

A General Framework of Quasi-probabilities and the Statistical Behaviour of Non-commuting Quantum Observables

Jaeha Lee and Izumi Tsutsui

Abstract We propose a general framework of the quantum/quasi-classical transformations by introducing the concept of quasi-joint-spectral distribution (QJSD). Specifically, we show that the QJSDs uniquely yield various pairs of quantum/quasi-classical transformations, including the Wigner-Weyl transform. We also discuss the statistical behaviour of combinations of generally non-commuting quantum observables by introducing the concept of quantum correlations and conditional expectations defined analogously to the classical counterpart. Based on these, Aharonov's weak value is given a statistical interpretation as one realisation of the quantum conditional expectations furnished in our formalism.

1 Introduction

Since the advent of quantum theory founded nearly a century ago, non-commutativity of quantum observables has undoubtedly been in the centrepiece of the theory marking its departure from classical theory. The hallmark of this is Heisenberg's uncertainty relation [1], which has later been elaborated from operational viewpoints by taking account of the measurement device by Ozawa [2, 3]. At the same time, the non-commutativity has been one of the major sources of troubles we face when we try to interpret their measurement outcomes in a sensible manner. This has naturally led to various attempts of 'quantisation' of classical systems, most notably in terms of non-commuting Hilbert space operators, or conversely of 'quasi-classical'

Jaeha Lee

National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, e-mail: jlee@post.kek.jp

Izumi Tsutsui

Theory Center, Institute of Particle and Nuclear Studies, High Energy Accelerator Research Organization (KEK), 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan, e-mail: izumi.tsutsui@kek.jp

interpretation of quantum systems in terms of commuting quantities familiar to us in classical theory.

The study on quantum and quasi-classical transformations has a long history dating back to the early days of quantum mechanics. Wigner and Weyl were among the prominent figures who have made much contribution in this effort bearing the theory of Wigner-Weyl transform [4, 5]. Historically, however, all these contributions in this area have been made more or less in a heuristic manner, and apparently their systematic treatment is still underdeveloped, not to mention a transparent overview of the relations among the various proposals of the transformations made so far.

On the other hand, in recent years we have witnessed the rise of interest in an issue related, at the roots, to the interpretation of measurement outcomes under non-commutativity. It is the novel quantity called the *weak value*,

$$A_w := \frac{\langle \psi', A \psi \rangle}{\langle \psi', \psi \rangle} \quad (1)$$

which has been proposed by Aharonov and co-workers [6] based on their time-symmetric formulation of quantum mechanics [7]. The weak value is a physical quantity that characterises the value of the observable A in the process specified by an initial (pre-selected) state $|\psi\rangle$ and a final (post-selected) state $|\psi'\rangle$. Unlike the standard measurement outcomes given by one of the eigenvalues of an observable A obtained in an ideal measurement, the weak value admits a definite value and is considered to be meaningful even for a set of non-commutable observables. The relation between the weak value and the quasi-classical transformations has been argued earlier, specifically with the Kirkwood-Dirac distribution [8, 9].

One of the aims of our present paper, expounded in Section 2, is to propose a general framework of the quantum/quasi-classical transformations by introducing the concept of *quasi-joint-spectral distribution* (QJSD). Specifically, we show that the QJSDs, of which definition shortly follows, uniquely yield various pairs of quantum/quasi-classical transformations, and that notable previous proposals of the transformations belong to this framework as special cases. Another aim, to which Section 3 is devoted, is to discuss the statistical behaviour of combinations of *generally non-commuting* quantum observables. Specifically, we introduce the concept of quantum correlations and conditional expectations, which are defined in analogue to the classical counterpart, and see how these concepts play together. Based on these, we finally endow Aharonov's weak value with a statistical interpretation as one realisation of the quantum conditional expectations furnished in our formalism.

Mathematical Notations Employed

Throughout this paper, we denote by \mathbb{K} either the real field \mathbb{R} or the complex field \mathbb{C} . Since our primary interest is on quantum mechanics, Hilbert spaces are always assumed to be complex. Conforming to the convention in physical literature, we denote the complex conjugate of a complex number $c \in \mathbb{C}$ by c^* , and an inner product

$\langle \cdot, \cdot \rangle$ defined on a complex linear space is anti-linear in its first argument and linear in the second. For simplicity, we adopt the natural units where we specifically have $\hbar = 1$, unless stated otherwise.

2 Quantisation and Quasi-classicalisation via QJSDs

For commuting quantum observables, a ‘trivial’ method of quantum and quasi-classical transformation is available, where the former is known as the functional calculus whereas the latter is known as the Born rule. These maps are both known to be characterised by the joint-spectral measure (JSM) of the observables concerned, and they are understood to be adjoint operations to each other. On the other hand, the problem becomes non-trivial when non-commuting observables are put in to consideration, primarily due to the lack of the JSM.

In this section, we first propose a novel approach to the problem of quantum/quasi-classical transformations by introducing the concept of quasi-joint-spectral distributions (QJSDs), which are intended as non-commuting generalisations to the JSMs of commuting observables. Just as the JSM induces a unique adjoint pair of quantum/quasi-classical transformation for commuting observables, QJSDs induce various adjoint pairs for non-commuting observables. Specifically, we see that there exists inherent indefiniteness in the possible definition of QJSDs, each leading to different possible transformations, in which the Wigner-Weyl transform belongs as a special case.

2.1 Preliminary Observations

As a prelude to our study, we first review some basic facts in quantum theory regarding the spectral theorem of self-adjoint operators, the functional calculus and the Born rule. In what follows, we consider a finite combination of simultaneously measurable observables, and observe that both the functional calculus and the Born rule can respectively be understood as the trivial realisation of *quantisation* and *quasi-classicalisation*, in the sense that the former allows us to map real functions (*i.e.*, classical observables) to self-adjoint Hilbert space operators (*i.e.*, quantum observables), and the latter defines a map from density operators (*i.e.*, quantum states) to probability distributions (*i.e.*, classical states). We then point out that the functional calculus and the Born rule are *adjoint* notions to each other. These rather trivial observations shall be the guiding line for our further study in considering the non-trivial problem of quantisation and quasi-classicalisation for the general case involving combination of non-commuting quantum observables.

2.1.1 Spectral Theorem for Self-adjoint Operators

A basic but important result of functional analysis (*i.e.*, linear algebra for arbitrary-dimensions) regarding self-adjoint operators is the *spectral theorem*, which states that for a self-adjoint operator A on a Hilbert space \mathcal{H} , there corresponds a unique *spectral measure* E_A such that

$$A = \int_{\mathbb{R}} a dE_A(a) \quad (2)$$

holds. In simple terms, the spectral measure

$$E_A(a) = \begin{cases} 0 & (a \text{ is not an eigenvalue of } A) \\ P_a & (a \text{ is an eigenvalue of } A) \end{cases}, \quad a \in \mathbb{R} \quad (3)$$

of A is a map from the eigenvalues a of A to the orthogonal projections P_a on the corresponding eigenspaces. For the case where the eigenvalues are all non-degenerate, the spectral measure is simply nothing but $E_A(a) = P_a = |a\rangle\langle a|$, where $|a\rangle\langle a|$ is the orthogonal projection on the 1-dimensional subspace of \mathcal{H} spanned by the eigenspace corresponding to the eigenvalue a . In this case, the integral (2) formally reduces to the familiar form

$$A = \int_{\mathbb{R}} a |a\rangle\langle a| da. \quad (4)$$

For the case in which the Hilbert space \mathcal{H} under consideration is moreover finite-dimensional, the spectral theorem is nothing but the eigendecomposition theorem

$$A = \sum_{i=1}^n a_i |a_i\rangle\langle a_i|$$

valid for Hermitian matrices A , and the spectral measure $E_A(a_i) = |a_i\rangle\langle a_i|$ reduces to the collection of orthogonal projections corresponding to the eigenvectors a_i of A . Simply put, spectral theorem is thus a generalisation of the eigendecomposition theorem for the infinite dimensional case, and the spectral measures are in turn the generalisation of orthogonal projections onto the corresponding eigenspaces. The primary advantage of this generalisation becomes apparent when infinite dimensional Hilbert spaces must be taken into consideration¹.

Joint-spectral Measures

Now, suppose one is given an ordered combination

¹ A self-adjoint operator defined on infinite dimensional Hilbert spaces may sometimes fail to have any eigenvalues in the sense that $A|\psi\rangle = a|\psi\rangle$ holds for some non-zero vector $|\psi\rangle \in \mathcal{H}$, as most famously exemplified by the position operator \hat{x} and the momentum operator \hat{p} of a free particle.

$$\mathbf{A} := (A_1, \dots, A_n), \quad 1 \leq n < \infty \quad (5)$$

of a finite number of pairwise strongly commuting² distinct self-adjoint operators (*i.e.*, simultaneously measurable quantum observables) on \mathcal{H} . An important fact regarding strongly commuting self-adjoint operators is that, one may uniquely construct the *joint-spectral measure* (JSM) of the combination

$$E_{\mathbf{A}}(\mathbf{a}) = \prod_{i=1}^n E_{A_i}(a_i), \quad \mathbf{a} := (a_1, \dots, a_n) \in \mathbb{R}^n \quad (6)$$

that fully describes their joint behaviour, where each E_{A_i} is the unique spectral measure corresponding to A_i , ($1 \leq i \leq n$). One then trivially has

$$A_i = \int_{\mathbb{R}^n} a_i dE_{\mathbf{A}}(\mathbf{a}), \quad 1 \leq i \leq n, \quad (7)$$

if one is to reclaim the original self-adjoint operators.

2.1.2 Functional Calculus

An important fact regarding spectral measures is that it induces a map from the space of functions to Hilbert space operators. Indeed, under the same situation as above, the JSM of the ordered combination (5) induces a map that maps a function f defined on \mathbb{R}^n to the operator

$$f_{E_{\mathbf{A}}} := \int_{\mathbb{R}^n} f(\mathbf{a}) dE_{\mathbf{A}}(\mathbf{a}) \quad (8)$$

on \mathcal{H} . The map $f \mapsto f_{E_{\mathbf{A}}}$ is one realisation of the *functional calculus*, which is a general term that points to a map from functions to operators satisfying certain algebraic properties. In fact, under some appropriate conditions, one finds that there is a one-to-one correspondence between functional calculi and spectral measures. In our context, we view the functional calculus as the trivial way to quantise classical observables (*i.e.*, real functions) into quantum observables (*i.e.*, self-adjoint operators). In what follows, we may occasionally write the image of the functional calculus by either of the following notations: $f_{E_{\mathbf{A}}} = f_{\mathbf{A}} = f(\mathbf{A}) = f(A_1, \dots, A_n)$.

² We say that a pair of self-adjoint operators A and B strongly commutes, if and only if they commute in the level of spectral measures. In a laxer notation, this is to say that

$$E_A(a)E_B(b) = E_B(b)E_A(a), \quad a, b \in \mathbb{R}$$

holds. Strictly speaking, strong commutativity is generally stronger than mere commutativity when unbounded operators are concerned, but we do not intend to delve into the intricacies, which are not essential for our discussion.

2.1.3 Born Rule

The Born rule is the corner stone of the probabilistic interpretation of quantum measurements, which states that, given a density operator (*i.e.*, mixed quantum state) ρ on \mathcal{H} and a combination (5) of simultaneously measurable observables, the joint behaviour of the measurement outcomes is described by the *joint-probability distribution*

$$\rho_{E_{\mathbf{A}}}(\mathbf{a}) := \text{Tr}[E_{\mathbf{A}}(\mathbf{a})\rho] \quad (9)$$

defined for the combination of simultaneously measurable observables concerned on the state. In our context, we view the Born rule as the trivial realisation of quasi-classicalisation of quantum states (*i.e.*, density operators) into classical states (*i.e.*, probability distributions). In what follows, we occasionally denote the resulting probability distribution by either of the following notations: $\rho_{E_{\mathbf{A}}} = \rho_{\mathbf{A}} = \rho_{(A_1, \dots, A_n)}$.

