Entropic measurement uncertainty relations for spin observables

Alberto Barchielli[1,2,3], Matteo Gregoratti[1,2]

[1] Politecnico di Milano, Dipartimento di Matematica,
Piazza Leonardo da Vinci 32, I-20133 Milano, Italy,
[2] Istituto Nazionale di Alta Matematica (INDAM-GNAMPA),
[3] Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Milano

23rd December 2019

Abstract

The information-theoretic formulation of quantum measurement uncertainty relations (MURs), based on the notion of relative entropy between measurement probabilities, is extended to the set of all the spin components for a generic spin s. For a physical class of approximate joint measurements of the spin components, we define the device information loss as the maximum loss of information per observable occurring in approximating the ideal incompatible components with the joint measurement at hand. By optimizing on the measuring device, we define the notion of minimum information loss. By using this notions, we show how to give a significant formulation of state independent MURs in the case of infinitely many target observables. The same construction works as well for finitely many observables, and we study the related MURs for two and three orthogonal spin components. The minimum information loss plays also the role of measure of incompatibility and in this respect it allows us to compare quantitatively the incompatibility of various sets of spin observables, with different number of involved components and different values of s.

Keywords: Measurement Uncertainty Relations; positive operator valued measures; spin s; information loss; relative entropy.

Contents

1 Introduction **2**

2 Approximate joint measurements of all spin components **4**
- 2.1 Target observables 4
- 2.2 Approximate joint measurements 4
  - 2.2.1 Post-processing. 5
  - 2.2.2 Classical noise and compatibility. 8
  - 2.2.3 Unbiased measurements. 9
- 2.3 Approximate joint measurements for spin 1/2, 1, 3/2 9
  - 2.3.1 Spin 1/2. 9
  - 2.3.2 Spin 1. 10
  - 2.3.3 Spin 3/2. 10

3 Entropic MURs for the set of all the spin components **11**
3.1 The device information loss . . . 11
3.2 The minimum information loss . . . 12
3.3 Entropic MURs . . . 13
3.4 Spin 1/2 . . . 14
3.5 Spin 1 . . . 15
3.6 Spin 3/2 . . . 17

4 MURs for two and three orthogonal components **18**
4.1 Target observables and approximate joint measurements . . . 18
4.1.1 Optimal cloning and approximate joint measurements. . . . 19
4.1.2 Spin 1/2. . . . 19
4.2 The information loss . . . 20
4.2.1 Entropic MURs. . . . 21
4.2.2 Spin 1/2. . . . 21
4.2.3 The bounds from optimal cloning. . . . 21
4.3 Some orderings and bounds . . . 22

5 Conclusions **23**

A Spin s: rotations and q-coefficients **24**
A.1 Properties of the Wigner small-d-matrix . . . 24
A.2 The q-coefficients . . . 24
A.2.1 Spin 1/2. . . . 25
A.2.2 Spin 1. . . . 25
A.2.3 Spin 3/2. . . . 25

B Orthogonal components **26**
B.1 Three orthogonal components . . . 26
B.2 Two orthogonal components . . . 26
B.3 Joint measurements from optimal cloning . . . 27

1 Introduction

Measurement uncertainty relations (MURs) quantify to which extent one can approximate a set of measurements of incompatible observables by a single joint measurement [1–17]. On the other side, one speaks of preparation uncertainty relations (PURs) when some lower bound is given on the "spreads" of the distributions of some observables measured in the same state [13–24]. An important point in both types of uncertainty relations is to arrive to formulate them for more than two observables [5, 13, 15, 19, 21, 22, 24].

The concept of MURs needs to introduce approximate joint measurements of incompatible observables, and this can be realized only by using the general notion of observable, represented by a positive operator valued measure; for a presentation of the modern theory of quantum measurements see, e.g., [14, 20]. Various approaches have been proposed to quantify the "errors" due to approximate measurements, such as distances for probability measures [8–10, 12–14] or conditional entropies [2–4]. Our approach is to see the joint measurement approximation of incompatible observables as a loss of information and to quantify it by the use of the relative entropy [25–27]. In information theory, the relative entropy is the notion which allows to quantify the loss of information due to the use of an approximate probability distribution instead of the true distribution. This quantification is independent of a dilation of the measurement units and of a reordering of the possible values. In this context it is possible to arrive to MURs for any set of observables and to quantify their amount of incompatibility.

In [25] we succeeded in formulating state independent MURs for any set of n general observables taking a finite number of possible values. The lower bound appearing in these MURs was named *entropic incompatibility degree*, and it was shown to play the role of an entropy-based measure of incompatibility. The generalization to position and momentum was given in [26]. However, the formulation given in these two articles does not extend to infinitely many observables. In [27] we treated the case of all the infinite components of a spin 1/2 system, by an approach based on a mean on the directions. However, this approach cannot be extended to sets of observables for which a natural mean does not exist, and, in any case, it is very difficult to apply it to higher spins.

Now our aim is to show that it is possible to modify the previous construction in a way that allows to formulate MURs for finite and infinite sets of target observables and to compare the "quantity of incompatibility" of different sets of observables, independently of the number of elements in the sets. We shall show how to reach this goal for spin observables in the case of all the components of the spin; moreover, we show that the same construction of MURs apply to 2 or 3 orthogonal components, and that the related lower bound allows the quantitative comparison of the various cases. The main difference between the present approach and the one introduced in [25] is that now our focus is on the worst loss of information per observable, while previously it was on the total loss of information. Moreover, in the special case $s = 1/2$, we easily obtain also a state dependent form of MURs.

The idea of formulating MURs for all the components of a generic spin s was introduced in [13]. There the approximation error is quantified by Wasserstein distances between target and approximating distributions, while we want to show how also this case can be treated by an information theoretical approach. An important point, already stressed in [13], is that a joint measurement of three orthogonal components is not equivalent to a joint measurement of all the components, in arbitrary directions, and that only the case of infinite components respect the rotation symmetry. So it is meaningful to enlighten the differences between the case of the spin components in all directions and the case of orthogonal components.

Scheme of the article. In Section 2 we present the construction of the class of approximate joint measurements for all the spin components. Such a construction is based on covariant generalized observables on the sphere (Section 2.2). Then, given a measure on the sphere, we process it into an approximate joint measurement of all the spin components by a suitable discretization procedure of its output (Section 2.2.1). After a discussion of the relevant properties for a generic spin s, more explicit results are given for small spins in Section 2.3. A bound, having the role of *minimum information loss*, is introduced in Section 3. Such an index represents a lower bound in the state independent MURs for all the spin components, formulated in Remarks 5 and 7. The numerical values of the minimum information loss are computed in Section 3.4 for $s = 1/2$, in Section 3.5 for $s = 1$ and in Section 3.6 for $s = 3/2$. The MURs for two and three orthogonal components and the corresponding bounds for these cases are introduced in Section 4. We show also that the minimum information loss has the role of figure of merit to quantify the incompatibility. The ordering from the least incompatible set to the more incompatible one is given in Section 4.3, for different number of spin components (including the case of infinite components) and different spin values s. Section 5 presents conclusions and outlooks.

2 Approximate joint measurements of all spin components

Let us fix a Cartesian system x, y, z determined by the orthogonal unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$. Let $S_x \equiv S_1$, $S_y \equiv S_2$, $S_z \equiv S_3$ be an irreducible representation of the commutation relations $[S_x, S_y] = \mathrm{i}S_z$ (and cyclic relations) in the Hilbert space $\mathcal{H} = \mathbb{C}^{2s+1}$, so that $S_x^2 + S_y^2 + S_z^2 = s(s+1)\mathbb{1}$, $s = 1/2, 1, 3/2, \ldots$. Let us denote by $\mathsf{X} \equiv \mathsf{X}_1$, $\mathsf{Y} \equiv \mathsf{X}_2$, $\mathsf{Z} \equiv \mathsf{X}_3$ the projection valued measures (pv-measures) associated with the self-adjoint operators S_x, S_y, S_z (respectively) and by $\mathcal{X}$ the set of possible eigenvalues m:

$$m \in \mathcal{X} := \{-s, -s+1, \ldots, s-1, s\}. \tag{1}$$

More in general, for a direction $\boldsymbol{n}$ ($\boldsymbol{n} \in \mathbb{R}^3$, $|\boldsymbol{n}| = 1$), we denote by $\mathsf{A}_{\boldsymbol{n}}(m)$ the eigen-projections of the spin component in the direction $\boldsymbol{n}$: $\boldsymbol{n} \cdot \boldsymbol{S} = \sum_{m\in\mathcal{X}} m\mathsf{A}_{\boldsymbol{n}}(m)$. As usual we shall identify $\boldsymbol{n} \cdot \boldsymbol{S}$ and $\mathsf{A}_{\boldsymbol{n}}$ by calling both them "spin component".

Let us introduce now the usual polar angles θ, ϕ in the fixed reference system and denote by $\boldsymbol{n}(\theta, \phi)$ the unit vector in the direction determined by the polar angles θ and ϕ:

$$\theta \in [0, \pi], \quad \phi \in [0, 2\pi), \qquad \boldsymbol{n}(\theta, \phi) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}. \tag{2}$$

In the following we shall need the rotation operator

$$V(\theta, \phi) := \exp\{-\mathrm{i}\theta S_\phi\}, \qquad S_\phi := S_y\cos\phi - S_x\sin\phi = \mathrm{e}^{-\mathrm{i}\phi S_z} S_y \mathrm{e}^{\mathrm{i}\phi S_z}, \tag{3}$$

corresponding to a counterclockwise rotation of an angle θ around the unit vector $\boldsymbol{n}\big(\pi/2, \phi + \pi/2\big)$, see Appendix A. Such a rotation brings the $\boldsymbol{k}$ axis to the $\boldsymbol{n}(\theta, \phi)$ one, so that

$$V(\theta, \phi) S_z V(\theta, \phi)^\dagger = \boldsymbol{n}(\theta, \phi) \cdot \boldsymbol{S}, \tag{4}$$

$$V(\theta, \phi)\mathsf{Z}(m)V(\theta, \phi)^\dagger = \mathsf{A}_{\boldsymbol{n}(\theta,\phi)}(m), \qquad m \in \mathcal{X}. \tag{5}$$

2.1 Target observables

We already fixed the Hilbert space by taking $\mathcal{H} = \mathbb{C}^{2s+1}$; the corresponding *state space* (the space of all the statistical operators on $\mathcal{H}$) will be denoted by $\mathcal{S}_s$. In particular, in some discussions, we shall need the maximally mixed state, given by

$$\rho_0 = \frac{\mathbb{1}}{2s+1}. \tag{6}$$

The set of observables which we want to approximate by joint measurements (the *reference or target observables*) consists of all the spin components:

$$\mathcal{A}_\infty := \left\{\mathsf{A}_{\boldsymbol{n}} : \boldsymbol{n} \in \mathbb{R}^3,\ |\boldsymbol{n}| = 1\right\}. \tag{7}$$

2.2 Approximate joint measurements

To introduce the approximate joint measurements of the spin components, we have to use the general notion of observable, a positive operator value measure (POVM) [14, Sects. 4.6, 9.3], called also *resolution of the identity* [20, Sect. 2.2]. We shall denote by $\mathcal{M}(\mathcal{Y})$ the set of all the POVMs with value space $\mathcal{Y}$; for instance, we have $\mathsf{A}_{\boldsymbol{n}} \in \mathcal{M}(\mathcal{X})$. The distribution of an observable A in a state ρ will be denoted by A^ρ.

The first step is to introduce the set $\mathcal{M}(\mathcal{A}_\infty)$ of the admissible approximate joint measurements of all the spin components $\mathsf{A}_{\boldsymbol{n}}$, $|\boldsymbol{n}| = 1$. Our approach is to approximate the target observables $\mathsf{A}_{\boldsymbol{n}}$ with compatible observables $\mathsf{M}_{\boldsymbol{n}}$ that share the same output space $\mathcal{X}$ as $\mathsf{A}_{\boldsymbol{n}}$, and that can be jointly got by processing the output $\boldsymbol{\xi}$ of a rotation covariant POVM, defined on the unit sphere

$$\mathbb{S}_2 = \left\{\boldsymbol{\xi} \in \mathbb{R}^3,\ |\boldsymbol{\xi}| = 1\right\}. \tag{8}$$

The most general covariant POVM on the spherical surface in $\mathbb{R}^3$ is given in [20, Sect. 4.10], [13, Eq. (109)]:

$$\begin{aligned} &\mathsf{F}_{\boldsymbol{\lambda}}(\mathrm{d}\theta\mathrm{d}\phi) = \sum_{\ell=-s}^{+s} \lambda_\ell \mathsf{F}_\ell(\mathrm{d}\theta\mathrm{d}\phi), \qquad \lambda_\ell \geq 0, \quad \sum_{\ell=-s}^{+s} \lambda_\ell = 1, \quad \boldsymbol{\lambda} = \{\lambda_\ell\}_{\ell\in\mathcal{X}}, \\ &\mathsf{F}_\ell(\mathrm{d}\theta\mathrm{d}\phi) = (2s+1)\, \mathsf{A}_{\boldsymbol{n}(\theta,\phi)}(\ell)\, \frac{\sin\theta \mathrm{d}\theta\mathrm{d}\phi}{4\pi}. \end{aligned} \tag{9}$$

In particular, the normalization of the measure $\mathsf{F}_{\boldsymbol{\lambda}}$ for any choice of the λ's implies the normalization of the measures F_ℓ, which means

$$\int_{\theta\in[0,\pi]} \int_{\phi\in[0,2\pi)} \mathsf{F}_\ell(\mathrm{d}\theta\mathrm{d}\phi) = \mathbb{1}, \qquad \forall \ell \in \mathcal{X}. \tag{10}$$

The covariance of the POVM (9) means that for any Borel subset of the sphere $B \subset \mathbb{S}_2$, and any rotation $R \in SO(3)$ we have $U(R)\mathsf{F}_{\boldsymbol{\lambda}}(B)U(R)^\dagger = \mathsf{F}_{\boldsymbol{\lambda}}(RB)$, where the representation $U(R)$ is introduced in Appendix A. Let us note that the choice of the z-axis is arbitrary.

