Lower bound on the dimension of a quantum system given measured data

Stephanie Wehner,[1] Matthias Christandl,[2] and Andrew C. Doherty[3]
[1]*Institute for Quantum Information, California Institute of Technology, 1200 E California Boulevard, Pasadena, California 91125, USA*
[2]*Arnold Sommerfeld Center for Theoretical Physics, Faculty of Physics, Ludwig-Maximilians-University Munich, Theresienstrasse 37, 80333 Munich, Germany*
[3]*School of Physical Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia*
(Received 29 August 2008; published 22 December 2008)

We imagine an experiment on an unknown quantum mechanical system in which the system is prepared in various ways and a range of measurements are performed. For each measurement M and preparation ρ the experimenter can determine, given enough time, the probability of a given outcome a: $p(a|M,\rho)$. How large does the Hilbert space of the quantum system have to be in order to allow us to find density matrices and measurement operators that will reproduce the given probability distribution? In this paper, we prove a simple lower bound for the dimension of the Hilbert space. The main insight is to relate this problem to the construction of quantum random access codes, for which interesting bounds on the Hilbert space dimension already exist. We discuss several applications of our result to hidden-variable or ontological models, to Bell inequalities, and to properties of the smooth min-entropy.

DOI: 10.1103/PhysRevA.78.062112 PACS number(s): 03.65.Ud, 03.65.Wj

I. INTRODUCTION

Loosely speaking, the dimension of the Hilbert space describing a quantum mechanical system limits the complexity or usefulness of the correlations that can be generated by experiments on the system. For example, it has been suggested that the primary resource for quantum computation is the Hilbert space dimension [1]. In practice though, when an experimentalist is faced with a real physical system, the dimension of the Hilbert space is often infinitely large in principle. The dimension of the Hilbert space that we use to describe the system of interest usually depends on the approximation used to describe the physics of the system and may well depend on how well the experiment has in fact been set up. For this reason it would be of interest to be able to use the correlations observed in experiment to find strict lower bounds on the dimension of the Hilbert space. Thus one could conclude based on experimental data that the Hilbert space dimension of some system of interest was necessarily large and that the system could not be effectively approximated by a smaller one. In this paper we show that it is certainly possible to derive very general lower bounds on Hilbert space dimension given experimental data.

The properties of quantum correlations have been best studied in the setting of the Bell experiment. Imagine two parties, Alice and Bob, who are given access to shared quantum states $|\Psi\rangle_{AB}$, but cannot communicate. Each of them now performs a randomly chosen measurement on $|\Psi\rangle_{AB}$ and records their measurement outcome. In order to obtain an accurate estimate for the correlation between their choice of measurement settings and measurement outcomes, Alice and Bob now perform this experiment many times, using an identically prepared state $|\Psi\rangle_{AB}$ in each round. Quantum mechanics imposes strict limits on the strength of such nonlocal correlations, and it has been shown that we can compute bounds on these correlations for any such experiment [2–4]. (These bounds generalize the well known Tsirelson inequalities [5,6] that apply to conventional Bell experiments that test the Clauser-Horne-Shimony-Holt inequality.) In particular, if we let $p(a,b|s,t)$ be the probability that Alice and Bob obtain measurement outcomes $a \in A$ and $b \in B$ when performing measurements indexed by $s \in S$ and $t \in T$, we can test using the methods of [2–4] whether there exists a shared state $|\Psi\rangle$ and measurement operators M_s^a and M_t^b for Alice and Bob such that

$$p(a,b|s,t) = \langle \Psi | M_s^a \otimes M_t^b | \Psi \rangle$$

for all a, b, s, and t. But how large does the dimension of the Hilbert space have to be such that we can find such a state and measurements?

Unfortunately, the methods of [2–4] do not give us any bound on the dimension in general. It is known that in the special case of two-party correlations, where Alice and Bob perform measurements using observables with eigenvalues ± 1 (also known as XOR games with $A=B=\{0,1\}$), the dimension of the entangled state does not need to be larger than $d=2^n$, where $n=\min(|S|,|T|)$ [5,6]. Results are also known for certain sets of two-outcome inequalities [7,8]. Very little is known otherwise. Even though one can construct an inequality with an infinite number of settings that requires an infinitely large entangled state [9], it is unknown whether there exist general experiments with a *finite* number of measurement settings for which an infinitely large entangled state is required to obtain the maximum possible quantum violation exactly.

