Imprecise probability for non-commuting observables

2015 New J. Phys. 17 085005

(http://iopscience.iop.org/1367-2630/17/8/085005)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
9 January 2015

REVISED
21 May 2015

ACCEPTED FOR PUBLICATION
10 June 2015

PUBLISHED
13 August 2015

PAPER

Imprecise probability for non-commuting observables

Armen E Allahverdyan

Yerevan Physics Institute, Alikhanian Brothers Street 2, Yerevan 375036, Armenia

E-mail: armen.allahverdyan@gmail.com

Keywords: imprecise probability, quantum joint probability, quasi-probability

Abstract

It is known that non-commuting observables in quantum mechanics do not have joint probability. This statement refers to the precise (additive) probability model. I show that the joint distribution of any non-commuting pair of variables can be quantified via upper and lower probabilities, i.e. the joint probability is described by an interval instead of a number (imprecise probability). I propose transparent axioms from which the upper and lower probability operators follow. The imprecise probability depend on the non-commuting observables, is linear over the state (density matrix) and reverts to the usual expression for commuting observables.

1. Introduction

Non-commuting observables in quantum mechanics do not have a joint probability [1–5] (see appendix A.1 for a reminder). This is the departure point of quantum mechanics from classical probabilistic theories [6]; it lies in the core of all quantum oddities. There are various quasi-probabilities (e.g., Wigner function) which have features of joint probability for (loosely defined) semiclassical states [7–9, 13]. Quasi-probabilities do have two problems: (i) they (must) get negative for a class of quantum states, thereby preventing *any* probabilistic interpretation for them[1]. (ii) Even if the quasi-probability is positive on a certain state, it is not unique, i.e. there can be other (equally legitimate) quasi-probability that is positive (and has other expected features of probability) on this state, e.g. there are Wigner function, P-function, Terletsky–Margenau–Hill function etc [15, 16]. Despite of the drawbacks, quasi-probabilities do have many applications [7–12], since they still possess certain features of joint probability, e.g. they reproduce the marginals [2, 4, 7–9, 12]. In particular, there are applications in equilibrium quantum statistical mechanics, where the Wigner quasi-probability and the Terletsky–Margenau–Hill function [15, 16] are routinely employed for studying equilibrium relations in the semi-classical domain[2][17]. Applications in non-equilibrium quantum statistical mechanics are even more known, since whole chapters of open-system dynamics are written in terms of quasi-probabilities; see e.g. [10]. Another application of the Terletsky–Margenau–Hill quasi-probability for quantum non-equilibrium thermodynamics was proposed recently in the context of fluctuations of work and fluctuation theorems [12].

As a possible alternative to quasi-probabilities, one can relax the requirement that the sought joint probability correctly reproduces the marginals[3]. This is done when studying joint measurements of non-commuting variables [4, 5, 21–26]. Such measurements have to be approximate, since they operate on an arbitrary initial state [4, 5]. They produce positive probabilities for the measurement results, but it is not clear to which extent these probabilities are intrinsic [27], i.e. to which extent they characterize the system itself, and not

[1] Negative probabilities were not found to admit a direct physical meaning [14] (what can be less possible, than the impossible?). In certain cases what seemed to be a negative probability was later on found to be a local value of a physical quantity, i.e. physically meaningful, but not a probability [14]. Mathematical meaning of negative probability is discussed in [18, 19].

[2] One should stress here that the usage of quasi-probabilities in statistical mechanics is frequently implicit, but is nevertheless essential. For instance, the routine introduction of symmetrized correlators of non-commuting variables [17] implies an implicit choice of the underlying Terletsky–Margenau–Hill quasi-probability, because the symmetrized correlators are the 'real' correlators with respect to this quasi-probability. This point is seen in the standard quantum fluctuation–dissipation theorem [17].

[3] Employing instead the unbiasedness: the averages of the non-commuting quantities are reproduced correctly [22].

approximate measurements employed. Alternatively, one can consider two consecutive measurements of the non-commuting observables [28, 29]. These two-time probabilities do not (generally) qualify for the joint probability of the non-commuting observables; see appendix A.2.

It is assumed that the sought joint probability is linear over the state (density matrix). If this condition is skipped, there are positive probabilities that correctly reproduce marginals for non-commuting observables [20, 24], e.g. simply the product of two marginals [13]. However, they do not reduce to the usual form of the joint quantum probability for *commuting* observables[4]; hence their physical meaning is unclear [13].

The statement on the non-existence of joint probability concern the usual *precise and additive* probability. This is not the only model of uncertainty. It was recognized since early days of probability theory [49] that the probability need not be precise: instead of being a definite number, it can be a definite interval [51–54]; see [55] for an elementary introduction[5].

Instead of a precise probability for an event E, the measure of uncertainty is now an interval $[\underline{p}(E),\ \overline{p}(E)]$, where $0 \leqslant \underline{p}(E) \leqslant \overline{p}(E)$ are called lower and upper probabilities, respectively. Qualitatively, $\underline{p}(E)\ (1-\overline{p}(E))$ is a measure of a sure evidence in favor (against) of E. The event E is surely more probable than E', if $\underline{p}(E) \geqslant \overline{p}(E')$. The usual probability is recovered for $\underline{p}(E) = \overline{p}(E)$. Two different pairs $[\underline{p}(E),\ \overline{p}(E)]$ and $[\underline{p}'(E),\ \overline{p}'(E)]$ can hold simultaneously (i.e. they are consistent), provided that $\underline{p}'(E) \leqslant \underline{p}(E)$ and $\overline{p}'(E) \geqslant \overline{p}(E)$ for all E. In particular, every imprecise probability is consistent with[6] $\underline{p}'(E) = 0,\ \overline{p}'(E) = 1$.

It is not assumed that for all E there is a true (precise, but unknown) probability that lies in $[\underline{p}(E),\ \overline{p}(E)]$. This assumption is frequently (but not always [34]) made in applications[7] [52, 53], and it did motivate the generalized Kolmogorovian axiomatics of imprecise probability [54]; see appendix B.1. Imprecise joint probabilities in quantum mechanics are to be regarded as fundamental entities, not reducible to a lack of knowledge. They do need an independent axiomatic ground.

My purpose here is to propose a transparent set of conditions (axioms) that lead to quantum lower and upper joint probabilities. They depend only on the involved non-commuting observables (and on the quantum state).

The next section discusses previous attempts to introduce imprecise probability in quantum physics. Section 3 recalls standard linear algebra notations employed in this work. Section 4 describes physical conditions that are imposed on the sought imprecise probability. Section 5 outlines the main linear-algebra tool (CS-representation for projection operators) that is employed for finding the imprecise probability operators. Details of this representation are outlined in appendix C. Section 6 presents the main result: formulas for upper and lower probability operators. Their detailed derivation can be followed in appendix D. Several physical features of these operators are discussed in section 7 and also in appendices E and F. Section 8 discusses upper and lower probabilities for coordinate and momentum. I summarize in the last section. There I also mention several open problems related to this research.

2. Previous work

In 1967 Prugovecki tried to describe the joint probability of two non-commuting observables in a way that resembles imprecise probabilities [30]. But his expression was not correct, since it still can be negative [13]; cf footnote 1 and see also [18] in this context.

In 1991 Suppes and Zanotti proposed a local upper probability model for the standard setup of Bell inequalities (two entangled spins) [31]; see also [32, 33]. The formulation was given in the classical event space of hidden variables, and it is not unique even for the particular case considered. It violates classical observability conditions for the imprecise probability [31, 34, 54]. In particular, no lower probability exists in this scheme. Despite of such drawbacks, the pertinent message of [31] is that one should attempt at quantum applications of the upper probabilities that go beyond its classical axioms.

More recently, Galvan attempted to empoy (classical) imprecise probabilities for describing quantum dynamics in configuration space [37]. For a general discussion on quantum versus classical probabilities see [38].

[4] Given two projectors P and Q and state ρ, this product is $\mathrm{tr}(\rho P)\,\mathrm{tr}(\rho Q)$, while the correct form for $PQ = QP$ is $\mathrm{tr}(\rho PQ)$.

[5] Ellsberg's paradox is an example in psychology, where the ordinary probability theory does not apply, while imprecise probabilities can be used fruitfully for explaining experimental results on human decision making [50].

[6] This 'nothing is known' situation cannot be represented by usual probabilities, the simplest example showing that imprecise probabilities can model types of uncertainty that are not captured by the precise model.

[7] The assumption is legitimate in statistics, where one bounds the unknown (additive) probability via a finite number of observations [52]. It is not forbidden in subjective theories, where one aims at quantifying an uncertain human opinion via probabilities [53].

3. Notations

All operators (matrices) live in a finite-dimensional Hilbert space $\mathbb{H}$. For two Hermitean operators Y and Z, $Y \geqslant Z$ means that all eigenvalues of $Y - Z$ are non-negative, i.e. $\langle\psi|(Y-Z)\psi\rangle \geqslant 0$ for any $|\psi\rangle \in \mathbb{H}$. The direct sum $Y \oplus Z$ of two operators refers to the following block-diagonal matrix:

$$Y \oplus Z = \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix}.$$

$\mathrm{ran}(Y)$ is the range of Y (set of vectors $Y|\psi\rangle$, where $|\psi\rangle \in \mathbb{H}$). I is the unity operator of $\mathbb{H}$. $\ker(Y)$ is the subspace of vectors $|\phi\rangle$ with $Y|\phi\rangle = 0$.

I_n and 0_n are the $n \times n$ unity and zero matrices, respectively.

In the direct sum of two sub-spaces, $\mathbb{H} \oplus \mathbb{G}$ it is always understood that $\mathbb{H}$ and $\mathbb{G}$ are orthogonal. The vector sum of (not necessarily orthogonal) sub-spaces $\mathbb{A}$ and $\mathbb{B}$ will be denoted as $\mathbb{A} + \mathbb{B}$. This space is formed by all vectors $|\psi\rangle + |\phi\rangle$, where $|\psi\rangle \in \mathbb{A}$ and $|\phi\rangle \in \mathbb{B}$.

4. Axioms for quantum imprecise probability

Existing axioms for imprecise probability are formulated on a classical event space with usual notions of con- and disjunction and complemention [51, 51–54]; see appendix B for a reminder. For quantum probability it is natural to start from a Hilbert space and introduce upper and lower probabilities as operators. The axioms below require only the most basic feature of upper and lower probability and demand its consistency with the quantum joint probability whenever the latter is well-defined.

The usual quantum probability can be defined over (Hermitean) projectors $P = P^2$ [39, 40]. A projector generalizes the classical notion of characteristic function. Each P uniquely relates to its eigenspace $\mathrm{ran}(P)$. P refers to a set of Hermitean operators $\{\mathcal{P}\}$:

$$[\mathcal{P}, P] \equiv \mathcal{P}P - P\mathcal{P} = 0. \tag{1}$$

P is a projector to an eigenspace of $\mathcal{P}$ or to a direct sum of such eigenspaces, i.e. P refers to an eigenvalue of $\mathcal{P}$ or to a union of several eigenvalues. The quantum (precise and additive) probability to observe $P = 1$ is $\mathrm{tr}[\rho P]$, where the density matrix $0 \leqslant \rho \leqslant I$ defines the quantum state [4, 5, 39, 40].