Remark 1 (Uniform distribution).
1. When $\lambda_\ell = \lambda_\ell^0 \equiv 1/(2s+1)$, $\forall \ell$, (5) and (9) imply that $\mathsf{F}_{\boldsymbol{\lambda}^0}(\mathrm{d}\theta\mathrm{d}\phi)$ is the uniform distribution on the sphere: $\mathsf{F}_{\boldsymbol{\lambda}^0}(\mathrm{d}\theta\mathrm{d}\phi) = \mathbb{1}\, \frac{\sin\theta}{4\pi}\, \mathrm{d}\theta\mathrm{d}\phi$.
2. Similarly, for any choice of the parameters λ_m we get the uniform distribution on the maximally mixed state (6): $\mathsf{F}_{\boldsymbol{\lambda}}^{\rho_0}(\mathrm{d}\theta\mathrm{d}\phi) = \frac{\sin\theta}{4\pi}\, \mathrm{d}\theta\mathrm{d}\phi$.

2.2.1 Post-processing.

Now we want to give a rule to process the result $\boldsymbol{\xi}$ obtained from a measurement of $\mathsf{F}_{\boldsymbol{\lambda}}$ on the system. Being $\boldsymbol{\xi}$ the observed value, for every direction $\boldsymbol{n}$ we want a value for the ideal spin component $\boldsymbol{n}\cdot\boldsymbol{S}$, obtained by a suitable discretization of $\boldsymbol{n}\cdot\boldsymbol{\xi}$. This discretization could be based on different criteria, such as angles of the same amplitude, or projections on $\boldsymbol{n}$ of the same length. In order to have a sufficiently large class of approximate measurements, we do not ask for such a restrictions; we ask only to have symmetry with respect to positive and negative values, so that we can identify $\boldsymbol{n}\cdot\boldsymbol{S}$ with $-\boldsymbol{n}\cdot\boldsymbol{S}$ up to a change of sign in the output value m.

Let us consider a set of angles dividing the interval $[0,\pi]$ into $2s+1$ pieces, symmetrically placed with respect to $\pi/2$:

$$\boldsymbol{\theta} = \{\theta_0, \theta_1, \ldots, \theta_{2s+1}\}, \qquad 0 = \theta_0 < \theta_1 < \cdots < \theta_{2s+1} = \pi, \qquad \theta_{2s+1-k} = \pi - \theta_k. \tag{11}$$

Let $\boldsymbol{\xi}$ be the result of the measurement $\mathsf{F}_{\boldsymbol{\lambda}}$ and $\boldsymbol{n}$ be a generic direction forming an angle α with $\boldsymbol{\xi}$. If we find $\alpha \in [\theta_{s-m}, \theta_{s-m+1})$ for $m = s, \ldots, -s+1$, or $\alpha \in [\theta_{2s}, \pi]$ for $m = -s$, we attribute the value $m \in \mathcal{X}$ to the spin component in direction $\boldsymbol{n}$.

In other terms, let $C_{\boldsymbol{n}}(m)$, $m \in \mathcal{X}$, be the $2s+1$ parts of the spherical surface obtained by using this discretization procedure around $\boldsymbol{n}$. For any choice of a finite number of directions $\boldsymbol{n}_1, \ldots, \boldsymbol{n}_k$, the approximate joint measurement of the spin components in that directions is represented by

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}_1,\boldsymbol{n}_2,\ldots,\boldsymbol{n}_k]}(m_1, m_2, \ldots, m_k) = \mathsf{F}_{\boldsymbol{\lambda}}\left(\bigcap_{i=1}^{k} C_{\boldsymbol{n}_i}(m_i)\right). \tag{12}$$

This expression defines a POVM belonging to $\mathcal{M}(\mathcal{X}^k)$.

Remark 2. By the construction we have followed, the POVMs (12) enjoy many properties; the most relevant properties are the following ones.

1. For a fixed $\boldsymbol{\lambda}$, the POVMs (12) are all compatible, because they are obtained by classical post-processing from a unique measure $\mathsf{F}_{\boldsymbol{\lambda}}$.

2. By the fact that we have a measure on the space of the directions (the sphere) and that the post-processing is described by the intersections in (12), the introduced POVMs are invariant under any permutation of the couples $(\boldsymbol{n}_1, m_1), \ldots, (\boldsymbol{n}_k, m_k)$, and they vanish any time the corresponding intersection is void.

3. The symmetry of the angles (11) implies that $C_{-\boldsymbol{n}}(m) = C_{\boldsymbol{n}}(-m)$ and, so, the symmetry property

$$\mathsf{M}_{\boldsymbol{\lambda},[-\boldsymbol{n}_1,\boldsymbol{n}_2,\ldots,\boldsymbol{n}_k]}(m_1, m_2, \ldots, m_k) = \mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}_1,\boldsymbol{n}_2,\ldots,\boldsymbol{n}_k]}(-m_1, m_2, \ldots, m_k). \tag{13}$$

4. By the covariance of $\mathsf{F}_{\boldsymbol{\lambda}}$, the POVMs (12) enjoy the symmetry property

$$U(R)\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}_1,\ldots,\boldsymbol{n}_k]}(m_1, \ldots, m_k)U(R)^\dagger = \mathsf{M}_{\boldsymbol{\lambda},[R\boldsymbol{n}_1,\ldots,R\boldsymbol{n}_k]}(m_1, \ldots, m_k). \tag{14}$$

The set of all these compatible POVMs implicitly defines a measure $\mathsf{M}_{\boldsymbol{\lambda}}$ for all the spin components; then, the measures (12) are k-dimensional marginals of $\mathsf{M}_{\boldsymbol{\lambda}}$.

Remark 3. The measure $\mathsf{M}_{\boldsymbol{\lambda}}$ depends on $2s + \lfloor s \rfloor$ free parameters: $2s$ parameters from the λ's and $\lfloor s \rfloor$ from the angles $\boldsymbol{\theta}$; $\lfloor s \rfloor$ is the integer part of s. The set of all these POVM's $\mathsf{M}_{\boldsymbol{\lambda}}$ is denoted by $\mathcal{M}(\mathcal{A}_\infty)$ and this is the set we take as physically sensible approximate joint measurements of all the spin components $\mathcal{A}_\infty$. To base a physical measurement $\mathsf{M}_{\boldsymbol{\lambda}}$ on a rotation covariant POVM is an idea coming from [13], while the descretization procedure is a peculiar feature of our approach based on the relative entropy.

The univariate marginal $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}$ represents the admissible approximation of $\mathsf{A}_{\boldsymbol{n}}$ and its expression turns out to be

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}(\theta,\phi)]}(m) = \mathsf{F}_{\boldsymbol{\lambda}}\big(C_{\boldsymbol{n}(\theta,\phi)}(m)\big) = V(\theta,\phi)\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}(m)V(\theta,\phi)^\dagger, \tag{15}$$

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}(m) = \int_{\theta\in[\theta_{s-m},\,\theta_{s-m+1})} \int_{\phi\in[0,2\pi]} \mathsf{F}_{\boldsymbol{\lambda}}(\mathrm{d}\theta\mathrm{d}\phi). \tag{16}$$

The compatible univariate POVMs $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}$ will be central in our formulation of the MURs and we shall call them "approximate spin components".

Remark 4. From (9) we see that $\mathsf{F}_{\boldsymbol{\lambda}}(\mathrm{d}\theta\mathrm{d}\phi)$ is a mixture of the POVMs $\mathsf{F}_\ell(\mathrm{d}\theta\mathrm{d}\phi)$; similarly, each $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}$ is a mixture, given by

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}(m) = \sum_{\ell=-s}^{+s} \boldsymbol{\lambda}_\ell \mathsf{M}_{\ell,[\boldsymbol{k}]}(m), \tag{17}$$

$$\mathsf{M}_{\ell,[\boldsymbol{k}]}(m) = (2s+1)\int_{\theta_{s-m}}^{\theta_{s-m+1}} \mathrm{d}\theta\, \frac{\sin\theta}{4\pi} \int_0^{2\pi} \mathrm{d}\phi\, \mathsf{A}_{\boldsymbol{n}(\theta,\phi)}(\ell). \tag{18}$$

In the same way, we have

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}(m) = \sum_{\ell=-s}^{+s} \lambda_\ell \mathsf{M}_{\ell,[\boldsymbol{n}]}(m), \qquad \mathsf{M}_{\ell,[\boldsymbol{n}(\theta,\phi)]}(m) = V(\theta,\phi)\mathsf{M}_{\ell,[\boldsymbol{k}]}(m)V(\theta,\phi)^\dagger. \tag{19}$$

In order to study the MURs for spin observables (Section 3), we need a more explicit form for $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}(m)$, for which the following probabilities are needed.

Definition 1 (q-coefficients)**.** We define

$$q_{\boldsymbol{\theta}}(m|\ell,h) := \mathsf{M}^{\rho_h}_{\ell,[\boldsymbol{k}]}(m) = \mathrm{Tr}\left\{\rho_h \mathsf{M}_{\ell,[\boldsymbol{k}]}(m)\right\}, \qquad \rho_h := \mathsf{Z}(h), \tag{20}$$

which is the probability of getting the result m in a measurement of $\mathsf{M}_{\ell,[\boldsymbol{k}]}$ when the system is in the eigen-state ρ_h of S_z. The vector $\boldsymbol{\theta}$ is the set of the discretization angles (11), defining $\mathsf{M}_{\ell,[\boldsymbol{k}]}$ by (18).

As stated by the following theorem, the q-coefficients involve the *Wigner small-d-matrix* [28, Sect. 3.6], defined by

$$d^{(s)}_{\ell,h}(\theta) := {}_z\langle \ell|\mathrm{e}^{-\mathrm{i}\theta S_y}|h\rangle_z, \qquad \ell,h\in\mathcal{X}, \tag{21}$$

where $|m\rangle_z$, $m\in\mathcal{X}$, is the normalized eigen-vector of S_z of eigen-value m.

Theorem 1. *Each admissible approximate measurement of $\boldsymbol{n}\cdot\boldsymbol{S}$* (17) *is diagonal in the basis of the eigen-vectors of $\boldsymbol{n}\cdot\boldsymbol{S}$; indeed, the approximate spin components* (19) *have the form*

$$\mathsf{M}_{\ell,[\boldsymbol{n}]}(m) = \sum_{h=-s}^{s} q_{\boldsymbol{\theta}}(m|\ell,h)\mathsf{A}_{\boldsymbol{n}}(h), \qquad \mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}(m) = \sum_{\ell,h=-s}^{s} q_{\boldsymbol{\theta}}(m|\ell,h)\lambda_\ell\mathsf{A}_{\boldsymbol{n}}(h), \tag{22}$$

where the q-coefficients (20) *appear. Moreover, these coefficients turn out to be given by*

$$q_{\boldsymbol{\theta}}(m|\ell,h) = \left(s+\frac{1}{2}\right)\int_{\theta_{s-m}}^{\theta_{s-m+1}} \mathrm{d}\theta\,\sin\theta\left|d^{(s)}_{\ell,h}(\theta)\right|^2, \tag{23}$$

where $d^{(s)}_{\ell,h}(\theta)$ is the Wigner small-d-matrix defined in (21).

Finally, the following properties hold: $\forall\, m,\,\ell,\,h\in\mathcal{X}$,

$$q_{\boldsymbol{\theta}}(m|\ell,h) > 0, \tag{24}$$

$$q_{\boldsymbol{\theta}}(m|\ell,h) = q_{\boldsymbol{\theta}}(m|h,\ell) = q_{\boldsymbol{\theta}}(m|-\ell,-h), \tag{25}$$

$$\sum_{\ell=-s}^{s} q_{\boldsymbol{\theta}}(m|\ell,h) = \sum_{h=-s}^{s} q_{\boldsymbol{\theta}}(m|\ell,h) = \left(s+\frac{1}{2}\right)\left(\cos\theta_{s-m} - \cos\theta_{s-m+1}\right). \tag{26}$$

Proof. By using the expressions (18) and (5) inside the probabilities (20) we get

$$\mathsf{M}^{\rho_h}_{\ell,[\boldsymbol{k}]}(m) = (2s+1)\int_{\theta_{s-m}}^{\theta_{s-m+1}} \mathrm{d}\theta\, \frac{\sin\theta}{4\pi}\int_0^{2\pi} \mathrm{d}\phi \operatorname{Tr}\left\{\mathsf{Z}(h)V(\theta,\phi)\mathsf{Z}(\ell)V(\theta,\phi)^\dagger\right\}.$$

By inserting the decomposition (92) of $V(\theta,\phi)$, we have that the dependence on ϕ disappears and (23) is obtained.

The structure of the integral in ϕ in the right hand side of (18) implies that $\mathsf{M}_{\ell,[\boldsymbol{k}]}(m)$ commutes with S_z and by the irreducibility of the spin representation it is a linear combination of the projections $\mathsf{Z}(h)$; by the previous result the coefficients in this expansion are the q's and we get $\mathsf{M}_{\ell,[\boldsymbol{k}]}(m) = \sum_{h=-s}^{s} q_{\boldsymbol{\theta}}(m|\ell,h)\mathsf{Z}(h)$. By (19) this proves (22).

As recalled in Appendix A.1, $\left|d^{(s)}_{\ell,h}(\theta)\right|^2$ is a polynomial in $\cos\theta$. As we asked $\theta_{s-m} < \theta_{s-m+1}$, the integral of this polynomial in (23) can vanish only if $\left|d^{(s)}_{\ell,h}(\theta)\right|^2 = 0$ for all θ, but this is impossible because we have

$$1 = \sum_m q_{\boldsymbol{\theta}}(m|\ell,h) = \left(s+\frac{1}{2}\right)\int_0^{\pi} \mathrm{d}\theta\,\sin\theta\left|d^{(s)}_{\ell,h}(\theta)\right|^2,$$

which follows from (23) and the fact that $q_{\boldsymbol{\theta}}(\bullet|\ell,h)$ is a probability. Therefore the strict positivity (24) holds.

Properties (25) follow immediately from the definition (20) and the symmetries (93).

