Revisiting the admissibility of non-contextual hidden variable models in quantum mechanics

Atul Singh Arora

Department of Physical Sciences, Indian Institute of Science Education & Research (IISER) Mohali, Sector 81 SAS Nagar, Manauli PO 140306 Punjab India.

Centre for Quantum Information and Communication, Ecole polytechnique de Bruxelles, CP 165, Université libre de Bruxelles (ULB), 1050 Brussels, Belgium

Kishor Bharti

Centre for Quantum Technologies (CQT), National University of Singapore (NUS) - Block S15, 3 Science Drive 2, Singapore 117543.

Arvind

Department of Physical Sciences, Indian Institute of Science Education & Research (IISER) Mohali, Sector 81 SAS Nagar, Manauli PO 140306 Punjab India.

Abstract

We construct a non-contextual hidden variable model consistent with all the kinematic predictions of quantum mechanics (QM). The famous Bell-KS theorem shows that non-contextual models which satisfy a further reasonable restriction are inconsistent with QM. In our construction, we define a weaker variant of this restriction which captures its essence while still allowing a non-contextual description of QM. This is in contrast to the contextual hidden variable toy models, such as the one by Bell, and brings out an interesting alternate way of looking at QM. The results also relate to the Bohmian model, where it is harder to pin down such features.

1. Introduction

Quantum mechanics (QM) has been one of the most successful theories in physics so far, however, there has not yet been a final word on its completeness and interpretation [?]. Einstein's [?] work on the incompleteness of QM and the subsequent seminal work of Bell [?], assessing the compatibility of a more complete model involving hidden variables (HV) and locality with QM, has provided deep insights into how the quantum world differs from its classical counterpart. In recent times, these insights have been of pragmatic utility in the area of quantum information processing (QIP), where EPR pairs are fundamental motifs of entanglement [? ? ?]. The work of Kochen Specker (KS) [?] and Gleason *et. al.* [? ? ?] broadened the schism between HV models and QM. They showed that it was contextuality and not non-locality which was at the heart of this schism and the incompatibility between HV models and QM can arise even for a single indivisible quantum system. Contextuality has thus been identified as a fundamental non-classical feature of the quantum world and experiments have also been proposed and conducted to this effect [? ?]. Contextuality, on the one hand has led to investigations on the foundational aspects of QM including attempts to prove the completeness of QM [? ?], and on the other hand has been harnessed for computation and cryptography [? ?]. While there have been generalisations, in this letter, we restrict ourselves to the standard notion of non-contextuality as used by KS [?].

Not all HV models (e.g. Bohm's model based on trajectories [? ?]), however, are incompatible with QM [?] and we, in this letter, present a new non-contextual HV model consistent with QM.

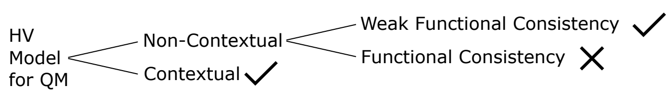

Figure 1: Non-contextuality is not inconsistent with kinematic predictions of QM

The KS theorem is applicable to non-contextual HV models which additionally satisfy *functional-consistency — algebraic constraints obeyed by commuting quantum observables must be satisfied by the HV model at the level of individual outcomes.* One of the main justifications for imposing this was the requirement that the observable $\hat{A}^2$ must depend on $\hat{A}$ even at the HV level. This sounds reasonable because otherwise one can construct HV models where these observables can get mapped to random variables with no relation to each other (see [?]). As we will show, one can weaken this requirement into what we call *weak functional-consistency* – *functional-consistency is demanded only when the observables in question have sharp values.* This entails that even in the HV model the observables will depend on each other where they should for consistency and not otherwise, thereby capturing the essential idea without undue constraints. The consequence of demanding only the weaker variant is that the KS theorem no longer applies and non-contextual HV models can be consistently constructed. This is in stark contrast with contextual HV models, such as the toy model[1] by Bell [?] because here we consider models where the algebraic constaints of QM are not obeyed in general by the HV model at the level of individual outcomes. We demonstrate this contrast by re-examining a 'proof of contextuality' using our model. This provides a new way of looking at the classical-quantum divide and at the foundations of quantum mechanics.

