Quantum Cryptographic Security from Contextuality

Gelo Noel M. Tabia [*]

Perimeter Institute for Theoretical Physics,
31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5

Institute for Quantum Computing and University of Waterloo,
200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1

25 November 2010

Summary

Quantum mechanics is *contextual*, that is, the outcome of a measurement does depend on how that value is measured. This somewhat bizarre, non-classical feature is a consequence of the Kochen-Specker theorem, which asserts that there is no non-contextual hidden variable theory that reproduces quantum theory.

Contextuality is a property of quantum systems with Hilbert spaces of $d > 2$, as it can be shown that a consistent non-contextual assignment of values is possible for qubits. Contextuality via the Kochen-Specker theorem also serves as a precondition for secure quantum key distribution in EPR-type protocols such as E91 and BBM92. In these cryptographic schemes, security is proven by violations of inequalities that also contradict assumptions in local realistic theories of the Kochen-Specker type.

Device-independent security from contextuality is also demonstrated by considering a distributed box of Peres-Mermin observables for the Kochen Specker theorem. The security of the corresponding key distribution protocol is certified by checking for not-too-strong violations of Bell inequalities.

[*] Email: gtabia@perimeterinstitute.ca

Table of Contents

1 Introduction

One of the landmark applications of quantum mechanics is quantum cryptography [1], where quantum resources are used by Alice and Bob in order to obtain a shared secret key, which they can later employ for encoding messages protected from any potential eavesdropping attack from Eve. The unconditional or information-theoretic security of quantum cryptographic protocols rests on the validity of physical laws that dictate several key properties of quantum systems [17]:

- indeterministic measurement outcomes for individual systems;
- complementarity, which implies the incompatibility of measuring non-commuting observables;
- coherent superpositions of classically mutually exclusive states;
- value indefiniteness and contextuality, as attested in the theorems by Bell, Kochen, and Specker, among others; and
- entanglement, with the ability of obtaining stronger-than-classical correlations.

The first quantum key distribution scheme, the famous BB84 protocol [2], was developed by Charles Bennett and Gilles Brassard. BB84 is based on the fundamental quantum feature that mutually non-orthogonal states are not perfectly distinguishable. This important property is a consequence of the no-cloning theorem. A proof of the theorem demonstrates that any linear cloning machine will be successful if and only if it generates copies of a single known state or any set of orthogonal states. Because quantum states include coherent superpositions, an arbitrary quantum state can't be cloned exactly. Therefore, Eve is not allowed to just create an extra copy of the quantum information Alice transmits to Bob. Eve can attempt to measure the signals transmitted by Alice through an intercept-resend attack of the quantum channel but each measurement she performs leaves a noticeable trace in the signals received by Bob, errors which Alice and Bob can reliably detect during classical post-processing. Thus, BB84 is secure because of the unavoidable trade-off between information gain and state disturbance.

Artur Ekert independently came up with a quantum key distribution protocol [3] that utilizes a maximally entangled qubit pair for the signal source. Alice and Bob check violations of Bell inequalities to see if Eve is present in the system. His seminal paper highlights two important ideas for cryptography:

- that quantum nonlocality is useful for generating a secure key, and
- that entanglement can be used to illustrate the compromise between gaining information about a state and disturbing it via measurement.

The latter concept has basically spawned the vast majority of quantum key distribution protocols that have been established. Only after the pioneering paper of Jonathan Barrett, Lucien Hardy and Adrian Kent [4] did the former concept became more appreciated and used for the latest generation of cryptographic protocols, those with so-called device-independent security [5]. Device-independent cryptography assumes that devices can be malicious and therefore should not be trusted by Alice and Bob. Two general approaches are involved: either one begins with the assumption that superluminal signalling is impossible, or one just assumes that quantum mechanical laws are valid without specifying models for Alice's signal source or Bob's measurement apparatus. In any approach, security is confirmed solely through measurement statistics.

In this project, generating a secret cryptographic key from contextuality is explored. The first two sections shows the significance of contextuality under the framework of hidden variable models for quantum theory, leading to a celebrated result in quantum foundations, the Kochen-Specker theorem. The next two sections demonstrate some indirect roles played by contextuality in quantum key distribution protocols. The main result on security from contextuality appears in Section 6.

2 Hidden variables and contextuality

Quantum theory has the peculiar feature that states provide only a statistical description of physical systems, reflected in the probabilistic nature of measurement outcomes. A natural conclusion from this would be that quantum mechanics is incomplete, in the sense that the theory may be supplemented with some additional components so that the resulting picture fits more closely with our classical intuitions.

Recall that in classical mechanics, the state ω of a system is associated with a point or region in phase space Ω, typically that of position and momentum, i.e.,

$$\omega = (x, p). \tag{1}$$

Some properties of the system such as mass or charge stays constant while others, called dynamical variables, change with time. For each dynamical property A, there is a mapping

$$f_A : \Omega \to \mathbb{R} \tag{2}$$

such that observable A has value $f_A(\omega)$ if the state is ω. Time evolution for a classical state is given by Hamilton's canonical equations:

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}, \qquad \frac{dp}{dt} = -\frac{\partial H}{\partial x}. \tag{3}$$

The key idea is that a system's properties can be ascribed deterministic values. A classical state is fully specified by listing down all relevant properties, something which seems impossible to do in quantum mechanics due to the intrinsic randomness of individual measurements.

In order to recover some semblance of classical determinism, the standard approach is to introduce a hidden variable model for quantum theory.

2.1 The philosophy behind hidden variables

Despite the many successes of quantum mechanics in explaining atomic phenomena, there is still some ongoing debate on the philosophical implications of its mathematical formalism. The problem lies in the fact that some paradoxical situations arise that are inconsistent with what one would expect to happen classically. Various attempts have been made to interpret quantum theory, some approaches which treat the quantum state as a state of knowledge while others consider it to be objectively real. In the latter situation, the hope is that a hidden variable reconstruction of quantum mechanics will allow for an essentially classical description of quantum systems.

A hidden variable model postulates that beneath the measurable quantities dealt with by the theory, there are other quantities inaccessible to measurement but whose values dictate individual outcomes obtained when measuring observables. The probabilistic outcomes are realized as average values of observables over hidden variables with deterministic values. In the papers of John Bell [6], Simon Kochen and Ernst Specker [7], they address the question of whether hidden variables can be fitted into quantum mechanics. Whether the hidden variables are contingently inaccessible or always hidden in principle is an independent issue that will not be relevant to us here.

What exactly does one wish to achieve with hidden variables? The foremost motivation behind hidden variables is to recover the notion of value definiteness [9], which is described as follows:

Definition 1 (Value definiteness)**.** *Every physical observable of a quantum system has a definite value at all times.*

Note that value definiteness is a prominent feature of classical physics. Aside from the fact that this seems like a nice property to also have in quantum mechanics, value definiteness is intimately connected to the philosophy of realism, here specifically for properties of a system. In contemporary philosophy, realism is the belief in the existence of certain objects independent of our thoughts, beliefs, or conceptions about them. A realist will typically profess to three fundamental beliefs [8]:

(a) Reality consists of everything that does exist. (This is meant to distinguish ontological or physical existence from Platonic realism.)

(b) Reality is independent of any act of observation.

(c) Some of the features of reality are accessible to our knowledge.

For scientists, realism is an attractive ideology because it agrees with the prevailing perception that science is more than just the mere accumulation of facts and observational data. Almost any scientist believes that a most crucial aspect of his profession involves explaining why nature behaves the way that it does, and the most straightforward way to do so is to contend that science explains objective reality. If properties are real, then any attempt to measure it merely reveals a pre-existing value, one that would be present whether or not any attempt has been made to measure it.

To make it concrete that property values are independent of measurement, a second assumption is typically introduced, called non-contextuality.

Definition 2 (Non-contextuality)**.** *A quantum observable has a value independent of how that value is eventually measured.*

This means that if a system possesses a given property, it does so independently of possessing other values pertaining to other arrangements. Thus, both assumptions incorporate the basic idea of physical reality existing independently of it being measured.

The result by Kochen and Specker show that a contradiction occurs when one supposes value definite, non-contextual hidden variables for quantum mechanics. One is logically forced to renounce either value definiteness or non-contextuality. But because the purpose of hidden variables is to retain some aspects of classical realism, the question then becomes, how do you come up with a consistent story for quantum mechanics that is value-definite but contextual? This is what generates substantial philosophical interest in the consequences of the Kochen-Specker theorem.

2.2 Constraints for hidden variable models of quantum theory

A fundamental doctrine in standard quantum mechanics states that a measurement does not simply uncover the pre-existing value of a measured property [11]. Rather, the act of measurement itself is supposed to create the outcome obtained in the measuring device. Most physicists simply accept this to be an empirical fact confirmed by the many experiments involving quantum systems. However, people in quantum foundations are interested in exploring this question further, trying to figure out whether the unpredictability in measurements is an epistemic or ontological constraint, that is, whether there are hidden variable models that account for the indeterminacy at a deeper level of reality or whether measurement outcomes are inherently random.

In his book on quantum theory [10], John von Neumann provided the first restrictions on the sort of hidden variable models one should allow. However, his 'no-go' theorem was later severely criticized for being 'silly' or too restrictive. In 1964, Bell himself rediscovered von Neumann's argument and developed more reasonable constraints on hidden variables, revolving around the concept of locality. Bell's theorem demonstrates that the local hidden variables necessitates some correlation conditions which are violated by measurements performed on entangled quantum systems. Further work by Bell in

1966 [6], and independently by Kochen and Specker in 1967 [7], showed that separate from the issue of non-locality, any hidden variable model must be contextual if it is to properly account for incompatible quantum observables. In order to describe the Kochen-Specker theorem, it is useful to first discuss some general characteristics of a hidden variable theory.

Quantum mechanics asserts that given a state $|\psi\rangle$ the probability of getting a particular outcome a_j when measuring observable $A = \sum_j a_j A_j$ is provided by the *Born rule*

$$p(j) = \langle\psi|A_j|\psi\rangle. \tag{4}$$

Hidden variables entertain the possibility that the results of a measurement are deterministic by associating the quantum state with an ensemble of systems where each member of the ensemble does have a specific value for every observable. The usual quantum mechanical rules that dictate the statistics of a measurement then stem from an averaging effect over some concealed attribute of individual members in an ensemble. This means that the uncertainty principle is a restriction not on the possible values of complementary properties of a system but rather is a limitation on the types of ensembles possible to prepare from individual systems due to disturbance effects from the state-preparation process. A physicist trained in standard quantum mechanics will be skeptical and dismissive of such a possibility but will realize that the mathematical formalism by itself does not necessarily preclude such a scenario. A no-go theorem for hidden variables attempts to provide a refutation of the situation described above, by making a few reasonable assumptions on how to make hidden variables self-consistent.

