFI is

$$\max_{|\psi_0\rangle} \mathcal{F}_Q(\theta) = \frac{2}{\Omega_\theta^4}\left[2\,\Omega_\theta^2\,t^2\theta^2 - \omega^2\cos(2\,\Omega_\theta t) + \omega^2\right] . \tag{2.2.75}$$

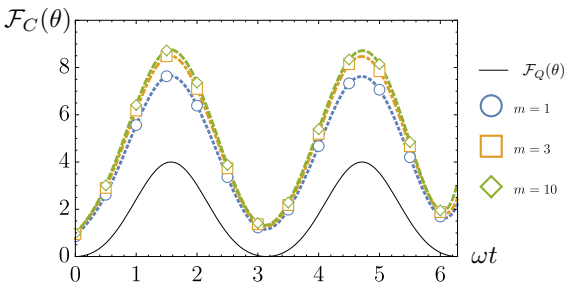
Since the eigenvectors of $\mathfrak{g}_\theta[S]$ satisfy condition (2.2.39), $\mathcal{G}(\theta)$ can be computed directly, whcih gives

$$\mathcal{G}(\theta) = \left(\frac{\omega}{\Omega_\theta^2} + \frac{\sqrt{2\left[2\,\Omega_\theta^2\,t^2\theta^2 - \omega^2\cos(2\,\Omega_\theta t) + \omega^2\right]}}{\Omega_\theta^2}\right)^2 . \tag{2.2.76}$$

A comparison similar to that of Fig. 2.3 is shown in Fig. 2.5.



(a) Fisher information of realistic controlled energy measurement, for different values of n and fixed $m = 5$.



(b) Fisher information of realistic controlled energy measurements, for different values of m and fixed $n = 6$.

Figure 2.4. Fisher information of realistic controlled energy measurements, for different values of n and m. Each marker represents the maximum Fisher information, computed numerically over the family of realistic controlled energy measurements for given n, m, τ and interrogation time t. The curves are obtained by interpolation. The thin solid curve corresponds to the QFI of Eq. (2.2.66). Both plots are obtained for $\omega = 1$ and $\tau = 0.1$ (in the natural units of the problem), while the true value of the parameter is taken to be $\theta = \pi/4$.

Example 2.2.6. *The parameter to be estimated is the magnitude of a weak magnetic field. The probing system is an NV-center in diamond. An NV center consists of a nitrogen atom (N) inside a diamond crystal lattice, having a vacancy (V) in one of its neighboring sites. Two different kinds of the defect are known: the neutral state NV_0 and the negatively-charged state NV_-. The latter is the most interesting for metrological applications. In fact, the NV_- form provides a spin triplet state which can be initialized, manipulated with long coherence time and read out by purely optical means [3]. Neglecting the interactions with the surrounding nuclear spins, the Hamiltonian H_{NV} for the triplet state is*

$$H_{NV} = \mu\, \boldsymbol{B} \cdot \boldsymbol{S} + D\, S_z^2 + E\, (S_x^2 - S_y^2)\,, \tag{2.2.77}$$

where $\boldsymbol{B}$ is the external magnetic field and $\boldsymbol{S} = (S_x, S_y, S_z)$ is a vector whose elements are the three spin 1 matrices:

$$S_x = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad S_y = \sqrt{2}i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad S_z = 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Moreover, $D \approx \pi \times 1.44\,\mathrm{GHz}$, $E \approx \pi \times 50\,\mathrm{kHz}$ and μ is the Bohr magneton. In the weak magnetic field regime, the transversal components B_x and B_y can be neglected compared to the component B_z, which is aligned along the NV-center defect axis. Relabelling B_z as θ, the Hamiltonian becomes

$$H_\theta = \mu\theta S_z + D\, S_z^2 + E\, (S_x^2 - S_y^2)\,. \tag{2.2.78}$$

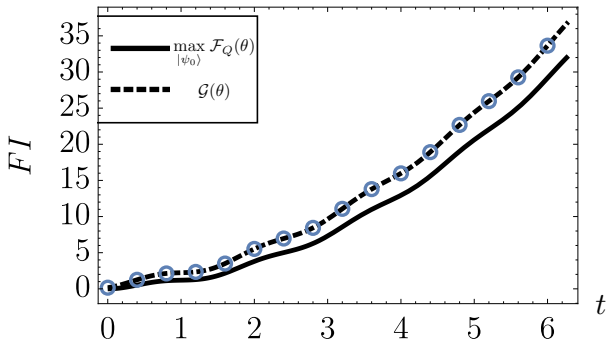


Figure 2.5. Comparison between the optimal Braunstein-Caves measurement and the optimal controlled energy measurement, for the estimation of one component of a magnetic field via a qubit probe. The circular marks correspond to values of $\mathcal{G}(\theta)$ computed by numerical optimization from its definition (2.2.2).

The maximum QFI is

$$\max_{|\psi_0\rangle} \mathcal{F}_Q(\theta) = \frac{8\mu^2 \left[2\theta^2\mu^2 t^2 \chi^2 + E^2 - E^2 \cos\left(4\chi t\right)\right]}{\chi^4} , \tag{2.2.79}$$

where $\chi := \sqrt{\theta^2\mu^2 + 4E^2}$. *Instead,* $\mathcal{G}(\theta)$ *is given by*

$$\mathcal{G}(\theta) = \left(\frac{2E\mu}{\chi^2} + 2\sqrt{2}\mu \frac{\sqrt{2\theta^2\mu^2 t^2 \chi^2 + E^2 - E^2 \cos\left(4\chi t\right)}}{\chi^2} \right)^2 . \tag{2.2.80}$$

A comparison is shown in Fig. 2.6.

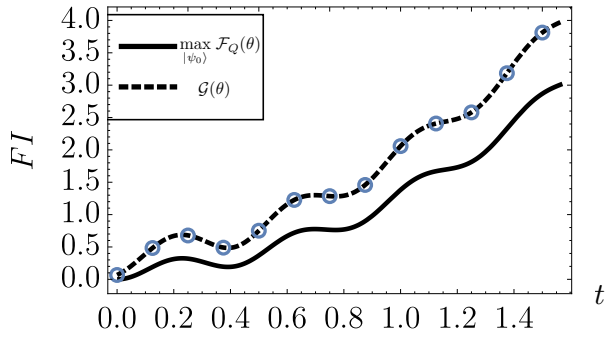


Figure 2.6. Comparison between the optimal Braunstein-Caves measurement and the optimal controlled energy measurement for the estimation of the magnitude of a weak magnetic field via an NV-center in diamond. The circular marks correspond to values of $\mathcal{G}(\theta)$ computed by numerical optimization from its definition (2.2.2).

2.3 Conclusions

In this chapter, we have considered non-regular measurements. In particular, we have introduced the family of controlled energy measurements and applied them to Hamiltonian parameter estimation problems. Each such measurement is obtained by applying a unitary control and then performing a projective energy measurement. It is non-regular whenever the Hamiltonian depends non-linearly on the parameter θ. We have then maximized the Fisher information over the set of controlled energy measurements and initial preparations (see Sect. 2.2.2). The maximum, denoted by $\mathcal{G}(\theta)$, can be computed by the closed-form expression given in Eq. (2.2.40). We have discussed how controlled energy measurements can be implemented in realistic scenarios, via an adaptation of the quantum phase estimation algorithm (see Sect. 2.2.4). Finally, we have worked out a collection of examples to clarify our methods (see Sect. 2.2.5). We have discovered that a precision enhancement, compared with regular measurements, is often possible. In particular, we have emphasized that, if the parameter is not a simple phase, the quantum Fisher information no longer necessarily embodies the ultimate precision limit.

Chapter 3

Information-disturbance trade-off in qubit thermometry

This chapter establishes the existence of a trade-off relation between the information on the temperature of a quantum thermometer, extractable via a quantum measurement, and the corresponding disturbance caused by the measurement [59].

3.1 Preliminaries

A fundamental feature of quantum mechanics is that extracting information from a system disturbs its original state. In general, the greater the amount of information extracted, the greater the disturbance caused by the measurement. This is the intuition behind the original derivation of the uncertainty principle by Heisenberg [44]. Over the years, the trade-off between information and disturbance has been given several quantitative formulations from a quantum information perspective [45–52]. In this chapter, we study the information/disturbance trade-off in the context of parameter estimation. In particular, we focus on the special case of estimating the temperature of a thermal bath via a two-level quantum system, i.e. qubit thermometry [53–56].

The typical thermometry protocol consists of letting the probe thermalize with the bath, performing a suitable measurement on the probe, and finally processing the outcome via an estimator. Given the Hamiltonian H of the system, the statistical model is the thermal model ρ_β of Eq. (1.4.3) and the parameter to be estimated is the inverse temperature $\beta \in \mathbb{R}^+$. Since β is a pure phase parameter, the ultimate precision is quantified by the quantum Fisher information $\mathcal{F}_Q(\beta)$, which evaluates to

$$\mathcal{F}_Q(\beta) = \mathrm{tr}[\rho_\beta \, (H - \langle H \rangle_\beta)^2] = \mathrm{Var}_{\rho_\beta}(H) \; . \tag{3.1.1}$$

The symmetric logarithmic derivative of ρ_β is given by $L_\beta = \langle H \rangle_\beta - H$, where $\langle H \rangle_\beta :=$

$\mathrm{tr}(\rho_\beta H)$ is the energy expectation value. The optimal Braunstein-Caves measurement is a projective energy measurement. Therefore, assuming one is interested only in minimizing the mean square error, the optimal thermometry protocol is as follows: wait until the probe thermalizes with the bath, perform an energy measurement, process the outcome via an efficient estimator.[1]

However, a projective energy measurement introduces a significant amount of disturbance. We have not yet defined precisely how to quantify disturbance, but it will turn out that a projective energy measurement maximizes all disturbance quantifiers we are going to employ. For example, a natural disturbance quantifier is the decrease of the quantum Fisher information between the post- and pre-measurement states (see Eq. (3.1.2) below). After an energy measurement, the system is found in an eigenstate of H, which is β-independent. Any repeated measurement on the probe thus extracts zero information. Therefore, an energy measurement is maximally disturbing with respect to this quantifier. After the measurement, the experimentalist has to put the thermometer in contact with the thermal bath and wait again for thermalization to set in.

On the other hand, a less disturbing and less time-consuming estimation protocol would involve implementing a measurement that, while suboptimal with respect to the QFI limit, nonetheless does not completely destroys all information on β. One expects a trade-off relation to exist between the Fisher information that is extracted by a measurement, and the disturbance it causes. By definition, *efficient* measurements are those measurements that, for given extracted information, minimize the disturbance, i.e. they make up the frontier of the information/disturbance trade-off region. In the following, we introduce four different disturbance quantifiers, that are variously employed in quantum information, quantum thermodynamics and quantum information geometry. We study their information/disturbance trade-off curves and, in particular, fully characterize the family of efficient measurements.

3.1.1 Disturbance quantifiers

In a parameter estimation context, there is a natural choice about how to quantify information: the Fisher information. It satisfies all the properties that are required of an information quantifier (e.g. monotonicity, invariance under sufficient statistics, additivity, covariance under reparametrization) and has a direct statistical interpretation (via the Cramér-Rao theorem 1.2.8). The choice of the disturbance quantifier is, however, less clear-cut. Different quantifiers have been proposed in the literature, which capture

[1] A thermal model is of the exponential type, therefore equality can be achieved in the classical Cramér-Rao bound even in the non-asymptotic regime.

different aspects of the disturbance caused by a measurement. We will make use of the following four quantifiers, denoted by $\mathcal{D}^{(\alpha)}(\beta)$, with $\alpha \in \{\Delta, F, \tau, \pi\}$:

(Q1) The Δ-disturbance $\mathcal{D}^{(\Delta)}(\beta)$ is the average *information loss* [57], i.e. the decrease of the quantum Fisher information induced by a measurement. Assuming the measurement outcomes take values in a sample space $\mathcal{I}$, with probabilities $\Pr_\beta(i) := \mathrm{tr}(\rho_\beta \Pi_i)$, one has

$$\mathcal{D}^{(\Delta)}(\beta) := \mathcal{F}_Q[\rho_\beta] - \langle \mathcal{F}_Q[\rho_{\beta|i}] \rangle \ , \tag{3.1.2}$$

where $\langle \mathcal{F}_Q[\rho_{\beta|i}] \rangle$ is the average QFI of the post-measurement state, i.e.

$$\langle \mathcal{F}_Q[\rho_{\beta|i}] \rangle := \sum_{i \in \mathcal{I}} \Pr_\beta(i) \mathcal{F}_Q[\rho_{\beta|i}] \ . \tag{3.1.3}$$

(Q2) The F-disturbance $\mathcal{D}^{(F)}(\beta)$ is the average fidelity-based distance [45] between the initial and the post-measurement state, i.e.

$$\mathcal{D}^{(F)}(\beta) := 1 - \sum_{i \in \mathcal{I}} \Pr_\beta(i) F^2(\rho_\beta, \rho_{\beta|i}) \ , \tag{3.1.4}$$

where $F(\rho_\beta, \rho_{\beta|i})$ is the fidelity,

$$F(\rho_\beta, \rho_{\beta|i}) = \mathrm{tr}\left[\sqrt{\sqrt{\rho_\beta}\, \rho_{\beta|i} \sqrt{\rho_\beta}} \right] \ . \tag{3.1.5}$$

(Q3) The τ-disturbance $\mathcal{D}^{(\tau)}(\beta)$ is the average spectral temperature variation, i.e.

$$\mathcal{D}^{(\tau)}(\beta) := \sum_{i \in \mathcal{I}} \Pr_\beta(i) \ \left| \beta - \tau(\rho_{\beta|i}) \right| \ , \tag{3.1.6}$$

where the spectral temperature $\tau(\rho)$ of a quantum state is defined as follows [58],

$$\tau(\rho) := \frac{1}{\sigma} \sum_{i=0}^{d-1} \left(\frac{\Pr_\beta(\xi_{i+1}) + \Pr_\beta(\xi_i)}{2} \right) \frac{\log(\Pr_\beta(\xi_i)/\Pr_\beta(\xi_{i+1}))}{\xi_{i+1} - \xi_i} \ ; \tag{3.1.7}$$

$\sigma := 1 - [\Pr_\beta(\xi_0) + \Pr_\beta(\xi_{d-1})]/2$ and $\{\xi_i\}_{i=0}^{d-1}$ is the energy spectrum of H, assumed to be non-degenerate. The spectral temperature generalizes the notion of thermodynamical temperature to out-of-equilibrium states.

(Q4) The π-disturbance $\mathcal{D}^{(\pi)}(\beta)$ is the average quantum relative entropy between the pre-measurement state and the I-projection of the post-measurement state [12]. Explicitly,

$$\mathcal{D}^{(\pi)}(\beta) := \sum_{i \in \mathcal{I}} \Pr_\beta(i)\, S(\rho_{\eta_i^{(\pi)}} || \rho_\beta) \ . \tag{3.1.8}$$

The I-projection $\rho_{\eta_i^{(\pi)}}$ is the thermal state whose temperature is defined as

$$\eta_i^{(\pi)} := \arg\min_\eta S(\rho_\eta || \rho_{\beta|i}) \ . \tag{3.1.9}$$

3.1.2 Qubit thermometry

A qubit thermometer is a two-level quantum system, with Hamiltonian $H = \delta\sigma_z/2$ (where δ is the energy gap). After thermal equilibrium sets in, the qubit is described by the thermal statistical model:

$$\rho_\beta = \frac{1}{Z_\beta}\,\mathrm{diag}(e^{-\beta\delta/2}, e^{\beta\delta/2})\ , \qquad \text{with} \quad Z_\beta := 2\cosh(\beta\delta/2)\ . \tag{3.1.10}$$

Information on β is extracted via a suitable measurement $\mathcal{M}$, yielding an outcome $i \in \mathcal{I}$. We are going to make the following assumptions on $\mathcal{M}$: the measurement has a binary sample space $\mathcal{I} = \{0, 1\}$, it is fine-grained (i.e. there is only one measurement operator for each measurement outcome) and bare (i.e. the post-measurement state is determined via the Lüders instrument). Such restrictions are quite natural and are suggested by considerations of simplicity. For instance, the latter two assumptions have no effect on the computation of the information for given probability operators, but simplify significantly the discussion about the corresponding disturbance. The family of all such measurements, denoted by $\mathscr{B}$, can be parametrized as in the following proposition.

Proposition 3.1.1. *The generic binary measurement $\mathcal{M} \in \mathscr{B}$ has probability operators $\{\Pi_0, \Pi_1\}$ of the form*

$$\Pi_0 = w\mathbb{1}_2 + x\sigma_x + y\sigma_y + z\sigma_z\ , \qquad \Pi_1 = \mathbb{1}_2 - \Pi_0\ , \tag{3.1.11}$$

where $0 < w \leq 1/2$ and $\sqrt{x^2+y^2+z^2} \leq w$. Its measurement operators are $\{M_0, M_1\}$, where each M_i is the principal square root of Π_i. The post-measurement state is given by the Lüders instrument, i.e.

$$\rho_{\beta|i} = \frac{M_i \rho_\beta M_i^\dagger}{\mathrm{Pr}_\beta(i)}\ , \qquad \textit{with} \quad \mathrm{Pr}_\beta(i) = \mathrm{tr}(\rho_\beta \Pi_i)\ . \tag{3.1.12}$$

Proof. Let us expand Π_0 as in Eq. (3.1.11) on the basis of $\mathsf{Her}_2(\mathbb{C})$ made up of the identity matrix $\mathbb{1}_2$ and the three Pauli matrices $\sigma_x, \sigma_y, \sigma_z$. Notice that $\Pi_1 = \mathbb{1}_2 - \Pi_0$. A necessary and sufficient condition for $\{\Pi_0, \Pi_1\}$ to be the probability operators of a physical measurement is that they are positive semi-definite matrices. For 2×2 matrices, this is equivalent to requiring that both their trace and determinant are nonnegative. By explicit computation, one obtains the constraints $0 < w \leq 1$, $\sqrt{x^2+y^2+z^2} \leq \min(w, 1-w)$. By assuming without loss of generality that $\mathrm{tr}(\Pi_0) \leq 1$, the previous constraints simplify to $0 < w \leq 1/2$ and $\sqrt{x^2+y^2+z^2} \leq w$. □

Therefore, each measurement $\mathcal{M} \in \mathscr{B}$ is represented by a quadruple (w, x, y, z), referred to as its cartesian coordinates. It is also convenient to introduce conical coordinates $(w, \lambda, \theta, \varphi)$, defined as follows:

$$x = \lambda w \sin\theta\cos\varphi\ , \qquad y = \lambda w \sin\theta\sin\varphi\ , \qquad z = \lambda w\cos\theta\ . \tag{3.1.13}$$

with the constraints $0 < w \leq 1/2,\ 0 \leq \lambda \leq 1,\ 0 \leq \theta \leq \pi,\ 0 \leq \varphi \leq 2\pi$. Graphically, the family of binary measurements $\mathscr{B}$ can be represented as a cone in the Euclidean space $\mathbb{R}^4$, with (w, x, y, z) as cartesian coordinates (see Fig. 3.1a). Each cross-section of the cone with a hyperplane of constant w is a 3-dimensional ball of radius w, with θ the polar angle and φ the azimuthal angle. For future convenience, we define the following special subfamilies of measurements:

(S1) *Projective measurements*: Let us define a purity parameter γ,

$$\gamma := \mathrm{tr}(\Pi_0^2) = 2(1 + \lambda^2)w^2 \ . \tag{3.1.14}$$

Projective measurements maximize γ. Their probability operators are rank-1 projectors. Graphically, they correspond to the sphere $w = 1/2$, $\lambda = 1$.

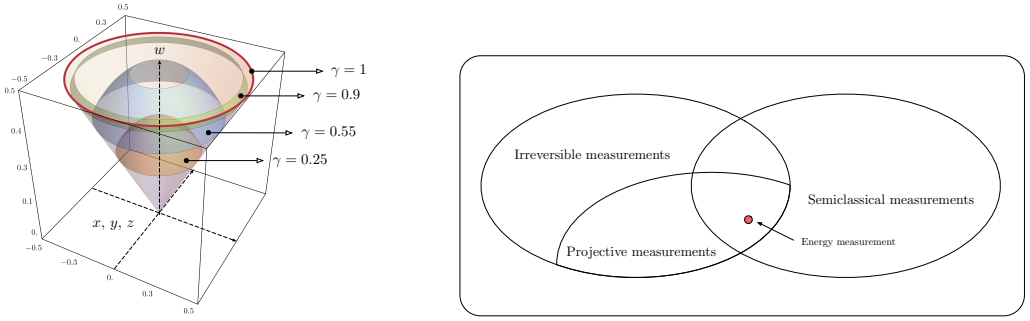
(S2) *Irreversible measurements*: A measurement is irreversible if at least one of its probability operators is a non-invertible matrix. Graphically, they correspond to the surface of the cone $\lambda = 1$.

(S3) *Semiclassical measurements*: A measurement is semiclassical if all its probability operators are diagonal in the eigenbasis of ρ_β, i.e. $[\Pi_i, \rho_\beta] = [\Pi_i, \Pi_{i'}] = 0$, $\forall i, i' \in \mathcal{I}$. One can prove that, as a consequence, the conditional state after sequential semiclassical measurements does not depend on the order they are performed. Let us introduce a non-commutativity parameter $\chi := \sin\theta$. It is easy to check that $[\Pi_i, \rho_\beta] = 0$ if and only if $\chi = 0$ (or, trivially, for *uninformative* measurements, which are defined by the condition $\lambda = 0$). Semiclassical measurements thus minimize χ, whereas *non-classical* measurements maximize it. Graphically, by fixing the values of λ and w one obtains a sphere, with the latter corresponding to its two poles and the former to its equator.

(S4) *Efficient measurements*: A measurement is said to be α-efficient, with respect to the disturbance quantifier $\mathcal{D}^{(\alpha)}(\beta)$, if it minimizes $\mathcal{D}^{(\alpha)}(\beta)$ for fixed value of the Fisher information.

3.2 Information/disturbance trade-off curves

In this section, we study the quantitative relation between information and disturbance, for each of the four disturbance quantifiers $\mathcal{D}^{(\alpha)}(\beta)$ ($\alpha \in \{\Delta, F, \tau, \pi\}$). We pay special attention, in particular, to the determination of the subset of α-efficient measurements.



(a) (b)

Figure 3.1. *Left*: Graphical representation of the set of binary measurements $\mathscr{B}$. Each probability operator Π_0 is represented by a point of the cone of equations $w^2 - x^2 - y^2 - z^2 \geq 0$ and $0 < w \leq 1/2$. The spherical caps are the loci of points with fixed value of the purity γ. *Right*: Relations between subfamilies of measurements. All projective measurements are irreversible. The optimal Braunstein-Caves measurement, i.e. an energy measurement, is both projective and semiclassical.

3.2.1 Information

Given a generic measurement $\mathcal{M} \in \mathscr{B}$, with conical coordinates $(w, \lambda, \theta, \varphi)$, its Fisher information is

$$\mathcal{F}_C(\beta) = \frac{\delta^2 \lambda^2 w \cos^2\theta \operatorname{sech}^4(\beta\delta/2)}{4Q\,[1 - wQ]}\,, \qquad Q := \left[1 - \lambda \cos\theta \tanh\left(\frac{\beta\delta}{2}\right)\right]\,. \tag{3.2.1}$$

With the help of Eq. (3.2.1), one can study how the information varies as a function of the non-commutativity parameter χ (*resp.*, the purity parameter γ). In Fig. 3.2, each measurement gives rise to a point in the plane $\mathcal{F}_C(\beta)$ vs χ (*resp.*, $\mathcal{F}_C(\beta)$ vs γ). The boundary curve corresponds to measurements that maximize the Fisher information, for fixed value of χ (*resp.*, γ). For instance, the measurement that extracts maximum information for given value of χ is obtained for $\lambda = 1$ and $w = 1/2$:

$$\Pi_0 = \frac{1}{2}\begin{pmatrix} 1 \pm \sqrt{1-\chi^2} & \chi\, e^{-i\varphi} \\ \chi\, e^{i\varphi} & 1 \mp \sqrt{1-\chi^2} \end{pmatrix}\,, \qquad \Pi_1 = \mathbb{I}_2 - \Pi_0\,. \tag{3.2.2}$$

Analytically, the boundary curve (see left panel of Fig. 3.2) is

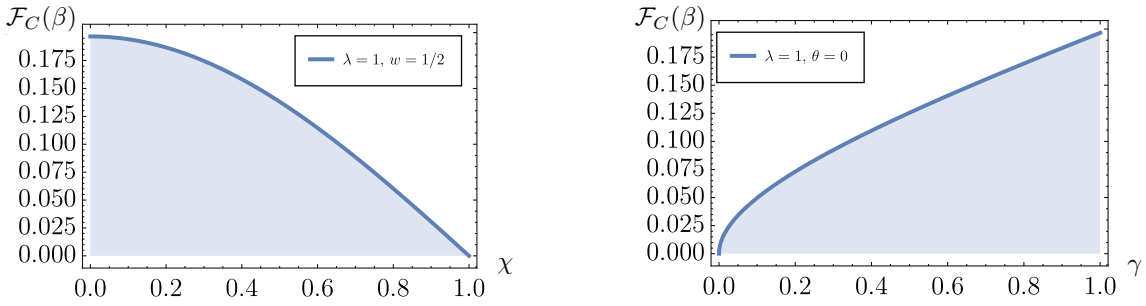
$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{F}_C^{(\chi)}(\beta)\,\big|_{\chi\,\text{const.}} = \frac{2\beta^2\,(1-\chi^2)}{4 - \chi^2 + 4\cosh(\beta\delta) + \chi^2\cosh(2\beta\delta)}\,. \tag{3.2.3}$$

Similarly, the measurement that extracts maximum information for given value of γ is obtained for $\lambda = 1$ and $\theta = 0$, i.e. it is the semiclassical measurement of the form:

$$\Pi_0 = \begin{pmatrix} \sqrt{\gamma} & 0 \\ 0 & 0 \end{pmatrix}\,, \qquad \Pi_1 = \mathbb{I}_2 - \Pi_0\,. \tag{3.2.4}$$

The boundary curve (see right panel of Fig. 3.2) is given by

$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{F}_C^{(\gamma)}(\beta)\,\Big|_{\gamma\,\text{const.}} = \frac{\beta^2\sqrt{\gamma}\,e^{2\beta\delta}}{(1+e^{\beta\delta})^2\,(1+e^{\beta\delta}-\sqrt{\gamma})}\,. \tag{3.2.5}$$



(a) **(b)**

Figure 3.2. *Left*: Fisher information and non-commutativity χ. *Right*: Fisher information and purity γ. Both plots are obtained upon fixing $\beta = \delta = 1$.

3.2.2 The Δ-disturbance

The Δ-disturbance caused by a generic measurement with conical coordinates $(w, \lambda, \theta, \varphi)$ can be computed as

$$\mathcal{D}_\beta^{(\Delta)}(\mathscr{M}) = \mathcal{F}_Q[\rho_\beta] - \Pr_\beta(0)\mathcal{F}_Q[\rho_{\beta|0}] - \Pr_\beta(1)\mathcal{F}_Q[\rho_{\beta|1}]\,, \tag{3.2.6}$$

where the QFI is $\mathcal{F}_Q[\rho_\beta] = \delta^2/(2+2\cosh\beta\delta)$. The outcomes probabilities are

$$\Pr_\beta(0) = w - \lambda w\cos\theta\tanh(\beta\delta/2)\,, \qquad \Pr_\beta(1) = 1 - \Pr_\beta(0)\,, \tag{3.2.7}$$

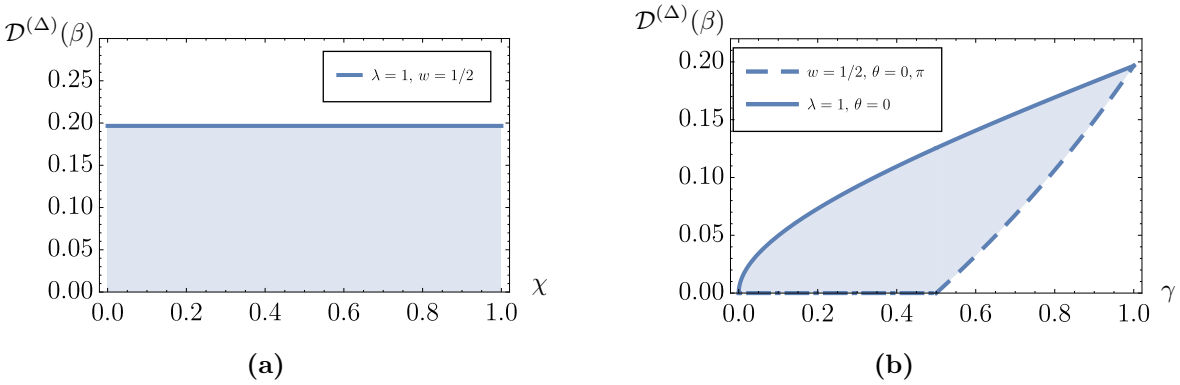
and the QFI of the conditional states

$$\mathcal{F}_Q[\rho_{\beta|0}] = \frac{\delta^2(1-\lambda^2)\,e^{\beta\delta}}{(K_+ + e^{\beta\delta}K_-)^2}\,, \qquad \mathcal{F}_Q[\rho_{\beta|1}] = \frac{\delta^2[(1-w)^2 - \lambda^2 w^2]\,e^{\beta\delta}}{[1-wK_+ + e^{\beta\delta}(1-wK_-)]^2}\,, \tag{3.2.8}$$

with $K_\pm := 1 \pm \lambda\cos\theta$. The Δ-disturbance is maximized by projective measurements and minimized by uninformative measurements.

If the non-commutativity χ is fixed, the measurement that maximizes $\mathcal{D}^{(\Delta)}(\beta)$ is obtained for $\lambda = 1$ and $w = 1/2$, i.e. it is the projective POVM of Eq. (3.2.2). The corresponding maximum is the QFI $\mathcal{F}_Q(\beta)$. On the other hand, for given value of γ, the maximally disturbing measurement is obtained for $\lambda = 1$ and $\theta = 0$, i.e. it is the semiclassical POVM of Eq. (3.2.4). The corresponding maximum is

$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(\Delta)}(\beta)\,\Big|_{\gamma\,\text{const.}} = \frac{\beta^2\sqrt{\gamma}\,e^{2\beta\delta}}{(1+e^{\beta\delta})^2(1+e^{\beta\delta}-\sqrt{\gamma})}\,. \tag{3.2.9}$$



(a) **(b)**

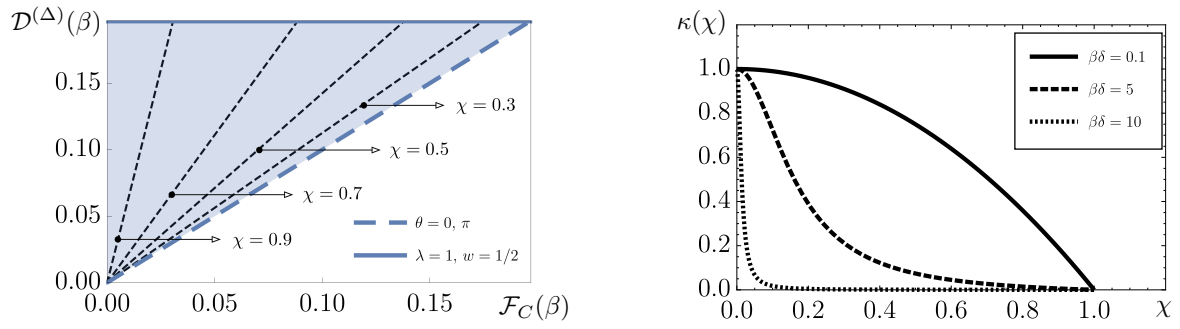
Figure 3.3. *Left*: Range of the Δ-disturbance as a function of χ. *Right*: Range of the Δ-disturbance as a function of γ. Both plots are obtained upon fixing $\beta = \delta = 1$.

