Memory cost for simulating all quantum correlations from the Peres–Mermin scenario

Gabriel Fagundes\textsuperscript{1,2} and Matthias Kleinmann\textsuperscript{2}

\textsuperscript{1} Departamento de Física, Universidade Federal de Minas Gerais UFMG, PO Box 702, 30123-970, Belo Horizonte, MG, Brazil
\textsuperscript{2} Department of Theoretical Physics, University of the Basque Country UPV/EHU, PO Box 644, 48080 Bilbao, Spain

E-mail: gabrielf@fisica.ufmg.br

Received 5 December 2016, revised 15 May 2017
Accepted for publication 21 June 2017
Published 17 July 2017

Abstract

Sequences of compatible quantum measurements can be contextual and any simulation with a classical model conforming with the quantum predictions needs to use internal memory. Kleinmann et al (2011 New J. Phys. \textbf{13} 113011) showed that simulating sequences from the Peres–Mermin scenario requires at least three different internal states in order not to contradict the deterministic predictions of quantum theory. We extend this analysis to probabilistic quantum predictions and ask how much memory is required to simulate correlations generated for sequences of compatible observables by any quantum state. We find that even in this comprehensive approach only three internal states are required for a perfect simulation of quantum correlations in the Peres–Mermin scenario.

Keywords: quantum information theory, foundations of quantum theory, quantum contextuality, memory cost

(Some figures may appear in colour only in the online journal)

1. Introduction

In the standard formulation of quantum theory (QT) the individual outcomes of measurements are, in general, not predetermined by the state of the system. Consequently, QT allows us to assess only the probability distribution over the measurement outcomes. Specker [1] noted that this is a fundamental property of QT and if quantum measurements had predetermined outcomes, this would imply that these values depend on the measurement context.
In this sense, QT is contextual and the mathematical formulation for this observation is the Kochen–Specker theorem [2].

Remarkable efforts have been undertaken to understand the connection between quantum contextuality and quantum information theory, for example, with respect to the advantage of quantum computing over classical computing [3, 4]. Similarly, a quantum system distributed over several parties can be used to reduce communication complexity with respect to what is possible with classical systems alone [5, 6] and the communication advantage has been identified as a resource [7–9]. A related concept is the memory cost in sequential measurements [10, 11], i.e. the memory needed to simulate the correlations occurring in sequences of quantum measurements by means of a classical automaton with memory. It has been found that the memory cost can exceed the amount of information that can be stored in the quantum system, yielding a quantum memory advantage [10–12]. We are here interested in the analysis of the memory cost with respect to quantum contextuality, i.e. sequences of compatible measurements [10]. In this strict form the question of whether there exists a quantum memory advantage due to contextuality is still open.

In this paper we investigate the situation for one of the most natural candidates for a quantum memory advantage, the Peres–Mermin square. We ask, what is the smallest memory with which a classical model can reproduce all contextuality predictions from the Peres–Mermin scenario, for any quantum state? Our focus here is to stay strictly in the regime of quantum contextuality, i.e. sequences of compatible measurements, and to take into account the probabilistic predictions of QT, while admitting the most versatile classical automaton models.

2. The Peres–Mermin square

A simple proof of the Kochen–Specker theorem was found by Peres [13] and Mermin [14]. It uses nine quantum observables arranged in the Peres–Mermin square,

\[
\begin{bmatrix}
A & B & C \\
\alpha & \beta & \gamma \\
\end{bmatrix} =
\begin{bmatrix}
\sigma_z \otimes 1 & 1 \otimes \sigma_z & \sigma_z \otimes \sigma_z \\
1 \otimes \sigma_x & \sigma_x \otimes 1 & \sigma_x \otimes \sigma_x \\
\sigma_x \otimes \sigma_x & \sigma_x \otimes \sigma_x & \sigma_x \otimes \sigma_x \\
\end{bmatrix},
\]  

(1)

where \(\sigma_x\), \(\sigma_y\), and \(\sigma_z\) are the Pauli operators. The proof of the theorem consists of the observations (i) that the operators within each row and each column form a context, i.e. they are mutually compatible, and (ii) that the condition

\[ABC = abc = \alpha\beta\gamma = A\alpha\alpha = B\beta\beta = -C\gamma\gamma = 1\]

holds. Therefore, according to QT, the expected value of the product of the outcomes of the observables in one context is always +1, with the exception \(\langle C\gamma\gamma \rangle = -1\). In order to obtain this behavior if the values of the observables are predetermined, at least one observable needs to have a context-dependent value, so that, for example, \(\gamma\) has a value of +1 in the context \(\alpha\beta\gamma\) but a value of −1 in the context \(C\gamma\gamma\).