In the earlier example, the case of a single property of a system with state $|\psi\rangle$ was considered. The more general situation involves a set of mutually commuting observables. Quantum theory tells us that these observables can be measured simultaneously, giving a joint probability distribution for the value of each observable. Since every observable is associated with a Hermitian operator and the possible outcomes of measuring it is given by its eigenvalues, then it is reasonable to think that a similar restriction extends to compatible observables. For instance, suppose there is a commuting set of observables A, B, C which obeys the functional identity

$$f(A, B, C) = 0 \tag{5}$$

then the set of values $v(A), v(B), v(C)$ of A, B, C also satisfies a similar relation:

$$f(v(A), v(B), v(C)) = 0. \tag{6}$$

Quite remarkably, the Kochen-Specker theorem shows that one arrives at a counterexample just by considering these constraints.

3 The Kochen-Specker theorem

Early generations of quantum mechanics practitioners believed that it was impossible to construct hidden-variable theories due to a proof by von Neumann [10]. His claim starts with a simple consequence of eq. (5) and (6). If two observables A and B commute, then the value of $C = A + B$ must obey

$$v(C) = v(A) + v(B). \tag{7}$$

Von Neumann argued that eq. (7) should hold for any hidden variable theory even when A and B don't commute. But non-commuting variables don't have simultaneous eigenvalues, meaning they can't be measured together and making such a constraint unreasonable. Note that for ensemble described by state $|\phi\rangle$,

$$\langle\phi|C|\phi\rangle = \langle\phi|A|\phi\rangle + \langle\phi|B|\phi\rangle. \tag{8}$$

However, this does not necessarily imply that it must hold for individual systems. Bell showed that a hidden-variable model is possible if one constrains only such expectation values. The Kochen-Specker theorem essentially corrects the defect in von Neumann's argument, and thus strengthens the case against hidden variable theories insofar as they assume eq. (6) holds only for sets of observables which are all mutually compatible. Later, it is shown that, at least for a certain class of deterministic hidden variable theories, such a special rule for commuting observables is, in fact, unnecessary.

3.1 The original argument by Kochen and Specker

The example considered by Kochen and Specker examines the observables for the angular momentum components for a particle of spin-1. The relevant observables are the squares of the components of the spin S_x^2, S_y^2, S_z^2 along orthogonal directions x, y, z. These observables have eigenvalues 0 or 1 (they are projectors) since each spin component of an $s = 1$ particle have values 0,1 or -1. The quantum mechanics of spin also tells us that the sums of squared spin components along orthogonal directions must obey

$$S_x^2 + S_y^2 + S_z^2 = s(s+1) = 2. \tag{9}$$

Observe that $\{S_x^2, S_y^2, S_z^2\}$ forms a mutually commuting set.

Suppose one is provided with a set of directions containing different orthogonal triples along with its associated squared spin components. Since each triple of rays involve compatible observables, they can be simultaneously measured. Any such measurement yields a combination of two values of 1 and one value of 0, in order to satify eq. (9). At first glance, it would appear that it is possible to assign a value of 0 or 1 to all such triples of rays that involve orthogonal projectors. However, for dimensions three (or larger), it is possible that two different orthogonal triples share a common observable, e.g.

$$\{S_v^2, S_w^2, S_x^2\} \text{ and } \{S_x^2, S_y^2, S_z^2\} \text{ can be both compatible sets.} \tag{10}$$

A consistent assignment of 0s and 1s should assign a fixed value to a particular direction. Kochen and Specker showed that this, in fact, is not possible for a particular set of 43 orthogonal triples constructed using 117 rays.

An explicit statement of the theorem goes as follows: [9]

Theorem 1 (Kochen-Specker Theorem). *Let $\mathcal{H}$ be a Hilbert space of dimension $d = 3$. There is a set S of observables on $\mathcal{H}$, containing N elements, such that the following two assumptions are contradictory:*

1. *All N members of S simultaneously have values, that is, they are unambiguously mapped onto some real numbers (designated by $v(A), v(B), v(C), \ldots$ for observables $A, B, C, \ldots$).*
2. *The values taken by observables in S conform to the following constraints:*
   - *If A, B, C are all compatible and $C = A + B$, then $v(C) = v(A) + v(B)$.*
   - *If A, B, C are all compatible and $C = AB$, then $v(C) = v(A)v(B)$.*

The first assumption corresponds to the condition of value-definiteness. The second assumption corresponds to functional identity constraints for commuting observables (generally called the sum rule and the product rule, respectively) and is a consequence of non-contextuality.

The Kochen-Specker result in $\mathbb{R}_3$ can be viewed as a solution to the coloring problem on the surface of a sphere. For example, if the color green is used to denote 0 and the color red to denote 1, this means that the surface of a sphere cannot be colored with red and green such that for every orthogonal triple

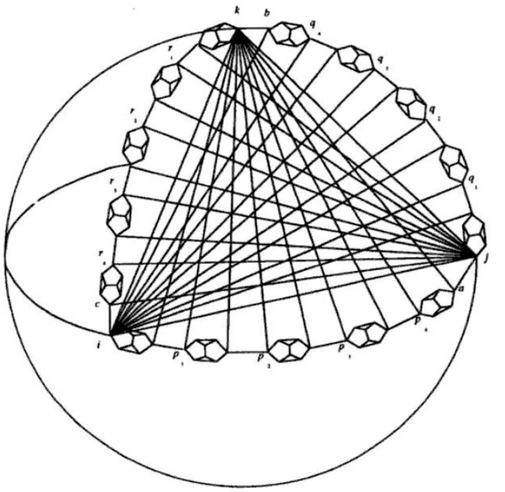


Figure 1: *Three dimensional version of the original Kochen-Specker diagram. [Image reproduced from A. Cabello, Fundamentals Problems in Quantum Physics, p.45. [14]]*

of points on the sphere (which specify three orthogonal directions in space), one and only one point is green while the other two are colored red. The impossibility of such a coloring of the sphere reflects a topological property which shows that two points of different colors can't be arbitrarily close to each other if they are to satisfy the functional identity constraint. The actual steps in the proof given by Kochen and Specker is quite complicated and not very illuminating so the details are omitted. Fig. (1) shows the $N = 117$ directions used by Kochen and Specker in their proof. The diagram contains 43 triangles representing the orthogonal triples of points yielding a uncolorable set. Note that although the proof uses observables for $d = 3$, it also applies for $d > 3$ because any higher-dimensional Hilbert space will contain a three-dimensional subspace where similar sets of directions and observables can be selected.

The Kochen-Specker result can also be thought of as a corollary of Gleason's theorem [12]. Gleason's theorem shows that the set of quantum states is complete, in the sense that all possible probability measures definable on the set of quantum propositions represented by Hilbert space projection operators $\{E_\alpha\}$ are generated by density operators ρ of pure or mixed states according to the Born rule:

$$p(\alpha) = \operatorname{tr}\{\rho E_\alpha\}. \tag{11}$$

The Kochen-Specker theorem follows from considering a spectral decomposition of ρ and attempting to assign non-contextual values to every projector in that decomposition.

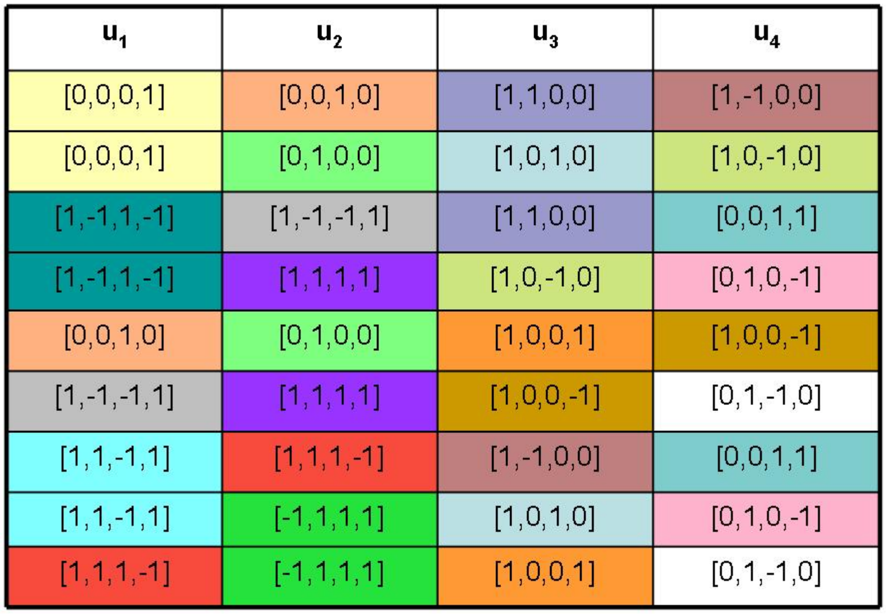
3.2 Simpler proofs of the Kochen-Specker theorem

Although the Kochen-Specker theorem can be proven for $d \geq 3$, two more elegant proofs were obtained by considering directions in four dimensions. First, let's examine the proof by Cabello, et al. [13] that uses 18 real vectors.

It is possible to think of the assignment of values to compatible observables (projectors) which are members of different sets (bases) as a set of propositions (which is either yes/no) belonging to different contexts. For example, consider the set P of orthogonal vectors

$$P = \{|1\rangle, |2\rangle, |3\rangle, |4\rangle\} = \{(0,0,0,1), (0,0,1,0), (1,1,0,0), (1,-1,0,0)\} \tag{12}$$

where we omit normalization factors since they are unimportant in the proof. Each vector has a corresponds to a projection operator, which becomes the relevant observable taking values 0 or 1. The set P forms a complete basis for $\mathcal{H} = \mathbb{R}_4$, which also means that the corresponding (orthogonal)



u_1	u_2	u_3	u_4
[0,0,0,1]	[0,0,1,0]	[1,1,0,0]	[1,-1,0,0]
[0,0,0,1]	[0,1,0,0]	[1,0,1,0]	[1,0,-1,0]
[1,-1,1,-1]	[1,-1,-1,1]	[1,1,0,0]	[0,0,1,1]
[1,-1,1,-1]	[1,1,1,1]	[1,0,-1,0]	[0,1,0,-1]
[0,0,1,0]	[0,1,0,0]	[1,0,0,1]	[1,0,0,-1]
[1,-1,-1,1]	[1,1,1,1]	[1,0,0,-1]	[0,1,-1,0]
[1,1,-1,1]	[1,1,1,-1]	[1,-1,0,0]	[0,0,1,1]
[1,1,-1,1]	[-1,1,1,1]	[1,0,1,0]	[0,1,0,-1]
[1,1,1,-1]	[-1,1,1,1]	[1,0,0,1]	[0,1,-1,0]