While for any given value of χ it is possible to find a zero-disturbance measurement, measurements with purity $\gamma > 1/2$ must destroy information. The minimum achievable information loss is attained by measurements having $w = 1/2$ and $\theta = 0$ or π, which identifies a subfamily of semiclassical measurements of the form

$$\Pi_0 = \frac{1}{2}\begin{pmatrix} 1 \pm \sqrt{2\gamma - 1} & 0 \\ 0 & 1 \mp \sqrt{2\gamma - 1} \end{pmatrix}, \qquad \Pi_1 = \mathbb{I}_2 - \Pi_0 \,, \tag{3.2.10}$$

where the upper choice of sign corresponds to the case $\theta = 0$ and the lower choice to $\theta = \pi$. The corresponding minimum is

$$\min_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(\Delta)}(\beta)\Big|_{\gamma>1/2\,\text{const.}} = \frac{\beta^2\,(2\gamma-1)\,\operatorname{sech}^2(\beta\delta/2)}{4(\gamma + \cosh\beta\delta - \gamma\cosh\beta\delta)} \,. \tag{3.2.11}$$



(a) **(b)**

Figure 3.4. *Left*: Information/disturbance trade-off region for $\mathcal{D}^{(\Delta)}_\beta$. *Right*: Plot of $\kappa(\chi)$ (see Proposition 3.2.1) for different values of the product $\beta\delta$.

Next, we discuss the trade-off relation between the information $\mathcal{F}_C(\beta)$ and the Δ-disturbance $\mathcal{D}^{(\Delta)}(\beta)$. The trade-off region is completely characterized by the following proposition.

Proposition 3.2.1. *The trade-off region in the plane $\mathcal{F}_C(\beta)$ vs $\mathcal{D}^{(\Delta)}(\beta)$ is the triangle of vertices $(0,0)$, $(0,\mathcal{F}_Q(\beta))$ and $(\mathcal{F}_Q(\beta),\mathcal{F}_Q(\beta))$. It can be foliated into one dimensional curves by fixing the value of the non-commutativity parameter χ. In fact, for given value of χ, the trade-off curves are line segments with endpoints $(0,0)$ and $(\kappa(\chi)\,\mathcal{F}_Q(\beta),\mathcal{F}_Q(\beta))$, where*

$$\kappa(\chi)=\frac{2(1-\chi^2)}{2-\chi^2+\chi^2\cosh(\beta\delta)}\ . \tag{3.2.12}$$

Proof. See Ref. [59]. □

From Prop. 3.2.1, it follows that Δ-efficient measurements line up on the curve $\kappa(\chi)=1$, which implies $\chi=0$. Thus, the family of Δ-efficient measurements coincides with the semiclassical family. This last statement can, in fact, be generalized to any finite-dimensional thermometer.

Proposition 3.2.2. *For a d-level thermometer, any semiclassical measurement is Δ-efficient.*

Proof. Since $\mathcal{D}^{(\Delta)}(\beta)\geq\mathcal{F}_C(\beta)$ [57], it will be enough to prove that, in fact, for a semiclassical measurement, equality holds. Let us consider a generic semiclassical measurement, with sample space $\mathcal{J}$ and probability operators $\{\Pi_j\}_{j\in\mathcal{J}}$. The statistical model is $\rho_\beta=\exp(-\beta H)/Z_\beta$, with $H\in\mathsf{Her}_d(\mathbb{C})$ the system's Hamiltonian and $Z_\beta:=\mathrm{tr}[\exp(-\beta H)]$. Its SLD is denoted by $L_\beta[\rho_\beta]$. If the outcome j is obtained, the post-measurement state is $\rho_{\beta|j}=\Pi_j\rho_\beta/\mathrm{Pr}_\beta(j)$, where $\mathrm{Pr}_\beta(j)=\mathrm{tr}(\rho_\beta\Pi_j)$. It easy to check that the SLD of $\rho_{\beta|j}$ is related to that of ρ_β as follows,

$$L_\beta[\rho_{\beta|j}]=L_\beta[\rho_\beta]-\partial_\beta\log\mathrm{Pr}_\beta(j)\,\mathbb{1}_d\ . \tag{3.2.13}$$

The average QFI of the post-measurement state evaluates to

$$\begin{aligned}\langle\mathcal{F}_Q[\rho_{\beta|x}]\rangle=\sum_{j\in\mathcal{J}}\{\mathrm{tr}(\rho_\beta L_\beta^2[\rho_\beta]\Pi_j)+\sum_{j\in\mathcal{J}}\mathrm{Pr}_\beta(j)[\partial_\beta\log\mathrm{Pr}_\beta(j)]^2+\\ -2\sum_{j\in\mathcal{J}}[\partial_\beta\log\mathrm{Pr}_\beta(j)]\,\mathrm{tr}(L_\beta[\rho_\beta]\rho_\beta\Pi_j)\}\ .\end{aligned} \tag{3.2.14}$$

The first summation on the RHS is equal to the QFI $\mathcal{F}_Q[\rho_\beta]$; the second to the FI for the measurement we are considering; the third is also equal to the FI (using the fact that $\mathrm{tr}(L_\beta[\rho_\beta]\rho_\beta\Pi_j)=\partial_\beta\mathrm{Pr}_\beta(j)$). It follows that the Fisher and the Δ-disturbance coincide, which proves our claim. □

3.2.3 The F-disturbance

For a generic measurement $\mathcal{M} \in \mathscr{B}$, the F-disturbance evaluates to

$$\mathcal{D}^{(F)}(\beta) = \frac{1+\cos^2\theta+\sin^2\theta\cosh(\beta\delta)}{2[1+\cosh(\beta\delta)]}\left[1-w\sqrt{1-\lambda^2}-\sqrt{(1-w)^2-w^2\lambda^2}\right] . \quad (3.2.15)$$

The subset of $\mathscr{B}$ that maximizes $\mathcal{D}^{(F)}(\beta)$ is made up of measurements that are simultaneously projective and non-classical. They have the following probability operators:

$$\Pi_0 = \frac{1}{2}\begin{pmatrix} 1 & e^{-i\varphi} \\ e^{i\varphi} & 1 \end{pmatrix} , \qquad \Pi_1 = \mathbb{I}_2 - \Pi_0 . \quad (3.2.16)$$

Moreover, the maximum F-disturbance is equal to $1/2$. On the other hand, the F-disturbance is minimized by uninformative measurements which, in fact, do not introduce any disturbance. However, if one fixes the value of the non-commutativity χ, the maximum F-disturbance is achieved by projective measurements. The corresponding maximum is

$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(F)}(\beta)\Big|_{\chi\,\text{const.}} = \frac{2-\chi^2+\chi^2\cosh(\beta\delta)}{2+2\cosh(\beta\delta)} . \quad (3.2.17)$$

If instead the purity γ is kept fixed, the maximum disturbance is attained by measurements that are simultaneously irreversible and non-classical. Explicitly, they have the form:

$$\Pi_0 = \frac{\sqrt{\gamma}}{2}\begin{pmatrix} 1 & e^{-i\varphi} \\ e^{i\varphi} & 1 \end{pmatrix} , \qquad \Pi_1 = \mathbb{I}_2 - \Pi_0 . \quad (3.2.18)$$

The corresponding disturbance is

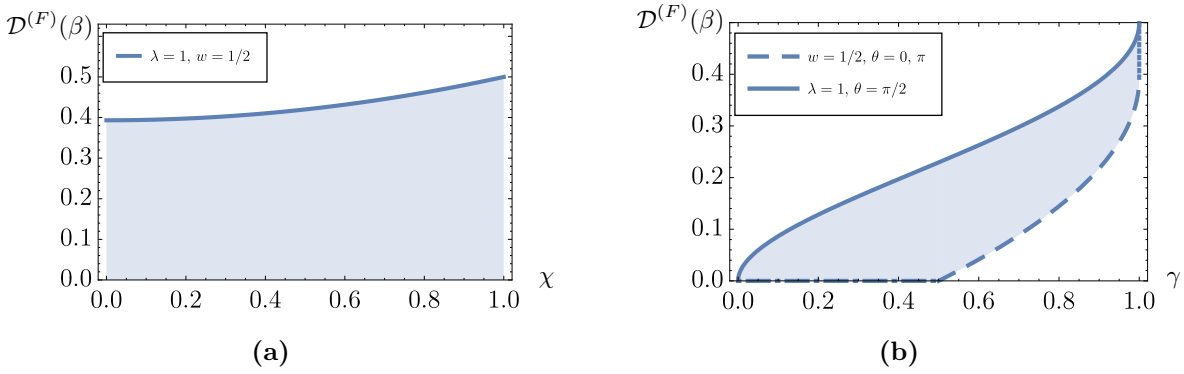
$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(F)}(\beta)\Big|_{\gamma\,\text{const.}} = \frac{1-\sqrt{1-\sqrt{\gamma}}}{2} . \quad (3.2.19)$$

For $\gamma > 1/2$, there is also a non-trivial lower bound,

$$\min_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(F)}(\beta)\Big|_{\gamma>1/2\,\text{const.}} = \frac{1-\sqrt{2-2\gamma}}{1+\cosh(\beta\delta)} , \quad (3.2.20)$$

which is achieved by a subfamily of semiclassical measurements satisfying the condition $w = 1/2$ (see Eq. (3.2.10)).

Finally, we study the trade-off relation between the Fisher information and the F-disturbance. The trade-off region in the plane $\mathcal{F}_C(\beta)$ vs $\mathcal{D}^{(F)}(\beta)$ is qualitatively similar to that for the Δ-disturbance. In particular, the F-efficient measurements are still of the semiclassical type. They are described by the conditions $\theta = 0$, $\lambda = \lambda^{\text{opt}}$ and $w = w^{\text{opt}}$. The latter two parameters can only be determined numerically, see the right panel of Fig. 3.6. In the left panel, the upper curve corresponds to projective measurements.

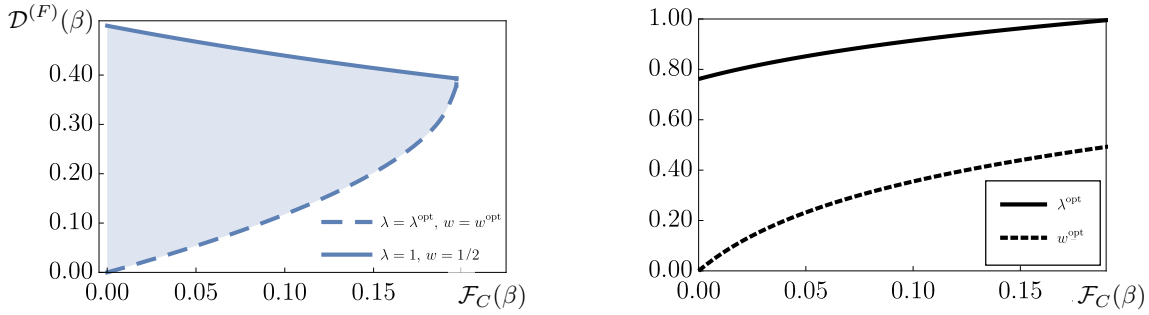


(a) **(b)**

Figure 3.5. *Left*: Range of the F-disturbance as a function of χ. *Right*: Range of the F-disturbance as a function of γ. Both plots are obtained upon fixing $\beta = \delta = 1$.

They maximize the F-disturbance for given information. The corresponding maximum is

$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(F)}(\beta)\Big|_{\mathcal{F}_C \text{ const.}} = \frac{\beta^2\delta^2}{2\beta^2\delta^2 - \mathcal{F}_C + \mathcal{F}_C\cosh(2\beta\delta)} \,. \tag{3.2.21}$$



(a) **(b)**

Figure 3.6. *Left*: Trade-off region for the F-disturbance. The F-efficient measurements (*dashed*) satisfy $\theta = 0$, $\lambda = \lambda^{\text{opt}}$ and $w = w^{\text{opt}}$. *Right*: Parameters λ^{opt} and w^{opt} as a function of $\mathcal{F}_C(\beta)$.

3.2.4 The τ-disturbance

In general, the post-measurement state of the thermometer is an out-of-equilibrium state and thus cannot be assigned a temperature in the conventional sense. However, the spectral temperature of Eq. (3.1.7) is defined for any state, it coincides with the standard temperature if evaluated on thermal states, and it shares many of its thermodynamical properties [58]. Thus, a natural disturbance quantifier is the average spectral temperature variation, i.e. the τ-disturbance. For a qubit thermometer, it evaluates to

$$\mathcal{D}^{(\tau)}(\beta) = \text{Pr}_\beta(0)\left|\beta - \tau(\rho_{\beta|0})\right| + \text{Pr}_\beta(1)\left|\beta - \tau(\rho_{\beta|1})\right| \,. \tag{3.2.22}$$

The spectral temperatures of the two conditional states are

$$\begin{aligned}\tau(\rho_{\beta|0}) &= \frac{1}{\delta}\,\log\left[\frac{2(1-\Lambda_+\Lambda_-)+e^{\beta\delta}\Lambda_-^\theta}{2(1-\Lambda_+\Lambda_-)\,e^{\beta\delta}+\Lambda_+^\theta}\right],\\ \tau(\rho_{\beta|1}) &= \frac{1}{\delta}\,\log\left[\frac{2(1-w-W_+W_-)+e^{\beta\delta}W_+^\theta}{2(1-w-W_+W_-)\,e^{\beta\delta}+W_-^\theta}\right],\end{aligned} \tag{3.2.23}$$

where we defined:

$$\begin{aligned}\Lambda_\pm &:= \sqrt{1\pm\lambda}\,, & \Lambda_\pm^\theta &:= \Big[\Lambda_\pm\,(\cot\theta+\csc\theta)+\Lambda_\mp\,\tan(\theta/2)\Big]^2,\\ W_\pm &:= \sqrt{1-w\pm\lambda w}\,, & W_\pm^\theta &:= \Big[W_\pm\,(\cot\theta+\csc\theta)+W_\mp\,\tan(\theta/2)\Big]^2.\end{aligned}$$

We now study the range of $\mathcal{D}^{(\tau)}(\beta)$ as a function of the non-commutativity χ and the purity γ. For given value of χ, the maximally disturbing measurements are the projective measurements. The corresponding maximum is

$$\max_{\mathcal{M}\in\mathscr{B}}\mathcal{D}^{(\tau)}(\beta)\,\Big|_{\chi\,\text{const.}} = \frac{\beta}{2}\left(H_-\widetilde{H}_- + H_+\widetilde{H}_+\right) \tag{3.2.24}$$

where

$$H_\pm := \left|1\pm\frac{1}{\beta\delta}\log\frac{1-\sqrt{1-\chi^2}}{1+\sqrt{1-\chi^2}}\right|, \qquad \widetilde{H}_\pm := \left(1\pm\sqrt{1-\chi^2}\tanh\frac{\beta\delta}{2}\right). \tag{3.2.25}$$

For given γ, instead, there is no finite upper-bound. For $\gamma > 1/2$, there is however a non-trivial lower bound, corresponding to measurements with coordinates $w = 1/2$ and $\theta = \pi/2$, or explicitly,

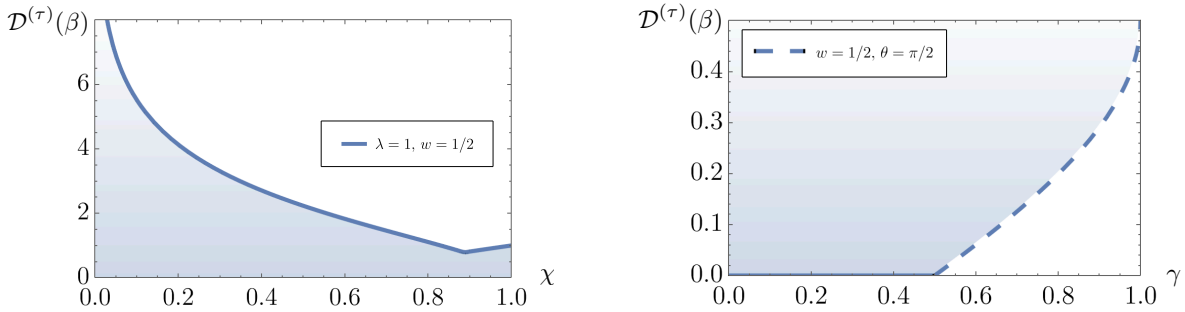
$$\Pi_0 = \frac{1}{2}\begin{pmatrix}1 & -\sqrt{2\gamma-1}\,e^{-i\varphi}\\ -\sqrt{2\gamma-1}\,e^{i\varphi} & 1\end{pmatrix}, \qquad \Pi_1 = \mathbb{I}_2-\Pi_0\,. \tag{3.2.26}$$

The corresponding disturbance is

$$\min_{\mathcal{M}\in\mathscr{B}}\mathcal{D}^{(\tau)}(\beta)\,\Big|_{\gamma>1/2\,\text{const.}} = \frac{1}{2\delta}\left|\beta\delta-\log\frac{L_+}{L_-}\right|, \qquad L_\pm := 1\pm\sqrt{2-2\gamma}\,\tanh\left(\frac{\beta\delta}{2}\right). \tag{3.2.27}$$

We conclude this section with a study of the information/disturbance trade-off region. First of all, the region is unbounded from above. That is, for any given value of the extracted information $\mathcal{F}_C(\beta)$, there always exists a measurement for which $\mathcal{D}^{(\tau)}(\beta)$ diverges. It can be explicitly constructed as follows. Take a measurement which is irreversible, semiclassical and whose w coordinate satisfies

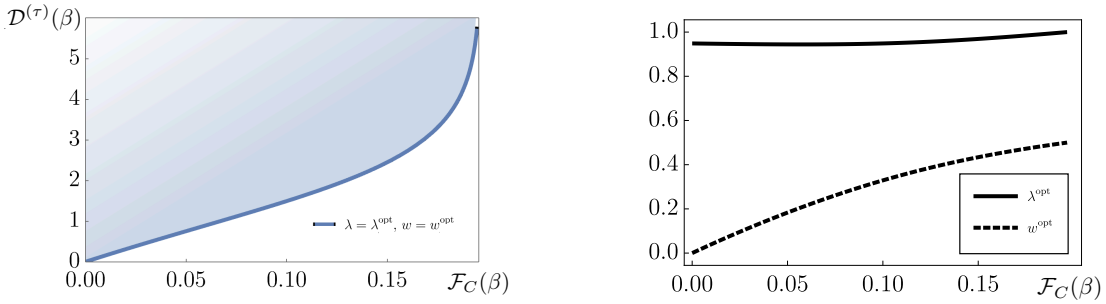
$$w = \frac{(1+e^{\beta\delta})^3\mathcal{F}_C(\beta)}{2[(1+e^{\beta\delta})^2\mathcal{F}_C(\beta)+\beta^2\delta^2\,e^{2\beta\delta}]}\,. \tag{3.2.28}$$



(a) **(b)**

Figure 3.7. *Left*: Range of the τ-disturbance as a function of χ. *Right*: Range of the τ-disturbance as a function of γ. Both regions are unbounded.

It is easy to check that the corresponding information is $\mathcal{F}_C(\beta)$, while the τ-disturbance diverges (since the probability $\Pr_\beta(0)$ vanishes). In contrast, there is a non-trivial lower curve which, in fact, corresponds to τ-efficient measurements. They form a subfamily of the semiclassical measurements, with $\theta = 0$ and parameters λ^{opt} and w^{opt} that must be determined numerically (see Fig. 3.8).



(a) **(b)**

Figure 3.8. *Left*: Trade-off region for the τ-disturbance. *Right*: Behavior of the optimal parameters λ^{opt} and w^{opt} for τ-efficient measurements.

3.2.5 The π-disturbance

The disturbance quantifier $\mathcal{D}^{(\pi)}(\beta)$ has a strong geometrical interpretation. In general, a measurement forces the state of the system out of the thermal manifold $S := \{\rho_\beta\}_{\beta\in\mathbb{R}^+}$. One can however project it back onto the manifold S by means of the I-projection defined in Eq. (3.1.9). On the thermal manifold, a natural way to quantify the distance between two states is given by the quantum relative entropy:

$$S_Q(\rho_\eta||\rho_\beta) = \text{tr}(\rho_\eta \log \rho_\eta) - \text{tr}(\rho_\eta \log \rho_\beta) = (\beta - \eta)\langle H\rangle_\eta + \log\left(Z_\beta/Z_\eta\right) \ . \tag{3.2.29}$$

The π-disturbance is, in fact, the average quantum relative entropy between the pre-measurement state and the projected post-measurement state. Even for qubits, computation of $\mathcal{D}^{(\pi)}(\beta)$ must, in general, be performed numerically.

Here, we summarize the main features of the behavior of $\mathcal{D}^{(\pi)}(\beta)$ as a function of the relevant parameters. For fixed value of the non-commutativity χ, the measurements that maximize $\mathcal{D}^{(\pi)}(\beta)$ are the projective ones. The corresponding disturbance is

$$\max_{\mathcal{M}\in\mathscr{B}} \mathcal{D}^{(\pi)}(\beta)\Big|_{\chi\,\text{const.}} = \frac{1}{2}\left[P_-^{\theta}P_+ + P_+^{\theta}P_-\right], \tag{3.2.30}$$

where

$$P_\pm := \log(1+e^{\pm\beta\delta}), \qquad P_\pm^{\theta} := 1 \pm \cos\theta\, \tanh\left(\frac{\beta\delta}{2}\right). \tag{3.2.31}$$

Instead, for fixed value of γ, the measurements that maximize $\mathcal{D}^{(\pi)}(\beta)$ are a subset of the irreversible ones, with conical coordinates $\lambda = 1$ and $\theta = \theta^{\text{opt}}$ (to be determined numerically). Let us remark that, in contrast to the three disturbance quantifiers previously studied, there is no non-trivial lower bound for $\gamma > 1/2$. In fact, a non-classical measurement scheme (i.e. $\theta = \pi/2$) always introduces a vanishing disturbance, but at the same time does not extract any useful information (see Eq. (3.2.1)). Finally, π-efficient measurements make up a subset of the semiclassical measurements. They have conical coordinates $\theta = 0$, $\lambda = \lambda^{\text{opt}}$ and $w = w^{\text{opt}}$; the latter two parameters must be determined numerically (see Fig. 3.10).

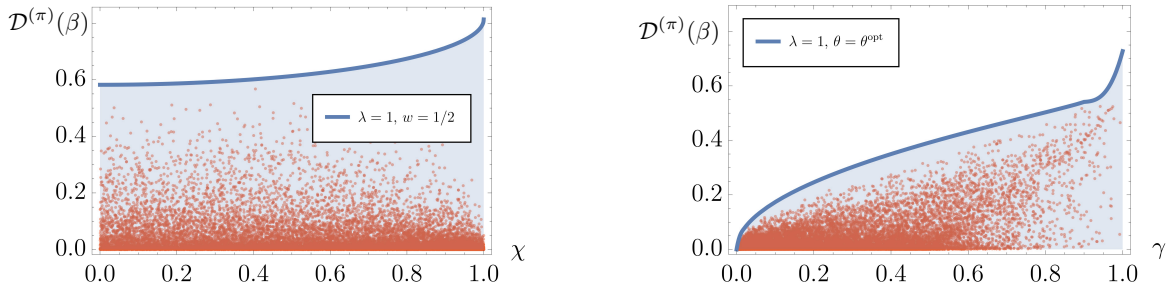
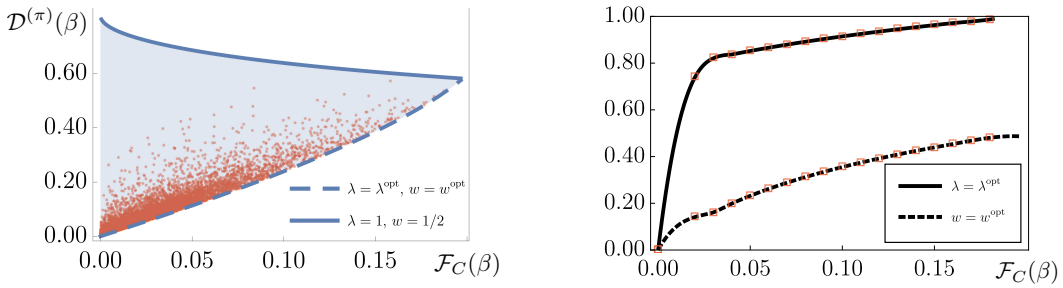


Figure 3.9. *Left*: Range of the π-disturbance as a function of χ. *Right*: Range of the π-disturbance as a function of γ. Dots represent the performance of randomly generated POVMs with parameters (w, λ, θ) chosen uniformly.

3.3 Conclusions

In this chapter, we have studied the trade-off relation between information and disturbance in the setting of qubit thermometry. Four different disturbance quantifiers have



(a) **(b)**

Figure 3.10. *Left*: Information/disturbance trade-off region for $\mathcal{D}_\beta^{(\pi)}$. Dots correspond to the performances of random measurements, whose POVMs have been generated with uniformly randomized parameters (λ, θ, w). *Right*: Behavior of the optimal parameters $\lambda^{\rm opt}$ and $w^{\rm opt}$ for π-efficient measurements.

been employed, which capture different aspects of the disturbance introduced by quantum measurements. For instance, the Δ-disturbance focuses on the information that cannot be retrieved by future measurements. The F-disturbance is based on the Hilbert space distance between the pre- and post-measurement states. The τ-disturbance has thermodynamical origins. The π-disturbance emphasizes the notion of statistical distinguishability between states before and after a measurement.

Despite their differences, some general features emerge from our analysis that appear to hold irrespective of how the disturbance is quantified. For example: for fixed value of non-commutativity, the measurements maximizing the disturbance are the projective ones; for fixed value of the purity, they are the irreversible ones; the efficient measurements are always of the semiclassical type. In particular, commutativity with the statistical model appears to be a necessary, but not sufficient condition for efficiency. It is both necessary and sufficient, however, in the case of the Δ-disturbance.

These commonalities suggest, on the one hand, that all four disturbance quantifiers are meaningful in their own and, on the other, that some of our conclusions may hold more generally, beyond the specifics of the model adopted here (e.g. for higher dimensional thermometers).

Chapter 4

Estimation of continuous-time quantum walks Hamiltonians on classical graphs

In this chapter, we study the problem of reconstructing the Hamiltonian of a quantum walker evolving continuously in time on a graph, focusing in particular on the interplay between the maximum achievable precision and the graph's topology. After setting up the problem in Sect. 4.1 and Sect. 4.2, we investigate the ultimate limits to precision in Sect. 4.3, while Sect. 4.4 analyzes the performance of a few realistic estimation strategies.

4.1 Introduction

The Hilbert space of a quantum system at low energies can often be truncated so that only a finite number of states is sufficient to describe its effective dynamics. In all such cases, the system can be considered as completely equivalent to a quantum particle that evolves continuously in time by transitioning between different nodes, or vertices, of a graph G. Each node of G is labeled by an integer j (with $j \in \{1, \ldots, n\}$) and is in one-to-one correspondence with a state $|j\rangle$ of the original system. If H_G denotes the system's Hamiltonian, any two vertices of G are connected by an edge whenever the transition amplitude $\langle i|H_G|j\rangle$ between the corresponding states is nonzero. One may therefore write H_G as

$$H_G = -\sum_{(i,j)\in E} \gamma_{ij}(|i\rangle\langle j| + |j\rangle\langle i|) + \sum_{j\in V} \epsilon_j |j\rangle\langle j| \ , \tag{4.1.1}$$

where E is the set of edges and V the set of vertices of G. Denoting by $|\psi_0\rangle$ the initial state at time $t = 0$, time evolution implies $|\psi_0\rangle \to |\psi_t\rangle = \exp(-itH_G)|\psi_0\rangle$. The system

is said to perform a continuous-time quantum walk on the graph G. The orthonormal states $\{|j\rangle\}$ are usually referred to as the position eigenstates of the walker.

Any quantum walk Hamiltonian of the form (4.1.1) can be simulated on a quantum computer of n qubits. Let the computational basis vectors be denoted by $|\chi_1\chi_2\ldots\chi_n\rangle := |\chi_1\rangle \otimes |\chi_2\rangle \otimes \cdots \otimes |\chi_n\rangle$, where $\chi_i \in \{\uparrow,\downarrow\}$. Let the Hamiltonian H_Q be

$$H_Q = -\sum_{(i,j)\in E} \gamma_{ij}(\sigma_i^{(-)} \otimes \sigma_j^{(+)} + \sigma_i^{(+)} \otimes \sigma_j^{(-)}) + \sum_{j\in V} \epsilon_j(1+\sigma_j^{(z)}) \,, \tag{4.1.2}$$

where $\sigma^{(-)} := |\downarrow\rangle\langle\uparrow|$, $\sigma^{(+)} := |\uparrow\rangle\langle\downarrow|$ and $\sigma^{(z)} := |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\uparrow|$. Each vertex state $|j\rangle$ can be encoded as a single-excitation state made up of one qubit in the up state $|\uparrow\rangle$ and all the remaining qubits in the down state $|\downarrow\rangle$, i.e. $|j\rangle = |\downarrow\rangle^{\otimes j-1} \otimes |\uparrow\rangle \otimes |\downarrow\rangle^{\otimes n-j}$. Notice that the Hamiltonian H_Q commutes with the total excitation number. Therefore, if the initial state is a single-excitation state, the system remains at all times in the single-excitation subspace. The restriction of the Hamiltonian H_Q to this subspace gives precisely the quantum walk Hamiltonian H_G of Eq. (4.1.1).

With the emergence of intermediate-scale quantum computers [60], quantum walks promise to be available in the near future for applications to quantum algorithms and quantum simulations. For instance, it is well known [61] that one-dimensional quantum walks have superior transport properties compared with classical random walks, which makes them extremely powerful for several algorithmic tasks [62]. Quantum walks may even be crucial to explain the coherent energy transport in photosynthetic complexes [63]. In this regard, characterization of the architecture on which the quantum walk takes place is a relevant problem. Full knowledge of the quantum walk Hamiltonian is in fact required to achieve complete quantum control of the system. In this chapter, we investigate the topic for several remarkable families of classical graphs. Some of the questions we aim to answer are: how precisely can the Hamiltonian be estimated via a suitable quantum measurement on the walker? How does the best achievable performance vary as a function of the topology of the underlying graph? How does it scale with the number of nodes? Is it possible to retrieve information on H_G by monitoring only a subset of nodes, or even a single node?

In the following, unless otherwise states, we will make the conventional choice that each on-site energy ϵ_j is proportional to the vertex degree d_j and that all transition amplitudes γ_{ij} are equal among themselves. That is, H_G can be written as

$$H_G = -\gamma \sum_{(i,j)\in E} (|i\rangle\langle j| + |j\rangle\langle i|) + \epsilon \sum_{j\in V} d_j |j\rangle\langle j| \,. \tag{4.1.3}$$

In matrix notation, with respect to the position eigenbasis:

$$(H_G)_{ij} = \begin{cases} \epsilon d_i & \text{if } i = j \\ -\gamma & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}, \tag{4.1.4}$$

with $\epsilon, \gamma \in \mathbb{R}^+$. Since both ϵ and γ appear linearly in the Hamiltonian, one may set e.g. ϵ to 1 and measure γ in units of ϵ. The coupling γ is then the only relevant parameter to be estimated.

Let us remark that our model is slightly more general than the one often employed in the quantum walk literature [64], which conventionally also assumes that $\gamma = \epsilon$. In particular, if $\gamma = \epsilon = 1$, the Hamiltonian H_G reduces to the Laplacian matrix $L = D - A$, where D is the diagonal degree matrix and A the adjacency matrix of G. Notice however that, in a physical architecture, the coupling strength γ and the on-site energy ϵ can in general be tuned independently from one another by the experimentalist.

The rest of the chapter is organized as follows. In Sect. 4.2, we introduce several families of graphs and discuss their corresponding quantum walk dynamics by solving for the eigenvalues and eigenvectors of the associated Hamiltonians. In Section 4.3, we quantify the best precision achievable by any protocol aimed at estimating H_G via arbitrary quantum measurements on the walker. Section 4.4 discusses a few specific measurements and analyzes their performance. Finally, in section 4.5 we draw our conclusions.

4.2 Quantum walks on special graph families

Let the energy eigenstates of H_G be denoted by $|\xi_j\rangle$, with $j \in \{0, \dots, n-1\}$. At time $t = 0$, the system is prepared in an initial state $|\psi_0\rangle = \sum_{j=0}^{n-1} \alpha_j |\xi_j\rangle$. After a time t, the state of the system is $|\psi_t\rangle = \sum_{j=0}^{n-1} \alpha_j e^{-i\xi_j t} |\xi_j\rangle$, which gives a complete description of any quantum walk on G. In this section, we compute the eigenvalues and eigenvectors of H_G for several remarkable graph families:

(**G1**) A *circulant* graph O_n is a simple graph having the following property: there exists a relabelling of its vertices which is 1) an isomorphism (i.e., any two vertices of the resulting graph are connected by an edge if and only if their pre-images are also connected) and 2) is a *cyclic* permutation of the vertices (the relabelling permutes a subset of the vertices in a cyclic fashion, while leaving fixed the remaining ones). Visually, a circulant graph can always be drawn in such a way that its n vertices are the corners of a regular n-gon and every rotational symmetry of the n-gon is also a symmetry of the drawing. The family of circulant graphs contains some

familiar examples (cycle graphs, complete graphs), as well as more exotic ones (cocktail-party graphs, Andrásfai graphs, antiprism graph, Paley graphs of prime order, Möbius ladders).

