Could a Classical Probability Theory Describe Quantum Systems?

Jinshan Wu[1], Shouyong Pei[2]
1. Department of Physics & Astronomy,
University of British Columbia,
Vancouver, B.C. Canada, V6T 1Z1
2. Department of Physics, Beijing Normal University
Beijing, P.R. China, 100875
(Dated: February 28, 2011)

Quantum Mechanics (QM) is a quantum probability theory based on the density matrix. The possibility of applying classical probability theory, which is based on the probability distribution function (PDF), to describe quantum systems is investigated in this work. In a sense this also addresses the question of the possibility of a Hidden Variable Theory (HVT) of quantum systems (QSs). Unlike Bell's inequality, which is respected or not by QSs needs to be checked experimentally, in this work HVT is ruled out by theoretical consideration. We propose five experimental facts of QSs as test stones of any quantum theories. Our approach here is to construct explicitly the most general HVT, which agrees with the five QS facts and to check its validity and acceptability. Our five facts include facts concerning subsequently repeated quantum measurements (in the sense of quantum non-demolition measurement). We show that those facts play an essential role at ruling out classical theories even on a single spin-$\frac{1}{2}$ quantum object. We also examine Bell's HVT and Bohm's HVT against the five facts and rule them out.

PACS numbers: 03.65.Ta, 03.65.Aa, 03.65.Ud

I. THE QUESTION AND THE COMMON GROUND TO START THE DISCUSSION

We regard quantum mechanics (QM), a theory based on wave amplitude $|\phi\rangle$ or density matrix ρ, as a quantum probability theory (QPT) as it possesses the following properties: first, for a given complete set of orthogonal vectors $\{|\mu\rangle\}$, it gives a classical probability distribution $\langle\mu|\,\rho\,|\mu\rangle$; second, for any other such vector sets related with the former one by unitary transformations, say $|\nu\rangle = \sum_{\mu} U_{\nu\mu}\,|\mu\rangle$, it also gives another classical probability theory, with probability distribution $\langle\nu|\,\rho\,|\nu\rangle$, which is related by the same unitary transformations, $\langle\nu|\,\rho\,|\nu\rangle = \langle\mu|\,U^{\dagger}\rho U\,|\mu\rangle$. Meanwhile, what we mean by classical probability theory (CPT), is a theory based on the classical probability distribution function (PDF) instead of the density matrix, which possesses only the first property.

The goal of this work is to examine the possibility of describing quantum systems (QSs) by a CPT. To put it another way, it is the possibility of replacing the QPT, which is a successful theory of QSs, equivalently by a CPT. The question is whether exists a map from density matrices, which can be off-diagonal, to probability distribution functions, which are diagonal. There are already famous answers to this question from two different aspects. First, the Bell's Theorem[1, 2] rules out local (or non-contextual) classical theories for QSs provided that it is an established (although currently it is still not) experimental fact that Bell's inequality is violated by QSs. Second, Theorem 7.1 in Ref.[3] proves that there is an equivalent HVT model for every statistical model, including QM, which is also a statistical model. In a sense these two already provide the ultimate answer to the above question: there are such CPTs for QSs, but they have to be non-local (contextual). However, some physicists are willing to accept contextual theories[4] of QSs since it is a philosophical matter, or a matter of taste. More physical significance of such CPTs for QSs, especially those concerning mathematical simplicity and concrete mathematical relations among physical quantities, should be investigated and revealed to check the acceptability of such CPTs. Furthermore, as we mentioned above ruling out local CPTs for QSs relies on further experiments. In this work, we will try to uncover more inevitable mathematical features of such CPTs and examine their physical meaning. In doing so, we build up our theories on only well-established experimental results. In the following, we have listed five QS facts (QS-I to QS-V) respected by all quantum measurement experiments. As demonstrated later, in this work we show that if only QS-I and QS-II are required it is not impossible to have a classical theory, but it is impossible to have a classical probability theory respecting all five facts and convex property of mixed states (CPMS, defined later). We will see next how we address this question and why we claim the answer is negative.

First of all, we clarify our terminology and establish an unambiguous language as a starting point of this discussion. Firstly, it is necessary to distinguish between the terms QM and QS. By QM we refer to the usual axiomatized system of quantum theory while QS is reserved to refer to systems showing quantum properties in experiments. We will explicitly define both later. We require theories to respect behavior of QS, not any axioms or derived theorem in QM. For example, the uncertainty principle, is regarded as a derived theorem in QM, but not an observed behavior of QS. Of course, one may argue about that, but we distinguish them explicitly for reasons that will become apparent soon. Secondly, in this work, we limit our attention to description of static quantum systems, excluding quantum evolution. Hence, only axioms about quantum measurement in QM and only quantum measurement experiments are the subjects we will focus on. For example, if we say CPT can describe a QS, it means CPT can explain all quantum measurement results of that QS. We wish to, at this point, avoid the discussion of evolution of QSs in a form of CPT because evolution is less nontrivial but more technically intense. Furthermore, we will focus on only projective measurements. It is not because that we think extending the projective measurement to Positive Operator Valued Measure (POVM)[5] is trivial, but that we believe when we aiming towards a classical theory of quantum systems we require that the theory at least is capable of describing the subset of quantum measurements — projective measurements. Thirdly, we deal only with two systems, namely a single $\frac{1}{2}$-spin system and an entangled two $\frac{1}{2}$-spin system. In QM language, they both have finite dimension. Discussion on the above systems can easily be generalized to general QSs including infinite systems with continuous variables.

The QSs are systems with the following properties:

QS-I There are a set of physical quantities associated with the system whose values we can measure. For each of them, when measurement is performed on a state of the quantum system, only finite outcomes will be observed. In addition, for every **single** measurement, only **one** specific outcome appears.

QS-II If the same state is prepared, i.e. different realizations of the system go through the same preparation procedure, and the same measurement is performed on this ensemble, there is a statistical limit for the chance of the appearance of a certain outcome.

QS-III If in some way the state of the quantum system is not destroyed as in the case of quantum non-demolition measurement and it can be measured again upon the resulted state of the previous measurement, then a **subsequent measurement of the same physical** quantity will give us **the same outcome** as in the previous measurement, with probability 1.

QS-IV If a **subsequent measurement** is made but **of a different physical quantity**, then still finite possible outcomes will be observed, with again **one outcome from each single measurement**, along with their existing statistical limits.

QS-V **States, which are represented by the same density matrix in the mathematical form of the usual QM**, can be prepared by possibly different procedures but **are not distinguishable** by any quantum measurements. One example, which can be generalized easily, is the following two states of a spin $\frac{1}{2}$: state I is prepared as following, with probability $\frac{1}{4}/\frac{3}{4}$, we use the apparatus to prepare the spin into the *up-/down-* state along z direction; state II is prepared with half possibility into the *up*-state along $\vec{r}_1 = \left(\theta = \frac{2\pi}{3}, \phi = \frac{\pi}{2}\right)$ direction and half possibility into the *up*-state along $\vec{r}_2 = \left(\theta = \frac{2\pi}{3}, \phi = \frac{3\pi}{2}\right)$ direction. Here θ, ϕ are the usual spherical coordinators. Let's assume we have an apparatus preparing a spin into any desired states. This gedanken experiment can be checked in real experiments and here we regard this as a fact. This requires that a proper theory of quantum systems should have the same representation for the above two states. Here we have used the notations of the usual QM just to define the states before we construct the formalism of the alternative theories.

The possibility or impossibility of measuring two physical quantities simultaneously, which while playing an important role in logic chains of usual QM, has not been taken as one of the QS facts here.

The usual QM realizes those properties of a QS through axioms:

QM-I States of a quantum object are $N \times N$ hermitian positive-defined normalized matrices ρ on a complex linear space $\mathcal{H}$ with dimension N, equipped with a definition of inner product. The set of such density matrices is denoted as $\mathcal{N}(\mathcal{H})$, the normalized positive operators over $\mathcal{H}$.

QM-II Physical quantities are hermitian operators on $\mathcal{H}$. Their set is denoted as $\mathcal{O}(\mathcal{H})$. Physical quantities are measurable. The measurement of A on a system at state ρ, results event α (meaning value of observable A is recorded as α) with probability p_α. α is one of the eigenvalues of A (assumed non-degenerate but could be trivially generalized) and $p_\alpha = \langle\alpha|\rho|\alpha\rangle$.

QM-III The state of the object after measurement, given the observed value is α, is $|\alpha\rangle\langle\alpha|$.

In a finite dimensional Hilbert space, the number of eigenvalues of an operator is finite. QM-II realizes both QS-I and QS-II. QM-III realizes QS-III in that if the same measurement is repeated subsequently, the outcome must be α and with probability 1.

In order to realize QS-IV, one needs to consider basis transformations in Hilbert space, that one vector could be expanded under difference bases. After measurement of A provided the outcome is event α, the system stays at $\rho = |\alpha\rangle\langle\alpha|$. If one then measures for example B with eigenvalues $\{\beta\}$, then according to QM-II, the event of a specific β will appear with probability $p_\beta = \langle\beta|\rho|\beta\rangle = \langle\beta|\alpha\rangle\langle\alpha|\beta\rangle$. This explain QS-IV.

The usual QM implements QS-V by using **convex property of mixed states** (CPMS): $\rho = p\rho_1 + (1-p)\rho_2$ for state (ρ) prepared by two mutually exclusive procedures (ρ_1, ρ_2) with respectively probability p and $(1-p)$. One can confirm that state II in QS-V leads to the same density matrix with state I in QS-V via the probability summation rule,

$$\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{3}{4} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \frac{1}{4} & i\frac{\sqrt{3}}{4} \\ -i\frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \frac{1}{4} & -i\frac{\sqrt{3}}{4} \\ i\frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix}. \quad (1)$$

Therefore no measurement can tell their difference. Conceptually we want our alternative theory, whatever it is, to respect CPMS. Our usual QM respects CPMS. Therefore QM realizes all the five experimental facts of a QS.