Let Q be another projector which refers to the set $\{\mathcal{Q}\}$ of observables. Generally, $[P, Q] \neq 0$. Given the density matrix ρ, we seek upper and lower joint probabilities of P and Q (i.e. of the corresponding eigenvalues of $\mathcal{P}$ and $\mathcal{Q}$):

$$\overline{p}(\rho; P, Q) = \mathrm{tr}(\rho\, \overline{\omega}(P, Q)), \quad \underline{p}(\rho; P, Q) = \mathrm{tr}(\rho\, \underline{\omega}(P, Q)), \tag{2}$$

where $\underline{\omega}(P, Q)$ and $\overline{\omega}(P, Q)$ are Hermitean operators. Note that the upper and lower probabilities in (2) are assumed to have the usual Born's form, as far as their dependence on ρ is concerned.

We impose the following conditions (axioms):

$$0 \leqslant \underline{\omega}(P, Q) \leqslant \overline{\omega}(P, Q) \leqslant I, \tag{3}$$

$$\underline{\omega}(P, Q) = \underline{\omega}(Q, P), \quad \overline{\omega}(P, Q) = \overline{\omega}(Q, P), \tag{4}$$

$$\underline{\omega}(P, Q) = \overline{\omega}(P, Q) = PQ, \quad \text{if } [P, Q] = 0, \tag{5}$$

$$\mathrm{tr}(\rho\, \underline{\omega}(P, Q)) \leqslant \mathrm{tr}(\rho\, PQ) \leqslant \mathrm{tr}(\rho\, \overline{\omega}(P, Q)), \quad \text{if } [P, \rho] = 0, \quad \text{or if } [\rho, Q] = 0, \tag{6}$$

$$[\omega(P, Q), Q] = [\omega(P, Q), P] = 0, \quad \omega = \underline{\omega}, \overline{\omega}. \tag{7}$$

Equation (2) implies that $\underline{p}$ and $\overline{p}$ depend on $\{\mathcal{P}\}$ and $\{\mathcal{Q}\}$ only through P and Q. This non-contextuality feature holds also for the ordinary (one-variable) quantum probability [43, 44]. Provided that the operators $\underline{\omega}$ and $\overline{\omega}$ are found, $\underline{p}$ and $\overline{p}$ can be determined in the usual way of quantum averages.

Conditions (3) stem from $0 \leqslant \underline{p}(\rho; P, Q) \leqslant \overline{p}(\rho; P, Q) \leqslant 1$ that are demanded for all density matrices ρ. Equation (4) is the symmetry condition necessary for the joint probability. Equation (5) is reversion to the commuting case. In particular, (5) ensures $\underline{\omega}(P, 0) = \overline{\omega}(P, 0) = 0$ and

$$\underline{\omega}(P, I) = \overline{\omega}(P, I) = P. \tag{8}$$

Since $Q = I$ means that Q is anywhere, (8) is the reproduction of the marginal probability. The latter cannot be recovered by summation, since the very probability model is not additive.

For $[P, Q] = 0$ the joint probability is $\mathrm{tr}(\rho QP) = \mathrm{tr}(\rho PQ)$. This expression is well-defined (i.e. positive, symmetric and additive) also for $[\rho, Q] = 0$ or $[\rho, P] = 0$ (but not necessarily $[P, Q] = 0$). If $[\rho, Q] = 0$, one obtains $\mathrm{tr}(\rho QP)$ by measuring Q (ρ is not disturbed) and then P. Alternatively, one can obtain it by measuring

the average of an Hermitean observable $\frac{1}{2}(PQ + QP)$. Thus (5), (6) demands that $\overline{p}(\rho; P, Q)$ and $\underline{p}(\rho; P, Q)$ are consistent with the joint probability $\mathrm{tr}(\rho\, PQ)$, whenever the latter is well-defined.

Finally, equation (7) means that $\omega(P, Q)$ ($\omega = \underline{\omega}, \overline{\omega}$) can be measured *simultaneously and precisely* with P or with Q (on any quantum state), a natural condition for the joint probability (operators)[8].

If there are several candidates satisfying (3)–(7) we shall naturally select the ones providing the largest lower probability and the smallest upper probability.

5. CS-representation

This representation will be our main tool. Given the projectors P and Q, Hilbert space $\mathbb{H}$ can be represented as a direct sum [45–47] (see appendix C)

$$\mathbb{H} = \mathbb{H}' \oplus \mathbb{H}_{11} \oplus \mathbb{H}_{10} \oplus \mathbb{H}_{01} \oplus \mathbb{H}_{00}, \tag{9}$$

where the sub-space $\mathbb{H}_{\alpha\beta}$ of dimension $m_{\alpha\beta}$ is formed by common eigenvectors of P and Q having eigenvalue α (for P) and β (for Q). Depending on P and Q every sub-space can be absent; all of them can be present only for $\dim\mathbb{H} \geqslant 6$. Now $\mathbb{H}_{11} = \mathrm{ran}(P) \cap \mathrm{ran}(Q)$ is the intersection of the ranges of P and Q. $\mathbb{H}'$ has even dimension $2m$ [46, 47], this is the only sub-space in (9) that is not formed by common eigenvectors of P and Q. There exists a unitary transformation

$$P = U\hat{P}U^{\dagger}, \quad Q = U\hat{Q}U^{\dagger}, \quad UU^{\dagger} = I, \tag{10}$$

so that $\hat{P}$ and $\hat{Q}$ get the following block-diagonal form related to (9) [46]:

$$\hat{Q} = Q' \oplus I_{m_{11}} \oplus I_{m_{10}} \oplus 0_{m_{01}} \oplus 0_{m_{00}}, \quad Q' \equiv \begin{pmatrix} I_m & 0_m \\ 0_m & 0_m \end{pmatrix}, \tag{11}$$

$$\hat{P} = P' \oplus I_{m_{11}} \oplus 0_{m_{10}} \oplus I_{m_{01}} \oplus 0_{m_{00}}, \quad P' \equiv \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}, \tag{12}$$

where C and S are invertible square matrices of the same size holding

$$C^2 + S^2 = I_m, \quad [C, S] = 0. \tag{13}$$

Now $\mathrm{ran}(P')$ and $\mathrm{ran}(Q')$ are sub-spaces of $\mathbb{H}'$. One has $C = \cos T$ and $S = \sin T$, where T is the operator analogue of the angle between two spaces. $\mathbb{H}_{m_{\alpha\beta}}$ are absent, if P and Q do not have any common eigenvector. This, in particular, happens in $\dim(\mathbb{H}) = 2$.

6. The main result

Note that if (3)–(7) holds for P and Q, they hold as well for $\hat{P}$ and $\hat{Q}$, because $\omega(P, Q) = U^{\dagger}\omega(\hat{P}, \hat{Q})U$ for $\omega = \underline{\omega}, \overline{\omega}$. Appendix D shows how to get $\overline{\omega}(\hat{P}, \hat{Q})$ and $\underline{\omega}(\hat{P}, \hat{Q})$ from (3)–(7) and (11), (12):

$$\overline{\omega}(\hat{P}, \hat{Q}) = \begin{pmatrix} C^2 & 0_m \\ 0_m & C^2 \end{pmatrix} \oplus I_{m_{11}} \oplus 0_{m_{10}+m_{01}+m_{00}}, \tag{14}$$

$$\underline{\omega}(\hat{P}, \hat{Q}) = 0_{2m} \oplus I_{m_{11}} \oplus 0_{m_{10}+m_{01}+m_{00}}. \tag{15}$$

Let $g(P, Q) = g(Q, P)$ be the projector onto intersection $\mathrm{ran}(P) \cap \mathrm{ran}(Q)$ of $\mathrm{ran}(P)$ and $\mathrm{ran}(Q)$. We now return from (10), (14) and (15) to original projectors P and Q (see appendix D) and obtain the main formulas:

$$\underline{\omega}(P, Q) = g(P, Q), \tag{16}$$

$$\overline{\omega}(P, Q) = I - (P - Q)^2 - g(I - P, I - Q). \tag{17}$$

For $[P, Q] = 0$, $g(P, Q) = PQ$, and we revert to $\underline{\omega}(P, Q) = \overline{\omega}(P, Q) = PQ$. Note that $[P, (P - Q)^2] = [Q, (P - Q)^2] = 0$.

7. Physical meaning of upper and lower probability operators

When looking for a joint probability defined over two projectors P and Q one wonders whether it is just not some (operator) mean of P and Q. For ordinary numbers $a \geqslant 0$ and $b \geqslant 0$ there are three means: arithmetic

[8] Without condition (7), I was not able to fix the upper probability operator, i.e. without (7) there are many operator candidates that are not consistent with each other, i.e. not related by operator analogues of larger or smaller.

$\frac{a+b}{2}$, geometric $\sqrt{ab}$ and harmonic $\frac{2ab}{a+b}$. Now (16) is precisely the operator harmonic mean of P and Q [57]

$$g(P,\,Q) = 2P(P+Q)^{-}\,Q = 2Q(P+Q)^{-}\,P, \tag{18}$$

where A^{-} is the inverse of A if it exists, otherwise it is the pseudo-inverse; see appendix E for various representations of $\underline{\omega}(P,\,Q)$ and $\overline{\omega}(P,\,Q)$. More familiar formula is

$$g(P,\,Q) = \lim_{n\to\infty} Q(PQ)^{n} = \lim_{n\to\infty} P(QP)^{n}. \tag{19}$$

The intersection projector $g(P,\,Q)$ appears in [39–43]. It was stressed that $g(P,\,Q)$ cannot be a joint probability for non-commutative P and Q [21]. Its meaning is clear by now: it is the lower probability for P and Q. Note that

$$g(P,\,Q) = 0, \quad \text{if} \quad [P,\,Q] \neq 0 \quad \text{and} \quad \mathrm{tr}(P) = \mathrm{tr}(Q) = 1, \tag{20}$$

since two different rays (P and Q) cross only at zero. Thus, $g(P,\,Q)$ is non-zero for $[P,\,Q] \neq 0$, only if $\mathrm{tr}(P) \geqslant 2$ (or $\mathrm{tr}(Q) \geqslant 2$). I consider this as a natural features of the quantum lower probability, because the classical case—where the lower probability is expected to be non-zero and close to the upper probability—can be generically reached due to the coarse-graining, i.e. due to $\mathrm{tr}(P)$ (or $\mathrm{tr}(Q)$, or both) being sufficiently larger than 1.

Let us now turn to $\overline{\omega}$. The transition probability between two pure states is determined by the squared cosine of the angle between them: $|\langle\psi|\phi\rangle|^2 = \cos^2\theta_{\phi\psi}$. Equation (14) shows that $\overline{\omega}(P,\,Q)$ depends on $C^2 = \cos^2 T$, where T is the operator angle between $\hat{P}$ and $\hat{Q}$. Note from (11), (12) that the eigenvalues λ of PQ, which hold $0 < \lambda < 1$ are the eigenvalues of C^2, and | as seen from (14)—they are also (doubly-degenerate) eigenvalues of $\overline{\omega}(P,\,Q)$. Thus we have a physical interpretation not only for $\mathrm{tr}(PQ)$ (transition probability), but also for eigenvalues of PQ (PQ and QP have the same eigenvalues).

Equations (10), (14) and (15) imply that the upper and lower probability operators can be measured simultaneously on any state (cf (7)):

$$[\underline{\omega}(P,\,Q),\,\overline{\omega}(P,\,Q)] = 0. \tag{21}$$

The operator $\overline{\omega}(P,\,Q) - \underline{\omega}(P,\,Q)$ quantifies the uncertainty for joint probability, the physical meaning of this characteristics of non-commutativity is new.

Appendix G calculates the upper and lower probabilities for several examples.

Note that the conditional (upper and lower) probabilities are straighforward to define, e.g. (cf (2)): $\overline{p}(\rho;\,P|Q) = \overline{p}(\rho;\,P,\,Q)/\mathrm{tr}(\rho Q)$.