2. Non-Contextuality and Functional Consistency

We introduce some notation and make the relevant notions precise to facilitate the construction of our model.

Notation. (a) $\psi \in \mathcal{H}$ represents a pure quantum mechanical state of the system in the Hilbert space $\mathcal{H}$, (b) $\hat{\mathcal{H}}$ is defined to mean the set of Hermitian observables for the system, (c) $[\mathcal{H}]$ is defined to mean $\{\mathcal{H}, \mathbb{R}^{\otimes}\}$, which represents the state space of the system including HVs, (d) $[\psi] \in [\mathcal{H}]$ will represent the state of the system including HVs, (e) a prediction map is $M : \{\hat{\mathcal{H}}, [\mathcal{H}]\} \to \mathbb{R}$, (f) a sequence map is $S : \{\hat{\mathcal{H}}, [\mathcal{H}], \mathbb{R}\} \to [\mathcal{H}]$, (g) f is an arbitrary map from $\{\hat{\mathcal{H}}, \hat{\mathcal{H}}, \dots \hat{\mathcal{H}}\} \to \hat{\mathcal{H}}$ constructed using multiplication and addition of compatible observables, and multiplication with complex numbers, (h) $\tilde{f}$ is a map constructed by replacing observables in f with real numbers.

Definition 1. A theory is non-contextual, if it provides a map $M : \{\hat{\mathcal{H}}, [\mathcal{H}]\} \to \mathbb{R}$ to explain measurement outcomes. A theory which is not non-contextual is contextual [?].

Remark 1.1. A prediction map of the form $M : \{\hat{\mathcal{H}}, [\mathcal{H}]\} \to \mathbb{R}$ itself can be called non-contextual.

Remark 1.2. Broader definitions in the literature have been suggested [?] which extend the notion to probabilistic HV models.

The idea is that any feature of a HV model that is not determined solely by the operational aspect of QM is defined to be a demonstration of contextuality. If, for instance, this distinction arises in the measurement procedure then it is termed as measurement contextuality. To maintain a distinction between different features of HV models, which is of interest here, we stick to the standard definition.

[1]introduced in connection with Gleason's Theorem

In addition to a HV model being non-contextual KS [?] demand *functional-consistency* to establish their no-go theorem which is defined below in our notation.

Definition 2. A prediction map M is *functionally-consistent* iff

$$M(f(\hat{B}_1, \hat{B}_2, \dots \hat{B}_N), [\psi]) = \tilde{f}(M(\hat{B}_1, [\psi]), M(\hat{B}_2, [\psi]), \dots M(\hat{B}_N, [\psi])),$$

where $\hat{B}_i \in \hat{\mathcal{H}}$ are arbitrary mutually commuting observables and $[\psi] \in [\mathcal{H}]$. A *non functionally-consistent* map is one that is not functionally-consistent.

Note that if M is taken to represent the measurement outcome (in QM), then for states of the system which are simultaneous eigenkets of $\hat{B}_i$s, M must clearly be *functionally-consistent.* It is, however, not obvious that this property must always hold. For example, consider two spin-half particles in the state $|1\rangle \otimes |1\rangle$ written in the computational basis and operators $\hat{B}_1 = \hat{\sigma}_x \otimes \hat{\sigma}_x$, $\hat{B}_2 = \hat{\sigma}_y \otimes \hat{\sigma}_y$ and $\hat{C} = \hat{B}_1\hat{B}_2 = -\hat{\sigma}_z \otimes \hat{\sigma}_z$ written in terms of Pauli operators. We must have $M(\hat{C}) = -1$ while $M(\hat{B}_1) = \pm 1$ and $M(\hat{B}_2) = \pm 1$ independently, according to QM, with probability half. Here *functional-consistency* clearly is not required to hold and indeed this is why it was demanded in addition to being consistent with QM by KS. Antithetically, it is clear that if one first measures $\hat{B}_1$ and subsequently measures $\hat{B}_2$, then the product of the results must be -1. This is consistent with measuring $\hat{C}$. In our treatment, instead of imposing *functional-consistency*, we demand its aforesaid weaker variant. It captures the same idea, however, only when it has a precise meaning according to QM. To this end we define *weak functional-consistency* as follows.