Next we construct a classical theory, which does not need to respect the QM axioms at all, but still respects QS-I, QS-II, QS-III, QS-IV and QS-V, also preferably CPMS. In a usual discussion of Hidden Variable Theory (HVT), only the first two, QS-I and QS-II, are required to be respected by the theory. We will see that if only these two are required it is not impossible to have a classical theory, but it is impossible to have a CPT respecting all five facts and CPMS.

We have to mention that firstly usually the state is destroyed after measurement, however, quantum nondemolition measurements[6] (QNM) allows us to make subsequently repeated measurements. Secondly, they are not valid for general POVM. However, as we discussed above here we consider only projective measurements. Therefore we take QS-III/IV also as an experimental fact. Also worth mentioning is that we did not include finite accuracy of real measurements into our experimental facts. Usual QM (QM-I, II, III) embraces non-zero commutators between operators so it supports the idea of the Uncertainty Principle. However, as argued by Bohm [7], on the fundamental level one could not tell if it is really impossible to measure some quantities simultaneously or it is just because of the problems of limited technology or accuracy of experiments. This gives us the possibility to relax non-commutation relations between physical quantities when necessary.

Explicitly what we refer to as a CPT means the following:

CPT-I States form set of event Ω. There is a map P from σ-Algebra $\mathcal{F}$ of Ω to $[0,1]$. P satisfies the Kolmogorov axioms of probability[8]. Only physical quantities corresponding to members of $\mathcal{F}$ are observable. A simpler case, which is quite often the case of a physical system, is that the set Ω is a set of countable simple events, which are mutually exclusive, and $\mathcal{F}$ is the trivial topology, set of all subsets of Ω. In our discussion, we only work with this simple case. For exclusive events, if $A \cap B = \phi, A, B \in \mathcal{F}$, then $P(A \cup B) = P(A) + P(B)$. And for independent events, $A \otimes B \in \mathcal{F}(\Omega_1 \otimes \Omega_2)$ where $A \in \mathcal{F}(\Omega_1), B \in \mathcal{F}(\Omega_2)$, then $P(A \otimes B) = P(A) \cdot P(B)$.

CPT-II Observables are $A \in \mathcal{F}$. When the measurement of any such A is performed, every value of $\omega \in A \subseteq \Omega$ can be observed, with corresponding probability $P(\omega)$.

CPT-III After the measurement, the system is at the observed state. Provided event ω is recorded, the state of the object after measurement is ω.

The validity of CPT-III is not really explicitly defined in the usual probability theory. CPT itself does not address this at all, but CPT is usually interpreted in this way. For example, imagine a truly random perfect die. After it is measured, say showing number 6, then one would like to say it is at the state with face value 6. However, what it really means experimentally, is if the die is measured again, it is guaranteed that one will observe the same value, here 6, with probability 1. Therefore, although CPT-III is quite natural, it can be altered if necessary. We also notice CPT respects CPMS as implied in CPT-I.

After clarifying the terminology, the question we seek to discuss is better defined. We are looking for a CPT, which follows CPT-I, CPT-II, CPT-III, of two systems: a single $\frac{1}{2}$-spin and two entangled $\frac{1}{2}$-spins, which both possess QS-I, QS-II, QS-III, QS-IV and QS-V. We will set the state of the single-spin system at $|\uparrow\rangle_x$ and the two-spin system at a singlet state $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. Although we aim at using CPT as an alternative theory for quantum systems, we will still use the usual QM language to denote their states. In other words, we admit QM is a theory for quantum systems but we seek to determine if QS can also be described by a classical theory such as CPT. One key difference between the above QM and CPT, is with/without superposition principle (or coherent summation), as it is usually called by physicists, that $(\mu + \nu) \in \mathcal{H}, \forall \mu, \nu \in \mathcal{H}$ but $\mu + \nu \notin \Omega, \forall \mu, \nu \in \Omega$. This is the central difference between QM and CPT. From this viewpoint, our work here is to find a theory without this superposition principle but still describing successfully all QS facts.

In fact, Spekkens has demonstrated a "classical" toy theory with the superposition principle enabled.[9] In that case, representations of the states, if put in forms of density matrices, depend on the basis vectors. Therefore, the density matrices could have non-zero diagonal elements. We believe that the key difference between a classical theory and a quantum theory is indeed the validity of the superposition principle, or in another word, existence of non-zero off-diagonal elements in the density matrices. Hence, in this sense, Spekkens's toy theory has extended the scope of classical theories. Throughout this work, we are looking at the possibility of describing quantum systems by a theory without the superposition principle. Therefore, discussion of Spekkens's toy theory and its alike is out of the scope of this work. In CPT-I, we have limited the scope of classical theories to be the classical probability theory built upon countable mutually exclusive simple events. There is no coherent summation among classical events.

To provide a unified language for both the classical and quantum theory in our discussion, in §II we will put both CPT and QPT formally into density-matrix forms. In section §III we discuss why we want to have such a map. After that we will present a CPT for a single-spin system and a CPT for a two-spin system, in section §V and §VI respectively. We will see that what kind of CPT is necessary to fully describe quantum systems. We will then discuss why our CPT violates Bell's inequality and what is the possible interpretation of such CPT. Finally in section §VII we conclude that if we are willing to accept all the prices we have to pay to have such a CPT for quantum system, our CPT could be the one. However, we will find that in that case it is even harder to be understood as compared with the usual QM.

II. DENSITY MATRIX LANGUAGE FOR BOTH CLASSICAL AND QUANTUM SYSTEMS

In density matrix language for QM, the state of a quantum object is represented by a density matrix $\rho^q(t)$. The evolution is described by a unitary transformation $U(t) \triangleq U(0,t)$ as

$$\rho^q(t) = U(t)\,\rho^q(0)\,U^\dagger(t)\,, \tag{2}$$

where generally $U(t)$ is determined by H, the Hamiltonian of the quantum object. For a pure initial state, the above density matrix formalism is equivalent with the usual wave function or right vector formalism, but it can also describe a mixed state. For example, we can consider an exclusive mixed state as used in Von Neumann's picture of quantum measurement[10],

$$\rho^q = \sum_i p_i \left|\phi_i\right\rangle\left\langle\phi_i\right|, \tag{3}$$

where $\{|\phi_i\rangle\}$ is a set of orthogonal normalized vectors. According to Von Neumann's picture, the meaning of such an exclusive state is that every sample of this object chooses one of $\{|\phi_i\rangle\}$ with probability p_i.

This interpretation reminds us of the PDF of a truly random classical object (TRCO), which generally should be included as an object of classical mechanics (CM). A state of a TRCO is a PDF $p(x)$ normalized over $\Omega = \{x\}$, the set of all its possible states. It's therefore possible to rewrite this PDF as a density matrix

$$\rho^c = \sum_{x\in\Omega} p(x)\left|x\right\rangle\left\langle x\right|, \tag{4}$$

which gives exactly the same information provided by a PDF. One could just regard this as another notation of a discrete PDF. For the purpose of normalization we also require that simple events are exclusive,

$$\langle x \mid x'\rangle = \delta(x - x')\,, \tag{5}$$

where $\delta(x-x')$ is the Kronecker delta for the discrete set Ω. In QM, generally a state of a quantum system is a full structure density matrix, while in CM, a state of a TRCO is a diagonal density matrix. Considering only discrete sets allows us to use the notation $|x\rangle$ and inner product $\langle x\,|y\rangle = \delta_{xy}$ without any problem. Although physicists use such notation also for continuous systems, mathematicians do not use $|x\rangle$ as a basis vector, or use Dirac δ function as a basis of function space, for continuous systems. Most expressions in physicists' notation can be mapped onto rigorous mathematicians' notation[11], but we do not want to discuss that in any more details here. We limit our description to discrete systems only. For example, a perfect die is such an object.

From this point forward, we are going to use density matrix notation for both QS and TRCO. Furthermore, we can even formally construct a similar dynamical theory to describe classical evolution processes. For example, if we denote the evolution process as a linear operator $\mathcal{T}$, then time evolution of such classical objects can be defined as

$$\rho^c(t) \triangleq \mathcal{T}(\rho^c(0)) = \sum_x p(x)\,\mathcal{T}(|x\rangle\langle x|)\,. \tag{6}$$

It is possible to construct the relation between $\mathcal{T}$ and the classical Hamiltonian H, although here we will not do so explicitly. Formally, we can use the evolution operator T as $\mathcal{T}(|x\rangle\langle x|) = (T|x\rangle)\left(\langle x|T^\dagger\right) = |x(t)\rangle\langle x(t)|$, so

$$\rho^c(t) = T\rho^c(0)\,T^\dagger. \tag{7}$$

Also $TT^\dagger = T^\dagger T = I$, which can be proven as follows: first, for a system fully determined by x, we have $\delta(x(t) - y(t)) = \delta(x-y) = \langle x \mid y\rangle$, then

$$\langle x|\,T^\dagger T\,|y\rangle = \langle x(t) \mid y(t)\rangle = \delta(x(t) - y(t))\,. \tag{8}$$

Therefore, both QM and CPT are unitary evolution theories of density matrices, while the difference between them is the existence of off-diagonal elements. From this point of view, our task in this work is to put a full-structure density matrix into a diagonal density matrix. We call this a question of finding a diagonalization map. The reason we introduce TRCOs is to help towards the understanding of classical objects and, later, quantum objects. By emphasizing "truly random", we are not referring to objects which behave randomly because of the uncertainty in their initial conditions. For a TRCO there is intrinsically no way, even for "God", to tell its real state before a measurement is performed. We only can say it stays in a classical mixed state. One may argue that a physical classical object is not a truly random object. Imagining such a TRCO, however, will help us to understand the classical and quantum measurement process.

