Rapid Communications

Excluding joint probabilities from quantum theory

Armen E. Allahverdyan and Arshag Danageozian
Yerevan Physics Institute, 2 Alikhanian Brothers Street, Yerevan 0036, Armenia

(Received 3 August 2017; published 14 March 2018)

Quantum theory does not provide a unique definition for the joint probability of two noncommuting observables, which is the next important question after the Born's probability for a single observable. Instead, various definitions were suggested, e.g., via quasiprobabilities or via hidden-variable theories. After reviewing open issues of the joint probability, we relate it to quantum imprecise probabilities, which are noncontextual and are consistent with all constraints expected from a quantum probability. We study two noncommuting observables in a two-dimensional Hilbert space and show that there is no precise joint probability that applies for any quantum state and is consistent with imprecise probabilities. This contrasts with theorems by Bell and Kochen-Specker that exclude joint probabilities for more than two noncommuting observables, in Hilbert space with dimension larger than two. If measurement contexts are included into the definition, joint probabilities are not excluded anymore, but they are still constrained by imprecise probabilities.

DOI: 10.1103/PhysRevA.97.030102

Open problems of quantum mechanics revolve around the notions of noncommutativity and probability [1–8]. In contrast to classical mechanics, where introducing probability relates to a limited control over the experimental situation, the quantum probability presents itself as a fundamental description of a single quantum observable, as well as pairs of commuting observables [1–3]. However, the structure and interpretation of the joint probability for noncommuting observables is an open problem, primarily because the standard machinery of quantum mechanics precludes a direct and precise joint measurement of such observables [1–3]. This need not exclude the definition of joint probability, since the latter may turn out to be a construct recovered indirectly through different measurements, or may even point out to a generalized theory beyond quantum mechanics. As seen below, many possible candidates for the joint probability were proposed. This subject has been active since the inception of quantum mechanics [9], and it is much alive now with the needs of quantum information and foundations [10]. Theorems by Bell and Kochen-Specker demonstrate the nonexistence of the joint probability for more than two variables living in a Hilbert space $\mathcal{H}$ with dimension $\dim\mathcal{H} > 2$ [5,6,11,12]. The Kochen-Specker theorem looks for a set of observables for which no joint probability exists for any state, while Bell's theorem is restricted to specific (entangled) states. Neither of these theorems restricts joint probability for two noncommuting observables, which is the next important case after the probability of a single observable. Recently Malley added to the Kochen-Specker setup the consistency with quantum conditional probability, which is derived assuming (additionally) the projection postulate [13]. This modified setup disallows joint probabilities for any pair of noncommuting observables for $\dim\mathcal{H} > 2$ [13].

Here we show in which sense the joint probability can be excluded (without assuming the projection postulate) already for the minimal situation, *viz.* any pair of noncommuting observables living in a two-dimensional Hilbert space. Note that the projection postulate is not necessary for the empiric applicability of quantum probability; e.g., there are interpretations of quantum mechanics that do not employ it [14,15]. We also study how our results depend on the standard assumption of a single probability space for noncommuting observables.

We illustrate the existing approaches towards defining joint probabilities for two noncommuting observables (Hermitian operators) A and B ($[A,B] \neq 0$) living in a Hilbert space $\mathcal{H}$. Let P_a (Q_b) be the eigenprojector of A (B) that corresponds to the eigenvalue a (b).

Our first condition is that a hypothetical joint probability p_{ab} of eigenvalues of a and b in a state with density matrix ρ holds marginality conditions:

$$\sum\nolimits_a p_{ab} = \mathrm{tr}(\rho Q_b), \quad \sum\nolimits_b p_{ab} = \mathrm{tr}(\rho P_a), \tag{1}$$

where the first (second) summation is taken over all *different* eigenvalues of A (B).

We embedded the joint probabilities of noncommuting observables into a single space; cf. (1). This assumption can be questioned, since noncommuting P_a and Q_b demand different measurement contexts [16]. But we adopt it (till the end of the paper), since it is a common point for all approaches–including the Bell and Kochen-Specker theorems–that look for the joint probability [5,6]. Below we give examples for a joint probability.