Figure 2: *18-vector proof of the Kochen-Specker theorem. Each row of vectors u_1, u_2, u_3, u_4 form a basis for $\mathbb{R}_4$ and corresponds to a set of orthogonal projectors that sum to the identity. Each vector appears exactly twice in the nine bases, as indicated by the color shading. Therefore, any value assignment of 0s and 1s to all vectors will yield an even sum for all bases, which contradicts the fact that there are an odd number of bases.*

projectors add up to the identity matrix, i.e. let $P_i = |i\rangle\langle i|$ be the projector associated with vector $|i\rangle$, then

$$P_1 + P_2 + P_3 + P_4 = \mathbb{1}_4. \tag{13}$$

In the language of propositions, the identity matrix corresponds to the trivially true statement, since for any decomposition of the identity, at least one of the elements must hold. Thus, according to the functional identity rule for compatible observables,

$$v(P_1) + v(P_2) + v(P_3) + v(P_4) = 1 \tag{14}$$

where one of the values $v(P_j)$ is 1 while the rest are zero. Cabello's Kochen-Specker proof uses the following nine 'interlocking' bases to arrive at a contradiction:

$$\begin{aligned}
v(0,0,0,1) + v(0,0,1,0) + v(1,1,0,0) + v(1,-1,0,0) &= 1,\\
v(0,0,0,1) + v(0,1,0,0) + v(1,0,1,0) + v(1,0,-1,0) &= 1,\\
v(1,-1,1,-1) + v(1,-1,-1,1) + v(1,1,0,0) + v(0,0,1,1) &= 1,\\
v(1,-1,1,-1) + v(1,1,1,1) + v(1,0,-1,0) + v(0,1,0,-1) &= 1,\\
v(0,0,1,0) + v(0,1,0,0) + v(1,0,0,1) + v(1,0,0,-1) &= 1,\\
v(1,-1,-1,1) + v(1,1,1,1) + v(1,0,0,-1) + v(0,1,-1,0) &= 1,\\
v(1,1,-1,1) + v(1,1,1,-1) + v(1,-1,0,0) + v(0,0,1,1) &= 1,\\
v(1,1,-1,1) + v(-1,1,1,1) + v(1,0,1,0) + v(0,1,0,-1) &= 1,\\
v(1,1,1,-1) + v(-1,1,1,1) + v(1,0,0,1) + v(0,1,-1,0) &= 1.
\end{aligned} \tag{15}$$

Observe that every vector appears twice on the left-hand side. This means that however the values 0

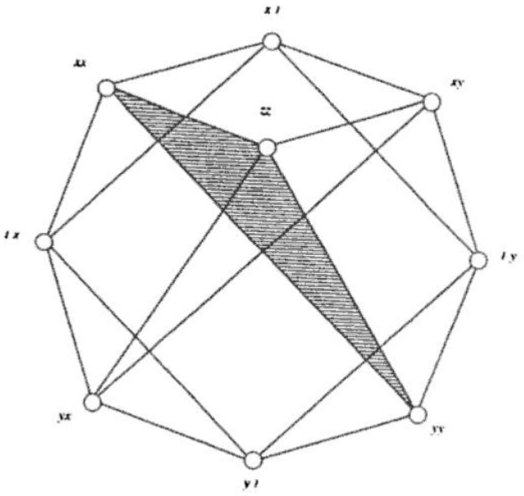


Figure 3: *Peres-Mermin example of the Kochen-Specker theorem. Each point represents one of the nine observables. Triangles connect points which are mutually commuting. The shaded triangle corresponds to the last column of observables in (16), which has product -1. [Image reproduced from A. Cabello, Fundamentals Problems in Quantum Physics, p.45. [14]]*

and 1 are distributed, the sum of all terms on the left-hand-side must be even. However, since there are nine equations, the sum of all terms on the right-hand-side is necessarily 9, which is odd. Therefore, assigning yes/no answers to all propositions can't be done independent of context. The nine bases are also displayed in fig. (2).

Another proof that will become crucial in a later discussion about acquiring a secret key from the Kochen-Specker paradox is often attributed to both Asher Peres and David Mermin [11]. It involves the following nine observables constructed from composing two spin-1/2 systems:

$$\begin{array}{c|c|c} \sigma_x^{(1)} & \sigma_x^{(2)} & \sigma_x^{(1)}\sigma_x^{(2)} \\ \hline \sigma_y^{(1)} & \sigma_y^{(1)} & \sigma_y^{(1)}\sigma_y^{(2)} \\ \hline \sigma_x^{(1)}\sigma_y^{(2)} & \sigma_y^{(1)}\sigma_x^{(2)} & \sigma_z^{(1)}\sigma_z^{(2)} \end{array} \tag{16}$$

Since these are Pauli operators for two independent subsystems 1 and 2, they satisfy

$$[\sigma_a^{(1)}, \sigma_b^{(2)}] = 0, \qquad \sigma_a^{(i)}\sigma_b^{(i)} = \delta_{ab} + i\epsilon_{abc}\sigma_c, \qquad i = 1, 2; a, b, c = x, y, z. \tag{17}$$

A careful examination of the observables in eq. (16) reveals the following:

1. The observables in each of the three rows and columns mutually commute. This is immediately obvious for the first two rows and columns. It is true for the last row and column because of a pair of anti-commutations that cancel each other.
2. The product of the observables in the last column is -1. The product of all other columns and all rows is -1.
3. Since the values assigned to compatible observables must obey functional identities satisfied by the observables themselves, the product rule in (2) must be followed.

But (3) is impossible to satisfy since the row identities require the product of the nine observables to be 1 whereas the column identities imply that the product should be -1. This concludes the Peres-Mermin proof. A proposed diagram for the Peres-Mermin proof is depicted in fig. (3).

3.3 Proposed experimental tests for contextuality

The differences between the proof of Bell's theorem and Kochen-Specker theorem lead to the question of what is the connection between them. In a later section, we will show how certain assumptions

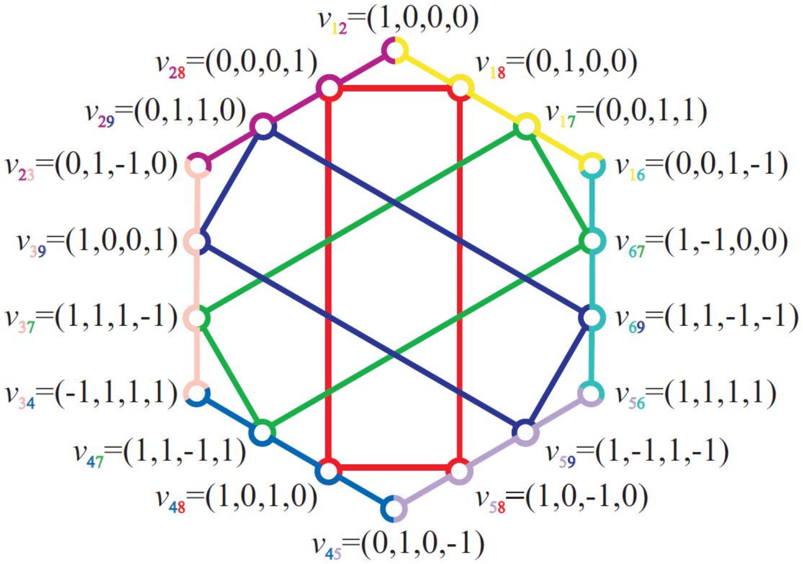


Figure 4: *A diagram for observables to test the Kochen-Specker theorem experimentally. Each dot represents a projector onto the vector v_{ij}. The sides of the regular hexagon and the three rectangles in the middle connect orthogonal projectors.*

of a deterministic hidden variable theory implies an essential equivalence in the consequences of both results. For the meantime, our interest lies on whether experimental tests similar to Bell's can be made for the Kochen-Specker theorem, or if they are even possible in the realm of finite measurement precision.

Adan Cabello proposed the following inequality as the Kochen-Specker equivalent for Bell-type experiments [15]: Suppose that A_{ij} is an observable with eigenvalues ±1. Two observables A_{ij} and A_{kl} commute if they share a subscript index (e.g., $j = k$ or $i = l$, etc.) Denote $\langle A_{ij}A_{kl}A_{pq}A_{rs}\rangle$ to be the average of the products of the indicated observables. Then according to any non-contextual hidden-variable theory with definite values for observables A_{ij}, the following inequality must be satisfied:

$$\begin{aligned}-\langle A_{12}A_{16}A_{17}A_{18}\rangle - \langle A_{12}A_{23}A_{28}A_{29}\rangle - \langle A_{23}A_{34}A_{37}A_{39}\rangle - \langle A_{34}A_{45}A_{47}A_{48}\rangle - \langle A_{45}A_{56}A_{58}A_{59}\rangle \\ -\langle A_{16}A_{56}A_{67}A_{69}\rangle - \langle A_{17}A_{37}A_{47}A_{67}\rangle - \langle A_{18}A_{28}A_{48}A_{58}\rangle - \langle A_{29}A_{39}A_{59}A_{69}\rangle \leq 7. \quad (18)\end{aligned}$$

The proof is rather straightforward: Define some parameter β to be equal to the left-hand-side of the inequality. A brute-force calculation of all possible values of β will yield a maximum of 7; thus, $\langle\beta\rangle \leq 7$. To measure β, subsets of compatible observables are measured on different sub-ensembles prepared in identical states, since the inequality holds for averages over each sub-ensemble. In contrast to Bell experiments that assume independence of space-like separated measurements, here our assumption is that the results will be independent of compatible measurements. The inequality (18) is clearly violated by any quantum state in a four-dimensional Hilbert space if we choose the observables to be

$$A_{ij} = 2|v_{ij}\rangle\langle v_{ij}| - \mathbb{1}, \quad (19)$$

where the vectors $|v_{ij}\rangle$ are defined in fig. (4). The observables are nothing more than the projectors for the 18 vectors forming nine bases in Cabello's proof of the Kochen-Specker theorem. Quantum mechanics predicts that the left-hand-side of eq. (18) must be 9 in any state.

The inequality can be tested experimentally as long as sequential compatible measurements are indeed compatible [16]. One way to achieve this would be to convert local contextuality constraints into quantum nonlocality conditions [33], after which, some Bell-type inequality can be checked.