The sum rules (26) follow from the property (94). □

By (22), the distribution of an approximate spin component $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}$ in a state ρ is given by the double mixture

$$\mathsf{M}^{\rho}_{\boldsymbol{\lambda},[\boldsymbol{n}]}(m) = \sum_{\ell,h=-s}^{s} q_{\boldsymbol{\theta}}(m|\ell,h)\lambda_\ell \mathsf{A}^{\rho}_{\boldsymbol{n}}(h). \tag{27}$$

2.2.2 Classical noise and compatibility.

Definition 1 says that the q-coefficients are probabilities with respect to m; then, the quantities $q_{\boldsymbol{\theta}}(\bullet|\ell,\bullet)$ and $\sum_{\ell=-s}^{s} q_{\boldsymbol{\theta}}(\bullet|\ell,\bullet)\lambda_\ell$ are transition matrices, independent of the system state ρ. Then, equations (22) and (27) can be interpreted by saying that the observables $\mathsf{M}_{\ell,[\boldsymbol{n}]}$ and $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}$ could be obtained from the target observables $\mathsf{A}_{\boldsymbol{n}}$ by perturbing them with some classical noise through a one-step stochastic evolution given by one of the transition matrices just introduced. As we have seen in Remark 2, the univariate POVMs $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}$ are all compatible because they are obtained by a classical post-processing from the unique POVM $\mathsf{F}_{\boldsymbol{\lambda}}$; the compatibility is not implied by the structure (22) alone.

Another approach [14, 29] to the construction of approximate joint measurements is to consider noisy versions of the target observables. In our case this would be to have

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}(m) = (1-\kappa)\mathsf{A}_{\boldsymbol{n}}(m) + \kappa p(m)\mathbb{1}, \qquad \kappa\in[0,1], \tag{28}$$

where $p(\cdot)$ is a classical probability, independent of the system state; the usual choice is to take uniform noise $p(m) = 1/(2s+1)$. By (22), equation (28) is equivalent to

$$\sum_{\ell\in\mathcal{X}} q(m|\ell,h)\lambda_\ell = (1-\kappa)\delta_{hm} + \kappa p(m), \tag{29}$$

which would be a very strong restriction.

We can say that the noise structure (22) is more general than the "noisy version" structure (28). We shall obtain an optimal measurement of the type "noisy version" in the case of spin $1/2$ (51), not in the case of spin 1 (59). Moreover, even the structure (22) could be too restrictive in different problems of approximating incompatible observables. If we consider only two non-orthogonal components of a spin 1/2, the "best" approximate joint measurement, even with different criteria, is not of the type (22) [9, Appendix B], [25, Sect. 3.2].

2.2.3 Unbiased measurements.

Sometimes, not only symmetries are used to restrict the class of possible approximate joint measurements of some incompatible target observables. In [10, 30, 31] spin measurements with *unbiased* marginals are considered; by this they mean that the outcomes of the measurement are uniformly distributed when the system is in the maximally mixed state. Note that in the field of inferential statistics this term has a different meaning, cf. [20, Chapt. 6].

By taking into account that our target observables $\mathsf{A}_{\boldsymbol{n}}$ are indeed unbiased in this sense, it could be reasonable to ask this restriction also for the approximating observables. In our case, by (27) and (26), to ask the uniform distribution $\mathsf{M}^{\rho_0}_{\boldsymbol{\lambda},[\boldsymbol{k}]}(m) = 1/(2s+1)$, in the maximally mixed state ρ_0 (6), implies immediately the strong restriction

$$\cos\theta_k - \cos\theta_{k+1} = \frac{2}{2s+1}, \qquad \text{i.e.} \qquad \cos\theta_k = \frac{2s+1-2k}{2s+1}. \tag{30}$$

This choice corresponds to discretize $\boldsymbol{n}\cdot\boldsymbol{\xi}$ by dividing the interval $[-1,1]$ into subintervals of equal length. By using the minimization of information loss as criterium of goodness, as done in Section 3, the best approximate joint measurement not always satisfies this restriction (see Sections 3.5, 3.6) and we do not ask unbiasedness. Let us note that the noisy spin observable (28) is unbiased if and only if $p(m) = 1/(2s+1)$.

2.3 Approximate joint measurements for spin 1/2, 1, 3/2

For small spins we can get explicit results by particularizing the discretization procedure of Section 2.2.1 and using the q-coefficients computed in Appendix A.2.

2.3.1 Spin 1/2.

In this case only three angles appear in the post-processing and they are completely determined by (11): $\theta_0 = 0$, $\theta_1 = \pi/2$, $\theta_2 = \pi$. So, no free parameter is introduced by the discretization of the directions and a single free parameter remains, coming from the λ's, see Remark 3. These angles automatically satisfy (30) and this means that for $s = 1/2$ any observable in $\mathcal{M}(\mathcal{A}_\infty)$ is unbiased in the sense of Section 2.2.3.

The most general expression of the approximate spin components (17), (18) has been already obtained in [27, Sect. 5], but it can be computed also from the explicit form of the q-coefficients given in (96):

$$\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}(m) = \frac{\mathbb{1}}{2} + \left(\lambda_{1/2} - \frac{1}{2}\right) 2mS_z, \qquad \lambda_{1/2} \in [0,1]. \tag{31}$$

By using $\mathsf{Z}(m) = \mathbb{1} - \mathsf{Z}(-m)$, we can rewrite (31) as

$$\begin{aligned} \mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}(m) &= \left(\frac{3}{2} - \lambda_{1/2}\right)\frac{\mathbb{1}}{2} + \left(\lambda_{1/2} - \frac{1}{2}\right)\mathsf{Z}(m) \\ &= \left(\frac{1}{2} + \lambda_{1/2}\right)\frac{\mathbb{1}}{2} + \left(\frac{1}{2} - \lambda_{1/2}\right)\mathsf{Z}(-m), \end{aligned} \tag{32}$$

from which we see that $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{k}]}$ is a noisy version of Z, i.e. it is in the form (28), only when $\lambda_{1/2} \geq \frac{1}{2}$.

For $s = 1/2$ the probabilities (27) can be easily computed. Firstly, any state can be parameterized as

$$\rho = \frac{1}{2}\left(\mathbb{1} + 2\boldsymbol{r} \cdot \boldsymbol{S}\right), \qquad r = |\boldsymbol{r}| \leq 1; \tag{33}$$

note that $2\boldsymbol{S}$ is the vector of the Pauli matrices. Then, by (5) and (31), we have

$$\mathsf{A}^{\rho}_{\boldsymbol{n}}(m) = \frac{1}{2} + m\,\boldsymbol{n} \cdot \boldsymbol{r}, \qquad \mathsf{M}^{\rho}_{\boldsymbol{\lambda},[\boldsymbol{n}]}(m) = \frac{1}{2} + \left(\lambda_{1/2} - \frac{1}{2}\right) m\,\boldsymbol{n} \cdot \boldsymbol{r}. \tag{34}$$

2.3.2 Spin 1.

The choice of the angles (11) gives $0 = \theta_0 < \theta_1 < \theta_2 = \pi - \theta_1 < \theta_3 = \pi$, and it introduces a single free parameter

$$a := \cos\theta_1, \qquad a \in (0,1). \tag{35}$$

Other two free parameters come from the λ's, see Remark 3. The q-coefficients are computed in Appendix A.2.2; then, the approximate spin components (22) take the expressions

$$\begin{aligned} &\mathsf{M}_{1,[\boldsymbol{k}]}(\pm 1) = \mathsf{M}_{-1,[\boldsymbol{k}]}(\mp 1) \\ &\quad = \left[1 - \frac{(1+a)^3}{8}\right]\mathsf{Z}(\pm 1) + \frac{2+a}{4}\,(1-a)^2\,\mathsf{Z}(0) + \frac{(1-a)^3}{8}\,\mathsf{Z}(\mp 1), \\ &\mathsf{M}_{1,[\boldsymbol{k}]}(0) = \mathsf{M}_{-1,[\boldsymbol{k}]}(0) = \frac{a}{2}\left(3 - a^2\right)\mathsf{Z}(0) + \frac{a}{4}\left(3 + a^2\right)\left[\mathsf{Z}(1) + \mathsf{Z}(-1)\right], \end{aligned} \tag{36}$$

$$\begin{aligned} \mathsf{M}_{0,[\boldsymbol{k}]}(\pm 1) &= \frac{2+a}{4}\,(1-a)^2\left[\mathsf{Z}(1) + \mathsf{Z}(-1)\right] + \frac{1-a^3}{2}\,\mathsf{Z}(0), \\ \mathsf{M}_{0,[\boldsymbol{k}]}(0) &= a^3\mathsf{Z}(0) + \frac{a}{2}\,(3-a^2)\left[\mathsf{Z}(1) + \mathsf{Z}(-1)\right]. \end{aligned} \tag{37}$$

To get unbiased marginals, according to (30) we would have to take $a = 1/3$. To get the marginal $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}$ to be an unbiased and noisy version of the target observable $\mathsf{A}_{\boldsymbol{n}}$ we would have to impose also (29), with an uniform probability distribution; this gives the further conditions $\lambda_{+1} = \frac{1}{3} + \frac{3}{2}\kappa$, $\lambda_0 = \frac{1}{3} - \frac{3}{2}\kappa$, $\lambda_{-1} = \frac{1}{3}$, $\kappa \in [0, 2/9]$. Both these conditions are too restrictive from the point of view of the loss of information of Section 3.

2.3.3 Spin 3/2.

For $s = 3/2$, the choice of the angles (11) gives

$$0 = \theta_0 < \theta_1 < \theta_2 = \frac{\pi}{2} < \theta_3 = \pi - \theta_1 < \theta_4 = \pi,$$

and it introduces a single free parameter: $a := \cos\theta_1$, $a \in (0,1)$. Other three free parameters come from the λ's, see Remark 3. The q-coefficients are computed in Appendix A.2.3; then, the approximate spin components are given by (22), (17) and the probability distribution by (27) (we to not write explicitly them, because the formulae are very long). To get unbiasedness, according to (30) we would have to take $a = 1/2$.

3 Entropic MURs for the set of all the spin components

The spin components are incompatible observables and a joint measurement can only approximate them. In information theory [32–34] the relative entropy is the quantity introduced to measure the error done when one uses an approximating probability distribution in place of the true one. Let us stress that the relative entropy is an intrinsic quantity: it is independent of the measure units of the involved observables and from renaming or reordering the possible values. Such a property does not hold for non entropic measures of the error.

In [25] we used as error function the sum of the relative entropies, each one involving a single target observable, because this sum represents the total loss of information; however this approach can not be extended to infinitely many observables. To overcome this difficulty, instead of the sum, we shall consider the maximum of the relative entropies over all target observables: this maximum represents the loss of information for the worst direction. Then, we consider the worst case also with respect to the system state. Finally, we shall optimize with respect to all approximating joint measurements. This is indeed the procedure used in [5, 10, 13], apart from the starting point (distances between distributions for them).

3.1 The device information loss

Let us recall that $\mathcal{A}_\infty$ (7) is the set of all the spin components (our target observables), and that $\mathcal{M}(\mathcal{A}_\infty)$ (Remark 3) is our class of covariant approximate joint measurements for all the spin components. If $\mathsf{A}_{\boldsymbol{n}} \in \mathcal{A}_\infty$ and $\mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty)$, we denote by $\mathsf{M}_{[\boldsymbol{n}]}$ the univariate marginal of M approximating $\mathsf{A}_{\boldsymbol{n}}$ and we call it the approximate spin component. With $\mathsf{A}_{\boldsymbol{n}}^\rho$ we denote the distribution of $\mathsf{A}_{\boldsymbol{n}}$ in the state ρ, and similar notation for the other observables.

To quantify the information loss due to the use of $\mathsf{M}_{[\boldsymbol{n}]}^\rho$, $\mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty)$, in place of the target distribution $\mathsf{A}_{\boldsymbol{n}}^\rho$ we take the relative entropy

$$S\big(\mathsf{A}_{\boldsymbol{n}}^\rho \big\| \mathsf{M}_{[\boldsymbol{n}]}^\rho\big) = \sum_{m\in\mathcal{X}} \mathsf{A}_{\boldsymbol{n}}^\rho(m) \log \frac{\mathsf{A}_{\boldsymbol{n}}^\rho(m)}{\mathsf{M}_{[\boldsymbol{n}]}^\rho(m)} \geq 0, \tag{38}$$

where the logarithm is with base 2: $\log \equiv \log_2$. Recall that the form $0 \log 0$ is taken to be zero and that the relative entropy can be $+\infty$ when the support of the second probability distribution is not contained in the support of the first one. By using the expression of $\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}^\rho$ given in terms of the λ's and the q-coefficients in (27), we have

$$S\big(\mathsf{A}_{\boldsymbol{n}}^\rho \| \mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}^\rho\big) = \sum_{m\in\mathcal{X}} \mathsf{A}_{\boldsymbol{n}}^\rho(m) \log \frac{\mathsf{A}_{\boldsymbol{n}}^\rho(m)}{\sum_{\ell,h} q_{\boldsymbol{\theta}}(m|\ell,h)\lambda_\ell \mathsf{A}_{\boldsymbol{n}}^\rho(h)}. \tag{39}$$

As all the q-coefficients are strictly positive (24), the relative entropy (39) is always finite.

The relative entropy (38) depends on the state and on the choice of the observable (the direction $\boldsymbol{n}$). To characterize an information loss due only to the measuring device, represented by the multi-observable M approximating all the observables in $\mathcal{A}_\infty$, we consider the worst case of (38) with respect to the system state and the measurement direction. So, we define the *device information loss* by

$$\Delta_s[\mathcal{A}_\infty \| \mathsf{M}] := \sup_{\rho\in\mathcal{S}_s,\ \boldsymbol{n}\in\mathbb{R}^3,\ |\boldsymbol{n}|=1} S\big(\mathsf{A}_{\boldsymbol{n}}^\rho \big\| \mathsf{M}_{[\boldsymbol{n}]}^\rho\big), \qquad \mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty). \tag{40}$$

This quantity is the analogue of the entropic divergence introduced in [25, Definition 2]); to use the worst case on the directions instead of the sum of the relative entropies, as done there, allows to consider also infinitely many target observables. Alternatively, in [27] we started from the mean of the relative entropies made over all the directions, but this approach gives rise to computations intractable outside the case $s = 1/2$, and without possible extensions in cases in which an invariant mean does not exist.