To conclude this section, we want to point out that our language of the diagonal density matrix for CPT and the non-diagonal density matrix for QM provides a unified description of classical and quantum mechanics such that relation between classical and quantum theory can be discussed in this language. Besides this, statistical operators[12] and C^*-algebra[13] also provide languages unifying description of classical and quantum systems. However, one should notice that whichever language is used, the fact that classical systems are described by PDFs and their equivalences, while quantum systems are described by density matrices and their equivalences, has never been changed. Only the forms of those equivalences are different. For example, they are $F^c(t,x,p)$ for classical systems and $F^q(t,x,y)$ for quantum systems in statistical operator formulation; and in C^*-algebra language[13], states are elements of convex set $\mathcal{S}$, which in the classical case becomes a usual set of simple events and their convex combination, but in the quantum case also allows summation – not only convex summation – of pure states. Another formal difference between CM and QM in C^*-algebra language is commutative/non-commutative relation between observables. From this viewpoint, in this work, we are searching for a theory based on commutative C^*-algebra for QSs. As we will see later, while different forms of diagonal density matrices are tried in searching for CPT for QSs, algebraic relation among observables is changed at the same time. However, in the present work we prefer density matrix language to this C^*-algebra language simply for consideration of familiarity of density matrices among physicists.

III. WHY ARE WE LOOKING FOR SUCH A MAP?

If we have a CPT as desired, it is an HVT of quantum systems. Quantum systems are no longer quantum but TRCOs. Therefore, one can understand quantum measurement if one can understand measurement of TRCOs. The most straightforward picture of a measurement is a measurement on a deterministic classical object. Assumed as a discrete system, it stays in state $|x\rangle\langle x|$ before it is measured but the observer does not know. After the measurement we get the information that it was in state $|x\rangle\langle x|$ and it remains in state $|x\rangle\langle x|$. The less straightforward picture of a measurement is a measurement on a statistically random classical object. Here the term "statistically random" means that the nature of this object is still deterministic, but with incomplete information it appears as a random object. For every given such object, we just do not know its state but it is already fixed. This also means its randomness is only meaningful as in an ensemble. This is called statistical interpretation of probability theory. Again for such an object, it is in state $|x\rangle\langle x|$ before measurement (the observer does not know) and after the measurement we get the information that it was in state $|x\rangle\langle x|$ and it remains in state $|x\rangle\langle x|$. Notice that although x can be one element of a large set, it is fixed according to probability $p(x)$ before the measurement is performed.

Measurement on a TRCO is less understandable. Assuming such an object really exists for the moment, its state is unknown before measurement. After the measurement we find with probability $p(x)$ that it was in state $|x\rangle\langle x|$ and it remains in state $|x\rangle\langle x|$ afterwards. Here we find that a phenomena so-called "collapse" of probability function has occurred. While this seems less understandable, both "statistical randomness" and "true randomness" give us the same measurement result. One could never distinguish which is the "real" one from measurements. It is a pure philosophical question to ask whether the state is statistically random or truly random and so from now on we will treat them as the same.

Now imagine we have two correlated TRCOs which have exactly the same states, but unknown. Since they are both TRCOs we do not know their states before the measurement. If we measure one of them, say we find that it stays in state $|x\rangle\langle x|$, then we immediately know the state of the other object is also $|x\rangle\langle x|$. In this sense, if we assume the existence of such TRCOs, "spooky action"[14] exists even in classical mechanics. Quantum "spooky action" in entangled systems is not stranger than its classical version at all. They are different just in that in the quantum singlet state of two spins, there is a freedom in choosing which direction ($\hat{n}$) to measure. This is related to basis transformation, which in turn is related again to the superposition principle.

Provided there is a TRCO description of quantum systems, the two problems of quantum measurement, namely collapse of the wave function and measurement of entangled states, become collapse of the probability function and measurement of classical correlated states in measurement of TRCOs. This implies that if one believes that measurement of TRCOs is understandable, then measurement of quantum systems is also understandable.

Here we assume TRCO Assumption: there is no difficulty or confusion in understanding measurement of TRCOs. Even if it is questionable, if TRCO can describe quantum systems, then we know the problem of quantum measurement comes from classical probability theory and has nothing to do with any other quantum nature. Of course the situation will be different if we find out that TRCOs can not describe quantum systems.

We can formally compare measurement of TRCOs and quantum systems. Here we include both auxiliary system m and object system o explicitly into our formal description. The measurement includes three steps:

CMeaure-I A classical correlated state is formed by an interaction process, so that from an initial state of

$$\rho^{c,o} = \sum_x p(x) |x\rangle \langle x| . \tag{9}$$

we get

$$\rho^{c,o} \otimes \rho^{c,m} \longrightarrow \rho^{c,om} = \sum_x p(x) |x \otimes M(x)\rangle \langle x \otimes M(x)| . \tag{10}$$

CMeaure-II When we only check the value recorded on the auxiliary system, we get the auxiliary system's partial distribution, which is

$$\rho^{c,m} \triangleq tr^o (\rho^{c,om}) = \sum_x p(x) |M(x)\rangle \langle M(x)| , \tag{11}$$

where tr^o means the trace is performed over the object's state space, a standard procedure in probability theory when only information on the partial distribution is needed.

CMeaure-III According to the exclusiveness nature of $|M(x)\rangle \langle M(x)|$ and CPT-III, the sampling process gives us one specific state $M(x^*)$. This happens with the desired probability $p(x^*)$, due to CPT-II. $M(x^*)$ on the auxiliary system means x^* on the measured object.

However, since the general quantum density matrix has non-zero off-diagonal terms, the picture of quantum measurement is slightly different:

QMeaure-I An interacting process evolves a quantum system to establish entanglement

$$\rho^{q,o} \otimes \rho^{q,m} \longrightarrow \rho^{q,om} = \sum_{\mu\nu} \rho_{\mu\nu} |\mu \otimes M(\mu)\rangle \langle \nu \otimes M(\nu)| , \tag{12}$$

where initially

$$\rho^{q,o} = \sum_{\mu\nu} \rho_{\mu\nu} |\mu\rangle \langle \nu| . \tag{13}$$

QMeaure-II When we check only the value recorded on the auxiliary system, we obtain the auxiliary system's partial distribution,

$$\rho^{q,m} \triangleq tr^o (\rho^{q,om}) = \sum_\lambda \rho_{\lambda\lambda} |M(\lambda)\rangle \langle M(\lambda)| . \tag{14}$$

QMeaure-III This then becomes a "classically" quantum state, which has exactly the same form of Eq. (11) under the given basis. Therefore, the sampling process gives us one specific state $M(x^*)$. This happens with the desired probability $p(x^*)$ and $M(x^*)$ on the auxiliary system means x^* on the measured object.

However, we should point out that for a quantum object, equ(14) is not a copy of equ(13), while equ(11) is an exact copy of equ(9) for a classical object. Therefore, if TRCOs could never describe quantum system, even with the TRCO Assumption, quantum measurement is still harder to understand than measurement of TRCOs. However, if we have a CPT for quantum system, then quantum measurement is just as understandable as measurement of a TRCO. Instead of making measurement as a coherent part of evolution, we have shown above that even classical measurement is different with classical evolution. Therefore the fact that the same holds for quantum measurement is not a surprise.

As we have seen, due to CPT-III, a classical measurement ends up with an exact copy of the object state. In fact, this is called a broadcasting and it has been proven that a quantum system can not be broadcast in quantum no-broadcasting theorem (QNBT)[15]. Unless the object system initially stays in one of a set of **known** orthogonal states, a quantum system can not be broadcast. In our language, this means when a system is in a classical probability combination of known orthogonal states, *i.e.* a diagonal density matrix under a known basis, it can be broadcast. This is a broadcast of a TRCO.

An arbitrary unknown state of a TRCO can be broadcast, or a diagonal density matrix state can be broadcast. Therefore, if the above diagonalization mapping exists, through it, a quantum system can also be broadcast. This would conflict with QNBT, which is proven in the language of usual QM. This leads to two possibilities: firstly, QNBT holds and diagonalization mapping does not exist; or secondly, QNBT is not valid and the mapping exists. Now we find that QNBT is also reduced to the existence of the diagonalization mapping. Therefore, it seems all the confusing and "extraordinary" problems in QM including quantum measurement, HVT and QNBT come down to one question, the existence of such diagonalization mapping.

The relation between QNBT and HVT can be shown more explicitly. A TRCO can be broadcast, by introducing a classical hidden variable. For example, let us use a perfect two-faced die as a TRCO. We introduce a classical signal λ, generated from a given PDF $\rho(\lambda)$ over $\Gamma = \{\lambda\}$. The state of the die is determined by this signal as follows,

$$\rho^{c,o} = \sum_{\lambda\in\Gamma} \rho_+(\lambda) |+\rangle \langle +| + \sum_{\lambda\in\Gamma} \rho_-(\lambda) |-\rangle \langle -| , \tag{15}$$

where it is required that

$$\int_\Gamma d\lambda \rho_+(\lambda) = \frac{1}{2} = \int_\Gamma d\lambda \rho_-(\lambda) , \tag{16}$$

where $\int$ could be interpreted as $\int$ or $\sum$ for continuous or discrete variables respectively. Similarly, later on $\sum$ should be understood in the same way. We then duplicate this hidden variable signal, send a copy to another die while the original signal is sent to the original die.

Each die determines its state respectively according to the value of its hidden variable. Now we get a broadcast of the die. In this sense, it is fair enough to say that the success of an HVT for CM makes it possible to broadcast a classical object. So what about an HVT for QM?

IV. A POSSIBLE TRCO AND UNDERSTANDING OF ITS MEASUREMENT

Consider a quantum system coupled with a large thermal bath and assume that we have the technology to quickly measure eigenvalues of Hamiltonian of the central quantum system. It is so quick that compared with the relaxation time of this composed system, it can be neglected. Our measurement is performed once in a while with the time interval between measurements being much longer than the typical relaxation time. The outcomes of such measurements will give us a sequence of eigenvalues whose probability of appearance follows classical Boltzmann distribution. Do we now believe that the system stays in one of the eigenstates before any measurements? Further, does our belief matter? It seems there is no difficulty in accepting the results from this measurement as is. From this example, we wish to argue that our assumption of the existence of TRCO and the validity of TRCO Assumption, which states there is no problem in understanding measurement of TRCO, is plausible.