Definition 3. A prediction map m has *weak functional-consistency* for a given sequence map s, iff

$$M(f(\hat{B}_1, \hat{B}_2, \dots \hat{B}_N), [\psi_1]) = \tilde{f}(M(\hat{B}_1, [\psi_{k_1}]), M(\hat{B}_2, [\psi_{k_2}]), \dots, M(\hat{B}_N, [\psi_{k_N}])),$$

where $\{k_1, k_2, \dots, k_N\} \in \{N! \text{ permutations of } k'\text{s}\}$, $\hat{B}_i \in \hat{\mathcal{H}}$ are arbitrary mutually commuting observables, $[\psi_i] \in [\mathcal{H}]$ and $[\psi_{k+1}] := S(\hat{B}_k, [\psi_k], M(\hat{B}_k, [\psi_k]))$, $\forall [\psi_i]$.

With these definitions we are now ready to discuss the 'proof of contextuality'. We first state the contextuality theorem in our notation:

Theorem. *Let a map $M : \hat{\mathcal{H}} \to \mathbb{R}$, be s.t. (a) $M(\hat{\mathbb{I}}) = 1$, (b) $M(f(\hat{B}_1, \hat{B}_2, \dots)) = \tilde{f}(M(\hat{B}_1), M(\hat{B}_2), \dots)$, for any arbitrary function f, where $\hat{B}_i$ are mutually commuting Hermitian operators. If m is assumed to describe the outcomes of measurements, then no M exists which is consistent with all predictions of QM.*

Proof. Peres Mermin [PM] ($|\mathcal{H}| \geq 4$) [? ?]: For a system composed of two spin-half particles consider the following set of operators

$$\hat{A}_{ij} \doteq \begin{bmatrix} \hat{\mathbb{I}} \otimes \hat{\sigma}_x & \hat{\sigma}_x \otimes \hat{\mathbb{I}} & \hat{\sigma}_x \otimes \hat{\sigma}_x \\ \hat{\sigma}_y \otimes \hat{\mathbb{I}} & \hat{\mathbb{I}} \otimes \hat{\sigma}_y & \hat{\sigma}_y \otimes \hat{\sigma}_y \\ \hat{\sigma}_y \otimes \hat{\sigma}_x & \hat{\sigma}_x \otimes \hat{\sigma}_y & \hat{\sigma}_z \otimes \hat{\sigma}_z \end{bmatrix}$$

which have the property that all operators along a row (or column) commute. Further, the product of rows (or columns) yields $\hat{R}_i = \hat{\mathbb{I}}$ and $\hat{C}_j = \hat{\mathbb{I}}\,(j \neq 3)$, $\hat{C}_3 = -\hat{\mathbb{I}}$, $(\forall\, i, j)$ where $\hat{R}_i := \prod_j \hat{A}_{ij}$, $\hat{C}_j := \prod_i \hat{A}_{ij}$. Let us assume that M exists. From property (b) of the map, to get $M(\hat{C}_3) = -1$ (as required by property (a)), we must have an odd number of -1 assignments in the third column. In the remaining columns, the number of -1 assignments must be even (for each column). Thus, in the entire square, the number of -1 assignments must be odd. Let us use the same reasoning, but along the rows. Since each $M(\hat{R}_i) = 1$, we must have even number of -1 assignments along each row. Thus, in the entire square, the number of -1 assignments must be even. We have arrived at a contradiction and therefore we conclude that M does not exist ☐

Remark. One could in principle assume M, to be s.t. (a) $M(\hat{\mathbb{I}}) = 1$, (b) $M(\alpha\hat{B}_i) = \alpha M(\hat{B}_i)$, for $\alpha \in \mathbb{R}$, (c) $M(\hat{B}_i^2) = M(\hat{B}_i)^2$, (d) $M(\hat{B}_i + \hat{B}_j) = M(\hat{B}_i) + M(\hat{B}_j)$, to deduce (d) $M(\hat{B}_i\hat{B}_j) = M(\hat{B}_i)M(\hat{B}_j)$ and that $M(\hat{B}_i) \in$ spectrum of $\hat{B}_i$. Effectively then, condition (b) listed in the theorem is satisfied as a consequence. Therefore, assuming (a)-(d) as listed above, rules out a larger class of M [**?**].