The distance between two probability intervals $[\underline{p},\,\overline{p}]$ and $[\underline{p}',\,\overline{p}']$ can be calculated via the Haussdorff metric [56]

$$\max\Big[\big|\underline{p} - \underline{p}'\big|,\,|\overline{p} - \overline{p}'|\Big], \tag{22}$$

which nullifies if and only if $\underline{p} = \underline{p}'$ and $\overline{p} = \overline{p}'$, and which reduces to the ordinary distance $|p - p'|$ for usual (precise) probabilities.

Let us see when we can use the notion of 'surely more probable'. Now

$$\mathrm{tr}(\rho\ \underline{\omega}(P_1,\,Q_1)) > \mathrm{tr}(\rho\ \overline{\omega}(P,\,Q)), \tag{23}$$

means that the pair of projectors $(P_1,\,Q_1)$ is surely more probable (on ρ) than $(P,\,Q)$; see appendix G for examples. Note from (16), (17) that if

$$\mathrm{tr}(\rho\ \omega(P,\,Q)) > \mathrm{tr}(\rho\ \omega(I-P,\,I-Q)), \tag{24}$$

holds for $\omega = \underline{\omega}$, then it also hods for $\omega = \overline{\omega}$ (and vice versa). Though in a weaker sense than (23), (24) means that P and Q together is more probable than neither of them together (which is the pair $(I-P,\,I-Q)$). equations (23), (24) are examples of comparative (modal) probability statements; see [36] in this context.

Further features of $\underline{\omega}$ and $\overline{\omega}$ are uncovered when looking at a monotonic change of their arguments; see appendix F. Appendix G discusses concrete examples that illustrate these features. Yet another example is provided below.

8. Coordinate and momentum

Coordinate x and momentum p operators, $[x,\,p] = \mathrm{i}$ ($\hbar = 1$) is the most known example of non-commutativity in quantum mechanics. Hence I shall illustrate the upper and lower probability for this example. In the (one-dimensional) x-representation, x-operator amounts to multiplication, while $p = -\mathrm{i}\dfrac{\mathrm{d}}{\mathrm{d}x}$. For intervals $x \in X = (X_1,\,X_2)$ and $p \in Y = (Y_1,\,Y_2)$ the corresponding projectors read in the coordinate representation

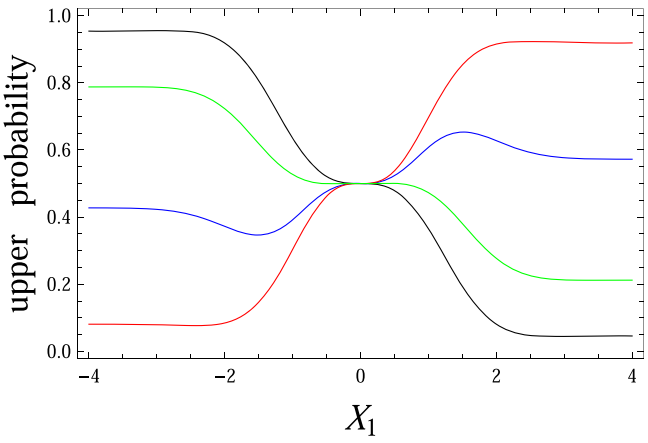

Figure 1. The upper probability $\langle\phi_1|\overline{\omega}\,(Q_X,\,P_Y)|\phi_1\rangle$, where $\overline{\omega}\,(Q_X,\,P_Y)$ is given by (28) with $X=(X_1,\,\infty)$ and $Y=[-\,0.5,\,0.5]$ (red curve), $Y=[-\,1,\,1]$ (blue curve), $Y=[-\,1.5,\,1.5]$ (green curve), $Y=[-\,2,\,2]$ (black curve). Here $\phi_1(x)=(4/\pi)^{1/4}xe^{-x^2/2}$ is the wave-function of the first excited level for the harmonic oscillator with Hamiltonian $H=(p^2+x^2)/2$. This example of ϕ_1 is chosen, because the Wigner function for the excited states of the harmonic oscillator is negative [8] and thus cannot serve for probabilistic reasoning. Note that for small values of X_1 (left plateau on the figure), $\langle\phi_1|\overline{\omega}\,(Q_X,\,P_Y)|\phi_1\rangle$ tends to $\langle\phi_1|P_Y|\phi_1\rangle$, while for large values of X_1 it tends to $1-\langle\phi_1|P_Y|\phi_1\rangle$. For the eigen-functions of $H=(p^2+x^2)/2$, the (marginal) distributions of the coordinate and momentum are equal, e.g. $\langle\phi_1|P_X|\phi_1\rangle=\langle\phi_1|Q_X|\phi_1\rangle$.

$$Q_X(x,\,x')=\delta(x-x')\chi_X(x),\qquad P_Y(x,\,x')=\frac{1}{2\pi}\int_{Y_1}^{Y_2}\mathrm{d}y\;\mathrm{e}^{\mathrm{i}y(x-x')}, \tag{25}$$

where $\chi_X(x)=\theta(x-X_1)\theta(X_2-x)$ is the characteristic function of interval X. Recall that Q_Y and P_Y are linked via the Fourier transform:

$$P_Y=\mathcal{F}^{+}Q_Y\mathcal{F}^{-},\quad [\mathcal{F}^{\pm}\phi](y)\equiv\frac{1}{\sqrt{2\pi}}\int\mathrm{d}x\;\mathrm{e}^{\pm\mathrm{i}xy}\phi(y). \tag{26}$$

The first thing to note is that if X and Y are finite intervals, then $g(Q_X,\,P_Y)=0$, i.e. $P_Y\phi=\phi$ and $Q_X\phi=\phi$ lead to $\phi=0$. This is a well-known result in the Fourier analysis; see, e.g. [61–63]. The simplest way to show it is to note (from (26)) that $Q_Y\mathcal{F}^{-}\phi$ has a finite support, hence $\mathcal{F}^{+}Q_Y\mathcal{F}^{-}\phi=\phi$ is analytic. On the other hand, ϕ should have a finite support. Thus $\phi=0$. This argument extends to the case, where (say) Y is semi-infinite, e.g. $Y=(Y_1,\,\infty)$, while X differs from $(-\infty,\,\infty)$ by a finite (or semi-infinite) interval [63]. Indeed, now $[P_Y\phi](x)=\frac{1}{\sqrt{2\pi}}\int\mathrm{d}\xi\;\mathrm{e}^{\mathrm{i}x\xi}[Q_Y\mathcal{F}^{-}\phi](\xi)$ is analytic for $\mathrm{Im}\;x>0$, while from $Q_X\phi=\phi$ it follows that $\phi(x)$ is zero in a finite interval at least.

Thus, for finite (or at least one semi-infinite) intervals X and Y the lower probability for the joint distribution of the coordinate and momentum is zero (cf (16))

$$\underline{\omega}\,(Q_X,\,P_Y)=g\,(Q_X,\,P_Y)=0. \tag{27}$$

However, if both X and Y are finite intervals, $g\,(I-Q_X,\,I-P_Y)\neq 0$, e.g. the above analiticity argument does not work. Moreover, $g\,(I-Q_X,\,I-P_Y)$ has a discrete spectrum, and its range is infinite-dimensional [61, 62]. We shall avoid this complication by looking at those case, where (at least) one of X and Y is semi-infinite. Then (27) still holds, while the upper probability operator $\overline{\omega}\,(Q_X,\,P_Y)$ reduces to (cf (17))

$$\overline{\omega}\,(Q_X,\,P_Y)=I-(Q_X-P_Y)^2, \tag{28}$$

and is straightforward to calculate via (25). Several examples of the upper probability calculated from (28) are presented in figure 1.

9. Summary and open problems

The main message of this work is that while joint precise probability for non-commuting observables does not exist, there are well-defined operator expressions for upper and lower imprecise probabilities. They are not additive, but otherwise they do satisfy a number of reasonable conditions: positivity, reproduction of correct marginals, direct observability via quantum averages, consistency with the (effectively) commuting case, where the joint probability is well-defined etc.

Several open questions are suggested by this research. First of all, it is not clear what is the suitable way of defining averages over the imprecise probability. This would be necessary for defining various correlation

functions. Recall that the average $\langle X\rangle$ of a random variable X that has a precise probability is defined via two conditions (see e.g. [60]): linearity, $\langle aX + bY\rangle = a\langle X\rangle + b\langle Y\rangle$, and monotonicity, $X \leqslant Y$ implies $\langle X\rangle \leqslant \langle Y\rangle$. Presumably, these conditions are to be modified for imprecise probability; in addition the sought average should reduce to the usual one when averaging over a single observable. This question should be clarified before the imprecise probability can be efficiently applied in quantum statistical mechanics.

Another open issue relates to the point (see (6)) that whenever the joint probability $\mathrm{tr}(\rho PQ)$ for non-commuting projectors P and Q is well defined due to e.g. $[\rho,\ P] = 0$ (see the discussion after (8)), the upper and lower probabilities $\overline{p}\,(\rho;\ P,\ Q)$ and $\underline{p}\,(\rho;\ P,\ Q)$ are merely consistent with the exact probability $\mathrm{tr}(\rho PQ)$, but are not equal to it, which would be a more desired outcome. It is thus not completely clear whether the found imprecise probabilities cannot be made more precise by looking at more general conditions (axioms), e.g. those that involve a nonlinear dependence on the density matrix ρ; cf (2). Such a dependence might however impede the direct observability of imprecise probabilities; the resulting issues need further investigations.

In a more remote perspective, one can ask about the joint imprecise probability of 3 (and more) non-commuting observables. In contrast to the previous two open problems, where the progress looks to be feasible, this is a difficult problem, because no analogue of the CS-representation for 3 (or more) non-commuting projectors seems to exists; see however [64].

Acknowledgments

I thank KV Hovhannisyan for discussions. I was supported by COST network MP1209.

Appendices

All 6 appendices can be read independently from each other.

Appendices A–C recall, respectively, the no-go statements for the joint quantum probability, generalized axiomatics for the imprecise probability and the CS-representation. This material is not new, but is presented in a focused form, adapted from several different sources.

Appendix D contains the derivation of the main result, while appendices E and F demonstrate various feature of quantum imprecise probability.

Appendix G illustrates it with simple physical examples.

Appendix A. Non-existence of (precise) joint probability for non-commuting observables

A.1. The basic argument

Given two sets of non-commuting Hermitean projectors:

$$\sum_{k=1}^{n_P} P_k = I, \quad P_k P_i = \delta_{ik} P_k, \quad n_P \leqslant n, \tag{A1}$$

$$\sum_{k=1}^{n_Q} Q_k = I, \quad Q_k Q_i = \delta_{ik} Q_k, \quad n_Q \leqslant n, \tag{A2}$$

we are looking for non-negative operators $\Pi_{ik} \geqslant 0$ such that for an arbitrary density matrix ρ

$$\sum_{ik} \mathrm{tr}(\rho\Pi_{ik}) = 1, \quad \sum_{i} \mathrm{tr}(\rho\Pi_{ik}) = \mathrm{tr}(\rho P_k), \quad \sum_{k} \mathrm{tr}(\rho\Pi_{ik}) = \mathrm{tr}(\rho Q_i). \tag{A3}$$

These relations imply

$$\sum_{ik} \Pi_{ik} = I, \quad \Pi_{ik} \leqslant Q_i, \quad \Pi_{ik} \leqslant P_k. \tag{A4}$$

Now the second (third) relation in (A4) implies $\mathrm{ran}(\Pi_{ik}) \subseteq \mathrm{ran}(Q_i)$ ($\mathrm{ran}(\Pi_{ik}) \subseteq \mathrm{ran}(P_k)$). Hence $\mathrm{ran}(\Pi_{ik}) \subseteq \mathrm{ran}(Q_i) \cap \mathrm{ran}(P_k)$.