The adjacency matrix of a circulant graph is a circulant matrix, i.e. each row is obtained by shifting the preceding one to the right. Circulant matrices are well studied in the mathematical literature [65] and, in particular, the analytical form of their eigenvalues and eigenvectors is known. Based on this fact, we will relax our previous assumption that all couplings are taken to be equal. Instead, we will allow the couplings γ_{ij} to be different, but with the constraint that they can be a function only of $|i-j| \bmod(n)$. For instance, $\gamma_{12} = \gamma_{21} = \gamma_{23} = \gamma_{32}$ but, in principle, γ_{12} is not equal to γ_{13} (they are distinct parameters). It follows that H_{O_n} is a circulant matrix, which may be written as

$$H_{O_n} = \begin{pmatrix} d\epsilon & -\gamma_1 & -\gamma_2 & \dots & -\gamma_{n-1} \\ -\gamma_{n-1} & d\epsilon & -\gamma_1 & \dots & -\gamma_{n-2} \\ -\gamma_{n-2} & -\gamma_{n-1} & d\epsilon & \dots & -\gamma_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\gamma_1 & -\gamma_2 & -\gamma_3 & \dots & d\epsilon \end{pmatrix} , \tag{4.2.1}$$

where d is the degree of each vertex and $\gamma_j = \gamma_{n-j}$, $j \in \{1, \dots, n-1\}$. Notice that there are a total of $\lfloor n/2 \rfloor$ independent couplings. Its eigenvalues are

$$\xi_j = d\epsilon - \sum_{k=1}^{n-1} \gamma_k \, e^{\frac{2\pi i j k}{n}} , \qquad j \in \{0, \dots, n-1\} . \tag{4.2.2}$$

With the help of the relation $\gamma_j = \gamma_{n-j}$, one may rewrite the previous equation as

$$\xi_j = \begin{cases} d\epsilon - 2\gamma_1 \cos\left(\frac{2\pi j}{n}\right) - \dots - 2\gamma_{(n-2)/2} \cos\left(\frac{(n-2)\pi j}{n}\right) - \gamma_{n/2} \cos(\pi j) & (n \text{ even}) \\ d\epsilon - 2\gamma_1 \cos\left(\frac{2\pi j}{n}\right) - \dots - 2\gamma_{(n-1)/2} \cos\left(\frac{(n-1)\pi j}{n}\right) & (n \text{ odd}) \end{cases} .$$

The spectrum is doubly degenerate, i.e. $\xi_j = \xi_{n-j}$. The corresponding eigenvectors are

$$|\xi_j\rangle = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} e^{\frac{2\pi i j (k-1)}{n}} \, |k\rangle \ . \tag{4.2.3}$$

Two special cases of circulant graphs are complete graphs and cycle graphs.

(**G1.1**) A *complete* graph K_n is a graph where each vertex is adjacent to any other vertex. It has n vertices, $n(n-1)/2$ edges and it is regular with degree $n-1$.

For complete graphs, we will return to assuming that all couplings are equal to the same constant γ. Explicitly, the associated quantum walk Hamiltonian can be written as

$$H_{K_n} = \begin{pmatrix} (n-1)\epsilon & -\gamma & -\gamma & \dots & -\gamma \\ -\gamma & (n-1)\epsilon & -\gamma & \dots & -\gamma \\ -\gamma & -\gamma & (n-1)\epsilon & \dots & -\gamma \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\gamma & -\gamma & -\gamma & \dots & (n-1)\epsilon \end{pmatrix} . \tag{4.2.4}$$

Making use of Eq. (4.2.2) and the identity $\sum_{k=0}^{n-1} \exp(2\pi jk/n) = 0$, the eigenvalues ξ_j of H_{K_n} can be written more compactly as

$$\xi_j = \begin{cases} (n-1)(\epsilon-\gamma) & \text{if } j = 0 \\ (n-1)\epsilon + \gamma & \text{if } j \neq 0 \end{cases} . \tag{4.2.5}$$

The least eigenvalue is ξ_0, whereas the remaining eigenvalues are all degenerate. The corresponding eigenvectors are the same as in Eq. (4.2.3).

(**G1.2**) A *cycle* graph C_n is a chain of n vertices (and as many edges). It is circulant and regular with degree 2.

The corresponding quantum walk Hamiltonian H_{C_n} is

$$H_{C_n} = \begin{pmatrix} 2\epsilon & -\gamma & 0 & \dots & -\gamma \\ -\gamma & 2\epsilon & -\gamma & \dots & 0 \\ 0 & -\gamma & 2\epsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\gamma & 0 & 0 & \dots & 2\epsilon \end{pmatrix} . \tag{4.2.6}$$

In the following, we will focus on the special case of a cycle graph with an *even* number of vertices (without loss of generality). Specializing some of the above formulas, one finds in particular that the least eigenvalue of H_{C_n} is given by $\xi_0 = 2(\epsilon - \gamma)$, while the largest is $\xi_{n/2} = 2(\epsilon + \gamma)$, with corresponding eigenvectors

$$|\xi_0\rangle = \frac{1}{\sqrt{n}}(1, 1 \dots, 1, 1)^t \qquad \text{and} \qquad |\xi_{n/2}\rangle = \frac{1}{\sqrt{n}}(1, -1 \dots, 1, -1)^t . \tag{4.2.7}$$

(**G2**) A *hypercube* graph Q_d is the regular graph formed from the vertices and edges of the d-dimensional hypercube. It has $n = 2^d$ vertices and $d \cdot 2^{d-1}$ edges; besides, each vertex has the same degree d. It can also be thought of as the d-fold Cartesian product of the complete graph K_2.

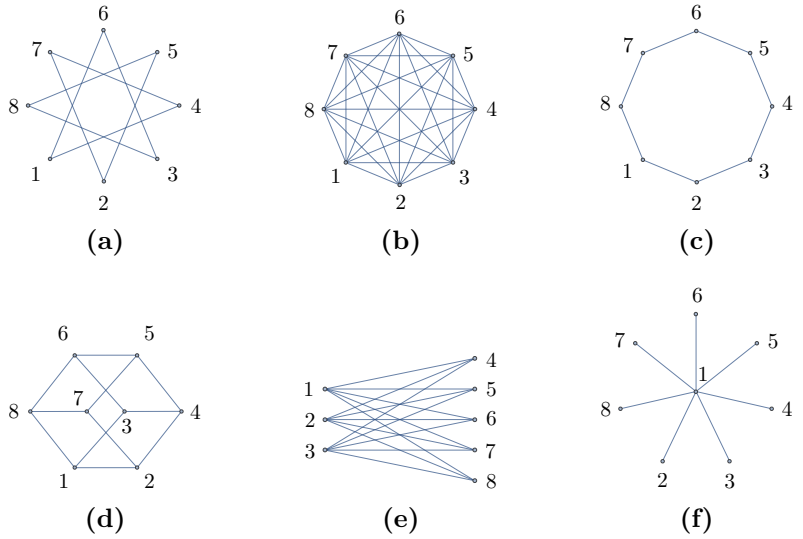


Figure 4.1. Families of graphs considered in the main text, with a conventional choice of labelling: **(a)** circulant graph, **(b)** complete graph, **(c)** cycle graph, **(d)** hypercube graph, **(e)** complete bipartite graph, **(f)** star graph.

The corresponding Hamiltonian matrix is $H_{Q_d} = d\epsilon\, \mathbb{I}_{2^d} - \gamma A^{(d)}$, where $A^{(d)}$ is the adjacency matrix of Q_d, which is defined recursively via the relation

$$A^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad A^{(d)} = \begin{pmatrix} A^{(d-1)} & \mathbb{I}_{2^{d-1}} \\ \mathbb{I}_{2^{d-1}} & A^{(d-1)} \end{pmatrix} . \tag{4.2.8}$$

The eigenvectors of H_{Q_d} are denoted by $|\xi_j^{(\kappa_j)}\rangle$, $j \in \{0, \ldots, d\}$, where κ_j is a degeneracy index ranging from 1 to $[d, j] := d!/j!(d-j)!$. The corresponding eigenvalues are

$$\xi_j = d\epsilon - \gamma(d - 2j) . \tag{4.2.9}$$

The eigenvectors $|\xi_j^{(\kappa_j)}\rangle$ coincide with the columns of a sequence of matrices F_d, indexed by the dimension d of the hypercube and defined recursively as follows:

$$F_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} , \qquad F_d = \frac{1}{\sqrt{2}} \begin{pmatrix} F_{d-1} & F_{d-1} \\ F_{d-1} & -F_{d-1} \end{pmatrix} . \tag{4.2.10}$$

By construction, each F_d is a Hadamard matrix. In particular, the ground state $|\xi_0\rangle$ and the highest excited state $|\xi_d\rangle$ can be written as

$$|\xi_0\rangle = \frac{1}{2^{d/2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{\otimes d} , \qquad |\xi_d\rangle = \frac{1}{2^{d/2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}^{\otimes d} . \tag{4.2.11}$$

Notice that they are non-degenerate, i.e. $\kappa_0 = \kappa_d = 1$, so we omit the degeneracy label.

(**G3**) A *complete bipartite* graph $K_{p,q}$ is such that its vertex set V can be split into two subsets, with the following property: all vertex of the first set are adjacent to every other vertex of the second, but no two vertices in either sets are adjacent among themselves. It has $n = p+q$ vertices, with $\{1, \dots, p\}$ and $\{p+1, \dots, p+q\}$ denoting the two partitions. A special case of a complete bipartite graph is the star graph $S_n = K_{1,n-1}$, corresponding to setting $p = 1$ and $q = n - 1$.

The corresponding Hamiltonian $H_{K_{p,q}}$ takes the following block form:

$$H_{K_{p,q}} = \left(\begin{array}{c|c} q\epsilon\, \mathbb{I}_p & -\gamma \mathbb{J}_{p\times q} \\ \hline -\gamma \mathbb{J}_{q\times p} & p\epsilon\, \mathbb{I}_q \end{array}\right) , \tag{4.2.12}$$

where $\mathbb{I}_p$ is the $p \times p$ identity matrix and $\mathbb{J}_{p\times q}$ is the $p \times q$ matrix made up of all ones. A natural ansatz for an eigenvector $|\xi\rangle$ of $H_{K_{p,q}}$ is in the form $|\xi\rangle = (\boldsymbol{x}, \boldsymbol{y})^t$, where $\boldsymbol{x}$ is a p-vector and $\boldsymbol{y}$ a q-vector. The eigenvalue equation $H_{K_{p,q}} |\xi\rangle = \xi \cdot |\xi\rangle$ implies the linear system of constraints

$$\begin{cases} (q\epsilon - \xi)\boldsymbol{x} = \gamma \mathbb{J}_{p\times q}\, \boldsymbol{y} \\ (p\epsilon - \xi)\boldsymbol{y} = \gamma \mathbb{J}_{q\times p}\, \boldsymbol{x} \end{cases} . \tag{4.2.13}$$

Multiplying by $\mathbb{J}_{q\times p}$ the first of the previous equations, and using the fact that $\mathbb{J}_{q\times p}\, \mathbb{J}_{p\times q} = p\, \mathbb{J}_{q\times q}$, one obtains

$$\mathbb{J}_{q\times q}\, \boldsymbol{y} = \frac{(q\epsilon - \xi)(p\epsilon - \xi)}{\gamma^2 p}\, \boldsymbol{y} , \tag{4.2.14}$$

which implies that $\boldsymbol{y}$ is an eigenvector of $\mathbb{J}_{q\times q}$. Similarly, applying $\mathbb{J}_{p\times q}$ to the second equation of (4.2.13), one gets

$$\mathbb{J}_{p\times p}\, \boldsymbol{x} = \frac{(q\epsilon - \xi)(p\epsilon - \xi)}{\gamma^2 q}\, \boldsymbol{x} , \tag{4.2.15}$$

so $\boldsymbol{x}$ is an eigenvector of $\mathbb{J}_{p\times p}$.

Let us recall that the spectrum of $\mathbb{J}_{p\times p}$ is made up of the eigenvalue 0 (with multiplicity $p - 1$ and corresponding eigenspace spanned by all p-vectors whose components sum to zero) and the eigenvalue p (with multiplicity 1 and corresponding eigenvector $\mathbf{1}_p$, the p-vector made up of ones). Eqs. (4.2.14) and (4.2.15) thus imply that the only possible values for ξ are $\mathrm{spec}(H_G) = \{\xi_1, \xi_2, \xi_+, \xi_-\}$, where $\xi_1 := q\epsilon$, $\xi_2 := p\epsilon$ and $\xi_\pm$ are the roots of the equation $(q\epsilon - \xi_\pm)(p\epsilon - \xi_\pm) = qp\,\gamma^2$, i.e.

$$\xi_\pm = \frac{(p+q)\epsilon \pm \sqrt{\Delta_{p,q}}}{2} , \qquad \text{with} \qquad \Delta_{p,q} := (p-q)^2\epsilon^2 + 4qp\,\gamma^2 . \tag{4.2.16}$$

Notice that the lowest eigenvalue is ξ_-, the highest is ξ_+, while ξ_1 and ξ_2 are always in between.

The corresponding eigenvectors can be found as follows. For the eigenvalue ξ_1, one finds that $\boldsymbol{y}$ must vanish and that $\boldsymbol{x} \in \ker \mathbb{J}_{p\times p}$, whereas for ξ_2, $\boldsymbol{x}$ must vanish and $\boldsymbol{y} \in \ker \mathbb{J}_{q\times q}$. Introducing two orthonormal basis $\boldsymbol{a}^{(\kappa)}$ and $\boldsymbol{b}^{(\kappa')}$, for $\ker \mathbb{J}_{p\times p}$ and $\ker \mathbb{J}_{q\times q}$ respectively,

$$|\xi_1^{(\kappa)}\rangle = (\boldsymbol{a}^{(\kappa)}, \boldsymbol{0}_q)^t \,, \qquad |\xi_2^{(\kappa')}\rangle = (\boldsymbol{0}_p, \boldsymbol{b}^{(\kappa')})^t \,, \tag{4.2.17}$$

where $\kappa \in \{1, \dots, p-1\}$ and $\kappa' \in \{1, \dots, q-1\}$. For the remaining two eigenvalues $\xi_\pm$, after substituting into the eigenvalue equation, one finally finds

$$|\xi_\pm\rangle = \eta_\pm \begin{pmatrix} \frac{p\epsilon - \xi_\pm}{\gamma p}\, \boldsymbol{1}_p \\ \boldsymbol{1}_q \end{pmatrix} \,, \qquad \text{with} \qquad \eta_\pm := \left[q \left(1 + \frac{p\epsilon - \xi_\pm}{q\epsilon - \xi_\pm} \right) \right]^{-1/2} \,. \tag{4.2.18}$$

Since Eqs. (4.2.17) and (4.2.18) already define a set of $p+q$ orthonormal eigenvectors, there are no additional eigenvectors.

4.3 Ultimate quantum limits to precision

The problem we are going to consider is to reconstruct the Hamiltonian H_G of a quantum walker on a graph G via repeated quantum measurements. In a parameter estimation framework, the best achievable precision is expressed by the quantum Fisher information. In the following, we compute it for the families of graphs introduced in the previous section. In particular, we study how the QFI scales with the number of vertices n and the interrogation time t, and maximize it over the initial preparation $|\psi_0\rangle$.

(G1) A *circulant graph* O_n, with Hamiltonian H_{O_n} of the same form as in Eq. (4.2.1), has n eigenvectors $|\xi_j\rangle$, which are given explicitly in Eq. (4.2.3). For the corresponding eigenvalues ξ_j, see Eq. (4.2.2). There are a total of $m := \lfloor n/2 \rfloor$ unknown parameters $\gamma_1, \dots, \gamma_m$.

We assume that the experimenter can prepare arbitrary superpositions of the walker's energy eigenstates. Therefore, at time $t = 0$, the initial state of the walker has the form $|\psi_0\rangle = \sum_{j=0}^{n-1} \alpha_j |\xi_j\rangle$, where the weights α_j are arbitrary except for the constraint $\sum_{j=0}^{n-1} |\alpha_j|^2 = 1$. The QFI metric is obtained via a standard computation. At the generic time t, it can be written compactly as:

$$[\mathcal{F}_Q(\gamma_1, \dots, \gamma_m)]_{kl} = 16t^2 \,\mathrm{Cov}\left[\cos(2\pi k \mathbb{X}/n),\ \cos(2\pi l \mathbb{X}/n)\right] \,, \tag{4.3.1}$$

where $\mathbb{X}$ is a random variable, taking values $j \in \{0, \dots, n-1\}$, with probabilities $\Pr(j) = |\alpha_j|^2$.

Let us now focus on just one coupling constant, e.g. γ_k. The corresponding QFI is $[\mathcal{F}_Q(\gamma_1, \dots, \gamma_m)]_{kk} = 16t^2 \operatorname{Var}[\cos(2\pi k\mathbb{X}/n)]$. We now make use of Popoviciu's inequality, see Eq. (1.4.44). The maximum value of $\cos(2\pi k\mathbb{X}/n)$ is 1, obtained e.g. for $\mathbb{X} = 0$ (or any other value φ_k such that $k\varphi_k = 0 \pmod n$). To find the minimum value, first define ϑ_k as follows

$$\vartheta_k := \operatorname*{arg\,min}_{\theta \in \{0,1,\dots,n-1\}} k\theta - \left\lfloor \frac{n}{2} \right\rfloor \pmod n . \tag{4.3.2}$$

The minimum value of $\cos(2\pi k\mathbb{X}/n)$ is $\cos(2\pi k\vartheta_k/n)$, obtained for $\mathbb{X} = \vartheta_k$. Thus, the following inequality holds

$$[\mathcal{F}_Q(\gamma_1, \dots, \gamma_m)]_{kk} \leq 4t^2[1 - \cos(2\pi k\vartheta_k/n)]^2 . \tag{4.3.3}$$

It is saturated when the initial state is a balanced superposition of $|\xi_0\rangle$ and $|\xi_{\vartheta_k}\rangle$, which is thus the optimal preparation. Let us remark that the optimal QFI with respect to the parameter γ_k is independent of the number of vertices n.

We now consider in detail the special cases of a complete graph K_n and of a cycle graph C_n:

(G1.1) For a complete graph K_n, there is only one independent parameter, γ. The initial preparation is $|\psi_0\rangle = \sum_{j=0}^{n-1} \alpha_j |\xi_j\rangle$, where $|\xi_0\rangle$ is the ground state and the other eigenstates form a degenerate subspace (the corresponding eigenvalues are given in Eq. (4.2.5)). The QFI evaluates to

$$\mathcal{F}_Q(\gamma) = 4t^2n^2 \ |\alpha_0|^2 \left(1 - |\alpha_0|^2\right) . \tag{4.3.4}$$

Maximizing over the initial preparation, one has that $\mathcal{F}_Q(\gamma) \leq n^2t^2$, which is saturated when $|\psi_0\rangle$ is equally distributed between the ground state and the excited energy subspace.

The QFI scales quadratically with the number of vertices n. The different scaling compared with the case of a general circulant graph O_n is explained by the fact that all couplings are known from the start to be equal among themselves.

(G1.2) For a cycle graph C_n, all couplings are zero except for $\gamma_1 = \gamma_{n-1} := \gamma$. The QFI can be found by specializing the general formula, Eq. (4.3.1), valid for any circulant graph O_n. In particular, Eq. (4.3.2) can be solved explicitly, which gives $\vartheta = \lfloor n/2 \rfloor$. The QFI, optimized over the initial preparation, is equal to $16t^2$ if n is even and $4t^2[1 + \cos(\pi/n)]^2$ if n is odd.

(G2) An *hypercube graph* Q_d has eigenvalues ξ_j, where $j \in \{0, \dots, d\}$, see Eq. (4.2.9). Each eigenvalue ξ_j has multiplicity $[d, j] := d!/j!(d-j)!$. The corresponding eigenstates $|\xi_j^{(\kappa_j)}\rangle$, with $\kappa_j \in \{1, \dots, [d, j]\}$, are constructed recursively for any d by means of Eq. (4.2.10). The most general initial preparation is

$$|\psi_0\rangle = \sum_{j=0}^{d} \sum_{\kappa_j=1}^{[d,j]} \alpha_j^{(\kappa_j)} |\xi_j^{(\kappa_j)}\rangle \ , \qquad \text{with} \qquad \sum_{j=0}^{d} \sum_{\kappa_j=1}^{[d,j]} |\alpha_j^{(\kappa_j)}|^2 = 1 \ . \tag{4.3.5}$$

For future convenience, let us denote by $p_{\xi_j} := \sum_{\kappa_j=1}^{[d,j]} |\alpha_j^{(\kappa_j)}|^2$ the total probability that an energy measurement returns the outcome ξ_j. After a standard computation, the QFI at time t can be written as

$$\mathcal{F}_Q(\gamma) = 4t^2 \operatorname{Var}(d - 2\mathbb{X}) \ , \tag{4.3.6}$$

where $\mathbb{X}$ is a random variable such that $\Pr(j) = p_{\xi_j}$, for $j \in \{0, \dots, d\}$. Since the maximum value of $d - 2\mathbb{X}$ is d (when $\mathbb{X} = 0$) and the minimum value is $-d$ (when $\mathbb{X} = d$), one has that $\mathcal{F}_Q(\gamma) \leq 4t^2 d^2$ by Popoviciu's inequality. The optimal QFI thus equals $4t^2 d^2$, which scales quadratically with the dimension d and the interrogation time t. The optimal preparation is a balanced superposition of the ground state $|\xi_0\rangle$ and the maximally excited state $|\xi_d\rangle$ (see Eq. (4.2.11)).

(G3) A *complete bipartite graph* $K_{p,q}$ has $n = p + q$ eigenvectors, of which only two, $|\xi_\pm\rangle$ given in Eq. (4.2.18), depend on the parameter γ. All other eigenvectors, as well as their eigenvalues, are independent of γ and no estimation strategy can fruitfully make use of them. As a consequence, the initial preparation is taken to be a superposition of $|\xi_\pm\rangle$ only, e.g. $|\psi_0\rangle = \alpha_- |\xi_-\rangle + \alpha_+ |\xi_+\rangle$. The corresponding QFI at the generic time t evaluates to

$$\mathcal{F}_Q(\gamma) = \frac{4\left(f_{p,q} - 4g_{p,q}^2\right)}{\Delta^2} \ , \tag{4.3.7}$$

where

$$f_{p,q} = pq[16p^2q^2\gamma^4t^2 + (p-q)^2(1 + 4pq\gamma^2t^2)\epsilon^2] \ , \tag{4.3.8}$$

$$g_{p,q} = (|\alpha_-|^2 - |\alpha_+|^2)\, pq\gamma t\sqrt{\Delta_{p,q}} + \Im\left(e^{it\sqrt{\Delta_{p,q}}}\bar{\alpha}_+\alpha_-\right)(p-q)\sqrt{pq}\,\epsilon \ . \tag{4.3.9}$$

The optimal initial preparation is such that $g_{p,q}$ vanishes, which is obtained when $|\alpha_-| = |\alpha_+|$, with a relative phase $\arg(\alpha_+/\alpha_-) = t\sqrt{\Delta_{p,q}}$. The maximum QFI is therefore equal to $4f_{p,q}/\Delta_{p,q}^2$. For fixed number of vertices n, one may further optimize over the cardinality of each bipartition p and q. The maximum is reached for $p = \lfloor n/2 \rfloor$, with corresponding scaling $\sim n^2t^2$, quadratic both in the number of vertices and the interrogation time.

We study in some detail the special case of a star graph S_n. The maximum QFI is obtained via the substitutions $p = 1$ and $q = n - 1$,

$$\max_{|\psi_0\rangle} \mathcal{F}_Q(\gamma) = \frac{4(n-1)[16(n-1)^2\,\gamma^4 t^2 + (n-2)^2[1 + 4(n-1)\gamma^2 t^2]\epsilon^2]}{[(n-2)^2\,\epsilon^2 + 4(n-1)\gamma^2]^2} \,. \quad (4.3.10)$$

It depends on n but does not grow indefinitely with the size of the graph. When $n \to \infty$, it saturates instead to a constant value. Therefore, an optimal number of nodes $n_{\rm opt}$ may exist. We solve for $n_{\rm opt}$ in the two opposite regimes of small and long times. For small times $\gamma t \ll 1$,

$$\max_{|\psi_0\rangle} \mathcal{F}_Q(\gamma) \sim \frac{4(n-1)(n-2)^2\,\epsilon^2}{[(n-2)^2\,\epsilon^2 + 4(n-1)\gamma^2]^2} \,, \quad (4.3.11)$$

which is maximized by

$$n_{\rm opt} \sim \frac{2[1 + \epsilon^2/\gamma^2 + \sqrt{1 + \epsilon^2/\gamma^2}]}{\epsilon^2/\gamma^2} \,. \quad (4.3.12)$$

For large times $\gamma t \gg 1$,

$$\max_{|\psi_0\rangle} \mathcal{F}_Q(\gamma) \sim \frac{16\gamma^2 t^2 (n-1)[4(n-1)^2\gamma^2 + (n-2)^2(n-1)\epsilon^2]}{[(n-2)^2\,\epsilon^2 + 4(n-1)\gamma^2]^2} \,. \quad (4.3.13)$$

If $\epsilon^2/\gamma^2 \leq 2$, then there is no optimal value of n (the optimal value is $n = \infty$). Instead, if $\epsilon^2/\gamma^2 > 2$, the optimal value of n is

$$n_{\rm opt} \sim \frac{2(\epsilon^2/\gamma^2 - 1)}{(\epsilon^2/\gamma^2 - 2)} \,. \quad (4.3.14)$$

Adding new vertices above $n_{\rm opt}$ will lower the maximum achievable precision.

4.4 Performance of a few selected measurements

Because the quantum Cramér-Rao inequality is tight, there always exists a quantum measurement whose FI is equal to the QFI computed in the previous section. However, the optimal Braunstein-Caves measurement may be quite exotic, or even depend on the true value of the parameter, so that it is not necessarily available to the experimentalist. In this section, we investigate the performance of a few specific measurements and discuss their efficiency with respect to the QFI upper bound.

One observable on the Hilbert space of the walker which has a clear-cut interpretation is its position. By definition, a position measurement corresponds to a projection onto the walker's position eigenstates, i.e. the position operator is $\hat{x} = \sum_{i=1}^{n} i\,|i\rangle\,\langle i|$. It is also useful to introduce *incomplete* position measurements. They correspond to coarse-grainings of

a position measurement. Explicitly, the probability operators of an incomplete position measurement of size m are the m rank-1 projectors onto m given position eigenstates (plus the projector onto the orthogonal complement of the subspace spanned by them). Incomplete measurements model a situation where one has experimental access only to a subset of nodes.

In the following, we analyze the performance of position measurements in the estimation of γ for different graph families, assuming that the initial preparation coincides with the optimal one maximizing the QFI.

(**G1**) We consider two particular examples of circulant graphs: an arbitrary complete graph K_n and a cycle graph C_n with n an even integer.

(**G1.1**) As shown in the previous section, for a complete graph K_n the optimal preparation is a balanced superposition of the ground state $|\xi_0\rangle$ and of any of the excited states, e.g. $|\xi_1\rangle$. We thus assume that the state of the walker at time $t = 0$ is $\sum_{j=0}^{n} \alpha_j |\xi_j\rangle$ with $\alpha_0 = \alpha_1 = 1/\sqrt{2}$ and all other α_j set to zero. After a time t, an incomplete position measurement of size m is performed. By definition, its probability operators are the projectors $|j\rangle\langle j|$, for $j \in \{1, \dots, m\}$ (plus the projector onto their orthogonal complement $\mathbb{1}_n - \sum_{j=0}^{m} |j\rangle\langle j|$). The corresponding Fisher information is denoted by $\mathcal{F}_C^{(m)}(\gamma)$. The efficiency is defined as the ratio of the FI to the QFI, optimized over the interrogation time t, i.e. $\eta^{(m)} := \max_t \mathcal{F}_C^{(m)}(\gamma)/\mathcal{F}_Q(\gamma)$.

The Fisher information $\mathcal{F}_C^{(m)}(\gamma)$ can be written as

$$\mathcal{F}_C^{(m)}(\gamma) = \sum_{j=1}^{m} \frac{(\partial_\gamma p_j)^2}{p_j} + \frac{(\partial_\gamma \bar{p})^2}{\bar{p}} \,, \tag{4.4.1}$$

where p_j is the probability of detecting the walker at node j, i.e.

$$p_j = |\langle j|\psi_t\rangle|^2 = \frac{2}{n} \cos^2\left[\frac{\gamma n t}{2} - \frac{\pi(j-1)}{n}\right] \tag{4.4.2}$$

and $\bar{p} := 1 - \sum_{j=1}^{m} p_j$ is the probability that the walker is measured outside the subset of the graph under control by the experimenter. It is possible to rewrite $\bar{p}$ as follows,

$$\bar{p} = \frac{n-m}{n} - \frac{1}{n} \cos\left[\gamma n t - \frac{(m-1)\pi}{n}\right] \frac{\sin\left(\frac{\pi m}{n}\right)}{\sin\left(\frac{\pi}{n}\right)} \,, \tag{4.4.3}$$

which can be proved by first transforming $\cos^2(x/2) = (1 + \cos x)/2$ and then simplifying the sum over the cosine terms via the geometric series identity

$$\sum_{j=0}^{m-1} e^{i\left(x - \frac{2\pi j}{n}\right)} = e^{ix} \left(\frac{1 - e^{-\frac{2\pi i m}{n}}}{1 - e^{-\frac{2\pi i}{n}}}\right) \,. \tag{4.4.4}$$

The final result for $\mathcal{F}_C^{(m)}(\gamma)$ is

$$\mathcal{F}_C^{(m)}(\gamma) = mnt^2 - nt^2 \cos\left[\gamma nt - \frac{(m-1)\pi}{n}\right] \frac{\sin\left(\frac{\pi m}{n}\right)}{\sin\left(\frac{\pi}{n}\right)} + \frac{t^2 \sin^2\left[\gamma nt - \frac{(m-1)\pi}{n}\right] \frac{\sin^2\left(\frac{\pi m}{n}\right)}{\sin^2\left(\frac{\pi}{n}\right)}}{\frac{n-m}{n} - \frac{1}{n}\cos\left[\gamma nt - \frac{(m-1)\pi}{n}\right] \frac{\sin\left(\frac{\pi m}{n}\right)}{\sin\left(\frac{\pi}{n}\right)}} . \tag{4.4.5}$$

For instance, if $m = 1$:

$$\mathcal{F}_C^{(1)}(\gamma) = \frac{2\sin^2\left(\frac{\gamma nt}{2}\right)}{n - 2\cos^2\left(\frac{\gamma nt}{2}\right)} n^2 t^2 . \tag{4.4.6}$$

The corresponding efficiency is $\eta^{(1)} = \frac{2}{n}$: a local measurement can at most extract a fraction $2/n$ of the maximum available information. Vice versa, for a complete position measurement, i.e. $m = n$, one finds from Eq. (4.4.5) that $\mathcal{F}_C^{(n)}(\gamma) = n^2 t^2$, which is also equal to the QFI: a complete position measurement is optimal.

For a generic incomplete measurement with $1 < m < n$,

$$\eta^{(m)} = \frac{m}{n} + \frac{1}{n} \frac{\sin\left(\frac{\pi m}{n}\right)}{\sin\left(\frac{\pi}{n}\right)} . \tag{4.4.7}$$

Using the inequality $\sin(m\pi/n) \leq m \sin(\pi/n)$, the efficiency $\eta^{(m)}$ is at most $2m/n$, i.e. twice the fraction of nodes that can be individually addressed. In particular, the upper bound $\eta^{(m)} \sim 2m/n$ is reached when $n \gg m$ and $n \gg 1$.