V. CPT FOR SINGLE-SPIN SYSTEM

The possibility of a CPT or an HVT for quantum system has been long investigated by many great physicists[1–3, 7, 16–18]. Bell's Theorem[1] says that all local HVT should obey the Bell's inequality, which is not respected by QM. Experimental tests suggest that QS does violate the Bell's inequality so QM is a preferred theory of QS[19]. But this statement has not yet been supported by all physicists. Others accept this but turn to construct explicitly contextual HVTs for QS[4]. In the following, we will try to answer this problem in another way. We are willing to go as far as possible to construct a CPT to give consistent results with quantum systems including all five QS facts. If this effort fails we will find where and why; or if it succeeds, we will check whether it is acceptable or not. We are not aiming at any explicitly contextual HVTs, but we will discuss such theories if they have to be so. If it succeeds, according to Bell's Theorem, it should be non-local. It would be interesting to show explicitly the place where non-locality (or contextuality) enters the theory. In fact, Theorem 7.1 in [3] proves that there is an equivalent HVT model for every statistical model, including the QM. QS-I and QS-II can be reproduced. Physical implications of such possible HVT is discussed further by the same author in [18], where it is shown to be not compatible with complementarity and nonseparability. Compared to those works, our work is different in that except the five QS facts, no additional presumably reasonable property of QS is assumed. For example, even CPMS and CPT-III are regarded as theoretical consideration only, so that they are allowed to be modified. We will clearly see later what are the physical implications of such an HVT. In [17], the author discussed a similar question of "How to make quantum mechanics look like a hidden-variable theory and vice versa" using the Wigner distribution. Also in more general coherent-state representation, an effective PDF can be constructed[20] to make the theory appear very similar with a CPT. However, such a PDF can have negative values so it it not really a CPT. Moreover, it does not respect even QS-I. Here in this paper, to discuss the same question, we insist Kolmogorov-type CPT and try to make it successful as far as possible. For the sake of simplicity of language, in this paper, we regard CPT and HVT of a quantum system as having the same meaning and later on just simply call them HVT.

According to our general framework, HVT is in a classical diagonal density matrix form,

$$\rho^{hvt} = \sum_{\lambda \in \Gamma} \rho\left(x\left(\lambda\right)\right) \left|x\left(\lambda\right)\right\rangle \left\langle x\left(\lambda\right)\right|, \tag{17}$$

where x is the dynamic variable, λ is the hidden random variable and $x\left(\lambda\right)$ is an onto mapping, $\rho\left(x\left(\lambda\right)\right)$ is a PDF over $\Gamma = \{\lambda\}$, a set of exclusive events,

$$\left\langle x\left(\lambda\right)\right| x\left(\lambda'\right)\rangle = \delta\left(\lambda - \lambda'\right). \tag{18}$$

One thing that is necessary to indicate here, is the parameter λ is abstract, not limited as a single variable. A successful HVT has to respect all QS facts. We will start from QS-I and II.

A. CPT based on exclusiveness of all elementary pure events

We first consider a single spin-$\frac{1}{2}$ as in Bohm's HVT[7], and then focus on an entangled object with two subsystems as discussed in Bell's inequality[1]. For simplicity, let's just consider a specific quantum state, a $\frac{1}{2}$-spin in the state of $\left|\uparrow\right\rangle_x$, the *up* state of S_x. In the language of QM, it's

$$\rho^{q} = \frac{1}{2}\left(\left|\uparrow\right\rangle_z \left\langle\uparrow\right|_z + \left|\uparrow\right\rangle_z \left\langle\downarrow\right|_z + \left|\downarrow\right\rangle_z \left\langle\uparrow\right|_z + \left|\downarrow\right\rangle_z \left\langle\downarrow\right|_z\right). \tag{19}$$

For an HVT, the first trial density matrix will naturally be,

$$\rho^{hvt} = \sum_{\lambda_z} \rho_+\left(\lambda_z\right) \left|\uparrow\right\rangle_z \left\langle\uparrow\right|_z + \sum_{\lambda_z} \rho_-\left(\lambda_z\right) \left|\downarrow\right\rangle_z \left\langle\downarrow\right|_z, \tag{20}$$

with the following requirement to give correct results for measurement on S_z,

$$\int_{\Gamma_z} d\lambda_z \rho_+\left(\lambda_z\right) = \frac{1}{2} = \int_{\Gamma_z} d\lambda_z \rho_-\left(\lambda_z\right). \tag{21}$$

However, this gives the consistent results with QS-I and II only for S_z measurement. We can also measure S_x. If we still respect the possible non-commutative relation between quantum operators S_x and S_z, then we need to do a basis transformation in $\mathcal{H}^q$ and do measurement of S_x. We get

$$\begin{aligned}\rho^{hvt} = & \tfrac{1}{2}\sum_{\Gamma_z}\left[\rho_+(\lambda_z)+\rho_-(\lambda_z)\right]\left(|\uparrow\rangle_x\langle\uparrow|_x+|\downarrow\rangle_x\langle\downarrow|_x\right) \\ & +\tfrac{1}{2}\sum_{\Gamma_z}\left[\rho_+(\lambda_z)-\rho_-(\lambda_z)\right]\left(|\uparrow\rangle_x\langle\downarrow|_x+|\downarrow\rangle_x\langle\uparrow|_x\right).\end{aligned} \tag{22}$$

We can see that, according to equ(21), the result of this measurement will be $\frac{1}{2}$ probability to get *up* and $\frac{1}{2}$ to get *down*. This is obviously wrong. We know for the specific state we chose above, the correct result of the S_x measurement is the *up* state only. This HVT does not realize QS-I and II.

There is one way to overcome this inconsistency with the price that not only one hidden variable, but also another hidden variable is needed. In order to get correct results for measurement on S_z and S_x, we need

$$\begin{aligned}\rho^{hvt} = \frac{1}{\mathcal{N}}\Bigg[&\sum_{\lambda_z}\rho_+(\lambda_z)|\uparrow\rangle_z\langle\uparrow|_z+\sum_{\lambda_z}\rho_-(\lambda_z)|\downarrow\rangle_z\langle\downarrow|_z \\ &+\sum_{\lambda_x}\rho_+(\lambda_x)|\uparrow\rangle_x\langle\uparrow|_x+\sum_{\lambda_x}\rho_-(\lambda_x)|\downarrow\rangle_x\langle\downarrow|_x\Bigg],\end{aligned} \tag{23}$$

with the requirement,

$$\int_{\Gamma_x} d\lambda_x\rho_+(\lambda_x)=1, \int_{\Gamma_x} d\lambda_x\rho_-(\lambda_x)=0. \tag{24}$$

$\mathcal{N}$ is a normalization constant to keep $tr(\rho)=1$ and here $\mathcal{N}=2$. With this density matrix, a measurement of S_x will give the *up* state only. We can similarly include S_y terms using another hidden variable λ_y. However, a successful HVT should respect QS-I and II for measurement on an arbitrary direction. For this purpose, will three hidden variables corresponding to S_x, S_y, S_z be enough? For example, for a measurement of

$$S_r=\sin\theta\cos\phi S_x+\sin\theta\sin\phi S_y+\cos\theta S_z, \tag{25}$$

on the above state, the possible outcomes are

$$s_r=\frac{1}{2}\left(\sin\theta\cos\phi\pm\sin\theta\sin\phi\pm\cos\theta\right). \tag{26}$$

This could be a continuous number, not only $\pm\frac{1}{2}$. We see that it does not respect QS-I. So QS-I requires one hidden variable for measurement on every direction and abandonment of the inherent relation between operators such as equ(25). Furthermore such a multi-hidden variable density matrix has one very important implication, that according to equ(18), all states (events) corresponding to arbitrary directions should all be exclusive events. Thus we call it an exclusive-event HVT (EHVT). This implies $S_{\vec{r}_1}S_{\vec{r}_2}=0$ and our CPT density matrix has to be

$$\rho^{hvt}=\frac{1}{\mathcal{N}}\sum_{\vec{r}}\left[p_\uparrow(\vec{r})|\uparrow\rangle_{\vec{r}}\langle\uparrow|_{\vec{r}}+p_\downarrow(\vec{r})|\downarrow\rangle_{\vec{r}}\langle\downarrow|_{\vec{r}}\right] \tag{27}$$

where

$$p_\uparrow(\vec{r})=\frac{1+r_x}{2}, p_\downarrow(\vec{r})=\frac{1-r_x}{2}. \tag{28}$$

There is a technical problem and another non-trivial conceptual problem with the above PDF. The technical problem is the value of $\mathcal{N}$. Since we need to keep $tr\left(\rho^{hvt}\right)=1$ and there is an infinite number of directions, $\mathcal{N}$ will be infinity if $tr\left(\rho^{hvt}\right)$ is simply,

$$tr\left(\rho^{hvt}\right)=\sum_{\vec{r}}\left[\langle\uparrow|_{\vec{r}}\rho^{hvt}|\uparrow\rangle_{\vec{r}}+\langle\downarrow|_{\vec{r}}\rho^{hvt}|\downarrow\rangle_{\vec{r}}\right]. \tag{29}$$

One way to define a "proper" $tr\left(\rho^{hvt}\right)$ to avoid such divergence is to decompose $\vec{r}=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ and treat

$$tr\left(\rho^{hvt}\right)=\int d\theta d\phi\sin\theta\left[\langle\uparrow|_{\vec{r}}\rho^{hvt}|\uparrow\rangle_{\vec{r}}+\langle\downarrow|_{\vec{r}}\rho^{hvt}|\downarrow\rangle_{\vec{r}}\right]. \tag{30}$$

In this case, $\mathcal{N}=4\pi$. This introduces additional relative probability between states corresponding to different $\vec{r}$. This may not be a proper definition, however, it is still possible to solve this technical question of divergent normalization constant by some other means. If only relative probability of a given direction $\vec{r}$ is concerned in real measurements, this problem does not affect the outcomes at all. Because of this, in the following, we simply drop the normalization constant $\mathcal{N}$.