In QT, the outcomes of all observables within a context can be obtained in a joint measurement. For the three dichotomic observables in each context of the Peres–Mermin square, joint measurement on two qubits has four distinct outcomes, taken from the set of eight possible combinations of outcomes \(((+1, +1, +1), (+1, +1, -1), \ldots, (-1, -1, -1))\). Alternatively, the outcomes can be obtained by sequentially measuring the observables in a context. This approach has been preferred in recent experiments on quantum contextuality [15–20]. When an observable \(X\) from the Peres–Mermin square is measured, then the quantum state \(\rho\) changes according to
\[ \rho \mapsto \frac{\Pi_{\delta|X}\rho\Pi_{\delta|X}}{\text{tr}(\rho\Pi_{\delta|X})}, \]

with \( \Pi_{\delta|X} = \frac{1}{2} (1 + xX) \) depending on the measurement outcome \( x = \pm 1 \) of \( X \). In a sense, sequential measurements with this Lüders transformation \([21]\) are a special way to implement a joint measurement. Since the quantum state changes according to the choice of observable and the measurement outcome, one can argue that the quantum state serves as a memory and contextual behavior is achieved due to the very presence of this memory.

However, in an extended variant of the Peres–Mermin square, it has been shown [10] that even if one takes this perspective, a classical model mimicking quantum behavior would need more than four internal states. This extended scenario uses quantum predictions for all combinations of Pauli matrices on two qubits, resulting in 15 dichotomic observables. The classical model must then reproduce the predictions from any sequence of compatible observables as well as respect conditions of compatibility and repeatability. The latter include conditions on sequences of incompatible measurements, and thus are outside the contextuality paradigm. Since the extended variant also operates on a quantum four-level system and such systems can carry at most two bits of classical information [22], this has been identified as an instance of memory advantage [10–12].

The analysis in [10] concerns classical models which reproduce the deterministic predictions of QT within a sequence. Such predictions include the product of outcomes in the sequence \( A, B, C \) being always +1 and the value of \( A \) being repeated in the sequence \( A, B, A \). For the case of the Peres–Mermin square and when any sequence of measurements consists of observables from one context, there is a classical model that is consistent with QT in this sense and only uses three internal states. This analysis does not cover the probabilistic predictions of QT—for example, that \( \langle A \rangle = 0 \) for certain quantum states—and it is not known how much memory is needed to also reproduce the probabilistic predictions of QT in the Peres–Mermin square. Since the Peres–Mermin scenario is tightly linked to contextuality, we only consider sequences of observables taken from one context. This includes predictions like \( \langle BBA \rangle = \langle A \rangle \), but excludes predictions involving incompatible observables as in \( \langle ABc \gamma \rangle = -1 \). Hence, in this paper our aim is to determine the smallest memory with which a classical model can reproduce the nondeterministic contextual quantum predictions from the Peres–Mermin scenario, for any quantum state.

### 3. Sequential correlations and stochastic automata

The outcomes of a sequence of quantum measurements may be viewed as an input–output process operating on a quantum system. The input is the choice of observable \( X \) and the output is the outcome \( x \) of the measurement of the observable. The overall probability for an output sequence \( x_1, x_2, \ldots \) for a given input sequence \( X_1, X_2, \ldots \) is \( P(x_1, x_2, \ldots | X_1, X_2, \ldots) \) and within standard QT only such correlations can be predicted.

The classical counterpart is modeled by an automaton which operates on classical memory. This memory is represented by a set \( M \) of internal memory states. In addition, the automaton has access to an external source of randomness, modeled by an external parameter \( \lambda \) which is fixed throughout a measurement sequence but randomly distributed among different sequences according to a distribution function \( p(\lambda) \). We use the model of a stochastic sequential automaton [23] where the output \( x \) and the state \( s' \in M \) after the output only depend on the input \( X \), the value of \( \lambda \), and the internal state \( s \in M \) before the output (see figure 1). The behavior of the automaton is hence summarized by the probability distribution \( p(x, s'|X, s, \lambda) \).