(**G1.2**) For a cycle graph C_n, with n an even integer (the case n odd can be dealt with in a totally analogous way), the optimal preparation is a balanced superposition of the ground state $|\xi_0\rangle$ and highest excited state $|\xi_{n/2}\rangle$, where $\xi_0 = 2(\epsilon - \gamma)$, $\xi_{n/2} = 2(\epsilon + \gamma)$, and the corresponding eigenvectors are defined in Eq. (4.2.7). The state of the walker at time t is

$$|\psi_t\rangle = \frac{1}{\sqrt{2}} e^{-i\xi_0 t} |\xi_0\rangle + \frac{1}{\sqrt{2}} e^{-i\xi_{n/2} t} |\xi_{n/2}\rangle = \frac{1}{\sqrt{2n}} e^{-2i(\epsilon-\gamma)t} \begin{pmatrix} 1 + e^{-4i\gamma t} \\ 1 - e^{-4i\gamma t} \\ \vdots \\ 1 + e^{-4i\gamma t} \\ 1 - e^{-4i\gamma t} \end{pmatrix} . \tag{4.4.8}$$

The probability p_j of measuring the walker at node j is

$$p_j = \frac{1}{n}\left[1 + (-1)^{j+1}\cos(4\gamma t)\right] = \begin{cases} \frac{2}{n}\cos^2(2\gamma t) := p_O & \text{if } j \text{ is odd} \\ \frac{2}{n}\sin^2(2\gamma t) := p_E & \text{if } j \text{ is even} \end{cases} . \tag{4.4.9}$$

We study the performance of an incomplete position measurement over a subset of the graphs's nodes, of which n_O have an odd label and n_E an even label. Introducing the notation $\beta_O := 2n_O/n$ and $\beta_E := 2n_E/n$ for, respectively, the fractions of odd and even nodes under control, the corresponding Fisher information $\mathcal{F}_C^{(\beta_0,\beta_E)}(\gamma)$ can be written as

$$\mathcal{F}_C^{(\beta_0,\beta_E)}(\gamma) = n_O\frac{(\partial_\gamma p_O)^2}{p_O} + n_E\frac{(\partial_\gamma p_E)^2}{p_E} + \frac{(\partial_\gamma \bar{p})^2}{\bar{p}} , \tag{4.4.10}$$

where $\bar{p} = 1 - n_O p_O - n_E p_E$. One finds

$$\mathcal{F}_C^{(\beta_0,\beta_E)}(\gamma) = 16t^2 \left[\frac{\beta_O + \beta_E - 2\beta_O\beta_E + (\beta_E - \beta_O)\cos(4\gamma t)}{2 - (\beta_O + \beta_E) + (\beta_E - \beta_O)\cos(4\gamma t)}\right] . \tag{4.4.11}$$

In spite of appearances, the previous expression is invariant under relabellings of the graph's nodes. A relabelling may change the parity of each vertex label, exchanging β_O with β_E, and p_O with p_E. It is now enough to make use of the relation $\cos(4\gamma t) = np_O - 1 = 1 - np_E$ to check the invariance.

Recalling that the optimal QFI is $\max_{|\psi_0\rangle} \mathcal{F}_Q(\gamma) = 16t^2$, the efficiency evaluates to $\eta^{(\beta_O,\beta_E)} = \max(\beta_0, \beta_E)$. In particular, a complete position measurement is always optimal. Incomplete measurements can also be optimal, e.g. a measurement of only the odd or only the even vertices still has unit efficiency (see also Fig. 4.2).

(**G2**) For a hypercube graph Q_d, the optimal initial preparation is a balanced superposition of the ground state $|\xi_0\rangle$ and the highest excited state $|\xi_d\rangle$. After a time t, an incomplete position measurement is performed on $|\psi_t\rangle = (e^{-i\xi_0 t}|\xi_0\rangle + e^{-i\xi_d t}|\xi_d\rangle)/\sqrt{2}$, where $\xi_0 = d(\epsilon - \gamma)$, $\xi_d = d(\epsilon + \gamma)$ and $|\xi_0\rangle$, $|\xi_d\rangle$ are defined in Eq. (4.2.11). We adopt the following notation: $\mathcal{F}_C^{(\delta)}(\gamma)$ denotes the Fisher information for an incomplete measurement having as probability operators the 2^δ rank-1 projectors over the nodes making up a δ-dimensional face of the hypercube (plus the projector onto their orthogonal complement). It can be computed as

$$\mathcal{F}_C^{(\delta)}(\gamma) = 2^{\delta-1}\frac{(\partial_\gamma p_+)^2}{p_+} + 2^{\delta-1}\frac{(\partial_\gamma p_-)^2}{p_-} , \tag{4.4.12}$$

where

$$p_+ := \frac{1}{2^{d-1}}\cos^2(d\gamma t) , \qquad p_- := \frac{1}{2^{d-1}}\sin^2(d\gamma t) . \tag{4.4.13}$$

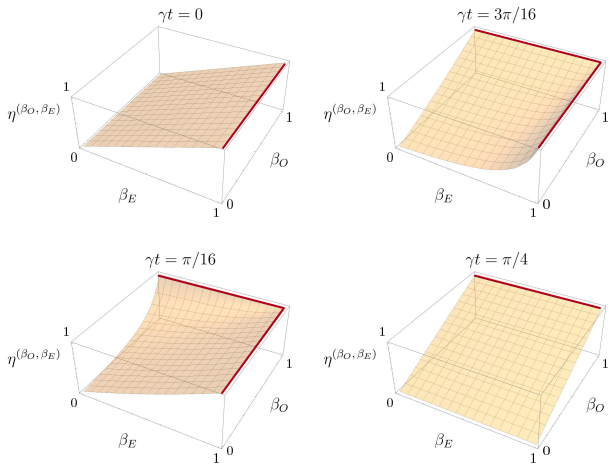


Figure 4.2. The efficiency $\eta^{(\beta_O,\beta_E)}$ of an incomplete measurement, as a function of the fractions of odd nodes β_O and even nodes β_E under experimental control, for different values of γt. Highlighted by a thick line (red), the optimal region of unit efficiency. For $\gamma t \neq m\pi/4$, $m \in \mathbb{Z}$, the optimal region consists of the two segments $(\beta_O, 1)$ and $(1, \beta_E)$. For even multiples of $\pi/4$, only the first segment is present, while for odd multiples only the second.

Notice that the probability of finding the walker in any of the accessible nodes is $2^{\delta-1}p_+ + 2^{\delta-1}p_- = 1/2^{d-\delta}$, which is, in particular, independent of γ. Thus, no term analogous to the last one on the left-hand-side of Eq. (4.4.1) appears in Eq. (4.4.12). The Fisher information evaluates to

$$\mathcal{F}_C^{(\delta)}(\gamma) = 2^{\delta-d+2}d^2t^2 \ . \tag{4.4.14}$$

Its efficiency is $\eta^{(\delta)} = 1/2^{d-\delta}$, i.e. the ratio between the number of nodes under individual control and the total number of nodes. In particular, a complete measurement (when $\delta = d$) is optimal.

(**G3**) Finally, we focus on one particular case of complete bipartite graph: a star graph.

For a star graph S_n, the optimal initial preparation is a balanced superposition of the two energy eigenstates $|\xi_\pm\rangle$ that can be read off from Eq. (4.2.18) after setting $p = 1$ and $q = n-1$, with a relative phase $\phi_{\text{opt}} = t\sqrt{\Delta_{1,n-1}}$. Since the optimal phase ϕ_{opt} depends on γ, an adaptive procedure is required in order to extract the maximum QFI. For the moment, we assume that the walker is prepared in the state $|\psi_0\rangle = (|\xi_-\rangle + e^{i\phi}|\xi_+\rangle)/\sqrt{2}$, where $\phi \in [0, 2\pi]$ is arbitrary.

At time t, an incomplete position measurement is performed. First, let us consider the case of an incomplete measurement monitoring only the central node, with as

probability operators the two projectors $|1\rangle\langle 1|$ and $\mathbb{I}_n - |1\rangle\langle 1|$. Its Fisher information coincides with the FI for a complete position measurement, denoted by $\mathcal{F}_C^{(\phi)}(\gamma)$ (the superscript indicates the dependence on the arbitrary phase ϕ of the initial state). The implication is that distinguishing outcomes corresponding to the walker being in one peripheral node or the other is useless for estimation purposes: one may as well monitor only the central node. The efficiency of a position measurement (either a complete measurement or an incomplete one, but including the central node) is

$$\eta^{(\phi)} = \frac{1}{16(n-1)^2t^2\gamma^4 + (n-2)^2[1+4(n-1)\gamma^2t^2]\epsilon^2} \times \frac{[(n-2)^2\epsilon^2\cos(\phi_{\rm opt}-\phi) - 4(n-1)\sqrt{\Delta_{1,n-1}}\,\gamma^2 t\sin(\phi_{\rm opt}-\phi)]^2}{4(n-1)\gamma^2\sin^2(\phi_{\rm opt}-\phi) + (n-2)^2\epsilon^2} . \tag{4.4.15}$$

Except for a few special cases (when $n = 2$, or $t = 0$ and $\phi = \phi_{\rm opt}$), a position measurement is suboptimal. For short interrogation time t, expanding for $\gamma t \ll 1$ and $\phi = \phi_{\rm opt}$ gives

$$\eta^{(\phi_{\rm opt})} = 1 - \frac{4(n-1)\Delta_{1,n-1}}{(n-2)^4\epsilon^4}\gamma^2t^2 + o(\gamma^2t^2) , \qquad (\gamma t \to 0) . \tag{4.4.16}$$

For large number of vertices $n \to \infty$ and $\phi = \phi_{\rm opt}$, one has instead

$$\eta^{(\phi_{\rm opt})} = \frac{1}{4\gamma^2t^2n} + o(1/n) , \qquad (n \to \infty), \tag{4.4.17}$$

i.e. the efficiency decreases linearly with the number of vertices. The reader is also referred to Fig. 4.3 for more details about the different possible regimes.

4.5 Conclusions

In this chapter, we have studied the problem of reconstructing the Hamiltonian H_G of a quantum walker evolving on a graph G. In particular, we have assumed a parameter estimation framework to infer the value of the coupling constants via repeated measurements on the walker. We have discovered that the graph topology can have dramatic effects on the maximum extractable information. For instance, the quantum Fisher information exhibits different scalings with the total number of vertices (see Table 4.1). For each graph family considered, we have maximized the quantum Fisher information over the initial preparation, determining in the process the optimal initial state of the walker, as well as the optimal measurement. Finally, we have studied in detail the performance of

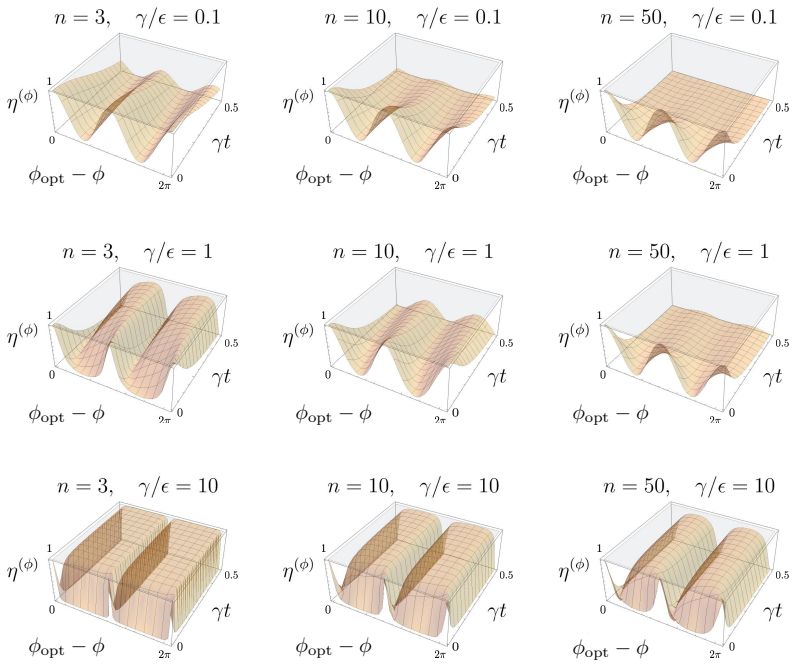


Figure 4.3. The efficiency $\eta^{(\phi)}$ of a complete measurement, as a function of the phase difference $\phi_{\text{opt}} - \phi$ and of the dimensionless time scale γt (for different values of n and γ/ϵ). In general, the efficiency is closer to one for higher values of γ/ϵ and for smaller values of γt and of n. Only for $t = 0$, there always exist a choice of ϕ which allows to reach unit efficiency; otherwise, a complete measurement is suboptimal.

position measurements. Complete position measurements perform quite well: they are often optimal, the only exception being (among the cases taken into account) that of complete bipartite graphs, e.g. star graphs. Incomplete position measurements still allow to extract a non-vanishing amount of information. Their efficiency (i.e. the ratio of the Fisher to quantum Fisher information) is closely related to the fraction β of the nodes under control by the experimentalist. The exception is again the case of star graphs, since monitoring only the central node yields the same information as monitoring each node separately.

	scaling of QFI with n	**optimal preparation**	**optimal measurement**	**η of incomplete position measurement**
K_n	$\sim n^2$	any balanced superposition of ground state and any other excited state	position	$\beta + 1/n \cdot \sin(\pi\beta)/\sin(\pi/n)$
C_n	independent of n	any balanced superposition of ground state and highest excited state	position	$\max(\beta_O, \beta_E)$
Q_d	$\sim (\log n)^2$	any balanced superposition of ground state and highest excited state	position	β
S_n	$\exists\, n_{\text{opt}}$	balanced superposition of $\lvert\xi_\pm\rangle$ with relative phase ϕ_{opt}	exotic	independent of β

Table 4.1. For each family of graphs considered in the main text, we report the scaling of the QFI with the total number of nodes n, the optimal measurement saturating the quantum Cramér-Rao bound, the optimal initial preparation and the efficiency of an incomplete position measurement η. Notice that β (*resp.*, β_O, β_E) denotes the fraction of the graphs's nodes (*resp.*, nodes with even labels, odd labels) under individual control by the experimentalist.

Part II

Applications of orthogonal arrays in quantum information

Abstract

In this second part of the thesis, we will focus on some applications of orthogonal arrays in quantum information. Introduced by Rao in 1947 [66], orthogonal arrays have been usefully employed in different fields, from cryptography and coding theory to the statistical design of experiments, software testing and quality control [67]. Moreover, they generalize some well-known classes of combinatorial designs, e.g. Graeco-Latin squares, Hadamard matrices and classical codes [68]. Remarkably, orthogonal arrays have also found application in quantum information theory and, in particular, in the study of quantum entanglement [81].

Besides being arguably the most conceptually profound feature of quantum mechanics, entanglement has also been recognized as a crucial resource in many information processing tasks, from quantum metrology to quantum teleportation and quantum cryptography [69]. Whereas for bipartite systems many results are known and a canonical theory may be said to have been established, for multipartite system there remain several unsolved problems. Here, the term multipartite refers to systems made up of $N > 2$ separate entities. In particular, an unfinished project is to classify all the inequivalent ways in which the different parties can be entangled among themselves. This has proved extremely challenging, not only with respect to local operations and classical communication protocols, but also with respect to their stochastic analogues. In the following chapter, we will study a toy version of the entanglement classification problem, which is set in the space of states that are one-to-one related to orthogonal arrays.

Subsystems of multipartite systems are described by reduced density matrices, also known as quantum marginals. Knowledge of only a few marginals is not enough to reconstruct uniquely the global state, but may still suffice to compute important physical quantities. Crucially, not all collections of marginals can arise. There are strong, non-trivial constraints among quantum marginals, such as the famous Pauli exclusion principles for many-body systems of fermions. The problem, given a set of compatible marginals, of reconstructing the global state corresponding to those marginals, is known as a quantum marginal problem. In the second chapter to follow, we will employ orthogonal arrays in order to solve marginal problems for a few low-dimensional settings.

* * *

In the following, Chapter 5 contains original material taken from Ref. [1], while Chapter 6 is based on unpublished work.

[1] L. Seveso, D. Goyeneche, and K. Życzkowski, "Coarse-grained entanglement classification through orthogonal arrays," *J. Math. Phys.*, vol. 59, no. 7, p. 072203, Jul. 2018.

Chapter 5

Coarse-grained entanglement classification via orthogonal arrays

Classification of entanglement in quantum systems is an unfinished task, solved so far only for bipartite systems and for systems made up of three and four qubits. In this chapter, after lying the necessary groundworks in Sect. 5.1, we develop a coarse-grained classification framework for systems of N qudits that is based on the properties of orthogonal arrays with N columns and d symbols [107].

5.1 Preliminaries

5.1.1 Pure multipartite entanglement classification

A pure state of a system of N distinguishable particles, with local dimensions $\{d_i\}_{i=1}^N$, is described by a vector in the Hilbert space $\mathcal{H} = \mathbb{C}^{d_1} \times \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_N}$. By definition, a state is said to be *separable* if it can be written in the tensor product form $|\psi\rangle = |\psi^{(1)}\rangle \otimes |\psi^{(2)}\rangle \otimes \cdots \otimes |\psi^{(N)}\rangle$. If a state is not separable, then it is *entangled*. An important problem in quantum information is to classify, for any given setting, all the ways in which the N parties can be entangled among themselves. This is done by introducing an equivalence relation and organizing states into equivalence classes. Several results are known in the bipartite ($N = 2$) case, whereas for $N > 2$ the problem is largely unsolved.

Different notions of equivalence for entangled states have been considered in the literature. The most fine-grained is based on *local unitary* (LU) equivalence. Two states $|\psi\rangle$ and $|\phi\rangle$ are said to be LU-equivalent if there exist unitaries $\{U_i \in U(d_i)\}_{i=1}^N$ such that

$$|\psi\rangle = (U_1 \otimes U_2 \otimes \cdots \otimes U_N)\,|\phi\rangle \ . \tag{5.1.1}$$

Clearly, a local change of basis can have no effect on the global entanglement properties

of a state. However, LU-equivalence proves to be a far too detailed criterion. Even for systems made up only of qubits, the number of parameters necessary to label different classes grows exponentially with N. This can be proved as follows. A normalized N-qubit state requires the specification of $2^{N+1}-2$ real parameters (disregarding its global phase). On the other hand, the group $G_{LU} = SU(2) \times \ldots SU(2)$ has $3N$ real parameters. Therefore, the dimension of the generic equivalence class, obtained as the orbit of some reference state $|\psi\rangle$ under the action of the group G_{LU}, cannot exceed $3N$ (it is strictly less than $3N$, if $|\psi\rangle$ has a non-trivial stabilizer). It follows that at least $2^{N-1} - 3N - 2$ real numbers are needed to characterize LU-equivalence classes of N qubits [70–72].

A coarser equivalence relation can be obtained by enlarging the family of operations that, by definition, preserve the entanglement of a quantum state. From a physical point of view, a natural criterion is based on the family of *local operations and classical communication* (LOCC). It is assumed that each of the N particles is kept in a separate laboratory under the control of an experimentalist. The experimentalists are allowed to communicate among themselves via a classical channel. A typical LOCC protocol is then made up of several rounds. In each round, one of the experimentalist performs a measurement on her particle, keeps the post-measurement state and broadcasts the outcome of the measurement to the other laboratories. Two states $|\psi\rangle$ and $|\phi\rangle$ are said to be LOCC-equivalent if they can be converted one into the other, after any number of such rounds. Despite the intuitive appeal, there is no established LOCC classification in a multipartite setting. This should be contrasted with the bipartite scenario, where Nielsen's theorem [73] provides a clear-cut answer: a state $|\psi\rangle$ can be obtained from $|\phi\rangle$ by means of LOCC operations if and only if the vector of Schmidt coefficients of $|\psi\rangle$ majorizes that of $|\phi\rangle$.

A still more coarse-grained criterion can be introduced by assuming that interconversion under LOCC operations is required to occur only with some nonzero probability. Two states are then said to be SLOOC-equivalent, where SLOCC stands for *stochastic* LOCC. The benefit of considering SLOCC operations is that a convenient mathematical characterization is available. It can be proved [74] that two states $|\psi\rangle$ and $|\phi\rangle$ are SLOCC-equivalent if and only if there exist matrices $\{M_i\}_{i=1}^N$ with unit determinant and a scalar $\lambda \in \mathbb{C}$, such that

$$(M_1 \otimes M_2 \otimes \cdots \otimes M_N)\,|\psi\rangle = \lambda\,|\phi\rangle \ . \tag{5.1.2}$$

The task of finding all possible SLOCC-equivalence classes for given number of particles has been solved only for a few low-dimensional settings [74, 75]. In this regard, a particularly fruitful idea has been to look for *invariant quantities*. One should distinguish between SL-invariants and SLOCC-invariants. The special linear group $\mathrm{SL}(d, \mathbb{C})$ is the

group of $d \times d$ complex matrices with unit determinant. An SL-invariant is a function $f(\psi)$ of the computational basis coefficients of a state $|\psi\rangle$, which does not change under SL-transformations of the form $|\psi\rangle \to (M_1 \otimes M_2 \otimes \cdots \otimes M_N)\,|\psi\rangle$, with $M_i \in \mathrm{SL}(d_i, \mathbb{C})$. In general, such transformations do not preserve the normalization of the state. Suppose however that f is an *homogeneous* SL-invariant of degree d, i.e. $f(\lambda\psi) = \lambda^d f(\psi)$, $\forall \lambda \in \mathbb{C}$. If $\tilde{f}$ is an SL-invariant of degree $\tilde{d}$, then an SLOCC-invariant can be constructed as follows: $g(\psi) := [f(\psi)]^{\tilde{d}}/[\tilde{f}(\psi)]^d$. In fact, if $|\psi\rangle$ and $|\phi\rangle$ are SLOCC-equivalent (see Eq.(5.1.2)), then $g(\psi) = g(\phi)$. Vice versa, if two states have different values for any SLOCC-invariant, then they must belong to different equivalence classes.

Entanglement invariants have been employed, for instance, to find all SLOCC classes of three qubits [74]. There are six classes in total. One class is represented by unentangled states, of the form $|\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle$. Three other classes correspond to states that possess only bipartite entanglement, such as $|\psi_1\rangle \otimes |\psi_{23}\rangle$ (and its permutations). The remaining two classes are genuinely entangled. That there exist at least two such classes can be proven by considering the following SL-invariant, known as the *Cayley hyperdeterminant*,

$$\mathrm{Det}(\psi) = \sum_{i_1,j_1,k_1} \sum_{i_2,j_2,k_2} \sum_{i_3,j_3,k_3} \sum_{i_4,j_4,k_4} c_{i_1 j_1 k_1} c_{i_2 j_2 k_2} c_{i_3 j_3 k_3} c_{i_4 j_4 k_4} \epsilon_{i_1 i_2} \epsilon_{i_3 i_4} \epsilon_{j_1 j_2} \epsilon_{j_3 j_4} \epsilon_{k_1 k_2} \epsilon_{k_3 k_4} \ , \tag{5.1.3}$$

where $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ has been expanded on the computational basis as $|\psi\rangle = \sum_{i,j,k} c_{ijk}\,|ijk\rangle$ and ϵ_{ij} is the totally antisymmetric Levi-Civita tensor. For the two states,

$$|GHZ_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \ , \qquad |W_3\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle) \ . \tag{5.1.4}$$

one obtains $\mathrm{Det}(W_3) = 0 \neq \mathrm{Det}(GHZ_3)$. The conclusion is that the GHZ state and the W state are inequivalent under SLOCC operations, and thus give rise to two distinct entanglement classes. It can be proven that no additional class exists [74]. The GHZ class is said to be *generic*, meaning that the W state can be approximated to arbitrary precision by states belonging to the GHZ class, but the vice versa does not hold [76]. However, entanglement of the W-type is more robust to particle loss: by tracing out any single party from a GHZ state, one obtains a bipartite unentangled mixed state, which is not true for the W state.

Already for four qubits there is an infinite number of SLOCC classes. They can be organized into 9 continuous inequivalent families [75]. For higher numbers of qubits, the situation gets even more complicated. In fact, one would need at least $2^{N+1} - 6N - 2$ real parameters to parametrize all possible SLOCC classes of N qubits [72]. Therefore, an exhaustive classification currently appears to be out-of-reach.

5.1.2 Maximally-entangled multipartite states

An *entanglement measure* allows to quantify the degree of entanglement of any given quantum state. To introduce one, a common approach is to proceed axiomatically, by making a list of desirable properties that any such measure should reasonably satisfy. An entanglement measure should take real nonnegative values and, in particular, should vanish on the set of separable states. Another necessary requirement is that of *monotonicity*: the value of an entanglement measure cannot increase under LOCC operations. Finally, a desirable property, though not necessary, is convexity. Many multipartite entanglement measures have been studied in the literature [69, 77–79]. While in the bipartite setting all measures are maximized by the same set of maximally entangled states, in the multipartite case different measures lead to different sets of states.

Here, we take as a measure of multipartite entanglement the average entanglement entropy over its possible partitions [108]. Given a multipartite state $|\psi\rangle$, made up of N, d-level particles, and a subset $K \subset \{1, 2, \ldots, N\}$, consider the bipartite entanglement between the reduced state of the particles whose labels are in K, and the reduced state of the particles whose labels are in its complement, $\{1, 2, \ldots, N\} \setminus K$. In the bipartite setting, a natural entanglement measure is the entanglement entropy, i.e. the von Neumann entropy of any of its reduced states. We therefore give the following definition.

Definition 5.1.1. ***(k-uniform state)*** *A k-uniform state of a system of N particles, with d levels each, is such that the reduced state, obtained after tracing out any $N - k$ parties, is the maximally mixed state, i.e.*

$$\rho_K = \mathrm{tr}_{\setminus K}(|\psi\rangle\langle\psi|) = \mathbb{I}_{d^k} \,, \tag{5.1.5}$$

for any subset $K \subset \{1, 2, \ldots, N\}$ of cardinality $|K| = k$.

Remark 5.1.2. *If one introduces an entanglement measure equal to the average entanglement entropy with respect to all possible partitions into two sets of k and $N - k$ particles, respectively, then k-uniform states are the maximally entangled states with respect to such a measure.*

Remark 5.1.3. *If a state is k-uniform, then it is also k' uniform, for any $k' < k$. In particular, a k-uniform state with $k = \lfloor N/2 \rfloor$ is called* absolutely maximally entangled *(AME) [80]. It maximizes the entanglement entropy, averaged over all possible partitions.*

5.1.3 Orthogonal arrays

Constructing k-uniform states is, in general, a non-trivial optimization problem. An efficient way to generate them, developed in Refs. [81, 82], makes use of a class of combinatorial designs, known as *orthogonal arrays*.

Definition 5.1.4. ***(orthogonal array)*** *An orthogonal array* $\mathrm{OA}(r, N, d, k)$ *is a rectangular arrangement of* $r < d^N$ *rows and* N *columns, made up of symbols from the alphabet* $\mathscr{A}_d := \{0, 1, \ldots, d-1\}$*, such that, for any subarray of* k *columns, every possible combination of* k *symbols is repeated the same number* λ *of times. The parameters of the array are usually called the number of* runs r*, the number of* factors N*, the number of levels* d *and the* strength k*, respectively.*

This terminology has its roots in experimental design theory [83]. Every column of an OA represents one factor, each factor having d possible values; every row represents a different run or treatment combination. By definition, a *full factorial design* contains all possible d^N treatment combinations. Unless d and N are reasonably small, it becomes exponentially costly to perform a full factorial experiment. One thus resorts to a *fractional factorial design*, which contains only a subset of all possible runs. An orthogonal array is a particular type of fractional factorial design, with well-studied combinatorial and statistical properties (see Ref. [68]).

Example 5.1.5. *As an example, the following is an orthogonal array with four rows, three columns, two symbols and strength two:*

$$\mathrm{OA}(4, 3, 2, 2) = \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix} . \tag{5.1.6}$$

Remark 5.1.6. *For the purposes of this thesis, we do not distinguish between arrays differing only in the order of their runs. Moreover, if two arrays can be mapped one into the other via a suitable combination of interchanges of rows, interchanges of columns and permutations of symbols within any given column, they will be said to be* isomorphic.

The strength k is one of the most important characteristics of an orthogonal array: the higher the strength, the better one can distinguish among separate factorial effects (see Thm. (11.3) of Ref. [68] for a precise statement). However, for given values of r, N, and d, there are in general many arrays having the same strength. To measure their different usefulness, statisticians have introduced a number of more fine-grained quantifiers, known as *quality factors*. One of the most commonly employed is the *generalized resolution* (GR). In order to compute the generalized resolution, each entry a_{ij} of the array is encoded as a complex root of unity, by mapping each symbol $s \in \mathscr{A}_d$ to $\omega_s = \exp(2\pi i s/d)$. Given the multi-index $I = j_1 \ldots j_n$, denoting a subset of $n < N$ columns, the corresponding J-characteristic of order n is defined as

$$J_n(I) := \left| \sum_{i=1}^{r} \omega_{a_{ij_1}} \cdot \omega_{a_{ij_2}} \ldots \omega_{a_{ij_n}} \right| = \left| \sum_{i=1}^{r} e^{\frac{2\pi i}{d} \sum_{j \in I} a_{ij}} \right| . \tag{5.1.7}$$

Let t be the smallest integer such that there exists at least one J-characteristic of order t different from zero (clearly, $t > k$ since all possible J-characteristics of order up to k are zero). Define $J_t^{(\max)} := \max_{I:|I|=t} J_t(I)$. Then, the generalized resolution is equal to

$$GR := t + 1 - \frac{J_t^{(\max)}}{r} \ . \tag{5.1.8}$$

Notice that $t < GR < t + 1$. For two-level OAs (i.e. $d = 2$), the above quantity is invariant under isomorphisms. For multi-level arrays, however, permutations of symbols within columns can change the value of the RHS of Eq. (5.1.8). Hence, the generalized resolution of a multi-level OA is defined as the maximum value of (5.1.8), taken over the family of all isomorphic arrays [84].

Any $\mathrm{OA}(r, N, d, k)$ encodes a pure quantum state $|\psi_{N,d,k}\rangle \in \mathcal{H}_d^{\otimes N}$ of N qudits (with $\mathcal{H}_d \simeq \mathbb{C}^d$ the local Hilbert space of each subsystem) via the following mapping [81]:

$$\tau : OA(r, N, d, k) = \begin{matrix} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ a_{r,1} & a_{r,2} & \dots & a_{r,N} \end{matrix} \implies |\psi_{N,d,k}\rangle = \sum_{i=1}^{r} |a_{i,1}a_{i,2}, \dots, a_{i,N}\rangle \ . \tag{5.1.9}$$

A quantum state which is the image of an orthogonal array under τ is said to be *array-based*. For example,

$$\tau : \quad OA(2, 2, 2, 1) = \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} \quad \implies \quad |\psi_{2,2,1}\rangle = |00\rangle + |11\rangle \ , \tag{5.1.10}$$

which is the Bell state $|\Phi^+\rangle$ of two qubits. For convenience, in this chapter we neglect to normalize quantum states.

Let the set of all orthogonal arrays $\mathrm{OA}(r, N, d, k)$ with fixed values of the parameters N, d and k, but an arbitrary number of runs $r \leq d^N$, be denoted by $\mathscr{OA}(N, d, k)$. Its image under τ is denoted by $\mathcal{OA}(N, d, k)$: it is the set of array-based quantum states belonging to $\mathcal{H}_d^{\otimes N}$. We give the following additional definition.

Definition 5.1.7. ***(k-irredundant OA)*** *An orthogonal array* $\mathrm{OA}(r, N, d, k)$ *is said to be k-irredundant, denoted* $\mathrm{IrOA}(r, N, d, k)$*, if every subarray made up of $N - k$ columns contains no repeated rows.*

Remark 5.1.8. *An orthogonal array* $\mathrm{OA}(r, N, d, k)$ *of strength k is also an orthogonal array with strength k', for any $k' < k$. Thus, an array* $\mathrm{OA}(r, N, d, k)$ *might not be k-irredundant, but it could be k'-irredundant for some $k' < k$. Clearly, a k'-irredundant array* $\mathrm{IrOA}(r, N, d, k')$ *is also a k''-irredundant array, for any $k'' < k'$.*

The following proposition links k-irredundant orthogonal arrays to k-uniform states.