(1) The best-known approach is that of quasiprobabilities [9,10]. They are linear over ρ, but they ought to become negative for certain states [9,10]. This limits their interpretation as probabilities [17]. In addition, there are many different quasiprobabilities, and it is not clear which one of them applies in a concrete situation, even if it is positive. Despite these issues, quasiprobabilities are widely used, e.g., in quasiclassics [9], information theory [10], signal processing [18,19], and statistical mechanics [20,21]. A good example of quasiprobability is the Terletsky-Margenau-Hill function [22]:

$$p^{\mathbf{T}}_{ab} = \mathrm{tr}[\rho\,(P_a Q_b + Q_b P_a)]/2. \tag{2}$$

$p^{\mathbf{T}}_{ab}$ is negative for certain ρ's, since $P_a Q_b + Q_b P_a$ has a negative eigenvalue if $[P_a, Q_b] \neq 0$. Equation (2) relates to weak values [23], and its negativity has a physical meaning [21]. $p^{\mathbf{T}}_{ab}$ is more convenient than the Wigner's function, which is defined only for specific pairs of observables.

The negativity of (2) for certain ρ's has deeper roots: any probability defined via

$$p^{\mathbf{C}}_{ab} = \mathrm{tr}(\rho\, \Pi_{ab}) \geqslant 0, \text{ due to } \Pi_{ab} \geqslant 0, \quad (3)$$

which holds (1) for any ρ (hence $\sum_b \Pi_{ab} = P_a$ and $\sum_a \Pi_{ab} = Q_b$) and which is forced to be positive by the non-negative definiteness of Hermitian Π_{ab}, will imply that A and B are commuting: $AB = BA$ [6,7].

(2) The linearity of (2), (3) over ρ has two implications. First, it means that within the standard measurement theory $p^{\mathbf{T}}_{ab}$ can be determined via measuring Hermitian $P_a Q_b + Q_b P_a$ in a state with an unknown ρ [1,3]. Second, consider the process of mixing: out of two ensembles with different ρ_1 and ρ_2, one makes up a new ensemble by taking ρ_k with probability μ_k ($k = 1,2$). The new ensemble has density matrix $\rho_{\rm mix} = \sum_{k=1}^{2} \mu_k \rho_k$. All quantities that are linear over density matrix (e.g., $p^{\mathbf{T}}_{ab}$) will depend directly on $\rho_{\rm mix}$ keeping no memory on separate preparations ρ_k that contributed into the mixing.

But the linearity is not obligatory: one can define a joint probability via sacrificing the linearity over ρ, so as to ensure the positivity for all ρ, keeping correct marginals as in (1) [19]. Such constructs cannot be measured via the standard approach, but they are determined theoretically once ρ is known. They are not unique, but they were employed for studying complementarity [18]. The simplest *nonlinear* example holding (1) is

$$p^{\mathbf{N}}_{ab} = \mathrm{tr}(P_a \sqrt{\rho}\, Q_b \sqrt{\rho}) \geqslant 0. \quad (4)$$

(3) Deterministic hidden-variable theories offer another definition of the joint probability [24]. Here is an example based on a hidden-variable theory proposed by Bell for $\dim \mathcal{H} = 2$ [24]. Recall the Bloch representation for any non-negative operator $R \geqslant 0$ (density matrix or projector) with $\mathrm{tr}\, R = 1$ in $\dim \mathcal{H} = 2$:

$$R = (1 + \vec{\beta}_R\, \vec{\sigma})/2, \quad \vec{\beta}_R = \mathrm{tr}(R\vec{\sigma}), \quad |\vec{\beta}_R| \leqslant 1, \quad (5)$$

where $\vec{\beta}_R = (\beta_{R,x}, \beta_{R,y}, \beta_{R,z})$ is a real vector and $\vec{\sigma}$ is the vector of Pauli matrices. Now the hidden variable is a real vector $\vec{m}$ with $|\vec{m}| = 1$. For any projector P we recover Born's rule via the integration over $\vec{m}$ with the characteristic function $\vartheta[\,\vec{\beta}_P(\vec{\beta}_\rho + \vec{m})]$ [24]:

$$\mathrm{tr}(\rho P) = \frac{1}{2}(1 + \vec{\beta}_\rho \vec{\beta}_P) = \int \frac{d\vec{m}}{4\pi}\, \vartheta[\,\vec{\beta}_P(\vec{\beta}_\rho + \vec{m})], \quad (6)$$

where $\vartheta[x]$ is the step function ($\vartheta[x \geqslant 0] = 1$ and $\vartheta[x < 0] = 0$), and $\int d\vec{m}$ integrates over all directions of three-dimensional hidden-variable space. Equation (6) is verified by going to spherical coordinates: $d\vec{m} = \sin(\theta)\, d\theta\, d\phi$. This model suggests the joint probability which holds (1):

$$p^{\mathbf{B}}_{ab} = \int \frac{d\vec{m}}{4\pi}\, \vartheta[\vec{\beta}_{P_a}(\vec{\beta}_\rho + \vec{m})]\, \vartheta[\,\vec{\beta}_{Q_b}(\vec{\beta}_\rho + \vec{m})] \geqslant 0. \quad (7)$$

Now (2) and (4) are noncontextual definitions, i.e., p_{ab} depends only on P_a and Q_b, and not on other projectors $\sum_a P_a = I$ and $\sum_a Q_b = I$ of A and B (I is the unity operator on $\mathcal{H}$). Equation (7) is also noncontextual, but its generalizations to $\dim \mathcal{H} > 2$ ought to be contextual [25].