It represents the probability of the output \( x \) and the subsequent transition to the internal state
s', given the input $X$, the current internal state $s$ and the value of the parameter $\lambda$. Similarly, the initial state of the automaton has a distribution depending on $\lambda$, which we write as $p(s_0|\lambda)$. With this model, the correlations achieved by the automaton are

$$P(x_1, x_2, \ldots | X_1, X_2, \ldots) = \sum_{\lambda, s_0, s_1, s_2, \ldots} p(\lambda) p(s_0|\lambda) p(x_1, s_1|X_1, s_0, \lambda) \times p(x_2, s_2|X_2, s_1, \lambda) \cdots.$$  \hspace{1cm} (4)

For a given automaton, i.e. $p(x, s'|X, x, \lambda)$ and $p(s_0|\lambda)$, the choice of $p(\lambda)$ yields different correlations, so that the correlations predicted by different quantum states can be reproduced using different choices of the probability distribution $p(\lambda)$.

Clearly, it is possible to reproduce all noncontextual correlations with only one internal state, $|M| = 1$, since in this case the right-hand side of equation (4) reduces to a hidden variable model [24]. $\sum_{\lambda} p(\lambda) p(x_1|X_1, \lambda) p(x_2|X_2, \lambda) \cdots$. The external parameter $\lambda$ is not always included in such an analysis—see, for example, the $\epsilon$-transducers studied in [25]. However, even noncontextual scenarios could then require memory, since, for example, measuring the sequence $\sigma_x, \sigma_x$ on an eigenstate of $\sigma_z$ gives a random outcome for the first measurement, but the second measurement has to repeat the value of the first measurement. Consequently, if $\lambda$ does not occur, the simulation requires two internal states, while when $\lambda$ can take two values, no memory is required. The automaton is allowed to be intrinsically random, i.e. the distributions $p(x, s'|X, x, \lambda)$ and $p(s_0|\lambda)$ may be nondeterministic. As evident from our analysis below, this intrinsic randomness is not required for simulating quantum correlations from the Peres–Mermin scenario.

4. A memory-optimal automaton for the Peres–Mermin scenario

As explained above, quantum contextuality is a feature of sets of compatible observables and we therefore only consider sequences of measurements where all observables are taken from one context. That is, the observables are taken from either one row or one column of
the Peres–Mermin square. Our first concern is the simulation of quantum measurements of a single sequence of compatible observables. According to QT, certain events can never occur; examples are the output \(+1, −1\) in the sequence \(A, A\) and the output \(+1, +1, −1\) in the sequence \(A, B, C\). In [10], it has been shown that any automaton which obeys all such quantum predictions must have memory with at least three internal states. An explicit example of such an automaton is given by [10]:

\[
\begin{align*}
\alpha_1 &= \begin{bmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & +1 & +1 \\
\end{bmatrix}, \\
\alpha_2 &= \begin{bmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & +1 & +1 \\
\end{bmatrix}, \\
\alpha_3 &= \begin{bmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & +1 & +1 \\
\end{bmatrix}, \\
\end{align*}
\]

\[
t_1 = \begin{bmatrix}
1 & 1 & 2 \\
1 & 1 & 3 \\
1 & 1 & 1 \\
\end{bmatrix}, \\
t_2 = \begin{bmatrix}
2 & 1 & 2 \\
2 & 2 & 2 \\
2 & 3 & 2 \\
\end{bmatrix}, \\
t_3 = \begin{bmatrix}
3 & 3 & 3 \\
1 & 3 & 3 \\
2 & 3 & 3 \\
\end{bmatrix}.
\]

This notation is supposed to be read as follows. Each matrix \(\alpha_i, i \in M\), represents the deterministic output for each of the three internal states \(M = \{1, 2, 3\}\). Similarly, the transition matrices \(t_i\) represent the internal state after the output. In terms of equation (4), the distribution \(p(x, s'|X, s)\) is 1 if the entry in the output matrix \(\alpha_i\) at the position of the observable \(X\) is \(x\) and the entry in the transition matrix \(t_i\) in the same position is \(s'\); the distribution is 0 otherwise. Here, \(x \in \{+1, −1\}\), \(s, s' \in M = \{1, 2, 3\}\), and \(X \in \{A, B, C, a, b, c, α, β, γ\}\). For example, if the automaton is in state \(s = 1\) and we provide the observable \(C\) as input, then the measurement outcome is \(x = +1\) and the automaton changes to the state \(s' = 2\). It is straightforward to verify that this automaton obeys all deterministic predictions of QT for any sequence of compatible observables [10] and for any initial state \(s_0\).