Theorem 2. *The device information loss* (40) *is strictly positive, the double supremum in its definition is a maximum, and we have*

$$0 < \Delta_s[\mathcal{A}_\infty \| \mathsf{M}] = \max_{\rho \in \mathcal{S}_s} S(\mathsf{A}_{\boldsymbol{n}}^\rho \| \mathsf{M}_{[\boldsymbol{n}]}^\rho) < +\infty, \qquad \forall \boldsymbol{n}, \quad \forall \mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty). \tag{41}$$

Moreover, the maximum over the states is realized in an eigen-projection of the spin component:

$$\Delta_s[\mathcal{A}_\infty \| \mathsf{M}] = \max_{m \in \mathcal{X}} S(\mathsf{A}_{\boldsymbol{n}}^{\rho_m^{\boldsymbol{n}}} \| \mathsf{M}_{[\boldsymbol{n}]}^{\rho_m^{\boldsymbol{n}}}), \qquad \rho_m^{\boldsymbol{n}} := \mathsf{A}_{\boldsymbol{n}}(m), \qquad \forall \mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty). \tag{42}$$

Finally, in terms of the q-coefficients (20)*, the device information loss* (40) *is given by*

$$\Delta_s[\mathcal{A}_\infty \| \mathsf{M}_{\boldsymbol{\lambda}}] = \log \left(\min_{m \in \mathcal{X}} \sum_{\ell} \lambda_\ell q_{\boldsymbol{\theta}}(m|\ell, m) \right)^{-1}, \qquad \forall \mathsf{M}_{\boldsymbol{\lambda}} \in \mathcal{M}(\mathcal{A}_\infty). \tag{43}$$

Proof. The relative entropy is equal to zero if and only if the two probability distributions coincide; by the incompatibility of the spin observables, the device information loss (40) is strictly positive.

In the double sup in (40) we can execute the supremum over the states first. By covariance, the quantity $\sup_{\rho \in \mathcal{S}_s} S(\mathsf{A}_{\boldsymbol{n}}^\rho \| \mathsf{M}_{[\boldsymbol{n}]}^\rho)$ is independent of $\boldsymbol{n}$ and we obtain

$$\Delta_s[\mathcal{A}_\infty \| \mathsf{M}] = \sup_{\rho \in \mathcal{S}_s} S(\mathsf{A}_{\boldsymbol{n}}^\rho \| \mathsf{M}_{[\boldsymbol{n}]}^\rho) = \sup_{\rho \in \mathcal{S}_s} S(\mathsf{Z}^\rho \| \mathsf{M}_{[\boldsymbol{k}]}^\rho).$$

By convexity, the supremum over the states of the expression (39) is a maximum among the $2s + 1$ eigen-states of S_z and we get (42), the equality in (41), and

$$\sup_{\rho \in \mathcal{S}_s} S(\mathsf{Z}^\rho \| \mathsf{M}_{[\boldsymbol{k}]}^\rho) = \max_{m \in \mathcal{X}} \log \left(\sum_{m'} \lambda_{m'} q(m|m', m) \right)^{-1}.$$

Then, the device information loss (40) can be written in the form (43), which is finite because of the strict positivity (24) of the q's. □

3.2 The minimum information loss

By optimizing over the approximate joint measurement M we get a lower bound for the device information loss

$$I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] := \inf_{\mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty)} \Delta_s[\mathcal{A}_\infty \| \mathsf{M}]; \tag{44}$$

we call it *minimum information loss*.

The quantity $I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)]$ has interesting properties; in particular, as shown in Theorem 3, it is strictly positive. Moreover, in the spin definition given in Section 2.1 we have used $\hbar = 1$, but (44) is independent of this choice, because of the invariance properties of the relative entropy. The minimum information loss will appear in the formulations of the MURs (Section 3.3) and it can be used as a measure of the incompatibility of the set of the target observables. The expression (44) can be elaborated and a more explicit form can be obtained.

Theorem 3. *The minimum information loss* (44) *can be expressed in terms of the q-coefficients* (20) *as*

$$I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] = \log{(K_s)}^{-1}, \qquad K_s := \sup_{\boldsymbol{\lambda},\boldsymbol{\theta}} \min_m \sum_\ell \lambda_\ell q_{\boldsymbol{\theta}}(m|\ell,m), \tag{45}$$

where $\boldsymbol{\theta}$ *is the set of angles satisfying the discretization conditions* (11) *and involved in the expression* (23) *of the q-coefficients. Moreover, the following bounds hold:*

$$0 < I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] \leq \log{(2s+1)}\,. \tag{46}$$

Proof. To get $I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)]$ from (43), one has to minimize over the λ's and the discretization angles:

$$I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] = \inf_{\boldsymbol{\lambda},\boldsymbol{\theta}} \log \left(\min_m \sum_\ell \lambda_\ell q_{\boldsymbol{\theta}}(m|\ell,m) \right)^{-1} = \log \left(\sup_{\boldsymbol{\lambda},\boldsymbol{\theta}} \min_m \sum_\ell \lambda_\ell q_{\boldsymbol{\theta}}(m|\ell,m) \right)^{-1};$$

this gives (45). Then, with the choice $\lambda_\ell = 1/(2s+1)$ and (30) for the angles, we have

$$K_s \geq \sup_{\boldsymbol{\theta}} \max_m \sum_\ell \frac{q_{\boldsymbol{\theta}}(m|\ell,m)}{2s+1} = \frac{1}{2}\left(\cos\theta_{s-m} - \cos\theta_{s-m+1}\right) = (2s+1)^{-1};$$

this proves the upper bound in (46).

To prove the first inequality in (46) we relay on the results of [25]. The entropic incompatibility degree for two target observables, defined in [25, (10)], is strictly positive when the two observables are incompatible [25, Theor. 2, point (v)]. Moreover, the class of the POVMs on $\mathcal{X}^2$, $\mathsf{M} \in \mathcal{M}(\mathcal{X}^2)$, is larger than the class of the bivariate marginals of measures in $\mathcal{M}(\mathcal{A}_\infty)$. By starting from two orthogonal spin components, X, Y, we get

$$\begin{aligned} 0 &\overset{(1)}{<} c_{\rm inc}(\mathsf{X},\mathsf{Y}) \overset{(2)}{=} \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{X}^2)} \sup_{\rho\in\mathcal{S}_s} \sum_{i=1}^2 S\big(\mathsf{X}_i^\rho \| \mathsf{M}_{[i]}^\rho\big) \overset{(3)}{\leq} \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{X}^2)} \sup_{\rho\in\mathcal{S}_s} 2 \max_{i=1,2} S\big(\mathsf{X}_i^\rho \| \mathsf{M}_{[i]}^\rho\big) \\ &\overset{(4)}{\leq} 2 \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{A}_\infty)} \sup_{\rho\in\mathcal{S}_s} \max_{i=1,2} S\big(\mathsf{X}_i^\rho \| \mathsf{M}_{[i]}^\rho\big) \overset{(5)}{\leq} 2 \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{A}_\infty)} \sup_{\rho\in\mathcal{S}_s,\, \boldsymbol{n}\in\mathbb{R}^3,\, |\boldsymbol{n}|=1} S\left(\mathsf{A}_{\boldsymbol{n}}^\rho \| \mathsf{M}_{[\boldsymbol{n}]}^\rho\right) \\ &\overset{(6)}{=} 2 I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)]. \end{aligned}$$

Here (1) is the result of [25], (2) is the definition of $c_{\rm inc}$, (3) is because we substitute the sum with two times the maximum, (4) is because we have restricted the class of approximating joint measurements in the infimum, (5) is because we enlarge the set of directions in the maximum, (6) is by our definition (40), (44). This ends the proof of the strict positivity. □

Let us remark that the last part of the proof, proving the strict positivity in (46), works for every class of approximate joint measurements one could use in the infimum, not only for our choice $\mathcal{M}(\mathcal{A}_\infty)$. The only point is that every spin component $\mathsf{A}_{\boldsymbol{n}}$ has to be approximated by a POVM $\mathsf{M}_{[\boldsymbol{n}]}$ on the same output space $\mathcal{X}$.

3.3 Entropic MURs

By the strict positivity of the minimum information loss and its definition (44), we get a first formulation of the MURs, in a state independent form, which is analogous to that given in [13, (11)].

Remark 5 (MURs, first version). For every physical approximate joint measurement M of all the spin components, the device information loss (40) is greater than a strictly positive lower bound:

$$\Delta_s[\mathcal{A}_\infty\|\mathsf{M}] \geq I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] > 0, \qquad \forall \mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty).$$

Remark 6. By the expression (42) of the device information loss, we can write (45) as

$$I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] = \inf_{\boldsymbol{\lambda},\boldsymbol{\theta}} \max_m S(\mathsf{A}_{\boldsymbol{n}}^{\rho_m^{\boldsymbol{n}}}\|\mathsf{M}_{\boldsymbol{\lambda},[\boldsymbol{n}]}^{\rho_m^{\boldsymbol{n}}}), \tag{47}$$

where $\rho_m^{\boldsymbol{n}}$ is the eigen-projection of $\boldsymbol{n}\cdot\boldsymbol{S}$ with respect to the eigen-value m and the discretization angles are implicitly contained in $\mathsf{M}_{\boldsymbol{\lambda}}$. When the infimum is realized in a point $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$, $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ we have that $\mathsf{M}_{\boldsymbol{\lambda}^*}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}$ plays the role of optimal approximate joint measurement.

The upper bound in (46) is surely non tight, but its role is at least to say that, when we have a device information loss greater than that, the approximating measurement is not optimal.

By the fact that the device information loss is a maximum and has the form (42), we have immediately the following formulation of the MURs.

Remark 7 (MURs, second version). The state independent MURs are

$$\forall \mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty) \quad \forall \mathsf{A}_{\boldsymbol{n}} \in \mathcal{A}_\infty \qquad \exists \rho \in \mathbb{S}_s: \ S(\mathsf{A}_{\boldsymbol{n}}^\rho\|\mathsf{M}_{[\boldsymbol{n}]}^\rho) \geq I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] > 0; \tag{48}$$

such a state ρ is one of the eigen-projections of $\boldsymbol{n}\cdot\boldsymbol{S}$.

So, in a physical approximate joint measurement M of all the spin components $\mathsf{A}_{\boldsymbol{n}}$, $\boldsymbol{n} \in \mathbb{S}_2$, the loss of information $S\big(\mathsf{A}_{\boldsymbol{n}}^\rho\big\|\mathsf{M}_{[\boldsymbol{n}]}^\rho\big)$ per direction $\boldsymbol{n}$ can not be arbitrarily reduced. It depends on the state ρ and on the direction $\boldsymbol{n}$, but for every $\boldsymbol{n}$ it can be potentially as large as $I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)]$.

By the comments after Theorem 3 we have that the MURs can be formulated also if we change the class of physical approximate joint measurements $\mathcal{M}(\mathcal{A}_\infty)$ with some other class; what can change is the value of the minimum information loss.

3.4 Spin 1/2

In this case no free parameter comes out from the angle discretization and the approximate spin components (32) are very simple.

Theorem 4. *The device information loss* (40) *and the minimum information loss* (44) *turn out to be given by*

$$\Delta_{1/2}[\mathcal{A}_\infty\|\mathsf{M}_{\boldsymbol{\lambda}}] = \log\frac{4}{1+2\lambda_{1/2}}, \tag{49}$$

$$I_{1/2}[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] = S\big(\mathsf{Z}^{\rho_m}\big\|\mathsf{M}_{1/2,[\boldsymbol{k}]}^{\rho_m}\big) = \log\frac{4}{3} \simeq 0.415037, \tag{50}$$

where $\rho_m = \mathsf{Z}(m)$.

The first equality in (50) *shows that* $\mathsf{M}_{1/2}$ *is the optimal measurement in the sense of Remark 6; its marginal in direction* $\boldsymbol{n}$ *is*

$$\mathsf{M}_{1/2,[\boldsymbol{n}]}(m) = \frac{1}{2}\left[\frac{\mathbb{1}}{2} + \mathsf{A}_{\boldsymbol{n}}(m)\right], \tag{51}$$

which is an unbiased noisy version of $\mathsf{A}_{\boldsymbol{n}}$ *(cf. Section 2.3.1).*

Proof. In this case, by (96) we have $q(m|\ell, m) = \frac{1+\ell}{2}$, independent of m; then, (49) follows from (43).

Directly from the definition (44) and the expression (49) we have

$$I_{1/2}[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] = \inf_{\lambda_{1/2}\in[0,1]} \Delta_{1/2}[\mathcal{A}_\infty \| \mathsf{M}_{\boldsymbol{\lambda}}] = \Delta_{1/2}[\mathcal{A}_\infty \| \mathsf{M}_{1/2}],$$

and the final expressions in (50) follow.

By the facts that there is no freedom in the choice of the θ's and that the infimum is reached for $\lambda_{1/2} = 1$, we get that $\mathsf{M}_{1/2}$ is the optimal measurement. Then, by (32) we get the form of the marginal (51). □

Let us remark that, actually, $\mathsf{M}_{1/2}$ enjoys an additional and useful property. By using the state representation (33) and the explicit expressions (34) for the probabilities, we have

$$S\big(\mathsf{A}^\rho_{[\boldsymbol{n}]} \big\| \mathsf{M}^\rho_{\boldsymbol{\lambda},[\boldsymbol{n}]}\big) = s\Big(\lambda_{1/2} - 1/2,\, \boldsymbol{n}\cdot\boldsymbol{r}\Big), \tag{52}$$

$$s(c,x) := \frac{1+x}{2}\log\frac{1+x}{1+cx} + \frac{1-x}{2}\log\frac{1-x}{1-cx}, \qquad |c|<1, \quad |x|\leq 1. \tag{53}$$

The parameter $\boldsymbol{r}$ is the Bloch vector characterizing the state ρ. By taking the c-derivative, we see that it is strictly negative, which implies that $s(c,x)$ decreases when c increases. This means that $\mathsf{M}_{1/2}$ minimizes (52) for any state ρ. This peculiarity of the case $s = 1/2$ makes possible to state that $\mathsf{M}_{1/2}$ is optimal even when we know the system state ρ and to easily formulate also a form of state dependent MURs.