Besides (1) there is another natural condition to which a physical joint probability p_{ab} should satisfy [11,12]:

$$p_{ab} = \mathrm{tr}(\rho P_a Q_b) \text{ if} \quad (8)$$

$$[P_a, \rho] = 0 \text{ or } [\rho, Q_b] = 0 \text{ or } [P_a, Q_b] = 0. \quad (9)$$

To explain (8) and (9), note that (9) forces $\mathrm{tr}(\rho P_a Q_b)$ to have features of joint probability, i.e., it is symmetric with respect to P_a and Q_b, non-negative and holds (1). Imposing (8) and (9) is especially obvious for $[P_a, Q_b] = 0$, where $P_a Q_b$ is a projector. For $[P_a, \rho] = 0$, $\mathrm{tr}(\rho P_a Q_b)$ can be recovered as the average of a Hermitian observable $(P_a Q_b + Q_b P_a)/2$. Alternatively, we note that measuring P_a does not change ρ statistically. Hence the joint probability is found by first measuring P and then Q: $\mathrm{tr}(\rho P_a Q_b) = \mathrm{tr}(P_a \rho P_a Q_b)$, and likewise for $[Q_b, \rho] = 0$.

Now (4) and (7)–which are non-negative for any ρ–do not hold (8) and (9). Equation (4) does not hold the third condition in (9), while (7) does not hold the first and second conditions in (9), as seen already in the simplest case $\rho = 1/2$. Equation (2) holds Eqs. (1), (8), and (9), but it is negative for certain ρ's. In this context we formulate the following.

Conjecture: There is *no* joint probability $p_{ab}(P_a, Q_b, \rho)$ that is non-negative for any ρ (i.e., quasiprobabilities are excluded), is noncontextual, and holds (8) and (9); cf. [8]. We stress that we do not require $p_{ab}(P_a, Q_b, \rho)$ to be linear over ρ. This conjecture is yet to be (in)validated.

Below we show that joint probabilities can be excluded from a different argument that relates to imprecise probabilities. In contrast to the usual joint probability, the imprecise probabilities are well defined given conditions of noncontextuality and correspondence with the commutative situation. The physical reason for this is that there exists a specific type of quantum uncertainty for two noncommuting observables that is captured by the quantum imprecise probability, which is consistent with all conditions expected from a quantum probability.

Before continuing, we comment on joint measurements schemes for noncommuting observables [1,2,26,27], a known method for characterizing noncommutativity. Generally, this method does not provide definitions for joint probabilities that are new compared with the above analysis. In particular, it is unclear to which extent the existing schemes for joint measurements produce intrinsic results that characterize the system itself and not approximate measurements employed [28]; e.g., they do not hold condition (1) of the joint probability [28]. Instead, they focus on different conditions, e.g., the unbiasedness [26] or stability [27].

Projectors are self-adjoint operators P with $P^2 = P$. Any projector P in a Hilbert space $\mathcal{H}$ bijectively relates to the subspace $\mathcal{S}_P$ of $\mathcal{H}$ [29]:

$$\mathcal{S}_P = \{|\psi\rangle \in \mathcal{H};\, P|\psi\rangle = |\psi\rangle\}. \quad (10)$$

Eigenvalues of P are 0 and/or 1, and it is a quantum analog of the characteristic function for a classical set [29]. Hence projectors define quantum probability: with a density matrix ρ, the probability of finding the eigenvalue 1 of P is given by Born's formula $\mathrm{tr}(\rho P)$.

The simplest projectors are 0 and I. We define

$$P \geqslant P' \text{ means } \langle\psi|P - P'|\psi\rangle \geqslant 0 \text{ for any } |\psi\rangle. \quad (11)$$