However, no quantum state gives deterministic predictions for all nine observables in the Peres–Mermin square, because these observables are not all mutually compatible and no common eigenstate can exist. In the following we extend the automaton from equation (5) to use the external parameter \(λ\), so that a statistical mixture \(p(λ)\) can reproduce the quantum predictions.

### 4.1. Other valid automata

Starting from the automaton in equation (5), there are several transformations which lead to other automata with the same properties. First, it is possible to flip the signs for the output under the constraint that for each row and each column either there is no flip of signs or there are exactly two flips of signs. This generates 15 additional automata. Second, it is possible to make any permutation of the rows or a permutation of the first and second columns. We restrict ourselves to three permutations of rows, which leave one row unchanged, and to the permutation of the first and second columns. This yields four additional automata, which when combined with the first set of transformations add up to \(16 \times 5 = 80\) automata. In addition, we are free to choose the initial state \(s_0\) and get in this way 240 different behaviors. We combine all these behaviors into a single automaton by allowing 240 different values for \(λ\), i.e. the value of \(λ\) determines the behavior of the automaton.
4.2. Example: singlet state

As an example, we reproduce all quantum correlations for the singlet state (the quantum state yielding $\langle C \rangle = \langle c \rangle = \langle \gamma \rangle = -1$) by choosing a distribution $p(\lambda)$ for $\lambda = 1, 2, \ldots, 240$. We choose $p(\lambda) = \frac{1}{4}$ if $\lambda \in \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ and $p(\lambda) = 0$ otherwise. For $\lambda_k$, $k = 1, 2, 3, 4$, the transition matrices $t^{(k)}_1, t^{(k)}_2, t^{(k)}_3$ are the same as in equation (5) and the outcome matrices $o^{(k)}_i$ are given by

\begin{align*}
  o^{(1)}_1 &= \begin{bmatrix} -1 & +1 & -1 \\ +1 & -1 & -1 \end{bmatrix}, \\  o^{(1)}_2 &= \begin{bmatrix} -1 & +1 & -1 \\ +1 & -1 & -1 \end{bmatrix}, \\  o^{(1)}_3 &= \begin{bmatrix} -1 & -1 & +1 \\ -1 & +1 & -1 \end{bmatrix}, \\
\end{align*}

(6a)

\begin{align*}
  o^{(2)}_1 &= \begin{bmatrix} -1 & +1 & -1 \\ +1 & +1 & +1 \\ -1 & +1 & -1 \end{bmatrix}, \\  o^{(2)}_2 &= \begin{bmatrix} -1 & +1 & -1 \\ +1 & -1 & -1 \end{bmatrix}, \\  o^{(2)}_3 &= \begin{bmatrix} -1 & -1 & +1 \\ +1 & +1 & +1 \end{bmatrix}. \\
\end{align*}

(6b)

\begin{align*}
  o^{(3)}_1 &= \begin{bmatrix} +1 & -1 & -1 \\ -1 & +1 & -1 \\ -1 & +1 & -1 \end{bmatrix}, \\  o^{(3)}_2 &= \begin{bmatrix} +1 & -1 & -1 \\ +1 & -1 & -1 \end{bmatrix}, \\  o^{(3)}_3 &= \begin{bmatrix} -1 & -1 & +1 \\ +1 & -1 & -1 \end{bmatrix}. \\
\end{align*}

(6c)

\begin{align*}
  o^{(4)}_1 &= \begin{bmatrix} +1 & -1 & -1 \\ +1 & +1 & +1 \\ +1 & -1 & -1 \end{bmatrix}, \\  o^{(4)}_2 &= \begin{bmatrix} +1 & -1 & -1 \\ -1 & +1 & -1 \end{bmatrix}, \\  o^{(4)}_3 &= \begin{bmatrix} +1 & +1 & +1 \\ +1 & +1 & +1 \end{bmatrix}. \\
\end{align*}

(6d)

The initial state for all four cases is $s_0 = 2$, i.e. we have $p(s_0|\lambda) = 1$ if $s_0 = 2$ and 0 otherwise. In principle one can now verify that for sequences of compatible observables, all quantum correlations from the singlet state are indeed reproduced. However, there is an infinite number of input sequences which needs to be considered and it is our next step to reduce the number of sequences to a finite set.
4.3. A finite set of sufficient input sequences

We show in this section that a finite number of input sequences suffices to allow us to determine all correlations for all sequences. Since we only consider sequences of observables from one context, as soon as two different observables occur in a sequence, it is already possible to predict the remainder of the sequence from the outcome of these two observables. This is because the product of the outcomes of the three observables of each context is always +1 or −1, depending on the context, and due to the requirement that repeated occurrences of an observable in a sequence produce repeated values.