Here M maybe viewed as a specific class of prediction maps that implicitly depends on the state $[\psi]$. It is clear that according to the theorem, non-contextual maps which are *functionally-consistent* must be incompatible with QM. This leaves open an interesting possibility that non-contextual maps which have *weak functional-consistency* could be consistent with QM. Before proceeding to do so explicitly we observe that *weak functional-consistency* is, in fact, a consequence of QM.

Proposition. *Let a quantum mechanical system be in a state, s.t. measurement of $\hat{C}$ yields repeatable results (same result each time). Then according to QM,* weak functional consistency *holds, where $\hat{C} := f(\hat{B}_1, \hat{B}_2, \dots \hat{B}_n)$, and $\hat{B}_i$ are as defined (in Definition* **??***)*

Proof. Without loss of generality we can take $\hat{B}_1, \hat{B}_2, \dots \hat{B}_n$ to be mutually compatible and a complete set of observables (operators can be added to make the set complete if needed). It follows that $\exists\, |\mathbf{b} = (b_1, b_2, \dots b_n)\rangle$ s.t. $\hat{B}_i |\mathbf{b}\rangle = b_i |\mathbf{b}\rangle$, and that $\sum_{\mathbf{b}} |\mathbf{b}\rangle \langle \mathbf{b}| = \hat{\mathbb{I}}$. Let the state of the system $|\psi\rangle$ be s.t. $\hat{C}|\psi\rangle = c|\psi\rangle$. For the statement to follow, one need only show that $|\psi\rangle$ must be made of only those $|\mathbf{b}\rangle$s, which satisfy $c = \tilde{f}(b_1, b_2, \dots b_n)$. This is the crucial step and proving this is straightforward. We start with $\hat{C}|\psi\rangle = c|\psi\rangle$ and take its inner product with $\langle \mathbf{b}|$ to get

$$\begin{aligned} c\langle \mathbf{b}|\psi\rangle &= \langle \mathbf{b}|\hat{C}|\psi\rangle \\ &= \langle \mathbf{b}|f(\hat{B}_1, \hat{B}_2, \dots \hat{B}_n)|\psi\rangle \\ &= \tilde{f}(b_1, b_2, \dots b_n)\langle \mathbf{b}|\psi\rangle \end{aligned}$$

Also, we have $|\psi\rangle = \sum_{\mathbf{b}} \langle \mathbf{b}|\psi\rangle |\mathbf{b}\rangle$, from completeness. If we consider $|\mathbf{b}\rangle$s for which $\langle \mathbf{b}|\psi\rangle \neq 0$, then we find that indeed $c = \tilde{f}(b_1, b_2, \dots b_n)$. The case when $\langle \mathbf{b}|\psi\rangle = 0$ is anyway irrelevant as the corresponding $|b\rangle$ does not contribute to $|\psi\rangle$. We can thus conclude that $|\psi\rangle$ is made only of those $|\mathbf{b}\rangle$s that satisfy the required relation. ☐

It is worth noting that in the Peres Mermin case, where $\hat{R}_i$ and $\hat{C}_j$ are $\pm\hat{\mathbb{I}}$, it follows that all states are their eigenstates. Consequently, for these operators *weak functional-consistency* must always hold.

3. Construction

We are now ready to describe our model explicitly. Let the state of a finite dimensional quantum system be $|\chi\rangle$. We wish to assign a value to an arbitrary observable $\hat{A} = \sum_a a\,|a\rangle\langle a|$, which has eigenvectors $\{|a_j\rangle\}$. The corresponding ordered eigenvalues are $\{a_j\}$ such that $a_{\min} = a_1$ and $a_{\max} = a_n$. Our HV model for QM assigns values in the following three steps:

1. **Initial HV:** Pick a number $c \in [0, 1]$, from a uniform random distribution.

2. **Assignment or Prediction:** The value assigned to $\hat{A}$ is given by finding the smallest a s.t. $c \leq \sum_{a'=a_{\min}}^{a} |\langle a'|\chi\rangle|^2$, viz. we have specified a prediction map, $M(\hat{A}, [\chi]) = a$.