Now apply (11) with $|\psi\rangle = |\psi_0\rangle$, where $P|\psi_0\rangle = 0$, and then with $|\psi\rangle = |\psi_1\rangle$, where $P'|\psi_1\rangle = |\psi_1\rangle$. Hence the eigenvalues of P and P' relate to each other leading to

$$PP' = P'P = P' \text{ if } P \geqslant P'. \quad (12)$$

Projectors (generally noncommuting) support logical operations [29]. *Negation* $P^{\perp} = I - P$ is a projector that has zeros (ones) whenever P has ones (zeros). *Conjunction* $P \wedge Q$ contains only those vectors that belong both to $\mathcal{S}_P$ and $\mathcal{S}_Q$. Thus $\mathcal{S}_{P\wedge Q} = \mathcal{S}_Q \cap \mathcal{S}_P$. *Disjunction* $P \vee Q$ cannot be defined via $\mathcal{S}_Q \cup \mathcal{S}_P$ (set-theoretic union), because the latter is not a Hilbert (linear) space. The minimal Hilbert space that contains $\mathcal{S}_Q \cup \mathcal{S}_P$, is made of all linear combinations of the vectors from $\mathcal{S}_Q \cup \mathcal{S}_P$:

$$\mathcal{S}_{P\vee Q} = \{|\psi\rangle_P + |\psi\rangle_Q;\ |\psi\rangle_P \in \mathcal{S}_P, |\psi\rangle_Q \in \mathcal{S}_Q\}. \quad (13)$$

There are alternative representations [29]:

$$P \wedge Q = \max_R[\, R \,|\, R^2 = R,\, R \leqslant P,\, R \leqslant Q\,], \quad (14)$$

$$P \vee Q = \min_R[\, R \,|\, R^2 = R,\, R \geqslant P,\, R \geqslant Q\,], \quad (15)$$

where the maximization and minimization go over projectors R [29]. Indeed, if $R \leqslant P$, and $R \leqslant Q$ in (14), then due to (10) and (12), $\mathcal{S}_R$ is a subspace of both $\mathcal{S}_P$ and $\mathcal{S}_Q$. The maximal such subspace is $\mathcal{S}_{P\wedge Q}$. Likewise, if $R \geqslant P$, and $R \geqslant Q$, then $\mathcal{S}_R$ has to contain both $\mathcal{S}_P$ and $\mathcal{S}_Q$. The minimal such subspace is $\mathcal{S}_{P\vee Q}$ as (13) shows.

Now $P \vee Q = 0$ only if $P = Q = 0$, but $P \wedge Q$ can be zero also for nonzero P and Q; e.g., for nonzero P and Q in $\dim\mathcal{H} = 2$, we have either $P = Q$ or $P \wedge Q = 0$ (and then $P \vee Q = I$).

The above three operations are related with each other and with a limiting process [29]:

$$P \vee Q = (P^{\perp} \wedge Q^{\perp})^{\perp},\ P \wedge Q = \lim_{n\to\infty}(PQ)^n. \quad (16)$$

They are well known in quantum logics [29], but we shall employ them without a specific logical interpretation. Equations (14), (15) were generalized to non-negative operators [30]. For $[P,Q] \equiv PQ - QP = 0$ we have from (12)–(16) ordinary features of classical characteristic functions:

$$P \wedge Q = PQ, \quad P \vee Q = P + Q - PQ. \quad (17)$$

Imprecise classical probability generalizes the usual (precise) probabilities [32]: the measure of uncertainty for an event E is 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. Now $\underline{p}(E)$ [resp. $1 - \overline{p}(E)$] is a measure of a sure evidence in favor (resp. 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*) if

$$\underline{p}'(E) \leqslant \underline{p}(E) \text{ and } \overline{p}'(E) \geqslant \overline{p}(E). \quad (18)$$

Every $[\underline{p}(E),\overline{p}(E)]$ is consistent with $\underline{p}'(E) = 0$, $\overline{p}'(E) = 1$. This noninformative situation is not described by the usual theory that inadequately offers for it the homogeneous probability [32]. Now $[\underline{p}(E),\overline{p}(E)]$ does not imply that there is an explicit (but possibly unknown) precise probability for E located between $\underline{p}(E)$ and $\overline{p}(E)$ [31].

There are various imprecise classical probability theories [32], from a rather weak structures called upper and lower measures in Ref. [33] to the Dempster-Shafer imprecise probability [34,35]. The latter has numerous applications e.g., in decision making and artificial intelligence [32]. Recently it was applied for describing aspects of the Bell's inequality [36–38]. The quantum imprecise probability is to be sought independently, along the physical arguments. Below we recall how it is determined.