Hence, it remains to consider sequences where initially one observable is measured repeatedly, for example, $X, X, Y$. In quantum mechanics we have

$$P(x, x, \ldots, x, y|X, \ldots, X, Y) = \text{tr}(\Pi_1|Y\Pi_2|Y\rho\Pi_3|Y|Y) = P(x, y|X, Y),$$

for any number of repetitions of the input $X$ and output $x$. However, for the automaton model we could have different values for outcome $y$ in the sequences $X, X, \ldots, X, Y$, depending on the number of repetitions of $x$, since the value of $Y$ does not need to be fixed until $Y$ is actually measured. Thus, we have to consider how our specific model behaves in this situation. For any value of $\lambda$, the behavior of our automaton is analogous to that of the automaton in equation (5) and for this automaton one observes that the internal state $s'$ after an $\ell$-fold measurement of $X$ does not depend on $\ell$, if $\ell \geq 1$. Hence, the outcome of $X, X, \ldots, X, Y$ is $x, x, \ldots, x, y$ if and only if $X, Y$ has outcome $x, y$.

In summary, our automaton with any choice of $p(\lambda)$ reproduces the quantum correlations for a state $\rho$ for all sequences of compatible observables if and only if it does so for all sequences of a length of two. For practical reasons, instead of dealing with the correlations $P(x, y|X, Y)$ we use the equivalent set of expectation values

$$\langle X \rangle = \sum_{x,y} xP(x, y|X, Y),$$

$$\langle XY \rangle = \sum_{x,y} yP(x, y|X, Y),$$

$$\langle Y \rangle = \sum_{x,y} xyP(x, y|X, Y),$$

where in the second equation the values of $X$ in the first and in the last position is the same. Note that while in QT we always have $\langle XY \rangle = \langle Y \rangle$, this does not hold for all ensembles $p(\lambda)$ in our automaton; for example, in equation (5) with the initial internal state $s = 1$, we have $\langle c \rangle = 1$, but $\langle CcC \rangle = -1$. However, we observe that $\langle XY \rangle = \langle YX \rangle$ for all $p(\lambda)$ and all compatible $X$ and $Y$, a relation that also holds in QT for any state.

Therefore, we have to take into account 9 values $\langle X \rangle$, 18 values $\langle XY \rangle$, and 36 values $\langle XYX \rangle$. We enumerate these values by $j = 1, \ldots, 63$ and collect for each $j$ the values for all 240 values of $\lambda$ in a vector $\vec{q}_j$. Then the expectation values $\vec{q} = (q_1, \ldots, q_{63})$ can be achieved if and only if $q_j = \vec{v}_j \cdot \vec{p}$ for some probabilities $\vec{p}$ with $p_\lambda \equiv p(\lambda)$. The set of achievable expectation values $\vec{q}$ is hence given by the polytope

$$P = \{ \vec{q}| q_j = \vec{v}_j \cdot \vec{p} \text{ for all } j \text{ and some } \vec{p} \}. \quad (9)$$
Similarly, for the quantum correlations we have 63 Hermitian operators $Z_j$, such that the expectation values $q_j$ can be obtained according to QT if and only if $q_j = \text{tr}(\rho Z_j)$ for all $j$ and some quantum state $\rho$. The set of achievable expectation values $\{q_j\}$ according to QT is consequently the convex set

$$Q = \{q_j|q_j = \text{tr}(\rho Z_j)\text{ for all }j\text{ and some }\rho\}. \quad (10)$$

This allows us to easily verify the correctness of the example in section 4.2, by comparing $\vec{v}_j \cdot \vec{p}$ with $\text{tr}(\rho Z_j)$ for all $j$ and for any quantum state $\rho$; finding a corresponding distribution $p(\lambda)$ reduces to finding probabilities $\vec{p}$ with $\vec{v}_j \cdot \vec{p} = \text{tr}(\rho Z_j)$ for all $j$. This can be solved by means of linear programming, which is in fact our method to find $p(\lambda)$ for the singlet state in section 4.2.