Proposition 5.1.9. *Given a k-irredundant orthogonal arrays* $\mathrm{IrOA}(r, N, d, k)$*, its image under the transformation τ, defined in Eq.* (5.1.9)*, is a k-uniform state.*

Proof. See Ref. [81]. □

Remark 5.1.10. *In particular, k-irredundant arrays of index unity give rise to k-uniform states with minimal support, i.e. k-uniform states minimizing the number of terms in their computational basis expansion. Thus, the number of runs r must be equal to d^k. Such arrays can be used to define a large class of classical codes, known as maximum distance separable (MDS) codes. MDS codes have the property that the Hamming distance between any two codewords, i.e. any two rows of* $\mathrm{IrOA}(d^k, N, d, k)$*, is a constant, taking the maximum possible value allowed by the Singleton bound. See Ref. [68] for more details.*

5.2 Orthogonal arrays and their Hilbert bases

Our aim is to provide a mathematical exploration of the spaces $\mathcal{OA}(N, d, k)$ of array-based quantum states. In particular, we will focus on a subset of arrays, referred to as *generating arrays*, that have special combinatorial properties. With the help of generating arrays, we will develop an algorithmic way to efficiently organize states into entanglement classes, for arbitrary values of the parameters N, d and k. Such array-based classification gives a coarse-grained picture of the different ways in which N qudits can be entangled among themselves in the Hilbert space $\mathcal{H}_d^{\otimes N}$ [107].

Our first task is to construct all possible array-based states for a given setting. We now introduce a systematic method which, in principle, allows to generate all elements of the set $\mathscr{OA}(N, d, k)$. In practice, computational complexity limits its applicability to low-dimensional cases [85].

Let us define the following composition function, denoted by $\oplus$,

$$\begin{array}{cccc} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ a_{r,1} & a_{r,2} & \dots & a_{r,N} \end{array} \oplus \begin{array}{cccc} a'_{1,1} & a'_{1,2} & \dots & a'_{1,N} \\ a'_{2,1} & a'_{2,2} & \dots & a'_{2,N} \\ \vdots & \vdots & & \vdots \\ a'_{s,1} & a'_{s,2} & \dots & a'_{s,N} \end{array} = \begin{array}{cccc} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ \vdots & \vdots & & \vdots \\ a_{r,1} & a_{r,2} & \dots & a_{r,N} \\ a'_{1,1} & a'_{1,2} & \dots & a'_{1,N} \\ \vdots & \vdots & & \vdots \\ a'_{s,1} & a'_{s,2} & \dots & a'_{s,N} \end{array} , \tag{5.2.1}$$

where it is assumed that $r + s \leq d^N$. That is, the operation $\oplus$ is a *partial* composition function: it is well-defined only for two arrays whose total number of runs is not greater

than d^N. Crucially, there always exists a subset of OAs in $\mathscr{OA}(N,d,k)$, called *generating arrays*, that are irreducible, i.e. they cannot be written as the composition of two arrays, whereas every other array can be obtained as a suitable linear combination of them. This can be shown as follows. From Eq. (5.1.9), an orthogonal array $\mathrm{OA}(r,N,d,k)$ gives rise to a quantum state $|\psi_{N,d,k}\rangle$,

$$|\psi_{N,d,k}\rangle = \sum_{i_1,\ldots,i_N=0}^{d-1} c_{i_1\cdots i_N} |i_1\cdots i_N\rangle \,, \tag{5.2.2}$$

where the coefficients $c_{i_1\cdots i_N}$ are nonnegative integers counting how many times the run $i_1\cdots i_N$ appears in the array. The coefficients $c_{i_1\cdots i_N}$ are not arbitrary: they satisfy a system of linear constraints to ensure orthogonality of the underlying array.

Example 5.2.1. *As an example, consider the case of $\mathscr{OA}(2,2,1)$. The coefficients of the state*

$$|\psi_{2,2,1}\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle \ . \tag{5.2.3}$$

are constrained to satisfy the linear system

$$\begin{cases} c_{00}+c_{01} = c_{10}+c_{11} \,, \\ c_{00}+c_{10} = c_{01}+c_{11} \,. \end{cases} \tag{5.2.4}$$

Eq. (5.2.4) simply states that each column of the array must contain the same number of zeros and ones, which implies that the underlying array gives rise to an OA of strength $k=1$. A further constraint is

$$c_{00}+c_{01}+c_{10}+c_{11} \leq 4 \,, \tag{5.2.5}$$

since the total number of runs cannot exceed that of a full factorial design.

As an additional example, in the case of $\mathscr{OA}(3,2,2)$, the coefficients of the state

$$\begin{aligned} |\psi_{3,2,2}\rangle \;=\; & c_{000}|000\rangle + c_{001}|001\rangle + c_{010}|010\rangle + c_{011}|011\rangle + \\ & c_{100}|100\rangle + c_{101}|101\rangle + c_{110}|110\rangle + c_{111}|111\rangle \,, \end{aligned} \tag{5.2.6}$$

must satisfy the linear system

$$\begin{cases} c_{000}+c_{001} = c_{010}+c_{011} = c_{100}+c_{101} = c_{110}+c_{111} \,, \\ c_{000}+c_{010} = c_{001}+c_{011} = c_{100}+c_{110} = c_{101}+c_{111} \,, \\ c_{000}+c_{100} = c_{001}+c_{101} = c_{010}+c_{110} = c_{011}+c_{111} \,, \end{cases} \tag{5.2.7}$$

together with a constraint similar to that of Eq. (5.2.5) *on the total number of runs. Eq.* (5.2.7) *expresses the fact that every two columns of the array must contain every possible pair of symbols the same number of times.*

Based on the previous examples, the following result can be established:

Proposition 5.2.2. *Any orthogonal array in $\mathscr{OA}(N,d,k)$ is uniquely associated to a quantum state*

$$|\psi_{N,d,k}\rangle = \sum_{i_1,\ldots,i_N=0}^{d-1} c_{i_1\cdots i_N}\,|i_1\cdots i_N\rangle, \tag{5.2.8}$$

with the coefficients $c_{i_1\cdots i_N}$ satisfying the linear constraints

$$\sum_{i_1,\ldots,i_{N-k}=0}^{d-1} c_{\sigma(i_1\cdots i_{N-k},\,j_1\cdots j_k)} = \sum_{i_1,\ldots,i_{N-k}=0}^{d-1} c_{\sigma(i_1\cdots i_{N-k},\,j'_1\cdots j'_k)}\ , \tag{5.2.9}$$

together with the further constraint

$$\sum_{i_1,\ldots,i_N=0}^{d-1} c_{i_1\cdots i_N} \leq d^N\ . \tag{5.2.10}$$

Here, $\sigma \in S_N$ is any permutation on the space of multi-indices of length N, and $j_1\cdots j_k$, $j'_1\cdots j'_k$ are any two multi-indices of length k.

Proof. Consider an arbitrary combination of k symbols, denoted by $j_1\ldots j_k$, where each symbol is taken from the alphabet $\mathscr{A}_d$. Then, the total number of times that such a combination appears in the last k columns of an array, denoted by $\#(j_1\ldots j_k)$, is obtained by summing over all possible combinations of symbols in the first $N-k$ columns of the array, i.e.

$$\#(j_1\ldots j_k) = \sum_{i_1,\ldots,i_{N-k}=0}^{d-1} c_{i_1\cdots i_{N-k},\,j_1\cdots j_k}\ . \tag{5.2.11}$$

By definition, any combination of k symbols occurs the same number of times in any k columns. In particular, this holds for the last k columns. Using Eq. (5.2.11) we have

$$\#(j_1\ldots j_k) = \sum_{i_1,\ldots,i_{N-k}=0}^{d-1} c_{i_1\cdots i_{N-k},\,j_1\cdots j_k} = \sum_{i_1,\ldots,i_{N-k}=0}^{d-1} c_{i_1\cdots i_{N-k},\,j'_1\cdots j'_k} = \#(j'_1\ldots j'_k)\ . \tag{5.2.12}$$

Imposing a similar condition for *any* possible choice of k columns leads us to Eq. (5.2.9), and thus proves our claim. Notice that Eq. (5.2.10) just expresses the fact that the number of runs r of the array is constrained to satisfy $r \leq d^N$. □

The linear system of Eq. (5.2.9) is made up of $(d^k-1)\times N!/[k!(N-k)!]$ equations constraining the coefficients of $|\psi_{N,d,k}\rangle$. If the coefficients are allowed to take nonnegative real values, each linearly independent constraint gives rise to an hyperplane in the positive orthant of $(\mathbb{R}^d)^{\otimes N}$. The intersection among all such hyperplanes produces a rational cone, denoted by $C(N,d,k)$, which is bounded because of the further constraint (5.2.10). The resulting convex polytope is denoted by $P(N,d,k)$. Points belonging to $P(N,d,k)$ with nonnegative integer coordinates are one-to-one related to orthogonal arrays in $\mathscr{OA}(N,d,k)$. We therefore arrive at the following proposition.

Proposition 5.2.3. *Any orthogonal array in $\mathscr{OA}(N,d,k)$ is uniquely associated to a point of the convex polytope $P(N,d,k)$ with nonnegative integer coordinates, where $P(N,d,k)$ is defined by the set of linear constraints* (5.2.9) *and* (5.2.10)*.*

The set $\mathscr{OA}(N,d,k)$ can thus be visualized as a d^N-dimensional integer lattice, which can be conveniently described with the help of the Gordan's lemma [86]. The lemma states that any integer point of a rational cone can be represented as a combination with nonnegative integer weights of a finite number of points. The requirement that such a set of points is as small as possible makes it unique. It is usually referred to as the *Hilbert basis* of the cone [87] and its computation is a well-studied problem: several software tools, such as the open-source package Normaliz [88] (to which we resort in the following), are available for this purpose.

We introduce the following terminology. The arrays in the set $\mathscr{OA}(N,d,k)$, which are represented by points of $P(N,d,k)$ belonging to the Hilbert basis of $C(N,d,k)$, are referred to as the *generating OAs*. The quantum states that are built on them are called *generating states*. For any given values of the parameters N, d and k, we denote the family of generating arrays by $\mathscr{G}(N,d,k)$ and the corresponding set of quantum states by $\mathcal{G}(N,d,k)$. The preceding discussion can be summarized via the following proposition.

Proposition 5.2.4. *The set $\mathscr{OA}(N,d,k)$ consists of all OAs of the form*

$$\mathrm{OA}(r,N,d,k)=\bigoplus_{i=1}^{m}\alpha_i\,\mathrm{OA}^{(i)}_{N,d,k}\,, \tag{5.2.13}$$

where $\alpha_i\in\mathbb{N}_0$, m is the cardinality of the Hilbert basis of the rational cone $C(N,d,k)$ and $\mathrm{OA}^{(i)}_{N,d,k}$ is the generating OA associated to the i^{th} element of the Hilbert basis. The numbers α_i are constrained to satisfy $\sum_{i=1}^{m}\alpha_i\,r^{(i)}\leq d^N$, where $r^{(i)}$ is the number of runs of $\mathrm{OA}^{(i)}_{N,d,k}$. Moreover, $\mathcal{OA}(N,d,k)$ consists of all states of the form

$$|\psi_{N,d,k}\rangle=\sum_{i=1}^{m}\alpha_i\;|g^{(i)}_{N,d,k}\rangle\;, \tag{5.2.14}$$

where $|g^{(i)}_{N,d,k}\rangle$ is the generating state based on $\mathrm{OA}^{(i)}_{N,d,k}$ and the α_i are constrained as above.

In Fig. 5.1, we illustrate Proposition 5.2.4 by showing a three-dimensional projection of the convex polytope $P(3,2,1)$, with the lattice points on its faces representing orthogonal arrays $\mathrm{OA}(r,3,2,1)$.

Example 5.2.5. *Let us return to the case of $\mathscr{OA}(2,2,1)$ discussed in Example 5.2.1. It can be checked that the Hilbert basis of $C(2,2,1)$ is made up of the following generating OAs,*

$$\mathrm{OA}^{(1)}_{2,2,1}=\begin{array}{cc}0&0\\1&1\end{array}\,,\qquad\qquad \mathrm{OA}^{(2)}_{2,2,1}=\begin{array}{cc}0&1\\1&0\end{array}\,. \tag{5.2.15}$$

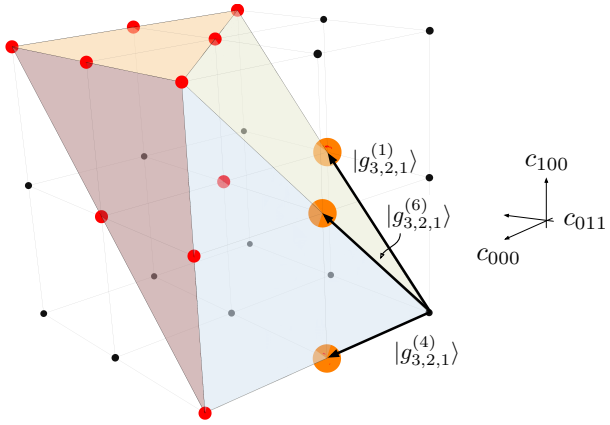


Figure 5.1. Three-dimensional projection of the convex polytope $P(3,2,1)$. The projection is obtained by imposing the constraint $c_{001} = c_{010} = 0$ and restricting the polytope to the box $[0,2]^{\times 3}$. It contains three out of the six generating states of the corresponding Hilbert basis, shown in orange *(larger points)*. Red *(medium)* points define quantum states associated with OAs that are convex superpositions of the generating states. Black *(smaller)* points in the lattice are outside the cone, and thus are not associated to orthogonal arrays.

The corresponding states are

$$|g^{(1)}_{2,2,1}\rangle = |00\rangle + |11\rangle \ , \qquad |g^{(2)}_{2,2,1}\rangle = |01\rangle + |10\rangle \ , \tag{5.2.16}$$

which, up to normalization, coincide with the Bell states $|\Phi^+\rangle$ *and* $|\psi^+\rangle$ *(the remaining two Bell states are not array-based, as they require a relative phase).*

Turning instead to the case of $\mathscr{OA}(3,2,2)$*, the Hilbert basis of the cone* $C(3,2,2)$ *consists of the orthogonal arrays*

$$\mathrm{OA}^{(1)}_{3,2,2} = \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{matrix} \ , \qquad \mathrm{OA}^{(2)}_{3,2,2} = \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix} \ , \tag{5.2.17}$$

giving rise to the states

$$|g^{(1)}_{3,2,2}\rangle = |001\rangle + |010\rangle + |100\rangle + |111\rangle \ , \qquad |g^{(2)}_{3,2,2}\rangle = |000\rangle + |011\rangle + |101\rangle + |110\rangle \ . \tag{5.2.18}$$

5.3 Entanglement classification for array-based states

In this section, we tackle the problem of classifying array-based states into entanglement classes. The first issue is to define the set of *free operations*. In the Hilbert space $\mathcal{H}_d^{\otimes N}$, the free operations are represented by local transformations of the form

$$|\psi\rangle \to (M_1 \otimes M_2 \otimes \cdots \otimes M_N)\,|\psi\rangle \ , \tag{5.3.1}$$

where $\{M_i\}_{i=1}^N$ are invertible complex matrices. However, such transformations, in general, are not *OA-preserving*: they do not necessarily map a generic array-based state into another array-based state.

Proposition 5.3.1. *The most general OA-preserving transformation is the local composition of transformations represented by invertible stochastic matrices with integer entries.*

Proof. Consider an invertible transformation acting on one party only: if $|\psi_{N,d,k}\rangle$ is an array-based quantum state, it is transformed to $T_{A_1}|\psi_{N,d,k}\rangle := (A_1 \otimes \mathbb{I}_d \otimes \cdots \otimes \mathbb{I}_d)|\psi_{N,d,k}\rangle$, with A_1 an invertible $d \times d$ matrix (without loss of generality the transformation is applied to the first party). For the transformation T_{A_1} to be OA-preserving, any two entries of A_1 must be in rational proportions. Since states are unnormalized, we may assume that all entries are integer. If some of the entries are negative, then $T_{A_1}|\psi_{N,d,k}\rangle$ is in general not array-based. Thus, we may further assume that all entries are nonnegative. Let us denote by $\alpha_{ij} \in \mathbb{N}_0$ the entries of A_1, with $i, j \in \mathscr{A}_d$. Then, by construction,

$$T_{A_1}|j\rangle \otimes |\phi\rangle = \sum_{i=0}^{d-1} \alpha_{ij}|i\rangle \otimes |\phi\rangle \ , \tag{5.3.2}$$

for arbitrary $|\phi\rangle \in \mathcal{H}_d^{\otimes N-1}$. Assuming that T_{A_1} is OA-preserving implies that

$$\sum_{j=0}^{d-1} \alpha_{i_1 j} = \sum_{j=0}^{d-1} \alpha_{i_2 j} \ , \tag{5.3.3}$$

for any choice of row indices i_1 and i_2, so that each symbol $s \in \mathscr{A}_d$ appears equally often along the first column of the array corresponding to $T_{A_1}|\psi_{N,d,k}\rangle$. Moreover, it is also necessary that each symbol appears equally often in any other column, which requires that

$$\sum_{i=0}^{d-1} \alpha_{i j_1} = \sum_{i=0}^{d-1} \alpha_{i j_2} \ , \tag{5.3.4}$$

for any choice of column indices j_1 and j_2. In conclusion, A_1 must have nonnegative integer entries that sum to the same constant c along either a row or a column, i.e. A_1 is an integer stochastic matrix or a generalized magic square with magic constant c [89]. In turn, if A_1 is an integer stochastic matrix, then $T_{A_1}|\psi_{N,d,k}\rangle$ is array-based, since the corresponding array is an OA with strength k at least 1. The most general OA-preserving transformation is obtained by composition, i.e. it is of the form $\otimes_j T_{A_j}$. □

Example 5.3.2. *As an example, let us consider the case of qubits. If A_j denotes a general invertible matrix of the form*

$$A_j = \begin{pmatrix} \alpha_{00} & \alpha_{01} \\ \alpha_{10} & \alpha_{11} \end{pmatrix} \ , \qquad \alpha_{00}\alpha_{11} - \alpha_{01}\alpha_{10} \neq 0 \ , \tag{5.3.5}$$

with $\alpha_{ij} \in \mathbb{N}_0$, in order to be integer stochastic, the following constraints must be satisfied:

$$\begin{aligned} \alpha_{00} + \alpha_{10} &= \alpha_{01} + \alpha_{11} \ , \\ \alpha_{00} + \alpha_{01} &= \alpha_{10} + \alpha_{11} \ . \end{aligned} \tag{5.3.6}$$

It follows that $\alpha_{00} = \alpha_{11} := \alpha$ and $\alpha_{01} = \alpha_{10} := \beta$, with $\alpha \neq \beta$. Thus, the most general OA-preserving transformation for qubits is the composition of transformations T_{A_j}, with

$$A_j = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \ , \qquad \alpha \neq \beta \ . \tag{5.3.7}$$

For instance, if $|\psi_{N,2,k}\rangle$ is the state based on OA$(r, N, 2, k)$, then

$$|\psi_{N,2,k}\rangle \to T_{A_1} |\psi_{N,2,k}\rangle = \alpha \, |\psi_{N,2,k}\rangle + \beta \, (\sigma_x \otimes \mathbb{I}_2 \cdots \otimes \mathbb{I}_2) \, |\psi_{N,2,k}\rangle \ , \tag{5.3.8}$$

or, for the corresponding orthogonal arrays,

$$OA(r, N, 2, k) \to \alpha \cdot OA(r, N, 2, k) \oplus \beta \cdot f_1[OA(r, N, 2, k)] \ , \tag{5.3.9}$$

where $f_j[OA]$ is obtained from OA by permuting symbols along the j^{th} column.

Transformations of the form (5.3.9) do not change the strength of the starting array. In fact, the following more general proposition can be established.

Proposition 5.3.3. *Any two OAs, which are pre-images with respect to the map τ of quantum states connected by an OA-preserving transformation, have the same strength.*

Proof. It is enough to check that, if $|\psi_{N,d,k}\rangle$ is based on OA(r, N, d, k), then $T_{A_1} |\psi_{N,d,k}\rangle$ is based on an OA having the same strength k. Here A_1 represents an integer stochastic matrix, with entries $\alpha_{ij} \in \mathbb{N}_0$ and magic constant $c = \sum_{i=0}^{d-1} \alpha_{ij} = \sum_{j=0}^{d-1} \alpha_{ij}$. For convenience, we introduce the following notation. Given a multi-index $I = j_1 j_2 \dots j_k$ with $|I| = k$, let π_I be the linear operator that acts on basis kets as follows: $\pi_I \, |i_1 i_2 \dots i_N\rangle = |i_{j_1} i_{j_2} \dots i_{j_k}\rangle$. For instance, $\pi_{23} \, |0123\rangle = |12\rangle$. Furthermore, if $s \in \mathscr{A}_d$, then ζ_s is the linear operator that acts on basis kets as $\zeta_s \, |i_1 i_2 \dots i_N\rangle = \delta_{s,i_1} \, |i_2 \dots i_N\rangle$. For instance, $\zeta_0 \, |0123\rangle = |123\rangle$, but $\zeta_1 \, |0123\rangle = 0$.

The fact that $|\psi_{N,d,k}\rangle$ is based on an OA of strength k is equivalent to the fact that, for any I with $|I| = k$,

$$\pi_I \, |\psi_{N,d,k}\rangle = \lambda \, \tau(\mathscr{F}_{k,d}) \ , \tag{5.3.10}$$

i.e. the projection of OA(r, N, d, k) to any k columns is a multiple of the full factorial design $\mathscr{F}_{k,d}$, with λ being the *index* of OA(r, N, d, k). Moreover, it also holds that, for any I with $|I| = k - 1$, and any $s \in \mathscr{A}_d$,

$$\pi_I \circ \zeta_s \, |\psi_{N,d,k}\rangle = \lambda \, \tau(\mathscr{F}_{k-1,d}) \ . \tag{5.3.11}$$

We have to check that $T_{A_1}|\psi_{N,d,k}\rangle$ is based on an OA of strength k. Write $|\psi_{N,d,k}\rangle$ in the following form,

$$\begin{aligned}|\psi_{N,d,k}\rangle &= \sum_{l=1}^{r/d}|0\rangle\otimes|\phi_{0l}\rangle + \sum_{l=1}^{r/d}|1\rangle\otimes|\phi_{1l}\rangle + \cdots + \sum_{l=1}^{r/d}|d-1\rangle\otimes|\phi_{d-1,l}\rangle \\ &= \sum_{l=1}^{r/d}\sum_{j=0}^{d-1}|j\rangle\otimes|\phi_{jl}\rangle \ ,\end{aligned}$$

where the $|\phi_{jl}\rangle$ are computational basis kets. Then, from Eq. (5.3.2), it follows that

$$T_{A_1}|\psi_{N,d,k}\rangle = \sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\sum_{i=0}^{d-1}\alpha_{ij}|i\rangle\otimes|\phi_{jl}\rangle \ . \tag{5.3.12}$$

Suppose that k is at least 2 (otherwise the statement is trivial). We have to check that for every I with $|I|=2$, the state $\pi_I|\psi_{N,d,k}\rangle$ is a multiple of $\tau(\mathscr{F}_{2,d})$. For instance, for $I=23$ one has

$$\begin{aligned}\pi_{23}\circ T_{A_1}\ |\psi_{N,d,k}\rangle &= \sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\sum_{i=0}^{d-1}\alpha_{ij}\,\pi_{23}(|i\rangle\otimes|\phi_{jl}\rangle) = \sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\left(\sum_{i=0}^{d-1}\alpha_{ij}\right)\pi_{12}|\phi_{jl}\rangle \\ &= c\sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\pi_{12}|\phi_{jl}\rangle = \ c\,\pi_{23}|\psi_{N,d,k}\rangle \\ &= \lambda c\,\tau(\mathscr{F}_{2,d}) \ .\end{aligned} \tag{5.3.13}$$

It is clear that the same holds for any choice of I not involving the first column. Consider instead $I=12$; then,

$$\begin{aligned}\pi_{12}\circ T_{A_1}\ |\psi_{N,d,k}\rangle &= \sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\sum_{i=0}^{d-1}\alpha_{ij}\,\pi_{12}(|i\rangle\otimes|\phi_{jl}\rangle) = \sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\sum_{i=0}^{d-1}\alpha_{ij}|i\rangle\otimes\pi_1|\phi_{jl}\rangle \\ &= \sum_{i=0}^{d-1}\sum_{j=0}^{d-1}\alpha_{ij}|i\rangle\otimes\pi_1\left(\sum_{l=1}^{r/d}|\phi_{jl}\rangle\right) = \sum_{i=0}^{d-1}\sum_{j=0}^{d-1}\alpha_{ij}|i\rangle\otimes\pi_1\circ\zeta_j|\psi_{N,d,k}\rangle \\ &= \lambda\sum_{i=0}^{d-1}\left(\sum_{j=0}^{d-1}\alpha_{ij}\right)|i\rangle\otimes\tau(\mathscr{F}_{1,d}) = \ \lambda c\sum_{i=0}^{d-1}|i\rangle\otimes\tau(\mathscr{F}_{1,d}) \\ &= \lambda c\,\tau(\mathscr{F}_{2,d}) \ .\end{aligned}$$

The same holds for any choice of I involving the first column and any other column.

If k were equal to 2, this would conclude the proof. If $k>2$, it is enough to notice that the previous computations rely on the two properties (5.3.10) and (5.3.11), which hold for any k. □

Therefore, OA-preserving transformations do not change the strength k. Moreover, they can only increase the generalized resolution of the corresponding array.

Proposition 5.3.4. *If a quantum state $|\psi\rangle$ is obtained from another array-based state $|\phi\rangle$ by an OA-preserving transformation, then the OA corresponding to $|\psi\rangle$ has generalized resolution not smaller than the array corresponding to $|\phi\rangle$.*

Proof. Define Ω to be the linear functional which acts on basis kets as follows: $\Omega\,|i_1 i_2 \dots i_n\rangle = \omega_{i_1}\omega_{i_2}\dots\omega_{i_n}$, where as before $\omega_s = \exp(2\pi i s/d)$ for $s \in \mathscr{A}_d$. Given the array-based state $|\psi_{N,d,k}\rangle$, the J-characteristic $J_n(I)$ of the OA $\tau^{-1}\,|\psi_{N,d,k}\rangle$ with multi-index I can be written as

$$J_n(I) = |\Omega \circ \pi_I\,|\psi_{N,d,k}\rangle|\ . \tag{5.3.14}$$

Suppose the transformation T_{A_1} is applied to $|\psi_{N,d,k}\rangle$. If our statement holds in this particular case, it also holds for the most general OA-preserving transformation. The J-characteristic of the OA $\tau^{-1}\circ T_{A_1}\,|\psi_{N,d,k}\rangle$ with multi-index I is denoted by $\tilde{J}_n(I)$. There are two cases: either I includes the first index or not. If not, from Eq. (5.3.12), it follows that:

$$\begin{aligned}\tilde{J}_n(I) &= |\Omega \circ \pi_I \circ T_{A_1}\,|\psi_{N,d,k}\rangle| = \left|\sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\left(\sum_{i=0}^{d-1}\alpha_{ij}\right)\Omega\circ\pi_I\,|\phi_{jl}\rangle\right| \\ &= c\,|\Omega\circ\pi_I\,|\psi_{N,d,k}\rangle| = c\,J_n(I)\ .\end{aligned} \tag{5.3.15}$$

We conclude that any J-characteristic $\tilde{J}_n(I)$, with I not involving the first column, is rescaled by c, the magic constant of the matrix A_1.

Now suppose I includes the first index, so that it may be written as $I = 1\cdot I'$, where $|I'| = n-1$. Then,

$$\begin{aligned}\tilde{J}_n(I) &= \left|\sum_{l=1}^{r/d}\sum_{j=0}^{d-1}\sum_{i=0}^{d-1}\alpha_{ij}\,\omega_i\,\Omega\circ\pi_{I'}\,|\phi_{jl}\rangle\right| = \left|\sum_{j=0}^{d-1}\left(\sum_{i=0}^{d-1}\omega_i\,\alpha_{ij}\right)\left(\sum_{l=1}^{r/d}\Omega\circ\pi_{I'}\,|\phi_{jl}\rangle\right)\right| \\ &= c\left|\sum_{i,j=0}^{d-1}\omega_i\beta_{ij}z_j\right|\ ,\end{aligned} \tag{5.3.16}$$

where we defined

$$\beta_{ij} := \frac{\alpha_{ij}}{c}\ , \qquad\qquad z_j := \sum_{l=1}^{r/d}\Omega\circ\pi_{I'}\,|\phi_{jl}\rangle\ . \tag{5.3.17}$$

Notice that β_{ij} is a bistochastic matrix of order d. In matrix notation, we may rewrite the summation on the right hand side of Eq. (5.3.17) as $\boldsymbol{\omega}^t\beta\mathbf{z}$, where $\boldsymbol{\omega}$ and $\mathbf{z}$ are complex column vectors. By Birkhoff theorem, the set of bistochastic matrices is given by the convex hull of all permutations of given size. Hence we can represent an arbitrary bistochastic matrix as a combination of permutation matrices σ_q. In particular one can write, $\beta = \sum_q \theta_q\sigma_q$, where the positive weights θ_q sum to unity, $\sum_q\theta_q = 1$. Then,

$$\tilde{J}_n(I) = c\left|\sum_q \theta_q\,\boldsymbol{\omega}^t\sigma_q\mathbf{z}\right| \le c\sum_q\theta_q\,|\boldsymbol{\omega}^t\sigma_q\mathbf{z}|\ . \tag{5.3.18}$$

Notice that $|\boldsymbol{\omega}^t \sigma_q \mathbf{z}|$ is the J-characteristic, for the same multi-index I, of the OA $\tau^{-1} \circ T_{\sigma_q^t} |\psi_{N,d,k}\rangle$ (which is isomorphic to the OA $\tau^{-1} |\psi_{N,d,k}\rangle$). This implies that

$$\tilde{J}_n(I) \;\leq\; c\mu \;, \qquad \text{with} \qquad \mu := \max_{\sigma} |\boldsymbol{\omega}^t \sigma_q \mathbf{z}| \;. \tag{5.3.19}$$

We conclude that any J-characteristic $\tilde{J}_n(I)$, with I involving the first column, is less or equal to c times the maximum of the J-characteristics, for the same multi-index I, of OAs obtained from $\tau^{-1} |\psi_{N,d,k}\rangle$ by a permutation of symbols in the first column.

In particular, our results imply that: if $J_n(I) = 0$, with I not involving the first index, then also $\tilde{J}_n(I) = 0$; if I involves the first index and $\mu = 0$, then also $\tilde{J}_n(I) = 0$. This in turn means that the quantity t appearing in definition (5.1.8) can never decrease under OA-preserving transformations. In particular, if it increases, then, recalling that $t < GR < t + 1$, also the generalized resolution increases. If it does not change, then one has to consider the maximum of the set of all J-characteristics of order t, taken over all possible multi-index I with $|I| = t$ and permutations of symbols within columns.

There are again two cases. Suppose that the multi-index achieving the maximum does not involve the first index. Then, from Eq. (5.3.15), $J_t^{(\max)}$ gets rescaled by c. However, since the number of rows is also rescaled by the same factor, the generalized resolution remains unchanged. Suppose instead that the multi-index achieving the maximum does not involve the first index. Then, we only know that $J_t^{(\max)}$ can not increase and, as a result, the resolution (5.1.8) can not decrease. □

By definition, the set of free operations is taken to consist of arbitrary compositions of OA-preserving transformations and their inverses, as well as relabellings of the local Hilbert spaces. Two array-based states are equivalent if they can be converted one into the other under free operations. All states equivalent under free operations form an entanglement class. Let us remark that isomorphic OAs always give rise to equivalent states. Indeed, given two isomorphic OAs, the transformation connecting their corresponding states is a composition of relabellings of the local Hilbert spaces (exchanges of columns) and local permutations (permutations of symbols within columns) – i.e., it is a free operation. However, it is important to emphasize that non-isomorphic OAs could produce equivalent states. For instance, the GHZ state for three qubit systems is based on the array

$$\mathrm{OA}(2,3,2,1) = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{matrix} . \tag{5.3.20}$$

This state is equivalent to the state based on the array

$$\mathrm{OA}(6,3,2,1) = \begin{matrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}, \tag{5.3.21}$$

via the OA-preserving transformation T_{A_1} with $A_1 = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$. However, from the point of view of orthogonal array theory, the OAs (5.3.20) and (5.3.21) are inequivalent.

The array-based entanglement classification just introduced manages to capture many genuinely different types of entanglement. In fact, for any choice of the parameters N, d and k, it always includes classes having the maximum and the minimum amounts of entanglement, as well as classes with different intermediate degrees of entanglement.

(i) There is always an entanglement class corresponding to the fully separable case, which is based on the full-factorial $\mathscr{F}_{N,d}$. This is because the orthogonal array $\mathrm{OA}(d^N, N, d, N)$ gives rise to the fully separable state $|+\rangle^{\otimes N}$, with $|+\rangle = |0\rangle + |1\rangle + \cdots + |d-1\rangle$.