2.1.4 Adjointness of the Functional Calculus and the Born Rule

An important observation we point out here that is crucial for our further discussion is that, both quantisation (functional calculus) and quasi-classicalisation (Born rule) are *adjoint* notions.

Dual Pair

To see this point, we first prepare some necessary terminologies and notations. Let $L(\mathcal{H})$ denote the space of all bounded linear operators on the Hilbert space \mathcal{H} , and let $N(\mathcal{H})$ denote the space of all nuclear operators (or, better known as trace-class operators) on \mathcal{H} . Bounded quantum observables $A \in L(\mathcal{H})$ and quantum states $\rho \in N(\mathcal{H})$ belong to the respective spaces. On the product space $L(\mathcal{H}) \times N(\mathcal{H})$ is defined a bilinear form

$$\langle X, N \rangle_Q := \text{Tr}[XN], \quad X \in L(\mathcal{H}), N \in N(\mathcal{H}) \quad (10)$$

that maps a pair of bounded linear operator and a nuclear operator to the trace of their product. In mathematics, a triple consisting of a pair of linear spaces X, Y and a bilinear form $\langle \cdot, \cdot \rangle : X \times Y \rightarrow \mathbb{K}$ satisfying the conditions

$$\begin{aligned} \forall x \in X \setminus \{0\}, \exists y \in Y \quad \langle x, y \rangle &\neq 0, \\ \forall y \in Y \setminus \{0\}, \exists x \in X \quad \langle x, y \rangle &\neq 0, \end{aligned} \quad (11)$$

is called a *dual pair*. The triple $(L(\mathcal{H}), N(\mathcal{H}), \langle \cdot, \cdot \rangle_Q)$ is one typical realisation of a dual pair, which are the familiar tools we use to describes quantum measurements.

On the other hand, let $\mathcal{S}(\mathbb{R}^n)$, $(1 \leq n < \infty)$ denote the *n*-dimensional *Schwartz space*, which is the space of all smooth functions that, even being multiplied by any polynomials after being differentiated arbitrarily many times, they ‘vanish at

infinity'. Also, let $\mathcal{S}'(\mathbb{R}^n)$ denote the continuous dual of the Schwartz space, which is called the space of *tempered distributions*. The space of tempered distributions is an extension of the familiar space of density functions or complex measures: every probability density function is a complex measure, and every complex measure is a tempered distribution, while the converses are not always true³. It reveals that the spaces of density functions or that of complex measures is not sufficient in properly handling quasi-probability distributions for non-commuting observables.

Now, if we allow ourselves some abuse of notation, we may formally treat a tempered distribution $\varphi \in \mathcal{S}'(\mathbb{R}^n)$ as a function $\varphi(\mathbf{x})$ on \mathbb{R}^n , and define the bilinear form

$$\langle f, \varphi \rangle_C := \int_{\mathbb{R}^n} f(\mathbf{x}) \varphi(\mathbf{x}) dm_n(\mathbf{x}), \quad f \in \mathcal{S}(\mathbb{R}^n), \varphi \in \mathcal{S}'(\mathbb{R}^n) \quad (12)$$

where we have introduced the renormalised Lebesgue-Borel measure

$$dm_n(\mathbf{x}) := (2\pi)^{-n/2} dx^n. \quad (13)$$

For brevity, we occasionally write $dm_1 = dm$ whenever there is no risk for confusion. This renormalisation is mostly of aesthetic purpose, whose advantage becomes apparent when we introduce the Fourier transformation later in our discussion. Equipped with the bilinear form, the triple $(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^n), \langle \cdot, \cdot \rangle_C)$ also qualifies as a dual pair that becomes a tool in describing classical measurements.

Adjointness of the Transformations

Given an ordered combination of simultaneously measurable quantum observables (5), let

$$\Phi_{E_A} : \mathcal{S}(\mathbb{R}^n) \rightarrow L(\mathcal{H}), f \mapsto f_{E_A} \quad (14)$$

denote the functional calculus⁴ of the ordered combination (5) of self-adjoint operators defined in (8), and in turn let

$$\Phi'_{E_A} : N(\mathcal{H}) \rightarrow \mathcal{S}'(\mathbb{R}^n), \rho \mapsto \rho_{E_A} \quad (15)$$

denote the Born rule⁵ described in (9). It can be demonstrated that both (14) and (15) are continuous linear maps between the respective spaces equipped with the usual

³ The space of density functions on \mathbb{R}^n is a proper subspace of the space of complex measures. An example of this is the delta measure, which is well-defined as a measure but not as a density function. The space of complex measures on \mathbb{R}^n is also a proper subspace of the space of tempered distributions. An example for this is the derivative of the delta measure, which is well-defined as a tempered distribution but not as a complex measure.

⁴ Note that the image of a Schwartz function under the functional calculus (14) is a bounded operator with the operator norm $\|f(A)\| \leq \sup_{x \in \mathbb{R}^n} |f(x)|$, hence the map is well-defined.

⁵ The image of a nuclear operator under the Born rule (15) belongs to the space of complex measures, which can be uniquely embedded into the space of tempered distributions. In this sense, we extend the codomain of the map (15) and understand their images to be tempered distributions, rather than complex measures, for later convenience.

topologies. One then readily observes by the following straightforward computation

$$\begin{aligned}
\langle \Phi_{E_{\mathbf{A}}}(f), \rho \rangle_Q &:= \text{Tr}[f E_{\mathbf{A}} \rho] \\
&= \int_{\mathbb{R}^n} f(\mathbf{a}) d\text{Tr}[E_{\mathbf{A}}(\mathbf{a}) \rho] \\
&= \int_{\mathbb{R}^n} f(\mathbf{a}) \rho_{E_{\mathbf{A}}}(\mathbf{a}) dm_n(\mathbf{a}) \\
&= \langle f, \Phi'_{E_{\mathbf{A}}}(\rho) \rangle_C, \quad f \in \mathcal{S}(\mathbb{R}^n), \rho \in N(\mathcal{H})
\end{aligned} \tag{16}$$

that the functional calculus (14) and the Born rule (15) are *adjoint maps* to each other. This relation can be illustrated by the following diagram:

$$\begin{array}{ccc}
L(\mathcal{H}) & \overset{\text{dual pair}}{\longleftrightarrow} & N(\mathcal{H}) \\
\uparrow \Phi_{E_{\mathbf{A}}} & & \downarrow \Phi'_{E_{\mathbf{A}}} \\
\mathcal{S}(\mathbb{R}^n) & \overset{\text{dual pair}}{\longleftrightarrow} & \mathcal{S}'(\mathbb{R}^n)
\end{array}$$

Quantisation (left vertical arrow), *Quasi-Classicalisation* (right vertical arrow)

Here, the top row denotes the dual pair of quantum observables and quantum states, whereas the bottom row depicts the classical counterpart. The left column consists of (quantum and classical) observables, whereas the right column consists of (quantum and classical) states.

2.2 Quantisation and Quasi-classicalisation via QJSDs

In the previous subsection, we have reviewed the very basics of the spectral theorem, the functional calculus and the Born rule defined for commuting observables, and have seen that the functional calculus (*i.e.*, quantisation) and the born rule (*i.e.*, quasi-classicalisation) are adjoint operations to each other. The next step is to generalise our whole arguments into the case for *non-commuting* observables.

2.2.1 Introducing Quasi-joint-spectral Distributions (QJSDs)

The key observation to make here is that, it was the JSM that uniquely gave rise to the adjoint pair of the desired maps. A straightforward idea for our current problem would be thus to introduce non-commuting analogues to the JSM.

Strong Commutativity in the Fourier Space

As a preparation to our further discussion, we review the characterisation for strong commutativity of spectral measures in their Fourier spaces. Let A be self-adjoint, and let E_A be its spectral measure. We call the operator valued function

$$(\mathcal{F}E_A)(s) := \int_{\mathbb{R}} e^{-isa} dE_A(a) = e^{-isA}, \quad s \in \mathbb{R}, \quad (17)$$

the *Fourier transform* of E_A . It is a basic fact of functional calculus that one may characterise the strong commutativity of self-adjoint operators by the Fourier transforms of their spectral measures: a pair of self-adjoint operators A and B strongly commutes if and only if the Fourier transforms of the respective spectral measures

$$e^{itA}e^{isB} = e^{isB}e^{itA}, \quad s, t \in \mathbb{R} \quad (18)$$

commute.

Fourier Transform of JSM

Let us first compute the Fourier transform of the JSMs. Given the JSM $E_{\mathbf{A}}$ of an ordered combination of strongly commuting self-adjoint operators, let

$$\begin{aligned} (\mathcal{F}E_{\mathbf{A}})(\mathbf{s}) &:= \int_{\mathbb{R}^n} e^{-i\langle \mathbf{s}, \mathbf{a} \rangle} dE_{\mathbf{A}}(\mathbf{a}) \\ &= e^{-i\langle \mathbf{s}, \mathbf{A} \rangle} \\ &= \prod_{i=1}^n e^{-is_i A_i}, \quad \mathbf{s} \in \mathbb{R}^n \end{aligned} \quad (19)$$

denote the *Fourier transform* of $E_{\mathbf{A}}$. Here, $\langle \mathbf{s}, \mathbf{a} \rangle := \sum_{i=1}^n s_i a_i$ denotes the standard inner product on \mathbb{R}^n , and $\langle \mathbf{s}, \mathbf{A} \rangle := \sum_{i=1}^n s_i A_i$. We also note that the last line of the above equality is due to the iterated application of the Lie-Trotter-Kato product formula.

Hashed Operators

From now on, we generalise the use of the notation \mathbf{A} so that it may also admit ordered combinations of generally *non-commuting* distinct self-adjoint operators.

To this end, we introduce, in a less formal way, the *hashed operator*

$$\hat{\#}_{\mathbf{A}}(\mathbf{s}) := \left\{ \begin{array}{l} \text{a suitable 'mixture' of the 'disintegrated'} \\ \text{components of the unitary groups } e^{-is_1 A_1}, \dots, e^{-is_n A_n} \end{array} \right\}, \quad \mathbf{s} \in \mathbb{K}^n \quad (20)$$

of the operators concerned. Here, by ‘disintegration’ we mean breaking each of the unitary operators $e^{-is_k A_k}$ into chunks of operator valued functions $T_{k, \iota_k}(s_k)$, ($\iota_k \in I_k$) such that

$$\prod_{\iota_k \in I_k} T_{k, \iota_k}(s_k) = e^{-is_k A_k} \quad (21)$$

holds. A straightforward example of this is provided by

$$T_{k, \iota_k}(s_k) := e^{-is_k \lambda_{\iota_k} A}, \quad (22)$$

where $\lambda_{\iota_k} \in \mathbb{R}$ are real numbers satisfying the normalisation $\sum_{\iota_k} \lambda_{\iota_k} = 1$. Then, by ‘mixture’ we imply that a hashed operator $\hat{\#}_{\mathbf{A}}(\mathbf{s})$ is constructed as a product of all components $T_{k, \iota_k}(s_k)$, ($\iota_k \in I_k, k = 1, \dots, n$) in arbitrary orders. We may also allow it to be given by convex combinations of two hashed operators $\hat{\#}_{\mathbf{A}}$ and $\hat{\#}'_{\mathbf{A}}$ as

$$\hat{\#}''_{\mathbf{A}}(\mathbf{s}) = \lambda \hat{\#}_{\mathbf{A}}(\mathbf{s}) + (1 - \lambda) \hat{\#}'_{\mathbf{A}}(\mathbf{s}). \quad (23)$$

Examples of the hashed operators pertaining to the simplest case $\mathbf{A} = (A, B)$ are thus given by

$$\hat{\#}_{\mathbf{A}}(s, t) = \begin{cases} e^{-itB} e^{-isA}, \\ e^{-isA} e^{-itB}, \\ \frac{1+\alpha}{2} \cdot e^{-itB} e^{-isA} + \frac{1-\alpha}{2} \cdot e^{-isA} e^{-itB}, & \alpha \in \mathbb{C} \\ \prod_{k=1}^N e^{-is \lambda_k A} e^{-it \mu_k B}, & (\sum_{k=1}^N \lambda_k = 1, \sum_{k=1}^N \mu_k = 1), \\ (e^{-isA/N} e^{-itB/N})^N, \\ e^{-i(sA+tB)} = \lim_{N \rightarrow \infty} (e^{-isA/N} e^{-itB/N})^N. \end{cases} \quad (24)$$

In general, either or both of the parameters s, t can be made to even admit complex numbers. One such example is given by

$$\hat{\#}_{\mathbf{A}}^{\kappa}(s, t) = e^{-i\langle \frac{1-\kappa}{2}, s \rangle A} e^{-itB} e^{-i\langle \frac{1+\kappa}{2}, s \rangle A}, \quad \kappa \in \mathbb{C}, s \in \mathbb{C}, t \in \mathbb{R}, \quad (25)$$

where $\langle \kappa, s \rangle = \kappa_1 s_1 + \kappa_2 s_2$ is understood as the standard inner product defined on \mathbb{R}^2 , where $\kappa = \kappa_1 + i\kappa_2$ ($\kappa_1, \kappa_2 \in \mathbb{R}$) and $s = s_1 + is_2$ ($s_1, s_2 \in \mathbb{R}$) are complex numbers identified as vectors on \mathbb{R}^2 .