In the context of bipartite Bell experiments, the question of determining the Hilbert space dimension from experimental data has been addressed in the recent work of Brunner *et al.* [10]. Their aim was to reproduce the statistics of an experiment performed by two separated observers on a single preparation of a bipartite quantum system. They introduce the concept of dimension witness, which is a certain kind of generalization of Bell inequalities that makes it possible to distinguish the strength of correlations that can be obtained in different dimensions. This very nice approach makes it possible to find interesting lower bounds on the dimension of the system in use and has recently been extended by Briët

et al. [11] for XOR games. Our work finds rather different bounds on the Hilbert space dimension that are obtained by a very different method. The bounds apply to quantum mechanical systems with any number of parties (even one), and apply also to the case where the experimental data refer to an arbitrary number of preparations of the system. Our bound is of particular significance if the number of measurement outcomes for each party is small.

The general problem we consider is this. Suppose we are given a set of preparations $\mathcal{S}$ of a given quantum system and a set of measurements $\mathcal{M}$, each of which has outcomes $a \in \mathcal{A}$.[1] We are given, perhaps as a result of experiments, probabilities $p(a|j,r)$ of obtaining outcome a when performing the measurement $M_j \in \mathcal{M}$ having prepared the system in state $r \in \mathcal{S}$. However, we do not know either an explicit density matrix ρ for the preparation r or a measurement operator M_j^a such that

$$p(a|j,r) = \mathrm{Tr}(\rho M_j^a) = p(a|M_j,\rho),$$

where we use M^a to denote the measurement operator corresponding to outcome a of measurement j, and will simply write $p(a|j,\rho) := p(a|j,r)$ from now on. How large does the dimension of the Hilbert space supporting the states ρ have to be?

This question was recently raised in [12,13], which determined the number of (hidden) variables in an ontological model necessary to reproduce the probabilities $p(a|j,r)$. In particular, it was shown that, if each measurement has only two outcomes, then for a particular ontological model the number of hidden variables must be greater than $N = \min(|\mathcal{S}|, 2^{|\mathcal{M}|})$. Here, we prove a simple lower bound that shows that in the quantum setting the dimension of our space scales as $2^{c \log_2 N}$, where c is a constant depending on the probabilities above. Thus, if the number of states $|\mathcal{S}|$ and the number of measurements $|\mathcal{M}|$ is large, the dimension of the quantum state that we need cannot be significantly smaller.

In the following, we first prove a simple lower bound for this general problem. We then examine how we can use this to lower bound the dimension of the entangled state in a Bell experiment, and provide a simple example. In the Appendix, we show that this example disproves that the smooth min-entropy is additive and that we can perform exact min-entropy splitting as for independent states, which is of interest in the noisy-quantum-storage model [14–16].

Throughout this paper, we use $h(p) := -p \log_2 p - (1-p)\log_2(1-p)$ to denote the binary entropy. We furthermore use $\mathcal{S}(\mathcal{H})$ to denote the set of all quantum states on the Hilbert space $\mathcal{H}$, and write $H(\rho) := -\mathrm{Tr}(\rho \log_2 \rho)$ for the *von Neumann entropy* of a state $\rho \in \mathcal{S}(\mathcal{H})$. Note that, if ρ is classical, this reduces to the Shannon entropy, and that $\log_2[\dim(\mathcal{H})] \geq H(\rho) \geq 0$ [17], Theorem 11.8, since we may equivalently write $H(\rho) = -\Sigma_j \lambda_j \log_2 \lambda_j$ where λ_j is the jth eigenvalue of ρ. We will also need the concept of a *cq-state* $\rho^{XQ} \in \mathcal{S}(\mathcal{H}^X \otimes \mathcal{H}^Q)$, a state that is part classical, part quantum, of the form

$$\rho^{XQ} = \sum_{x \in \mathcal{X}} P_X(x)|x\rangle\langle x| \otimes \rho_x^Q,$$

where P_X is a probability distribution over $\mathcal{X}$ and $\langle x|x'\rangle = 0$ for $x \neq x'$. Let $\rho^X = \mathrm{Tr}_Q(\rho^{XQ})$ and $\rho^Q = \mathrm{Tr}_X(\rho^{XQ})$ be the reduced states on systems $\mathcal{H}^X$ and $\mathcal{H}^Q$, respectively. The *conditional von Neumann entropy* is defined as $H(X|Q) := H(\rho^{XQ}) - H(\rho^Q)$. We will also use the shorthands $H(X) := H(\rho^X)$ and $H(XQ) := H(\rho^{XQ})$ and $[\ell] := \{1, \ldots, \ell\}$.