Thus, if $\mathrm{ran}(Q_i) \cap \mathrm{ran}(P_k) = 0$ (e.g. when P_k and Q_i are one-dimensional (1D)), then $\Pi_{ik} = 0$, which means that the sought joint probability does not exist.

If $\mathrm{ran}(Q_i) \cap \mathrm{ran}(P_k) \neq 0$, then the largest Π_{ik} that holds the second and third relation in (A4) is the projection $g(P_k,\ Q_i)$ on $\mathrm{ran}(Q_i) \cap \mathrm{ran}(P_k) = 0$. However, the first relation in (A4) is still impossible to satisfy (for $[P_i,\ Q_k] \neq 0$), as seen from the superadditivity feature (F1):

$$\sum_{ik} g(P_i, Q_k) \leqslant \sum_k g\left(\sum_i P_i, Q_k\right) = \sum_k g(I, Q_k) = \sum_k Q_k = I. \tag{A5}$$

A.2. Two-time probability (as a candidate for the joint probability)

Given (A1), (A2), we can carry out two successive measurements. First (second) we measure a quantity, whose eigen-projections are $\{P_k\}$ ($\{Q_i\}$). This results to the following joint probability for the measurement results (ρ is the density matrix)

$$\mathrm{tr}(\ Q_i P_k \rho P_k\ ). \tag{A6}$$

Likewise, if we first measure $\{Q_i\}$ and then $\{P_k\}$, we obtain a quantity that generally differs from (A6):

$$\mathrm{tr}(\ P_k Q_i \rho Q_i\ ). \tag{A7}$$

If we attempt to consider (A7) (or (A6)) as a joint additive probability for P_i and Q_k, we note that (A7) (and likewise (A6)) reproduces correctly only one marginal:

$$\sum_{i=1}^{n_Q} \mathrm{tr}(Q_i P_k \rho P_k) = \mathrm{tr}(P_k \rho), \quad \text{but} \quad \sum_{i=1}^{n_P} \mathrm{tr}(Q_i P_k \rho P_k) \neq \mathrm{tr}(Q_i \rho). \tag{A8}$$

One can attempt to interpret the mean of (A6), (A7)

$$\mu(\rho; P_k, Q_i) = \frac{1}{2}[\mathrm{tr}(P_k Q_i \rho Q_i) + \mathrm{tr}(Q_i P_k \rho P_k)] = \mathrm{tr}\left(\rho \frac{P_k Q_i P_k + Q_i P_k Q_i}{2}\right), \tag{A9}$$

as a non-additive probability. This object is linear over ρ, symmetric (with respect to interchanging P_k and Q_i), non-negative, and reduces to the additive joint probability for $[P_k, Q_i] = 0$. The relation $\mu(\rho; P_k, I) = \mathrm{tr}(\rho P_k)$ can be interpreted as consistency with the correct marginals (once $\mu(\rho; P_k, Q_i)$ is regarded as a non-additive probability, there is no point in insisting that the marginals are obtained in the additive way).

However, the additive joint probability $\mathrm{tr}(\rho P_k Q_i)$ is well-defined also for $[\rho, P_k] = 0$ (or for $[\rho, Q_i] = 0$). If $[\rho, P_k] = 0$ holds, $\mu(\rho; P_k, Q_i)$ is not consistent with $\mathrm{tr}(\rho P_k Q_i)$, i.e. depending on ρ, P_k and Q_i both

$$\mu(\rho; P_k, Q_i) > \mathrm{tr}(\rho P_k Q_i) \quad \text{and} \quad \mu(\rho; P_k, Q_i) < \mathrm{tr}(\rho P_k Q_i) \tag{A10}$$

are possible.

To summarize, the two-time measurement results do not qualify as the additive joint probability, first because they are not unique (two different expressions (A6) and (A7) are possible), and second because they do not reproduce the correct marginals. If we take the mean of two expressions (A6) and (A7) and attempt to interpret it as a non-additive probability, it is not compatible with the joint probability, whenever the latter is well-defined.

Appendix B. Axioms for classical imprecise probability

B.1. Generalized Kolmogorov's axioms

Given the full set of events Ω, $\overline{p}(.)$ and $\underline{p}(.)$ defined over sub-sets A, B, ... of Ω (including the empty set $\{0\}$) satisfy [52–54]:

$$\underline{p}(\{0\}) = 0, \tag{B1}$$

$$\overline{p}(\Omega) = 1, \tag{B2}$$

$$\overline{p}(A) = 1 - \underline{p}(\Omega - A), \tag{B3}$$

$$\underline{p}(A \bigcup B) \geqslant \underline{p}(A) + \underline{p}(B), \quad \text{if } A \bigcap B = \{0\}, \tag{B4}$$

$$\overline{p}(A \bigcup B) \leqslant \overline{p}(A) + \overline{p}(B), \quad \text{if } A \bigcap B = \{0\}, \tag{B5}$$

where $\Omega - A$ includes all elements of Ω that are not in A, and where $A \bigcap B$ means intersection of two sets; $A \bigcap B = \{0\}$ holds for elementary events.

Here are some direct implications of (B1)–(B5)

$$A \supseteq B \implies \overline{p}(A) \geqslant \overline{p}(B), \tag{B6}$$

$$\underline{p}(A) \geqslant \underline{p}(B). \tag{B7}$$

Equation (B7) follows directly from (B4). Equation (B6) follows from (B4), (B3). Next relation:

$$\overline{p}(A \bigcup B) \geqslant \overline{p}(A) + \underline{p}(B) \geqslant \underline{p}(A \bigcup B), \quad \text{if } A \bigcap B = 0, \tag{B8}$$

which, in particular, implies

$$\overline{p}(A) \geqslant \underline{p}(A). \tag{B9}$$

To derive (B8), note that (B4), (B3) imply $\overline{p}(\Omega - A - B) \leqslant \overline{p}(\Omega - A) - \underline{p}(B)$ or $\overline{p}(\Omega - A) \geqslant \underline{p}(B) + \overline{p}(\Omega - A - B)$, which is the first inequality in (B8). The second inequality is derived via (B5), (B3).

The following inequality generalizes the known relation of the additive probability theory

$$\underline{p}(A) + \underline{p}(B) \leqslant \overline{p}(A \bigcup B) + \underline{p}(A \bigcap B). \tag{B10}$$

To prove (B10), we denote $A' = A - A \cap B$, which means $A' \cap B = \{0\}$. Now

$$\overline{p}(A \bigcup B) + \underline{p}(A \bigcap B) = \overline{p}(A' \bigcup B) + \underline{p}(A \bigcap B)$$

$$\geqslant \overline{p}(A') + \underline{p}(A \bigcap B) + \underline{p}(B) \tag{B11}$$

$$\geqslant \underline{p}(A) + \underline{p}(B), \tag{B12}$$

where in (B11) (resp. in (B12)) we applied the first (resp. the second) inequality in (B8).

Note that the (non-negative) difference $\Delta p(A) = \overline{p}(A) - \underline{p}(A)$ between the upper and lower probabilities also holds the super-additivity feature (cd (B5))

$$\Delta p(A \bigcup B) \leqslant \Delta p(A) + \Delta p(B), \quad \text{if } A \bigcap B = 0. \tag{B13}$$

Employing (B8) one can derive [58] for arbitrary A_1 and A_2:

$$\underline{p}(A_1 \bigcup A_2) + \underline{p}(A_1 \bigcap A_2) \leqslant \underline{p}(A_1) + \overline{p}(A_2) \leqslant \overline{p}(A_1 \bigcup A_2) + \overline{p}(A_1 \bigcap A_2), \tag{B14}$$

$$\underline{p}(A_1) + \underline{p}(A_2) \leqslant \underline{p}(A_1 \bigcup A_2) + \overline{p}(A_1 \bigcap A_2) \leqslant \overline{p}(A_1) + \overline{p}(A_2), \tag{B15}$$

$$\underline{p}(A_1) + \underline{p}(A_2) \leqslant \overline{p}(A_1 \bigcup A_2) + \underline{p}(A_1 \bigcap A_2) \leqslant \overline{p}(A_1) + \overline{p}(A_2). \tag{B16}$$

B.2. Joint probability

The joint probabilities of A and B are now defined as

$$\underline{p}(A, B) = \underline{p}(A \bigcap B), \quad \overline{p}(A, B) = \overline{p}(A \bigcap B). \tag{B17}$$

Employing the distributivity feature

$$(A_1 \bigcup A_2) \bigcap A_3 = (A_1 \bigcap A_3) \bigcup (A_2 \bigcap A_3), \tag{B18}$$

which holds for any triple A_1, A_2, A_3, we obtain from (B4), (B5) for $B \cap C = \{0\}$

$$\underline{p}(A, B \bigcup C) = \underline{p}((A \bigcap B) \bigcup (A \bigcap C)) \geqslant \underline{p}(A, B) + \underline{p}(A, C), \tag{B19}$$

$$\overline{p}(A, B \bigcup C) = \overline{p}((A \bigcap B) \bigcup (A \bigcap C)) \leqslant \overline{p}(A, B) + \overline{p}(A, C). \tag{B20}$$

B.3. Dominated upper and lower probability

The origin of (B1)–(B5) can be related to the *simplest* scheme of hidden variable(s) [52]. One imagines that there exists a precise probability $P_\theta(A)$, where the parameter θ is not known. Only the extremal values over the parameter are known:

$$\overline{p}(A) = \max_\theta[P_\theta(A)], \quad \underline{p}(A) = \min_\theta[P_\theta(A)], \tag{B21}$$

which satisfy (B1)–(B5).

However, it is generally not true that (B1)–(B5) imply the existence of a precise probability $P_\theta(A)$ that holds (B21) [53].

Appendix C. Derivation of the CS-representation

C.1. The main theorem

Let $\mathbb{Q}'$ and $\mathbb{P}'$ are two subspaces of Hilbert space $\mathbb{H}'$ that hold ($\mathbb{Q}'^{\perp}$ is the orthogonal complement of $\mathbb{Q}'$)

$$\mathbb{Q}' \bigcap \mathbb{P}' = 0, \quad \mathbb{Q}' \bigcap \mathbb{P}'^{\perp} = 0, \quad \mathbb{Q}'^{\perp} \bigcap \mathbb{P}' = 0, \quad \mathbb{Q}'^{\perp} \bigcap \mathbb{P}'^{\perp} = 0. \tag{C1}$$

The simplest example realizing (C1) is when $\mathbb{Q}'$ and $\mathbb{P}'$ are 1D subspaces of a two-dimensional (2D) $\mathbb{H}'$.

Let $\hat{Q}'$ and $\hat{P}'$ be projectors onto $\mathbb{Q}'$ and $\mathbb{P}'$ respectively. Now $I - \hat{Q}'$ is the projector of $\mathbb{P}'^{\perp}$, and let $g(\hat{P}', \hat{Q}')$ be the projector $\mathbb{Q}' \cap \mathbb{P}'$. Employing the known formulas (see e.g. [47])

$$\mathrm{tr}(\hat{Q} - g(\hat{Q}, I - \hat{P})) = \mathrm{tr}(\hat{P} - g(\hat{P}, I - \hat{Q})), \tag{C2}$$

we get from (C1)

$$\dim \mathbb{P}' = \dim \mathbb{Q}' = \dim \mathbb{P}'^{\perp} = \dim \mathbb{Q}'^{\perp} = \frac{1}{2}\dim \mathbb{H}' = m, \tag{C3}$$

which means that $\dim \mathbb{H}'$ should be even for (C1) to hold[9].