The precise definition as to what types of construction one may allow for hashed operators will depend on the context and the properties that one would like them to retain. For example, the evaluation of the operator norm of the hashed operator

$$\|\hat{\#}_{\mathbf{A}}(\mathbf{s})\| = 1 \quad (26)$$

can be guaranteed if we restrict ourselves to the specific type of construction that only allows hashed operators to be defined in the form of products of (22), since in that case $\hat{\#}_{\mathbf{A}}(\mathbf{s})$ becomes a unitary operator. If we moreover consider their convex combinations (23), the equality in (26) generally becomes an inequality $\|\hat{\#}_{\mathbf{A}}(\mathbf{s})\| \leq$

1. If we moreover allow complex numbers to be chosen as λ in (23), the operator norm of $\hat{\#}_{\mathbf{A}}(\mathbf{s})$, while still being bounded, may exceed 1.

Hashed Operator of Commuting Observables

One readily realises that the hashed operators (20) are, while differing in their representations, unique *if and only if* the self-adjoint operators concerned are all simultaneously measurable. In that case, it is easy to see that the hashed operators all reduce to the same Fourier transform (19) of the JSM.

Quasi-joint-spectral Distributions

Under the same conditions as above, let us choose any hashing $\hat{\#}_{\mathbf{A}}$ introduced in (20), and introduce the *quasi-joint-spectral distribution*⁶ (QJSD) of the ordered pair \mathbf{A} defined by its inverse Fourier transform

$$\begin{aligned} \#_{\mathbf{A}}(\mathbf{a}) &:= (\mathcal{F}^{-1} \hat{\#}_{\mathbf{A}})(\mathbf{a}) \\ &:= \int_{\mathbb{K}^n} e^{i(\mathbf{a}, \mathbf{s})} \hat{\#}_{\mathbf{A}}(\mathbf{s}) dm_n(\mathbf{s}), \quad \mathbf{a} \in \mathbb{K}^n. \end{aligned} \quad (27)$$

Due to the bijectivity of the Fourier transformation, to each QJSD corresponds a unique hashed operator, and hence QJSDs are highly non-unique in the case a given ordered combination \mathbf{A} is non-commutative. The QJSD is unique if and only if \mathbf{A} admits simultaneous measurability, and in such case, the unique QJSD actually reduces to the JSM itself

$$\#_{\mathbf{A}} = E_{\mathbf{A}}. \quad (28)$$

By construction and the observations made above, one may surmise that the QJSDs serve as generalisations of the JSM to generally non-commuting observables. Indeed, QJSDs share some of the basic properties one finds in common with the standard JSM. The primary fact we mention is the normalisation property: the total integration of any QJSD reduces to the identity Id, as one readily finds through the following formal computation

$$\begin{aligned} \int_{\mathbb{K}^n} \#_{\mathbf{A}}(\mathbf{a}) dm_n(\mathbf{a}) &= \int_{\mathbb{K}^n} e^{-i(\mathbf{0}, \mathbf{a})} \#_{\mathbf{A}}(\mathbf{a}) dm_n(\mathbf{a}) \\ &= (\mathcal{F} \#_{\mathbf{A}})(\mathbf{0}) \\ &= \hat{\#}_{\mathbf{A}}(\mathbf{0}) = \text{Id}. \end{aligned} \quad (29)$$

⁶ The reason for our choice of the nomination quasi-joint-spectral *distributions*, rather than *measures*, lies in the fact that, in contrast to the JSMs, QJSDs does not necessarily lie in the space of operator valued measures (OVMS). In fact, we understand them as members of the operator valued distributions (OVDs), which is an operator analogue of generalised functions (distributions). The space of OVDs is larger than the space of OVMS, and the latter can be embedded into the former.

In fact, this is actually a corollary to a more stronger property regarding the marginals

$$\int_{\mathbb{K}} \#_{\mathbf{A}}(\mathbf{a}) dm(a_k) = \#_{\mathbf{A}_k}(a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n), \quad (30)$$

where $\mathbf{A}_k := (A_1, \dots, A_{k-1}, A_{k+1}, \dots, A_n)$, $1 \leq k \leq n$, denotes the ordered combination of the self-adjoint operators that lacks the k th component of the original ordered combination \mathbf{A} , and $\#_{\mathbf{A}_k}$ denotes the QJSD of \mathbf{A}_k defined by

$$\hat{\#}_{\mathbf{A}_k}(s_1, \dots, s_{k-1}, s_{k+1}, \dots, s_n) := \hat{\#}_{\mathbf{A}}(s_1, \dots, s_{k-1}, 0, s_{k+1}, \dots, s_n), \quad (31)$$

which corresponds to the hashing constructed by ‘taking away’ all the disintegrated components of the k th member $e^{-is_k A_k}$ from the original hashing $\hat{\#}_{\mathbf{A}}$. To see this, let H_k denote the l. h. s. of (30). The Fourier transform of H_k then reads

$$\begin{aligned} & (\mathcal{F}H_k)(s_1, \dots, s_{k-1}, s_{k+1}, \dots, s_n) \\ &= \int_{\mathbb{K}^{n-1}} \sum_{i=1}^{k-1} \sum_{i=k+1}^n e^{-is_i a_i} \left(\int_{\mathbb{K}} \#_{\mathbf{A}}(\mathbf{a}) dm(a_k) \right) dm_{n-1}(a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n) \\ &= \int_{\mathbb{K}^n} \sum_{i=1}^{k-1} \sum_{i=k+1}^n e^{-is_i a_i} e^{-i0 \cdot a_n} \#_{\mathbf{A}}(\mathbf{a}) dm_n(\mathbf{a}) \\ &= \hat{\#}_{\mathbf{A}}(s_1, \dots, s_{k-1}, 0, s_{k+1}, \dots, s_n) \\ &= (\mathcal{F}\hat{\#}_{\mathbf{A}_k})(s_1, \dots, s_{k-1}, s_{k+1}, \dots, s_n), \end{aligned} \quad (32)$$

and by the injectivity of the Fourier transformation, one concludes $H_k = \#_{\mathbf{A}_k}$.

Reclaiming the JSMs

A straightforward but important corollary to the above property is the following observation. Let

$$\mathbf{B} = (A_{i_1}, \dots, A_{i_k}) \quad (33)$$

be an order-preserving (*i.e.*, $1 \leq i_1 < \dots < i_k \leq n$) subset of \mathbf{A} consisting of k ($1 \leq k \leq n$) numbers of *pairwise strongly commuting* distinct members, and let

$$\mathbf{B}^c := (A_{j_1}, \dots, A_{j_{n-k}}) \quad (34)$$

denote its order-preserving (*i.e.*, $1 \leq j_1 < \dots < j_{n-k} \leq n$) complement consisting of $n - k$ numbers of the members of \mathbf{A} that do not belong to \mathbf{B} . Then, an iterated application of (30) leads to

$$E_{\mathbf{B}}(\mathbf{b}) = \int_{\mathbb{K}^{n-k}} \#_{\mathbf{A}}(\mathbf{a}) dm_{n-k}(\mathbf{b}^c), \quad (35)$$

where $\mathbf{b} := (a_{i_1}, \dots, a_{i_k})$ and $\mathbf{b}^c := (a_{j_1}, \dots, a_{j_{n-k}})$ denotes the variables corresponding to the respective order-preserving subsets⁷. This implies that, if one ‘integrate-outs’ all the variables corresponding to the complement \mathbf{B}^c , one may reclaim the authentic JSM of B .

2.2.2 Quantisation and Quasi-classicalisation

Now that we have constructed the QJSDs, which could be understood as non-commutative analogues to the standard JSM, we shall now embark on the construction of quantisation and quasi-classicalisation regarding combination of observables that may fail to be measured simultaneously.

Quantisation of classical Observables

For an ordered combination of (generally non-commuting) distinct quantum observables \mathbf{A} , let $\#_{\mathbf{A}}$ be any QJSD of one’s choice. Guided by a straightforward analogy of the functional calculus originally defined for the commutative case, we thus define the map

$$\Phi_{\#_{\mathbf{A}}} : f \mapsto f_{\#_{\mathbf{A}}} := \int_{\mathbb{K}^n} f(\mathbf{a}) \#_{\mathbf{A}}(\mathbf{a}) \, dm_n(\mathbf{a}) \quad (36)$$

that maps a Schwartz function $f \in \mathcal{S}(\mathbb{K}^n)$ to a bounded⁸ linear operator $f_{\#_{\mathbf{A}}} \in L(\mathcal{H})$. We call the map (36) the *quantisation* pertaining to the QJSD $\#_{\mathbf{A}}$, and in turn call the image $\Phi_{\#_{\mathbf{A}}}(f)$ the quantisation of f . Occasionally, we denote the quantisation of f by either of the following notations: $f_{\#_{\mathbf{A}}} = f(\#_{\mathbf{A}})$. We may sometimes

⁷ Here, we adopt the convention

$$\int_{\mathbb{K}^{n-k}} \#_{\mathbf{A}}(\mathbf{a}) \, dm_{n-k}(\mathbf{b}^c) = \#_{\mathbf{A}}(\mathbf{a})$$

for the case $k = n$.

⁸ The quantisation of a Schwartz function f is formally defined as a unique bounded map $f_{\#_{\mathbf{A}}}$ such that the equality

$$\mathrm{Tr}[f_{\#_{\mathbf{A}}}\rho] := \int_{\mathbb{R}^n} \check{f}(\mathbf{s}) \mathrm{Tr}[\hat{\#}_{\mathbf{A}}(\mathbf{s})\rho] \, dm_n(\mathbf{s})$$

holds for all $\rho \in N(\mathcal{H})$, where \check{f} denotes the inverse Fourier transform of f . The r. h. s of the above equation is well-defined, since

$$\left| \int_{\mathbb{R}^n} \check{f}(\mathbf{s}) \mathrm{Tr}[\hat{\#}_{\mathbf{A}}(\mathbf{s})\rho] \, dm_n(\mathbf{s}) \right| \leq \|\check{f}\|_1 \cdot M \|\rho\|_{\mathrm{nuk}}$$

is finite. Here, $\|\cdot\|_1$ and $\|\cdot\|_{\mathrm{nuk}}$ respectively denote the L^1 norm of integrable functions and the nuclear norm (*alias* trace norm) of nuclear operators, and $\|\hat{\#}_{\mathbf{A}}(\mathbf{s})\| \leq M$, $\mathbf{s} \in \mathbb{K}^n$ is the upper bound of the operator norm of the hashed operator. The existence and uniqueness of such an operator $f_{\#_{\mathbf{A}}} \in L(\mathcal{H})$ is due to the fact that the space of continuous linear functionals on $N(\mathcal{H})$ is isomorphic to the space of bounded operators $N(\mathcal{H})' \cong L(\mathcal{H})$. The continuity of the linear functional $\rho \mapsto \mathrm{Tr}[f_{\#_{\mathbf{A}}}\rho]$ follows directly from the above evaluation.

even omit \mathbf{A} and write $f_{\#}$, when the observables concerned are obvious from the context.