Imprecise joint quantum probability is sought for two noncommuting projectors P and Q. We look for upper $\overline{\omega}(P,Q)$ and lower $\underline{\omega}(P,Q)$ non-negative probability operators. The respective upper and lower probabilities in a state with density matrix ρ are given by Born's rule:

$$\overline{p}(P,Q) = \text{tr}[\rho\,\overline{\omega}(P,Q)], \quad \underline{p}(P,Q) = \text{tr}[\rho\,\underline{\omega}(P,Q)]. \quad (19)$$

The linearity of $\overline{p}(P,Q)$ and $\underline{p}(P,Q)$ over ρ can be motivated as in (2) above. We determine $\overline{\omega}(P,Q)$ and $\underline{\omega}(P,Q)$ from the following conditions [39]:

$$0 \leqslant \underline{\omega}(P,Q) = \underline{\omega}(Q,P) \leqslant \overline{\omega}(P,Q) = \overline{\omega}(Q,P) \leqslant I. \quad (20)$$

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

$$\underline{\omega}(P,Q) = \overline{\omega}(P,Q) = PQ \quad \text{if} \quad [P,Q] = 0. \quad (22)$$

$$\text{tr}[\rho\,\underline{\omega}(P,Q)] \leqslant \text{tr}(\rho\,PQ) \leqslant \text{tr}[\rho\,\overline{\omega}(P,Q)] \quad \text{if (9) holds.} \quad (23)$$

Equations (20) and (19) force $0 \leqslant \underline{p}(P,Q) \leqslant \overline{p}(P,Q) \leqslant 1$ for any ρ. Equation (20) also demands symmetry with respect to P and Q that is natural for the joint probability.

Now $\underline{\omega}(P,Q)$ and $\overline{\omega}(P,Q)$ are noncontextual in the sense that they depend only on P and Q. Even a stronger feature holds: Eq. (21) shows that both $\underline{\omega}(P,Q)$ and $\overline{\omega}(P,Q)$ can be measured together with either P or Q; see also (25). Hence the imprecise probability of two noncommuting observables does not lead to additional noncommutativity [40].

For $[P,Q] = 0$ we revert to the usual joint probability; see (22). For $Q = I$ we get from Eqs. (22) and (19) the marginal and precise probability of P. Equations (23) and (9) also refer to the consistency with the precise probability [cf. (18)], because the latter is well defined not only for $[P,Q] = 0$, but also under conditions (9), where it amounts to $\text{tr}(\rho PQ)$.

Equations (20)–(23) suffice for deducing [39]

$$\underline{\omega}(P,Q) = P \wedge Q, \quad \overline{\omega}(P,Q) = P \vee Q - (P - Q)^2, \quad (24)$$

where $\underline{\omega}(P,Q)$ and $\overline{\omega}(P,Q)$ are (resp.) the largest and the smallest positive operators holding (20)–(23). Now $\underline{\omega}(P,Q)$ is a projector, while $\overline{\omega}(P,Q)$ is generally just a non-negative operator. For $[P,Q] = 0$, we have from (17) and (24): $\underline{\omega}(P,Q) = \overline{\omega}(P,Q) = PQ$, as required by (22).

Equations (24) imply (21), because–as follows from (14) and (15) and checked directly–$P \wedge Q$, $P \vee Q$, and $(P - Q)^2$ commute with each other and with P and Q. Hence

$$[\,\overline{\omega}(P,Q), \underline{\omega}(P,Q)\,] = 0. \quad (25)$$

The origin of (24) is understood from (20) and (21) and (14) and (15), i.e., $P \wedge Q$ and $P \vee Q$ qualify as certain (resp.) lower and upper probability operators, while the factor $(P - Q)^2$ in (24) is needed to ensure (22).

In (19) we stress that if we would search for imprecise probability without assuming the linear dependence on ρ [but still assuming the analogues of (20)–(23)], we can obtain only more precise [in the sense of (18)] quantities than $\overline{p}(P,Q)$ and $\underline{p}(P,Q)$ in (19). The same argument is given for (21): relaxing it (but keeping other features) will lead to more precise probability. Thus conclusions obtained via linear $\overline{p}(P,Q)$ and $\underline{p}(P,Q)$ will stay intact.

A geometric feature of (24) is that both PQP (i.e., the restriction of Q into $\mathcal{S}_P$) and QPQ hold:

$$\underline{\omega}(P,Q) \leqslant PQP,\ QPQ\ \leqslant \overline{\omega}(P,Q). \tag{26}$$

Now $\underline{\omega}(P,Q) \leqslant PQP$ is shown from $P \wedge Q \leqslant Q$ [see (14)], which implies $P \wedge Q = P(P \wedge Q)P \leqslant PQP$. And $PQP \leqslant \overline{\omega}(P,Q)$ follows from $\overline{\omega}(P,Q) - PQP = \overline{\omega}(P,Q) - P\overline{\omega}(P,Q)P = \overline{\omega}(P,Q)(I-P) \geqslant 0$ recalling that $[\overline{\omega}(P,Q),P] = 0$ from (21) and (24).