3. **Update:** After measuring an operator, the state is updated (collapsed) in accordance with the rules of QM. This completely specifies the sequence map S.

The above HV model works for all quantum systems, however, to illustrate its working, consider the example of a spin-half particle in the state $|\chi\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle$ and the observable $\hat{A} = \hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1|$. Now, according to the postulates of this theory, $M(\hat{A}, [\chi]) = +1$ if the randomly generated value for $c \leq \cos^2\theta$ else $\hat{A}$ is assigned -1; it then follows, from c being uniformly random in $[0, 1]$ that the statistics agree with the predictions of QM *i.e.* the Born rule. The assignment described by the prediction map M is non-contextual since, given an operator and a state (alongwith the HV c), the value is uniquely assigned. The map M is, however, *non functionally-consistent*.

To see this *non functional-consistency* explicitly in our model and to see its applicability to composite systems, we apply the model to the Peres Mermin situation of two spin-half particles. Consider the initial state of the system $|\psi_1\rangle = |00\rangle$. Assume we obtained $c = 0.4$ as a random choice. To arrive at the assignments, note that $|00\rangle$ is an eigenket of only $\hat{R}_i, \hat{C}_j$ and $\hat{A}_{33} = \hat{\sigma}_z \otimes \hat{\sigma}_z$. Thus, in the first iteration, all these should be assigned their respective eigenvalues. The remaining operators must be assigned -1 as one can readily verify by explicitly finding the smallest a as described in postulate 2 of the model (see Table **??**).

For the next iteration, $i = 2$, after the first measurement is over say the random number generator yielded the value $c = 0.1$. Since $|\psi\rangle$ is also unchanged the assignment remains invariant (in fact any of $c < 0.5$ would yield the same result as evident from the previous exercise). For the final step we choose to measure $\hat{A}_{23}(= \hat{\sigma}_y \otimes \hat{\sigma}_y)$, to proceed with sequentially measuring $\hat{C}_3$. To simplify calculations, we note

$$|00\rangle = \frac{1}{\sqrt{2}} \left[\left(\left|\tilde{+}\tilde{-}\right\rangle + \left|\tilde{-}\tilde{+}\right\rangle \right)/\sqrt{2} + \left(\left|\tilde{+}\tilde{+}\right\rangle + \left|\tilde{-}\tilde{-}\right\rangle \right)/\sqrt{2} \right],$$

with $\left|\tilde{\pm}\right\rangle = (|0\rangle \pm i\,|1\rangle)/\sqrt{2}$ (eigenkets of $\hat{\sigma}_y$). Since $|00\rangle$ is manifestly not an eigenket of $\hat{A}_{23}$, we must find an appropriate eigenket $\left|a_{23}^{-}\right\rangle$ s.t. $\hat{A}_{23}\left|a_{23}^{-}\right\rangle = -\left|a_{23}^{-}\right\rangle$, since $c = 0.1$ and $\left\langle a_{23}^{-}|00\right\rangle$ is already > 0.1. It is evident that $\left|a_{23}^{-}\right\rangle = \left(\left|\tilde{+}\tilde{-}\right\rangle + \left|\tilde{-}\tilde{+}\right\rangle \right)/\sqrt{2} = (|00\rangle + |11\rangle)/\sqrt{2}$, which becomes the final state.

For the final iteration, $i = 3$, say we obtain $c = 0.7$. So far, we have $M_1(\hat{A}_{33}) = 1$ and $M_2(\hat{A}_{23}) = -1$. We must obtain $M_3(\hat{A}_{13}) = 1$, independent of the value of c, to be consistent. Let us check that. Indeed, according to postulate 2, since $\hat{\sigma}_x \otimes \hat{\sigma}_x (|00\rangle + |11\rangle)/\sqrt{2} = 1\,(|00\rangle + |11\rangle)/\sqrt{2}$, $M_3(\hat{A}_{13}) = 1$ for all allowed values of c. It is to be noted that $M_2(\hat{A}_{33}) = M_3(\hat{A}_{33})$ and $M_2(\hat{A}_{23}) = M_3(\hat{A}_{23})$, which essentially expresses the compatibility of these observables; i.e. once measured, the values of observables compatible with $\hat{A}_{13}$ are not affected by the measurement of $\hat{A}_{13}$.