II. LOWER BOUND

We first state the intuition behind our simple lower bound, based on quantum random access codes. A quantum (m,q,p) random access code is an encoding of an m-bit string x into a q-qubit state ρ_x such that for any $i \in [m]$ we can retrieve the bit x_i from ρ_x with probability p. Note that we are interested only in retrieving a single bit of the original string x from ρ_x. In general, it is unlikely that we will be able to retrieve more than a single bit. For such encodings it is not hard to prove a lower bound on the number of qubits q [18] if the distribution over the strings is uniform and the probability of decoding each bit is the same.

Now note that our problem has a very similar flavor. Suppose we were given states $\rho_1, \ldots, \rho_\ell$ and measurements $M_1, \ldots, M_m$ that give us the desired probabilities $p(a|M_j,\rho_x)$. For simplicity, assume for now that $\ell = 2^m$ and $a \in \{0,1\}$. Then the states $\rho_1, \ldots, \rho_\ell$ form a generalized quantum random access code, where each state represents an encoding of an m-bit string x and we think of M_j as the measurement that we can apply to extract bit x_j with probability $p(a|M_j,\rho_x)$. Once we realize this viewpoint it is indeed very intuitive that we should be able to apply techniques similar to the ones used for quantum random access codes also in the present setting.

A. Tools

We first state a general lemma from which our bound later follows by constructing an appropriate mapping that associates a string x with a state ρ_x. Our proof is a straightforward extension of the techniques employed for the random access code lower bound [18–20] to more generalized distributions and alphabets.

Lemma 1. Let $\mathcal{X} = \mathcal{A}^{\times m}$ denote the set of strings of length m, let P_X be a probability distribution over $\mathcal{X}$, and let $X = X_1, \ldots, X_m$ denote a random variable chosen from $\mathcal{X}$ according to the distribution P_X. Let $\mathcal{H}^Q$ be a Hilbert space supporting an ensemble of states $\{P_X(x), \rho_x | \rho_x \in \mathcal{S}(\mathcal{H}^Q), x \in \mathcal{X}\}$ and positive operator valued measures (POVMs) $E_j = (E_j^{z_j})$, $j \in [m]$, with outcomes $z_j \in \mathcal{A}$. Let Z_j be the random variable corresponding to the decoding of X_j by performing the measurement E_j on $\mathcal{H}^Q$ where we use $P_{Z_j|X_j}(z_j|x_j) := \mathrm{Tr}(E_j^{z_j}\rho_x)$ to denote the conditional probability distribution of a random variable Z_j over $\mathcal{A}$. Then

$$\dim(\mathcal{H}^Q) \geq 2^{H(X) - \Sigma_j H(X_j|Z_j)}.$$

Proof. Consider a cq-state $\rho^{XQ} \in \mathcal{S}(\mathcal{H}^X \otimes \mathcal{H}^Q)$ of the form

[1] Without loss of generality, we will take all measurements to have the same number of outcomes, as we may extend them otherwise.

$$\rho^{XQ} = \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_x^Q,$$

where $\rho_x^Q := \rho_x$. We have

$$\begin{aligned}\log_2[\dim(\mathcal{H}^Q)] &\geqslant H(Q)\\ &\geqslant H(Q) - \sum_x P_X(x)H(\rho_x^Q)\\ &= H(X) + H(Q) - H(XQ)\\ &= H(X) - H(X|Q)\\ &\geqslant H(X) - \sum_{j=1}^m H(X_j|Q),\end{aligned}$$

where the first inequality follows from $H(Q) \leqslant \log_2[\dim(\mathcal{H}^Q)]$ [17], Theorem 11.8.2, the second from the fact that for all ρ_x^Q we have $H(\rho_x^Q) \geqslant 0$ [17], Theorem 11.8.1, the third equality from $H(XQ) = H(X) + \Sigma_x P_X(x) H(\rho_x^Q)$ [17], Theorem 11.8.5, the fourth from the definition of the conditional von Neumann entropy, and the last inequality from its strong subadditivity $H(X_1, \ldots, X_m|Q) = H(X|Q) \leqslant \Sigma_{j=1}^m H(X_j|Q)$ [17], Theorem 11.16, where X_j is the random variable corresponding to the jth entry of X.