Remark 8 (State dependent MURs). The following state dependent bound holds:

$$\begin{gathered} S\big(\mathsf{A}^\rho_{\boldsymbol{n}} \big\| \mathsf{M}^\rho_{[\boldsymbol{n}]}\big) \geq S\big(\mathsf{A}^\rho_{\boldsymbol{n}} \big\| \mathsf{M}^\rho_{1/2,[\boldsymbol{n}]}\big) = \sum_{\epsilon=\pm1} \frac{1+\epsilon\boldsymbol{n}\cdot\boldsymbol{r}}{2}\log\frac{1+\epsilon\boldsymbol{n}\cdot\boldsymbol{r}}{1+\frac{\epsilon}{2}\boldsymbol{n}\cdot\boldsymbol{r}}, \\ \forall\rho\in\mathcal{S}_s, \qquad \forall\mathsf{M}\in\mathcal{M}(\mathcal{A}_\infty), \qquad \forall\boldsymbol{n}\in\mathbb{R}^3, \quad |\boldsymbol{n}|=1. \end{gathered} \tag{54}$$

3.5 Spin 1

In this case there is a single parameter (35) coming from the angle discretization; then, the minimum information loss and the optimal measurement can be computed.

Theorem 5. *Let us set $\rho_m = \mathsf{Z}(m)$; then,*

$$I_1[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] = S\big(\mathsf{Z}^{\rho_m} \big\| \mathsf{M}^{\rho_m}_{1,[\boldsymbol{k}]}\big)\Big|_{a=a_0} = \log\frac{2}{a_0\left(3-a_0^2\right)} \simeq 0.682505. \tag{55}$$

The quantity a_0 is the real solution of the equation

$$a^3 - a^2 - 5a + \frac{7}{3} = 0, \tag{56}$$

which is given by

$$a_0 = \frac{1}{3}\left(1+8\cos\alpha\right), \qquad \cos(3\alpha-\pi) = \frac{1}{8}, \qquad \alpha\in(0,\pi/2). \tag{57}$$

This gives also

$$a_0 \simeq 0.444703; \qquad \cos 3\alpha = -\frac{1}{8}, \qquad (\cos\alpha)^3 = \frac{1}{4}\left(3\cos\alpha - \frac{1}{8}\right). \tag{58}$$

Proof. From (97) we have

$$\sum_{\ell} \lambda_\ell q_a(\pm 1|\ell, \pm 1) = \lambda_+ \left[1 - \frac{(1+a)^3}{8}\right] + \lambda_- \frac{(1-a)^3}{8} + \lambda_0 \frac{(2+a)(1-a)^2}{4},$$

$$\sum_{\ell} \lambda_\ell q_a(0|\ell, 0) = (\lambda_+ + \lambda_-)\frac{a}{2}\left(3 - a^2\right) + \lambda_0 a^3.$$

One can check that both these expressions have an absolute maximum in $\lambda_+ = 1$ for all $a \in (0,1)$. Then, (45) gives

$$K_1 \leq \sup_{a\in(0,1)} \min_m \sup_{\boldsymbol{\lambda}} \sum_{\ell} \lambda_\ell q_a(m|\ell, m) = \sup_{a\in(0,1)} \min_m q_a(m|1, m).$$

On the other side, by eliminating the supremum over the λ's and choosing $\lambda_\ell = \delta_{\ell,1}$ in (45), we get $K_1 \geq \sup_{a\in(0,1)} \min_m q_a(m|1,m)$; so, the equality holds and we have

$$K_1 = \sup_a \min_m q_a(m|1,m) = \sup_a \min\left\{1 - \frac{(1+a)^3}{8},\ \frac{a}{2}\left(3-a^2\right)\right\}.$$

The first term in the minimum decreases with a and the second one increases; this means that the supremum over a is reached when these two terms are equal, which happens when (56) holds. This proves (55). It is possible to check that (57) is the unique real solution of (56) and that this gives the properties (58). □

Theorem 6 (The optimal measurement)**.** *Equation* (55) *gives in particular that the optimal measurement is* $\mathsf{M}_1\big|_{a=a_0}$. *The marginal along* $\boldsymbol{k}$ *of this measurement is given by*

$$\begin{aligned} \mathsf{M}_{1,[\boldsymbol{k}]}(m)\big|_{a=a_0} &= \kappa_1 \mathsf{Z}(m) + \kappa_2 \frac{\mathbb{1}}{3} + \kappa_3 \mathsf{N}(m), \\ \mathsf{N}(0) = \mathsf{Z}(1) + \mathsf{Z}(-1), &\qquad \mathsf{N}(\pm 1) = \frac{1}{2}\mathsf{Z}(0). \end{aligned} \tag{59}$$

The weights κ_i *are positive and sum to one; their explicit expressions are*

$$\kappa_1 = \frac{3}{4}\left(1 - a_0^2\right) = \frac{1}{3}\left[2 - 4\cos\alpha - 16(\cos\alpha)^2\right] \simeq 0.601679, \tag{60}$$

$$\kappa_2 = \frac{3}{4}\left(\frac{5}{3} - 4a_0 + a_0^2\right) = \frac{1}{3}\left(1 - 20\cos\alpha + 16(\cos\alpha)^2\right) \simeq 0.064211, \tag{61}$$

$$\kappa_3 = 3a_0 - 1 = 8\cos\alpha \simeq 0.334110. \tag{62}$$

Proof. Equations (59) can be rewritten as

$$\mathsf{M}_{1,[\boldsymbol{k}]}(\pm 1)\big|_{a=a_0} = \left(\kappa_1 + \frac{\kappa_2}{3}\right)\mathsf{Z}(\pm 1) + \frac{\kappa_2}{3}\mathsf{Z}(\mp 1) + \left(\frac{\kappa_3}{2} + \frac{\kappa_2}{3}\right)\mathsf{Z}(0),$$

$$\mathsf{M}_{1,[\boldsymbol{k}]}(0)\big|_{a=a_0} = \left(\kappa_1 + \frac{\kappa_2}{3}\right)\mathsf{Z}(0) + \left(\kappa_3 + \frac{\kappa_2}{3}\right)\left(\mathsf{Z}(1) + \mathsf{Z}(-1)\right).$$

On the other side, (36) and (57) give

$$\mathsf{M}_{1,[\boldsymbol{k}]}(\pm 1)\big|_{a=a_0} = \left(\frac{7}{6} - a_0 - \frac{a_0^2}{2}\right)\mathsf{Z}(\pm 1) + \left(\frac{5}{12} - a_0 + \frac{a_0^2}{4}\right)\mathsf{Z}(\mp 1) + \left(\frac{a_0^2}{4} + \frac{a_0}{2} - \frac{1}{12}\right)\mathsf{Z}(0),$$

$$\mathsf{M}_{1,[\boldsymbol{k}]}(0)\big|_{a=a_0} = \left(\frac{7}{6} - a_0 - \frac{a_0^2}{2}\right)\mathsf{Z}(0) + \left(\frac{a_0^2}{4} + 2a_0 - \frac{7}{12}\right)\big(\mathsf{Z}(1) + \mathsf{Z}(-1)\big).$$

By identifying the coefficients and using (57), we get Equations (60)-(62). Finally, we have

$$\kappa_1 + \kappa_2 + \kappa_3 = \frac{1}{3}\big[2 - 4\cos\alpha - 16(\cos\alpha)^2 + 1 - 20\cos\alpha + 16(\cos\alpha)^2\big] + 8\cos\alpha = 1.$$

□

Remark 9. Differently from the case $s = 1/2$, for $s = 1$ the marginal $\mathsf{M}_{1,[\boldsymbol{k}]}\big|_{a=a_0}$ of the optimal measurement is not a noisy version of Z and it is not unbiased because $a_0 \neq 1/3$. Indeed, on the maximally mixed state ρ_0, the relative entropy is not zero and its value is

$$S\big(\mathsf{A}_{[\boldsymbol{n}]}^{\rho_0}\big\|\mathsf{M}_{1,[\boldsymbol{n}]}^{\rho_0}\big)_{a=a_0} = \frac{2}{3}\log\frac{6}{4+3a_0(1-a_0)} + \frac{1}{3}\log\frac{6}{3a_0(10-a_0)-5} \simeq 0.103607.$$

3.6 Spin 3/2

Theorem 7. *Let us set $\rho_m = \mathsf{Z}(m)$; then, we have*

$$I_{3/2}[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] = S\big(\mathsf{Z}^{\rho_m}\big\|\mathsf{M}_{3/2,[\boldsymbol{k}]}^{\rho_m}\big)_{a=a_0} = \log\frac{32}{45 - 24a_0 - 24a_0^2 - 8a_0^3} \simeq 0.88615563; \quad (63)$$

$\mathsf{M}_{3/2}\big|_{a=a_0}$ *is the optimal measurement. The quantity a_0 is the unique real solution in $(0,1)$ of the equation*

$$a^4 - 6a^2 - 8a + \frac{15}{2} = 0, \quad (64)$$

which gives

$$a_0 \simeq 0.6461537831. \quad (65)$$

Proof. From Appendix A.2.3 we get

$$\max_\ell q_a(\pm 3/2|\ell, \pm 3/2) = q_a(\pm 3/2|3/2, \pm 3/2) = \frac{1}{16}\left(15 - 4a - 6a^2 - 4a^3 - a^4\right),$$

a quantity which decreases with a from $\frac{15}{16}$ to 0, and

$$\max_\ell q_a(\pm 1/2|\ell, \pm 1/2) = q_a(\pm 1/2|3/2, \pm 1/2) = \frac{1}{16}\left(12a + 6a^2 - 4a^3 - 3a^4\right),$$

a quantity which increases with a from 0 to $\frac{11}{16}$. Then, as in the proof of Theorem 5, we get

$$\begin{aligned} K_{3/2} &= \sup_{a\in(0,1)}\min_m q_a(m|3/2, m) \\ &= \sup_{a\in(0,1)}\frac{1}{16}\min\left\{15 - 4a - 6a^2 - 4a^3 - a^4,\ 12a + 6a^2 - 4a^3 - 3a^4\right\}. \end{aligned}$$

By equating these two expressions we get equation (64), whose solution (65) is computed numerically. As we have

$$\min_m q_a(m|3/2, m) = \begin{cases} q_a(\pm 1/2|3/2, \pm 1/2) & \text{for } a \le a_0, \\ q_a(\pm 3/2|3/2, \pm 3/2) & \text{for } a \ge a_0, \end{cases}$$

(45) gives

$$I_{3/2}[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] = \log \left(q_{a_0}(m|3/2, m)\right)^{-1};$$

by using also (64), the final expression in (63) follows. By Theorem (3), the optimal measurement is identified and the intermediate expression in (63) follows. □

By direct computations one can check that the optimal measurement is biased and that on the maximally mixed state ρ_0 it gives

$$S\big(\mathsf{Z}^{\rho_0} \big\| \mathsf{M}^{\rho_0}_{3/2,[\boldsymbol{k}]}\big)_{a=a_0} = \frac{1}{2} \log[4a_0(1-a_0)]^{-1} \simeq 0.0644281. \tag{66}$$

Remark 10. The results we have found for small spin values give

$$0 < I_{1/2}[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] < I_1[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)] < I_{3/2}[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)]. \tag{67}$$

This chain of inequalities suggests the conjecture that $I_s[\mathcal{A}_\infty \| \mathcal{M}(\mathcal{A}_\infty)]$ could grow with s: in some sense the minimum information loss grows with the complexity of the spin system.

4 MURs for two and three orthogonal components

In this section we study the MURs for the cases of two and three orthogonal spin components. In [13] the authors remark that it is not possible to get the case of infinite components from the case of three orthogonal components; only the case of infinite components respects the rotation symmetry, while in the other case the three directions are fixed. The cases of orthogonal components involve less symmetries and there is more freedom in the construction of the approximate joint measurements; so it is meaningful to enlighten the differences between the case of the spin components in all directions and the case of orthogonal components. In principle also a few non-orthogonal components could be considered; in [25] we already considered two non-orthogonal spin components with $s = 1/2$, but with the sum of relative entropies as starting point.

The cases of orthogonal components allow to show how the minimum information loss and the related MURs can be introduced also for other sets of observables by adapting the construction of Section 3. Moreover, the minimum information loss can be used as quantification of the incompatibility of the target observables and allows to compare different sets of observables. In the cases of spin components we shall obtain orderings for different numbers of target observables and different values of s, which are not at all trivial or intuitive.

4.1 Target observables and approximate joint measurements

The first set of target observables we consider is $\mathcal{A}_3 = \{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$, which is covariant with respect to the octahedron group O, see Appendix B.1. Then, $\mathcal{M}(\mathcal{A}_3)$ is the set of observables with value space $\mathcal{X}^3$ and O-covariant in the sense of (103). By using the notation (12) and the covariance properties (13), (14), (103) we have that

$$\mathsf{M} \in \mathcal{M}(\mathcal{A}_\infty) \quad \Rightarrow \quad \mathsf{M}_{[\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}]} \in \mathcal{M}(\mathcal{A}_3). \tag{68}$$

The other set of target observables is $\mathcal{A}_2 = \{\mathsf{X}, \mathsf{Y}\}$, which is covariant with respect to the dihedral group D_4, see Appendix B.2. Then, $\mathcal{M}(\mathcal{A}_2)$ is the set of observables with value space $\mathcal{X}^2$ and D_4-covariant in the sense of (105). By using the notation (12) and the covariance properties (105), (103) we have that

$$\mathsf{M} \in \mathcal{M}(\mathcal{A}_3) \quad \Rightarrow \quad \mathsf{M}_{[\boldsymbol{i},\boldsymbol{j}]} \in \mathcal{M}(\mathcal{A}_2). \tag{69}$$

Note that the implications above are one-sided: there are elements in $\mathcal{M}(\mathcal{A}_2)$ which are not marginals of elements in $\mathcal{M}(\mathcal{A}_3)$ and the same for $\mathcal{M}(\mathcal{A}_3)$ with respect to $\mathcal{M}(\mathcal{A}_\infty)$.

We obtained the explicit form of a covariant approximate joint measurement, for two and three orthogonal components, only in the case of a spin 1/2. For a generic spin s we can give only particular covariant approximate joint measurements, such as the ones based on optimal cloning.

4.1.1 Optimal cloning and approximate joint measurements.

As approximate joint measurement of the spin components A_h, $h = 1, \ldots, \mathtt{r}$, a significant multi-observable $\mathsf{M}_{\mathrm{cl}} \in \mathcal{M}(\mathcal{X}^{\mathtt{r}})$ can be constructed by using the so called *optimal cloning* [29, 35, 36]; its univariate marginals are given by (106). Let us stress that the marginal of the multi-observable constructed by optimal cloning is always a noisy version of the target observable.