Equations (20) and (23) can be deduced from (24) and (26), which also imply a version of sub- and superadditivity:

$$\sum\nolimits_a \overline{\omega}(P_a,Q) \geqslant Q \geqslant \sum\nolimits_a \underline{\omega}(P_a,Q), \quad \sum\nolimits_a P_a = I. \tag{27}$$

Thus the additive marginalization leads to an upper bound $\sum_a \overline{\omega}(P_a,Q)$ [and lower bound $\sum_a \underline{\omega}(P_a,Q)$] for the correct marginal probability $\overline{\omega}(I,Q) = \underline{\omega}(I,Q)$. We also note from (27) that non-negative operators $\overline{\omega}(P_a,Q_b)$ and $\underline{\omega}(P_a,Q_b)$ do not hold a semispectra resolution, e.g., $\sum_{a,b} \overline{\omega}(P_a,Q_b) \geqslant I$, hence they cannot be interpreted via generalized measurements [1,2]. Note as well that the monotonicity does not hold: $\overline{\omega}(P,Q) \nleqslant \overline{\omega}(I,Q) = Q$, though $P \leqslant I$; cf. (30).

Consistency with quantum conditional probabilities: Two-time (conditional) quantum probabilities are defined as follows [1]: first, measure P and assume the validity of the projection postulate. Now the result $P = 1$ implies the postmeasurement density matrix $P\rho P/\mathrm{tr}(\rho P)$. Then measuring Q leads to conditional probability $p_{Q=1|P=1} = \mathrm{tr}(\rho PQP)/\mathrm{tr}(\rho P)$ for the $Q = 1$. Likewise, we obtain $p_{P=1|Q=1} = \mathrm{tr}(\rho QPQ)/\mathrm{tr}(\rho Q)$ when measuring first Q and then P. The two-time probabilities are not usual conditional probabilities, since they do not lead to joint probabilities, e.g., applying the usual formulas does not generally lead to unique results due to $p_{P=1|Q=1}\ \mathrm{tr}(\rho Q) = \mathrm{tr}(\rho QPQ) \neq \mathrm{tr}(\rho PQP)$. However, imprecise joint probabilities (19) can lead to conditional imprecise probabilities. They are defined via the usual formulas, because the marginal [i.e., $\mathrm{tr}(\rho P)$ and $\mathrm{tr}(\rho Q)$] probabilities are precise. Equations (26) show that two-time probabilities are bound by the conditional imprecise probability, e.g.,

$$\frac{\mathrm{tr}[\rho\underline{\omega}(P,Q)]}{\mathrm{tr}(\rho P)} \leqslant \frac{\mathrm{tr}(\rho PQP)}{\mathrm{tr}(\rho P)} \leqslant \frac{\mathrm{tr}[\rho\overline{\omega}(P,Q)]}{\mathrm{tr}(\rho P)}. \tag{28}$$

Inconsistency with precise joint probabilities: We saw above that imprecise probabilities (19) are consistent [in the sense of (18)] with all instances, where quantum mechanics provides reasonable definitions of joint or conditional probability; cf. (23) and (28). Hence we *assume* that the reasonable definition of precise quantum joint probability should also be consistent with (19).

Thus we study a joint probability p_{ab} under two conditions. First, we require (1). Second, we demand that p_{ab} is consistent with (24) in the sense of (18) for any density matrix ρ and all a and b:

$$0 \leqslant \mathrm{tr}[\rho\,\underline{\omega}(P_a,Q_b)] \leqslant p_{ab} \leqslant \mathrm{tr}[\rho\,\overline{\omega}(P_a,Q_b)]. \tag{29}$$

We do not demand that p_{ab} depends only on P_a, Q_b; i.e., for p_{ab} we allow contextuality. We also do not demand that its dependence on ρ is linear. Since for p_{ab} we require $p_{ab} \geqslant 0$ for all ρ, quasiprobabilities are naturally excluded from consideration. Any theory that generalizes quantum mechanics and predicts joint probability for any preparation will be constrained by (29). Note that (29) is not stronger or weaker than (8) and (9).

Let two noncommuting observables A and B with (resp.) eigenprojectors $P_1 + P_2 = I$ and $Q_1 + Q_2 = I$ live in a two-dimensional Hilbert space. Due to $P_a \vee Q_b = I$ and $P_a \wedge Q_b = 0$, Eqs. (24) simplify as

$$\underline{\omega}(P_a,Q_b) = 0, \quad \overline{\omega}(P_a,Q_b) = \mathrm{tr}(P_aQ_b), \tag{30}$$

i.e., both probability operators reduce to numbers [41].