(ii) Irredundant OAs of strength k give rise to k-uniform states. Since every irredundant OA has a number of runs $r < d^N$, all irredundant OAs of strength k are in $\mathscr{OA}(N,d,k)$. In particular, AME states, i.e. IrOAs with strength $k = \lfloor N/2 \rfloor$, maximize entanglement.

We now state a few simple results. Equivalent states in our classification are also equivalent according to the standard SLOCC classification, but the opposite does not hold. By Prop. 5.3.3 the strength k is the same within each class. Because of Prop. 5.3.4, each entanglement class contains a state such that the corresponding array has the least number of runs, and thus also generalized resolution at least as small as any other array in the same class. Such a state is essentially unique, in the sense that all other states are based on isomorphic arrays. This is because an integer stochastic transformation increases the number of runs by a factor c, equal to its magic constant, and an integer stochastic transformation with $c = 1$ is a permutation matrix. Thus, any two OA-based states in the same entanglement class having the same number of runs must be isomorphic.

We call such a state a *representative state* of the class. The corresponding OA minimizes the number of runs within each class. Every state based on a generating array is a representative state (but the opposite does not necessarily hold). This follows from the fact that a generating OA is always irreducible, by definition, whereas all states which are

not representative states are based on OAs which are reducible. Therefore, states based on non-isomorphic arrays among the generating OAs belong to different entanglement classes. The remaining representative states must be constructed by taking all allowed compositions of the generating OAs, and then identifying states connected via free operations. The array with the least number of runs within each class is the representative state of that class. Because there is only a finite number of allowed compositions, there is also a finite number of classes, for any number of parties N and any local dimension d.

5.4 Entanglement classes of generating states

In this section, we focus on the space $\mathcal{G}(N,d,k)$ of generating states and classify them into entanglement classes. Generating arrays tend to display higher degrees of entanglement than other states. For instance, $\mathcal{G}(N,d,k)$ contains all k-uniform states of minimal support, as stated by the following proposition.

Proposition 5.4.1. *Every maximum distance separable code gives rise to a quantum state based on a generating OA in the Hilbert basis of $C(N,d,k)$, for some values of the parameters N, d and k.*

Proof. Every vector $|\psi_{N,d,k}\rangle$ associated to an $\mathrm{OA}(r,N,d,k)$ can be obtained as a non-negative integer combination of the generating states, i.e. $|\psi_{N,d,k}\rangle = \sum_{i=1}^{m} \alpha_i \, |g^{(i)}_{N,d,k}\rangle$ with $\alpha_i \in \mathbb{N}_0$. If the OA associated to the i^{th} generator $|g^{(i)}_{N,d,k}\rangle$ has index $\lambda^{(i)}_{N,d,k}$ then the OA associated to the state $|\psi_{N,d,k}\rangle$ has index $\lambda = \sum_{i=i}^{m} \alpha_i \lambda^{(i)}_{N,d,k}$. Since OAs associated to MDS codes have index unity, i.e. $\lambda = 1$, and $\alpha_i \geq 0$, we conclude that $|\psi_{N,d,k}\rangle$ has to be a generator state $|g^{(i)}_{N,d,k}\rangle$, for a suitable value of $i \in [1,\ldots,m]$. □

In principle, the problem of finding all inequivalent classes of generating states can be solved for arbitrary values of the parameters N, d and k. However, considerations of computational complexity make our methods applicable only in a few lower-dimensional cases. Two distinct steps are involved: *(i)* computing the Hilbert basis of the cone $C(N,d,k)$, i.e. the set $\mathscr{G}(N,d,k)$ of generating arrays; *(ii)* finding within $\mathscr{G}(N,d,k)$ the non-isomorphic arrays, i.e. the representative states of the different classes. The first task is a standard integer programming problem. The second task can be efficiently carried out by encoding orthogonal arrays as vertex-colored graphs, in such a way that two arrays are isomorphic if and only if their graphs are isomorphic [90]. Testing for graph isomorphism can then be performed via specialized software (e.g. the open-source tool `nauty` [91]).

5.4.1 Three qubits

We begin by considering the case of three qubits ($N = 3$). Since $\mathscr{OA}(N,d,k) \subset \mathscr{OA}(N,d,k')$ whenever $k' < k$, we take the lowest possible value of the strength ($k = 1$). There are only two non-isomorphic arrays in $\mathscr{G}(3,2,1)$:

$$\mathrm{OA}^{(\mathrm{I})}_{3,2,1} = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{matrix}, \qquad \mathrm{OA}^{(\mathrm{II})}_{3,2,1} = \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix}. \tag{5.4.1}$$

Therefore, there are only two classes, with representative states:

$$\begin{aligned} |\phi^{(\mathrm{I})}_{3,2,1}\rangle &= |000\rangle + |111\rangle \ , \\ |\phi^{(\mathrm{II})}_{3,2,1}\rangle &= |000\rangle + |011\rangle + |101\rangle + |110\rangle \ . \end{aligned} \tag{5.4.2}$$

One can can also compute the entanglement classes of states belonging to $\mathcal{OA}(3,2,1)$. There are 9 classes, with representative states $|\psi^{(\gamma)}_{3,2,1}\rangle$:

$$\begin{aligned} |\psi^{(\mathrm{I})}_{3,2,1}\rangle &= |000\rangle + |111\rangle \ , \\ |\psi^{(\mathrm{II})}_{3,2,1}\rangle &= |000\rangle + |011\rangle + |101\rangle + |110\rangle \ , \\ |\psi^{(\mathrm{III})}_{3,2,1}\rangle &= 2\,|000\rangle + |001\rangle + |110\rangle + 2\,|111\rangle \ , \\ |\psi^{(\mathrm{IV})}_{3,2,1}\rangle &= 2\,|000\rangle + |011\rangle + |101\rangle + |110\rangle + |111\rangle \ , \\ |\psi^{(\mathrm{V})}_{3,2,1}\rangle &= 3\,|000\rangle + |011\rangle + |101\rangle + |110\rangle + 2\,|111\rangle \ , \\ |\psi^{(\mathrm{VI})}_{3,2,1}\rangle &= 2\,|000\rangle + |001\rangle + |011\rangle + |101\rangle + 2\,|110\rangle + |111\rangle \ , \\ |\psi^{(\mathrm{VII})}_{3,2,1}\rangle &= 2\,|000\rangle + |001\rangle + |010\rangle + |101\rangle + |110\rangle + 2\,|111\rangle \ , \\ |\psi^{(\mathrm{VIII})}_{3,2,1}\rangle &= |000\rangle + |001\rangle + |110\rangle + |111\rangle \ , \\ |\psi^{(\mathrm{IX})}_{3,2,1}\rangle &= |000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle \ . \end{aligned} \tag{5.4.3}$$

The first two states are based on the OAs of Eq. (5.4.1) and are unitarily equivalent to the GHZ state $|GHZ_3\rangle$ of three qubits. The states $|\psi^{(\gamma)}_{3,2,1}\rangle$ (with $\gamma = \mathrm{III}, \ldots, \mathrm{VI}$) are of the GHZ-type, but not unitarily equivalent to $|GHZ_3\rangle$. $|\psi^{(\mathrm{VII})}_{3,2,1}\rangle$ is of the W-type, since its 3-tangle vanishes. The state $|\psi^{(\mathrm{VIII})}_{3,2,1}\rangle$ has only bipartite entanglement, since it can be written as $|GHZ_2\rangle \otimes |+\rangle$. The state $|\psi^{(\mathrm{IX})}_{3,2,1}\rangle$ is fully separable and corresponds to the full factorial $\mathfrak{F}_{3,2}$. We have therefore been able to reproduce all qualitatively different types of entanglement for three qubits [74]. Let us remark that states that are inequivalent in the array-based classification can nonetheless be equivalent in $\mathcal{H}_d^{\otimes N}$, in which the set of free operations is larger.

5.4.2 Four qubits

We now move on to the case of 4 qubits. For $k = 1$, there are three non-isomorphic arrays in $\mathscr{G}(4, 2, 1)$:

$$\mathrm{OA}^{(\mathrm{I})}_{4,2,1} = \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{matrix}, \qquad \mathrm{OA}^{(\mathrm{II})}_{4,2,1} = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{matrix}, \qquad \mathrm{OA}^{(\mathrm{III})}_{4,2,1} = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{matrix}.$$

Hence, there are three entanglement classes in $\mathcal{G}(4, 2, 1)$, with representative states:

$$\begin{aligned} |\phi^{(\mathrm{I})}_{4,2,1}\rangle &= |0000\rangle + |1111\rangle \ , \\ |\phi^{(\mathrm{II})}_{4,2,1}\rangle &= |0000\rangle + |0011\rangle + |1101\rangle + |1110\rangle \ , \\ |\phi^{(\mathrm{III})}_{4,2,1}\rangle &= 2\,|0000\rangle + |0111\rangle + |1011\rangle + |1101\rangle + |1110\rangle \ . \end{aligned} \tag{5.4.4}$$

The first one coincides with the GHZ state $|GHZ_4\rangle$. The second is a 1-uniform state, but is inequivalent under SLOCC operations to the GHZ state. The third is a genuinely entangled state, with average purity of its single-party reductions equal to ~ 0.531. None of them is 2-uniform. In fact, it can be proven that there exists no 2-irredundant orthogonal array having 4 columns and 2 symbols, i.e. there is no AME state of 4 qubits.

Finding all classes in $\mathcal{OA}(4, 2, 1)$ is significantly more complicated than for 3 qubits. It turns out that there are a total of 1110 different classes. In general, it is reasonable to expect the number of classes to grow exponentially with the number of parties N, even if they are always a finite number by construction.

5.4.3 Five qubits and beyond

Finally, we consider systems made up of five qubits. For $k = 1$, there are eleven non-isomorphic arrays in $\mathscr{G}(5, 2, 1)$, giving rise to eleven entanglement classes in $\mathcal{G}(5, 2, 1)$. Below, we report a possible choice of representative states:

$$\begin{aligned} |\phi^{(\mathrm{I})}_{5,2,1}\rangle &= |00000\rangle + |11111\rangle \ , \\ |\phi^{(\mathrm{II})}_{5,2,1}\rangle &= |00111\rangle + |01000\rangle + |10100\rangle + |11011\rangle \ , \\ |\phi^{(\mathrm{III})}_{5,2,1}\rangle &= |00111\rangle + |01000\rangle + |10011\rangle + |11100\rangle \ , \\ |\phi^{(\mathrm{IV})}_{5,2,1}\rangle &= |00100\rangle + |00001\rangle + |00010\rangle + |11000\rangle + 2\,|11111\rangle \ , \\ |\phi^{(\mathrm{V})}_{5,2,1}\rangle &= |00000\rangle + |00110\rangle + |01001\rangle + |10011\rangle + |11101\rangle + |11110\rangle \ , \end{aligned}$$

$$
\begin{aligned}
|\phi^{(\mathrm{VI})}_{5,2,1}\rangle &= |00100\rangle + |01000\rangle + |00011\rangle + |10011\rangle + |11101\rangle + |11110\rangle \ , \\
|\phi^{(\mathrm{VII})}_{5,2,1}\rangle &= |01000\rangle + |00011\rangle + |00101\rangle + |00110\rangle + |11011\rangle + |11101\rangle + |11110\rangle + |10000\rangle \ , \\
|\phi^{(\mathrm{VIII})}_{5,2,1}\rangle &= |00000\rangle + 2\,|01111\rangle + |10101\rangle + |10110\rangle + |11000\rangle + |11011\rangle \ , \\
|\phi^{(\mathrm{IX})}_{5,2,1}\rangle &= |01000\rangle + |00000\rangle + 2\,|00111\rangle + |11011\rangle + |11101\rangle + |11110\rangle + |10000\rangle \ , \\
|\phi^{(\mathrm{X})}_{5,2,1}\rangle &= |01000\rangle + |00100\rangle + |00001\rangle + |00010\rangle + |10000\rangle + 3\,|11111\rangle \ , \\
|\phi^{(\mathrm{XI})}_{5,2,1}\rangle &= 3\,|00000\rangle + 2\,|01111\rangle + |11100\rangle + |11001\rangle + |11010\rangle + 2\,|10111\rangle \ .
\end{aligned}
$$

For six qubits, the problem of computing the Hilbert basis of the cone $C(6,2,k)$ is computationally out of reach for $k < 4$. We discuss the case $k = 4$, since the case $k = 5$ is included in it. There are three non-isomorphic generating arrays in $\mathscr{G}(6,2,4)$, giving rise to as many classes, with the following representative states:

$$
\begin{aligned}
|\phi^{(\mathrm{I})}_{6,2,4}\rangle =& |000000\rangle + |000001\rangle + |000110\rangle + |000111\rangle + |001010\rangle + |001011\rangle + |001100\rangle \\
&+ |001101\rangle + |010010\rangle + |010011\rangle + |010100\rangle + |010101\rangle + |011000\rangle + |011001\rangle \\
&+ |011110\rangle + |011111\rangle + |100010\rangle + |100011\rangle + |100100\rangle + |100101\rangle + |101000\rangle \\
&+ |101001\rangle + |101110\rangle + |101111\rangle + |110000\rangle + |110001\rangle + |110110\rangle + |110111\rangle \\
&+ |111010\rangle + |111011\rangle + |111100\rangle + |111101\rangle \ , \\
|\phi^{(\mathrm{II})}_{6,2,4}\rangle =& |000001\rangle + |000010\rangle + |000100\rangle + |000111\rangle + |001000\rangle + |001011\rangle + |001101\rangle \\
&+ |001110\rangle + |010000\rangle + |010011\rangle + |010101\rangle + |010110\rangle + |011001\rangle + |011010\rangle \\
&+ |011100\rangle + |011111\rangle + |100000\rangle + |100011\rangle + |100101\rangle + |100110\rangle + |101001\rangle \\
&+ |101010\rangle + |101100\rangle + |101111\rangle + |110001\rangle + |110010\rangle + |110100\rangle + |110111\rangle \\
&+ |111000\rangle + |111011\rangle + |111101\rangle + |111110\rangle \ , \\
|\phi^{(\mathrm{III})}_{6,2,4}\rangle =& 2\,|000000\rangle + |000001\rangle + |000010\rangle + |000011\rangle + |000100\rangle + |000101\rangle + |000110\rangle \\
&+ 2\,|000111\rangle + 2\,|001001\rangle + 2\,|001010\rangle + |001011\rangle + 2\,|001100\rangle + |001101\rangle \\
&+ |001110\rangle + |001111\rangle + |010000\rangle + |010001\rangle + |010010\rangle + 2\,|010011\rangle + |010100\rangle \\
&+ 2\,|010101\rangle + 2\,|010110\rangle + 2\,|011000\rangle + |011001\rangle + |011010\rangle + |011011\rangle \\
&+ |011100\rangle + |011101\rangle + |011110\rangle + 2\,|011111\rangle + |100000\rangle + |100001\rangle + |100010\rangle \\
&+ 2\,|100011\rangle + |100100\rangle + 2\,|100101\rangle + 2\,|100110\rangle + 2\,|101000\rangle + |101001\rangle \\
&+ |101010\rangle + |101011\rangle + |101100\rangle + |101101\rangle + |101110\rangle + 2\,|101111\rangle + |110000\rangle \\
&+ 2\,|110001\rangle + 2\,|110010\rangle + 2\,|110100\rangle + 3\,|110111\rangle + |111000\rangle + |111001\rangle \\
&+ |111010\rangle + 2\,|111011\rangle + |111100\rangle + 2\,|111101\rangle + 2\,|111110\rangle \ . \qquad (5.4.5)
\end{aligned}
$$

Let us point out that the absolutely maximally entangled states of 5 and 6 qubits cannot be obtained from the generating states listed above, since AME$(5,2)$ and AME$(6,2)$ are not array-based [109].

We conclude this section by remarking that our methods can also be applied to *heterogeneous* systems [82], i.e. systems made up of particles with different numbers of internal levels. For instance, we have considered the families *(i)* $\mathscr{OA}(2^2 3^1, k)$ and *(ii)* $\mathscr{OA}(2^1 3^2, k)$, with $k = 1, 2$. Case *(i)* corresponds to system of two qubits and one qutrit. Vice versa, case *(ii)* corresponds to systems of one qubit and two qutrits. Here, we report the representative states for the entanglement classes of $\mathcal{G}(2^2 3^1, 1)$ and $\mathcal{G}(2^1 3^2, 1)$. For two qubits and one qutrit, there are six classes, with representative states:

$$\begin{aligned}
|\phi^{(\mathrm{I})}_{2^2 3^1,1}\rangle &= |000\rangle + |110\rangle + |001\rangle + |012\rangle + |102\rangle + |111\rangle \ , \\
|\phi^{(\mathrm{II})}_{2^2 3^1,1}\rangle &= |000\rangle + |010\rangle + |101\rangle + |112\rangle + |111\rangle + |002\rangle \ , \\
|\phi^{(\mathrm{III})}_{2^2 3^1,1}\rangle &= 2\,|000\rangle + |011\rangle + |101\rangle + |112\rangle + |112\rangle \ , \\
|\phi^{(\mathrm{IV})}_{2^2 3^1,1}\rangle &= |000\rangle + |110\rangle + 2\,|001\rangle + |112\rangle + |112\rangle \ , \\
|\phi^{(\mathrm{V})}_{2^2 3^1,1}\rangle &= |000\rangle + |110\rangle + |001\rangle + |002\rangle + |111\rangle + |112\rangle \ , \\
|\phi^{(\mathrm{VI})}_{2^2 3^1,1}\rangle &= 2\,|000\rangle + |011\rangle + |102\rangle + |112\rangle + |111\rangle \ .
\end{aligned} \tag{5.4.6}$$

Instead, for systems of one qubit and two qutrits, one finds fifteen classes, with the following representative states:

$$\begin{aligned}
|\phi^{(\mathrm{I})}_{2^1 3^2,1}\rangle &= |000\rangle + |112\rangle + |121\rangle + |100\rangle + |012\rangle + |021\rangle \ , \\
|\phi^{(\mathrm{II})}_{2^1 3^2,1}\rangle &= |000\rangle + |111\rangle + |102\rangle + |120\rangle + |012\rangle + |021\rangle \ , \\
|\phi^{(\mathrm{III})}_{2^1 3^2,1}\rangle &= 2\,|000\rangle + |111\rangle + |112\rangle + |121\rangle + |022\rangle \ , \\
|\phi^{(\mathrm{IV})}_{2^1 3^2,1}\rangle &= |000\rangle + |121\rangle + |112\rangle + |110\rangle + |002\rangle + |021\rangle \ , \\
|\phi^{(\mathrm{V})}_{2^1 3^2,1}\rangle &= |000\rangle + |120\rangle + |112\rangle + |111\rangle + |002\rangle + |021\rangle \ , \\
|\phi^{(\mathrm{VI})}_{2^1 3^2,1}\rangle &= 2\,|000\rangle + 2\,|111\rangle + |122\rangle + |022\rangle \ , \\
|\phi^{(\mathrm{VII})}_{2^1 3^2,1}\rangle &= |000\rangle + |111\rangle + |102\rangle + |120\rangle + |011\rangle + |022\rangle \ , \\
|\phi^{(\mathrm{VIII})}_{2^1 3^2,1}\rangle &= |000\rangle + |112\rangle + |111\rangle + |101\rangle + |020\rangle + |022\rangle \ , \\
|\phi^{(\mathrm{IX})}_{2^1 3^2,1}\rangle &= |000\rangle + |111\rangle + |112\rangle + |120\rangle + 3\,|101\rangle + 2\,|010\rangle + 3\,|022\rangle \ , \\
|\phi^{(\mathrm{X})}_{2^1 3^2,1}\rangle &= |000\rangle + 3\,|111\rangle + 3\,|102\rangle + 3\,|020\rangle + |021\rangle + |012\rangle \ , \\
|\phi^{(\mathrm{XI})}_{2^1 3^2,1}\rangle &= 4\,|000\rangle + 3\,|111\rangle + 3\,|122\rangle + |012\rangle + |021\rangle \ , \\
|\phi^{(\mathrm{XII})}_{2^1 3^2,1}\rangle &= 3\,|000\rangle + |120\rangle + |101\rangle + 2\,|111\rangle + 2\,|112\rangle + |021\rangle + 2\,|022\rangle \ , \\
|\phi^{(\mathrm{XIII})}_{2^1 3^2,1}\rangle &= |000\rangle + |112\rangle + |121\rangle + 2\,|110\rangle + 2\,|101\rangle + |011\rangle + 2\,|022\rangle + |002\rangle + |020\rangle \ , \\
|\phi^{(\mathrm{XIV})}_{2^1 3^2,1}\rangle &= 2\,|000\rangle + |112\rangle + |121\rangle + 2\,|110\rangle + 2\,|101\rangle + |011\rangle + 3\,|022\rangle \ , \\
|\phi^{(\mathrm{XV})}_{2^1 3^2,1}\rangle &= |000\rangle + 2\,|112\rangle + 2\,|121\rangle + |110\rangle + |101\rangle + |011\rangle + 2\,|002\rangle + 2\,|020\rangle \ .
\end{aligned} \tag{5.4.7}$$

5.5 Conclusions

In this chapter, we have introduced a family of multipartite states of N qudits that are one-to-one related to orthogonal arrays. In particular, we have focused our attention on a subfamily of states, referred to as generating states, with a remarkable combinatorial property: every orthogonal array with given parameters can be obtained as a composition of the arrays associated to them. The generating states include all irredundant orthogonal arrays having minimum support or, equivalently, all MDS codes. Next, we have tackled the problem of classifying the generating states into entanglement classes. This task requires to identify the non-isomorphic arrays within the set of generating arrays. To illustrate our methods, we have worked out all entanglement classes of generating arrays for systems of up to five qubits, as well as heterogeneous systems of two qubits + one qutrit, and one qubit + two qutrits.

Chapter 6

Orthogonal arrays and quantum marginal problems

In this chapter, we establish a connection between orthogonal array theory and quantum marginal problems, which is then exploited to provide constructive solutions, applicable to infinitely many settings, including the already solved case of N-qubit systems for arbitrary N. In particular, Sect. 6.1 provides a concise introduction to quantum marginal problems, emphasizing their relevance in the context of quantum information theory. Subsequent sections, starting from Sect. 6.2, are based on original unpublished work.

6.1 Quantum marginal problems

The global Hilbert space $\mathcal{H}$ for a system of N particles, each one having d possible levels, has the tensor product structure $\mathcal{H} = \mathcal{H}_d^{\otimes N}$, where $\mathcal{H}_d = \mathbb{C}^d$. A pure state $|\psi_{N,d}\rangle \in \mathcal{H}$ is thus specified by d^N complex numbers, an amount exponentially growing with the number of parties. In order to escape the curse of dimensionality, a natural idea is to do without a complete multi-body description and work instead only with the reduced density matrices describing few-body subsets of particles. If O_K denotes the generic operator acting non-trivially only on a subset $K \subset \{1, 2, \ldots, N\}$ of subsystems, then the *reduced density matrix* $\rho_K \in \mathsf{Her}^+_{|K|d}(\mathbb{C})$ is defined as the matrix such that $\langle O_K \rangle := \langle \psi_{N,d} | O_K | \psi_{N,d} \rangle = \mathrm{tr}(\rho_K O_K)$. It can be proven that such matrix is unique and that it can be obtained as $\rho_K = \mathrm{tr}_{\setminus K}(|\psi_{N,d}\rangle \langle \psi_{N,d}|)$, where $\mathrm{tr}_{\setminus K}(\cdot)$ denotes the partial trace over the Hilbert spaces of particles whose labels are not in K. Knowledge of all reduced density matrices $\{\rho_K\}_{|K|=n}$ is clearly sufficient to compute the expectation values of all possible operators acting only on at most n particles at a time.

The partial trace operation is the quantum version of marginalization in classical

probability. Given N real random variables $\mathbb{X}_1, \ldots \mathbb{X}_N$, with joint probability density function p, for any subset $K \subset \{1, \ldots, N\}$ the marginal distribution of $\{\mathbb{X}_i\}_{i \in K}$ is defined as $p_I = \int_{\setminus K} dx\, p(x_1, \ldots, x_n)$, where $\int_{\setminus K}$ denotes integration over the sample spaces of all random variables $\{\mathbb{X}_i\}_{i \notin K}$. The multi-body wavefunction $\psi_{N,d}$ can thus be considered as the analogue of the joint probability density function p and the reduced density matrix ρ_K as the analogue of the marginal p_K. Hence, ρ_K is often referred to as a *quantum marginal* and from now on we use the terms quantum marginal and reduced density matrix interchangeably.

Knowledge of how quantum marginals relate to their global states allows to recover fundamental properties of the system under study. For instance, consider a multi-body system with Hamiltonian H that contains only pairwise interactions, i.e. $H = \sum_{i<j} h_{ij}$. The ground state energy E_0 of the system may be computed as

$$E_0 = \min_{|\psi_{N,d}\rangle \in \mathcal{H}} \langle \psi_{N,d}|H|\psi_{N,d}\rangle = \min_{\rho_{ij} \text{ comp.}} \sum_{i<j} \mathrm{tr}(h_{ij}\rho_{ij}) \ ,$$

where ρ_{ij} denotes the two-body quantum marginal of the i^{th} and j^{th} subsystems. Crucially, the second minimization is performed over all possible two-body density matrices that are *compatible* with the existence of a global N-body pure state $|\psi_{N,d}\rangle$, i.e. they may be obtained from $|\psi_{N,d}\rangle$ by a suitable partial trace operation over $N-2$ parties. Seemingly, an exponential speed-up can be achieved by minimizing the objective function $\langle H \rangle$ over the set of compatible two body marginals (whose dimension scales as $N^2 d^4$), instead of the set of pure N-body states (whose dimension scales as d^N). This however would be true assuming one were able to solve efficiently the following general problem: given a collection of density matrices $\{\rho_{K_i}\}_{i=1}^m$, where the sets $K_i \subset \{1, \ldots, N\}$ may or may not have empty intersection between each other, are they compatible? That is, is there a pure state $|\psi_{N,d}\rangle$ such that

$$\rho_{K_i} = \mathrm{tr}_{\setminus K_i}(|\psi_{N,d}\rangle \langle \psi_{N,d}|) \ , \quad \forall i \in \{1, \ldots, m\} \quad ?$$

Such problems are often referred to as quantum marginal problems (QMPs). First formulated within the quantum chemistry community in the 1960s [92], they have recently come into focus from a quantum information perspective (see e.g. [93] for a review).

One should distinguish between two main types of quantum marginal problems: either the marginals are allowed to overlap or not. The first case has proven extremely difficult to tackle. In fact, from our previous argument, if an efficient algorithm were to be discovered to solve such problem, then one could efficiently compute the ground state energy of any multi-body system – a task that appears out of reach. It is known that even a universal quantum computer would not be able to solve efficiently the overlapping version of the

marginal problem: its algorithmic complexity belongs to the QMA-complete class [94], the quantum analogue of the NP-complete class [95]. A simpler version of the problem is when the marginals are *non-overlapping*. Without loss of generality, one may even assume that all assigned density operators are single-body, i.e. the problem is *univariate*.

Finally, an extension of the quantum marginal problem which is also of great interest is the following. Given a set of compatible marginals, explicitly construct the global state having those marginals. The task is known to be computationally hard and, except for a few solutions to univariate problems in low-dimensional settings, analytical progress has been quite limited.

For the purposes of this thesis, we adopt the following terminology:

- *Non-overlapping QMP*: any two marginals are assumed to involve subsets of particles that do not overlap; otherwise the problem is said to be overlapping.

- *Univariate QMP*: all marginals are assumed to be 1-body marginals; otherwise the problem is said to be multivariate.

- *Non-constructive QMP*: given N d-level particles and $n \in \mathbb{N}$, find all possible sets $\{\rho_K\}_{|K|=n}$ of compatible n-body marginals.

- *Constructive QMP*: given any set of compatible n-body marginals $\{\rho_K\}_{|K|=n}$, construct the global state $|\psi_{N,d}\rangle$ having those marginals.

Up to now, it has been implicitly assumed that the system is made up of *distinguishable* particles. It is also possible to consider quantum marginal problems for *indistinguishable* particles. One has to add the further requirement that the global state lies within either the symmetric subspace $\vee^{\otimes N}\mathcal{H}_d$ (for bosons) or the totally antisymmetric one $\wedge^{\otimes N}\mathcal{H}_d$ (for fermions).

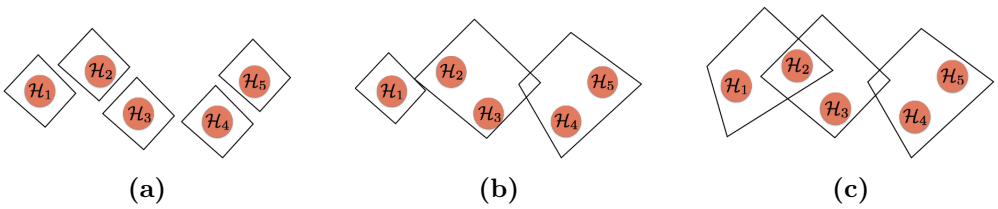


Figure 6.1. Different types of quantum marginal problems: **(a)** univariate, non-overlapping, **(b)** multivariate, non-overlapping, **(c)** multivariate, overlapping.

6.1.1 Univariate QMPs: overview of solutions

Quantum marginal problems of the univariate type are the only ones that have proved tractable. In the following, we will thus restrict to this scenario. The basic reason for its greater manageability is that the problem has a unitary invariance that disappears in the overlapping case, as illustrated by the following proposition.

Proposition 6.1.1. *If $\{\rho_k\}_{k=1}^N$ is a set of compatible 1-body marginals, then $\{U_k\rho_k U_k^\dagger\}_{k=1}^N$, where each $U_k \in U(d)$ is an arbitrary unitary matrix, is also a set of compatible marginals.*

Proof. Since the marginals $\{\rho_k\}_{k=1}^N$ are compatible, there exists a global state $|\psi_{N,d}\rangle$ having those marginals. It is easy to check that $\otimes_{k=1}^N U_k \, |\psi_{N,d}\rangle$ is then the global state having marginals $\{U_k\rho_k U_k^\dagger\}_{k=1}^N$. □

As a consequence, compatibility depends only on the local spectra. In particular, one may assume that each 1-body marginal ρ_k is given in diagonal form, with the eigenvalues $\{\lambda_k^{(j)}\}_{j=0}^{d-1}$ ordered non-increasingly along its main diagonal. This simplification will always be made use of in the following. Notice that, in $\lambda_k^{(j)}$, a subscript denotes the party under consideration (ranging from 1 to N), while a superscript labels the eigenvalues of each party in non-increasing order (from 0 to $d-1$).

It is natural to embed the eigenvalues corresponding to a set of compatible marginals as a point in $\mathbb{R}_+^{Nd}$. That is, each such set gives rise to a point $\boldsymbol{\lambda} \in \mathbb{R}_+^{Nd}$ of the general form $\boldsymbol{\lambda} = (\lambda_1^{(0)}, \ldots, \lambda_1^{(d-1)}, \ldots, \lambda_N^{(0)}, \ldots, \lambda_N^{(d-1)})^t$, which is referred to as a λ-point. We denote by $\Lambda_{N,d}$ the set of all possible λ-points, for given number of parties N and local dimension d. It is a corollary of a theorem by Kirwan [96] that $\Lambda_{N,d}$ is always a convex polytope, which is referred to as the *moment polytope*, or *spectral polytope*. The 1-body marginal problem can thus be given the following geometrical representation. If compatibility of quantum marginals introduced no non-trivial constraints, λ-points would fill the whole *Weyl polytope* $\Omega_{N,d}$, which is the convex set defined via the following system of linear constraints

$$\begin{cases} \lambda_k^{(0)} \geq \lambda_k^{(1)} \geq \cdots \geq \lambda_k^{(d-1)} \geq 0 \\ \sum_{j=0}^{d-1} (\lambda_k^{(j)})^2 = 1 \end{cases} \qquad \text{with } k \in \{1, \ldots, N\}, \tag{6.1.1}$$

where the inequality constraints introduce a conventional ordering among the eigenvalues of each reduction and the equality constraints enforce normalization. However, compatibility usually *does* introduce non-trivial constraints among the local eigenvalues, so that $\Lambda_{N,d}$ is a strict and convex subset of $\Omega_{N,d}$. The spectral polytope can be characterized either as the convex hull of its extremal points (its V-representation), or in terms of a finite list of linear inequalities of the form $\boldsymbol{\kappa} \cdot \boldsymbol{\lambda} + \kappa_0 \leq 0$ (its H-representation). For

given N and d, solving a non-constructive QMP is equivalent to determining the spectral polytope $\Lambda_{N,d}$. Instead, a constructive QMP requires finding, for the generic λ-point $\boldsymbol{\lambda} \in \Lambda_{N,d}$, at least one pure state $|\psi_{N,d}\rangle \in \mathcal{H}$ having local eigenvalues $\boldsymbol{\lambda}$. Notice that there is, in general, no guarantee of uniqueness and, in fact, many different pure states exist all having the same local spectra.