When the target observables are $\mathcal{A}_3 = \{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$, we get the multi-observable $\mathsf{M}^3_{\mathrm{cl}}$, whose univariate marginals (106) take the form

$$\mathsf{M}^3_{\mathrm{cl[i]}}(m) = \frac{1}{3(s+1)} \left[\mathbb{1} + (s+2)\, \mathsf{X}_i(m)\right], \qquad i = 1, 2, 3, \quad m \in \mathcal{X}. \tag{70}$$

Obviously $\mathsf{M}^3_{\mathrm{cl}} \in \mathcal{M}(\mathcal{X}^3)$, but one has also $\mathsf{M}^3_{\mathrm{cl}} \in \mathcal{M}(\mathcal{A}_3)$, as shown in Appendix B.3.

When the target observables are $\mathcal{A}_2 = \{\mathsf{X}, \mathsf{Y}\}$, the optimal cloning gives the bi-observable $\mathsf{M}^2_{\mathrm{cl}} \in \mathcal{M}(\mathcal{X}^2)$ and (106) becomes

$$\mathsf{M}^2_{\mathrm{cl[i]}}(m) = \frac{1}{4(s+1)} \left[\mathbb{1} + (2s+3)\, \mathsf{X}_i(m)\right], \qquad i = 1, 2, \quad m \in \mathcal{X}. \tag{71}$$

Again one has also $\mathsf{M}^2_{\mathrm{cl}} \in \mathcal{M}(\mathcal{A}_2)$, as shown in Appendix B.3.

4.1.2 Spin 1/2.

For a spin $1/2$ the explicit expressions of the general element in $\mathcal{M}(\mathcal{A}_3)$ and $\mathcal{M}(\mathcal{A}_2)$ have been obtained in [25, Proposition 5, Theorem 10] and used also in [27]. Then, the most general covariant joint measurement in $\mathcal{M}(\mathcal{A}_3)$ [27, Eq. (11)] can be written as

$$\mathsf{M}_c(m_1, m_2, m_3) = \frac{\mathbb{1}}{8} + \frac{c}{2}\left(m_1 S_x + m_2 S_y + m_3 S_z\right), \quad |c| \leq \frac{1}{\sqrt{3}}. \tag{72}$$

Similarly, the most general element in $\mathcal{M}(\mathcal{A}_2)$ has the expression [27, Eq. (7)]

$$\mathsf{M}_c(m_1, m_2) = \frac{\mathbb{1}}{4} + c\left(m_1 S_x + m_2 S_y\right), \qquad |c| \leq \frac{1}{\sqrt{2}}. \tag{73}$$

Remark 11. In both the cases of two and three orthogonal components, the univariate marginals have the expression

$$\mathsf{M}_{c[i]}(m) = \frac{\mathbb{1}}{2} + 2cmS_i = \begin{cases} c\mathsf{X}_i(m) + (1-c)\,\frac{\mathbb{1}}{2}, & c \geq 0, \\ |c|\,\mathsf{X}_i(-m) + (1-|c|)\,\frac{\mathbb{1}}{2}, & c < 0; \end{cases} \tag{74}$$

the only difference is the maximally possible value for $|c|$: $|c| \leq 1/\sqrt{3}$ in the case of three components and $|c| \leq 1/\sqrt{2}$ in the case of two components. Also the marginal of the optimal measurement (51) for infinite components has the form (74) with $c = 1/2$.

Remark 12. By particularizing (70) and (71) to $s = 1/2$, we obtain that the marginals of the joint measurements from optimal cloning have again the form (74) with $c = 5/9$ in the case of three components and $c = 2/3$ in the case of two components. As we have $1/2 < 5/9 < 1/\sqrt{3} < 2/3 < 1/\sqrt{2}$, there is an increase of minimum noise in going from the case of two orthogonal components, to cloning of two components, three components, cloning of three components, infinite components.

4.2 The information loss

Analogously to what is done in Section 3, also in the case of orthogonal spin components it is possible to define the device information loss and the minimum information loss. The *device information loss* of M is defined as in (40); then, exactly as for (41), after the supremum on the states, the covariance implies the independence from the direction. So, we have: for $\mathtt{r} = 2, 3$,

$$\Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}] := \sup_{\rho\in\mathcal{S}_s,\ i:i\leq\mathtt{r}} S\left(\mathsf{X}_i^\rho\|\mathsf{M}_{[i]}^\rho\right) = \sup_{\rho\in\mathcal{S}_s} S\left(\mathsf{X}_i^\rho\|\mathsf{M}_{[i]}^\rho\right), \qquad \mathsf{M}\in\mathcal{M}(\mathcal{A}_\mathtt{r}). \tag{75}$$

By optimizing over the approximate joint measurement M we get the *minimum information loss*

$$I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{A}_\mathtt{r})] := \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{A}_\mathtt{r})} \Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}] = \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{A}_\mathtt{r})} \sup_{\rho\in\mathcal{S}_s} S\left(\mathsf{X}_i^\rho\|\mathsf{M}_{[i]}^\rho\right), \qquad \mathtt{r} = 2, 3. \tag{76}$$

As in [12, 25], we can extend the previous definitions to non-symmetric approximate joint measurements, without changing the final conclusions. Firstly, we introduce the device information loss for general measurements:

$$\Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}] = \sup_{\rho\in\mathcal{S}_s,\ i:i\leq\mathtt{r}} S\left(\mathsf{X}_i^\rho\|\mathsf{M}_{[i]}^\rho\right), \qquad \mathsf{M}\in\mathcal{M}(\mathcal{X}^\mathtt{r}), \qquad \mathtt{r} = 2, 3. \tag{77}$$

Obviously, now we cannot eliminate the maximum over the directions as in (75), because this follows from the covariance. Then, we optimize over all these measurements by defining

$$I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{X}^\mathtt{r})] := \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{X}^\mathtt{r})} \Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}] = \inf_{\mathsf{M}\in\mathcal{M}(\mathcal{X}^\mathtt{r})} \sup_{\rho\in\mathcal{S}_s,\ i:i\leq\mathtt{r}} S\left(\mathsf{X}_i^\rho\|\mathsf{M}_{[i]}^\rho\right), \qquad \mathtt{r} = 2, 3. \tag{78}$$

Next proposition shows that this extension does not change the value of the minimum information loss and that this value grows with the increasing complexity of the set of observables, i.e. going from $\mathcal{A}_2$, to $\mathcal{A}_3$, and then to $\mathcal{A}_\infty$

Proposition 8. *The two definitions* (76) *and* (78) *are equivalent, as we have*

$$I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{X}^\mathtt{r})] = I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{A}_\mathtt{r})], \qquad \mathtt{r} = 2, 3. \tag{79}$$

Moreover, the minimum information loss is strictly positive and finite and we have

$$0 < I_s[\mathcal{A}_2\|\mathcal{M}(\mathcal{A}_2)] \leq I_s[\mathcal{A}_3\|\mathcal{M}(\mathcal{A}_3)] \leq I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] < +\infty. \tag{80}$$

Proof. The proof of (79) is a very slight modification of what is done in [25]. Let us use the notation $G_3 = O$ and $G_2 = D_4$ for the two groups introduced in AppendicesB.1 and B.2; the actions of these two groups on the POVMs, as given in the two appendices, can be seen to satisfy the hypotheses of Theorem 9 of [25], as done in [25, Sections B.2, B.4]. We denote by $g\mathsf{M}$ the action of an element $g \in G_\mathtt{r}$ on the POVM $\mathsf{M} \in \mathcal{M}(\mathcal{X}^\mathtt{r})$ and by $\mathsf{M}_{G_\mathtt{r}} \in \mathcal{M}(\mathcal{A}_\mathtt{r})$ the *covariant version* of M as done in [25, Sections 3.1, 4.1]. Thanks to the hypotheses on the group action of [25, Theorem 9], by substituting the sum of the relative entropies by their maximum, we get that the results on the *entropic divergence* of Theorems 4 and 9 of [25] go into analogous results on the device information loss. In this way one proves that, for $\mathtt{r} = 2, 3$,

$$\Delta_s[\mathcal{A}_\mathtt{r}\|g\mathsf{M}] = \Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}], \qquad \forall g \in G_\mathtt{r}, \qquad \forall \mathsf{M} \in \mathcal{M}(\mathcal{X}^\mathtt{r}),$$

$$\Delta_s\big[\mathcal{A}_\mathtt{r}\big\|\mathsf{M}_{G_\mathtt{r}}\big] \leq \Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}], \qquad \forall \mathsf{M} \in \mathcal{M}(\mathcal{X}^\mathtt{r}).$$

As $\mathsf{M}_{G_\mathtt{r}} \in \mathcal{M}(\mathcal{A}_\mathtt{r})$, by taking the infimum we get (79).

To prove (80), note that, by (68) and (69), the definition (76) gives the ordering among the three information losses $I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{A}_\mathtt{r})]$, $\mathtt{r} = 2, 3, \infty$. We already proved the last inequality in Theorem 3, cf. the upper bound in (46). The proof of the strict positivity is analogous to the proof of the strict positivity in (46). Exactly as in the final part of the proof of Theorem 3 we obtain $0 < c_{\rm inc}(\mathsf{X},\mathsf{Y}) \leq 2I_s[\mathcal{A}_2\|\mathcal{M}(\mathcal{A}_2)]$, where $c_{\rm inc}(\mathsf{X},\mathsf{Y})$ is defined in [25, (10)]. □

4.2.1 Entropic MURs.

By the definition and the strict positivity of the minimum information loss we get the state independent MURs in a formulation involving the device information loss:

$$\Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}] \geq I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{A}_\mathtt{r})] > 0, \qquad \forall \mathsf{M} \in \mathcal{M}(\mathcal{X}^\mathtt{r}) \supset \mathcal{M}(\mathcal{A}_\mathtt{r}). \tag{81}$$

We have used (79) to extend the set of possible measurements M. This form of MURs is the analog of what is done in Remark 5 for the case of infinitely many components.

By proving that the supremum over the states in (75) reduces to a maximum, we could get a MUR formulation analogue of the one in Remark 7, but we skip this.

4.2.2 Spin 1/2.

By using the state representation (33) and the univariate measure (74), we can compute the relative entropies, as done in equations (52) and (53). Then, by taking the supremum over the states, we get

$$\Delta_{1/2}[\mathcal{A}_\mathtt{r}\|\mathsf{M}_c] = S\big(\mathsf{X}_i^{\rho_i}\big\|(\mathsf{M}_c)_{[i]}^{\rho_i}\big) = \log\frac{2}{1+c}, \qquad |c| \leq \frac{1}{\sqrt{\mathtt{r}}}, \quad \mathtt{r} = 2, 3. \tag{82}$$

Here, the measurement M_c is given by (72) for $\mathtt{r} = 3$ or by (73) for $\mathtt{r} = 2$, while the state ρ_i is anyone of the two eigen-projections of S_i.

By the definition (76) and the explicit expression (82), we obtain

$$I_{1/2}[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{A}_\mathtt{r})] = \inf_{c\in[-1/\sqrt{\mathtt{r}},1/\sqrt{\mathtt{r}}]} S\big(\mathsf{X}_i^{\rho_i}\big\|(\mathsf{M}_c)_{[i]}^{\rho_i}\big) = \log\frac{2}{1+1/\sqrt{\mathtt{r}}}\,. \tag{83}$$

Let us note that there is an optimal POVM, the one with $c = 1/\sqrt{\mathtt{r}}$, the same of the one appearing in [5, 16, 25], where different optimality criteria where used. By using this measurement it would be possible to give a state dependent version of the MURs as done in Remark 8.

4.2.3 The bounds from optimal cloning.

For $s > 1/2$ we can get a bound on the minimal information loss by using the POVM obtained from optimal cloning, because by construction we have

$$I_s[\mathcal{A}_\mathtt{r}\|\mathcal{M}(\mathcal{A}_\mathtt{r})] \leq \Delta_s[\mathcal{A}_\mathtt{r}\|\mathsf{M}^\mathtt{r}_{\rm cl}], \qquad \mathtt{r} = 2, 3. \tag{84}$$

Three orthogonal components. Let us set $p_m := \mathsf{X}^\rho(m)$; then, by (70) and (38), we get

$$\mathsf{M}^{3,\rho}_{\mathrm{cl}[1]}(m) = \frac{1+(s+2)\,p_m}{3(s+1)}, \qquad S\big(\mathsf{X}^\rho \big\| \mathsf{M}^{3,\rho}_{\mathrm{cl}[1]}\big) = \sum_{m=-s}^{s} p_m \log \frac{3\,(s+1)\,p_m}{1+(s+2)\,p_m}.$$

This gives the device information loss

$$\Delta_s[\mathcal{A}_3\|\mathsf{M}^3_{\mathrm{cl}}] = \sup_\rho S\big(\mathsf{X}^\rho \big\| \mathsf{M}^{3,\rho}_{\mathrm{cl}[1]}\big) = \log\frac{3(s+1)}{s+3}. \tag{85}$$

Two orthogonal spin components. By the same definition of p_m and using (71) instead of (70), in a similar way we get

$$\mathsf{M}^{2,\rho}_{\mathrm{cl}[1]}(m) = \frac{1+(2s+3)\,p_m}{4(s+1)}, \qquad S\big(\mathsf{X}^\rho \big\| \mathsf{M}^{2,\rho}_{\mathrm{cl}[1]}\big) = \sum_{m=-s}^{s} p_m \log \frac{4\,(s+1)\,p_m}{1+(2s+3)\,p_m},$$

$$\Delta_s[\mathcal{A}_2\|\mathsf{M}^2_{\mathrm{cl}}] = \sup_\rho S\big(\mathsf{X}^\rho \big\| \mathsf{M}^{2,\rho}_{\mathrm{cl}[1]}\big) = \log\frac{2(s+1)}{s+2}. \tag{86}$$

Note that the device information losses (85) and (86) grow with s and that they enjoy some unexpected relations, such as

$$\Delta_1[\mathcal{A}_2\|\mathsf{M}^2_{\mathrm{cl}}] > \Delta_{1/2}[\mathcal{A}_3\|\mathsf{M}^3_{\mathrm{cl}}], \qquad \Delta_2[\mathcal{A}_2\|\mathsf{M}^2_{\mathrm{cl}}] = \Delta_1[\mathcal{A}_3\|\mathsf{M}^3_{\mathrm{cl}}],$$

$$\lim_{s\to+\infty} \Delta_s[\mathcal{A}_2\|\mathsf{M}^2_{\mathrm{cl}}] = \Delta_3[\mathcal{A}_3\|\mathsf{M}^3_{\mathrm{cl}}].$$

For instance, the first relation says that, for the devices constructed by optimal cloning, the information loss for the case of two orthogonal components and $s=1$ is greater than the information loss for the case of three orthogonal components and $s=1/2$.