The other problem is rather serious. Due to the full exclusiveness between all events, the meaning of a measurement changes. "Measuring $S_{\vec{r}}$" for a specifically given $\vec{r}$ is no longer a pure elementary event but a compound event. A pure elementary event instead would be "measuring S", with no specific direction given. The result of such a measurement will be one direction, which got randomly picked up during the measurement process, and an *up*- or *down*- state, would be recorded correspondingly with the right probability. In this way, there is no guarantee that the randomly picked-up direction will be the desired direction of an observer.

A classical die would be a good example of a classical probability distribution based on all exclusive events. From a perfect 6-faced die, we wish to only measure the relative probability between face 1 and face 2. We could still get all 6 numbers, but we discard all the other four if they turn out to be the outcomes of our measurement. Therefore, effectively, we will find out that the state of the die within the subspace is,

$$\rho^c=\frac{1}{2}\left(|1\rangle\langle 1|+|2\rangle\langle 2|\right). \tag{31}$$

Similarly measurement of S on our ρ^{hvt} will be one of all of the exclusive events, and during our analysis of the results, we can discard all irrelevant events. In real quantum measurements, however, we never find such irrelevant and redundant outcomes. If we measure S_x, according to the all-exclusive nature, in a classical measurement of the above state, sometimes our apparatus detects nothing and sometimes it detects the right state – *up*. However, in a real quantum measurement, assuming there are no further experimental accuracy limits, such detecting-nothing events never happen. What we get is only the *up*- or *down*- state of a given direction. This shows that, in fact, the above state based on exclusiveness is not the desired state. Or if it is, then this is only possible if the system somehow knows the intention of the observer during the process.

This "contextual" relation between system and observer is unexpected, however, some physicists may still be willing to accept such a theory since it is a problem about interpretation not about any predictions from the theory.

Now we will try to make EHVT compatible with QS-III and QS-IV. CPT-III tells us that if all events are exclusive, then after a measurement, for example along the x direction, given the *up* state is recorded, its state is simply $\left|\uparrow\right\rangle_x \left\langle\uparrow\right|_x$. A subsequently repeated measurement along the x direction results in event *up* again. This is the expected result stated in QS-III. However, if the subsequently repeated measurement is along the z direction, for state $\left|\uparrow\right\rangle_x \left\langle\uparrow\right|_x$, the exclusiveness tells us, that there is not any such events as measurement along z direction. So we would again get a detecting-nothing event. This conflicts with QS-IV. Furthermore, the

initial state is the *up* state along the x direction, therefore after a measurement along the x direction, nothing changes. If CPT-III holds, we see the state before and after the measurement is respectively,

$$\rho^{hvt}_{before} = \sum_{\vec{r}} \left[p_\uparrow(\vec{r}) \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + p_\downarrow(\vec{r}) \left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}\right] \tag{32}$$

and

$$\rho^{hvt}_{after} = \left|\uparrow\right\rangle_x \left\langle\uparrow\right|_x, \tag{33}$$

where

$$p_\uparrow(\vec{r}) = \frac{1+r_x}{2}, p_\downarrow(\vec{r}) = \frac{1-r_x}{2}. \tag{34}$$

Eq.(32) and Eq.(33) are obviously different. The state stays the same before and after the measurement, however, we find its expressions are different before and after the measurement. This means CPT-III is wrong. A state after it has been revealed in a measurement is not the state corresponding to the measurement result. We will have to also sacrifice CPT-III after abandoning Eq.(25).

CPT-III$'$ After a measurement, the object stays at the state which guarantees a subsequently repeated measurement in accordance with QS-III and QS-IV. For example, for spin-$\frac{1}{2}$ after a measurement on $S_{\vec{r}_0}$ and a *up*-state being recorded, the state is,

$$\rho^{hvt}_{after} = \sum_{\vec{r}} \left[\frac{1+\vec{r}\cdot\vec{r_0}}{2} \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + \frac{1-\vec{r}\cdot\vec{r_0}}{2} \left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}\right]. \tag{35}$$

If a *down*-state is recorded after measurement of $\sigma_{\vec{r}_0}$, we can simply replace $\vec{r}_0$ with $-\vec{r}_0$ in Eq.(35).

This CPT-III$'$ could not be understood easily.

Furthermore, this is not a consistent theory. We already know that an *up*-state along the x direction, before and after measurements of S_x, is Eq.(32). Given this state if we want to calculate the probability of observing the *up*-state along the x direction, we will do

$$\begin{cases} p_{x_{up}} = \left\langle\uparrow\right|_x \rho^{hvt}_{before} \left|\uparrow\right\rangle_x = 1, \\ p_{x_{down}} = \left\langle\downarrow\right|_x \rho^{hvt}_{before} \left|\downarrow\right\rangle_x = 0. \end{cases} \tag{36}$$

This gives the correct answer that the relative probability between *up*- and *down*-state is 1. But notice that we times $\left\langle\uparrow\right|_x$ from the left and $\left|\uparrow\right\rangle_x$ the right to a density matrix to get the probability of $p_{x_{up}}$. In doing so we assume that vector $\left|\uparrow\right\rangle_x$ stands for the event of an *up*-state along the x direction, but it is different from Eq.(32), which is the expression standing for the event of an *up*-state along the x direction as we previously pointed out. We have now two different expressions for the same state in a theory.

Therefore we conclude that the first rescue of HVT, based on the assumption of the exclusiveness among all $\{S_{\vec{r}}\}$, failed to achieve a consistent theory satisfying simultaneously QS-I, QS-II, QS-III and QS-IV. To do so we will not have inherent relation between operators as in Eq.(25), we will have to put a twist on CPT-III and allow "contextual" communication between object and observer. Even after all these, we would not be able to get a self-consistent theory. We will now try another more plausible construction of CPT for QS, based on independence of all pure elementary events. In C^*-algebra language, we have just tried the trivial commutative multiplication $S_{\vec{r}_1} S_{\vec{r}_2} = 0$ among operators $\{S_{\vec{r}}\}$ for the exclusive case. Next let us consider another commutative algebra, direct product $S_{\vec{r}_1} \otimes S_{\vec{r}_2}$ assuming independence among operators $\{S_{\vec{r}}\}$.

B. CPT based on independence of pure elementary events

Although the idea of exclusive events fails, in fact, there is another way to save the idea of HVT, being that λ_z and λ_x are independent events, so that an HVT density matrix could be,

$$\rho^{hvt} = \sum_{\{\vec{\lambda}\}} \rho\left(\vec{\lambda}\right) \left|\ldots, x_{\vec{r}}(\lambda_{\vec{r}}), \ldots\right\rangle \left\langle\ldots, x_{\vec{r}}(\lambda_{\vec{r}}), \ldots\right|, \tag{37}$$

where $\lambda_{\vec{r}}$ is a random variable for direction $\vec{r}$ and $x_{\vec{r}}(\lambda_{\vec{r}}) =\uparrow,\downarrow$. Notation $\vec{\lambda}$ refers to an infinite dimensional vector $(\lambda_x,\cdots,\lambda_y,\cdots,\lambda_z,\cdots)$. Under this independent event assumption, measurement on every direction is done independently. This is only possible if all $S_{\lambda_{\vec{r}}}$, operators corresponding to all directions $S_{\vec{r}}$ are commutative and every operator could be treated independently. We call this independent-event HVT (IHVT).

IHVT requires too abandoning inherent relation as in Eq.(25) and the non-commutative relation between quantum operators. A valid multiplication between operators is the direct product, $S_{\vec{r}_1}\otimes S_{\vec{r}_2}$. A common basis is $\left|(\uparrow_x \text{ or } \downarrow_x),\cdots,(\uparrow_y \text{ or } \downarrow_y),\cdots,(\uparrow_z \text{ or } \downarrow_z),\cdots\right\rangle$. We have an infinite number of hidden variables to represent all directions of measurement. in principle, by choosing appropriate $\rho\left(\vec{\lambda}\right)$ one can always fulfill QS-I and II. For example, the following scheme gives the correct results on measurement of $S_{\vec{r}}$. For direction $\vec{r}=(r_x,r_y,r_z)$, we choose $\lambda_{\vec{r}}\in\left\{-\frac{1}{2},\frac{1}{2}\right\}$, a two-value discrete random variable as the hidden variable. Then we require, $tr^{-\vec{r}}(\cdot)$, the partial trace except direction $\vec{r}$ of ρ^{hvt} gives,

$$\rho_{\vec{r}}^{hvt}\triangleq tr^{-\vec{r}}\left(\rho^{hvt}\right)=\frac{1+r_x}{2}\left|\uparrow\right\rangle_{\vec{r}}\left\langle\uparrow\right|_{\vec{r}}+\frac{1-r_x}{2}\left|\downarrow\right\rangle_{\vec{r}}\left\langle\downarrow\right|_{\vec{r}}, \tag{38}$$

for example, by requiring

$$\rho^{hvt}=\Pi_{\vec{r}}\otimes\rho_{\vec{r}}^{hvt}. \tag{39}$$

Notice the product state in Eq.(39) is just one example, not necessary required, while Eq.(38) is a strict requirement. There are many more density matrices in the form of Eq.(37) and satisfying Eq.(38). Independence of pure elementary events does not lead to independent product states. However, from now on we will simply use Eq.(39) to as an representative of all possible forms of density matrices.

One can check that this satisfies QS-I and II for measurement on an arbitrary $\vec{r}$ direction. The outcomes could be $\uparrow$ or $\downarrow$ with probability of $\frac{1+r_x}{2}$ and $\frac{1-r_x}{2}$ respectively. Furthermore, IHVT does not require contextuality between object and observer. After the partial trace only the desired direction will survive. The partial trace is a standard procedure for an independent random variable. The above explicitly constructed density matrix gives the correct results for measurement on any directions. We see that IHVT is at least as valid as Bell's HVT on a spin $\frac{1}{2}$ object[2] as they both respect QS-I and QS-II. It is less controversial than our former EHVT.