In the rest of this section, we give a quick historical review of the literature of known solutions. First, we focus on the case of distinguishable particles and thereafter consider indistinguishable particles.

For distinguishable particles, the following cases have been solved.

- $N = 2$, *d arbitrary*: Any pure bipartite state can be written as

$$|\psi_{2,d}\rangle = \sum_{j=0}^{d-1} \sqrt{\lambda^{(j)}}\, |\phi_j\rangle \otimes |\chi_j\rangle \ , \tag{6.1.2}$$

  where $\{|\phi_j\rangle\}_{j=0}^{d-1}$ and $\{|\chi_j\rangle\}_{j=0}^{d-1}$ are two local orthonormal basis. It follows that the 1-body reductions ρ_1 and ρ_2 have the same spectra. This solves the non-constructive version of the problem. The constructive version is also easily solved: if $\{\lambda^{(j)}\}_{j=0}^{d-1}$ denotes the spectrum of either of the two marginals, the solution is given by Eq. (6.1.2), where $\{|\phi_j\rangle\}_{j=0}^{d-1}$ and $\{|\chi_j\rangle\}_{j=0}^{d-1}$ are the the eigenvectors of ρ_1 and ρ_2, respectively.

- *N arbitrary,* $d = 2$: The case of N qubits has been completely solved by Higuchi *et al.* [97]. To conform to the conventions of Ref. [97], let us denote by $\lambda_k := \min \mathrm{spec}(\rho_k)$ the minimum eigenvalue of each reduction; the remaining one can always be recovered as $1 - \lambda_k$. Then, the spectral polytopes $\Lambda_{N,2}$ are the subsets of the Weyl polytopes $\Omega_{N,2}$ cut out by the *polygonal inequalities*

$$\lambda_k \leq \sum_{k' \neq k} \lambda_{k'} \ , \qquad k \in \{1, \ldots, N\} \ . \tag{6.1.3}$$

  Higuchi *et al.* also managed to solve the constructive version of the problem as follows. Let $\{\lambda_k\}_{k=1}^{3}$ denote a generic triple of compatible eigenvalues satisfying $0 \leq \lambda_k \leq 1/2$, as well as the polygonal inequalities of Eq. (6.1.3). The task is thus to construct a state $|\psi_{3,2}\rangle \in \mathcal{H}_2^{\otimes 3}$ with 1-body reductions having minimum eigenvalues $\{\lambda_k\}_{k=1}^{3}$. A solution is

$$|\psi_{3,2}\rangle = a\,|100\rangle + b\,|010\rangle + c\,|001\rangle + d\,|111\rangle \ , \tag{6.1.4}$$

  where from now on $\{|i\rangle\}_{i=0}^{d-1}$ denotes the computational basis of each local Hilbert space $\mathcal{H}_d$ and

$$\begin{aligned} a &= \sqrt{(\lambda_2 + \lambda_3 - \lambda_1)/2}\ , \quad b = \sqrt{(\lambda_3 + \lambda_1 - \lambda_2)/2}\ , \\ c &= \sqrt{(\lambda_1 + \lambda_2 - \lambda_3)/2}\ , \quad d = \sqrt{(2 - \lambda_1 - \lambda_2 - \lambda_3)/2}\ . \end{aligned} \tag{6.1.5}$$

Starting from $|\psi_{3,2}\rangle$ one can recursively construct solutions for higher numbers of qubits. The procedure is most easily explained in the simple case of $N = 4$, from which one can generalize to arbitrary N. Let $\{\lambda_k\}_{k=1}^4$ be the minimum eigenvalues of the single-body reductions of $|\psi_{4,2}\rangle \in \mathcal{H}_2^{\otimes 4}$. Without loss of generality, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. Define $\Delta_1 = \lambda_1 - \lambda_4$. Then $(\Delta_1, \lambda_2, \lambda_3)^t$ is a λ-point in $\Lambda_{3,2}$ (i.e. it satisfies the polygonal inequalities). Thus, there exists a corresponding pure state $|\tilde{\psi}_{3,2}\rangle$ which has the same form as in Eq. (6.1.4) but with Δ_1 taking the place of λ_1. For suitable (unnormalized) states $|\phi\rangle$ and $|\chi\rangle$ in $\mathcal{H}_2^{\otimes 2}$, one may write it as $|\tilde{\psi}_{3,2}\rangle = |0\rangle \otimes |\phi\rangle + |1\rangle \otimes |\chi\rangle$. It is then easy to check that the desired state corresponding to the λ-point $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)^t \in \Lambda_{4,2}$ can be taken as

$$|\psi_{4,2}\rangle = |0\rangle \otimes |\phi\rangle \otimes |1\rangle + \sin\theta\, |0\rangle \otimes |\chi\rangle \otimes |0\rangle + \cos\theta\, |1\rangle \otimes |\chi\rangle \otimes |1\rangle \ , \tag{6.1.6}$$

where θ is fixed so that λ_1 and λ_4 are the minimum eigenvalues of ρ_1 and ρ_4, respectively, which leads to

$$\theta = \arcsin\left(\frac{\lambda_4}{1 - \lambda_1 + \lambda_4}\right) \ . \tag{6.1.7}$$

Explicitly, one finds

$$|\psi_{4,2}\rangle = a\,|0000\rangle + b\,|0011\rangle + c\,|0101\rangle + d\,|0110\rangle + e\,|1001\rangle + f\,|1111\rangle \ , \tag{6.1.8}$$

where

$$\begin{aligned}
a =& \sqrt{\lambda_4(\lambda_2 + \lambda_3 + \lambda_4 - \lambda_1)/[2(1 - \lambda_1 + \lambda_4)]} \ , \\
b =& \sqrt{(\lambda_1 + \lambda_2 - \lambda_3 - \lambda_4)/2} \ , \\
c =& \sqrt{(\lambda_1 + \lambda_3 - \lambda_2 - \lambda_4)/2} \ , \\
d =& \sqrt{\lambda_4(2 + \lambda_4 - \lambda_1 - \lambda_2 - \lambda_3)/[2(1 - \lambda_1 + \lambda_4)]} \ , \\
e =& \sqrt{(1 - \lambda_1)(\lambda_2 + \lambda_3 + \lambda_4 - \lambda_1)/[2(1 - \lambda_1 + \lambda_4)]} \ , \\
f =& \sqrt{(1 - \lambda_1)(2 + \lambda_4 - \lambda_1 - \lambda_2 - \lambda_3)/[2(1 - \lambda_1 + \lambda_4)]} \ .
\end{aligned} \tag{6.1.9}$$

Proceeding along the same steps, one can produce recursively a solution $|\psi_{N,2}\rangle$ for any set of compatible eigenvalues and any number of qubits N.

- $N = 3$, $d = 3$: The case of three qutrits has been tackled by Higuchi in Ref. [98]. The spectral polytope $\Lambda_{3,3}$ is the subset of the Weyl polytope $\Omega_{3,3}$ cut out by the

following system of inequalities:

$$\begin{cases} \lambda_a^{(2)} + \lambda_a^{(1)} \leq \lambda_b^{(2)} + \lambda_b^{(1)} + \lambda_c^{(2)} + \lambda_c^{(1)} \\ \lambda_a^{(3)} + \lambda_a^{(1)} \leq \lambda_b^{(2)} + \lambda_b^{(1)} + \lambda_c^{(3)} + \lambda_c^{(1)} \\ \lambda_a^{(2)} + \lambda_a^{(3)} \leq \lambda_b^{(2)} + \lambda_b^{(1)} + \lambda_c^{(2)} + \lambda_c^{(3)} \\ 2\lambda_a^{(2)} + \lambda_a^{(1)} \leq 2\lambda_b^{(2)} + \lambda_b^{(1)} + 2\lambda_c^{(2)} + \lambda_c^{(1)} \\ 2\lambda_a^{(1)} + \lambda_a^{(2)} \leq 2\lambda_b^{(2)} + \lambda_b^{(1)} + 2\lambda_c^{(1)} + \lambda_c^{(2)} \\ 2\lambda_a^{(2)} + \lambda_a^{(3)} \leq 2\lambda_b^{(2)} + \lambda_b^{(1)} + 2\lambda_c^{(2)} + \lambda_c^{(3)} \\ 2\lambda_a^{(2)} + \lambda_a^{(3)} \leq 2\lambda_b^{(1)} + \lambda_b^{(2)} + 2\lambda_c^{(3)} + \lambda_c^{(2)} \end{cases} , \tag{6.1.10}$$

where (a, b, c) is any permutation of $(1, 2, 3)$.

- *N arbitrary, d arbitrary*: In a breakthrough paper [99], Klyachko has developed an algorithm to produce the H-representation of the spectral polytopes $\Lambda_{N,d}$ for arbitrary values of N and d. In particular, the above solutions for non-constructive marginal problems can be recovered from Klyachko's results. There are, however, two caveats. First, the algorithm has a high computational complexity, so that in practice it has been possible to apply it only in relatively low-dimensional settings ($N \leq 10$, $d \leq 5$). Second, while it solves in principle any non-constructive (univariate) marginal problem for indistinguishable particles, it gives no clue about the constructive version.

We now focus on marginal problems for indistinguishable particles. For bosons, the problem trivializes. Let $\{\lambda^{(j)}\}_{j=0}^{d-1}$ denote the local eigenvalues of any 1-body reduced state $\rho^{(\vee)}$ of a system of N bosons. Notice that no subscript label is necessary since, by definition, all marginals are the same for all particles, as a consequence of indistinguishability. Then, the state

$$|\psi_{N,d}^{(\vee)}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \sqrt{\lambda^{(j)}}\, |j\rangle^{\otimes N} \tag{6.1.11}$$

is totally symmetric and moreover has the desired local spectra.

In contrast, the problem for fermions is non-trivial. It is also highly significant from a physical standpoint. The reason is the following. Adhering to computational chemistry conventions, the 1-body reduced state $\rho^{(\wedge)}$ for a system of N fermions over $d > N$ modes will now be normalized so that $\mathrm{tr}[\rho^{(\wedge)}] = N$. Thus its eigenvalues $\{\lambda^{(j)}\}_{j=0}^{d-1}$ can be interpreted as occupation numbers. As a consequence of antisymmetry, the Pauli exclusion principle holds, that is the eigenvalues are constrained so that $0 \leq \lambda^{(j)} \leq 1$, $\forall j \in \{0, \ldots, d-1\}$. However, antisymmetry is a condition far stronger than implied by the Pauli principle. For fermions, each λ-point of compatible local eigenvalues can be

embedded in $\mathbb{R}^d$, giving rise to the spectral polytope $\Lambda_{N,d}^{(\wedge)}$. It is a convex subset of the Pauli polytope Π_d, which is defined via the system of constraints

$$\begin{cases} 1 \geq \lambda^{(0)} \geq \lambda^{(1)} \geq \cdots \geq \lambda^{(d-1)} \geq 0 \\ \sum_{j=0}^{d-1} \lambda^{(j)} = N \end{cases} . \tag{6.1.12}$$

The Pauli polytope Π_d plays the same role as the Weyl polytope $\Omega_{N,d}$ for indistinguishable particles. If the Pauli principle were the only constraint, then λ-points would fill the entire Pauli polytope Π_d. This, however, is not the case. The study of marginal problems for fermions thus leads to *generalized Pauli constraints* [100]. Here, we summarize the literature of known solutions.

- $N = 2$*, d arbitrary*: A necessary and sufficient condition on the local spectrum is that it must be doubly degenerate (plus possibly a single unpaired vanishing eigenvalue if d is odd). The corresponding global state can be constructed as follows. Let the 1-body marginal $\rho^{(\wedge)}$ have the following local spectrum:

$$\mathrm{spec}(\rho^{(\wedge)}) = \begin{cases} \{\lambda^{(0)}, \ldots, \lambda^{(d/2-1)}\} & \text{if } d \text{ even} \\ \{\lambda^{(0)}, \ldots, \lambda^{(\lfloor d/2 \rfloor - 1)}, 0\} & \text{if } d \text{ odd} \end{cases} . \tag{6.1.13}$$

  Each non-vanishing eigenvalues has multiplicity 2. Then, the state

$$|\psi_{N,d}^{(\wedge)}\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{\lfloor d/2 \rfloor - 1} \sqrt{\lambda^{(j)}}\, |2j\rangle \wedge |2j+1\rangle \ , \tag{6.1.14}$$

  where by definition $|j\rangle \wedge |j'\rangle := |j\rangle \otimes |j'\rangle - |j'\rangle \otimes |j\rangle$ is the desired solution. The case of $d-2$ fermions over d modes can be treated similarly, which is a consequence of *particle-hole duality* [101]: if the spectral polytope $\Lambda_{N,d}^{(\wedge)}$ is known, the spectral polytope $\Lambda_{d-N,d}^{(\wedge)}$ can be obtained via the mapping $\lambda^{(j)} \to \lambda^{(d-j+1)}$.

- $N = 3$*,* $d = 6$: Because of particle-hole duality, this is the lowest dimensional setting which is non-trivial. The spectral polytope $\Lambda_{3,6}^{(\wedge)}$ was first determined numerically by Dennis & Borland in 1972 [102], while a rigorous proof was published only in 2007 by Ruskai [103]. It is the subset of Π_6 cut out by the constraints

$$\begin{cases} \lambda^{(0)} + \lambda^{(5)} = \lambda^{(1)} + \lambda^{(4)} = \lambda^{(2)} + \lambda^{(3)} \\ \lambda^{(4)} + \lambda^{(5)} \geq \lambda^{(3)} \end{cases} . \tag{6.1.15}$$

- N *arbitrary,* d *arbitrary*: In complete parallel with the distinguishable particles case, Klyachko has proposed an algorithm to determine spectral polytopes for arbitrary (univariate) fermionic marginal problems [104]. Computational complexity arguments limit its usefulness to low-dimensional settings. Moreover, Klyachko's solution offers no hint regarding constructive problems.

6.2 Connecting the quantum marginal problem with orthogonal arrays

In this section, we illustrate how orthogonal array theory can be used to obtain solutions for constructive marginal problems. The basic idea is that a balanced superposition of the states associated to the runs of an $\mathrm{IrOA}(r, N, d, 1)$ always gives rise to a 1-uniform state. 1-uniform states are the solution to a particular type of marginal problems, i.e. when all marginals are required to equal the maximally mixed state $\mathbb{1}_d/d$. The next natural step is to consider arbitrary superpositions of the runs of a 1-irredundant orthogonal array. If the array is chosen appropriately, one may hope to obtain a general solution to the problem for arbitrary *compatible* marginals, upon a judicious choice of the relative weights.

To better understand this point, let us consider an arbitrary pure state of N qudits $|\psi_{N,d}\rangle \in \mathcal{H} := \mathcal{H}_d^{\otimes N}$. Let $\{|I\rangle\}_{I\in\{\mathscr{A}_d\}^{\times N}}$ be a basis for $\mathcal{H}$ (where $\mathscr{A}_d := \{0, 1, \ldots, d-1\}$ and the following multi-index notation is employed: given $I = i_1\, i_2 \ldots\, i_N$, $|I\rangle := |i_1\rangle \otimes \cdots \otimes |i_N\rangle$). Then, we may expand $|\psi_{N,d}\rangle$ as

$$|\psi_{N,d}\rangle = \sum_{I\in\mathscr{A}_d^{\times N}} c_I\, |I\rangle \ , \tag{6.2.1}$$

with the normalization constraint $\sum_I |c_I|^2 = 1$. The 1-body marginal corresponding to the k^{th} subsystem reads

$$\rho_k = \mathrm{tr}_{\setminus k}(|\psi_{N,d}\rangle\langle\psi_{N,d}|) = \sum_{j=0}^{d-1}\sum_{j'=0}^{d-1} \sum_{I_1\in\mathscr{A}_d^{\times(k-1)}} \sum_{I_2\in\mathscr{A}_d^{\times(N-k)}} c_{I_1\cdot j\cdot I_2}\, \bar{c}_{I_1\cdot j'\cdot I_2}\ |j\rangle\langle j'| \ . \tag{6.2.2}$$

Notice that composition of two multi-indices $I = i_1 \ldots i_m$ and $I' = i_{m+1} \ldots i_{m+n}$ is naturally defined as follows: $I \cdot I' = i_1 \ldots i_m i_{m+1} \ldots i_{m+n}$. As argued before, each marginal ρ_k can be assumed to be given in diagonal form, with eigenvalues $\{\lambda_k^{(j)}\}_{j=0}^{d-1}$ ordered non-increasingly. For every $k \in \{1, \ldots, N\}$, we thus have to impose that

$$\sum_{I_1\in\mathscr{A}_d^{\times(k-1)}} \sum_{I_2\in\mathscr{A}_d^{\times(N-k)}} |c_{I_1\cdot j\cdot I_2}|^2 = \lambda_k^{(j)} \ , \qquad \text{for } j \in \mathscr{A}_d \ ; \tag{6.2.3}$$

and

$$\sum_{I_1\in\mathscr{A}_d^{\times(k-1)}} \sum_{I_2\in\mathscr{A}_d^{\times(N-k)}} c_{I_1\cdot j\cdot I_2}\, \bar{c}_{I_1\cdot j'\cdot I_2} = 0 \ , \qquad \text{for } j \neq j' \in \mathscr{A}_d \ . \tag{6.2.4}$$

Geometrically, a non-constructive marginal problem is equivalent to identifying a high-dimensional algebraic variety embedded in $\mathbb{C}^{d^N}$, which is defined via the system of equations (6.2.3) and (6.2.4).

Orthogonal arrays prove useful in the following way. Suppose that, instead of starting from the most general state $|\psi_{N,d}\rangle$ as in Eq. (6.2.1), one considers arbitrary superpositions of the runs of an IrOA(r, N, d, k), denoted by α, i.e.

$$|\varphi_{N,d}\rangle = \sum_{j=0}^{r} c_j \ |s_{j1} \dots s_{jN}\rangle \ , \tag{6.2.5}$$

where the underlying array is

$$\alpha = \begin{matrix} s_{11} & s_{12} & \dots & s_{1N} \\ s_{21} & s_{22} & \dots & s_{2N} \\ \vdots & \vdots & & \vdots \\ s_{r1} & s_{r2} & \dots & s_{rN} \end{matrix} . \tag{6.2.6}$$

Because of irredundancy, all equations of the form (6.2.4) are automatically satisfied by construction. One is left to solve a linear system in the variables $|c_I|^2$, which can be done efficiently. Of course, there is no guarantee that a single array α will be sufficient to obtain a solution for every set of compatible eigenvalues. Indeed, the linear system may have no solution. Clearly, however, one should start from an array which has as many runs as possible, compatibly with the requirement of irredundancy.

Definition 6.2.1. *(**maximal IrOA**) An irredundant orthogonal array IrOA(r, N, d, k) is called maximal if it maximizes the number of rows r, among all irredundant orthogonal arrays with the same number of columns N, number of symbols d and strength k.*

We now describe a recursive construction that allows to build, for given d, maximal 1-irredundant arrays with any number of columns N. For $N = 2$, consider the IrOA$(d, 2, d, 1)$:

$$\xi_{2,d} = \begin{matrix} 0 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ d-1 & d-1 \end{matrix} . \tag{6.2.7}$$

Let $\xi'_{2,d}$ be the extended array

$$\xi'_{2,d} = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ d-1 & d-1 & 0 \end{matrix} . \tag{6.2.8}$$

Let $f_s(\xi'_{2,d})$ denote the array obtained by adding $s \in \mathscr{A}_d$ mod(d), to each entry in the first and last columns of $\xi'_{2,d}$:

$$f_s(\xi'_{2,d}) = \begin{array}{ccc} 0+s \bmod(d) & 0 & s \\ 1+s \bmod(d) & 1 & s \\ \vdots & \vdots & \vdots \\ d-1+s \bmod(d) & d-1 & s \end{array}. \tag{6.2.9}$$

Then, the array $\xi_{3,d} := \oplus_{s\in\mathscr{A}_d} f_s(\xi'_{2,d})$ is a maximal 1-irredundant IrOA$(d^2, 3, d, 1)$. From $\xi_{3,d}$, repeat the same procedure as many times as needed in order to have an array with N columns, denoted by $\xi_{N,d}$, which is a maximal 1-irredundant IrOA$(d^{N-1}, N, d, 1)$.

The proof is elementary. First, let us check irredundancy. If one imagines eliminating the last column, then the resulting array can have no repetition of rows: any two rows belonging to the same block $f_s(\xi'_{N-1,d})$ are different (because $\xi_{N-1,d}$ is irredundant), while rows belonging to different blocks always differ at least in the first entry. If one imagines eliminating any other column, then there can be no duplication of rows, either because they differ in the last entry or, if not, because they must differ in one of the first $N-2$ entries. Finally, $\xi_{N,d}$ is clearly maximal, since it has d^{N-1} rows, i.e. eliminating any one column gives the full factorial design with $N-1$ columns, so that no irredundant array with more rows can exist.

Let us remark that the maximal irredundant array $\xi_{N,d}$ just constructed is non-unique. In fact, any array isomorphic to $\xi_{N,d}$ is also maximal. For a binary alphabet, however, one can prove that there are no other maximal IrOAs ($\xi_{N,2}$ is unique up to isomorphisms). This is a consequence of the fact that any maximal IrOA with N columns over a binary alphabet is associated to a proper 2-coloring of a hypercube graph Q_N. Since Q_N is uniquely colorable, the statement follows. However, for $d > 2$, other maximal IrOAs, inequivalent to $\xi_{N,d}$, may exist.

Example 6.2.2. *Let us consider the case $d = 2$ and go through the steps of the previous construction. The starting point is*

$$\xi_{2,2} = \begin{array}{cc} 0 & 0 \\ 1 & 1 \end{array}. \tag{6.2.10}$$

By definition,

$$f_0(\xi'_{2,2}) = \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 0 \end{array}, \qquad f_1(\xi'_{2,2}) = \begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}. \tag{6.2.11}$$

Thus, for $N = 3$ columns, one finds the following maximal 1-irredundant orthogonal

array:

$$\xi_{3,2} = f_0(\xi'_{2,2}) \oplus f_1(\xi'_{2,2}) = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{matrix} . \tag{6.2.12}$$

The runs of the IrOA $\xi_{3,2}$ *span the subspace*

$$|\varphi_{3,2}\rangle = c_1 |000\rangle + c_2 |110\rangle + c_3 |101\rangle + c_4 |011\rangle \ , \tag{6.2.13}$$

with $\sum_{j=1}^{4} |c_j|^2 = 1$. *Because of irredundancy, all single-body reductions are in diagonal form:*

$$\begin{aligned} \rho_1 &= \mathrm{diag}(|c_1|^2 + |c_4|^2, |c_2|^2 + |c_3|^2) \ , \\ \rho_2 &= \mathrm{diag}(|c_1|^2 + |c_3|^2, |c_2|^2 + |c_4|^2) \ , \\ \rho_3 &= \mathrm{diag}(|c_1|^2 + |c_2|^2, |c_3|^2 + |c_4|^2) \ , \end{aligned} \tag{6.2.14}$$

Letting $\lambda_k^{(0)}$ *and* $\lambda_k^{(1)}$ *denote the eigenvalues of each reduction* ρ_k, *the task is to solve the linear system* $M_{3,2} \cdot \boldsymbol{c} = \boldsymbol{\lambda}$, *where* $\boldsymbol{c} = (|c_1|^2, |c_2|^2, |c_3|^2, |c_4|^2)^t$ *belongs to the standard simplex* Δ_3, $\boldsymbol{\lambda} = (\lambda_1^{(0)}, \lambda_1^{(1)}, \lambda_2^{(0)}, \lambda_2^{(1)}, \lambda_3^{(0)}, \lambda_3^{(1)})^t$ *belongs to the spectral polytope* $\Lambda_{3,2}$ *and* $M_{3,2}$ *is the matrix*

$$M_{3,2} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}^t . \tag{6.2.15}$$

A solution can be obtained for any possible choice of $\boldsymbol{\lambda} \in \Lambda_{3,2}$, *if and only if the intersection of the Weyl polytope* $\Omega_{3,2}$ *with the image of* $M_{3,2}$ *acting on the standard simplex* Δ_3 *coincides with the spectral polytope* $\Lambda_{3,2}$. *A quick computation reveals that indeed* $\Xi_{3,2} := \mathrm{im}_{\Delta_3}(M_{3,2}) \cap \Omega_{3,2} = \Lambda_{3,2}$ *(one may for instance compute their V-representations and check that they are the same). By solving explicitly the linear system above, one finds*

$$\begin{aligned} c_1 &= \sqrt{(2 - \lambda_1^{(1)} - \lambda_2^{(1)} - \lambda_3^{(1)})/2} \ , \qquad c_2 = \sqrt{(\lambda_1^{(1)} + \lambda_2^{(1)} - \lambda_3^{(1)})/2} \ , \\ c_3 &= \sqrt{(\lambda_1^{(1)} + \lambda_3^{(1)} - \lambda_2^{(1)})/2} \ , \qquad c_4 = \sqrt{(\lambda_2^{(1)} + \lambda_3^{(1)} - \lambda_1^{(1)})/2} \ . \end{aligned} \tag{6.2.16}$$

Substituted in Eq. (6.2.13), *Eq.* (6.2.16) *provide a constructive solution to the marginal problem for 3 qubits, which is unitarily equivalent to the one by Higuchi* et al..

In the remaining part of this section, we discuss the case of fermionic systems. Notice that 1-irredundant orthogonal arrays give rise to 1-uniform states of *distinguishable* particles. To obtain fermionic 1-uniform states, the starting array must have the following antisymmety properties (besides orthogonality and irredundancy): no two rows may

contain repeated symbols; moreover, if a particular row appears in the array, then all its permutations must also appear. A maximal fermionic irredundant orthogonal array $\xi^{(\wedge)}_{N,d}$ may thus be constructed as follows: its unique rows (up to permutations) are a maximum set of subsets of N elements, all different, taken from the alphabet $\mathscr{A}_d$, such that every two subsets have at most $N-2$ symbols in common. For arbitrary N and d, $\xi^{(\wedge)}_{N,d}$ can be generated algorithmically in an efficient way. The following simple example helps clarifying our methods.

Example 6.2.3. *Let us consider the setting $N=2$, $d=4$, i.e. 2 fermions over 4 modes. A maximal set of subsets of 2 elements (taken from the alphabet $\mathscr{A}_4=\{0,1,2,3\}$ and all different), with no elements in common, is e.g. $\{\{0,1\},\{2,3\}\}$. The corresponding maximal irredundant orthogonal array is*

$$\xi_{2,4}=\begin{matrix}0 & 1\\ 1 & 0\\ 2 & 3\\ 3 & 2\end{matrix}, \tag{6.2.17}$$

which gives rise to the totally antisymmetric states $|\varphi^{(\wedge)}_{2,4}\rangle = c_1\,|0\rangle\wedge|1\rangle + c_2\,|2\rangle\wedge|3\rangle$. All 1-body reductions are the same and equal to $\rho^{(\wedge)}=\mathrm{diag}(|c_1|^2,|c_1|^2,|c_2|^2,|c_2|^2)$. As recalled before, the most general single-particle spectrum for bipartite fermionic systems is doubly-degenerate. Suppose, without loss of generality, that the local eigenvalues are $\boldsymbol{\lambda}=(\lambda^{(0)}\lambda^{(0)},\lambda^{(1)},\lambda^{(1)})$. Then the desired global state is

$$|\varphi^{(\wedge)}_{2,4}\rangle=\sqrt{\lambda^{(0)}}\,|0\rangle\wedge|1\rangle+\sqrt{\lambda^{(1)}}\,|2\rangle\wedge|3\rangle\ . \tag{6.2.18}$$

6.3 Construction of solutions via orthogonal arrays

In this section, we formalize the procedure followed in the previous examples to obtain constructive solutions of (univariate) quantum marginal problems, for arbitrary values of N and d. Given the vector of local eigenvalues $\boldsymbol{\lambda}$, the construction succeeds only if $\boldsymbol{\lambda}$ belongs to a convex subset $\Xi_{N,d}$ of the spectral polytope $\Lambda_{N,d}$. In the case of qubits, however, we will prove that $\Xi_{N,2}$ and $\Lambda_{N,2}$ coincide, thus obtaining an alternative solution to the one by Higuchi *et al.* [97]. For distinguishable particles, the construction is based on the following steps:

(S1) Using the recursive construction given above, obtain the maximal 1-irredundant orthogonal array $\xi_{N,d}$.

(S2) Compute the V-representation of the convex polytope

$$\Xi_{N,d} := \mathrm{im}_{\Delta_\delta}(M_{N,d}) \cap \Omega_{N,d} \,, \tag{6.3.1}$$

where Δ_δ is the standard simplex of dimension $\delta := d^{N-1} - 1$, $M_{N,d}$ is a $Nd \times d^{N-1}$ matrix and $\Omega_{N,d}$ is the Weyl polytope. The matrix $M_{N,d}$ can be obtained from $\xi_{N,d}$ as follows: 1) replace each symbol $s \in \mathscr{A}_d$ of the array with d symbols, all of them 0, except a single 1 in position s; 2) interpret the resulting array as a matrix and 3) transpose it.[1] By a well-known lemma, the vertices of the linear image of a convex polytope are a subset of the images of the vertices of the starting polytope. Thus, one may also write:

$$\mathrm{im}_{\Delta_\delta}(M_{N,d}) = \mathrm{conv}(\mathrm{col}(M_{N,d})) \,, \tag{6.3.3}$$

the convex-hull of the columns of $M_{N,d}$.

(S3) Check whether the vector of local eigenvalues $\boldsymbol{\lambda}$ belongs to $\Xi_{N,d}$. If not, the construction cannot proceed.

(S4) Find at least one solution of the linear system $M_{3,2} \cdot \boldsymbol{c} = \boldsymbol{\lambda}$. Denote such solution as $\tilde{\boldsymbol{c}}$. The global state having the assigned marginals is

$$|\varphi_{N,d}\rangle = \sum_{j=1}^{d^{N-1}} \sqrt{\tilde{c}_j}\, |I_j\rangle \,, \tag{6.3.4}$$

where $\{I_j\}_{j=1}^{d^{N-1}}$ are the rows of $\xi_{N,d}$.

On the basis of Example 6.2.3, the interested reader may adapt the above procedure to the case of fermionic systems.

6.3.1 The case of qubits

As anticipated, the case of N qubits is sufficiently special that the construction given in the previous section can be proven to be exhaustive.

Proposition 6.3.1. *For any number of qubits N, the set of λ-points $\Xi_{N,2}$ for which a solution can be constructed from the maximal IrOA $\xi_{N,d}$ coincides with the spectral polytope $\Lambda_{N,2}$.*

[1] For instance, for $d = 2$, one has to make the substitutions $0 \to 1\ 0$ and $1 \to 0\ 1$ into the array $\xi_{2,2}$, which produces the matrix

$$\xi_{2,2} = \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} \quad \to \quad M_{2,2}^t = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \,. \tag{6.3.2}$$

Proof. The set $\Xi_{N,2} := \mathrm{im}_{\Delta_\delta}(M_{N,2}) \cap \Omega_{N,2}$, with $\delta = 2^{N-1} - 1$, is a convex polytope (it is, in fact, the intersection of two convex polytopes). To establish the above proposition, we will thus check that $\Xi_{N,2}$ has the same vertices as the spectral polytope $\Lambda_{N,2}$.