4.4. Simulation of the correlations of any quantum state

We are now equipped with the necessary tools to prove that the correlations of any quantum state can be simulated with a construction analogous to the one in section 4.2. According to our previous analysis, the question of whether quantum predictions can be simulated by an appropriate distribution $p(\lambda)$ reduces to the question of whether the convex set $Q$ is contained in the polytope $P$. In order to make this question tractable, we use an equivalent representation of the polytope, which is written as a finite intersection of half-spaces \cite{26} parametrized by vectors $\vec{h}_\ell$ and numbers $\alpha_\ell$, so that

$$P = \{\vec{q}|\vec{h}_\ell \cdot \vec{q} \leq \alpha_\ell \text{ for all }\ell\}. \quad (11)$$

According to this half-space representation, $P$ contains $Q$ if and only if $\vec{h}_\ell \cdot \text{tr}(\rho Z) \leq \alpha_\ell$ for all $\ell$ and all $\rho$. By writing

$$W_\ell = \alpha_\ell \mathbb{1} - \vec{h}_\ell \cdot \vec{Z}, \quad (12)$$

we further simplify this to $\text{tr}(\rho W_\ell) \geq 0$ for all $\ell$ and all $\rho$. That is, $Q \subset P$ holds if and only if all $W_\ell$ are positive semidefinite. Conversely, if we find a state with $\text{tr}(\rho W_\ell) < 0$ for some $\ell$, and hence $W_\ell$ is not positive semidefinite, then our automaton cannot simulate all quantum predictions for this state.

In principle, this can be tested directly. However, since the polytope $P$ is given in the form of equation (9), we need to compute the half-space representation in equation (11). This can be achieved by using Fourier–Motzkin elimination, but it is known to be a computationally hard task and for our problem we are not able to find a direct solution. The central observation to solve the problem nonetheless is that $Q$ spans a rather low-dimensional affine space. In particular, $Q$ is contained in the affine space $\vec{a} + U \equiv \{\vec{a} + \vec{u}|\vec{u} \in U\}$, where $a_j = \text{tr}(\rho Z_j)$ for some fixed $\rho_0$ (for example, $\rho_0 = \frac{1}{2}\mathbb{1}$) and $U$ is a linear space $U = \{\vec{u}|\vec{u} = \text{tr}(GZ_j)\text{ for some }G\}$ with $G$ denoting any Hermitian operator obeying $\text{tr}(\rho_0 G) = 0$. This holds true since we can always write $\rho = \rho_0 + G$ for some $G$. The dimension of the linear space $U$ is only $\text{dim }U = 9$, as can be found by using the linear independence relations of the operators $Z_j$. Therefore, $Q \subset P$ is equivalent to $Q \subset P \cap (\vec{a} + U)$ and our problem reduces to calculate a half-space representation for the polytope $P \cap (\vec{a} + U)$. This problem is easily tractable, as we discuss in appendix. We obtain 24 nonzero operators $W_\ell$, each of which is positive semidefinite. This proves $Q \subset P$ and thus our automaton can simulate the quantum correlations for any quantum state. We mention that the nonzero operators $W_\ell$ are, up to an arbitrary positive factor, exactly those 24 projectors of unit rank which commute with all observables from one of the six contexts in the Peres–Mermin square.
5. Conclusions

Quantum contextuality is considered as one of the key differences between the microscopic world and the world governed by classical mechanics. Recent experimental demonstrations of this phenomenon proceed by measuring sequences of observables and yield a contradiction to the assumption of noncontextuality, i.e. the assumption that the value of an observable does not depend on which other compatible observables are measured alongside it. We revisit this conclusion for the case of the Peres–Mermin scenario using classical models which utilize internal memory in order to reproduce the quantum behavior. We show that for this scenario an automaton using only three internal states can reproduce the quantum correlations from any quantum state for any sequence of compatible observables. This model is also optimal, since a lower bound of three internal states has already been established [10]. The memory cost of the Peres–Mermin scenario is therefore actually lower than that of canonical quantum implementation, which requires two qubits. Since for quantum correlations involving sequences of incompatible observables the memory cost can also be greater than the memory of the quantum system, this leaves open the question of whether there can be a quantum memory advantage when restricted to sequences of compatible observables and if so, which contextuality scenario this occurs in.