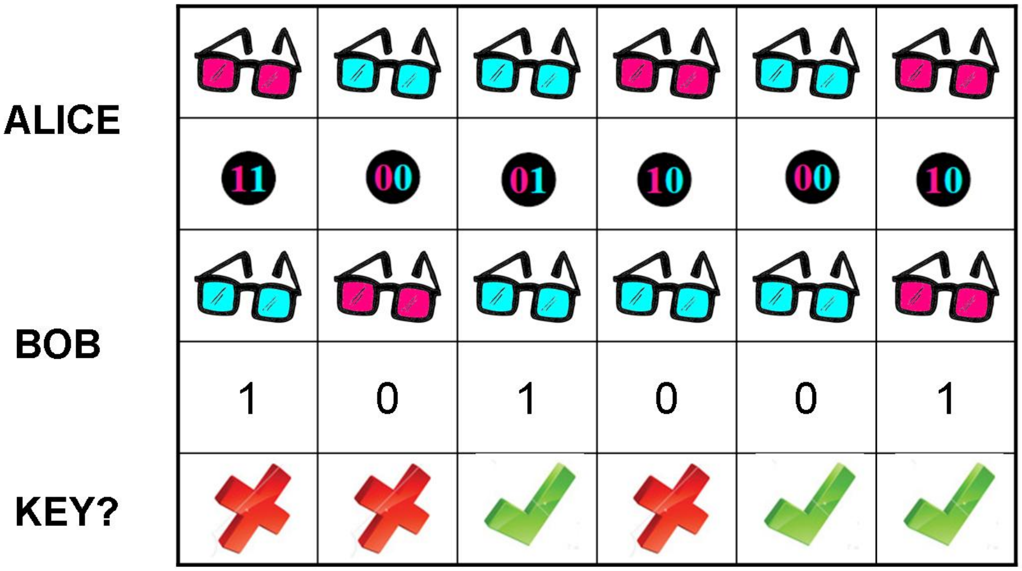


Figure 5: *BB84 with chocolate balls. Quantum states are replaced by chocolate balls carrying two complementary bit values, here colored red and blue. Alice and Bob can only read off the value for the colored eyeglass they choose.*

4 Non-contextual cryptography: BB84 with chocolate balls

It is rarely acknowledged that, when it comes to contextuality, there is definitely a difference between two and three dimensional Hilbert spaces [17]. This difference can be explained easily in terms of conjugate bases: different qubit bases are fully disjoint and separated (trivially sharing only the origin) whereas for higher dimensional spaces, various orthogonal bases can share common elements. It is the interlocking of various bases that allows for both Gleason's theorem [12] and the Kochen-Specker theorem [7] to be true. Thus, the contextuality of quantum mechanics holds only if the quantum system in question is at least three-dimensional. In particular, a non-contextual hidden variable model is possible for qubits. To illustrate this, we use the BB84 protocol as an example and demonstrate that one can provide a quasi-classical picture of what goes on.

Recall that in BB84 [2], Alice randomly selects signals from two conjugate bases $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$, assigning vectors from each basis the bit value 0 or 1. Alice sends her signals one at a time to Bob and Bob measures the signals either in one of the two bases. After transmission, Alice and Bob talk in an authenticated classical channel and compare bases that they used. They discard those results where their bases disagree and those that remain will have matching bit values, which form a random bit string and can be subsequently used as a cryptographic key.

In BB84 with chocolate balls [18], the same steps are used except that instead of using quantum signals, one envisions a source of chocolate balls painted with two symbols having two possible values (0 or 1) in two colors (red or blue), many copies of which are randomly distributed in an urn (this is actually an example of the so-called generalized urn models [19]). The idea behind the colored symbols is that only one of the two symbols is accessible–only the one with a color that matches that of a viewing eyeglass. In essence, choosing one of the colors decides which one of two complementary observables 'red' or 'blue' is measured. The bit value is then given by the value associated with the color selected. The revised protocol runs in the following manner:

1. Alice selects at random a colored eyeglass.
2. Alice draws a chocolate ball from the urn and uses her chosen eyeglass to read off the bit value and records it in her lab manual.

3. Alice then sends the chocolate ball over to Bob.
4. Bob also randomly chooses a colored eyeglass with which he measures a bit value from the chocolate ball.
5. After transmission, Alice and Bob compare eyeglasses used and keep those results where their choices agree. The corresponding bit string can be used as a secure random key.

These steps are illustrated in fig. (5).

The lesson from the example above is this: since the dimension of Hilbert space determines the number of mutually exclusive outcomes in quantum mechanics, a necessary condition for a quantum system to be protected by contextuality is that the quantum system must have at least three possible orthogonal states [17]. Of course, for the purposes of cryptography it might not mean much because the BB84 protocol is still protected by complementarity. Also, any quantum cryptographic protocol that uses three [20] or more orthogonal quantum states is implicitly protected by contextuality.

5 Quantum key agreement and Kochen-Specker realism

Bell's theorem [21] states that a local hidden variable theory produces correlations which must satisfy some natural constraints often expressed in terms of inequalities (sometimes referred to as the Bell-Clauser-Horne inequalities). Bell goes on to show that such a local realistic model fails to reproduce all statistical predictions made by quantum mechanics. The usual way to demonstrate the violation is to check the correlations on outcomes obtained from measuring entangled systems. In Ekert's E91 protocol, the security of the scheme is guaranteed by violations of such Bell inequalities.

Koji Nagata stated that the violation of Ekert's inequality and also the violation of a similar inequality in the EPR version of the BB84 protocol called the BBM92 scheme (by Charles Bennett, Gilles Brassard, and David Mermin) constitutes a denial of non-contextual hidden variables, or Kochen-Specker realism, for the scheme [25]. Therefore, another link between no-go theorems and success in quantum key distribution can be made via the Kochen-Specker theorem.

5.1 Deterministic hidden variables and Kochen-Specker realism

To establish the connection between quantum cryptographic security and no-hidden-variable proofs, it is instructive to first specify the details of a deterministic (or factorizable stochastic) hidden variable model [22, 23, 24]. In particular, the following result is desired: Kochen-Specker realism is equivalent to a situation where all quantum observables commute [26].

Denote a quantum system with state ρ to be $Q = Q(\mathcal{H}, \rho, \Theta)$ where the Hilbert space $\mathcal{H}$ specifies its dimension and Θ corresponds to a set of observables for Q.

Let $\Lambda = \Lambda(\Omega, \Sigma(\Omega), \mu)$ be a classical probability space, where Ω is a nonempty set, $\Sigma(\Omega)$ is a Boolean algebra of subsets of Ω and μ is a probability measure on $\Sigma(\Omega)$.

A deterministic hidden variable model for quantum system Q makes one or more of the following assumptions

(a) Given $\omega \in \Omega, A \in \Theta$ there is a mapping f from $(\omega, A) \in (\Omega, \Theta)$ to $\mathbb{R}$ such that

$$f_\omega(A) = f(\omega, A) = a, \qquad \text{where } A|a\rangle = a|a\rangle, \tag{20}$$

that is, the function f yields an eigenvalue a of A.

(b) For any two commuting observables $A, B \in \Theta$, the mapping f is such that

$$f(\omega, A+B) = f(\omega, A) + f(\omega, B). \tag{21}$$

(c) The probability measure μ gives the marginal probabilities for observable A, i.e., for any real Borel set S, μ is such that

$$\mathrm{tr}\left\{\rho \Pi_A(S)\right\} = \int f(\omega, \Pi_A(S)) d\mu \tag{22}$$

where $\Pi_A(S)$ is the projector in the spectral resolution of A associated with set S.

(d) For any two commuting observables A, B the measure μ yields conditional probabilities in the following manner: for real Borel sets S, T

$$\mathrm{tr}\left\{\rho \Pi_A(S) \Pi_B(T)\right\} = \int f(\omega, \Pi_A(S), \Pi_B(T)) d\mu \tag{23}$$

where Π_A, Π_B are the projectors corresponding to S, T in the spectral decomposition of A, B.

Arthur Fine [22] defines the set of conditions (a), (c), and (d) as a deterministic hidden-variable model (equivalently, a factorizable stochastic model). He also showed that these conditions are entirely equivalent to choosing instead (a), (b), and (d). In fact, the former set essentially corresponds to Bell's hidden-variable model while the latter choice corresponds to Kochen-Specker's conditions.

Note that the product rule for functional identity constraint follows from the Borel function rule

$$f(v(A)) = v(f(A)), \tag{24}$$

which is a natural consequence of $f(A) = \sum_j f(a_j) A_j$ and taking

$$v'(A) = \log_2(v(2^A)). \tag{25}$$

Thus, for compatible observables A, B

$$v'(A) + v'(B) = \log_2(v(2^A)) + \log_2(v(2^A)) = \log_2(v(2^A)v(2^B)) = \log_2(v(AB)). \tag{26}$$

where the product rule is used in the last step. In practice, it is more convenient to specify both sum and product rule, although only one of them is strictly necessary.

Assume that there is a classical probability space such that the outcomes for projectors $A = A^2, B = B^2$ is described by the joint distribution μ. In quantum theory, for some state ρ, the conditional probability of A given B is defined by

$$\Pr[A|B] = \frac{\mathrm{tr}\left\{AB\rho BA\right\}}{\mathrm{tr}\left\{B\rho B\right\}} = \frac{\mathrm{tr}\left\{\rho BAB\right\}}{\mathrm{tr}\left\{\rho B\right\}}. \tag{27}$$

The conditional probability rule is said to hold if the distribution defined above coincides with the distribution given by measure μ. For any projector P, define

$$P^{-1}(1) = \{\omega \in \Lambda \mid P(\omega) = 1\}, \tag{28}$$

i.e., it gives the set of all elements in Λ that yield a value of 1 for observable X. This result can be stated as a theorem:

Theorem 2 (Conditional probability rule)**.** *Suppose that* $\dim(\mathcal{H}) \geq 3$ *and that conditions (a), (c),* *and (d) hold. The conditional probability rule states that for one-dimensional projectors A, B, if*

$$\mu[a|b] = \frac{\mu[a \cap b]}{\mu[b]} = \frac{\mathrm{tr}\left\{\rho BAB\right\}}{\mathrm{tr}\left\{\rho B\right\}}, \tag{29}$$

where $a = A^{-1}(1), b = B^{-1}(1)$.

The next theorem describes the key result:

Theorem 3. *Assume* $\dim(\mathcal{H}) \geq 3$ *and a deterministic hidden-variable model for quantum events. Then all quantum observables commute.*

Proof: Let A, B be a pair of quantum observables, which can be thought of as one-dimensional projectors without loss of generality. It must be the case that

$$[A, B] = 0 \iff [A_i, B_j] = 0 \tag{30}$$

for all A_i, B_j in the spectral decomposition of A, B. All projectors may be expressed as a sum of one-dimensional ones (though the decomposition is not necessarily unique.)

Using theorem 2:

$$\mu[a, b] = \mu[a|b]\mu[b] = \frac{\mathrm{tr}\left\{\rho BAB\right\}}{\mathrm{tr}\left\{\rho B\right\}} \mathrm{tr}\left\{\rho B\right\} = \mathrm{tr}\left\{\rho BAB\right\}. \tag{31}$$

Also,

$$\mu[a, b] = \mu[b|a]\mu[a] = \frac{\mathrm{tr}\left\{\rho ABA\right\}}{\mathrm{tr}\left\{\rho A\right\}} \mathrm{tr}\left\{\rho A\right\} = \mathrm{tr}\left\{\rho ABA\right\}. \tag{32}$$

Thus, for any state ρ, since $A^2 = A, B^2 = B$

$$BAB = ABA \implies [AB, BA] = 0, BAB^2 = (AB)^2, (BA)^2 = ABA^2 \implies (AB - BA)^2 = 0. \tag{33}$$

Because $C = [A, B]$ is skew-Hermitian, i.e., $C^\dagger = [B, A] = -[A, B]$,

$$C^2 = 0 \implies C = [A, B] = 0. \tag{34}$$

Thus, a hidden-variable model of the Kochen-Specker (or Bell) type is proven to be equivalent to simultaneous commutativity of quantum observables. In particular, because conditions (b) and (c) are effectively interchangeable, it implies that the sum rule can be valid for non-commuting observables provided other conditions are present—namely, conditions (a) and (d). Therefore, von Neumann's assumption in his original no-go proof is shown to be no less physical than Bell's or Kochen-Specker's seemingly less restrictive conditions.