Will IHVT realize QS-III and QS-IV? The answer is "yes" for QS-III. According to CPT-III, after measurement, the system stays at the state corresponding to the observed value of the observable and all the others states corresponding to other observables remain at the same states. For example, when we measure S_z with the outcome being *up*, the state after measurement is

$$\rho_{after}^{hvt}=\frac{\left|\uparrow\right\rangle_z\left\langle\uparrow\right|_z\otimes\left\langle\uparrow\right|_z\rho_{before}^{hvt}\left|\uparrow\right\rangle_z}{tr^{-z}\left(\left\langle\uparrow\right|_z\rho_{before}^{hvt}\left|\uparrow\right\rangle_z\right)}. \tag{40}$$

If measured on S_z again the outcome is still *up*. What if the second measurement is on a different direction, say S_x? We have,

$$\begin{aligned}p_{z_{up},x_{down}} &= tr\left(\left|\downarrow\right\rangle_x\left\langle\downarrow\right|_x\rho_{after}^{hvt}\right)\\ &= \frac{tr^{-x,-z}\left(\left\langle\uparrow\right|_z\left\langle\downarrow\right|_x\rho_{before}^{hvt}\left|\downarrow\right\rangle_x\left|\uparrow\right\rangle_z\right)}{tr^{-z}\left(\left\langle\uparrow\right|_z\rho_{before}^{hvt}\left|\uparrow\right\rangle_z\right)}\\ &= 0,\end{aligned} \tag{41}$$

where we make use of the fact that the state is initially x direction *up*, $\left\langle\downarrow\right|_x\rho_{before}^{hvt}\left|\downarrow\right\rangle_x=0$. However, this number is expected to be $\frac{1}{2}$ according to QM prediction. This shows our HVT does not respect QS-IV if CPT-III holds. Thus we need to modify CPT-III to the following,

CPT-III$''$ After a measurement, the object stays at the state which guarantees that, a subsequently repeated measurement gives the right result, stated in QS-III and QS-IV. For example, for spin-$\frac{1}{2}$ after a measurement on $S_{\vec{r}_0}$ and a *up*-state is recorded, it stays at, ρ^{hvt} satisfying,

$$\rho_{\vec{r}}^{hvt}=\frac{1+\vec{r}\cdot\vec{r}_0}{2}\left|\uparrow\right\rangle_{\vec{r}}\left\langle\uparrow\right|_{\vec{r}}+\frac{1-\vec{r}\cdot\vec{r}_0}{2}\left|\downarrow\right\rangle_{\vec{r}}\left\langle\downarrow\right|_{\vec{r}}. \tag{42}$$

If a *down*-state is recorded after measurement of $S_{\vec{r}_0}$, we can simply replace $\vec{r}_0$ with $-\vec{r}_0$ in Eq.(42).

IHVT implies it is in principle possible to measure S along several directions simultaneously, which is impossible in current practical quantum measurement. If it is proven that it is impossible then IHVT should be discarded, and QM should be the only choice of description of QS.

This version of CPT-III has the same inconsistency as the last exclusive-event HVT. Given a x direction *up*-state, represented by Eq.(38), if we want to calculate the probability of x direction *up*-state, we do

$$p_{x_{up}}=tr\left(\left|\uparrow\right\rangle_x\left\langle\uparrow\right|_x\rho^{hvt}\right)=1. \tag{43}$$

This gives us the correct result, however, with the assumption that $\left|\uparrow\right\rangle_x\left\langle\uparrow\right|_x$ refers to the x direction *up*-state, which is not the state which really means the x direction *up*-state as in Eq.(38). Although this theory suffers all the above difficulties, it is less controversial than the former theory. Now let us try to make it consistent with QS-V.

C. QS-V and twisted CPMS

We have shown that Eq.(39) satisfies QS-I, II, III and IV with a twisted CPT-III. Let us now check if it also respects QS-V and CPMS.

According to CPMS, the first preparation gives us state,

$$\rho^I = \frac{1}{4}\rho^{11} + \frac{3}{4}\rho^{12}. \quad (44)$$

Here ρ^{11} is a product state of $\rho^{11}_{\vec{r}}$ that $\rho^{11} = \Pi_{\vec{r}} \otimes \rho^{11}_{\vec{r}}$,

$$\rho^{11}_{\vec{r}} = \frac{1+r_z}{2} \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + \frac{1-r_z}{2} \left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}. \quad (45)$$

ρ^{12} is a product state of $\rho^{12}_{\vec{r}}$,

$$\rho^{12}_{\vec{r}} = \frac{1-r_z}{2} \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + \frac{1+r_z}{2} \left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}. \quad (46)$$

Similarly, the second preparation gives us state,

$$\rho^{II} = \frac{1}{2}\rho^{21} + \frac{1}{2}\rho^{22}. \quad (47)$$

Here ρ^{21} is a product state of $\rho^{21}_{\vec{r}}$,

$$\rho^{21}_{\vec{r}} = \frac{1+r_y\frac{\sqrt{3}}{2}+r_z\frac{1}{2}}{2} \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + \frac{1-r_y\frac{\sqrt{3}}{2}-r_z\frac{1}{2}}{2} \left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}. \quad (48)$$

ρ^{22} is a product state of $\rho^{22}_{\vec{r}}$,

$$\rho^{22}_{\vec{r}} = \frac{1-r_y\frac{\sqrt{3}}{2}+r_z\frac{1}{2}}{2} \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + \frac{1+r_y\frac{\sqrt{3}}{2}-r_z\frac{1}{2}}{2} \left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}. \quad (49)$$

We can see that

$$\rho^I \neq \rho^{II}. \quad (50)$$

So they are different density matrices thus our HVT does not respect QS-V.

However, we notice that if we compare the results after a partial trace $tr^{-\vec{r}}$ on ρ^I and ρ^{II}, they are the same,

$$tr^{-\vec{r}}\left(\rho^I\right) = \rho^{11}_{\vec{r}} + \rho^{12}_{\vec{r}} = \rho^{21}_{\vec{r}} + \rho^{22}_{\vec{r}} = tr^{-\vec{r}}\left(\rho^{II}\right), \quad (51)$$

$\forall \vec{r} \in \mathbb{R}^3$. QS-V is satisfied. However, now CPMS is twisted that it holds only on the level of reduced density matrices, $tr^{-\vec{r}}(\rho)$, not for the whole density matrices ρ. In Ref.[21], Srivinas concluded similarly that HVT is not compatible with CPMS. A subtle difference is that we claim that IHVT violates CPMS too but not QS-V. While QS-V is a fact, CPMS is only a seem-natural theoretical requirement. So it is not impossible to break CPMS.

D. Test of Bell's HVT and Bohm's HVT on the five QS facts

We have shown that, at some huge cost, our EHVT and IHVT could respect the five QS facts. Now let us also check Bell's HVT and Bohm's HVT. Imagine we are given one of the five states below and a measurement device as will be explained. We then are asked to find out if there is one system which could not be distinguished from the first one.

QM: A quantum spin-$\frac{1}{2}$ at $\rho_0 = \left|\uparrow\right\rangle_x \left\langle\uparrow\right|_x$.

CM: A classical two-faced die at state $\rho_0 = \frac{1}{2}\left|\uparrow\right\rangle_z \left\langle\uparrow\right|_z + \frac{1}{2}\left|\downarrow\right\rangle_z \left\langle\downarrow\right|_z$.

EHVT: A classical vector pointing to arbitrary directions with probability $\rho_0 = \frac{1}{4\pi}\iint d\theta d\phi \sin\theta \frac{1+\sin\theta\cos\phi}{2} \left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}}$. If we rewrite state A in a spin coherent basis, we will get the same distribution. The only difference is that here in treating it like a classical object, we further assume that the basis is orthogonal. It is a state in the form of an exclusive-event HVT (EHVT).

IHVT: A classical object at state $\rho_0 = \Pi_{\vec{r}\in\mathbb{D}} \otimes \rho^{\vec{r}}_0$, where $\rho^{\vec{r}}_0 = \frac{1+\sin\theta\cos\phi}{2}\left|\uparrow\right\rangle_{\vec{r}} \left\langle\uparrow\right|_{\vec{r}} + \frac{1-\sin\theta\cos\phi}{2}\left|\downarrow\right\rangle_{\vec{r}} \left\langle\downarrow\right|_{\vec{r}}$. Here $(\theta,\phi) \in D = \left([0,\frac{\pi}{2}) \otimes [0,2\pi)\right) \cup \left(\{\frac{\pi}{2}\} \otimes [0,\pi)\right)$, which denotes half of all direction vector $\vec{r}$. This is a state in the form of IHVT.

BellHVT: Bell's hidden variable theory of spin-$\frac{1}{2}$ object[2] uniformly distributed with hidden variable $\lambda \in \left[-\frac{1}{2},\frac{1}{2}\right]$: given a specific λ, measurement on Pauli matrix $\vec{\beta}\cdot\vec{S}$ on direction $\vec{\beta}$ yields, $sign\left(\lambda + \frac{1}{2}\beta_x\right) sign\left(X\right)$, where $X = \beta_x$ if $\beta_x \neq 0$, $X = \beta_y$ if $\beta_y \neq 0, \beta_x = 0$ and $X = \beta_z$ if $\beta_z \neq 0, \beta_x = 0, \beta_y = 0$. Here we changed the expression accordingly to represent the x direction *up* state.

The measurement device has an indicator showing a positive/negative value if the object is along the same/opposite direction with its internal direction. One can control the internal direction of the device. Only when its direction is parallel or opposite to the object's direction, will it be activated. Assume this device is sharp so that it will not respond to even a slight mis-matching. The device works on both classical and quantum systems.