The *non functional-consistency* is manifest, for $M_1(\hat{C}_3) = 1 \neq M_1(\hat{A}_{13})M_1(\hat{A}_{23})M_1(\hat{A}_{33}) = -1$, where the subscript refers to the iteration number. More precisely, $M_1(\hat{O}) := M(\hat{O}, [|\psi_1\rangle = |00\rangle])$ where the complete state $[|\psi_1\rangle]$ implicitly refers to both the quantum state $|00\rangle$ and the HV $c = 0.4$. The model, however, obeys the *weak functional-consistency* requirement, namely,

$$\begin{aligned} & M_1(\hat{C}_3) = M_1(\hat{A}_{33})M_2(\hat{A}_{23})M_3(\hat{A}_{13}), \\ \text{where} \quad & M_2 := M(\hat{O}, [|\psi_2\rangle]), \\ & M_3 := M(\hat{O}, [|\psi_3\rangle]) \end{aligned}$$

and $|\psi_2\rangle, |\psi_3\rangle$ are obtained from postulate 3. Note that for each iteration, a new HV is generated.

4. Implications and Remarks

The model demonstrates *non function-consistency* as an alternative signature of quantumness as opposed to contextuality. This view is not just an artifact of the simplicity of the model. It has implications to Bohm's HV model (BHVM), where if the initial conditions are precisely known the entire trajectory of a particle (guided by the wavefunction) can be predicted including the individual outcome of measurements. BHVM applied to a single spin-half particle in a Stern-Gerlach experiment, can predict opposite results for two measurements of the same operator. This observation which is often used as a demonstration of "contextuality" in BM, does not involve overlapping sets of compatible measurements to provide two different contexts [**?**] and is therefore not of interest here. It turns out that, if one constructs a one-one map between an experiment and an observable (by following a certain convention), this so-called "contextuality" can be removed from BHVM. However, the prediction map so obtained from observables to measurement outcomes turns out to be *non functionally-consistent*, suggesting that *non functional-consistency* is a more suitable explanation of non-classicality of BHVM. In fact, our model when appropriately extended to continuous variables yields Bohmian trajectories in the single particle case which suggests that it can be used as a starting point for constructing more interesting families of HV theories that take quantum time dynamics into account as well.

Bell himself had constructed a deceptively similar toy model[2] to demonstrate a HV construction that is not ruled out by his contextuality no-go theorem. However, his model was contextual (and *functionally-consistent*) which is in contrast with our construction which is non-contextual (and has *weak functional-consistency*).

We end with two short remarks. First, we illustrate how *non functional-consistency* gives rise to situations which could get confused with the presence of contextuality. Imagine for two spin 1/2 particles

$$\begin{aligned} & \hat{B}_1 = \hat{\sigma}_z \otimes \hat{\mathbb{I}} = \\ & |00\rangle\langle 00| + |01\rangle\langle 01| - [|10\rangle\langle 10| + |11\rangle\langle 11|], \\ & \hat{B}_2 = \hat{\mathbb{I}} \otimes \hat{\sigma}_z = \\ & |10\rangle\langle 10| + |11\rangle\langle 11| - [|00\rangle\langle 00| + |01\rangle\langle 01|], \\ & \hat{C} = f(\{\hat{B}_i\}) \\ & \quad = 1.\,|00\rangle\langle 00| + 2.\,|01\rangle\langle 01| + \\ & \quad 3.\,|10\rangle\langle 10| + 4.\,|11\rangle\langle 11|. \end{aligned}$$

$\hat{C}$ maybe viewed as a function of $\hat{B}_1$, $\hat{B}_2$ and other operators $\hat{B}_i$ which are constructed to obtain a maximally commuting set. A measurement of $\hat{C}$, will collapse the state into one of the states which are simultaneous eigenkets of $\hat{B}_1$ and $\hat{B}_2$. Consequently, from the observed value of $\hat{C}$, one can deduce the values of $\hat{B}_1$ and $\hat{B}_2$. Now consider $\sqrt{2}\,|\chi\rangle = |10\rangle + |01\rangle$, for which $M_1(\hat{B}_1) = 1$, and $M_1(\hat{B}_2) = 1$, using our model, with $c < 0.5$. However, $M_1(\hat{C}) = 1$, from which one can deduce that $\hat{B}_1$ was $+1$, while $\hat{B}_2$ was -1. This property itself, one may be tempted call contextuality, viz. the value of $\hat{B}_1$ depends on whether it is measured alone or with the remaining $\{\hat{B}_i\}$. However, it must be noted that $\hat{B}_1$ has