5.2 Success in the E91 and BBM92 protocols revisited

Part of the testing process in any quantum key agreement protocol checks for the identity of the source state, sometimes referred to as a candidate state for Alice and Bob. The idea is that the presence of Eve will introduce some noise into the quantum state of the signals received by Alice and Bob. For example, in the E91 protocol [3], Alice and Bob expect a pair of qubits in a singlet state

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}} \left(|10\rangle - |01\rangle\right). \tag{35}$$

In the presence of noise, either from Eve or from imperfections in the quantum channel, they may

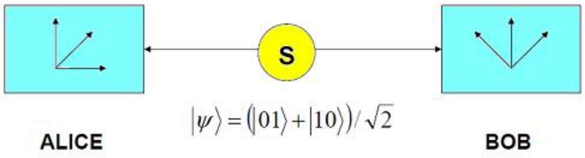


Figure 6: *The E91 protocol. A source of photon pairs in the singlet state sends a qubit each to Alice and Bob. Alice and Bob measure the polarization along the three directions:* $a_1 = 0°, a_2 = 45°, a_3 = 90°$ *for Alice and* $b_1 = 45°, a_2 = 90°, a_3 = 135°$ *for Bob.*

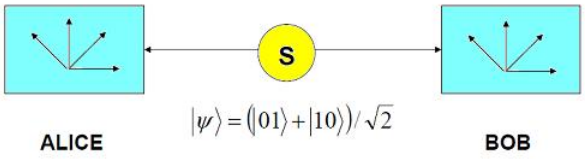


Figure 7: *The BBM92 protocol. Alice and Bob get qubits from a singlet source. Alice and Bob measure the polarization along the two bases: the horizontal-vertical and diagonal bases. Afterwards, they compare bases they used and keep those results with matching basis choices.*

instead observe a Werner state ρ

$$\rho = (1-\epsilon)|\Psi_-\rangle\langle\Psi_-| + \epsilon\frac{\mathbb{1}}{4}. \tag{36}$$

if, for example, the noise is completely unbiased. If the level of noise ϵ is too large, Alice and Bob abort the protocol. Fig. (6) summarizes how the E91 protocol is implemented. Alice and Bob verify the security of the key by testing if the correlations in their measurement results obey Bell's theorem.

In 1992, Charles Bennett, Gilles Brassard, and David Mermin argued that Bell's theorem is not an essential part of cryptographic security and proposed a simpler, EPR-type version of BB84, the BBM92 protocol [27], shown in fig. (7). They established a similar inequality for measurement correlations but concluded that the EPR effect is superficial for successful quantum key distribution. Koji Nagata then showed in 2005 that for both the E91 and BBM92 protocols, the criterion of success depends on the fidelity to an EPR state [25]. This leads to an explicit construction of inequalities valid for the non-contextual hidden variables as discussed previously. The violation of this Kochen-Specker realism becomes a necessary condition for the security of both protocols.

First, let us analyze the Ekert inequality. Suppose ρ is a candidate for the source state used by Alice and Bob. Let $E(A)$ be the expectation value of observable A with respect to state ρ:

$$E(A) = \mathrm{tr}\left\{\rho A\right\}. \tag{37}$$

Then the Ekert inequality is given by

$$|E(a_1b_1) - E(a_1b_1) + E(a_1b_1) + E(a_1b_1)| \leq \sqrt{2}. \tag{38}$$

The observables a_i, b_j can be written as

$$a_1 = \sigma_x^{(A)}, \quad a_3 = \sigma_y^{(A)}, \quad b_1 = \frac{\sigma_x^{(B)} + \sigma_y^{(B)}}{\sqrt{2}}, \quad b_3 = \frac{\sigma_y^{(B)} - \sigma_x^{(B)}}{\sqrt{2}}, \tag{39}$$

where $\sigma_\alpha^{(i)}, \alpha = x, y, z$ are Pauli operators for Alice and Bob ($i = A, B$), respectively.

Substituting eq. (39) into eq. (38)

$$\left| E\left(\sigma_x^{(A)}\sigma_x^{(B)}\right) + E\left(\sigma_y^{(A)}\sigma_y^{(B)}\right)\right| \leq 1. \tag{40}$$

It is straightforward to verify that

$$\mathrm{tr}\left\{\rho\left(\sigma_x^{(A)}\sigma_x^{(B)} + \sigma_y^{(A)}\sigma_y^{(B)}\right)\right\} = 2\mathrm{tr}\left\{\rho\left(|\Psi^+\rangle\langle\Psi^+| - |\Psi^-\rangle\langle\Psi^-|\right)\right\}, \tag{41}$$

where $|\Psi^{\pm}\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. This implies that

$$\left|\langle\Psi^{+}|\rho|\Psi^{+}\rangle - \langle\Psi^{-}|\rho|\Psi^{-}\rangle\right| \leq \frac{1}{2}. \tag{42}$$

Thus, it follows that the violation of the Ekert inequality in the form of eq. (42) is the same as saying that

$$\langle\Psi^{+}|\rho|\Psi^{+}\rangle > \frac{1}{2} \text{ or } \langle\Psi^{-}|\rho|\Psi^{-}\rangle > \frac{1}{2}. \tag{43}$$

That is, if the fidelity to either EPR state (equivalently, the singlet fraction) of the state ρ is such that ρ corresponds to a distillable entangled two-qubit state, then the Ekert inequality is violated and the E91 protocol is secure.

The original form of Ekert's inequality can also be derived if one supposes that Alice and Bob's correlations are induced by a separable state of the form

$$\rho = \int p(\vec{n}_A, \vec{n}_B)\rho^{(A)} \otimes \rho^{(B)} d\vec{n}_A d\vec{n}_B, \tag{44}$$

where

$$\rho^{(i)} = \frac{1}{2}\left(\mathbb{1} + \vec{n}_A \cdot \vec{\sigma}^{(i)}\right), \qquad \vec{\sigma}^{(i)} = (\sigma_x^{(i)}, \sigma_y^{(i)}, \sigma_z^{(i)}), i = A, B. \tag{45}$$

The choice of $\vec{n}_i$ corresponds to some eavesdropping strategy for Eve. Introducing $\vec{a}_k, \vec{b}_l$ such that

$$\vec{a}_k \cdot \vec{\sigma}^{(A)} = a_k, \qquad \vec{a}_k \cdot \vec{\sigma}^{(A)} = a_k. \tag{46}$$

Define

$$S = \operatorname{tr}\left\{\rho\left(a_1 b_1 - a_1 b_3 + a_3 b_1 + a_3 b_3\right)\right\}. \tag{47}$$

Since a_1, a_3 and b_1, b_3 anti-commute in the E91 scheme, $|S| \leq \sqrt{2}$ if ρ is separable. In fact, one gets Ekert's original expression [3] for S if the state ρ (44) is plugged into eq. (47).

One can do a very similar analysis with the BBM92 protocol. The BBM inequality is given by

$$\left|E\left(\sigma_x^{(A)}\sigma_x^{(B)}\right) + E\left(\sigma_z^{(A)}\sigma_z^{(B)}\right)\right| \leq 1. \tag{48}$$

Since

$$\operatorname{tr}\left\{\rho\left(\sigma_x^{(A)}\sigma_x^{(B)} + \sigma_y^{(A)}\sigma_y^{(B)}\right)\right\} = 2\operatorname{tr}\left\{\rho\left(|\Phi^{+}\rangle\langle\Phi^{+}| - |\Psi^{-}\rangle\langle\Psi^{-}|\right)\right\}, \tag{49}$$

where $|\Phi^{+}\rangle = \left(|00\rangle + |11\rangle\right)/\sqrt{2}$, it follows that

$$\left|\langle\Phi^{+}|\rho|\Phi^{+}\rangle - \langle\Psi^{-}|\rho|\Psi^{-}\rangle\right| \leq \frac{1}{2}. \tag{50}$$

This means that a BBM inequality violation implies that

$$\langle\Phi^{+}|\rho|\Phi^{+}\rangle > \frac{1}{2} \text{ or } \langle\Psi^{-}|\rho|\Psi^{-}\rangle > \frac{1}{2}. \tag{51}$$

Introducing $\vec{x}, \vec{z}$ such that

$$\vec{x} \cdot \vec{\sigma}^{(i)} = \sigma_x^{(i)}, \quad \vec{z} \cdot \vec{\sigma}^{(i)} = \sigma_z^{(i)}, \qquad i = A, B, \tag{52}$$

then one can define a correlation T:

$$T = \operatorname{tr}\left\{\rho\left(\sigma_x^{(A)}\sigma_x^{(B)} + \sigma_z^{(A)}\sigma_z^{(B)}\right)\right\} \tag{53}$$

and substitute eq. (44) into ρ to get

$$|T| \leq 1, \tag{54}$$

which yields the same inequality in the BBM92 paper [27].

Let us finally establish the connection between the two inequality violations and the refutation of Kochen-Specker realism. Recall that the Kochen-Specker proof shows a contradiction in the following way: Non-contextual realistic functions are represented by non-commuting operators in the Hilbert space formalism of quantum theory. However, the product rule and the uniqueness property of Gleason's theorem imply that all quantum operators should simultaneously commute.

Consider some realistic function $f_\omega(A)$ of some hidden variable $\omega \in \Omega$ and observable A onto the algebra $\sigma(\Omega)$ with normalized measure μ. Then,

$$E(A) = \mathrm{tr}\,\{\rho A\} = \int_{\omega\in\Omega} \mu(d\omega) f_\omega(A), \qquad \forall A \in \Theta. \tag{55}$$

For observables in terms of Pauli operators $\sigma_\alpha^{(i)}$ ($\alpha = x, y, z; i = A, B$),

$$f_\omega(\sigma_\alpha^{(i)}) = \pm 1 \tag{56}$$

since these are their eigenvalues.