Quasi-classicalisation of quantum States

Conversely, we also allow ourselves to be guided by a straightforward analogy of the Born rule and intend to extend it to the non-commutative case. We thus define the map

$$\Phi'_{\#\mathbf{A}} : \rho \mapsto \rho_{\#\mathbf{A}}(\mathbf{a}) := \text{Tr}[\#\mathbf{A}(\mathbf{a})\rho], \quad \mathbf{a} \in \mathbb{K}^n \quad (37)$$

that maps a density operator $\rho \in N(\mathcal{H})$ to a tempered distribution $\rho_{\#\mathbf{A}} \in \mathcal{S}'(\mathbb{K}^n)$. We call the (image $\rho_{\#\mathbf{A}}$ of the) map (37) the *quasi-classicalisation* (of ρ) pertaining to the QJSD $\#\mathbf{A}$. Specifically, for a density operator $\rho \in N(\mathcal{H})$, *i.e.*, a positive nuclear operator with the normalisation condition $\text{Tr}[\rho] = 1$, we occasionally call the distribution $\rho_{\#\mathbf{A}}$ the *quasi-joint-probability* (QJP) distribution of \mathbf{A} on ρ pertaining to the QJSD $\#\mathbf{A}$. Note that, in general, the corresponding QJP distributions may be negative or even complex valued.

Adjointness of Quantisation and Quasi-classicalisation

As one may surmise, quantisation (36) and quasi-classicalisation (37) are *adjoint* operations to each other, as one may readily check by the formal computation

$$\begin{aligned} \langle \Phi_{\#\mathbf{A}}(f), \rho \rangle_Q &:= \text{Tr}[f_{\#\mathbf{A}}\rho] \\ &= \int_{\mathbb{K}^n} f(\mathbf{a}) \text{Tr}[\#\mathbf{A}(\mathbf{a})\rho] \, dm_n(\mathbf{a}) \\ &= \int_{\mathbb{K}^n} f(\mathbf{a}) \rho_{\#\mathbf{A}}(\mathbf{a}) \, dm_n(\mathbf{a}) \\ &= \langle f, \Phi'_{\#\mathbf{A}}(\rho) \rangle_C, \quad f \in \mathcal{S}(\mathbb{R}^n), \rho \in N(\mathcal{H}). \end{aligned} \quad (38)$$

This relation can be illustrated by the following diagram:

$$\begin{array}{ccc} L(\mathcal{H}) & \overset{\text{dual pair}}{\longleftrightarrow} & N(\mathcal{H}) \\ \uparrow \Phi_{\#\mathbf{A}} \text{ Quantisation} & & \downarrow \Phi'_{\#\mathbf{A}} \text{ Quasi-classicalisation} \\ \mathcal{S}(\mathbb{K}^n) & \overset{\text{dual pair}}{\longleftrightarrow} & \mathcal{S}'(\mathbb{K}^n) \end{array}$$

Note again that, as maps, quantisation (36) and quasi-classicalisation (37) are uniquely dictated by the choice of the QJSD $\#_{\mathbf{A}}$. Even for the same classical observable $f \in \mathcal{S}(\mathbb{K}^n)$, its quantisation generally differs $f_{\#_{\mathbf{A}}} \neq f_{\tilde{\#}_{\mathbf{A}}}$ given a distinct choice of the QJSD $\#_{\mathbf{A}} \neq \tilde{\#}_{\mathbf{A}}$, and the same is also true for the quasi-classicalisation of quantum states $\rho \in N(\mathcal{H})$.

2.3 Quantum/Quasi-classical Representations

Since quantisation and quasi-classicalisation are adjoint notions to each other, they are different facets of a single entity. In this sense, we occasionally use the term *quantum/quasi-classical representations* or *transformations* referring to the adjoint pair. In general, these representations are non-unique, and each of the representations can be specified by the choice of the QJSD, whose indefiniteness originates directly from the non-commutative nature of the observables concerned.

2.3.1 Transformation of Representations

We may define transformations of QJSDs in various manners. In some cases, it may occur that a group of quantisation/quasi-classical representations could be understood as being *equivalent* to each other in the sense that they can be mutually transformed into one another. Given an ordered combination \mathbf{A} of quantum observables, it is an interesting question to ask ourselves how many QJSDs there are (or in other words, the way of ordering of non-commuting observables) that are essentially distinct to each other up to isomorphisms.

The Simplest Case

The simplest case for this is when two QJSDs, while being distinct in their form as hashed operators, are identical. Needless to say, this is trivially always the case when all the members of the ordered combination \mathbf{A} pairwise strongly commute. As an example of less trivial cases, let $\mathbf{X} = (Q, P)$ be an ordered pair of self-adjoint operators such that the members satisfy the Weyl representation

$$e^{-isQ}e^{-itP} = e^{-ist}e^{-itP}e^{-isQ}, \quad s, t \in \mathbb{R} \quad (39)$$

of the canonical commutation relation (CCR). Then, by a simple observation, one verifies that the hashed operators of the form

$$\begin{aligned}
e^{-i\frac{t}{2}P} e^{-isQ} e^{-i\frac{t}{2}P} &= e^{-i\frac{s}{3}Q} e^{-i\frac{t}{2}P} e^{-i\frac{s}{3}Q} e^{-i\frac{t}{2}P} e^{-i\frac{s}{3}Q} \\
&= e^{-i\frac{t}{4}P} e^{-i\frac{s}{3}Q} e^{-i\frac{t}{4}P} e^{-i\frac{s}{3}Q} e^{-i\frac{t}{4}P} e^{-i\frac{s}{3}Q} e^{-i\frac{t}{4}P} \\
&= \dots \\
&= e^{-i\overline{(sQ+tP)}} \\
&= \dots \\
&= e^{-i\frac{s}{4}Q} e^{-i\frac{t}{3}P} e^{-i\frac{s}{4}Q} e^{-i\frac{t}{3}P} e^{-i\frac{s}{4}Q} e^{-i\frac{t}{3}P} e^{-i\frac{s}{4}Q} \\
&= e^{-i\frac{t}{3}P} e^{-i\frac{s}{2}Q} e^{-i\frac{t}{3}P} e^{-i\frac{s}{2}Q} e^{-i\frac{t}{3}P} \\
&= e^{-i\frac{s}{2}Q} e^{-itP} e^{-i\frac{s}{2}Q}, \quad s, t \in \mathbb{R},
\end{aligned} \tag{40}$$

are all identical. Here, the equality in the center is due to the product formula proven by Lie-Trotter-Kato, and the overline on the operator $sQ+tP$ denotes its self-adjoint extension. Although they differ in their representation as mixtures, their corresponding QJSDs are identical, and thus all the quantisation/quasi-classical representations induced could be naturally understood as being equivalent.

Affine Transformations

As another example, let $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ be a linear map, and consider the affine map

$$T_{\mathbf{b}} : \mathbf{a} \mapsto T\mathbf{a} + \mathbf{b} \tag{41}$$

defined for $\mathbf{b} \in \mathbb{K}^n$. We then introduce the *affine transform* of a QJSD $\#_{\mathbf{A}}$ with respect to the affine map (41) by

$$(T_{\mathbf{b}}\#_{\mathbf{A}})(\mathbf{a}) := \#_{\mathbf{A}}(T_{\mathbf{b}}^{-1}\mathbf{a}), \quad \mathbf{a} \in \mathbb{K}^n. \tag{42}$$

If T is a bijection (*i.e.*, $\det T \neq 0$), both $T_{\mathbf{b}}\#_{\mathbf{A}}$ and $\#_{\mathbf{A}}$ are invertible to one another and thus essentially contain the same information of the combination \mathbf{A} of the observables concerned. It is to be noted that the affine transform of a QJSD is generally not a member of the QJSDs. Indeed, while the total integration of (42) reduces to the identity Id by definition, it most importantly fails to satisfy the marginal properties (30) in general. Even so, affine transforms of a QJSD give rise to adjoint pairs of quantisation and quasi-classicalisation in an extended sense, which are respectively defined by

$$\Phi_{(T_{\mathbf{b}}\#_{\mathbf{A}})} : f \mapsto f_{(T_{\mathbf{b}}\#_{\mathbf{A}})} := \int_{\mathbb{K}^n} f(\mathbf{a}) (T_{\mathbf{b}}\#_{\mathbf{A}})(\mathbf{a}) dm_n(\mathbf{a}) \tag{43}$$

and

$$\Phi'_{(T_{\mathbf{b}}\#_{\mathbf{A}})} : \rho \mapsto \rho_{(T_{\mathbf{b}}\#_{\mathbf{A}})}(\mathbf{a}) := \text{Tr}[(T_{\mathbf{b}}\#_{\mathbf{A}})(\mathbf{a})\rho]. \tag{44}$$

The adjointness of these operations

$$\langle f_{(T_{\mathbf{b}}\#_{\mathbf{A}})}, \rho \rangle_Q = \langle f, \rho_{(T_{\mathbf{b}}\#_{\mathbf{A}})} \rangle_C \tag{45}$$

can be readily confirmed by a simple computation. To see how the quantum/quasi-classical representations corresponding to affine transforms relate to the original representation, we first observe

$$\begin{aligned}
f_{(T_{\mathbf{b}}\#_{\mathbf{A}})} &:= \int_{\mathbb{K}^n} f(\mathbf{a}) (T_{\mathbf{b}}\#_{\mathbf{A}})(\mathbf{a}) dm_n(\mathbf{a}) \\
&= \int_{\mathbb{K}^n} f(T_{\mathbf{b}}\mathbf{a})\#_{\mathbf{A}}(\mathbf{a}) dm_n(\mathbf{a}) \\
&= (f \circ T_{\mathbf{b}})\#_{\mathbf{A}} \\
&=: (T_{\mathbf{b}}^* f)\#_{\mathbf{A}}
\end{aligned} \tag{46}$$

and

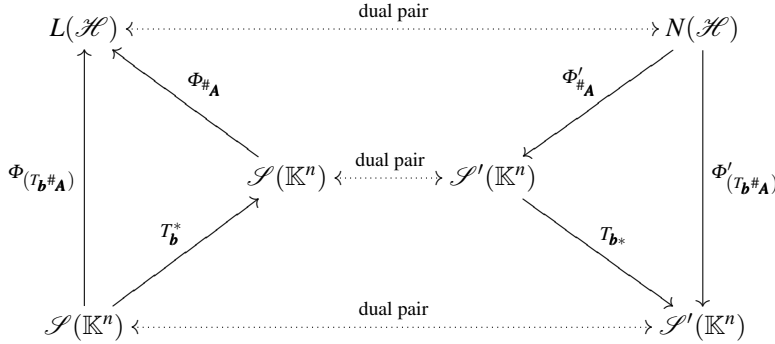
$$\begin{aligned}
\rho_{(T_{\mathbf{b}}\#_{\mathbf{A}})}(\mathbf{a}) &= \rho_{\#_{\mathbf{A}}}(T_{\mathbf{b}}^{-1}\mathbf{a}), \\
&=: T_{\mathbf{b}*}\rho_{\#_{\mathbf{A}}}
\end{aligned} \tag{47}$$

where

$$T_{\mathbf{b}}^* f := f \circ T_{\mathbf{b}}, \tag{48}$$

$$T_{\mathbf{b}*}\rho := \rho \circ T_{\mathbf{b}}^{-1}, \tag{49}$$

respectively denote the pullback of a function f and the pushforward of a distribution ρ by the affine map $T_{\mathbf{b}}$. The relation can thus be illustrated by the following diagram