Acknowledgments

We thank Costantino Budroni, Adán Cabello, Marcelo Terra Cunha, Jan-Åke Larsson, Marco Túlio Quintino, and Géza Tóth for their discussions. This work was supported by CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil, by the FQXi Large Grant ’The Observer Observed: A Bayesian Route to the Reconstruction of Quantum Theory’, by the EU (ERC Starting Grant GEDENTQOPT), and by the DFG (Forschungsstipendium KL 2726/2–1).

Appendix. Low-dimensional section of a polyhedral cone

A central step in section 4.4 is to compute the half-space representation of the polytope $P \cap (\vec{a} + U)$, where $P$ is a polytope, $\vec{a} \in P$ is a vector and $U$ is a linear subspace of low dimensions.

We first consider the equivalent problem for a polyhedral cone $P = \left\{ A\vec{r} \geq 0 \right\}$, where $A\vec{r} \geq 0$ abbreviates $r_k \geq 0$ for all $k$ and $A$ is some matrix with real entries. For a matrix $K$, let $F$ be a matrix the range of which is the kernel of $KA$. We have

\[
P \cap \ker(K) = \left\{ A\vec{r} \geq 0, \ K(A\vec{r}) = 0 \right\} \\
= \left\{ A\vec{r} \geq 0, \ \vec{r} = F\vec{s} \text{ for some } \vec{s} \right\} \\
= \left\{ AF\vec{s} \geq 0 \right\} \\
= AF\vec{s} \geq 0, \quad (A.1)
\]

where $KA\vec{r} = 0$ implies $\vec{r} = F\vec{s}$ for some $\vec{s}$ and, conversely, $(KA)F\vec{s} = 0$ for any $\vec{s}$. It follows that if we can obtain a matrix $F'$, such that $\left\{ F'\vec{s} \geq 0 \right\} = \left\{ \vec{s} | F\vec{s} \geq 0 \right\}$, then $P \cap \ker(K) = \left\{ AFF'\vec{s} \geq 0 \right\}$.

For our case, we extend the polytope $P$ from equation (9) to a polyhedral cone $P$ by adding $\vec{e} = (1, 1, \ldots, 1)$ to the vectors $\vec{v}_i$ and by dropping the constraint $\sum p_i = 1$, i.e. $P = \left\{ A\vec{r} \geq 0 \right\}$, where $A$ is the matrix with rows $[e, v_1, \ldots, v_{63}]$. Then $(1, \vec{q}) \in P$ if and only
if \( \vec{q} \in P \). Similarly, we define the linear subspace \( \mathcal{U} = \{ (\lambda, \lambda \vec{a} + \vec{u}) | \lambda \in \mathbb{R} \text{ and } \vec{u} \in U \} \), so that \((1, \vec{x}) \in \mathcal{U}\) is equivalent to \(\vec{x} \in \vec{a} + U\).

In order to apply equation (A.1), we choose some matrix \(K\) such that \(\ker(K) = \mathcal{U}\) and some matrix \(F\) with range \(\ker(KA)\). Despite \(F^T\) being a larger matrix than \(A\), we find that \(F'\) is rather easy to compute. The matrix \(B = A F F'\) is then only of rank \(\dim(\mathcal{U}) = 10\) and a matrix \(B'\) with \(\{B' y \geq 0\} = \{\vec{y} B' y \geq 0\}\) can be computed at an instance. We use the software cddlib [27] to generate the matrices \(F'\) and \(B'\) and iml [28] to compute \(K\) and \(F\). Both packages work with unlimited exact integer arithmetic and hence our computation of \(B'\) is exact. We independently verify our results by using para [29] to compute \(K, F, F'\) and \(B'\).

Finally, we have \(\vec{q} \in P\) and \(\vec{q} \in \vec{a} + U\) if and only if \((1, \vec{q}) \in P \cap \mathcal{U}\), i.e. if and only if \(B'_{\ell,1} + \sum_j B'_{\ell,j+1} a_j \geq 0\) for all \(\ell\). Therefore, the operators \(W_\ell\) defined in equation (12) are given by

\[
W_\ell = B'_{\ell,1} 1 - \sum_j B'_{\ell,j+1} Z_j.
\]

As we show in the main text, \(Q \subset P \cap (\vec{a} + U)\) is equivalent to all \(W_\ell\) being positive semidefinite. In our analysis, all operators \(W_\ell\) satisfy this condition.