First, we introduce the following convenient abbreviations:

$$A := \frac{1}{2}, \frac{1}{2} , \qquad B := 1, 0 , \qquad C := 0, 1 , \tag{6.3.5}$$

which allow to rewrite certain special vectors in compact form. For instance, the vector $(1/2, 1/2, 1, 0, 1/2, 1/2)^t$ is represented by the string $ABBA$. The set of all strings of N symbols e.g., A and B, is denoted by $\mathcal{S}(A, B)$, and the number of occurrences in a string of e.g., A, is denoted by $\#(A)$. From the polygonal inequalities of Eq. (6.1.3), it follows that the vertices of $\Lambda_{N,2}$ are

$$\mathrm{vert}(\Lambda_{N,2}) = \{\text{all } N\text{-strings} \in \mathcal{S}(A, B) \text{ such that } \#(A) \neq 1\} . \tag{6.3.6}$$

Instead, the vertices of the Weyl polytope $\Omega_{N,2}$ are given by

$$\mathrm{vert}(\Omega_{N,2}) = \{\text{all } N\text{-strings} \in \mathcal{S}(A, B)\} . \tag{6.3.7}$$

Note in particular that the vertices of the spectral polytope $\Lambda_{N,2}$ are a proper subset of the vertices of the Weyl polytope $\Omega_{N,2}$ (this does not hold in general for $d > 2$).

We now focus on $\mathrm{im}_{\Delta_\delta}(M_{N,2})$. Recall that $\mathrm{im}_{\Delta_\delta}(M_{N,2})$ is the convex hull of the columns of $M_{N,2}$ or, equivalently,

$$\mathrm{im}_{\Delta_\delta}(M_{N,2}) = \mathrm{conv}\left(\{\text{all } N\text{-strings} \in \mathcal{S}(B, C) \text{ such that } \#(C) \text{ is even}\}\right) , \tag{6.3.8}$$

In fact, the array $\xi_{N,2}$ has 2^{N-1} rows, all different by construction. Moreover, each row of $\xi_{N,2}$ contains an even number of 1s (each row of $\xi_{2,2}$ clearly has the previous property and the operations involved in the recursive construction of Sect. 6.2 preserve such property, so that it holds for any N). Therefore, the rows of $\xi_{N,2}$ are precisely all possible binary strings containing an even number of 1s. Since $M_{N,2}$ is built from $\xi_{N,2}$ via the substitutions $0 \to B$, $1 \to C$ (and matrix transposition), Eq. (6.3.8) follows.

We will prove that all vertices of $\Lambda_{N,2}$ are also vertices of $\Xi_{N,2}$. This is sufficient because, by construction, $\Xi_{N,2} \subseteq \Lambda_{N,2}$, so that if $\Xi_{N,2}$ had any other vertex, it would strictly contain $\Lambda_{N,2}$, which is absurd. By a simple lemma, every vertex of $\Omega_{N,2}$ that belongs to $\mathrm{im}_{\Delta_\delta}(M_{N,2})$ is also a vertex of $\Xi_{N,2}$ (if a point is an extremal point of a convex set, it is also extremal with respect to any convex subset). Therefore, the task is to prove that all vertices of $\Omega_{N,2}$ of the form (6.3.6) belong to $\mathrm{im}_{\Delta_\delta}(M_{N,2})$. Or, equivalently, that any string $s \in \mathcal{S}(A, B)$ with $n := \#(A) \neq 1$ can be obtained as a convex combination of strings in $\mathcal{S}(B, C)$ with $\#(C)$ even. If n is even, the statement is trivial, since s is equal

to $1/2$ times a string of only Bs, plus $1/2$ times a string of n Cs, at the same positions as the As in s, while the rest are all Bs. For instance, $AABAA = (CCBCC + BBBBB)/2$. Therefore, let us suppose n is odd. Then s can be written as an equal weights combination of all strings having an even number of Cs, at the same positions as the As in s. For instance,

$$AAAB = \frac{1}{4}\left(BBBB + CCBB + CBCB + BCCB\right) . \tag{6.3.9}$$

This works because $B + C = 2A$ and, at any given position where an A occurs in s, one has a sum of the form m_B times B plus m_C times C, where:

$$\begin{aligned} m_C &= \binom{n-1}{1} + \binom{n-1}{3} + \cdots + \binom{n-1}{n-2} = 2^{n-2} , \\ m_B &= \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{n-1} - m_C = 2^{n-1} - 2^{n-2} = 2^{n-2} , \end{aligned} \tag{6.3.10}$$

so $m_B = m_C$, as desired. □

For instance, applied to the case of 4 qubits, our methods produce the following solution

$$|\varphi_{4,2}\rangle = c_1 \,|0000\rangle + c_2 \,|1100\rangle + c_3 \,|1010\rangle + c_4 \,|0110\rangle + c_5 \,|1001\rangle \ , \tag{6.3.11}$$

where[2]

$$\begin{aligned} c_1 &= \sqrt{(2 - \lambda_1^{(1)} - \lambda_2^{(1)} - \lambda_3^{(1)} - \lambda_4^{(1)})/2} , \\ c_2 &= \sqrt{(\lambda_1^{(1)} + \lambda_2^{(1)} - \lambda_3^{(1)} - \lambda_4^{(1)})/2} , \\ c_3 &= \sqrt{(\lambda_1^{(1)} - \lambda_2^{(1)} + \lambda_3^{(1)} - \lambda_4^{(1)})/2} , \\ c_4 &= \sqrt{(-\lambda_1^{(1)} + \lambda_2^{(1)} + \lambda_3^{(1)} + \lambda_4^{(1)})/2} , \\ c_5 &= \sqrt{\lambda_4^{(1)}} , \end{aligned} \tag{6.3.12}$$

which should be compared with the solution by Higuchi *et al.* of Eqs. (6.1.8) and (6.1.9). The two solutions can be proven to be non-trivially related (in contrast to the case of three qubits), e.g. by computing their respective polynomial invariants [105].

6.3.2 Beyond qubits

For local dimension $d > 2$, it is not guaranteed that the construction of Sect. 6.3 is exhaustive: the polytope $\Xi_{N,d}$ is in general only a strict subset of the spectral polytope $\Lambda_{N,d}$. We discuss in some detail the setting of three qutrits.

[2]Here it is assumed that, without loss of generality, the Hilbert space labels are assigned so that $\lambda_1^{(1)} \geq \lambda_2^{(1)} \geq \lambda_3^{(1)} \geq \lambda_4^{(1)}$.

Example 6.3.2. *We start from the maximal* $\mathrm{IrOA}(9,3,3,1)$*:*

$$\xi_{3,3} = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \\ 2 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{matrix} . \tag{6.3.13}$$

By mapping $0 \to 1\,0\,0$*,* $1 \to 0\,1\,0$*,* $2 \to 0\,0\,1$*, the matrix* $M_{3,3}$ *is found to be*

$$M_{3,3} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} . \tag{6.3.14}$$

The polytope $\Xi_{3,3}$ *is the convex hull of the columns of* $M_{3,3}$*, intersected with the Weyl polytope* $\Omega_{3,3}$*. We report below its V-representation, computed with the help of the Parma Polyhedra Library (PPL) [106]. First, let us introduce the following abbreviations:*

$$\begin{gathered} A := 1,\, 0,\, 0, \qquad B := \frac{1}{2},\, \frac{1}{2},\, 0, \qquad C := \frac{1}{3},\, \frac{1}{3},\, \frac{1}{3}, \qquad D := \frac{2}{3},\, \frac{1}{6},\, \frac{1}{6}, \\ E := \frac{1}{2},\, \frac{1}{4},\, \frac{1}{4}, \qquad F := \frac{3}{4},\, \frac{1}{4},\, 0, \qquad G := \frac{2}{3},\, \frac{1}{3},\, 0\ . \end{gathered} \tag{6.3.15}$$

From Ref. [98], the spectral polytope $\Lambda_{3,3}$ *has 33 vertices:* AAA*,* BBB*,* CCC*,* ABB*,* ACC*,* BBC*,* BCC*,* BDD*,* CGG*,* BCD*,* BEF*, together with all their possible permutations. In contrast,* $\Xi_{3,3}$ *has only 26 vertices, which are a subset of the vertices of* $\Lambda_{3,3}$*: the 7 missing vertices are* BBB*,* CGG*,* ABB*,* BEF*,* BFE*,* FEB*,* FBE*.*

We remark that, starting from a suitable array isomorphic to $\xi_{3,3}$*, it is actually possible to cover such missing vertices. However, there is no single array whose associated polytope coincides with the spectral polytope.*

6.3.3 A few fermionic cases

In this section, we discuss a few low-dimensional fermionic scenarios, beginning with the Dennis-Borland setting ($N = 3$, $d = 6$).

Example 6.3.3. *We start from the array*

$$\xi_{3,6}^{(\wedge)} = \alpha(012) \oplus \alpha(034) \oplus \alpha(135) \oplus \alpha(245) \ , \tag{6.3.16}$$

where, e.g.,

$$\alpha(012) = \begin{matrix} 0 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \\ 2 & 1 & 0 \end{matrix} . \tag{6.3.17}$$

To it, one associates the family of states

$$|\varphi_{3,6}^{(\wedge)}\rangle = c_1 \, |\alpha(012)\rangle + c_2 \, |\alpha(034)\rangle + c_3 \, |\alpha(135)\rangle + c_4 \, |\alpha(245)\rangle \ , \tag{6.3.18}$$

with the normalization condition $|c_1|^2 + |c_2|^2 + |c_3|^2 + |c_4|^2 = 1/2$. *Notice that, e.g.,*

$$|\alpha(012)\rangle := |0\rangle \wedge |1\rangle \wedge |2\rangle = |012\rangle - |021\rangle - |102\rangle + |120\rangle - |210\rangle + |201\rangle \ . \tag{6.3.19}$$

The matrix $M_{3,6}^{(\wedge)}$ *is*

$$M_{3,6}^{(\wedge)} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{6.3.20}$$

The set of λ*-points for which a solution can be constructed is* $\Xi_{3,6}^{(\wedge)} = \mathrm{conv}(\mathrm{col}(M_{3,6}^{(\wedge)})) \cap \Pi_{3,6}$. *Its vertices are*

$$(1,1,1,0,0,0)^t \ , \quad \left(\frac{3}{4},\frac{3}{4},\frac{3}{4},\frac{3}{4},\frac{3}{4},\frac{3}{4}\right)^t \ , \quad \left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)^t \ , \quad \left(1,\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},0\right)^t \ ,$$

which are also the vertices of the spectral polytope $\Lambda_{3,6}^{(\wedge)}$.

As a last example, we study the case of three 7-dimensional fermions ($N = 3$, $d = 7$).

Example 6.3.4. *The starting point is the array*

$$\xi_{3,7} = \alpha(012) \oplus \alpha(034) \oplus \alpha(056) \oplus \alpha(135) \oplus \alpha(146) \oplus \alpha(236) \oplus \alpha(245) \ . \tag{6.3.21}$$

The corresponding polytope, $\Xi^{(\wedge)}_{3,7}$*, has the following vertices:*

$$\begin{gathered}
(1,\,1,\,1,\,0,\,0,\,0,\,0)^t\ , \qquad \left(\tfrac{3}{5},\,\tfrac{3}{5},\,\tfrac{3}{5},\,\tfrac{3}{5},\,\tfrac{1}{5},\,\tfrac{1}{5},\,\tfrac{1}{5}\right)^t\ , \qquad \left(1,\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3}\right)^t\ , \\
\left(1,\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,0,\,0\right)^t\ , \qquad \left(\tfrac{2}{3},\,\tfrac{2}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3}\right)^t\ , \qquad \left(\tfrac{5}{7},\,\tfrac{5}{7},\,\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{1}{7},\,\tfrac{1}{7}\right)^t\ , \\
\left(\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,0\right)^t\ , \qquad \left(\tfrac{2}{3},\,\tfrac{2}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3},\,\tfrac{1}{3}\right)^t\ , \qquad \left(\tfrac{3}{4},\,\tfrac{3}{4},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{4},\,\tfrac{1}{4},\,0\right)^t\ , \\
\left(\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{3}{7},\,\tfrac{3}{7}\right)^t\ , \qquad \left(\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{2},\,\tfrac{1}{4},\,\tfrac{1}{4}\right)^t\ ,
\end{gathered} \tag{6.3.22}$$

which are the same as the vertices of the moment polytope $\Lambda^{(\wedge)}_{3,7}$*.*

In both examples, the construction of Sect. 6.3 is exhaustive. However, we have not been able to prove that this is the case in general, or at least for some infinite family of settings.

6.4 Conclusions

In this chapter, we have established a connection between orthogonal array theory and the quantum marginal problem. In particular, in Sect. 6.3, we have developed a systematic construction that allows to obtain efficient solutions to constructive QMPs, for any number of parties N and number of levels d. Our methods are completely general, but have the drawback that reconstruction is possible only if the local eigenvalues belong to a convex subset $\Xi_{N,d}$ of the spectral polytope $\Lambda_{N,d}$. For some simple enough low-dimensional settings, such as the case of qubits and of three six- and seven-dimensional fermions, the construction is, in fact, exhaustive and the two polytopes coincide. In higher-dimensional settings, the polytopes $\Xi^{(\wedge)}_{N,d}$ provide only an inner approximation to the spectral polytopes $\Lambda^{(\wedge)}_{N,d}$.

Appendix A

List of Notation

$\mathbb{N}$	Set of positive integers
$\mathbb{N}_0$	Set of nonnegative integers
$\mathbb{R}$	Set of real numbers
$\overline{\mathbb{R}}$	Extended set of real numbers
$\mathbb{R}_+$	Set of nonnegative real numbers
$\mathbb{C}$	Set of complex numbers
$\lvert S\rvert$	Cardinality of a set
$\mathscr{P}(S)$	Power set of S
$\mathrm{conv}(S)$	Convex hull of a set of points S
$\mathrm{vert}(\Pi)$	Set of vertices of a convex polytope Π
M_{ij}	Element ij of M
M^t	Transpose of a matrix M
$\mathrm{spec}(M)$	Spectrum of a matrix M
$\mathrm{rk}(M)$	Rank of a matrix M
$\sigma(M)$	Spectral gap of a matrix M
$\mathrm{col}(M)$	Set made up of the columns of a matrix M
$\mathrm{diag}(\{\lambda_i\}_{i=1}^n)$	Diagonal matrix, with diagonal elements $\{\lambda_i\}_{i=1}^n$
$\mathrm{im}_S(M)$	Image of a matrix M on a set S
$\mathcal{M}_{n,m}(\mathbb{K})$	Set of $n \times m$ matrices over a field $\mathbb{K}$
$\mathsf{Her}_n(\mathbb{K})$	Set of $n \times n$ Hermitian matrices over a field $\mathbb{K}$
$\mathsf{Her}_n^+(\mathbb{K})$	Set of $n \times n$ positive semi-definite Hermitian matrices over a field $\mathbb{K}$
$\mathbb{I}_n$	$n \times n$ identity matrix
$\mathbb{0}_n$	$n \times n$ zero matrix
$\mathbb{J}_{p\times q}$	$p \times q$ matrix made up of all ones
$\mathbb{X}$, $\mathbb{Y}\ldots$	Classical random variables

$\mathrm{E}(\mathbb{X})$	Expectation value of $\mathbb{X}$
$\mathrm{Var}(\mathbb{X})$	Variance of $\mathbb{X}$
$\mathrm{Cov}(\mathbb{X}, \mathbb{Y})$	Covariance of $\mathbb{X}$ and $\mathbb{Y}$

THESIS BIBLIOGRAPHY

Bibliography

[1] C. W. Helstrom, *Quantum detection and estimation theory.* Academic Press, 1976.

[2] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, vol. 1. Springer Science & Business Media, 2011.

[3] L. Rondin, J. Tetienne, T. Hingant, J. Roch, P. Maletinsky, and V. Jacques, "Magnetometry with nitrogen-vacancy defects in diamond," *Rep. Prog. Phys.*, vol. 77, no. 5, p. 056503, 2014.

[4] V. B. Braginsky, V. B. Braginsky, and F. Y. Khalili, *Quantum measurement.* Cambridge University Press, 1995.

[5] A. De Pasquale and T. M. Stace, "Quantum Thermometry," *arXiv:1807.05762*, July 2018.

[6] M. G. Paris, "Quantum estimation for quantum technology," *Int. J. Quantum Inf.*, vol. 7, pp. 125–137, 2009.

[7] C. M. Caves, "Quantum-mechanical noise in an interferometer," *Phys. Rev. D*, vol. 23, no. 8, p. 1693, 1981.

[8] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, "Using entanglement improves the precision of quantum measurements," *Phys. Rev. Lett.*, vol. 87, p. 270404, Dec. 2001.

[9] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum metrology," *Phys. Rev. Lett.*, vol. 96, no. 1, p. 010401, 2006.

[10] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nat. Photonics*, vol. 5, no. 4, pp. 222–229, 2011.

[11] R. A. Fisher, "Theory of statistical estimation," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 22, pp. 700–725, Cambridge University Press, 1925.

[12] S. Amari and H. Nagaoka, *Methods of information geometry*, vol. 191. American Mathematical Society, 2007.

[13] O. E. Barndorff-Nielsen, R. D. Gill, and P. E. Jupp, "On quantum statistical inference," *J. Royal Stat. Soc. Series B Stat. Methodol.*, vol. 65, no. 4, pp. 775–804, 2003.

[14] M. Hayashi, *Asymptotic theory of quantum statistical inference: selected papers*. World Scientific, 2005.

[15] H. Cramér, *Mathematical Methods of Statistics (PMS-9)*, vol. 9. Princeton University Press, 2016.

[16] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," in *Breakthroughs in Statistics*, pp. 235–247, Springer, 1992.

[17] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," *Phys. Rev. Lett.*, vol. 72, no. 22, p. 3439, 1994.

[18] L. L. Campbell, "An extended Čencov characterization of the information metric," *Proc. Am. Math. Soc.*, vol. 98, p. 135, Sept. 1986.

[19] N. Čencov, "Algebraic foundation of mathematical statistics ²," *Ser. Stat.*, vol. 9, pp. 267–276, Jan. 1978.

[20] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer, "Information geometry and sufficient statistics," *Probab. Theory Relat. Fields*, vol. 162, pp. 327–364, June 2015.

[21] G. Casella and R. L. Berger, *Statistical Inference*, vol. 2. Duxbury Pacific Grove, CA, 2002.

[22] A. M. Gleason, "Measures on the closed subspaces of a Hilbert space," *J. Math. Mech.*, pp. 885–893, 1957.

[23] K. Kraus, *States, effects and operations: fundamental notions of quantum theory*. Springer, 1983.

[24] M. Ozawa, "Quantum measuring processes of continuous observables," *J. Math. Phys.*, vol. 25, pp. 79–87, Jan. 1984.

[25] D. Petz, "Monotone metrics on matrix spaces," *Linear Algebra Appl.*, vol. 244, pp. 81–96, 1996.

[26] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2017.

[27] D. Petz and C. Sudár, "Geometries of quantum states," *J. Math. Phys.*, vol. 37, no. 6, pp. 2662–2673, 1996.

[28] L. Jing, J. Xiao-Xing, Z. Wei, and W. Xiao-Guang, "Quantum Fisher information for density matrices with arbitrary ranks," *Commun. Theor. Phys.*, vol. 61, no. 1, p. 45, 2014.

[29] H. Nagaoka, "An asymptotically efficient estimator for a one-dimensional parametric model of quantum statistical operators," in *Proc. Int. Symp. on Inform. Theory*, vol. 198, pp. 577–82, 1988.

[30] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzanski, "Compatibility in multiparameter quantum metrology," *Phys. Rev. A*, vol. 94, no. 5, p. 052108, Nov. 2016.

[31] K. Matsumoto, "A new approach to the Cramér-Rao-type bound of the pure-state model," *J. Phys. A: Math. Gen.*, vol. 35, pp. 3111–3123, Apr. 2002.

[32] R. D. Gill and S. Massar, "State estimation for large ensembles," *Phys. Rev. A*, vol. 61, no. 4, p. 042312, Mar. 2000.

[33] H. Yuen and M. Lax, "Multiple-parameter quantum estimation and measurement of nonselfadjoint observables," *IEEE Trans. Inf. Theory*, vol. 19, pp. 740–750, Nov. 1973.

[34] T. Popoviciu, "Sur les équations algébriques ayant toutes leurs racines réelles," *Mathematica*, vol. 9, pp. 129–145, 1935.

[35] L. Seveso, M. A. Rossi, and M. G. A. Paris, "Quantum metrology beyond the quantum Cramér-Rao theorem," *Phys. Rev. A*, vol. 95, no. 1, p. 012111, Jan. 2017.

[36] M. Akahira and K. Takeuchi, *Non-regular statistical estimation*. Springer Science & Business Media, 2012.

[37] S. Pang and T. A. Brun, "Quantum metrology for a general Hamiltonian parameter," *Phys. Rev. A*, vol. 90, no. 2, p. 022117, 2014.

[38] L. Seveso and M. G. A. Paris, "Estimation of general Hamiltonian parameters via controlled energy measurements," *Phys. Rev. A*, vol. 98, no. 3, p. 032114, 2018.

[39] S. Nakayama, A. Soeda, and M. Murao, "Quantum algorithm for universal implementation of the projective measurement of energy," *Phys. Rev. Lett.*, vol. 114, no. 19, p. 190501, 2015.

[40] Y. Matsuzaki, S. Nakayama, A. Soeda, S. Saito, and M. Murao, "Projective measurement of energy on an ensemble of qubits with unknown frequencies," *Phys. Rev. A*, vol. 95, no. 6, p. 062106, 2017.

[41] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, "Quantum Metropolis sampling," *Nature*, vol. 471, no. 7336, pp. 87–90, 2011.

[42] A. Riera, C. Gogolin, and J. Eisert, "Thermalization in nature and on a quantum computer," *Phys. Rev. Lett.*, vol. 108, no. 8, p. 080402, 2012.

[43] A. Y. Kitaev, "Quantum measurements and the Abelian stabilizer problem," *Electr. Coll. Comp. Compl.*, vol. TR96, p. 3, 1996.

[44] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," in *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pp. 478–504, Springer, 1985.

[45] C. A. Fuchs and A. Peres, "Quantum-state disturbance versus information gain: uncertainty relations for quantum information," *Phys. Rev. A*, vol. 53, no. 4, p. 2038, 1996.

[46] K. Banaszek, "Fidelity balance in quantum operations," *Phys. Rev. Lett.*, vol. 86, no. 7, p. 1366, 2001.

[47] C. A. Fuchs and K. Jacobs, "Information-tradeoff relations for finite-strength quantum measurements," *Phys. Rev. A*, vol. 63, no. 6, p. 062305, 2001.

[48] G. M. D'Ariano, "On the Heisenberg principle, namely on the information-disturbance trade-off in a quantum measurement," *Fortschritte Physik*, vol. 51, no. 4-5, pp. 318–330, 2003.

[49] L. Maccone, "Information-disturbance tradeoff in quantum measurements," *Phys. Rev. A*, vol. 73, no. 4, p. 042307, 2006.

[50] S. Olivares and M. G. A. Paris, "Improving information/disturbance and estimation/distortion trade-offs with non-universal protocols," *J. Phys. A*, vol. 40, no. 28, p. 7945, 2007.

[51] M. Ozawa, "Uncertainty relations for noise and disturbance in generalized quantum measurements," *Ann. Phys.*, vol. 311, no. 2, pp. 350–416, 2004.

[52] M. G. Genoni and M. G. A. Paris, "Optimal quantum repeaters for qubits and qudits," *Phys. Rev. A*, vol. 71, no. 5, p. 052307, 2005.

[53] S. Campbell, M. G. Genoni, and S. Deffner, "Precision thermometry and the quantum speed limit," *Quantum Sci. Technol.*, vol. 3, no. 2, p. 025002, 2018.

[54] S. Jevtic, D. Newman, T. Rudolph, and T. Stace, "Single-qubit thermometry," *Phys. Rev. A*, vol. 91, no. 1, p. 012331, 2015.

[55] S. Campbell, M. Mehboudi, G. De Chiara, and M. Paternostro, "Global and local thermometry schemes in coupled quantum systems," *New J. Phys.*, vol. 19, no. 10, p. 103003, 2017.

[56] L. A. Correa, M. Mehboudi, G. Adesso, and A. Sanpera, "Individual quantum probes for optimal thermometry," *Phys. Rev. Lett.*, vol. 114, no. 22, p. 220405, 2015.

[57] T. Shitara, Y. Kuramochi, and M. Ueda, "Trade-off relation between information and disturbance in quantum measurement," *Phys. Rev. A*, vol. 93, no. 3, p. 032134, 2016.

[58] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics*. Springer, Berlin, 2004.

[59] L. Seveso and M. G. A. Paris, "Trade-off between information and disturbance in qubit thermometry," *Phys. Rev. A*, vol. 97, no. 3, p. 032129, Mar. 2018.

[60] J. Preskill, "Quantum Computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018.

[61] O. Muelken and A. Blumen, "Efficiency of quantum and classical transport on graphs," *Phys. Rev. E*, vol. 73, no. 6, p. 066117, June 2006.

[62] A. Ambainis, "Quantum walks and their algorithmic applications," *arXiv:0403120*, Mar. 2004.

[63] N. Lambert, Y.-N. Chen, Y.-C. Cheng, C.-M. Li, G.-Y. Chen, and F. Nori, "Quantum biology," *Nature Physics*, vol. 9, pp. 10–18, Jan. 2013.

[64] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum algorithm for the Hamiltonian NAND tree," *arXiv:0702144*, Feb. 2007.

[65] R. Aldrovandi, *Special matrices of mathematical physics: stochastic, circulant, and Bell matrices*. World Scientific, 2001.

[66] C. R. Rao, "Factorial experiments derivable from combinatorial arrangements of arrays," *Suppl. J. Royal Stat. Soc.*, vol. 9, no. 1, pp. 128–139, 1947.

[67] M. Grindal, J. Offutt, and S. F. Andler, "Combination testing strategies: a survey," *Softw. Test. Verification Reliab.*, vol. 15, no. 3, pp. 167–199, 2005.

[68] A. S. Hedayat, N. J. A. Sloane, and J. Stufken, *Orthogonal arrays: theory and applications*. Springer Science & Business Media, 2012.

[69] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Rev. Mod. Phys.*, vol. 81, no. 2, p. 865, 2009.

[70] N. Linden, S. Popescu, and S. Popescu, "On Multi-Particle entanglement," *Fortschritte Physik: Prog. Phys.*, vol. 46, no. 4-5, pp. 567–578, 1998.

[71] H. Carteret, A. Higuchi, and A. Sudbery, "Multipartite generalization of the Schmidt decomposition," *J. Math. Phys.*, vol. 41, no. 12, pp. 7932–7939, 2000.

[72] M. Walter, D. Gross, and J. Eisert, "Multipartite entanglement," *arXiv:1612.02437*, Dec. 2016.

[73] M. A. Nielsen, "Conditions for a class of entanglement transformations," *Phys. Rev. Lett.*, vol. 83, no. 2, p. 436, 1999.

[74] W. Dür, G. Vidal, and J. I. Cirac, "Three qubits can be entangled in two inequivalent ways," *Phys. Rev. A*, vol. 62, no. 6, p. 062314, 2000.

[75] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, "Four qubits can be entangled in nine different ways," *Phys. Rev. A*, vol. 65, no. 5, p. 052112, 2002.

[76] P. Vrana and M. Christandl, "Asymptotic entanglement transformation between W and GHZ states," *Journal of Mathematical Physics*, vol. 56, no. 2, p. 022204, 2015.

[77] G. Vidal, "Entanglement monotones," *J. Mod. Opt.*, vol. 47, no. 2-3, pp. 355–376, 2000.

[78] M. B. Plenio and S. Virmani, "An introduction to entanglement measures," *arXiv:0504163*, 2005.

[79] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, "Entanglement in many-body systems," *Rev. Mod. Phys.*, vol. 80, no. 2, p. 517, 2008.

[80] W. Helwig and W. Cui, "Absolutely Maximally Entangled States: Existence and Applications," *arXiv:1306.2536*, June 2013.

[81] D. Goyeneche and K. Życzkowski, "Genuinely multipartite entangled states and orthogonal arrays," *Phys. Rev. A*, vol. 90, no. 2, p. 022316, 2014.

[82] D. Goyeneche, J. Bielawski, and K. Życzkowski, "Multipartite entanglement in heterogeneous systems," *Phys. Rev. A*, vol. 94, no. 1, p. 012346, 2016.

[83] D. C. Montgomery, *Design and analysis of experiments*. John Wiley & Sons, 2017.

[84] S.-W. Cheng and K. Q. Ye, "Geometric isomorphism and minimum aberration for factorial designs with quantitative factors," *Ann. Stat.*, vol. 32, no. 5, pp. 2168–2185, 2004.

[85] E. Carlini and G. Pistone, "Hilbert bases for orthogonal arrays," *J. Stat. Theory Pract.*, vol. 1, no. 3-4, pp. 299–309, 2007.

[86] W. Bruns and J. Gubeladze, *Polytopes, rings, and K-theory*, vol. 27. Springer, 2009.

[87] D. Hilbert, *Theory of algebraic invariants*. Cambridge University Press, 1993.

[88] W. Bruns and B. Ichim, "Normaliz: algorithms for affine monoids and rational cones," *J. Algebra*, vol. 324, no. 5, pp. 1098–1113, 2010.

[89] R. P. Stanley, *Combinatorics and commutative algebra*, vol. 41. Springer Science & Business Media, 2007.

[90] D. A. Bulutoglu and F. Margot, "Classification of orthogonal arrays by integer programming," *J. Stat. Plan. Inference*, vol. 138, no. 3, pp. 654–666, 2008.

[91] B. D. McKay and A. Piperno, "Practical graph isomorphism, II," *J. Symb. Comput.*, vol. 60, pp. 94–112, 2014.

[92] A. J. Coleman, "Structure of fermion density matrices," *Rev. Mod. Phys.*, vol. 35, pp. 668–686, July 1963.

[93] T. Tyc and J. Vlach, "Quantum marginal problems," *Eur. Phys. J. D*, vol. 69, no. 9, p. 209, Sept. 2015.

[94] Y.-K. Liu, M. Christandl, and F. Verstraete, "Quantum computational complexity of the N-representability problem: QMA complete," *Phys. Rev. Lett.*, vol. 98, no. 11, p. 110503, Mar. 2007.

[95] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and quantum computation*. American Mathematical Society, 2002.

[96] F. Kirwan, "Convexity properties of the moment mapping, III," *Invent. Math.*, vol. 77, no. 3, pp. 547–552, Oct. 1984.

[97] A. Higuchi, A. Sudbery, and J. Szulc, "One-qubit reduced states of a pure many-qubit state: polygon inequalities," *Phys. Rev. Lett.*, vol. 90, Mar. 2003.

[98] A. Higuchi, "On the one-particle reduced density matrices of a pure three-qutrit quantum state," *arXiv:0309186*, Sept. 2003.

[99] A. Klyachko, "Quantum marginal problem and representations of the symmetric group," *arXiv:0409113*, Sept. 2004.

[100] A. A. Klyachko, "The Pauli exclusion principle and beyond," *arXiv:0904.2009*, Apr. 2009.

[101] M. Altunbulak and A. Klyachko, "The Pauli principle revisited," *Commun. Math. Phys.*, vol. 282, no. 2, pp. 287–322, Sept. 2008.

[102] R. E. Borland and K. Dennis, "The conditions on the one-matrix for three-body fermion wavefunctions with one-rank equal to six," *J. Phys. B: At. Mol. Phys.*, vol. 5, no. 1, pp. 7–15, Jan. 1972.

[103] M. B. Ruskai, "Connecting N-representability to Weyl's problem: The one particle density matrix for $N = 3$ and $R = 6$," *J. Phys. A: Math. Theor.*, vol. 40, no. 45, pp. F961–F967, Nov. 2007.

[104] A. Klyachko, "Quantum marginal problem and N-representability," *J. Phys.: Conf. Ser.*, vol. 36, no. 1, pp. 72–86, Apr. 2006.

[105] J.-G. Luque and J.-Y. Thibon, "The polynomial invariants of four qubits," *Phys. Rev. A*, vol. 67, no. 4, Apr. 2003.

[106] R. Bagnara, P. M. Hill, and E. Zaffanella, "The Parma Polyhedra Library: Toward a complete set of numerical abstractions for the analysis and verification of hardware and software systems," *arXiv:0612085*, Dec. 2006.

[107] L. Seveso, D. Goyeneche, and K. Życzkowski, "Coarse-grained entanglement classification through orthogonal arrays," *J. Math. Phys.*, vol. 59, no. 7, p. 072203, Jul. 2018.

[108] P. Facchi, G. Florio, G. Parisi, and S. Pascazio, "Maximally multipartite entangled states," *Phys. Rev. A*, vol. 77, no. 6, p. 060304, Jun. 2008.

[109] A. J. Scott, "Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions," *Phys. Rev. A*, vol. 69, no. 5, p. 052330, May 2004.