Here is the statement of the CS-representation [46]: after a unitary transformation $\hat{Q}'$ and $\hat{P}'$ can be presented as

$$\hat{Q}' = \begin{pmatrix} I_m & 0_m \\ 0_m & 0_m \end{pmatrix}, \quad \hat{P}' = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}, \quad C^2 + S^2 = I_m, \quad \ker[C] = \ker[S] = 0, \tag{C4}$$

where all blocks in (C4) have the same dimension m.

To prove (C4), note that $\hat{Q}'$ and $\hat{P}'$ can be written as (cf (C3))

$$\hat{Q}' = \begin{pmatrix} I_m & 0_m \\ 0_m & 0_m \end{pmatrix}, \quad \hat{P}' = \begin{pmatrix} K' & B \\ B^{\dagger} & L \end{pmatrix}, \quad K' \geqslant 0, \ L \geqslant 0, \quad \mathrm{tr}(K' + L) = m. \tag{C5}$$

Next, let us show that

$$\ker[\, B \,] = 0. \tag{C6}$$

Since $\hat{Q}'\hat{P}'(I - \hat{Q}') = \begin{pmatrix} 0_m & B \\ 0_m & 0_m \end{pmatrix}$, we need to show that for any $|\psi\rangle \in \mathbb{Q}'^{\perp}$, $\hat{Q}'\hat{P}'|\psi\rangle = 0$ means $|\psi\rangle = 0$. Indeed, we have $\hat{P}'|\psi\rangle \in \mathbb{Q}'^{\perp}$, which together with $\mathbb{Q}'^{\perp} \cap \mathbb{P}' = 0$ (see (C1)) leads to $|\psi\rangle = 0$. Equation (C6) implies that there is the well-defined polar decomposition ($\hat{B}$ is Hermitean, while V is unitary)

$$B = V\hat{B}, \quad \hat{B} = \sqrt{B^{\dagger}B} = \hat{B}^{\dagger}, \quad V = B(B^{\dagger}B)^{-1/2} = V^{\dagger\,-1}. \tag{C7}$$

We transform as

$$\begin{pmatrix} V^{\dagger} & 0_m \\ 0_m & I_m \end{pmatrix}\hat{Q}'\begin{pmatrix} V & 0_m \\ 0_m & I_m \end{pmatrix} = \hat{Q}', \quad \begin{pmatrix} V^{\dagger} & 0_m \\ 0_m & I_m \end{pmatrix}\hat{P}'\begin{pmatrix} V & 0_m \\ 0_m & I_m \end{pmatrix} = \begin{pmatrix} K & \hat{B} \\ \hat{B} & L \end{pmatrix}, \tag{C8}$$

where $K = V^{\dagger}V'U$. We shall now employ the fact that the last matrix in (C8) is a projector:

$$K = K^2 + \hat{B}^2, \quad L = L^2 + \hat{B}^2, \quad \hat{B} = \hat{B}K + L\hat{B}. \tag{C9}$$

The first and second relations in (C9) show that $[K, \hat{B}] = [L, \hat{B}] = 0$. Then the third relations produces $\hat{B}(K + L - 1) = 0$. Since $\hat{B} > 0$ (due to $\ker(B) = 0$), we conclude that $K + L = 1$. The rest is obvious.

C.2. Joint commutant for two projectors

Given (C4), we want to find matrices that commute both with $\hat{P}'$ and $\hat{Q}'$ [46]. Matrices that commute with $\hat{Q}'$ read

$$\begin{pmatrix} X & 0_m \\ 0_m & Y \end{pmatrix}. \tag{C10}$$

Employing (C4), we get that (C10) commutes with $\hat{P}'$ if

$$XC^2 = C^2X, \tag{C11}$$

$$YS^2 = S^2Y, \tag{C12}$$

$$XCS = CSY. \tag{C13}$$

Since C and S are invertible, (C11), (C12) imply that $[X, S] = [X, C] = [Y, S] = [Y, C] = 0$. And then (C13) implies that $X = Y$. Hence

$$X = Y, \quad [X, C] = [X, S] = 0. \tag{C14}$$

C.3. General form of the CS representation

The above derivation of (C4) assumed conditions (C1). More generally, the Hilbert space $\mathbb{H}$ can be represented as a direct sum [45–47]

[9] Equation (C3) can be derived by noting that $\mathbb{Q}'^{\perp} \cap \mathbb{P}' = 0$ implies $\ker(\hat{Q}'\hat{P}') = 0$. Indeed, if $Q|p\rangle = 0$, where $|p\rangle \in \mathbb{P}'$, then $\mathbb{Q}'^{\perp} \cap \mathbb{P}' \neq 0$. Hence $|p\rangle = 0$. Let us mention for completeness that $\mathrm{ran}(\hat{Q}'\hat{P}') \cap \mathbb{Q}'^{\perp} = 0$. Indeed, let us assume that $|f\rangle \in \mathbb{Q}'$, $|g\rangle \in \mathbb{P}'$ and $\langle f|\hat{Q}'g\rangle = 0$. Then $\langle \hat{Q}'f|g\rangle = \langle f|g\rangle = 0$. The last relation means that either $f = 0$, or $\mathbb{Q}'^{\perp} \cap \mathbb{P}' \neq 0$, which contradicts to (C1).

$$\mathbb{H} = \mathbb{H}' \oplus \mathbb{H}_{11} \oplus \mathbb{H}_{10} \oplus \mathbb{H}_{01} \oplus \mathbb{H}_{00}, \tag{C15}$$

where the sub-space $\mathbb{H}_{\alpha\beta}$ of dimension $m_{\alpha\beta}$ is formed by common eigenvectors of P and Q having eigenvalue α (for P) and β (for Q). Depending on P and Q every sub-space can be absent; all of them can be present only for $\dim\mathbb{H} \geqslant 6$. Now $\mathbb{H}_{11} = \mathrm{ran}(P) \cap \mathrm{ran}(Q)$. $\mathbb{H}'$ has even dimension $2m$ [46, 47], this is the only sub-space that is not formed by common eigenvectors of P and Q.

After a unitary transformation

$$P = U\hat{P}U^{\dagger}, \quad Q = U\hat{Q}U^{\dagger}, \quad UU^{\dagger} = I, \tag{C16}$$

$\hat{P}$ and $\hat{Q}$ get the following block-diagonal form that is related to (C15) [46] and that generalizes (C4):

$$\hat{Q} = Q' \oplus I_{m_{11}} \oplus I_{m_{10}} \oplus 0_{m_{01}} \oplus 0_{m_{00}}, \quad Q' \equiv \begin{pmatrix} I_m & 0_m \\ 0_m & 0_m \end{pmatrix}, \tag{C17}$$

$$\hat{P} = P' \oplus I_{m_{11}} \oplus 0_{m_{10}} \oplus I_{m_{01}} \oplus 0_{m_{00}}, \quad P' \equiv \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}, \tag{C18}$$

where C and S are invertible square matrices of the same size holding

$$C^2 + S^2 = I_m, \quad [C, S] = 0. \tag{C19}$$

$\mathbb{H}'$ refers to P' and Q'. If $\hat{P}$ and $\hat{Q}$ do not have any common eigenvector, $\hat{P} = P'$ and $\hat{Q} = Q'$.

Appendix D. Derivation of the main result (equations (16), (17) of the main text)

We start with representation (C17), (C18) and axioms (3)–(7) of the main text. These axioms hold for $\hat{P}$, $\hat{Q}$ and $\hat{\rho} = U\rho U^{\dagger}$ (see (C16)) instead of P, Q and ρ, because $\omega(\hat{P}, \hat{Q}) = U^{\dagger}\omega(P, Q)U$ for $\omega = \underline{\omega}, \overline{\omega}$ (recall that $\underline{\omega}(\hat{P}, \hat{Q})$ and $\overline{\omega}(\hat{P}, \hat{Q})$ are Taylor expandable). Hence we now search for $\underline{\omega}(\hat{P}, \hat{Q})$ and $\overline{\omega}(\hat{P}, \hat{Q})$.

The block-diagonal form (C17), (C15) remains intact under addition and multiplication of $\hat{P}$ and $\hat{Q}$. Hence $\underline{\omega}(\hat{P}, \hat{Q})$ and $\overline{\omega}(\hat{P}, \hat{Q})$ have the block-diagonal form similar to (C17), where the diagonal blocks are to be determined. Let now $\Pi_{\alpha\beta}$ be the projector on $\mathbb{H}_{\alpha\beta}$. We get $[\alpha, \beta = 0, 1]$

$$\Pi_{\alpha\beta}\omega(\hat{P}, \hat{Q})\Pi_{\alpha\beta} = \omega(\alpha, \beta)\Pi_{\alpha\beta}, \quad \omega = \underline{\omega}, \overline{\omega}. \tag{D1}$$

Hence condition (5) of the main text implies [for $\omega = \underline{\omega}, \overline{\omega}$ and $\omega' = \underline{\omega}', \overline{\omega}'$]

$$\omega(\hat{P}, \hat{Q}) = \omega' \oplus I_{m_{11}} \oplus 0_{m_{10}+m_{01}+m_{00}}, \quad \omega' = \begin{pmatrix} \omega'_{11} & \omega'_{12} \\ \omega'^{\dagger}_{12} & \omega'_{22} \end{pmatrix}. \tag{D2}$$

Aiming to apply (6) of the main text, we write down (C17) explicitly as

$$\hat{Q} = \begin{pmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{D3}$$

The most general density matrix $\hat{\rho}$ that commutes with $\hat{Q}$ reads (in the same block-diagonal form)

$$\hat{\rho} = \begin{pmatrix} a_{11} & 0 & a_{12} & a_{13} & 0 & 0 \\ 0 & b_{11} & 0 & 0 & b_{12} & b_{13} \\ a_{12}^{\dagger} & 0 & a_{22} & a_{23} & 0 & 0 \\ a_{13}^{\dagger} & 0 & a_{23}^{\dagger} & a_{33} & 0 & 0 \\ 0 & b_{12}^{\dagger} & 0 & 0 & b_{22} & b_{23} \\ 0 & b_{13}^{\dagger} & 0 & 0 & b_{23}^{\dagger} & b_{33} \end{pmatrix}. \tag{D4}$$

Now $\hat{\rho}\hat{Q} = \hat{Q}\hat{\rho}$ is seen from the fact that after permutations of rows and columns, $\hat{Q}$ and $\hat{\rho}$ become $I_{m+m_{11}+m_{10}} \oplus 0_{m+m_{01}+m_{00}}$ and $a \oplus b$, respectively. Note that $a_{ii} \geqslant 0$ and $b_{ii} \geqslant 0$.