Our main result is that for given noncommutative A and B, there is a density matrix that violates (29) or (1). In this sense the precise joint probability does not exist already in $\dim\mathcal{H} = 2$. Indeed, take $p_{22} \leqslant \mathrm{tr}(P_2Q_2) = \mathrm{tr}(P_1Q_1)$ from (29), (30), $p_{21} \leqslant \mathrm{tr}(\rho Q_1)$ from (1), and employ them in $p_{22} + p_{21} = 1 - \mathrm{tr}(\rho P_1)$ from (1):

$$\mathrm{tr}(P_1Q_1) + \mathrm{tr}(\rho P_1) + \mathrm{tr}(\rho Q_1) - 1 \geqslant 0. \tag{31}$$

For given P_1 and Q_1, there is a density matrix ρ for which the left-hand-side of (31) is negative. Indeed, its positivity amounts in the Bloch representation (5) to $\vec{\beta}_{P_1}\vec{\beta}_{Q_1} + \vec{\beta}_\rho(\vec{\beta}_{P_1} + \vec{\beta}_{Q_1}) \geqslant -1$, and it can be violated e.g., as follows. If $\vec{\beta}_{P_1}\vec{\beta}_{Q_1} = \cos\alpha$, then we can choose $\vec{\beta}_\rho \to 1$ and $\vec{\beta}_\rho\vec{\beta}_{P_1} = \vec{\beta}_\rho\vec{\beta}_{Q_1} = -\cos\frac{\alpha}{2}$ producing $\cos\alpha - 2\cos\frac{\alpha}{2} < -1$ for $0 < \alpha < \pi$, i.e., for those values of α, where $[A,B] \neq 0$.

Thus, no precise joint probability is consistent with the quantum imprecise probability for all states.

Summary and outlook: We defined a setup of searching for a joint probability for two noncommuting observables. This is the next problem after the Born's probability for a single observable. An open aspect of this problem was formulated as a conjecture. We show that there is no joint probability for two noncommuting observables in a two-dimensional Hilbert space, if should this probability apply to any state and should be consistent with the quantum imprecise probability. This statement does not apply, if measurement contexts are introduced. Now we look for the joint probability generalizing out condition (1). Redefine the Born's probabilities as conditional ones $p_{a|P_a} = \mathrm{tr}(\rho P_a)$ and $p_{b|Q_b} = \mathrm{tr}(\rho Q_b)$, where conditioning can account for different devices needed to measure P_a and Q_b. We cannot deduce $p_{a|P_a}$ and $p_{b|Q_b}$ from a single joint probability, i.e., marginality condition (1) does not apply anymore, and is generalized as follows. We postulate two different joint probabilities $p_{ab|P_a}$ and $p_{ab|Q_b}$ holding $\sum_a p_{ab|Q_b} = \mathrm{tr}(\rho Q_b)$ and $\sum_b p_{ab|P_a} = \mathrm{tr}(\rho P_a)$, where $\sum_b p_{ab|Q_b} = p_{a|Q_b}$ and $\sum_a p_{ab|P_a} = p_{b|P_a}$ are well defined, but they need not be given by Born's formulas. Equations (21) and (25) mean that $\underline{\omega}(P_a, Q_b)$ and $\overline{\omega}(P_a, Q_b)$ can be measured simultaneously with each other and with either P_a or Q_b. Hence they can be used in both contexts P_a and Q_b. We constrain $p_{ab|P_a}$ and $p_{ab|Q_b}$ demanding their consistency with (19) for any ρ [cf. (30) and (29)]: $\mathrm{tr}(\rho\, \underline{\omega}(P_a Q_b)) \leqslant p_{ab|P_a}$, $p_{ab|Q_b} \leqslant \mathrm{tr}[\rho\, \overline{\omega}(P_a Q_b)]$. Together with generalized marginality these conditions define a set of probabilities that contains $\hat{p}_{ab|P_a} = \mathrm{tr}(\rho P_a Q_b P_a)$ and $\hat{p}_{ab|Q_b} = \mathrm{tr}(\rho Q_b P_a Q_b)$; cf. (26). Elsewhere we shall explore this approach, noting that descriptions via sets of probabilities are well known in mathematical statistics [32].