The Kochen-Specker paradox occurs when both realistic functions for deterministic hidden-variables exists and the product rule holds:

$$f_\omega(A) f_\omega(B) = f_\omega(AB), \tag{57}$$

which is assumed to hold for every hidden variable ω (more complicated scenarios can be envisioned [22]). There are three cases to consider. Let $x, y = \pm 1$. Then the three scenarios are

$$\begin{aligned} &(1) \quad 1 + x + y - xy = \pm 2, \\ &(2) \quad 1 - x - y - xy = \pm 2, \\ &(3) \quad 1 + x - y + xy = \pm 2. \end{aligned} \tag{58}$$

For each scenario, we can define a function $W_j(\omega), j = 1, 2, 3$ such that

$$x = f_\omega(\sigma_x^{(A)}\sigma_x^{(B)}), \qquad y = f_\omega(\sigma_y^{(A)}\sigma_y^{(B)}), \tag{59}$$

from which $W_j = \pm 2$ implies

$$E(W_j) = \int_{\omega\in\Omega} \mu(d\omega) W_j(\omega) \leq 2, \qquad j = 1, 2, 3. \tag{60}$$

From the product rule,

$$f_\omega(\sigma_x^{(A)}\sigma_x^{(B)}) f_\omega(\sigma_y^{(A)}\sigma_y^{(B)}) = f_\omega(\sigma_x^{(A)}\sigma_y^{(B)}) f_\omega(\sigma_y^{(A)}\sigma_x^{(B)}) = f_\omega(\sigma_z^{(A)}\sigma_z^{(B)}). \tag{61}$$

Hence,

$$\begin{aligned} E(W_1) \leq 2 &\iff \quad 1 + E(\sigma_x^{(A)}\sigma_x^{(B)}) + E(\sigma_y^{(A)}\sigma_y^{(B)}) - E(\sigma_z^{(A)}\sigma_z^{(B)}) \leq 2, \\ E(W_2) \leq 2 &\iff \quad 1 - E(\sigma_x^{(A)}\sigma_x^{(B)}) - E(\sigma_y^{(A)}\sigma_y^{(B)}) - E(\sigma_z^{(A)}\sigma_z^{(B)}) \leq 2, \\ E(W_3) \leq 2 &\iff \quad 1 + E(\sigma_x^{(A)}\sigma_x^{(B)}) - E(\sigma_y^{(A)}\sigma_y^{(B)}) + E(\sigma_z^{(A)}\sigma_z^{(B)}) \leq 2 \end{aligned} \tag{62}$$

These conditions can be rewritten, respectively, as

$$\begin{aligned} &(1) \qquad \mathrm{tr}\,\{\rho|\Psi^+\rangle\langle\Psi^+|\} \leq 2, \\ &(2) \qquad \mathrm{tr}\,\{\rho|\Psi^-\rangle\langle\Psi^-|\} \leq 2, \\ &(3) \qquad \mathrm{tr}\,\{\rho|\Phi^+\rangle\langle\Phi^+|\} \leq 2, \end{aligned} \tag{63}$$

since for example, for case (2),

$$|\Psi^-\rangle\langle\Psi^-| = \frac{1}{4}\left(\mathbb{1} - \vec{\sigma}^{(A)} \cdot \vec{\sigma}^{(B)}\right). \tag{64}$$

Observe that conditions in eq. (63) are exactly the conditions necessary for satisfying the Ekert and BBM inequalities. The implication is that violation of either inequality is not possible is a non-contextual realistic model is valid in the Hilbert space of the quantum system in question. Thus, the Kochen-Specker theorem provides a precondition for secure quantum key agreement in these two EPR-type protocols.

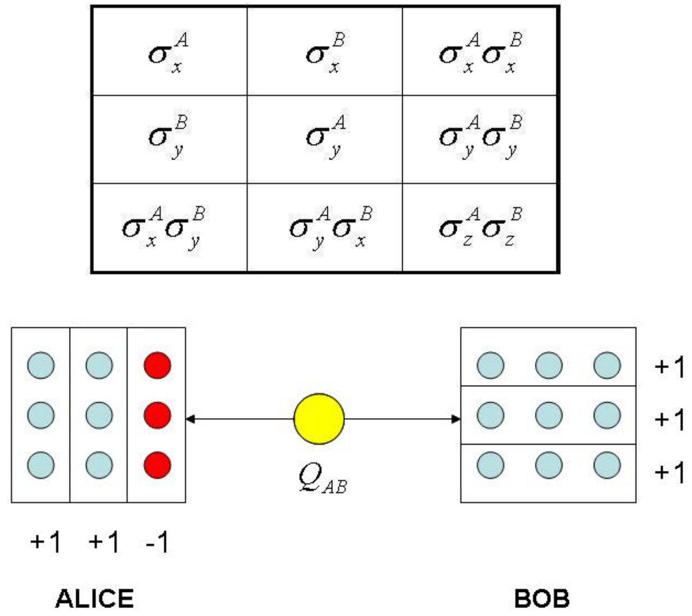
Nagata [25] also demonstrated that one can replace the Kochen-Specker product rule with a locality constraint on the hidden variables, leading to a set of Bell inequalities, the violation of which also implies security in the E91 and BBM92 protocols.

6 Device-independent security from contextuality

Quantum key agreement is usually based on three basic assumptions: (i) that the laws of quantum mechanics are correct, (ii) that no information leaks from Alice's and Bob's laboratories (since this is potentially accessible to Eve), and (iii) Alice and Bob can accurately characterize how their devices operate. The last assumption is crucial because if Alice and Bob do not know completely how their devices work, the protocol might be compromised. For instance, in the BB84 protocol, if Alice and Bob share ququarts instead of qubits then it the scheme becomes insecure [28].

At first glance, it would seem that the control of devices is an unavoidable assumption. Remarkably, this is not the case. It is possible to prove that a cryptographic protocol secure by making no assumptions about how devices work or on what quantum states they operate. The physical basis for such device-independent security [5] lies on the fact that measurements on entangled systems provide non-local correlations, i.e., correlations not reproducible by classical shared randomness that Alice and Bob can exploit for generating a secret key. Device-independent quantum key distribution protocols face many experimental challenges but its practical implementations will certainly be more robust against technological limitations.

In the BB84 protocol, the phenomenon of information gain vs. state distrbance is mainly responsible for the secret key. However, the traditional scheme as it stands makes use of a particular set of states and measuring devices, which means in practice, the security of practical implementations of BB84 depend on how trustworthy Alice and Bob's devices will be. A rather fundamental question arises: is it possible to utilize the trade-off between information gain and disturbance for device-independent security? Karol Horodecki, et al. [34] showed that contextuality plays a major role in allowing the trade-off to be operational, leading to a quantum key distribution protocol similar in many ways to the BBM92 protocol.



σ_x^A	σ_x^B	$\sigma_x^A\sigma_x^B$
σ_y^B	σ_y^A	$\sigma_y^A\sigma_y^B$
$\sigma_x^A\sigma_y^B$	$\sigma_y^A\sigma_x^B$	$\sigma_z^A\sigma_z^B$

Figure 8: *The Peres-Mermin box. The table shows the observables for each box, where compatible observables lie on the rows and columns. A source of pairs of Peres-Mermin boxes distributes one each to Alice and Bob is also shown, where Alice measures the rows, while Bob measures the columns. The product of outcomes for rows and columns are also indicated. There is enough intrinsic (non-local) randomness in the setup to generate 0.44 bits of secret key per pair.*

6.1 Peres-Mermin boxes and the Kochen-Specker paradox

In order to use contextuality in cryptography, consider the notion of a 'box' in the Popescu-Rohrlich [29] sense—a family of probability distributions. For a given input observable, the box returns some output whose statistics is described by some distribution. In this case, the box will respect quantum mechanics; in fact, it will generate distributions concerning Kochen-Specker effects, i.e., the impossibility of jointly measuring incompatible observables. Because of this, the box is called a Kochen-Specker box. The relevant version of the Kochen-Specker theorem will be that of Peres and Mermin.

A Peres-Mermin box is a set of 6 joint probability distributions where the nine input observables $\{X_{ij}\}$ are depicted in fig. (8). In the 3×3 array, the Kochen-Specker product rule applies to the mutually commuting observables in the rows and columns:

$$\text{[rows]} \prod_{j=1}^{3} X_{ij} = 1 \text{ for } i = 1,2,3; \qquad \text{[columns]} \prod_{i=1}^{3} X_{ij} = 1 \text{ for } j = 1,2, \quad \prod_{i=1}^{3} X_{i3} = -1. \tag{65}$$

One can envision a distributed version where a source prepares pairs of Peres-Mermin boxes that are perfectly correlated, as illustrated in fig.(8). The bipartite box has the following properties:

- Local outcomes satisfy Kochen-Specker constraints in the Peres-Mermin version.
- if Alice and Bob measure the same observable, the results are perfectly correlated.

The idea now would be for Alice to measure along rows and Bob to measure along columns of their own respective Peres-Mermin box. Note that this is different situation from what is used to prove the Kochen-Specker theorem, since here the contexts for the measurement are predetermined for Alice and Bob. In addition we assume perfect correlations for corresponding observables in each of the boxes, so that identical outcomes are obtained in Alice and Bob's sides. There is also a no-signalling assumption, which says that Alice's local distributions are not influenced by Bob's choice of measurement, and vice-versa. The no-signalling condition allows us to talk meaningfully about local (i.e., marginal) probability distributions for Alice and Bob.

It should be emphasized that the distributed Peres-Mermin box necessarily exhibits quantum non-

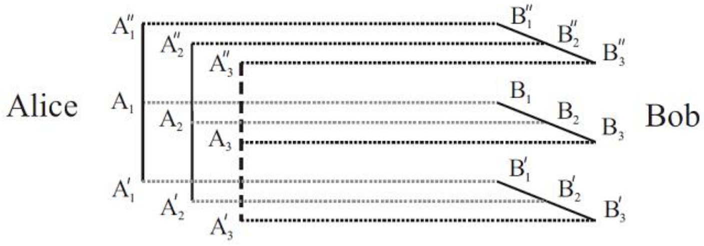


Figure 9: *The distributed Peres-Mermin box. Solid lines correspond to an even number of -1s while the dashed line corresponds to an odd number -1s. The dotted lines indicate the Alice-Bob correlations.*

locality, as would any distributed version of the Kochen-Specker paradox [30]. Indeed, what the distributed Peres-Mermin box achieves is that it translates local contextuality into non-locality [33], which is well-known to be a necessary condition for quantum cryptographic security.

Let us describe the distributed Peres-Mermin box more specifically. Formally, it consists of a family of nine conditional probability distributions $\Pr(a, b|A, B)$ where $A = 1, 2, 3$ runs over columns of Alice's box and $B = 1, 2, 3$ runs over rows of Bob's box, and $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3)$ denotes the triples of outcomes for a row/column. The family of distributions is constrained by the following conditions:

(a) (Kochen-Specker conditions) For $A = 1, 2; B = 1, 2, 3$ the product of outcomes is +1, that is,

$$a, b \in \{(+, +, +), (-, -, +), (-, +, -), (+, -, -)\}. \tag{66}$$

For $A = 3$, the outcomes multiply to -1, so

$$a[A_3] \in \{(-, -, -), (-, +, +), (+, -, +), (+, +, -)\}. \tag{67}$$

(b) (Alice-Bob correlations) Corresponding observables for Alice and Bob's boxes yield identical outcomes, $a_j = b_j, \forall j$.