As a simple concrete example, we consider the simplest case $\mathbf{A} = (A, B)$. Below, we see that all the members $\#_{\mathbf{A}}^{\kappa}$ of the subfamily of the QJSDs of the form (25) are linear transforms of $\#_{\mathbf{A}}^i$ for the specific choice $\kappa = i$. To see this, consider the matrix

$$\tilde{T}_{\kappa} := \begin{pmatrix} 1 & \kappa_1 \\ 0 & \kappa_2 \end{pmatrix} \tag{50}$$

defined for each $\kappa = \kappa_1 + i\kappa_2$, ($\kappa_1, \kappa_2 \in \mathbb{R}$), and the linear transformation $T_{\kappa} := \tilde{T}_{\kappa} \times \text{Id}$ on $\mathbb{C} \times \mathbb{R}$ defined by $T_{\kappa}(a, b) := (\tilde{T}_{\kappa}a, b)$. Then, a simple computation yields

$$\begin{aligned}
\mathcal{F}(T_\kappa \#_{\mathbf{A}}^i)(s, t) &= \int_{\mathbb{C} \times \mathbb{R}} e^{-i(s, T_\kappa a)} e^{-itb} \#_{\mathbf{A}}^i(a, b) dm_2(a, b) \\
&= \hat{\#}_{\mathbf{A}}^i(T'_\kappa s, t) \\
&= \mathcal{F}\#_{\mathbf{A}}^K(s, t),
\end{aligned} \tag{51}$$

where T'_κ denotes the adjoint matrix of T_κ . We thus conclude

$$T_\kappa \#_{\mathbf{A}}^i = \#_{\mathbf{A}}^K. \tag{52}$$

Since $\det T_\kappa = \kappa_2$, it is straightforward to see that all the members $\#_{\mathbf{A}}^K$ of the QJSDs for the choice $\text{Im } \kappa \neq 0$ are equivalent to one another by linear transformations.

Convolutions

As another important class of transformations, we consider the *convolution*

$$(h * \#_{\mathbf{A}})(\mathbf{a}) := \int_{\mathbb{K}^n} h(\mathbf{a} - \mathbf{a}') \#_{\mathbf{A}}(\mathbf{a}') dm_n(\mathbf{a}') \tag{53}$$

of a QJSD $\#_{\mathbf{A}}$ and a function h with the total integration of unity. The Fourier transform of the convolution reads

$$\begin{aligned}
\mathcal{F}(h * \#_{\mathbf{A}})(\mathbf{s}) &= \hat{h}(\mathbf{s}) \hat{\#}_{\mathbf{A}}(\mathbf{s}) \\
&= \int_{\mathbb{K}^n} h(\mathbf{k}) e^{-i(\mathbf{s}, \mathbf{k})} \hat{\#}_{\mathbf{A}}(\mathbf{s}) dm_n(\mathbf{k}) \\
&= \int_{\mathbb{K}^n} h(\mathbf{k}) \hat{\#}_{\mathbf{A} + \mathbf{k}}(\mathbf{s}) dm_n(\mathbf{k}),
\end{aligned} \tag{54}$$

where $\mathbf{A} + \mathbf{k} := (A_1 + k_1 \cdot \text{Id}, \dots, A_n + k_n \cdot \text{Id})$ denotes the ordered combination of the normal operators defined as the parallel translation of \mathbf{A} towards the direction $\mathbf{k} \in \mathbb{K}^n$. From this, the convolution

$$h * \#_{\mathbf{A}} = \int_{\mathbb{K}^n} h(\mathbf{k}) \#_{\mathbf{A} + \mathbf{k}} dm_n(\mathbf{k}) \tag{55}$$

could be understood as the ‘weighted average’ of the family of QJSDs $\#_{\mathbf{A} + \mathbf{k}}$ of the ordered combination $\mathbf{A} + \mathbf{k}$ of normal operators, or equivalently, the parallel translation

$$\#_{\mathbf{A} + \mathbf{k}} = \tau_{\mathbf{k}} \#_{\mathbf{A}}, \quad (\tau_{\mathbf{k}} \mathbf{a} := \mathbf{a} + \mathbf{k}) \tag{56}$$

of the original QJSD $\#_{\mathbf{A}}$, with respect to the ‘weight function’ h .

It is important to note that, in general, the convolution (53) itself is not necessarily a QJSD of \mathbf{A} . Indeed, consider the most extreme case where all the members of \mathbf{A} pairwise strongly commute. In such case, the unique QJSD of \mathbf{A} is the JSM $E_{\mathbf{A}}$, so in order for the convolution $h * E_{\mathbf{A}} = E_{\mathbf{A}}$ to be the unique QJSD of \mathbf{A} , we must have $h = \delta_{\mathbf{0}} \Leftrightarrow \hat{h} = 1$. In a more general setting, a necessary condition for $h * \#_{\mathbf{A}}$ to satisfy

the marginal condition (35) is given by⁹

$$\int_{\mathbb{K}^{n-k}} h(\mathbf{a}) dm_{n-k}(\mathbf{b}^c) = \delta_{\mathbf{0}}(\mathbf{b}), \quad (57)$$

where \mathbf{B} is an order-preserving subset (33) of \mathbf{A} consisting of k ($1 \leq k \leq n$) numbers of its pairwise strongly commuting distinct members, \mathbf{B}^c the order-preserving complement (34) of \mathbf{B} , and $\mathbf{b} \in \mathbb{K}^k$, $\mathbf{b}^c \in \mathbb{K}^{n-k}$ are their corresponding variables. Here, $\delta_{\mathbf{a}}$, denotes the *delta distribution* centred at $\mathbf{a} \in \mathbb{K}^n$, which is a generalised function symbolically defined as

$$\delta_{\mathbf{a}}(\mathbf{x}) = \begin{cases} \infty, & (\mathbf{x} = \mathbf{a}) \\ 0, & (\mathbf{x} \neq \mathbf{a}) \end{cases} \quad (58)$$

in the usual manner.

On the other hand, the convolution still gives rise to the adjoint pair of representations in an extended sense in a similar manner to the affine transforms of a QJSD. In order to see its relation to the original representation, we first observe that

$$\begin{aligned} f_{(h*\#_{\mathbf{A}})} &:= \int_{\mathbb{K}^n} f(\mathbf{a})(h*\#_{\mathbf{A}})(\mathbf{a}) dm_n(\mathbf{a}) \\ &= \int_{\mathbb{K}^n} (\tilde{h}*f)(\mathbf{a})\#_{\mathbf{A}}(\mathbf{a}) dm_n(\mathbf{a}) \\ &=: (\tilde{h}*f)_{\#_{\mathbf{A}}}, \end{aligned} \quad (59)$$

where $\tilde{h}(\mathbf{a}) := h(-\mathbf{a})$ denotes the transpose of h . One also finds

$$\rho_{(h*\#_{\mathbf{A}})} = h*\rho_{\#_{\mathbf{A}}}. \quad (60)$$

The adjointness of these operations

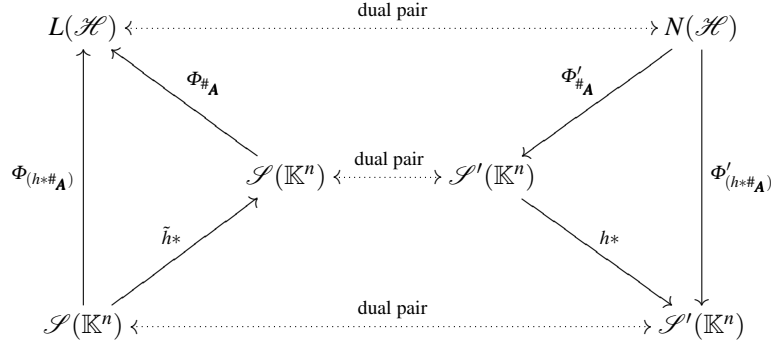
$$\langle f_{(h*\#_{\mathbf{A}})}, \rho \rangle_Q = \langle f, \rho_{(h*\#_{\mathbf{A}})} \rangle_C \quad (61)$$

can be readily confirmed by a simple computation. The diagram

⁹ Here, we adopt the convention

$$\int_{\mathbb{K}^{n-k}} h(\mathbf{a}) dm_{n-k}(\mathbf{b}^c) = h(\mathbf{a})$$

for the case $k = n$.



gives a visual summary as to how the quantum/quasi-classical representation corresponding to the convolution of the QJSD $\#_{\mathbf{A}}$ with the function h relates to the original representation. One readily sees that the two representations corresponding to $\#_{\mathbf{A}}$ and $h^*\#_{\mathbf{A}}$ are equivalent if the map $h^* : f \mapsto h^* f$ is a bijection.

2.3.2 Faithfulness of the Representations

For a fixed \mathbf{A} , we say that the quantum/quasi-classical representation pertaining to a given choice of QJDS $\#_{\mathbf{A}}$ is *faithful*, if either of the following equivalent¹⁰ conditions are met:

1. The quantisation $\Phi_{\#_{\mathbf{A}}}$ has a dense range, *i.e.*,

$$\overline{\text{ran } \Phi_{\#_{\mathbf{A}}}} = L(\mathcal{H}). \quad (62)$$

2. The quasi-classicalisation $\Phi'_{\#_{\mathbf{A}}}$ is injective, *i.e.*,

$$\rho_{\#_{\mathbf{A}}} = \sigma_{\#_{\mathbf{A}}} \Leftrightarrow \rho = \sigma, \quad (\rho, \sigma \in N(\mathcal{H})) \quad (63)$$

In physical terms, this is to say that every quantum operator $X = \lim f_i(\#_{\mathbf{A}})$ can be represented by limits of quantised operators if and only if no two quantum states give rise to the same QJP distribution. In general, the larger the number of observables belonging to the ordered combination \mathbf{A} becomes, the closer the representation approaches to faithfulness.

2.3.3 Realness of the Representations

Among the various candidates of representations, ‘real’ representations are often-times prized. To state the precise definition, we first need some preparations. Given a QJSD $\#_{\mathbf{A}}$, the *conjugate* of $\#_{\mathbf{A}}$ is formally defined by

¹⁰ The proof for the equivalence of the conditions can be carried out by applying the Hahn-Banach theorem on locally convex spaces.

$$\#_{\mathbf{A}}^*(\mathbf{a}) := \#_{\mathbf{A}}(\mathbf{a})^*, \quad \mathbf{a} \in \mathbb{K}^n, \quad (64)$$

where the asterisk on the r. h. s. denotes the adjoint of $\#_{\mathbf{A}}(\mathbf{a})$. The Fourier transform of the conjugate of a QJSD reads

$$\begin{aligned} (\mathcal{F}\#_{\mathbf{A}}^*)(\mathbf{s}) &:= \int_{\mathbb{K}^n} e^{-i\langle \mathbf{s}, \mathbf{a} \rangle} \#_{\mathbf{A}}(\mathbf{a})^* dm_n(\mathbf{a}) \\ &= \left(\int_{\mathbb{K}^n} e^{-i\langle -\mathbf{s}, \mathbf{a} \rangle} \#_{\mathbf{A}}(\mathbf{a}) dm_n(\mathbf{a}) \right)^* \\ &= \hat{\#}_{\mathbf{A}}(-\mathbf{s})^* \\ &=: \hat{\#}_{\mathbf{A}}^\dagger(\mathbf{s}), \end{aligned} \quad (65)$$

where we have introduced the *involution* $\hat{\#}_{\mathbf{A}}^\dagger$ of the hashed operator $\hat{\#}_{\mathbf{A}}$ in the last equality. It is not difficult to see that involutions of hashed operators are again hashed operators, hence a conjugate of a QJSD is a QJSD.