[1] W. M. de Muynck, *Foundations of Quantum Mechanics: An Empiricist Approach* (Kluwer, Dordrecht, 2002).
[2] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer, Berlin, 1996).
[3] A. E. Allahverdyan, R. Balian and T. M. Nieuwenhuizen, Phys. Rep. **525**, 1 (2013).
[4] D. Jennings and M. Leifer, Contemp. Phys. **57**, 60 (2016).
[5] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[6] A. Fine, J. Math. Phys. **23**, 1306 (1982).
[7] A. Fine, Brit. J. Phil. Soc. **24**, 1 (1973).
[8] S. Bugajski, Z. Naturforsch. A **33**, 1383 (1978).
[9] M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, Phys. Rep. **106**, 121 (1984).
[10] C. Ferrie, Rep. Prog. Phys. **74**, 116001 (2011).
[11] L. Accardi and A. Fedullo, Lett. Nuovo Cimento **34**, 161 (1982).
[12] S. P. Gudder, Found. Phys. **14**, 997 (1984).
[13] J. D. Malley, Phys. Rev. A **69**, 022118 (2004).
[14] F. J. Tipler, Proc. Natl. Acad. Sci. USA **111**, 11281 (2014).
[15] R. B. Griffiths, Found. Phys. **41**, 705 (2011).
[16] L. E. Szabo, Found. Phys. Lett. **8**, 417 (1995); G. Bana and T. Durt, *ibid.* **27**, 1355 (1997); T. M. Nieuwenhuizen, AIP Conf. Proc. **1101**, 127 (2009); D. A. Slavnov, Theor. Math. Phys. **136**, 1273 (2003).
[17] W. Mückenheim, Phys. Rep. **133**, 337 (1986); J. Maddox, Nature (London) **320**, 481 (1986); D. E. Parry, *ibid.* **321**, 644 (1986); W. Mückenheim, *ibid.* **324**, 307 (1986).
[18] L. Cohen, *Time Frequency Analysis: Theory and Applications* (Prentice Hall, Upper Saddle River, New Jersey, 1995).
[19] L. Cohen and Y. I. Zaparovanny, J. Math. Phys. **21**, 794 (1980).
[20] A. E. Allahverdyan, Phys. Rev. E **90**, 032137 (2014).
[21] M. Lostaglio, Phys. Rev. Lett. **120**, 040602 (2018).
[22] J. G. Kirkwood, Phys. Rev. **44**, 31 (1933); Y. P. Terletsky, Journ. Exper. Theor. Phys. **7**, 1290 (1937); P. A. M. Dirac, Rev. Mod. Phys. **17**, 195 (1945); H. Margenau and R. N. Hill, Prog. Theor. Phys. **26**, 722 (1961); A. O. Barut, Phys. Rev. **108**, 565 (1957).
[23] J. B. Hartle, Phys. Rev. A **70**, 022104 (2004); C. Bamber and J. S. Lundeen, Phys. Rev. Lett. **112**, 070405 (2014); L. M. Johansen, arXiv:0804.4379.
[24] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[25] S. P. Gudder, J. Math. Phys. **11**, 431 (1970).
[26] E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).
[27] A. E. Allahverdyan, R. Balian and T. M. Nieuwenhuizen, Phys. Rev. Lett. **92**, 120402 (2004).
[28] J. Uffink, Int. J. Theor. Phys. **33**, 199 (1994).
[29] J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, 1968).
[30] T. Moreland and S. Gudder, Lin. Alg. Appl. **286**, 1 (1999); H. Du, C. Deng, and Q. Li, Sci. China A **49**, 545 (2006).
[31] A. Papamarcou and T. L. Fine, Ann. Prob. **14**, 710 (1986).
[32] T. Augustin, F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes, editors, *Introduction to Imprecise Probabilities* (John Wiley & Sons, New York, 2014).
[33] P. Suppes and M. Zanotti, Erkenntnis **31**, 323 (1989).
[34] A. P. Dempster, Ann. Math. Stat. **38**, 325 (1967).
[35] G. Shafer, *The Mathematical Theory of Evidence* (Princeton University Press, Princeton, 1976).
[36] P. Suppes and M. Zanotti, Found. Phys. **21**, 1479 (1991).
[37] A. Vourdas, J. Math. Phys. **55**, 082107 (2014).
[38] B. H. Feintzeig and S. C. Fletcher, Found. Phys. **47**, 294 (2017)
[39] A. E. Allahverdyan, New J. Phys. **17**, 085005 (2015).
[40] Another possibility to require a joint measurability is to insist on $\underline{\omega}(P,Q) + \overline{\omega}(P,Q) + P \leqslant 1$. Then the averages over ρ of $\underline{\omega}(P,Q)$, $\overline{\omega}(P,Q)$, and P can be found via a single POVM measurement [1]. This condition cannot work already for $PQ = QP$, as seen from (21).
[41] For $P \to Q$, $\overline{\omega}(P,Q)$ jumps from $\approx I$ to P for $P = Q$; cf. (22). A similar jumping from 0 to P is seen for $\underline{\omega}(P,Q)$; see (30). Such jumps present no problems, since $[\underline{p} = 0, \overline{p} = 1]$ is consistent with any other probability.