Define the activation ratio Q as the ratio between times when the device is activated out of the total times the device is used, and define the *up*-state probability P as the ratio between the number of positive values out of the times when the device is activated. We want to check if the above five states give us different values of Q and P during measurements. First, assume the device is along the z direction. We see from Table I that from the values

TABLE I. Values of Q and P with device along the z direction

	QM	CM	EHVT	IHVT	BellHVT
Q	1	1	$\ll 1$[22]	1	1
P	0.5	0.5	0.5	0.5	0.5

of Q, state $EHVT$ is different from state QM.

Next we adjust the device to the x direction. We see from Table II that from the values of Q, state CM is different from state QM. However, those measurements do not differentiate state QM, $IHVT$ and $BellHVT$. The

TABLE II. Values of Q and P with device along the x direction

	QM	CM	EHVT	IHVT	BellHVT
Q	1	0	$\ll 1$[22]	1	1
P	1	NA	1	1	1

fact that those two states $IHVT$ and $BellHVT$ both respect QS-I and QS-II, makes them very good counterexamples of Von Neumann's proof of impossibility of HVT[10]. This is exactly made possible by the fact that operators in those two theories do not obey Eq.(25) the linear relation between operators even when operators' averages have those linear relation. Such relation between operators is too restrictively assumed in Von Neumann's proof and leads to impossibility[10, 16].

In order to differentiate state QM, $IHVT$ and $BellHVT$, we have to perform subsequently repeated measurements, say first along the z direction and then along the x direction. In dealing with subsequently repeated measurements, we need some rules to determine the object's state right after the first measurement. Here we first assume both CPT-III and QM-III hold. In the following table, we list only values of Q and P after the second measurement. From Table III the values of P there we find that state $IHVT$ and $BellHVT$ are different from state QM. That is, we can distinguish a quantum state with Bell's HVT and our IHVT state by measurements if CPT-III/QM-III holds. As for state $IHVT$, this can be seen from,

TABLE III. Values of Q and P during the second measurement in a subsequently repeated measurement with device along the z and then the x direction, assuming both CPT-III and QM-III hold

	QM	CM	EHVT	IHVT	BellHVT
Q_2	1	0	0	1	1
P_2	0.5	NA	NA	1	1

$$\rho_1 = \frac{\left|\uparrow\right\rangle_z \left\langle\uparrow\right|_z \otimes \left\langle\uparrow\right|_z \rho_0 \left|\uparrow\right\rangle_z}{tr^{-z}\left(\left\langle\uparrow\right|_z \rho_0 \left|\uparrow\right\rangle_z\right)}, \tag{52}$$

and

$$P_2 = \frac{tr^{-x,-z}\left(\left\langle\uparrow\right|_z \left\langle\uparrow\right|_x \rho_0 \left|\uparrow\right\rangle_x \left|\uparrow\right\rangle_z\right)}{tr^{-z}\left(\left\langle\uparrow\right|_z \rho_0 \left|\uparrow\right\rangle_z\right)} = 1. \tag{53}$$

As for state $BellHVT$, let's assume $\lambda = \lambda^*$ after the first measurement and that the description of the state stays the same just with the specific λ^*, then for the second measurement, one will get,

$$sign\left(\lambda^* + \frac{1}{2}\beta_x\right) sign\left(\beta_x\right) = sign\left(\lambda^* + \frac{1}{2}\right) = 1, \forall \lambda^*. \tag{54}$$

If we are allowed to relax CPT-III then it is always possible to adjust ρ_1 for state $IHVT$ and adjust the proposed measurement result for state $BellHVT$ after the first measurement to make $P_2 = 0.5$. We have done so for state $IHVT$ in CPT-III$''$. And here we can adjust state $BellHVT$ to satisfy the requirement. That is if we get the $up/down$-state in the first measurement, for arbitrary second measurement of $\vec{\beta}\cdot\vec{S}$

$$sign\left(\lambda \pm \frac{1}{2}\beta_z\right) sign\left(X\right), \tag{55}$$

where

$$X = \begin{cases} \beta_z \text{ if } \beta_z \neq 0 \\ \beta_x \text{ if } \beta_z = 0, \beta_x \neq 0 \\ \beta_y \text{ if } \beta_z = 0, \beta_x = 0, \beta_y \neq 0 \end{cases}. \tag{56}$$

In that case as we see in Table IV, state $IHVT$ and $BellHVT$ are indistinguishable from state QM under all measurements, yet state $IHVT$ and $BellHVT$ are classical states and state QM is a quantum state.

TABLE IV. Values of Q and P during the second measurement in a subsequently repeated measurement with device along z and then x direction, with CPT-III adjusted accordingly

	QM	CM	EHVT	IHVT	BellHVT
Q_2	1	0	$\ll 1$[22]	1	1
P_2	0.5	NA	0.5	0.5	0.5

From the above comparison, we see that when subsequently repeated measurements are taken into consideration and CPT-III holds, none of the above four theories other than QM respects all five QS facts. Only when we relax CPT-III, both Bell's HVT and our IHVT provide alternative theories for quantum systems. Unlike Bell's HVT theory, in our IHVT state, we have explicitly written down the state in a density matrix form, so it can be generalized for any objects not only spin-$\frac{1}{2}$ particles. In this sense this work can be seen as a development of Bell's HVT. Another thing we would like to point out is the relation between our IHVT and Bohm's HVT[16]: in the following sense, our IHVT provides exactly the explicit form of a state of Bohm's HVT.

Originally Bohm's HVT gave only a classical HVT based interpretation of the measurement process on a single direction. Since we are free to choose an arbitrary direction, we need to generalize the theory a bit. Basically it says during the measurement process of a specific direction $\vec{r}$, system evolves according to the following equation system,

$$\begin{cases} \frac{dJ^1_{\vec{r}}}{dt} = 2\gamma\left(R^1 - R^2\right) J^1_{\vec{r}} J^2_{\vec{r}} \\ \frac{dJ^2_{\vec{r}}}{dt} = 2\gamma\left(R^2 - R^1\right) J^2_{\vec{r}} J^1_{\vec{r}} \end{cases}, \tag{57}$$

where $R^i = \frac{|J^i_{\vec{r}}|^2}{|\xi^i_{\vec{r}}|^2}$ and $\xi^i_{\vec{r}}$ are those hidden variables. Instead of quantum wavefunction ψ here we take $J^i_{\vec{r}}$ as our fundamental variable since only $J^1_{\vec{r}} = |\langle\uparrow_{\vec{r}} \mid \psi\rangle|^2$ and $J^2_{\vec{r}} = |\langle\downarrow_{\vec{r}} \mid \psi\rangle|^2$ are used in those equations. Then, if we focus on the state representing this direction only, it can be written down as

$$\rho_{\vec{r}} = J^1_{\vec{r}} \left|\uparrow_{\vec{r}}\right\rangle\left\langle\uparrow_{\vec{r}}\right| + J^2_{\vec{r}} \left|\downarrow_{\vec{r}}\right\rangle\left\langle\downarrow_{\vec{r}}\right|. \tag{58}$$

From this point of view, Eq.(57) provides an explanation of the process that the above state in Eq.(58) turns into $|\uparrow_{\vec{r}}\rangle\langle\uparrow_{\vec{r}}|$ or $|\downarrow_{\vec{r}}\rangle\langle\downarrow_{\vec{r}}|$ at probability respectively J_1 or J_2. Now let us consider a separate measurement along another direction $\vec{r}'$. One possible approach is to start from the quantum wavefunction ψ again to calculate $J^i_{\vec{r}'}$ and redo the above procedure. However, in this way this theory never totally eliminates the quantum wavefunction. There is however another way to recover the right prediction of measurement on $\vec{r}'$ and it eliminates the quantum wavefunction totally. That is to assume that the HVT state is in fact,

$$\rho = \Pi_{\vec{r}} \otimes \rho_{\vec{r}}, \tag{59}$$

while $\rho_{\vec{r}}$ is given by Eq.(58) for a specific direction $\vec{r}$ with proper predefined $J^i_{\vec{r}}$. We see that this is exactly state $IHVT$, our IHVT state. We have shown that we can make this state agree with the quantum mechanical state QM on everything.

However, There is a price to do so. This IHVT is far from a standard CPT. To summarize, IHVT satisfies QS-I and QS-II easily but CPT-III needs to be modified to make it satisfy QS-III and QS-IV. IHVT does not require contextuality between object and observer as EHVT does. But both suffer from the same inconsistency problem: two different expressions are used to represent the same state for two different purposes. Furthermore, both discard inherent relation among operators as in Eq.(25) by treating operators independently or exclusively. Also CPMS need to be twisted too: a state resulted from two exclusive procedures are not a probability summation of the two states respectively corresponding to the two individual procedures. We find that all of the above has made HVT less understandable than the usual QM, which has none of above problems and respects all five QS facts. We would like to conclude that we have ruled out HVT just from theoretical consideration of a spin-$\frac{1}{2}$ object.

If one is still willing to pay all the prices mentioned above, then we are also willing to go a little further to show that this IHVT conceals something else which one may not want in a theory of physics. We will apply this theory onto the description of the entangled singlet state, which was also used in the discussion of Bell's inequality.

VI. CPT FOR TWO-SPIN SYSTEM

The quantum density matrix form of a singlet state is

$$\rho^q = \frac{1}{2}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right)\left(\langle\uparrow\downarrow| - \langle\downarrow\uparrow|\right), \tag{60}$$

where $|\uparrow\downarrow\rangle$ can be regarded as eigenstate on an arbitrary direction. For this specific example, QS-I and QS-II reads, with vector $\vec{r}_1$ and $\vec{r}_2$ on spin 1 and 2 respectively, on two-spin measurements,

$$\langle S_{\vec{r}_1} S_{\vec{r}_1}\rangle = -\frac{1}{4}\vec{r}_1 \cdot \vec{r}_2 \tag{61}$$

and in each measurement, $S_{\vec{r}_1} S_{\vec{r}_2} = \pm\frac{1}{4}$; and on single-spin measurements,

$$\langle S_{\vec{r}}\rangle = 0 \tag{62}$$

and in each measurement $S_{\vec{r}} = \pm\frac{1}{2}$. A successful HVT theory should respect the above two results. For a subsequently repeated measurement, HVT should also give the correct results depending on the outcome from the first measurement. Although Bell's theorem[2] has generally proven that through local classical theory it is impossible to achieve this, here, we will construct one such state, in the form of a classical density matrix, that does, in fact, achieve this. We will then find out the cost of such a theory.