References

[1] Specker E 1960 Die Logik nicht gleichzeitiger entscheidbarer Aussagen Dialectica 14 239–46
[2] Kochen S and Specker E P 1967 The problem of hidden variables in quantum mechanics J. Math. Mech. 17 59–87
[3] Raussendorf R 2013 Contextuality in measurement-based quantum computation Phys. Rev. A 88 022322
[4] Howard M, Wallman J, Veitch V and Emerson J 2014 Contextuality supplies the ‘magic’ for quantum computation Nature 510 351–5
[5] Toner B F and Bacon D 2003 Communication cost of simulating Bell correlations Phys. Rev. Lett. 91 187904
[6] Buhrman H, Cleve R, Massar S and de Wolf R 2010 Nonlocality and communication complexity Rev. Mod. Phys. 82 665–98
[7] Gallego R, Wrulinger L E, Acín A and Navascués M 2012 Operational framework for nonlocality Phys. Rev. Lett. 109 070401
[8] de Vicente J I 2014 On nonlocality as a resource theory and nonlocality measures J. Phys. A: Math. Theor. 47 424017
[9] Barrett J, Linden N, Massar S, Pironio S, Popescu S and Roberts D 2005 Nonlocal correlations as an information-theoretic resource Phys. Rev. A 71 022101
[10] Kleinmann M, Gühne O, Portillo J R, Larsson J-A and Cabello A 2011 Memory cost of quantum contextuality New J. Phys. 13 113011
[11] Brierley S, Kosowski A, Markiewicz M, Paterek T and Przyśiężna A 2015 Nonclassicality of temporal correlations Phys. Rev. Lett. 115 120404
[12] Galván E F and Hardy L 2003 Substituting a qubit for an arbitrarily large number of classical bits Phys. Rev. Lett. 90 087902
[13] Peres A 1990 Incompatible results of quantum measurements Phys. Lett. A 151 107–8
[14] Mermin N D 1990 Simple unified form for the major no-hidden-variables theorems Phys. Rev. Lett. 65 3373–6
[15] Kirchmair G, Zähringer F, Gerritsma R, Kleinmann M, Gühne O, Cabello A, Blatt R and Roos C F 2009 State-independent experimental test of quantum contextuality Nature 460 494–7
[16] Amselem E, Ràdmark M, Bourennane M and Cabello A 2009 State-independent quantum contextuality with single photons Phys. Rev. Lett. 103 160405
[17] Lapkiewicz R, Li P, Schueff C, Langford N K, Ramelow S, Wieśniak M and Zeilinger A 2011 Experimental non-classicality of an indivisible quantum system Nature 474 490–3
[18] Zhang X, Um M, Zhang J, An S, Wang Y, Deng D-L, Shen C, Duan L-M and Kim K 2013 State-independent experimental test of quantum contextuality with a single trapped ion Phys. Rev. Lett. 110 070401
[19] D’Ambrosio V, Herbauts I, Amselem E, Nagali E, Bourennane M, Sciarrino F and Cabello A 2013 Experimental implementation of a Kochen–Specker set of quantum tests Phys. Rev. X 3 011012
[20] Jerger M 2016 Contextuality without nonlocality in a superconducting quantum system Nat. Commun. 7 12930
[21] Lüders G 1951 Über die Zustandsänderung durch den Meßprozeß Ann. Phys. 443 322–8
[22] Holevo A S 1973 Bounds for the quantity of information transmitted by a quantum communication channel Probl. Inf. Trans. 9 177–83
[23] Paz A 1971 Introduction to Probabilistic Automata (New York: Academic)
[24] Bell J S 1964 On the Einstein Podolsky Rosen paradox Physics 1 195–200
[25] Barnett N and Crutchfield J P 2015 Computational mechanics of input–output processes: structured transformations and the ε-transducer J. Stat. Phys. 161 404–51
[26] Grünbaum B 2003 Convex Polytopes 2nd edn (New York: Springer)
[27] Fukuda K 2015 cddlib v094h https://inf.ethz.ch/personal/fukudak/cdd_home/
[28] Chen Z, Storjohann A and Fletcher C 2015 Integer matrix library v1.0.5 https://cs.uwaterloo.ca/~astorjoh/iml.html
[29] Christof T and Löbel A 2009 Polyhedron representation transformation algorithm v1.4.1 http://porta.zib.de/