Equations (D2)–(D4) imply

$$\mathrm{tr}(\hat{Q}\hat{\rho}\hat{P}) = \mathrm{tr}(\hat{Q}\hat{\rho}\hat{Q}\hat{P}) = \mathrm{tr}(C^2 a_{11} + a_{22}), \tag{D5}$$

$$\mathrm{tr}(\underline{\omega}(\hat{P}, \hat{Q})\hat{\rho}) = \mathrm{tr}(\underline{\omega}'_{11} a_{11} + \underline{\omega}'_{22} a_{44} + a_{22}), \tag{D6}$$

$$\mathrm{tr}(\overline{\omega}(\hat{P}, \hat{Q})\hat{\rho}) = \mathrm{tr}(\overline{\omega}'_{11} a_{11} + \overline{\omega}'_{22} a_{44} + a_{22}). \tag{D7}$$

Condition (7) (of the main text) and (C14) imply

$$\omega'_{11} = \omega'_{22}, \quad \omega'_{12} = 0, \quad \text{for} \quad \omega' = \overline{\omega}', \underline{\omega}'. \tag{D8}$$

Recall condition (6) of the main text. It amounts to (D5) $\geqslant$ (D6) that should hold for arbitrary a_{ik} and b_{ik}. Hence we deduce: $\underline{\omega}'_{22} = 0$ and hence $\underline{\omega}'_{11} = \underline{\omega}'_{12} = 0$. Likewise, (D7) $\geqslant$ (D5) leads to $\overline{\omega}'_{11} = \overline{\omega}'_{22} = C^2$, $\overline{\omega}'_{12} = 0$; recall that we want the smallest upper probability. Now (4) (of the main text) holds, since

$$\begin{pmatrix} C^2 & 0_m \\ 0_m & C^2 \end{pmatrix} = \begin{pmatrix} I_m & 0_m \\ 0_m & I_m \end{pmatrix} - (P' - Q')^2. \tag{D9}$$

Thus (cf (C15))

$$\overline{\omega}(\hat{P}, \hat{Q}) = \begin{pmatrix} C^2 & 0_m \\ 0_m & C^2 \end{pmatrix} \oplus I_{m_{11}} \oplus 0_{m_{10}+m_{01}+m_{00}}, \tag{D10}$$

$$\underline{\omega}(\hat{P}, \hat{Q}) = 0_{2m} \oplus I_{m_{11}} \oplus 0_{m_{10}+m_{01}+m_{00}}. \tag{D11}$$

Now $g(\hat{P}, \hat{Q}) = g(\hat{Q}, \hat{P})$ is the projector onto $\mathrm{ran}(\hat{P}) \cap \mathrm{ran}(\hat{Q})$. To return from (D10), (D11) to original projectors P and Q, we note via (C17), (C18) (recall that $\Pi_{\alpha\beta}$ is the projector onto $\mathbb{H}_{\alpha\beta}$):

$$\Pi_{11} = g(\hat{P}, \hat{Q}), \quad \Pi_{10} = g(\hat{P}, I - \hat{Q}), \quad \Pi_{01} = g(I - \hat{P}, \hat{Q}), \quad \Pi_{00} = g(I - \hat{P}, I - \hat{Q}), \tag{D12}$$

$$\underline{\omega}(\hat{P}, \hat{Q}) = g(\hat{P}, \hat{Q}), \tag{D13}$$

$$P' = \hat{P} - g(\hat{P}, \hat{Q}) - g(\hat{P}, I - \hat{Q}), \tag{D14}$$

$$Q' = \hat{Q} - g(\hat{P}, \hat{Q}) - g(\hat{Q}, I - \hat{P}), \tag{D15}$$

$$\overline{\omega}(\hat{P}, \hat{Q}) = I - g(\hat{P}, I - \hat{Q}) - g(I - \hat{P}, \hat{Q}) - g(I - \hat{P}, I - \hat{Q}) - \Big(P - Q - g(\hat{P}, I - \hat{Q}) + g(I - \hat{P}, \hat{Q}) \Big)^2. \tag{D16}$$

We act back by U, e.g. $g(\hat{P}, \hat{Q}) = U^\dagger g(P, Q) U$, and get finally

$$\underline{\omega}(P, Q) = g(P, Q), \tag{D17}$$

$$\overline{\omega}(P, Q) = I - (P - Q)^2 - g(I - P, I - Q). \tag{D18}$$

Appendix E. Various representations of upper and lower probability operators

E.1. Representations for the upper probability operator

Let us turn into a more detailed investigation of (D18). Note from (C17), (C18) and (D12) that $I - g(I - P, I - Q)$ is the projector to $\mathrm{ran}(P) + \mathrm{ran}(Q)$, where $\mathrm{ran}(P) + \mathrm{ran}(Q)$ is the vector sum of two subspaces. Note the following representation [48]:

$$I - g(I - P, I - Q) = (P + Q)^- (P + Q) = (P + Q)(P + Q)^-, \tag{E1}$$

where A^- is the pseudo-inverse of Hermitean A, i.e. if $A = V(a \oplus 0)V^\dagger$ (where V is unitary: $VV^\dagger = I$), then $A^- = V(a^{-1} \oplus 0)V^\dagger$.

The third equality in (E1) is the obvious feature of the pseudo-inverse. The first equality in (E1) follows from the fact that $(P + Q)^- (P + Q)$ is the projector on $\mathrm{ran}(P + Q)$ and the known relation [48]:

$$\mathrm{ran}(P + Q) = \mathrm{ran}(P) + \mathrm{ran}(Q). \tag{E2}$$

Employing $(P - Q)^2 = I - (I - P - Q)^2$, $\overline{\omega}(P, Q)$ can be presented as a function of $P + Q$ (cf (E1)):

$$\overline{\omega}(P, Q) = (I - P - Q)^2 - I + (P + Q)(P + Q)^-. \tag{E3}$$

Note another representation for the projector to $\mathrm{ran}(P) + \mathrm{ran}(Q)$ [40]

$$I - g(I - P, I - Q) = \min[A \mid A \geqslant Q, P], \tag{E4}$$

where $I - g(I - P, I - Q)$ equals to the minimal Hermitean operator A that holds two conditions after |.

E.2. Representations for the lower probability operator

Let us first show that the projector $g(P, Q)$ into $\mathrm{ran}(P) \cap \mathrm{ran}(Q)$ holds [57]

$$g(P, Q) = 2P(P + Q)^- Q = 2Q(P + Q)^- P, \tag{E5}$$

where A^- is the pseudo-inverse of A (cf (E1)).

The last equality in (E5) follows from the fact that $(Q+P)(P+Q)^{-}$ is the projector to $\mathrm{ran}(P)+\mathrm{ran}(Q)$ (see (E1), E2)), which then leads to $P(Q+P)(P+Q)^{-}=(Q+P)(P+Q)^{-}P$.

The first equality in (E5) is shown as follows. Let $|\psi\rangle\in\mathrm{ran}(P)\cap\mathrm{ran}(Q)$. Then using (E5):

$$2P(P+Q)^{-}Q|\psi\rangle=(P+Q)(P+Q)^{-}|\psi\rangle=|\psi\rangle.$$

Thus, $\mathrm{ran}(2P(P+Q)^{-}Q)\supseteq(\mathrm{ran}(P)\cap\mathrm{ran}(Q))$. On the other hand, $\mathrm{ran}(2P(P+Q)^{-}Q)\subseteq\mathrm{ran}(P)$ and $\mathrm{ran}(2P(P+Q)^{-}Q)\subseteq\mathrm{ran}(Q)$, where the first relation follows from the implication: if $|\psi\rangle\notin\mathrm{ran}(P)$, then $|\psi\rangle\notin\mathrm{ran}(2P(P+Q)^{-}Q)$.

There are two other (more familiar) representations of $g(P,Q)$ (see e.g. [40, 48]):

$$g(P,Q)=\lim_{n\to\infty}Q(PQ)^{n}, \tag{E6}$$

$$=\max[A\mid 0\leqslant A\leqslant Q,P]. \tag{E7}$$

Equation (E6) can be interpreted as a result of (infinitely many) successive measurements of P and Q. Equation (E7) should be compared to (E4).

Yet another representation is useful in calculations, since it explicitly involves a 2×2 block-diagonal representation [57]:

$$P=\begin{pmatrix}P_{11}&P_{12}\\P_{21}&P_{22}\end{pmatrix},\quad Q=\begin{pmatrix}I_{n_1}&0\\0&0_{n_2}\end{pmatrix}, \tag{E8}$$

$$g(P,Q)=\begin{pmatrix}P_{11}-P_{12}P_{22}^{-}P_{21}&0\\0&0\end{pmatrix}, \tag{E9}$$

where P_{11}, P_{12}, P_{21} and P_{22} are, respectively, $n_1\times n_1$, $n_1\times n_2$, $n_2\times n_1$, $n_2\times n_2$ matrices.

E.3. Direct relation between the eigenvalues of $P-Q$ and PQ

We can now prove directly (i.e. without employing the CS representation) that there is a direct relation between the eigenvalues of $\overline{\omega}(P,Q)$ and PQ. Let $|x\rangle$ be the eigenvector of Hermitean operator $P-Q$:

$$(P-Q)|x\rangle=\lambda|x\rangle, \tag{E10}$$

where $-1\leqslant\lambda\leqslant 1$ is the eigenvalue. Multiplying both sides of (E10) by P (by Q) and using $P^2=P$ ($Q^2=Q$) we get

$$QP|x\rangle=(1+\lambda)Q|x\rangle,\quad PQ|x\rangle=(1-\lambda)P|x\rangle, \tag{E11}$$

which then implies

$$PQ(P|x\rangle)=(1-\lambda^2)(P|x\rangle),\quad QP(Q|x\rangle)=(1-\lambda^2)(Q|x\rangle). \tag{E12}$$

Thus $P|x\rangle$ ($Q|x\rangle$) is an eigenvector of PQ (QP) with eigenvalue $1-\lambda^2$.

As seen from (E11), the 2D linear space $\mathrm{Span}(P|x\rangle,\ Q|x\rangle)$ formed by all superpositions of $P|x\rangle$ and $Q|x\rangle$ remains invariant under action of both $\hat{P}$ and $\hat{Q}$. Together with $\mathrm{tr}(P-Q)=0$ this means that if (E10) holds, then $P-Q$ has eigenvalue $-\lambda$ with the eigen-vector living in $\mathrm{Span}(P|x\rangle,\ Q|x\rangle)$.

Further details on the relation between PQ and $P-Q$ can be looked up in [59].

Appendix F. Additivity and monotonicity

We discuss here the behavior of $\underline{\omega}(P,Q)$ and $\overline{\omega}(P,Q)$ (given by (D17), (D18)) with respect to a monotonic change of their arguments. For two projectors Q' and Q, $Q'\geqslant Q$ means $Q'=Q+K$, where $K^2=K$ and $QK=0$. Now (D17), (D18) and (E7) imply that $\underline{\omega}(P,Q)$ is operator superadditive

$$\underline{\omega}(P,Q+K)\geqslant\underline{\omega}(P,K)+\underline{\omega}(P,Q). \tag{F1}$$

Likewise, $\overline{\omega}(P,Q)$ is operator subadditive, but under an additional condition:

$$\overline{\omega}(P,Q+K)\leqslant\overline{\omega}(P,K)+\overline{\omega}(P,Q),\quad \text{if } Q+K=I. \tag{F2}$$

They are the analogues of classical features (B4) and (B5), respectively. Note that (F1) and (B4) are valid under the same conditions, since $QK=0$ is the analogue of $A\cap B=\{0\}$. In that sense the correspondence between (F2) and (B5) is more limited, since $Q+K=I$ is more restrictive than $QK=0$.

We focus on deriving (F1), since (F2) is derived in the same way. Note from (E7) that $g(P,Q)\leqslant Q$ and $g(P,K)\leqslant K$ imply $g(P,Q)+g(P,K)\leqslant Q+K$. Since $QK=0$, $g(P,Q)+g(P,K)\leqslant P$. Using (E7) for $g(P,Q+K)$ we obtain (F1).