To denote a state in IHVT, one example of an equivalent class is used to represent the whole class. We should check if the following state respects all the QS facts. A reduced density matrix for two spins on $\vec{r}_1$ and $\vec{r}_2$ is

$$\begin{aligned}\rho_{\vec{r}_1,\vec{r}_2} = & \tfrac{1-\vec{r}_1\cdot\vec{r}_2}{4}\left(|\uparrow_{\vec{r}_1}\uparrow_{\vec{r}_2}\rangle\langle\uparrow_{\vec{r}_1}\uparrow_{\vec{r}_2}| + |\downarrow_{\vec{r}_1}\downarrow_{\vec{r}_2}\rangle\langle\downarrow_{\vec{r}_1}\downarrow_{\vec{r}_2}|\right) \\ & + \tfrac{1+\vec{r}_1\cdot\vec{r}_2}{4}\left(|\uparrow_{\vec{r}_1}\downarrow_{\vec{r}_2}\rangle\langle\uparrow_{\vec{r}_1}\downarrow_{\vec{r}_2}| + |\downarrow_{\vec{r}_1}\uparrow_{\vec{r}_2}\rangle\langle\downarrow_{\vec{r}_1}\uparrow_{\vec{r}_2}|\right)\end{aligned}. \tag{63}$$

The whole density matrix is

$$\rho^{hvt} = \prod_{\vec{r}_1\vec{r}_2} \otimes \rho_{\vec{r}_1\vec{r}_2}, \tag{64}$$

The reduced density matrix satisfies equ(61) and equ(62), hence QS-I and QS-II. For QS-III and QS-IV, although we will skip the details here, an after-measurement state can be constructed easily to satisfy the two. Twisting CPMS again solves the problem of QS-V. We have successfully constructed a classical theory for two-spin quantum system. It is a classical theory but it violates Bell's inequality. As we argued above, we already know that, due to inconvenience and inconsistency, this theory should not be preferred. However, we can still ask how can such a classical theory succeed to give all expected results from QM? The answer is, it includes non-local (or contextual) information.

In [23], Bell's inequality was proven more generally with only the locality assumption, their equ(2′) uses

$$p_{1,2}(\lambda, a, b) = p_1(\lambda, a)\, p_2(\lambda, b), \tag{65}$$

where λ **is a hidden variable independent of** a, b to express the idea of measurement-independent reality of a quantum system. Since our IHVT violates Bell's inequality, we want to check if it respects the above equation. Consider the situation where we measure direction a and b on those two spins respectively.

$$\begin{aligned}\left\langle \hat{S}^1 \hat{S}^2 \right\rangle (a,b) &= tr\left(\hat{S}^1(a)\,\hat{S}^2(b)\,\rho\left(\vec{\lambda}\right)\right) \\ &= \sum_{\lambda_{ab}} \langle\lambda_{ab}|\,\hat{S}^1(a)\,\hat{S}^2(b)\,|\lambda_{ab}\rangle\, f(\lambda_{ab}) \\ &= \sum_{\lambda_{ab}} s^1(a,\lambda_{ab})\, s^2(b,\lambda_{ab})\, f(\lambda_{ab})\end{aligned} \tag{66}$$

The left hand side can be regarded as

$$\left\langle \hat{S}^1 \hat{S}^2 \right\rangle (a,b) = \sum_{\lambda_{ab}} s^1 s^2 (a,b,\lambda_{ab})\, f(\lambda_{ab}). \tag{67}$$

From the core of the integral, we see that

$$s^1 s^2 (a,b,\lambda_{ab}) = s^1(a,\lambda_{ab})\, s^2(b,\lambda_{ab}), \tag{68}$$

or generally,

$$s^1 s^2 \left(a,b,\vec{\lambda}\right) = s^1\left(a,\vec{\lambda}\right) s^2\left(b,\vec{\lambda}\right). \tag{69}$$

Compared with equ(65), equ(69) does look like an expression of locality, with the difference that a single hidden variable is replaced by many hidden variables. However, it is this replacement that introduces non-local information, because the particular effective one of $\vec{\lambda}$ is λ_{ab}, which does depend on both a and b, directions of measurements on both spins. During the measurement process, a sample should be drawn from an effective probability distribution. And the effective one has to be determined through information from both directions a and b together. It is as if the system has to know both directions to make its decision. It is definitely contextual.

We have explicitly shown the place where non-locality comes into QM. When the classical theory is used to describe QM, we have to require non-local information. If this non-locality is unacceptable, then we should rule out this particular HVT. However, this never means QM in its own language requires non-local information. This is a topic which has never been addressed in this paper.

VII. CONCLUSION AND DISCUSSION

We test theories against nothing else but the five QS facts, which are all established experimental facts not any kind of presumably reasonable conditions. In a summary, to find a classical theory respecting all five QS facts, our conclusion is: first, single variable HVT is incompatible with non-commutative relation between operators; second, even if all operators are commutative, the inherent relation between them has to be abandoned; third, the EHVT requires contextuality between object and observer; fourth, both EHVT and IHVT suffer from the inconsistency problem: the expression used to denote the state is different with the one used to recover probability; and last, IHVT is shown to imply contextuality between the two distant spins of a singlet. We find the price is unreasonably high: even after we accept the contextuality, both CPT-III and CPMS need to be twisted.

Due to this twist, such a classical system could no longer be broadcast, noticing CPT-III is essential to make it possible to broadcast a classical system. The possibility of being broadcast is one key fact in understanding classical measurement. This makes is impossible to fulfill the initial motivation of HVT, which is to understand better the quantum measurement using the picture of classical measurement. Therefore, even theoretically, not depending on the experimental test of Bell's inequality, the idea of HVT should be discarded from theories of quantum systems. Besides our own HVTs we have also examined Bell's HVT and Bohm's HVT and ruled them out based on subsequently repeated measurements and validity of CPT-III.

In other words, under reasonable consideration it is impossible to map a full-structure density matrix to a diagonal density matrix. With this conclusion in mind, we may say that although the current language of QM may not be the ultimate one, any equivalent language should include the existence of off-diagonal elements of the density matrix and allow vectors to be transformed from one basis to another, which is only possible when operators do not always commute with each other. We then know quantum measurement is not equal to classical measurement of a TRCO. Classical measurement of TRCO creates a broadcast, but quantum measurement does not.

Finally, we are not saying those are all the possibilities of CPT for QS. From the commutative C^*-algebra point of view, what we have tried here are just two examples of commutative multiplications between operators, $S_{\vec{r}_1} S_{\vec{r}_2} = 0$ for the exclusive case and direct product $S_{\vec{r}_1} \otimes S_{\vec{r}_2}$ for the independent case. There may be other kinds of algebras among operators. If we assume symmetry among S operators on all directions, then those two are the only choices.

VIII. ACKNOWLEDGMENT

Thanks should first be given to Dr. Leslie Ballentine and Dr. Ian Affleck for their inspiring courses on quantum mechanics, giving a detailed and insightful check of the basis of quantum mechanics, and a fascinating introduction into the fundamental problems in quantum mechanics. Thanks should also be given to Dr. Mona Berciu for her comments and discussions on this work, and Dr. Robert Peter and Janelle von Dongen for reading through the manuscript. At last, we acknowledge the anonymous referee gratefully for the comments and suggestions during the revision.

[1] J.S. Bell, Physics **1**, 195(1964).

[2] J.S. Bell, Rev. Mod. Phys. **38**, 447(1966).

[3] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland 1982).

[4] A. Khrennikov, *Contextual Approach to Quantum Formalism*(Springer, 2009).

[5] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).

[6] A. Lupascu, S. Saito, T. Picot, P. C. De Groot,C. J. P. M. Harmans, J. E. Mooij, **3**, 119(2007).

[7] D. Bohm, Phys. Rev. **85**, 166(1952).

[8] A. Gut, *Probability: A Graduate Course* (Springer-Verlag, 2005).

[9] R. Spekkens, Phys. Rev. A **75**, 032110(2007).

[10] J. von, Neumann *Mathematical Foundations of Quantum Mechanics* (Princeton, Princeton Univ. Press, 1955).

[11] L, Ballentine, *Quantum Mechanics: A Modern Development* (World Scientific, 1998).

[12] D.Y. Petrina, *Mathematical foundations of quantum statistical mechanics*, (Kluwer Acad. Publ. 1995); D. Y. Petrina, V. I. Gerasimenko and P. V. Malyshev, *Mathematical foundations of classical statistical mechanics* (Taylor & Francis, 2002).

[13] N.P. Landsman, *Lecture notes on C*-algebras, Hilbert C*-modules, and quantum mechanics*, arXiv:math-ph/9807030v1, 1998.

[14] A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. **47**, 777(1935).

[15] H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa and B. Schumacher, Phys. Rev. Lett, **76**2818(1996).

[16] D. Bohm and J. Bub, Rev. Mod. Phys. **38**, 453(1966).

[17] M.O. Scully, Phys. Rev. D, **28**, 2477(1983).

[18] A.S. Holevo, *Statistical Structure of Quantum Theory* (Springer, 2001).

[19] A. Aspect, Nature **398**, 189(1999).

[20] C.W. Gardiner and P. Zoller, *Quantum Noise*(Springer, 2000).

[21] M.D. Srinivas, Pramana **19**, 159-173(1982).

[22] If we allow contextual communication between object and observer then Q could be equal to 1. Or Q could also be equal to 1 if our device is totally unsharp.

[23] J.F. Clauser and M.A. Horne, Phys. Rev. D**10**, 526(1974).