[2] as referred to earlier

$i=1:\ c=0.4,\ \lvert\psi_{\text{init}}\rangle = \lvert 00\rangle$	$i=2:\ c=0.1,\ \lvert\psi_{\text{init}}\rangle = \lvert 00\rangle$	$i=3:\ c=0.7,\ \lvert\psi_{\text{init}}\rangle = \frac{1}{\sqrt{2}}(\lvert 00\rangle + \lvert 11\rangle)$
$M_1(\hat{A}_{ij}) \doteq \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & +1 \end{bmatrix}$	$M_2(\hat{A}_{ij}) \doteq \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & +1 \end{bmatrix}$	$M_3(\hat{A}_{ij}) \doteq \begin{bmatrix} +1 & +1 & +1 \\ +1 & +1 & -1 \\ +1 & +1 & +1 \end{bmatrix}$
$M_1(\hat{R}_i), M_1(\hat{C}_j) = +1\,(j \neq 3)$	$M_2(\hat{R}_i), M_2(\hat{C}_j) = +1\,(j \neq 3)$	$M_3(\hat{R}_i), M_3(\hat{C}_j) = +1\,(j \neq 3)$
$M_1(\hat{C}_3) = -1$	$M_2(\hat{C}_3) = -1$	$M_3(\hat{C}_3) = -1$
$\hat{A}_{33} = \hat{\sigma}_z \otimes \hat{\sigma}_z; M_1(\hat{A}_{33}) = +1$	$\hat{A}_{23} = \hat{\sigma}_y \otimes \hat{\sigma}_y; M_2(\hat{A}_{23}) = -1$	$\hat{A}_{13} = \hat{\sigma}_x \otimes \hat{\sigma}_x; M_3(\hat{A}_{13}) = +1$
$\lvert\psi_{\text{final}}\rangle = \lvert 00\rangle$	$\lvert\psi_{\text{final}}\rangle = \frac{1}{\sqrt{2}}(\lvert 00\rangle + \lvert 11\rangle)$	$\lvert\psi_{\text{final}}\rangle = \frac{1}{\sqrt{2}}(\lvert 00\rangle + \lvert 11\rangle)$

Table 1: HV model applied to the Peres Mermin situation

a well defined value, and so does $\hat{C}$. Thus by our accepted definition, there is no contextuality. It is just that $M_1(\hat{C}) \neq \tilde{f}(M_1(\hat{B}_1), M_1(\hat{B}_2), \dots)$, viz. the theory is *non functionally-consistent.* Note that after measuring $\hat{C}$, however, $M_2(\hat{B}_1) = +1$ and $M_2(\hat{B}_2) = -1$ (for any value of c) consistent with those deduced by measuring $\hat{C}$. Evidently, *functional-consistency* must hold for the common eigenkets of $\hat{B}_i$'s. Consequently, any violation of *functional-consistency* must arise from states that are super-positions or linear combinations of these eigenkets.

Second, note that entanglement is not necessary to demonstrate *non functional-consistency*; for instance the Peres Mermin test is a state independent test where a separable state can be used to arrive at a contradiction. On the other hand if everything is *functionally-consistent* in a situation, can we have violation of Bell's inequality or nonlocality? The answer is no and, therefore, one can say that Bell's inequalities bring out non-local consequences of *non functionally-consistent* prediction maps and the notion of *non functional-consistency* is more basic.

5. Conclusion

In this letter we have presented *non functional-consistency* as an alternative to contextuality and as an essential signature of quantumness at the kinematic level. Our result points to a (quantum) dynamical exploration of contextuality which so far has effectively been studied kinematically only. We expect our result to provide new insights that will be useful in the areas of foundations of QM and QIP.

Acknowledgments: ASA and Arvind acknowledge funding from KVPY and DST Grant No. EMR/2014/000297 respectively.