We say that a QJSD $\#_{\mathbf{A}}$ is *real* if $\#_{\mathbf{A}}^* = \#_{\mathbf{A}}$ holds. By the bijectivity of the Fourier transformation, a QJSD is real if and only if its Fourier transform (*i.e.*, the corresponding hashed operator) is a self-involution. Examples of real QJSDs of $\mathbf{A} = (A, B)$ are those whose corresponding hashed operators read

$$\hat{\#}(s, t) = \begin{cases} \frac{1}{2} \cdot (e^{-isB} e^{-isA} + e^{-isA} e^{-itB}) \\ e^{-i\frac{s}{2}A} e^{-itB} e^{-i\frac{s}{2}A} \\ e^{-i\frac{t}{2}B} e^{-isA} e^{-i\frac{t}{2}B} \\ e^{-i\frac{s}{3}A} e^{-i\frac{t}{2}B} e^{-i\frac{s}{3}A} e^{-i\frac{t}{2}B} e^{-i\frac{s}{3}A} \\ e^{-i\frac{t}{4}B} e^{-i\frac{s}{3}A} e^{-i\frac{t}{4}B} e^{-i\frac{s}{3}A} e^{-i\frac{t}{4}B} e^{-i\frac{s}{3}A} e^{-i\frac{t}{4}B} \\ \overline{e^{-i(sA+tB)}} \end{cases} \quad s, t \in \mathbb{R},$$

where the overline on the operator $sA + tB$ denotes its unique self-adjoint extension. Colloquially speaking, a QJSD is real if its corresponding hashed operator is ‘symmetric’ in its form. We call the quantum/quasi-classical representation real, if the corresponding QJSD is real. Real representations have the following convenient properties:

1. The quantisation $f_{\#_{\mathbf{A}}}$ of a real function $f \in \mathcal{S}(\mathbb{K}^n)$ is always self-adjoint.
2. The quasi-classicalisation $\rho_{\#_{\mathbf{A}}}$ of a density operator $\rho \in N(\mathcal{H})$ is always real.

Real representations thus have a formal advantage of taking a classical observables into self-adjoint operators, and a quantum state into real QJP distribution, which some may find favourable above non-real representations.

2.3.4 Relation to some prior Works

The study on quantum-classical transformation has a long history. In this passage, we investigate the relation of the formalism of QJSDs presented in this paper to

some of the prior works on this topic. In this passage, we are specifically interested in the choice $\mathbf{X} = (Q, P)$, where Q and P are operators satisfying the Weyl representation (39) of the CCR.

The Theory of Wigner and Weyl

We first point out that the Weyl map and the Wigner map are respectively the quantisation $\Phi_{\#\mathbf{X}}$ and the quasi-classicalisation $\Phi'_{\#\mathbf{X}}$ pertaining to the QJSD $W_{\mathbf{X}}$ of \mathbf{X} corresponding to the hashed operator of the form

$$\hat{W}_{\mathbf{X}}(s, t) := e^{-i\overline{(sQ+tP)}}. \quad (66)$$

In fact, since the hashed operator (66) is equivalent to any choice of the form (40), the Wigner-Weyl transformation is the quantum/quasi-classical representation pertaining to any of the choices. It is easy to see from the self-involutive form of the hashed operator (66) that the Wigner-Weyl transformation is real. It is widely known that the Wigner map is injective, which is equivalent for its adjoint map (*i.e.*, the Weyl map) to have a dense range: in the terminology of this paper, the Wigner-Weyl transformation is faithful.

In order to confirm the claim, we first compute the quantisation of functions with respect to the QJSD corresponding to (66). The quantisation of $f \in \mathcal{S}(\mathbb{R}^2)$ reads

$$\begin{aligned} f_{W_{\mathbf{X}}} &:= \int_{\mathbb{R}^2} f(q, p) W_{\mathbf{X}}(q, p) dm_2(q, p) \\ &= \int_{\mathbb{R}^4} f(q, p) e^{i(qs+pt)} \hat{W}_{\mathbf{X}}(s, t) dm_2(q, p, s, t) \\ &= \int_{\mathbb{R}^4} f(q, p) e^{-is(Q-q)} e^{-it(P-p)} e^{ist/2} dm_2(q, p, s, t) \\ &=: W_{\mathbf{X}}(f), \end{aligned} \quad (67)$$

where the last equality is precisely the definition of the Weyl quantisation of f . Here, note that we have used the relation

$$e^{-itP/2} e^{-isQ} e^{-itP/2} = e^{ist/2} e^{-isQ} e^{-itP} \quad (68)$$

in order to obtain the third equality. As for the quasi-classical representation of a quantum state, we assume without loss of generality that $\mathcal{H} = L^2(\mathbb{R})$, and that $Q = \hat{x}$, $P = \hat{p}$ are respectively the familiar position and momentum operators. For better readability, we moreover restrict ourselves to the case that $\rho = |\psi\rangle\langle\psi|$ for some wave-function $\psi \in L^2(\mathbb{R})$. Then, the Fourier transform of $\rho_{W_{\mathbf{X}}}$ reads

$$\begin{aligned}
(\mathcal{F}\rho_{W_{\mathbf{X}}})(s,t) &= \text{Tr} [\hat{W}_{\mathbf{X}}(s,t)\rho] \\
&= \langle e^{itP/2}\psi, e^{-isQ}e^{-itP/2}\psi \rangle \\
&= \int_{\mathbb{R}} e^{-isq}\psi^*(q+t/2)\psi(q-t/2) dm(x) \\
&= \int_{\mathbb{R}} e^{-i(sq+tp)} \left(\int_{\mathbb{R}} \psi^*(q+t/2)\psi(q-t/2)e^{ipt} dm(t) \right) dm_2(q,p) \\
&= (\mathcal{F}W_{\mathbf{X}}^{\rho})(s,t), \tag{69}
\end{aligned}$$

where

$$W_{\mathbf{X}}^{\rho}(q,p) := \int_{\mathbb{R}} \psi^*(q+t/2)\psi(q-t/2)e^{ipt} dm(t) \tag{70}$$

is the Wigner function of the wave function $\psi \in L^2(\mathbb{R})$. Due to the injectivity of the Fourier transformation, we conclude $\rho_{W_{\mathbf{X}}} = W_{\mathbf{X}}^{\rho}$. The proof for the general case goes essentially the same.

QSJDs generated by Convolution

We next consider the problem of generating a family of QJSDs of \mathbf{X} by means of the convolution of $W_{\mathbf{X}}$ with functions. This setting has been previously examined by Cohen et al. [10] with the intention of constructing a family of generalised phase space distribution functions. For the convolution $h * W_{\mathbf{X}}$ to be a QJSD of \mathbf{X} , one sees from the result (57) that the condition $\hat{h}(0,t) = \hat{h}(s,0) = 1$, $s,t \in \mathbb{R}$ is necessary. Specifically, we demonstrate below that, under the assumption that the Fourier transform

$$\hat{h}(s,t) = \hat{g}(st/2), \quad s,t \in \mathbb{R} \tag{71}$$

of the function h can be represented by a function $g : \mathbb{R} \rightarrow \mathbb{R}$ with the total integration of unity, the convolution (55) becomes a QJSD of the pair \mathbf{X} . Indeed, observe that since

$$\mathcal{F}(h * W_{\mathbf{X}})(s,t) = \int_{\mathbb{R}} g(\kappa)e^{-i\kappa st/2}\hat{W}_{\mathbf{X}}(s,t) dm(\kappa), \tag{72}$$

we have

$$h * W_{\mathbf{X}} = \int_{\mathbb{R}} g(\kappa)\#_{\mathbf{X}}^{\kappa}(s,t) dm(\kappa), \tag{73}$$

where we have used the parametrised family $\#_{\mathbf{X}}^{\kappa}$, $\kappa \in \mathbb{R}$, of the QJSD of \mathbf{X} corresponding to the hashed operators of the form

$$\begin{aligned}
\hat{\#}_{\mathbf{X}}^{\kappa}(s,t) &:= e^{-i\kappa st/2}\hat{W}_{\mathbf{X}}(s,t) \\
&= e^{-i\frac{1-\kappa}{2}sQ}e^{-itP}e^{-i\frac{1+\kappa}{2}sQ}, \tag{74}
\end{aligned}$$

which was originally introduced in (25) for general $\mathbf{A} = (A,B)$. Hence, as the ‘weighted average’ of the parametrised families of the QJSDs of \mathbf{X} , the convolution

(73) itself is a QJSD of \mathbf{X} . Each choice of the function g yields distinct representation. Among the well-known representations are those proposed by:

1. Weyl [4] and Wigner [5]

$$g(k) = \delta_0(k) \Leftrightarrow \hat{g}(\omega) = 1 \quad (75)$$

2. Kirkwood [11] and Dirac [12]

$$g(k) = \delta_{\pm 1}(k) \Leftrightarrow \hat{g}(\omega) = e^{\mp i\omega} \quad (76)$$

3. Margenau and Hill [13]

$$g(k) = \frac{(\delta_{-1}(k) + \delta_{+1}(k))}{2} \Leftrightarrow \hat{g}(\omega) = \cos \omega \quad (77)$$

4. Born and Jordan [14]

$$g(k) = \begin{cases} \frac{1}{2}, & (|k| \leq 1) \\ 0, & (|k| > 1) \end{cases} \Leftrightarrow \hat{g}(k) = \frac{\sin \omega}{\omega} \quad (78)$$

which all belongs to the same class generated by convolutions.

The Theory of Hushimi and Glauber-Sudarshan

In the previous paragraph, we have considered the problem of constructing various QJSDs of \mathbf{X} by means of convolution. In a broader perspective, however, the convolution $h * W_{\mathbf{X}}$ itself need not be a QJSD of \mathbf{X} in the sense that it gives rise to the adjoint pair of representations in an extended sense. If the map $h * : f \mapsto h * f$ is bijective, both the original QJSD and the convolution contain the same amount of information of \mathbf{X} , and a sufficient condition for this is $\hat{h} > 0$.

As for the choice of the function h , which may from the result (55) be interpreted as the ‘weight function’ of the family of QJSDs, we typically consider the normal distribution (*i.e.*, Gaußian function)

$$G(\mathbf{x}) := e^{-\|\mathbf{x}\|^2/2^n}, \quad \mathbf{x} \in \mathbb{R}^n, \quad (79)$$

where $\|\mathbf{x}\| := \sqrt{\sum_{i=1}^n |x_i|^2}$ denotes the Euclidean norm of an n -dimensional vector $\mathbf{x} \in \mathbb{R}^n$. Since the normal distribution never satisfies the condition (57), the convolution $G * \#_{\mathbf{X}}$ of a normal distribution with a QJSD of \mathbf{X} is not a QJSD of \mathbf{X} in general. However, since the Fourier transform of a normal distribution is another normal distribution, hence $\hat{G} > 0$, the original QJSD $\#_{\mathbf{X}}$ and the convolution $G * \#_{\mathbf{X}}$ are equivalent to each other. We thus introduce two operator valued distributions (OVDs) $H_{\mathbf{X}}$ and $GS_{\mathbf{X}}$ that are uniquely specified through the relations

$$H_{\mathbf{X}} = G * W_{\mathbf{X}}, \quad (80)$$

$$W_{\mathbf{X}} = G * GS_{\mathbf{X}}. \quad (81)$$

By the above argument, we see that all $W_{\mathbf{X}}$, $H_{\mathbf{X}}$ and $GS_{\mathbf{X}}$ contain the same information in the sense that they can be transformed to one another by convolution with normal distributions, and thus the adjoint pairs of quantum/quasi-classical representations which they yield are all equivalent. In this paper, we casually call the QJSD $W_{\mathbf{X}}$ the *Wigner-Weyl* type, and the OVDs $H_{\mathbf{X}}$ and $GS_{\mathbf{X}}$ the *Hushimi* and the *Glauber-Sudarshan* type, respectively. From the result (60), one sees that the quasi-classical representation of a density operator $\rho \in N(\mathcal{H})$ pertaining to the respective representations are related to each other by

$$H_{\mathbf{X}}^{\rho} = G * W_{\mathbf{X}}^{\rho}, \quad (82)$$

$$W_{\mathbf{X}}^{\rho} = G * GS_{\mathbf{X}}^{\rho}. \quad (83)$$

As we have seen in the previous passage that $W_{\mathbf{X}}^{\rho}(q, p)$ is the Wigner function of ρ , we here learn that the distributions $H_{\mathbf{X}}^{\rho}(q, p)$ and $GS_{\mathbf{X}}^{\rho}(q, p)$ are precisely the *Hushimi Q-function* [15] and the *Glauber-Sudarshan P-function* [16, 17] of the quantum state ρ , respectively.