Note as well that both $\overline{\omega}(P,\,Q)$ and $\underline{\omega}(P,\,Q)$ are monotonous [$\omega = \underline{\omega},\,\overline{\omega}$]:

$$\omega(P',\,Q') \geqslant \omega(P,\,Q), \quad \text{if } P' \geqslant P,\,Q \text{ and } Q' \geqslant P,\,Q. \tag{F3}$$

Equation (F3) for $\omega = \underline{\omega}$ follows from (F1). For $\omega = \overline{\omega}$ it is deduced as follows (cf (E4), (E7)):

$$\overline{\omega}(P',\,Q') \geqslant \underline{\omega}(P',\,Q') = g(P',\,Q') \geqslant I - g(I-P,\,I-Q) \geqslant \overline{\omega}(P,\,Q). \tag{F4}$$

Let us now discuss whether (F2) can hold under the same condition $QK = 0$ as (F1). Now

$$\overline{\omega}(P,\,Q+K) \leqslant \overline{\omega}(P,\,Q) + \overline{\omega}(P,\,K), \tag{F5}$$

amounts to

$$g(I-P,\,I-Q) + g(I-P,\,I-K) \leqslant I - P + g(I-P,\,I-Q-K). \tag{F6}$$

First of all note that for $Q + K = 1$ and $QK = 0$ we get $(I-Q)(I-K) = 0$ and (F6) does hold for the same reason as (F1).

For $QK = 0$, equation (F6) is invalid in 3D space (as well as for larger dimensional Hilbert spaces). Indeed, let us assume that Q and K are 1D:

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1. \end{pmatrix} \tag{F7}$$

Given $I - P$ as

$$I - P = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12}^{\star} & a_{22} & a_{23} \\ a_{13}^{\star} & a_{23}^{\star} & a_{33} \end{pmatrix}, \tag{F8}$$

we get (cf (E9))

$$g(I-P,\,1-K) = \begin{pmatrix} a'_{11} & a'_{12} & 0 \\ a'_{21} & a'_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} a_{33}^{-} (\,a_{31}\;a_{32}\,),$$

where a_{33}^{-} is the pseudo-inverse of a_{33}.

Likewise

$$g(I-P,\,I-Q) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a'_{22} & a'_{23} \\ 0 & a'_{32} & a'_{33} \end{pmatrix}, \quad \begin{pmatrix} a'_{22} & a'_{23} \\ a'_{32} & a'_{33} \end{pmatrix} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - \begin{pmatrix} a_{21} \\ a_{31} \end{pmatrix} a_{11}^{-} (\,a_{12}\;a_{13}\,).$$

Now $g(I-P,\,I-Q-K) = 0$, since $I-Q-K$ is a 1d projector. We can now establish that generically

$$g(I-P,\,I-Q) + g(I-P,\,I-K) \not\leqslant I \tag{F9}$$

(let alone (F6)), because the difference has both positive and negative eigenvalues.

The message (F9) is that the function $I - g(I-P,\,I-Q)$ is not sub-additive.

Now consider (F5), (F6), but under additional condition that $PK = 0$. Now (F6) amounts to

$$g(I-P,\,I-Q) \leqslant K + g(I-P,\,I-Q-K), \tag{F10}$$

which holds as equality since $\mathrm{ran}\,(K) \subseteq \mathrm{ran}\,(P^{\perp}) \cap \mathrm{ran}\,(Q^{\perp})$.

Appendix G. Upper and lower probabilities for simple examples

G.1. 2D Hilbert space

It should be clear from (D10), (D11) that in 2D Hilbert space, any lower probability operator $\underline{\omega}(P,\,Q)$ is zero (since two rays overlap only at zero), while the upper probability operator $\overline{\omega}(P,\,Q) = \overline{p}(\rho;\,P,\,Q)$ just reduces to the transition probability (i.e. to a number) $\mathrm{tr}\,(PQ)$. Thus for the present case both $\overline{p}$ and $\underline{p}$ do not depend on ρ.

G.2. Spin 1

G.2.1. Projectors. The 3×3 matrices for the spin components read

$$L^x=\frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix},\quad L^y=\frac{i}{\sqrt{2}}\begin{pmatrix}0&-1&0\\1&0&-1\\0&1&0\end{pmatrix},\quad L^z=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{G1}$$

Now $P^a_{\pm 1,0}$ for $a=x,\,y,\,z$ are the 1D projectors to the eigenspace with eigenvalues ± 1 or 0 of L^a:

$$P^x_1=\frac{1}{4}\begin{pmatrix}1&\sqrt{2}&1\\\sqrt{2}&2&\sqrt{2}\\1&\sqrt{2}&1\end{pmatrix}\quad P^x_0=\frac{1}{2}\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix},\quad P^x_{-1}=\frac{1}{4}\begin{pmatrix}1&-\sqrt{2}&1\\-\sqrt{2}&2&-\sqrt{2}\\1&-\sqrt{2}&1\end{pmatrix}, \tag{G2}$$

$$P^y_1=\frac{1}{4}\begin{pmatrix}1&-i\sqrt{2}&-1\\i\sqrt{2}&2&-i\sqrt{2}\\-1&i\sqrt{2}&1\end{pmatrix}\quad P^y_0=\frac{1}{2}\begin{pmatrix}1&0&1\\0&0&0\\1&0&1\end{pmatrix},\quad P^y_{-1}=\frac{1}{4}\begin{pmatrix}1&i\sqrt{2}&-1\\-i\sqrt{2}&2&i\sqrt{2}\\-1&-i\sqrt{2}&1\end{pmatrix}, \tag{G3}$$

$$P^z_1=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}\quad P^z_0=\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix},\quad P^z_{-1}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&1\end{pmatrix}, \tag{G4}$$

where the zero components are orthogonal to each other:

$$P^x_0P^y_0=P^x_0P^z_0=P^z_0P^y_0=0. \tag{G5}$$

Other overlaps are simple as well ($\alpha\neq\beta$)

$$\begin{aligned}\mathrm{tr}\left(P^\alpha_jP^\beta_k\right)&=1/4\ \text{ if }\ j\neq 0\ \text{and}\ k\neq 0,\\&=1/2\ \text{ if }\ j=0\ \text{or}\ k=0\ \text{but not both},\\&=0\ \text{ if }\ j=0\ \text{and}\ k=0.\end{aligned} \tag{G6}$$

Given two projectors P and Q, we defined $g(P,\,Q)$ as the projector on $\mathrm{ran}(P)\cap\mathrm{ran}(Q)$. For calculating $g(P,\,Q)$ we employ (E5).

G.2.2. Fine-grained joint probabilities for $P^z_{\pm1,0}$ and $P^x_{\pm1,0}$. Here are upper probability operators for joint values of $P^z_{\pm1,0}$ and $P^x_{\pm1,0}$:

$$\overline{\omega}(P^z_0,\,P^x_0)=0, \tag{G7}$$

$$\overline{\omega}\left(P^x_{\pm1},\,P^z_1\right)=\begin{pmatrix}\frac{1}{4}&0&0\\0&\frac{1}{6}&\frac{\pm1}{6\sqrt{2}}\\0&\frac{\pm1}{6\sqrt{2}}&\frac{1}{12}\end{pmatrix}, \tag{G8}$$

$$\overline{\omega}\left(P^x_{\pm1},\,P^z_{-1}\right)=\begin{pmatrix}\frac{1}{12}&\frac{\pm1}{6\sqrt{2}}&0\\\frac{\pm1}{6\sqrt{2}}&\frac{1}{6}&0\\0&0&\frac{1}{4}\end{pmatrix}, \tag{G9}$$

$$\overline{\omega}\left(P^x_0,\,P^z_{\pm1}\right)=\begin{pmatrix}\frac{1}{2}&0&0\\0&0&0\\0&0&\frac{1}{2}\end{pmatrix}, \tag{G10}$$

$$\overline{\omega}\left(P^x_{\pm1},\,P^z_0\right)=\begin{pmatrix}\frac{1}{4}&0&-\frac{1}{4}\\0&\frac{1}{2}&0\\-\frac{1}{4}&0&\frac{1}{4}\end{pmatrix}. \tag{G11}$$

Since $P^z_{\pm1,0}$ and $P^x_{\pm1,0}$ are 1D projectors, all the lower probability operators nullify. Equation (G7) means that the precise probability of P^z_0 and P^x_0 is zero; cf (G5).

We now get from (G7)–(G11)

$$\sum_{k=\pm,0}\sum_{i=\pm,0}\overline{\omega}\,(P^x_k,\,P^z_i)=\begin{pmatrix}\frac{13}{6} & 0 & -\frac{1}{2}\\ 0 & \frac{5}{3} & 0\\ -\frac{1}{2} & 0 & \frac{13}{6}\end{pmatrix}. \tag{G12}$$

This matrix is larger than I, since its eigenvalues are $\frac{8}{3}$, $\frac{5}{3}$ and $\frac{5}{3}$.

Note from (C17), (C18) that for 3×3 matrices $\dim\mathcal{H}'=2$, while $\dim\mathcal{H}_{11}=1$ (if this sub-space is present at all). Hence the eigenvalues of $\overline{\omega}$ relate to transition probabilities (G6). Indeed, the eigenvalues of matrices in (G8), (G9) (resp. in (G10), (G11)) is $(\frac{1}{4},\,\frac{1}{4},\,0)$ (resp. $(\frac{1}{2},\,\frac{1}{2},\,0)$). Hence the maximal probability interval $[\frac{1}{4},\,0]$ that can be generated by (G8), (G9) is smaller than the maximal interval $[\frac{1}{2},\,0]$ generated by (G10), (G11). As an example, let us take the upper probabilities generated on eigenstates of $L^y(\epsilon,\,\eta,\,\chi=1,\,0,\,-1)$:

$$\begin{aligned}\mathrm{tr}\Big[\,\overline{\omega}\left(P^x_\epsilon,\,P^z_\eta\right)P^y_\chi\Big]&=1/6\quad\text{if}\quad\epsilon\eta\neq0,\\ &=1/4\quad\text{if}\quad\epsilon\eta=0,\,(1-\epsilon)(1-\eta)\neq1,\,\chi\neq0,\\ &=1/2\quad\text{if}\quad\epsilon\eta=0,\,(1-\epsilon)(1-\eta)\neq1,\,\chi=0,\\ &=0\quad\text{if}\quad\epsilon=\eta=\chi=0.\end{aligned} \tag{G13}$$