(c) (No-signalling) The marginal probabilities of Alice do not depend on Bob's choice of columns to measure and the marginal probabilities of Bob do not depend on Alice's choice of rows to measure, and vice-versa:

$$\Pr(a|A, B) = \Pr(a|A), \qquad \Pr(b|A, B) = \Pr(b|B). \tag{68}$$

6.2 Intrinsic randomness and Bell inequalities for distributed Peres-Mermin boxes

The distributed box obeys quantum laws but what observables and states are involved in how it is implemented does not need to be specified—the condition for device-independent security [34]. Under such an assumption, the outcome of a fixed row or column possesses about 0.44 bits of intrinsic randomness. To make the discussion more definite, the key produced from the first row in Bob's system is considered.

Let us first demonstrate that on Bob's side, the outcome of the first rows and columns can not all have definite values. Suppose that these observables have deterministic outcomes. Using the notation indicated in fig. (9), assume without loss of generality that for the observables in question

$$B'_1, B_1, B''_1, B''_2, B''_3 = +1. \tag{69}$$

Due to perfect Alice-Bob correlations,

$$A'_1, A_1, A''_1, A''_2, A''_3 = +1. \tag{70}$$

Using both Alice-Bob correlations and the Kochen-Specker conditions, one obtains extra-strong correlations for the following set of observables:

$$a = A_2, \quad a' = A'_3, \quad b = B'_2, \quad b' = B_3. \tag{71}$$

This can be seen by noting that both constraints imply

$$A_2 = A'_2 = B'_2 = B_3, \quad A_3 = -A'_3 = -B'_3 = -B_3, \quad B_2 = B_3, \quad B'_2 = B'_3, \tag{72}$$

where since the first two observables in each equation are commuting,

$$|\langle ab\rangle + \langle ab'\rangle + \langle a'b\rangle - \langle a'b'\rangle| = 4. \tag{73}$$

Such a correlation violates the well-known Clauser-Horne-Shimony-Holt inequality, which has a value of 2 on the right-hand-side, and also the Tsirelson bound for quantum mechanics, which is $2\sqrt{2}$.

Now let us show that the first row of Bob's side can not itself have definite values. Assume that $A''_1 = A''_2 = A''_3 = +1$. Due to Kochen-Specker conditions,

$$A_1 = A'_1, \qquad A_2 = A'_2, \qquad A_3 = -A'_3 \tag{74}$$

so that the product of outcomes is correct. On the other hand, the perfect correlations between Alice and Bob imply that

$$A_i = B_i, \qquad A'_i = B'_i, \qquad A''_i = B''_i. \tag{75}$$

Therefore, the following correlations hold for the bipartite system

$$\begin{array}{ccccccccc} A_1 & = & B_1, & A_1 & = & B_2, & A_3 & = & B_3, \\ A_1 & = & B'_1, & A_2 & = & B'_2, & A_3 & = & -B'_3. \end{array} \tag{76}$$

This leads us to consider the following Bell inequality:

$$\gamma(A,B) = \langle A_1B_1\rangle + \langle A_2B_2\rangle + \langle A_3B_3\rangle + \langle A_1B'_1\rangle + \langle A_2B'_2\rangle - \langle A_3B'_3\rangle \leq 4. \tag{77}$$

Correlations (76) says that

$$\gamma(A,B) = 6. \tag{78}$$

Supposing that γ is not necessarily 6, let us derive the constraints for Bob's probability distribution for outcomes in his first row. In this regard, consider a related Bell inequality involving probabilities

$$\begin{aligned} \beta(A,B) \quad = \quad & \Pr[A_1 = B_1] + \Pr[A_2 = B_2] + \Pr[A_3 = B_3] \\ & + \Pr[A_1 = B'_1] + \Pr[A_2 = B'_2] + \Pr[A_3 \neq B'_3] \leq 5 \end{aligned} \tag{79}$$

where

$$\beta = \frac{1}{2}\left(\gamma(A,B) + 6\right). \tag{80}$$

From Alice-Bob correlations, it must be true that

$$\Pr[A_1 = B_1] = \Pr[A_2 = B_2] = \Pr[A_3 = B_3] = 1. \tag{81}$$

For the remaining three probabilities in β, define the following probabilities for outcomes in Bob's first row:

$$q_0 = \Pr(+,+,+), \quad q_1 = \Pr(+,-,-), \quad q_2 = \Pr(-,+,-), \quad q_3 = \Pr(-,-,+). \tag{82}$$

where it can be checked that $\sum_j q_j = 1$.

Consider the marginal distribution

$$p_i = \Pr[B''_i = +1]. \tag{83}$$

Then for each observable in Bob's first row, we have the following probabilities of getting +1:

$$p_1 = q_0 + q_1, \qquad p_2 = q_0 + q_2, \qquad p_3 = q_0 + q_3. \tag{84}$$

Observe that the p_is do not sum up to 1 since they represent three separate probability distributions $(p_i, 1-p_i)$ for Bob's first row observables. Thus,

$$\begin{array}{rclrcl} q_0 & = & \frac{1}{2}\left(-1+p_1+p_2+p_3\right), & q_1 & = & \frac{1}{2}\left(1+p_1-p_2-p_3\right), \\ q_2 & = & \frac{1}{2}\left(1-p_1+p_2-p_3\right), & q_3 & = & \frac{1}{2}\left(1-p_1-p_2+p_3\right). \end{array} \tag{85}$$

Not all p_is are allowed by the Kochen-Specker constraints so that $q_i \geq 0$ always. Due to Alice-Bob correlations, the marginal distribution for Alice's first row have the same probabilities:

$$\Pr[A_1 = A'_1] = p_1 = \Pr[A''_1 = +1], \quad \Pr[A_2 = A'_2] = p_2, \quad \Pr[A_3 = A'_3] = p_3. \tag{86}$$

Consider the events $X \cap Y \subset Z$ where

$$X = \{A_1 = A'_1\}, \quad Y = \{A'_1 = B'_1\}, \quad Z = \{A_1 = B'_1\}. \tag{87}$$

and using the trivial identity

$$\Pr(Z) \geq \Pr(X \cap Y) \geq \Pr(X) + \Pr(Y) - 1 \tag{88}$$

leads to

$$\Pr[A_1 = B'_1] \geq p_1 \tag{89}$$

since $\Pr(Y) = 1$ from the Alice-Bob correlations.

Similarly,

$$\Pr[A_2 = B'_2] \geq p_2, \qquad \Pr[A_3 \neq B'_3] \geq p_3. \tag{90}$$

Thus,

$$p_1 + p_2 + p_3 \leq \beta_{\text{QM}} - 3. \tag{91}$$

The above result was obtained by letting Bob's first row be $(+,+,+)$ with some non-zero probability. Doing the same for the combinations $(+,-,-), (-,+,-), (-,-,+)$:

$$\begin{array}{rcl} p_1 + (1-p_2) + (1-p_3) & \leq & \beta_{\text{QM}} - 3, \\ (1-p_1) + p_2 + (1-p_3) & \leq & \beta_{\text{QM}} - 3, \\ (1-p_1) + (1-p_2) + p_3 & \leq & \beta_{\text{QM}} - 3. \end{array} \tag{92}$$

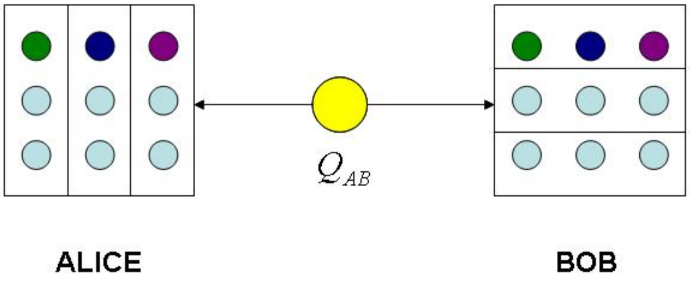


Figure 10: *Quantum key distribution using distributed Peres-Mermin boxes. Alice and Bob receives box pairs from a source. They split the boxes into three samples: one for checking Kochen-Specker constraints on individual boxes, another to check first-row correlations as indicated in the diagram. If both conditions are satisfied, the correlated outcomes in the first row of the remaining samples constitute the shared secret key.*

Using eq. (85) and eq. (80) gives

$$q_k \leq \frac{1}{2}(\beta_{\mathrm{QM}} - 4) = \frac{1}{4}(\gamma_{\mathrm{QM}} - 2). \tag{93}$$

The nontrivial quantum mechanical bound $\gamma_{\mathrm{QM}} = 5.6364$ was obtained numerically using a semi-definite program following the generalized Tsirelson approach of [31, 32].

Therefore,

$$q_k \leq \mu \approx 0.9091. \tag{94}$$

Note that $\mu \geq 1/2$ since $\gamma = 4$ is a classically achievable bound. Recall that the Kochen-Specker result for Peres-Mermin observables suggest that $\gamma = 6$ is possible. This means that the distributed Peres-Mermin box offers security by not allowing for extra-strong violations of Bell inequalities—that is, correlations stronger than what even quantum mechanics can achieve.

6.3 Secure key from distributed Peres-Mermin boxes

Suppose Alice and Bob share a bipartite box Q_{AB} which Eve might decompose into N boxes:

$$Q_{AB} = \sum_{e=0}^{3} q_e Q_{AB}^{(e)} \tag{95}$$

which is a result of a joint box Q_{ABE} where Eve hands the part Q_{AB} to Alice and Bob. This situation is analogous to giving Eve access to the purification of Alice and Bob's bipartite state.

In this case, the bipartite box will be distributed Peres-Mermin boxes, to be used in the following cryptographic protocol:

1. Alice and Bob share many distributed boxes, possibly given to them by Eve from a joint box Q_{ABE}.
2. Alice and Bob divide the boxes into three groups, two for testing purposes and one for the secret key.
3. In the first sample, Alice measures random rows of observables to check for Kochen-Specker conditions. Bob does the same thing by measuring random columns. If the outcomes they get do not correspond to a distributed Peres-Mermin box, they abort the protocol. Otherwise they proceed. (In the presence of noise, they check for a certain level of errors in the outcomes, rather than expecting all results to satisfy the product rule.)
4. From our previous analysis, Alice and Bob would like to get a key from the first row (of course, they could use any other row, or a column for that matter but Bob's first row is chosen for definiteness). They use the second sample of box pairs to check for perfect correlations (as in

fig. (10)). Note that because Alice initially had devices for measuring columns, she has to switch to a different setup for measuring rows. There are no assumptions made about devices so they might be potentially malicious. Nonetheless, Bob's first row outcomes have been shown to be secure so it is enough that Alice is able to get perfect correlations in her first row to verify that the Alice-Bob correlations hold.

5. If both Kochen-Specker conditions and Alice-Bob correlations are met, then Alice and Bob go ahead and measure the first row of the remaining box pairs, and their identical outcomes constitute a raw key.
6. Alice and Bob can then perform error correction and privacy amplification to get a shorter but more secure key. Note that error correction is not needed in the absence of noise in the system.