3 Some Applications: Quantum Correlations, Conditional Expectations and the Weak Value

We now seek application of the formalisms we have constructed in the previous chapter. By its nature, the framework of quantisation/quasi-classicalisation by QJSDs should become useful in analysing problems where non-commutativity of quantum observables is concerned. In this section, we specifically focus on ‘correlations’ and ‘conditioning’ between (generally non-commuting) quantum observables induced by QJSDs. Due to the non-commuting nature of quantum observables, there are various candidates of quantum correlations and conditional expectations. Specifically, we see that Aharonov’s weak value [7] can be identified as one realisation among various candidates of quantum conditional expectations.

Complex Parametrised Subfamily

In handling relatively abstract objects as QJSDs of quantum observables, concrete examples are always of use. To this, we occasionally consider the simplest case where only two self-adjoint operators $\mathbf{A} = (A, B)$ are concerned, and make use of the complex parametrised subfamily of hashed operators

$$\hat{\#}_{\mathbf{A}}^{\alpha}(s, t) := \frac{1 + \alpha}{2} \cdot e^{-itB} e^{-isA} + \frac{1 - \alpha}{2} \cdot e^{-isA} e^{-itB}, \quad \alpha \in \mathbb{C}, \quad (84)$$

and the resulting subfamily $\#_{\mathbf{A}}^{\alpha} := \mathcal{F}^{-1} \hat{\#}_{\mathbf{A}}^{\alpha}$ of QJSDs for demonstration¹¹.

3.1 Correlations of Quantum Observables

In classical probability theory, *correlation* has been an important quantity in various aspects. The definition of the quantum counterpart, however, is not so obvious, when non-commutative observables are taken into account. In what follows, we define a family of quantum correlation based on our framework of QJSDs of quantum observables, and observe their very basic properties.

3.1.1 Sesquilinear Forms induced by QJP Distributions

As usual, let $\mathbf{A} = (A_1, \dots, A_n)$, $n \geq 1$, be an ordered combination of self-adjoint operators on a Hilbert space \mathcal{H} , and choose a QJSD $\#_{\mathbf{A}}$ and a density operator $\rho \in N(\mathcal{H})$. In what follows, in order to refrain ourselves from dealing with unessential mathematical intricacies, we assume that the resulting QJP distribution $\rho_{\#_{\mathbf{A}}}$ can be

¹¹ Do not confuse the subfamily introduced above with that of (25). We have used different superscript characters κ , α as parameters to make the distinction more easier.

represented by a density function¹². This allows us to introduce a sesquilinear form

$$\langle g, f \rangle_{\rho_{\#_A}} := \int_{\mathbb{K}^n} g^*(\mathbf{a}) f(\mathbf{a}) \rho_{\#_A}(\mathbf{a}) dm_n(\mathbf{a}), \quad f, g \in L^2(\rho_{\#_A}) \quad (85)$$

defined on the space of square-integrable functions with respect to the complex density function $\rho_{\#_A}$. By definition, we have

$$\langle f, g \rangle_{\rho_{\#_A}} = \langle g^*, f^* \rangle_{\rho_{\#_A}} \quad (86)$$

and

$$\langle f, g \rangle_{\rho_{\#_A}^*} = \langle g, f \rangle_{\rho_{\#_A}}^*, \quad (87)$$

where the superscript asterisk denotes the complex conjugate. The following observations are direct consequences of the above properties.

1. The quantum correlation (85) is *symmetric* (Hermitian)

$$\langle f, g \rangle_{\rho_{\#_A}} = \langle g, f \rangle_{\rho_{\#_A}}^* \quad (88)$$

if and only if the QJP distribution $\rho_{\#_A}$ is *real*. This is guaranteed for every $\rho \in N(\mathcal{H})$ if and only if the choice of the QJPD $\#_A$ is self-adjoint.

2. The quantum correlation (85) is *positive definite*¹³

$$\begin{aligned} \forall f \in L^2 \quad \langle f, f \rangle_{\rho_{\#_A}} &\geq 0, \\ \langle f, f \rangle_{\rho_{\#_A}} = 0 &\Leftrightarrow f = 0 \end{aligned} \quad (89)$$

if and only if the QJP distribution $\rho_{\#_A}$ is *positive*. This is guaranteed for every $\rho \in N(\mathcal{H})$ if and only if the QJPD $\#_A$ is positive.

Note that the second condition (*i.e.*, positive definiteness) is stronger than the first condition (*i.e.*, symmetricity), for indeed the positiveness of the QJP distributions trivially implies its realness, and in parallel, the positiveness of the QJPDs implies its self-adjointness.

3.1.2 Quantum Correlations

We next introduce the concept of quantum correlations based on the sesquilinear forms defined above. In what follows, in order to ease our arguments, we confine ourselves to the simplest case $\mathbf{A} = (A, B)$ without loss of generality. We also write $\# := \#_A$ for better readability. Now, let $f(A), g(B)$ be operators respectively defined from the functions $f(a)$ and $g(b)$ by means of the functional calculus. We then define the *quantum correlations* or *quasi-correlations* between the operators $f(A)$,

¹² We say that a tempered distribution $u \in \mathcal{S}'(\mathbb{K}^n)$ admits representation by a density function if $u(x)$ is actually an integrable function.

¹³ Here, the equality $f = 0$ in the second line of (89) is meant to hold $\rho_{\#_A}$ -almost everywhere.

$g(B)$ by

$$\langle g(B), f(A) \rangle_{\rho_{\#}} := \int_{\mathbb{K}^2} g^*(b) f(a) \rho_{\#}(a, b) dm_2(a, b) \quad (90)$$

By construction, quantum correlations are dependent on the choice of the QJP distributions $\rho_{\#}$, which is generally non-unique due to the indefiniteness of the QJSDs. When A and B happen to be simultaneously measurable, indefiniteness of the QJSDs vanishes, and the quantum correlation reduces to the unique classical correlation in the standard sense.

3.1.3 Quantum Covariances

Now that we have introduced the concept of quantum correlations, we next introduce the concept of quantum covariances. Under the same assumptions, we introduce the *quantum covariance* of the pair with respect to the QJP distribution $\rho_{\#}$ defined as

$$\begin{aligned} \mathbb{C}\mathbb{V}[f(A), g(B); \rho_{\#}] &:= \langle g(B) - \mathbb{E}[g(B); \rho], f(A) - \mathbb{E}[f(A); \rho] \rangle_{\rho_{\#}} \\ &= \langle g(B), f(A) \rangle_{\rho_{\#}} - \mathbb{E}[f(A); \rho] \cdot \mathbb{E}[g(B); \rho], \end{aligned} \quad (91)$$

where $\mathbb{E}[X; \rho] := \text{Tr}[X\rho]$ denotes the expectation value of $X \in L(\mathcal{H})$ on a density operator $\rho \in N(\mathcal{H})$ as usual. The quantum covariance serves as a natural extension to the standard covariance in classical probability theory, and they indeed coincide when the pair of self-adjoint operators $f(A)$ and $g(B)$ strongly commute.

Example

As an example, let $\#^{\alpha}$ be a complex parametrised QJSD for $\alpha \in \mathbb{C}$ introduced in (84). By a simple computation, the quantum correlation of the operators A and B reads

$$\begin{aligned} \langle B, A \rangle_{\rho_{\#}} &:= \int_{\mathbb{R}^2} ba \rho_{\#}(a, b) dm_2(a, b) \\ &= (i\partial_{s_1})(i\partial_{s_2})(\mathcal{F}\rho_{\#})(\mathbf{s})|_{\mathbf{s}=0} \\ &= \text{Tr} \left[\frac{\{A, B\}}{2} \rho \right] + i\alpha \text{Tr} \left[\frac{[A, B]}{2i} \rho \right], \end{aligned} \quad (92)$$

where $\{X, Y\} := XY + YX$ and $[X, Y] := XY - YX$ respectively denotes the anti-commutator and the commutator of X and Y as usual. This computation leads to

$$\mathbb{C}\mathbb{V}[A, B; \rho_{\#^{\alpha}}] = \mathbb{C}\mathbb{V}_S[A, B; \rho] + i\alpha \mathbb{C}\mathbb{V}_A[A, B; \rho], \quad (93)$$

where we have introduced the *standard symmetric* and *standard anti-symmetric* quantum covariances

$$\mathbb{C}\mathbb{V}_S[A, B; \rho] := \text{Tr} \left[\frac{\{A, B\}}{2} \rho \right] - \mathbb{E}[A; \rho] \cdot \mathbb{E}[B; \rho], \quad (94)$$

$$\mathbb{C}\mathbb{V}_A[A, B; \rho] := \text{Tr} \left[\frac{[A, B]}{2i} \rho \right], \quad (95)$$

for better readability.

3.2 Conditioning by Quantum Observables

In the previous passage, we have introduced quantum analogues of correlations by means of QJP distributions. Closely related to these are quantum analogues of conditional expectations.

3.2.1 Introducing Quantum Conditional Expectations

In what follows, in order to avoid distraction by unessential mathematical intricacies, we impose an additional assumption that, for the given choice of the density operator $\rho \in \mathcal{N}(\mathcal{H})$, the probability of finding the outcome of B is always positive $\rho_B(b) > 0$ on its spectrum¹⁴. The quantum correlation of both the operators $f(A)$ and $g(B)$ then reads

$$\begin{aligned} \langle g(B), f(A) \rangle_{\rho_{\#}} &:= \int_{\mathbb{K}^2} g^*(b) f(a) \rho_{\#}(a, b) dm_2(a, b) \\ &= \int_{\mathbb{K}} g^*(b) \mathbb{E}[f(A)|B = b; \rho_{\#}] \rho_B(b) dm(b) \\ &= \langle g(B), \mathbb{E}[f(A)|B; \rho_{\#}] \rangle_{\rho_B}, \end{aligned} \quad (96)$$

where we have introduced the *quantum conditional expectation* or *conditional quasi-expectation*

$$\mathbb{E}[f(A)|B = b; \rho_{\#}] := \frac{\int_{\mathbb{K}} f(a) \rho_{\#}(a, b) dm(a)}{\rho_B(b)}. \quad (97)$$

of the operator $f(A)$ given the outcome of B under the QJP distribution $\rho_{\#}$. Note that the quantum conditional expectation $\mathbb{E}[f(A)|B; \rho_{\#}]$ is defined as an (equivalence class of) complex function(s) rather than a scalar. The normal operator

$$\mathbb{E}[f(A)|B; \rho_{\#}] := \int_{\mathbb{K}} \mathbb{E}[f(A)|B = b; \rho_{\#}](b) dE_B(b) \quad (98)$$

¹⁴ The spectrum of a self-adjoint operator A is defined as the largest closed subset $J \subset \mathbb{R}$ such that $E_B(J) = \text{Id}$ holds.

in the last equation is the image of the functional calculus of the (equivalence class of) function(s) (97).

Some ‘statistical’ Properties and its Interpretation

The key observation to make here is that the quantum correlation of an operator $f(A)$ with any operator $g(B)$ generated by B can be reproduced by the *authentic* correlation of the quantum conditional expectation $\mathbb{E}[f(B)|B; \rho_{\#}]$ with $g(B)$, which we reiterate for emphasis as:

$$\langle g(B), \mathbb{E}[f(A)|B; \rho_{\#}] \rangle_{\rho_B} = \langle g(B), f(A) \rangle_{\rho_{\#}}, \quad \forall g \in L^2(\rho_B). \quad (99)$$

Also, by taking the constant function $g = 1$, we have

$$\mathbb{E}[\mathbb{E}[f(A)|B; \rho_{\#}]; \rho] = \mathbb{E}[f(A); \rho]. \quad (100)$$

The above two equalities show that the quantum conditional expectation serve as the ‘approximation’ of the original operator $f(A)$ by operators generated by B , and that it is unique in the sense that it precisely reproduces the quantum correlation with any other operators generated by B in place of the original $f(A)$.