4.3 Some orderings and bounds

As we already said, the minimum information loss can be interpreted as a quantification of the incompatibility of the set of target observables. So, we can take the results obtained on $I_s[\mathcal{A}_{\mathtt{r}}\|\mathcal{M}(\mathcal{A}_{\mathtt{r}})]$, $\mathtt{r} = 2,3,\infty$, $s = 1/2, 1, 3/2, \ldots$, to compare different sets of spin observables (even in different Hilbert spaces) from the point of view of incompatibility; as we shall see, some non intuitive relations appear.

First of all we have the inequalities (67) in the case of all the components and small s; for the same s and different $\mathtt{r}$ we have the inequalities (80).

By the optimal cloning bound (84) and the growing with s of the expressions (85) and (86), we get the bounds

$$\begin{aligned} &I_s[\mathcal{A}_2\|\mathcal{M}(\mathcal{A}_2)] \leq 1, \qquad I_s[\mathcal{A}_3\|\mathcal{M}(\mathcal{A}_3)] \leq \log 3, \qquad s \geq \frac{1}{2}, \\ &I_s[\mathcal{A}_3\|\mathcal{M}(\mathcal{A}_3)] \leq 1, \qquad \frac{1}{2} \leq s \leq 3. \end{aligned} \tag{87}$$

By the bound (84) again, and the fact the we have the numerical value of $I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)]$ for $s = 1,\, 3/2$, see equations (55) and (63), we obtain

$$\begin{aligned} &I_s[\mathcal{A}_2\|\mathcal{M}(\mathcal{A}_2)] \leq I_1[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)] < 1, \qquad 1/2 \leq s \leq 3, \\ &I_s[\mathcal{A}_3\|\mathcal{M}(\mathcal{A}_3)] < I_1[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)], \qquad s = 1/2,\, 1, \\ &I_s[\mathcal{A}_2\|\mathcal{M}(\mathcal{A}_2)] < I_{3/2}[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)], \qquad 1/2 \leq s \leq 11, \\ &I_s[\mathcal{A}_3\|\mathcal{M}(\mathcal{A}_3)] < I_{3/2}[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)], \qquad 1/2 \leq s \leq 2. \end{aligned} \tag{88}$$

For instance, the second-last inequality says that two orthogonal components for $s = 11$ are less incompatible than the set of all components for $s = 3/2$; similar interpretations hold for the other inequalities.

5 Conclusions

The entropic formulation of MURs has the advantage of being well based on information theory (in particular on the notion of information loss) and independent of the measurement units of the observed physical quantities and from a reordering of their possible values [25–27]. By using the case of the spin components, in this article we have shown that the approach based on the relative entropy can be extended so to treat on the same foot finitely or infinitely many observables.

By introducing the worst information loss with respect to the target observables and the system states, we have defined the *device information loss* in the various cases (40), (75), (77). Then, by optimizing with respect to the approximating joint measurements we have defined the *minimum information loss* (44), (76), (78). These two quantities allow for a clear formulation of state independent MURs, see Sections 3.3 and 4.2.1.

To realize the minimum information loss one needs also to optimize the approximating measurement; an interesting point is that the "best" approximating measurement of a target spin observable is not necessarily a noisy version of the target, with additive classical noise, but most general noise structures can be involved, as discussed in Section 2.2.2 and in Theorem 6.

Moreover, the lower bound appearing in the state independent MURs, the minimum information loss, plays also the role of measure of incompatibility and allows to order different sets of target observables according to increasing incompatibility, as done in the inequalities (67), (80), (88).

However, the computations of the two "information losses" need to solve difficult optimization problems and we have done these computations only for small values of s, Sections 3.4, 3.5, 3.6, 4.2.2. To compute the minimum information loss for other values of the spin also numerical computations should be surely involved.

Another open problem is the conjecture given after inequality (67): is it true that the minimum information loss grows with s? For the cases of two and three orthogonal components we proved that the minimum information loss is upper bounded by a value independent from s, see (87). However, for the case of infinitely many components we proved only the existence of the upper bound (46), which grows with s; the problem of the asymptotic behaviour of $I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)]$ for large s is open.

As we remarked at the end of Section 3.3, the proof of MURs is independent of the choice of the class of approximating joint measurements. Anyway, the value of the minimum information loss can depend on this choice. So, another open problem is to study if the lower bound remains $I_s[\mathcal{A}_\infty\|\mathcal{M}(\mathcal{A}_\infty)]$ even when other classes of measurements are considered, different from $\mathcal{M}(\mathcal{A}_\infty)$.

A Spin s: rotations and q-coefficients

Let us consider the rotation group in $\mathbb{R}^3$: a counterclockwise rotation of the angle α around the unit vector $\boldsymbol{u}$ is denoted by

$$R_{\boldsymbol{u}}(\alpha) \in SO(3), \qquad |\boldsymbol{u}| = 1, \quad \alpha \in [0, 2\pi). \tag{89}$$

Then, we introduce the unitary representation of $SO(3)$ on $\mathcal{H} = \mathbb{C}^{2s+1}$, given by

$$U\big(R_{\boldsymbol{u}}(\alpha)\big) := \exp\left\{-\mathrm{i}\alpha\, \boldsymbol{u} \cdot \boldsymbol{S}\right\}. \tag{90}$$

Such a representation is an essential tool in our whole construction; this representation and its main properties can be found, e.g., in [28, Sect. 3.5], [20, Sect. 3.11].

By comparing equations (90) and (3), we have the identification

$$V(\theta, \phi) = U\big(R_{\boldsymbol{u}(\phi)}(\theta)\big), \qquad \boldsymbol{u}(\phi) = (-\sin\phi, \cos\phi, 0) = \boldsymbol{n}(\pi/2, \phi + \pi/2), \tag{91}$$

the unit vector $\boldsymbol{n}(\theta, \phi)$ is defined in (2). Moreover, the following decompositions hold:

$$U\big(R_{\boldsymbol{n}(\theta,\phi)}(\alpha)\big) = V(\theta, \phi) U(R_{\boldsymbol{k}}(\alpha)) V(\theta, \phi)^\dagger, \qquad V(\theta, \phi) = \mathrm{e}^{-\mathrm{i}\phi S_z} \mathrm{e}^{-\mathrm{i}\theta S_y} \mathrm{e}^{\mathrm{i}\phi S_z}. \tag{92}$$

A.1 Properties of the Wigner small-d-matrix

An explicit, but complicated, form of the Wigner small-d-matrix (21) has been obtained [28, (3.65)]; in particular, the explicit expressions for $s = 1/2,\ 1,\ 3/2,\ 2$ can be found in [38, Fig. 44.1][1]. Moreover, the following properties hold [28, (3.80)-(3.82), (3.125)-(3.126)]:

$$d^{(s)}_{m',m} = (-1)^{m-m'} d^{(s)}_{m,m'} = d^{(s)}_{-m,-m'}, \tag{93}$$

$$\sum_{m=-s}^{s} d^{(s)}_{m_1,m} d^{(s)}_{m_2,m} = \sum_{m=-s}^{s} d^{(s)}_{m,m_1} d^{(s)}_{m,m_2} = \delta_{m_1,m_2}. \tag{94}$$

From [28, (3.65)] one sees that the form of the matrix elements is sufficiently simple when one of the indices takes the maximal value and one gets

$$\left| d^{(s)}_{s,m}(\theta) \right|^2 = \frac{(2s)!}{(s+m)!\,(s-m)!} \left(\frac{1+x}{2}\right)^{s+m} \left(\frac{1-x}{2}\right)^{s-m}, \qquad x = \cos\theta; \tag{95}$$

we reported only the square modulus, because we need only this, see (23).

Finally, as one sees from [28, (3.72)], the quantity $\left| d^{(s)}_{\ell,m}(\theta) \right|^2$ is a polynomial in $\cos\theta$.

A.2 The q-coefficients

By using the expressions given in [38, Fig. 44.1] we can compute the q-coefficients in the cases $s = 1/2,\ 1,\ 3/2$.

[1] The table can be downloaded from http://pdg.lbl.gov/2019/reviews/rpp2018-rev-clebsch-gordan-coefs.pdf

A.2.1 Spin 1/2.

In this case we have $\left|d^{(1/2)}_{\ell,h}(\theta)\right|^2 = \frac{1}{2} + 2h\ell\cos\theta$. Then, from the definition (20) we obtain

$$q(m|\ell,h) = \frac{1}{2} + 2\ell hm; \tag{96}$$

we suppressed the index $\boldsymbol{\theta}$, because there is no arbitrariness in these indices, as recalled in Section 2.3.1. By (17), (22), we get (31).

A.2.2 Spin 1.

In this case we have

$$\left|d^{(1)}_{0,0}(\theta)\right|^2 = x^2, \qquad \left|d^{(1)}_{0,\pm1}(\theta)\right|^2 = \left|d^{(1)}_{\pm1,0}(\theta)\right|^2 = \frac{1-x^2}{2},$$
$$\left|d^{(1)}_{\pm1,1}(\theta)\right|^2 = \left|d^{(1)}_{\mp1,-1}(\theta)\right|^2 = \frac{(1\pm x)^2}{4}, \qquad x := \cos\theta.$$

From (20), by direct computations, we get the explicit expressions of the q-coefficients, with a given in (35); by using this parameter as index, instead of $\boldsymbol{\theta}$, we have

$$\begin{aligned}
&q_a(\pm1|1,\pm1) = q_a(\pm1|-1,\mp1) = 1 - \frac{(1+a)^3}{8},\\
&q_a(\mp1|1,\pm1) = q_a(\mp1|-1,\mp1) = \frac{(1-a)^3}{8},\\
&q_a(1|0,\pm1) = q_a(1|\pm1,0) = q_a(-1|0,\pm1) = q_a(-1|\pm1,0) = \frac{2+a}{4}\,(1-a)^2,\\
&q_a(\pm1|0,0) = \frac{1-a^3}{2}, \qquad q_a(0|0,\pm1) = q_a(0|\pm1,0) = \frac{a}{2}\left(3-a^2\right),\\
&q_a(0|0,0) = a^3, \qquad q_a(0|1,\pm1) = q_a(0|-1,\mp1) = \frac{a}{4}\left(3+a^2\right).
\end{aligned} \tag{97}$$

A.2.3 Spin 3/2.

In this case we have, with $x = \cos\theta$,

$$\left|d^{(3/2)}_{\pm3/2,3/2}(\theta)\right|^2 = \left|d^{(3/2)}_{\mp3/2,-3/2}(\theta)\right|^2 = \frac{(1\pm x)^2}{8},$$
$$\left|d^{(3/2)}_{\pm3/2,1/2}(\theta)\right|^2 = \left|d^{(3/2)}_{\mp3/2,-1/2}(\theta)\right|^2 = \left|d^{(3/2)}_{1/2,\pm3/2}(\theta)\right|^2 = \left|d^{(3/2)}_{-1/2,\mp3/2}(\theta)\right|^2 = \frac{3}{8}\,(1\pm x)\,(1-x^2),$$
$$\left|d^{(3/2)}_{\pm1/2,1/2}(\theta)\right|^2 = \left|d^{(3/2)}_{\mp1/2,-1/2}(\theta)\right|^2 = \frac{1\pm x}{8}\,(3x\mp1)^2.$$

From (20), by direct computations, we get the explicit expressions of the q-coefficients, with a given in Section 2.3.3; by using this parameter as index, instead of $\boldsymbol{\theta}$, we have

$$\begin{aligned}
&q_a(\pm3/2|\pm3/2,3/2) = \frac{1}{16}\left(15-4a-6a^2-4a^3-a^4\right),\\
&q_a(\pm3/2|\pm3/2,-3/2) = \frac{1}{16}\left(1-4a+6a^2-4a^3+a^4\right),\\
&q_a(\pm3/2|\pm3/2,1/2) = \frac{1}{16}\left(11-12a-6a^2+4a^3+3a^4\right),\\
&q_a(\pm3/2|\pm3/2,-1/2) = \frac{1}{16}\left(5-12a+6a^2+4a^3-3a^4\right),
\end{aligned} \tag{98}$$

$$q_a(\pm 3/2|\pm 1/2, 1/2) = \frac{1}{16}\left(7 - 4a + 10a^2 - 4a^3 - 9a^4\right),$$
$$q_a(\pm 3/2|\pm 1/2, -1/2) = \frac{1}{16}\left(9 - 4a - 10a^2 - 4a^3 + 9a^4\right), \tag{99}$$

$$q_a(\pm 1/2|\pm 3/2, 3/2) = \frac{a}{16}\left(4 + 6a + 4a^2 + a^3\right),$$
$$q_a(\pm 1/2|\pm 3/2, -3/2) = \frac{a}{16}\left(4 - 6a + 4a^2 - a^3\right),$$
$$q_a(\pm 1/2|\pm 3/2, 1/2) = \frac{a}{16}\left(12 + 6a - 4a^2 - 3a^3\right),$$
$$q_a(\pm 1/2|\pm 3/2, -1/2) = \frac{a}{16}\left(12 - 6a - 4a^2 + 3a^3\right), \tag{100}$$

$$q_a(\pm 1/2|\pm 1/2, 1/2) = \frac{a}{16}\left(4 - 10a + 4a^2 + 9a^3\right),$$
$$q_a(\pm 1/2|\pm 1/2, -1/2) = \frac{a}{16}\left(4 + 10a + 4a^2 - 9a^3\right); \tag{101}$$

the other coefficients are obtained by the symmetry properties (25).