To estimate the secret key rate, let us consider the random variables (A, B, E) associated with the outcomes of Alice and Bob's first row observables and Eve's ensemble choice. Using the Csiszar-Korner formula

$$R_\infty \geq I(A:B) - I(B:E) \tag{96}$$

provides a lower bound for the asymptotic key rate, where $I(A:B)$ is the Shannon information between variables A, B. Since $I(A:B) = H(B) - H(B|A)$ and that in the ideal case, $H(B|A) = 0$, the lower bound on the rate depends only on how Eve splits the boxes. One can show that the distribution that gives the smallest entropy is given by

$$(q_0, q_1, q_2, q_3) = (\mu, 1-\mu, 0, 0), \qquad \mu \approx 0.9091, \tag{97}$$

which gives

$$R_\infty \geq H(B|E) \geq 0.439 \tag{98}$$

which gives the secret key rate per distributed box.

So far, only the ideal case has been considered. One can extend the analysis to the case where some noise is present, characterized an error rate ϵ. Even with imperfect distributed boxes, Kochen-Specker conditions can still be satisfied all time. This is because Alice and Bob can force it by measuring only two of three observables in a column or row, respectively, and then simply fabricating the last outcome to what it should be to obey the constraints. The noise, therefore, only affects the correlations between Alice and Bob for corresponding observables. Here Alice and Bob's outcomes on the test sample are assumed to be correlated with probability $1-\epsilon$ for each observable.

One can carefully calculate [34] the new constraints on the q_ks to be

$$q_k \leq \mu - \frac{9}{2}\epsilon = \nu, \qquad k = 0, 1, 2, 3. \tag{99}$$

With this value, the bound on $H(B|E)$ is given by

$$H(B|E) \geq \sup_{\delta > 0} \left(1 - \frac{\epsilon}{\delta}\right) h\left(x + \frac{9}{2}\epsilon\right), \tag{100}$$

and

$$H(B|A) \leq h\left(\frac{3}{2}\epsilon\right) + \frac{3}{2}\epsilon \log_2 3 \tag{101}$$

where $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function.

Inserting these values into the Csiszar-Korner formula with $\delta = 1.8$ yields

$$R_\infty > 0 \text{ for } \epsilon < 0.68\%. \tag{102}$$

This error threshold is smaller compared to usual thresholds obtained from Clauser-Korne-Shimony-Holt inequalities, which are of the order of 2%. However, this value is just a rough estimate since

$$H(B|E) = \inf_{q_i,\epsilon_i} \sum_i q_i H(B_i), \qquad \sum_i q_i \epsilon_i = \epsilon, \tag{103}$$

where $H(B_i)$ is the first row entropy of box $Q_{AB}^{(i)}$, has yet to be optimized to achieve a better key rate.

So far, only abstract distributed Peres-Mermin boxes have been considered. To apply the results, the boxes have to be realized experimentally in the lab. Indeed, this is the case: the distributed boxes can be simulated by measuring the Peres-Mermin version of Kochen-Specker observables on pairs of qubits in the maximally entangled state

$$|\psi\rangle_{AB} = |\Psi^-\rangle_{13} \otimes |\Psi^-\rangle_{24}, \tag{104}$$

where Alice receives qubits 1,2 and Bob receives qubits 3,4. The same state has been used in deriving non-locality from contextuality [33].

The above reasoning proves that the key obtained from the distributed boxes protocol is secure under individual attacks: Eve couples to each box independently and measures before Alice and Bob do classical post-processing. Horodecki et al. [34] conjecture that stronger statements can be made if one uses more advanced techniques for showing information-theoretic security.

7 Concluding remarks

The Kochen-Specker theorem states that there is no consistent way of assigning definite, non-contextual answers to a set of yes-no questions regarding an individual quantum system. Conceptually, this means that quantum theory can not be interpreted using a deterministic hidden variable model that assumes value-definiteness and non-contextuality.

The invalidation of such Kochen-Specker form of classical realism is implicitly exhibited in the security of quantum key agreement protocols, in particular, for EPR-type schemes such as E91 and BBM92. The violation of certain Bell-type inequalities in both schemes entails a certification of the Kochen-Specker theorem while simultaneously guaranteeing the secrecy of Alice and Bob's shared key.

It has also been shown that contextuality in quantum mechanics holds only when the quantum system of interest has a Hilbert space dimension of larger than 2. The BB84 protocol with chocolate balls show that qubits are not protected from value definiteness. This does not imply that protocols such as BB84 are insecure since complementarity still holds but it does indicate a fundamental difference between two-level quantum systems and higher-dimensional ones.

In quantum key distribution, two features of quantum mechanics are often employed for security: (i) the trade-off between information gain and disturbance of non-orthogonal states and (ii) quantum non-locality. It has been demonstrated that device-independent cryptography is possible by exploiting contextuality through a distributed version of the Peres-Mermin observables for the Kochen-Specker paradox. Note that any distributed version of the Kochen-Specker paradox translates local contextuality constraints into non-locality conditions. In the case of distributed Peres-Mermin boxes, security is guaranteed by showing that a part of the total system can not exhibit too strong a non-locality—which would lead to extra-strong correlations not attainable even in quantum theory.

References

[1] M. Dusek, N. Lutkenhaus, M. Hendrych, “Quantum cryptography,” *Progress in Optics* **49**, edited by E. Wolf (Elsevier, Amsterdam, 2006) 381-454. Also available as arXiv:quant-ph/0601207 (2006).

[2] C. H. Bennett and G. Brassard, “Quantum Cryptography: Public key distribution and coin tossing,” *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore* (1984) 175.

[3] A. Ekert “Quantum cryptography based on Bell’s theorem,” *Physical Review Letters* **67** (1991) 661-3.

[4] J. Barrett, L. Hardy, A. Kent “No signalling and quantum key distribution,” *Physical Review Letters* **95** (2005) 010503.

[5] S. Pironio, A. Acin, N. Brunner, N. Gisin, S. Massar, V. Scarani, “Device-independent quantum key distribution secure against collective attacks,” *New Journal of Physics* **11** (2009) 045021. Also available as arXiv:quant-ph/0903.4460 (2009).

[6] J. S. Bell, “On the problem of hidden variables in quantum mechanics,” *Reviews of Modern Physics* **38** (1966) 447-452.

[7] S. Kochen, E. Specker, “The problem of hidden variables in quantum mechanics,” *Journal of Mathematics and Mechanics* **17** (1967) 59-87.

[8] B. D’Espagnat, *Conceptual Foundations of Quantum Mechanics* 2nd ed. (Benjamin, Reading, 1976). Reprinted (Advanced Book Classics, Westview, 1999).

[9] C. Held, “The Kochen-Specker Theorem,” *Stanford Encyclopedia of Philosophy* (last revised 25 Dec 2006).

[10] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, New Jersey, 1955). Reprinted (Princeton, 1996).

[11] N. D. Mermin, “Hidden variables and the two theorems of John Bell,” *Reviews of Modern Physics* **65** 3 (1993) 803-15.

[12] A. Gleason, “Measures on the Closed Subspaces of a Hilbert Space,” *Indiana University Mathematics Journal* **6** 4 (1957) 885893.

[13] A. Cabello, J. M. Estebaranz, G. G. Alacaine, “Bell-Kochen-Specker theorem: A proof with 18 vectors,” *Physics Letters A* **212** (1996) 183-92. Also available as arXiv:quant-ph/9706009 (1997).

[14] A. Cabello, “Kochen-Specker diagram of the Peres-Mermin example” in *Fundamental Problems in Quantum Physics*, edited by M. Ferrero, A. van der Merwe (Kluwer, Dordrecht, 1995) 43-46.

[15] A. Cabello, “Experimentally testable state-independent quantum contextuality,” *Physical Review Letters***101** (2008) 210401. Also available as arXiv:quant-ph/0808.2456 (2008).

[16] O. Gunhe, M. Kleinmann, A. Cabello, J.-A. Larsson, G. Kirchmair, F. Zahringer, R. Gerritsma,C. Roos, “Compatibility and noncontextuality for sequential measurements,” *Physical Review A* **81** (2010) 022121. Also available as arXiv:quant-ph/0912.4846 (2010).

[17] K. Svozil, “Bertlmann’s chocolate balls and quantum type cryptography,” arXiv:quant-ph/0903.0231 (2009).

[18] K. Svozil, “Staging quantum cryptography with chocolate balls,” *American Journal of Physics* **74**(2006) 800-03. Also available as arXiv:physics/0510050 (2005).

[19] R. Wright, “Generalized urn models,” *Foundations of Physics* **20** (1990) 881-903.

[20] H. Bechmann-Pasquinucci, A. Peres, “Quantum cryptography with 3-state systems,” *Physical Review Letters* **85** (2000) 3313-23. Also available as arXiv:quant-ph/0001083 (2000).

[21] J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics* **1** (1964) 195-200.

[22] A. Fine, “Joint distributions, quantum correlations, and commuting observables,” *Journal of Mathematical Physics* **23** 7 (1982) 1306-1310.

[23] A. Fine, P. Teller, “Algebraic constraints on hidden variables,” *Foundations of Physics* **8** (1978) 629-36.

[24] A. Fine, J. Malley, “Noncommuting observables and local realism,” arXiv.org:quant-ph/0505016 (2005).

[25] K. Nagata, “Kochen-Specker theorem as a precondition for secure quantum key distribution,” *Physical Review A* **72** (2005) 012325.

[26] J. Malley, “All quantum observables in a hidden-variable model must commute simultaneously,” *Physical Review A* **69** (2004) 022118.

[27] C. Bennett, G. Brassard, N. D. Mermin, “Quantum cryptography without Bell’s theorem,” *Physical Review Letters* **68** (1992) 557-9.

[28] A. Acin, N. Gisin, L. Masanes, “From Bell’s Theorem to secure quantum key distribution,” *Physical Review Letters* **97** (2006) 120405. Also available as arXiv:quant-ph/0510094 (2005).

[29] S. Popescu, D. Rohrlich, “Quantum nonlocality as an axiom,” *Foundations of Physics* **24** (1994) 379-85.

[30] A. Stairs, “Quantum logic, realism, and value definiteness,” *Philosophy of Science* **50** (1983) 578-602.

[31] M. Navascues, S. Pironio, A. Acin, “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations,” *New Journal of Physics* **10** (2008) 073013. Also available as arXiv:quant-ph/0803.4290 (2008).

[32] S. Wehner, “Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities,” *Physical Review A* **73** (2006) 022110. Also available as arXiv:quant-ph/0510076 (2006).

[33] A. Cabello, “Proposal for revealing quantum nonlocality via local contextuality,” *Physical Review Letters* **104** (2010) 220401. Also available as arXiv:quant-ph/0910.5507 (2010).

[34] K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, M. Pawlowski, M. Bourennane “Contextuality offers device-independent security,” arXiv:quant-ph/1006.0468 (2010).