In physical terms, the quantum conditional expectation can be interpreted as the quantum analogue of conditional expectations of the operator $f(A)$ given the outcome b of B under the hypothetical ‘joint’ distribution (*i.e.*, QJP distribution) $\rho_{\#}$. If the combination of the observables $\mathbf{A} = (A, B)$ happens to be simultaneously measurable, the quantum conditional expectation simply reduces to the conditional expectation in the classical sense that is familiar to us. On the other hand, if some of the pair of observables fail to admit simultaneous measurability, the quantum conditional expectation becomes a hypothetical quantity, whose definition becomes non-unique due to the indefiniteness of the choice of the QJPDs.

Examples

We next provide some concrete examples to actually compute the quantum conditional expectations. To this end, let $\#^{\alpha}$ be the parametrised subfamily of QJSDs of the ordered pair $\mathbf{A} = (A, B)$ of self-adjoint operators introduced in (84). For a given function $f(a)$, let

$$f(a) = \sum_{m=0}^{\infty} f_m a^m \quad (101)$$

denote its Taylor expansion. If we let $\varphi(b)$ denote the numerator of (97), we then have

$$\begin{aligned}
(\mathcal{F}\varphi)(t) &= \int_{\mathbb{R}} e^{-itb} \left(\int_{\mathbb{R}} \sum_{m=0}^{\infty} f_m a^m \rho_{\#\alpha}(a, b) dm(a) \right) dm(b) \\
&= \sum_{m=0}^{\infty} f_m \left(\int_{\mathbb{R}^2} e^{-i(0a+tb)} a^m b^0 \rho_{\#\alpha}(a, b) dm_2(a, b) \right) \\
&= \sum_{m=0}^{\infty} f_m ((i\partial_s)^m (i\partial_t)^0 (\mathcal{F}\rho_{\#\alpha})(0, t)) \\
&= \sum_{m=0}^{\infty} f_m \left(\frac{1+\alpha}{2} \cdot \text{Tr} [e^{-itB} A^m e^{-i0A} \rho] + \frac{1-\alpha}{2} \cdot \text{Tr} [A^m e^{-i0A} e^{-itB} \rho] \right) \\
&= \frac{1+\alpha}{2} \cdot \text{Tr} [e^{-itB} f(A) \rho] + \frac{1-\alpha}{2} \cdot \text{Tr} [f(A) e^{-itB} \rho]. \tag{102}
\end{aligned}$$

Observing that the Fourier transform of the function $b \mapsto \text{Tr} [E_B(b) f(A) \rho]$ and $b \mapsto \text{Tr} [f(A) E_B(b) \rho]$ are respectively $\text{Tr} [e^{-itB} f(A) \rho]$ and $\text{Tr} [f(A) e^{-itB} \rho]$, one finds that the injectivity and the linearity of the Fourier transformation leads to

$$\int_{\mathbb{R}} f(a) \rho_{\#\alpha}(a, b) dm(a) = \frac{1+\alpha}{2} \cdot \text{Tr} [E_B(b) f(A) \rho] + \frac{1-\alpha}{2} \cdot \text{Tr} [f(A) E_B(b) \rho] \tag{103}$$

by combining the results. We thus finally have

$$\begin{aligned}
\mathbb{E}^\alpha [A|B = b; \rho] &:= \mathbb{E}[f(A)|B = b; \rho_{\#\alpha}] \\
&= \frac{1+\alpha}{2} \cdot \frac{\text{Tr} [E_B(b) f(A) \rho]}{\text{Tr} [E_B(b) \rho]} + \frac{1-\alpha}{2} \cdot \frac{\text{Tr} [f(A) E_B(b) \rho]}{\text{Tr} [E_B(b) \rho]} \\
&= \text{Re} \left[\frac{\text{Tr} [E_B(b) f(A) \rho]}{\text{Tr} [E_B(b) \rho]} \right] + i\alpha \text{Im} \left[\frac{\text{Tr} [E_B(b) f(A) \rho]}{\text{Tr} [E_B(b) \rho]} \right], \tag{104}
\end{aligned}$$

where we have introduced an abbreviated symbol (the first line) for better readability.

3.3 The Weak Value as a Quantum Conditional Expectation

As the simplest case of the examples provided in the previous Section 3.2, let us consider the quantum conditional expectation of the function $f(a, b) = a$. Based on the formula (104), one obtains

$$\begin{aligned}
\mathbb{E}^\alpha [A|B = b; \rho] &= \frac{1+\alpha}{2} \cdot A_w^\rho(b) + \frac{1-\alpha}{2} \cdot A_w^\rho(b)^* \\
&= \text{Re} [A_w^\rho(b)] + i\alpha \text{Im} [A_w^\rho(b)], \quad \alpha \in \mathbb{C}, \tag{105}
\end{aligned}$$

where we have introduced the Aharonov's weak value

$$A_w^\rho(b) := \mathbb{E}^1[A|B = b; \rho] = \frac{\text{Tr}[E_B(b)A\rho]}{\text{Tr}[E_B(b)\rho]}.$$

The quantity (105), defined as the complex convex combination of the weak value, was initially proposed in [18], in which it was called the *two state value*. Provided that the density operator $\rho = |\psi\rangle\langle\psi|$ is an orthogonal projection onto the 1-dimensional linear subspace spanned by a unit vector $|\psi\rangle \in \mathcal{H}$, weak value reduces to its familiar form

$$A_w^\psi(b) = \frac{\langle b|A|\psi\rangle}{\langle b|\psi\rangle}. \quad (106)$$

In this respect, the weak value admits an interpretation as one manifestation of the possible family of quantum conditional expectations corresponding to the specific choice $\#\alpha$, $\alpha = 1$, of the subfamily of the QJSDs.

The interpretation of the weak value, as one of the possible candidates of quantum analogues of classical conditional expectations, has been proposed earlier in several literatures. Relevant to our framework presented here is [19], in which the inherent non-uniqueness of the QJP distributions for non-commuting observables is particularly emphasised, a fact which is oftentimes overlooked in discussing the weak value. There, quantum conditional expectations are computed not only for the Kirkwood-Dirac distribution (76), but also for several other types of QJP distributions, including the Wigner-Weyl distribution (75) and the Margenau-Hill distribution (77).

4 Summary and Discussion

In the former part of this paper (Section 2), we focused on the problem of *quantisation* of classical observables and *quasi-classicalisation* of quantum states. For the simplest case in which the observables concerned are all simultaneously measurable, we reviewed that the joint-spectral measure (JSM) uniquely attributed to the commuting observables gives rise to the unique pair of functional analysis and the Born rule, which could respectively be considered as the trivial realisation of quantisation and quasi-classicalisation. Specifically, by taking the duality relation between observables and states into account, we saw that quantisation and quasi-classicalisation are actually adjoint to one another as maps, and thus the JSM, quantisation and quasi-classicalisation are all equivalent as an entity, although they may differ in concept. In this sense, we occasionally referred to the pair of maps as quantum/quasi-classical transformations or representations.

We next considered the general case in which observables concerned are arbitrary. To this, we let ourselves be guided by the observation above, and introduced the concept of *quasi-joint-spectral distributions* (QJSDs), which could be interpreted as non-commutative analogues to the standard JSM. In contrast to the commutative case, QJSDs attributed to a given set of non-commuting observables are non-unique, and thus they give rise to various distinct pairs of quantisations and

quasi-classicalisations. We also discussed the basic properties of QJSDs and the transformations between them. An important implication of this framework is that, although there may be countless possible ways to construct quantisation and quasi-classicalisation, there is a precise one-to-one correspondence between them. These realisations help us to understand the relation between various proposals made historically, including those proposed by Wigner-Weyl, Kirkwood-Dirac, Margenau-Hill, Born-Jordan, Hushimi and Glauber-Sudarshan.

As an application to this framework, the latter portion of this paper (Section 3) focused on the problem of constructing quantum analogues to the classical concept of correlation and conditioning. We proposed a framework to this problem by means of QJSDs introduced earlier, and demonstrated that some of the statistical properties familiar in classical probability theory are still preserved even under the quantum counterpart, especially the relation between correlation and conditional expectation. We finally mentioned that Aharonov's weak value could be interpreted as one manifestation of quantum conditional expectations. One of the virtues of this interpretation is that it reveals a novel aspect of the uncertainty relations [20], in the sense that the weak value appears as the optimal choice of approximation: quantum conditional expectations are best approximations of an observable by another observable, just as classical conditional expectation is the best approximation of a random variable by means of another.

The framework of the quantum/quasi-classical transformation proposed in this paper may find a variety of applications. In fact, it should be obvious from our arguments that one can always draw an analogy to various concepts and results in classical probability theory when one considers the quantum counterparts obtained by this method. Naturally, this will allow for an intuitive treatment of the latter based on the statistical and geometric structures present in classical probability theory.

References

1. W. K. Heisenberg. Über den anschaulichen Inhalt der Quantenmechanik. *Z. Phys.*, 43:172, 1927.
2. M. Ozawa. Universally valid reformulation of the heisenberg uncertainty principle on noise and disturbance in measurement. *Phys. Rev. A*, 67:042105, 2003.
3. M. Ozawa. Uncertainty relations for joint measurements of noncommuting observables. *Phys. Lett. A*, 320:367, 2004.
4. H. Weyl. Quantenmechanik und Gruppentheorie. *Zeitschrift für Physik*, 46(1):1–46, 1927.
5. E. Wigner. On the Quantum Correction For Thermodynamic Equilibrium. *Phys. Rev.*, 40:749, 1932.
6. Y. Aharonov, D. Z. Albert, and L. Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351, 1988.
7. Y. Aharonov, P. G. Bergmann, and L. Lebowitz. Time Symmetry in the Quantum Process of Measurement. *Phys. Rev.*, 134:B1410, 1964.
8. M. Ozawa. Universal uncertainty principle, simultaneous measurability, and weak values. *AIP Conf. Proc.*, 1363:53–62, 2011.
9. H. F. Hofmann. Reasonable conditions for joint probabilities of non-commuting observables. *Quantum Stud.: Math. Found.*, 1:39, 2014.

10. M. O. Scully and L. Cohen. *The Physics of Phase Space Nonlinear Dynamics and Chaos Geometric Quantization, and Wigner Function*, volume 278 of *Lecture Notes in Physics*, pages 253–260. Springer-Verlag, 1987.
11. J. G. Kirkwood. Quantum Statistics of Almost Classical Assemblies. *Phys. Rev.*, 44:31, 1933.
12. P. A. M. Dirac. On the Analogy Between Classical and Quantum Mechanics. *Rev. Mod. Phys.*, 17:195–199, 1945.
13. H. Margenau and N. R. Hill. Correlations between measurements in quantum theory. *Prog. Theoret. Phys.*, 26:722, 1961.
14. M. Born and P. Jordan. Zur Quantenmechanik. *Z. Phys.*, 34:858, 1925.
15. K. Husimi. Some formal properties of the density matrix. *Proc. Phys. Math. Soc. Jap.*, 22:264, 1940.
16. R. J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131:2766, 1963.
17. E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Phys. Rev. Lett.*, 10:277–279, 1963.
18. T. Morita, T. Sasaki, and I. Tsutsui. Complex probability measure and Aharonov’s weak value. *Prog. Theor. Exp. Phys.*, 2013(053A02), 2013.
19. Justin Dressel. Weak values as interference phenomena. *Phys. Rev. A*, 91(032116), 2015.
20. J. Lee and I. Tsutsui. Uncertainty relations for approximation and estimation. *Phys. Lett. A*, 380:2045, 2016.