B Orthogonal components

B.1 Three orthogonal components

The set of the three orthogonal spin components $\mathcal{A}_3 = \{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ is invariant under the action of the order 24 octahedron group $O \subset SO(3)$ [25, Appendix B.4], generated by the 90° rotations around the three coordinate axes: $S_O = \{R_{\boldsymbol{i}}(\pi/2),\ R_{\boldsymbol{j}}(\pi/2),\ R_{\boldsymbol{k}}(\pi/2)\}$. Let us denote the three generators of O by $g_1 = R_{\boldsymbol{i}}(\pi/2)$, $g_2 = R_{\boldsymbol{j}}(\pi/2)$, $g_3 = R_{\boldsymbol{k}}(\pi/2)$; then we have the covariance relations

$$\begin{aligned} U_{g_1}\mathsf{X}(x)U_{g_1}^\dagger &= \mathsf{X}(x), & U_{g_1}\mathsf{Y}(y)U_{g_1}^\dagger &= \mathsf{Z}(y), & U_{g_1}\mathsf{Z}(z)U_{g_1}^\dagger &= \mathsf{Y}(-z), \\ U_{g_2}\mathsf{X}(x)U_{g_2}^\dagger &= \mathsf{Z}(-x), & U_{g_2}\mathsf{Y}(y)U_{g_2}^\dagger &= \mathsf{Y}(y), & U_{g_2}\mathsf{Z}(z)U_{g_2}^\dagger &= \mathsf{X}(z), \\ U_{g_3}\mathsf{X}(x)U_{g_3}^\dagger &= \mathsf{Y}(x), & U_{g_3}\mathsf{Y}(y)U_{g_3}^\dagger &= \mathsf{X}(-y), & U_{g_3}\mathsf{Z}(z)U_{g_3}^\dagger &= \mathsf{Z}(z). \end{aligned} \tag{102}$$

Then, $\mathsf{M} \in \mathcal{M}(\mathcal{A}_3)$ is a POVM on $\mathcal{X}^3$ with the same covariance properties:

$$\begin{gathered} U_{g_1}\mathsf{M}(x,y,z)U_{g_1}^\dagger = \mathsf{M}(x,-z,y), \qquad U_{g_2}\mathsf{M}(x,y,z)U_{g_2}^\dagger = \mathsf{M}(z,y,-x), \\ U_{g_3}\mathsf{M}(x,y,z)U_{g_3}^\dagger = \mathsf{M}(-y,x,z). \end{gathered} \tag{103}$$

B.2 Two orthogonal components

Here the set of target observables is $\mathcal{A}_2 = \{\mathsf{X}, \mathsf{Y}\}$. Their symmetry group is the dihedral group $D_4 \subset SO(3)$, the order 8 group of the 90° rotations around the $\boldsymbol{k}$-axis, together with the 180° rotations around $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{n}_1 := \boldsymbol{n}(\pi/2,\ \pi/4)$, and $\boldsymbol{n}_2 := \boldsymbol{n}(\pi/2,\ 3\pi/4)$. Note that $D_4 \subset O$. The two rotations $S_{D_4} = \{R_{\boldsymbol{i}}(\pi), R_{\boldsymbol{n}_1}(\pi)\}$ generate D_4, as we have

$$\begin{gathered} R_{\boldsymbol{j}}(\pi) = R_{\boldsymbol{n}_1}(\pi)R_{\boldsymbol{i}}(\pi)R_{\boldsymbol{n}_1}(\pi), \qquad R_{\boldsymbol{n}_2}(\pi) = R_{\boldsymbol{i}}(\pi)R_{\boldsymbol{n}_1}(\pi)R_{\boldsymbol{i}}(\pi), \\ R_{\boldsymbol{k}}(\pi/2) = R_{\boldsymbol{n}_2}(\pi)R_{\boldsymbol{j}}(\pi). \end{gathered}$$

As discussed in [25, Appendix B.2], the covariance relations are: $\forall (x,y) \in \mathcal{X}^2$,

$$\begin{aligned} U\big(R_{\boldsymbol{i}}(\pi)\big)\mathsf{X}(x)U\big(R_{\boldsymbol{i}}(\pi)\big)^\dagger &= \mathsf{X}(x), & U\big(R_{\boldsymbol{i}}(\pi)\big)\mathsf{Y}(y)U\big(R_{\boldsymbol{i}}(\pi)\big)^\dagger &= \mathsf{Y}(-y), \\ U\big(R_{\boldsymbol{n}_1}(\pi)\big)\mathsf{X}(x)U\big(R_{\boldsymbol{n}_1}(\pi)\big)^\dagger &= \mathsf{Y}(x), & U\big(R_{\boldsymbol{n}_1}(\pi)\big)\mathsf{Y}(y)U\big(R_{\boldsymbol{n}_1}(\pi)\big)^\dagger &= \mathsf{X}(y). \end{aligned} \tag{104}$$

Then, $\mathsf{M} \in \mathcal{M}(\mathcal{A}_2)$ is a POVM on $\mathcal{X}^2$ with the same covariance properties:

$$\begin{aligned} U\big(R_{\boldsymbol{i}}(\pi)\big)\mathsf{M}(x,y)U\big(R_{\boldsymbol{i}}(\pi)\big)^\dagger &= \mathsf{M}(x,-y), \\ U\big(R_{\boldsymbol{n}_1}(\pi)\big)\mathsf{M}(x,y)U\big(R_{\boldsymbol{n}_1}(\pi)\big)^\dagger &= \mathsf{M}(y,x). \end{aligned} \tag{105}$$

B.3 Joint measurements from optimal cloning

A technique to construct good multi-observables approximating a set of incompatible observables is based on optimal cloning [29, 35, 36]; we already applied it to the context of MURs in [25]. Let us consider a system with Hilbert space $\mathcal{H}$, of dimension $\dim(\mathcal{H}) = d$, and let $\mathcal{S}(\mathcal{H})$ denote its state space; then, the optimal approximate $\mathtt{r}$-cloning channel is the map

$$\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}^{\otimes \mathtt{r}}), \qquad \Phi(\rho) = \frac{d!\mathtt{r}!}{(d+\mathtt{r}-1)!}\,\Pi_\mathtt{r}(\rho \otimes \mathbb{1}^{\otimes(\mathtt{r}-1)})\Pi_\mathtt{r},$$

where $\Pi_\mathtt{r}$ is the orthogonal projection of $\mathcal{H}^{\otimes \mathtt{r}}$ onto its symmetric subspace $\mathrm{Sym}(\mathcal{H}^{\otimes \mathtt{r}})$ [36]. Let $\{\mathsf{A}_1, \ldots, \mathsf{A}_\mathtt{r}\}$ be a set of observables, possibly incompatible; then, by using the adjoint channel we get the reasonably approximate multi-observable $\mathsf{M}_{\mathrm{cl}} = \Phi^*(\mathsf{A}_1 \otimes \cdots \otimes \mathsf{A}_\mathtt{r})$, whose marginals are given by [35]

$$\mathsf{M}_{\mathrm{cl}[h]}(x) = \lambda_{d,\mathtt{r}}\mathsf{A}_h(x) + (1-\lambda_{d,\mathtt{r}})\,\frac{\mathbb{1}}{d}, \qquad \lambda_{d,\mathtt{r}} = \frac{d+\mathtt{r}}{\mathtt{r}\,(d+1)}. \tag{106}$$

The multi-observable M_{cl} turns out to have the same symmetry properties of the set of observables $\{\mathsf{A}_1, \ldots, \mathsf{A}_\mathtt{r}\}$. Indeed, let U be a unitary operator on $\mathcal{H}$; by using the commutation property $U^{\otimes \mathtt{r}}\Pi_\mathtt{r} = \Pi_\mathtt{r}U^{\otimes \mathtt{r}}$, it is possible to prove the transformation rule

$$U\mathsf{M}_{\mathrm{cl}}(x_1, \ldots, x_\mathtt{r})U^\dagger = \Phi^*\big(U\mathsf{A}_1(x_1)U^\dagger \otimes \cdots \otimes U\mathsf{A}_\mathtt{r}(x_\mathtt{r})U^\dagger\big).$$

We shall use this construction for 2 or 3 orthogonal spin components; so, we have $d = 2s+1$ and $\mathtt{r} = 2, 3$. The property above implies immediately that $\Phi^*(\mathsf{X}, \mathsf{Y}, \mathsf{Z})$ satisfies the covariance properties (103) and $\Phi^*(\mathsf{X}, \mathsf{Y})$ the covariance properties (105).

References

[1] Ozawa M, *Uncertainty relations for joint measurements of noncommuting observables*, Phys. Lett. A **320** (2004) 367–374.

[2] Buscemi F, Hall M J W, Ozawa M, Wilde M M, *Noise and disturbance in quantum measurements: An information-theoretic approach*, Phys. Rev. Lett. **112** (2014) 050401.

[3] Coles P J, Furrer F, *State-dependent approach to entropic measurement-disturbance relations*, Phys. Lett. A **379** (2015) 105–112.

[4] Abbot A A, Branciard C, *Noise and disturbance of Qubit measurements: An information-theoretic characterisation*, Phys. Rev. A **94** (2016) 062110.

[5] Qin H H, Zhang T G, Jost L, Sun C P, Li-Jost X, Fei S M, *Uncertainties of genuinely incompatible triple measurements based on statistical distance*, Phys. Rev. A **99** (2019) 032107.

[6] Watanabe D, Narikiyo O, *Entropic approach to error-disturbance tradeoff in quantum measurements*, J. Phys. Soc. Jpn. **86** (2017) 095001.

[7] Busch P, Heinosaari T, *Approximate joint measurements of qubit observables*, Quantum Inf. Comp. **8** (2008) 797–818.

[8] Busch P, Lahti P, Werner R F, *Quantum root-mean-square error and measurement uncertainty relations*, Rev. Mod. Phys. **86** (2014) 1261–1281.

[9] Busch P, Lahti P, Werner R F, *Heisenberg uncertainty for qubit measurements*, Phys. Rev. A **89** (2014) 012129.

[10] Bullock T, Busch P, *Measurement uncertainty relations: characterizing optimal error bounds for qubits*, J. Phys. A: Math. Theor. **51** (2018) 283001.

[11] Demirel B, Sponar S, Abbott A A, Branciard C, Hasegawa Y, *Experimental test of an entropic measurement uncertainty relation for arbitrary qubit observables*, New J. Phys. **21** (2019) 013038.

[12] Reardon-Smith O, *Simplifying measurement uncertainty with quantum symmetries*, arXiv:1909.05055 (2019).

[13] Dammeier L, Schwonnek R, Werner R F, *Uncertainty relations for angular momentum*, New J. Phys. **17** (2015) 093046.

[14] Busch P, Lahti P, Pellonpää J-P, Ylinen K, *Quantum Measurement*, Theoretical and Mathematical Physics (Springer International Publishing, Switzerland, 2016).

[15] Coles P J, Berta M, Tomamichel M, Whener S, *Entropic uncertainty relations and their applications*, Rev. Mod. Phys. **89** (2017) 015002.

[16] Heinosaari T, Miyadera T, Ziman M, *An invitation to quantum incompatibility* J. Phys. A Math. Theor. **49** (2016) 123001.

[17] Werner R F, Farrelly T, *Uncertainty from Heisenberg to today*, Found. Phys. **49** (2019) 460–491.

[18] Maassen H, Uffink J B M, *Generalized entropic uncertainty relations*, Phys. Rev. Lett. **60** (1988) 1103–1106.

[19] Wehner S, Winter A, *Entropic uncertainty relations—A survey*, New J. Phys. **12** 2010 025009.

[20] Holevo A S, *Probabilistic and Statistical Aspects of Quantum Theory*; Quaderni della Normale; Edizioni della Normale: Pisa, Italy, 2011.

[21] Abbott A A, Alzieu P-L, Hall M J W, Branciard C, *Tight state-independent uncertainty relations for qubits*, Mathematics, **4** (2016) 8.

[22] Riccardi A, Macchiavello C, Maccone L, *Tight entropic uncertainty relations for systems with dimension three to five*, Phys. Rev. A **95** (2017) 032109.

[23] de Guise H, Maccone L, Sanders B C, Shukla N, *State-independent uncertainty relations*, Phys. Rev. A **98** (2018) 042121.

[24] A. Ketterer, O. Gühne, *Entropic uncertainty relations from quantum designs*, arXiv:1911.07533 (2019).

[25] Barchielli A, Gregoratti M, Toigo A, *Measurement uncertainty relations for discrete observables: Relative entropy formulation*, Commun. Math. Phys. (2018) pp. 1–52.

[26] Barchielli A, Gregoratti M, Toigo A, *Measurement uncertainty relations for position and momentum: Relative entropy formulation*, Entropy **19** (2017) 301.

[27] Barchielli A, Gregoratti M, *Uncertainty relations and information loss for spin 1/2 measurements*, (2018) arXiv:1805.03919, to appear in *Proceedings of the 38th International Conference on Infinite Dimension Analysis, Quantum Probability and Related Topics* (World Scientific).

[28] Biedenharn L and Louck J, *Angular Momentum in Quantum Physics* (Reading, MA: Addison-Wesley, 1981).

[29] Heinosaari T, Schultz J, Toigo A, Ziman M, *Maximally incompatible quantum observables*, Phys. Lett. A **378** (2014) 1695–1699.

[30] Yu S, Liu N, Li L, Oh C H, *Joint measurement of two unsharp observables of a qubit*, Phys. Rev. A **81** (2010) 062116.

[31] Yu S, Oh C H, *Quantum contextuality and joint measurement of three observables of a qubit*, arXiv:1312.6470 (2013).

[32] Vedral V, *The role of relative entropy in quantum information theory*, Rev. Mod. Phys. **74** (2002) 197–234.

[33] Burnham K P, Anderson D R, *Model Selection and Multi-Model Inference*, 2nd edition (Springer, New York, 2002).

[34] Cover T M, Thomas J A, *Elements of Information Theory*, 2nd edition (Wiley, Hoboken, New Jersey, 2006).

[35] Werner R F, *Optimal cloning of pure states*, Phys. Rev. A **58** (1998) 1827–1832.

[36] Keyl M, Werner R F, *Optimal cloning of pure states, testing single clones*, J. Math. Phys. **40** (1999) 3283–3299.

[37] Pedersen G K, *Analysis now* (Springer-Verlag, New York, 1989).

[38] Tanabashi M et al. (Particle Data Group), *Review of Particle Physics*, Phys. Rev. D 98 (2018) 030001.