G.2.3. Coarse-grained joint probabilities for $P^z_{\pm1,0}$ and $P^x_{\pm1,0}$. Let us now turn to joint probabilities, where the lower probability is non-zero

$$\underline{\omega}\left(P^x_0+P^x_{\pm1},\,P^z_1+P^z_0\right)=\begin{pmatrix}\frac{2}{3} & \pm\frac{\sqrt{2}}{3} & 0\\ \pm\frac{\sqrt{2}}{3} & \frac{1}{3} & 0\\ 0 & 0 & 0\end{pmatrix},\quad \overline{\omega}\left(P^x_0+P^x_{\pm1},\,P^z_1+P^z_0\right)=\begin{pmatrix}\frac{3}{4} & \pm\frac{1}{2\sqrt{2}} & 0\\ \pm\frac{1}{2\sqrt{2}} & \frac{1}{2} & 0\\ 0 & 0 & \frac{1}{4}\end{pmatrix}, \tag{G14}$$

$$\underline{\omega}\left(P^x_0+P^x_{\pm1},\,P^z_{-1}+P^z_0\right)=\begin{pmatrix}0 & 0 & 0\\ 0 & \frac{1}{3} & \pm\frac{\sqrt{2}}{3}\\ 0 & \pm\frac{\sqrt{2}}{3} & \frac{2}{3}\end{pmatrix},\quad \overline{\omega}\left(P^x_0+P^x_{\pm1},\,P^z_{-1}+P^z_0\right)=\begin{pmatrix}\frac{1}{4} & 0 & 0\\ 0 & \frac{1}{2} & \pm\frac{1}{2\sqrt{2}}\\ 0 & \pm\frac{1}{2\sqrt{2}} & \frac{3}{4}\end{pmatrix}, \tag{G15}$$

$$\underline{\omega}\left(P^x_1+P^x_{-1},\,P^z_{\pm1}+P^z_0\right)=\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},\quad \overline{\omega}\left(P^x_1+P^x_{-1},\,P^z_{\pm1}+P^z_0\right)=\begin{pmatrix}\frac{1}{2} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \frac{1}{2}\end{pmatrix}, \tag{G16}$$

$$\underline{\omega}\left(P^x_0+P^x_{\pm1},\,P^z_1+P^z_{-1}\right)=\begin{pmatrix}\frac{1}{2} & 0 & -\frac{1}{2}\\ 0 & 0 & 0\\ -\frac{1}{2} & 0 & \frac{1}{2}\end{pmatrix},\quad \overline{\omega}\left(P^x_0+P^x_{\pm1},\,P^z_1+P^z_{-1}\right)=\begin{pmatrix}\frac{3}{4} & 0 & -\frac{1}{4}\\ 0 & \frac{1}{2} & 0\\ -\frac{1}{4} & 0 & \frac{3}{4}\end{pmatrix}, \tag{G17}$$

$$\underline{\omega}\,(P^x_1+P^x_{-1},\,P^z_1+P^z_{-1})=\overline{\omega}\,(P^x_1+P^x_{-1},\,P^z_1+P^z_{-1})=\begin{pmatrix}\frac{1}{2} & 0 & \frac{1}{2}\\ 0 & 0 & 0\\ \frac{1}{2} & 0 & \frac{1}{2}\end{pmatrix}. \tag{G18}$$

Now $\overline{\omega}-\underline{\omega}$ for (G14), (G15) has eigenvalues $(\frac{1}{4},\,\frac{1}{4},\,0)$, while for for (G16), (G17) this matrix has eigenvalues $(\frac{1}{2},\,\frac{1}{2},\,0)$ (the last case (G18) refers to the commutative situation). Hence the probabilities for (G16), (G17) are more uncertain.

Next, let us establish whether certain combinations can be (surely) more probable than others. Note that

$$\text{Eigenvalues}\,[\underline{\omega}\,(P_0^x+P_{-1}^x,\,P_1^z+P_0^z)-\overline{\omega}\,(P_0^x+P_1^x,\,P_1^z+P_0^z)]=\left(\frac{\pm\sqrt{393}-3}{24},\,-\frac{1}{4}\right). \tag{G19}$$

Once there is (one) positive eigenvalue, there is a class of states ρ for which

$$\mathrm{tr}(\rho\,\underline{\omega}\,(P_0^x+P_{-1}^x,\,P_1^z+P_0^z))>\mathrm{tr}(\rho\,\overline{\omega}\,(P_0^x+P_1^x,\,P_1^z+P_0^z)), \tag{G20}$$

i.e. $P^x=0$ or -1 and $P^z=0$ or 1 is more probable than $P^x=0$ or 1 and $P^z=0$ or 1. Note that

$$[\underline{\omega}\,(P_0^x+P_{-1}^x,\,P_1^z+P_0^z),\,\overline{\omega}\,(P_0^x+P_1^x,\,P_1^z+P_0^z)]\neq 0. \tag{G21}$$

Such examples can be easily continued, e.g.

$$\text{Eigenvalues}\,[\underline{\omega}\,(P_0^x+P_1^x,\,P_1^z+P_0^z)-\overline{\omega}\,(P_1^x+P_{-1}^x,\,P_1^z+P_0^z)]=\left(\frac{\pm\sqrt{57}-3}{12},\,-\frac{1}{2}\right). \tag{G22}$$

References

[1] Wigner E P 1971 *Perspectives in Quantum Theory* ed W Yourgrau and A van der Merwe (Cambridge, MA: MIT Press)
Srinivas M D and Wolf E 1975 *Phys. Rev.* D **11** 1477
Kruszynski P and de Muynck W M 1987 *J. Math. Phys.* **28** 1761
Cohen L 1966 *Phil. Sci.* **33** 317
Margenau H and Cohen L 1968 *Quantum Theory and Reality* ed M Bunge (Berlin: Springer)
[2] Hudson R 1974 *Rep. Math. Phys.* **6** 249
Soto F and Claverie P 1983 *J. Math. Phys.* **24** 97
[3] Gudder S P 1968 *J. Math. Mech.* **18** 325 (https://zbmath.org/?q=an:0241.60092)
[4] de Muynck W M 2002 *Foundations of Quantum Mechanics, an Empiricist Approach* (Dordrecht: Kluwer Academic)
[5] Busch P, Lahti P J and Mittelstaedt P 1996 *The Quantum Theory of Measurement* (Berlin: Springer)
[6] Fine A 1982 *Phys. Rev. Lett.* **48** 291
Malley J D 2004 *Phys. Rev.* A **69** 022118
Malley J D and Fine A 2005 *Phys. Lett.* A **347** 51
[7] Hillery M, O'Connell R F, Scully M O and Wigner E P 1984 *Phys. Rep.* **106** 121
[8] Tatarskii V I 1983 *Sov. Phys.—Usp.* **26** 311
[9] Ferrie C 2011 *Rep. Prog. Phys.* **74** 116001
[10] Gardiner C W and Zoller P 2004 *Quantum Noise* (Berlin: Springer)
[11] Blakie P B, Bradley A S, Davis M J, Ballagh R J and Gardiner C W 2008 *Adv. Phys.* **57** 363
[12] Allahverdyan A E 2014 *Phys. Rev.* E **90** 032137
[13] Ballentine E 1970 *Rev. Mod. Phys.* **42** 358
[14] Mückenheim W 1986 *Phys. Rep.* **133** 337
[15] Kirkwood J G 1933 *Phys. Rev.* **44** 31
Terletsky Y P 1937 *Zh. Eksp. Teor. Fiz.* **7** 1290
Dirac P A M 1945 *Rev. Mod. Phys.* **17** 195
Margenau H and Hill R N 1961 *Prog. Theor. Phys.* **26** 722
[16] Barut A O 1957 *Phys. Rev.* **108** 565
[17] Landau L D and Lifshitz E M 1978 *Statistical Physics I* (Oxford: Pergamon)
[18] Khrennikov A 1997 *Can. J. Phys.* **75** 291
Bulinski A and Khrennikov A 2004 *Stat. Prob. Lett.* **70** 49–58
[19] Bartlett M S 1945 *Math. Proc. Camb. Phil. Soc.* **41** 71
Székely G J 2005 *Wilmott Magazine* **50** 66–8 (http://www.wilmott.com/pdfs/100609_gjs.pdf)
[20] Cohen L and Zaparovanny Y I 1980 *J. Math. Phys.* **21** 794
[21] de Muynck W M, Janssen P A E M and Santman A 1979 *Found. Phys.* **9** 71
[22] Arthurs E and Goodman M S 1988 *Phys. Rev. Lett.* **60** 2447
[23] Allahverdyan A E, Balian R and Nieuwenhuizen Th M 2013 *Phys. Rep.* **525** 1
[24] Yukalov V I and Sornette D 2013 *Laser Phys.* **23** 105502
[25] Heinosaari T, Kiukas J and Reitzer D 2014 *J. Phys. A: Math. Theor.* **47** 225301
[26] Allahverdyan A E, Balian R and Nieuwenhuizen Th M 2004 *Phys. Rev. Lett.* **92** 120402
[27] Uffink J 1994 *Int. J. Theor. Phys.* **33** 199
[28] Bub J 1977 *J. Phil. Log.* **6** 381
[29] Hughes R I G 1989 *The Structure and Interpretation of Quantum Mechanics* (Cambridge: Harvard University Press)
[30] Prugovecki E 1967 *Can. J. Phys.* **45** 2173
[31] Suppes P and Zanotti M 1991 *Found. Phys.* **21** 1479
[32] de Barros J A and Suppes P 2000 Some conceptual issues involving probability in quantum mechanics (arXiv:quant-ph/0001017)
[33] Hartmann S and Suppes P 2010 Entanglement, upper probabilities and decoherence in quantum mechanics *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association* ed M Suaraz *et al* vol 2 (Berlin: Springer) pp 93–103
[34] Fine T L 1994 *in Patrick Suppes: Scientific Philosopher* (Dordrecht: Kluwer Academic) pp 109–133
[35] Fine T L 1974 *Synthese* **29** 187
[36] Ochs W 1985 *Quantum Probability and Applications II* (Berlin: Springer) pp 388–396
[37] Galvan B 2008 *J. Stat. Phys.* **131** 1155
[38] Khrennikov A 2009 *Contextual Approach to Quantum Formalism* (Berlin: Springer)
[39] Jauch J M 1974 *Synthese* **29** 131

[40] Jauch J M 1968 *Foundations of Quantum Mechanics* (Reading, MA: Addison-Wesley)
[41] Jauch J M and Piron C 1963 *Helv. Phys. Acta* **36** 827 (http://cds.cern.ch/record/345316/files/CM-P00056869.pdf)
[42] Busch P 1984 *J. Math. Phys.* **25** 1794
[43] Bell J S 1966 *Rev. Mod. Phys.* **38** 447
[44] Logiurato F and Smerzi A 2012 *J. Mod. Phys.* **3** 1802
[45] Dixmier J 1948 *Rev. Sci.* **86** 387–99 (in French)
[46] Halmos P 1969 *Trans. Am. Math. Soc.* **144** 381
[47] Hardegree G M 1977 *Found. Phys.* **7** 495
[48] Piziak R, Odell P L and Hahn R 1999 *Comput. Math. Appl.* **37** 67
[49] Shafer G 1978 *Arch. Hist. Exact Sci.* **19** 309–70
[50] Ellsberg D 1961 *Q. J. Econ.* **75** 643
[51] Good I J 1962 *Logic, Methodology and Philosophy of Science: Proc. of the 1960 Int. Congress* ed E Nagel, P Suppes, A Tarski *et al* (Stanford: Stanford University Press)
[52] Kuznetsov V P 1991 *Interval Statistical Models* (Moscow: Radio i Svyaz Publ.) (in Russian)
[53] Walley P 1991 *Statistical Reasoning with Imprecise Probabilities* (London: Chapman and Hall)
[54] Walley P and Fine T L 1982 *Ann. Stat.* **10** 741
[55] FAQs of imprecise probability (http://-old.me.gatech.edu/~ywang/research/interval/WhyImpreciseProbability.html)
[56] Alefeld G and Herzberger J 1983 *Introduction to interval computations* (New York: Academic)
[57] Anderson W N and Duffin R J 1969 *J. Math. Anal. Appl.* **26** 576
[58] Halpern J Y and Pucella R 2002 *J. Artif. Intell. Res.* **17** 57
[59] Anderson W N, Harner E J and Trapp G E 1985 *Linear Multilinear Algebr.* **17** 295
[60] Halpern J Y and Pucella R 2014 Reasoning about expectation (arXiv:1407.7184)
[61] Lenard A 1972 *J. Funct. Anal.* **10** 410
[62] Amrein W O and Berthier A M 1977 *J. Funct. Anal.* **24** 258
[63] Busch P and Lahti P J 1986 *Phys. Lett.* A **115** 259
[64] Sunder V S 1988 *Can. J. Math.* **40** 38