Finally, note that we can express the effects of a measurement M on $\mathcal{H}^Q$ by performing a unitary operation U on $\rho^{XQ} \otimes |z_0\rangle\langle z_0| \in \mathcal{S}(\mathcal{H}^X \otimes \mathcal{H}^Q \otimes \mathcal{H}^{Z_j})$ with $|z_0\rangle\langle z_0|$ an initial pure state of $\mathcal{H}^{Z_j}$, followed by tracing out the ancilla $\mathcal{H}^{Z_j}$ holding the measurement outcome [17]. We then have $H(X_j|Q) = H(X_j|QZ_j)$ since U is unitary, and $H(X_j|QZ_j) \leqslant H(X_j|Z_j)$ since conditioning reduces entropy ([17], Theorem 11.15.1), from which the claim follows. ■

This means that if we want to encode a string of n dits[2] into a number of qubits and attempt to recover the jth dit with the jth measurement, then we need at least $H(X) - \Sigma_j H(X_j|Z_j)$ qubits. $H(X_j|Z_j)$ quantifies the uncertainty about the jth bit given the outcome of the jth measurement. For instance, if the n dits are drawn uniformly and independently ($H(X) = n \log_2 d$ where $d = |\mathcal{A}|$) and we wish to recover them perfectly, ($H(X_j|Z_j) = 0$ for all j) then we need $n \log_2 d$ qubits to do so. In the other extreme, where Z_j holds no information about X_j, i.e., our recovery probability is no better than guessing, we have $H(X_j|Z_j) = H(X_j) = \log_2 d$ for all j, meaning that we need no qubits at all for the encoding.

Corollary 1. For the definitions as given in Lemma 1, it furthermore holds that

$$\dim(\mathcal{H}^Q) \geqslant 2^{H(X) - \Sigma_j [h(p_j) + (1-p_j)\log_2(|\mathcal{A}|-1)]},$$

where $p_j = \Sigma_{x \in \mathcal{X}} P_X(x) \mathrm{Tr}(E_j^{x_j} \rho_x)$ is the average recovery probability of the jth entry of X when measuring E_j on $\mathcal{H}^Q$.

Proof. The statement follows immediately from Lemma 1 and Fano's inequality giving $H(X_j|Z_j) \leqslant h(p_j) + (1-p_j)\log_2(|\mathcal{A}|-1)$, where p_j is the average probability of correctly decoding the jth bit of X given access to Z_j. ■

Note that the bound further simplifies to

$$\dim(\mathcal{H}) \geqslant 2^{m[1-h(p)]}$$

in the case where X_j is binary and P_X is the uniform distribution for which $H(X_j) = 1$ for all j and the recovery probability for each bit is lower bounded by $p \geqslant \frac{1}{2}$. This last bound was first noted in the context of random access codes. Lemma 1 does in general give a better bound than Corollary 2, although it may be harder to apply since it requires more information about the distributions and will be less convenient for us when considering nonlocal games where we may have limited information. Fano's inequality is tight for a distribution where the most likely outcome has probability p_j and all others have probability $p_j/(|\mathcal{A}|-1)$ and in this case Corollary 1 gives exactly the same bound as Lemma 1.

B. Dimension bound

We are now ready to use these tools to prove a lower bound for our problem. Intuitively, we let the states $\rho_1, \ldots, \rho_\ell$ corresponding to the preparations $r_1, \ldots, r_\ell$ represent encodings of m-element strings x chosen according to a probability distribution P_X from $\mathcal{X} = \mathcal{A}^{\times m}$. If $\ell < |\mathcal{A}|^m$, then this just means that some strings have zero probability of occurring. If $\ell > |\mathcal{A}|^m$, then there are more elements in our string than we wish to extract, in which case our lower bound will not be any stronger than could be obtained by letting $\ell = |\mathcal{A}|^m$. There is some freedom in applying the above bound to our setting, since we are in general free to associate strings with states in any way we like, pick any of our available measurements M_j to decode x_j, and finally also choose any prior distribution $P_X(x)$, since our bound should hold for any such prior. First, we associate strings $x \in \mathcal{X}$ with states $\rho \in \mathcal{S}$ by constructing a map $g_{T,R}$ as follows. Recall that, without loss of generality, we may order the states in lexicographic order $\rho_1, \ldots, \rho_\ell$. Let $T \subseteq [\ell]$ such that $|T| = \min(\ell, |\mathcal{X}|)$, let $R \subseteq \mathcal{X}$ such that $|R| = |T|$, and consider the set of one-to-one maps

$$\mathcal{G}_{T,R} := \{g_{T,R} : R \to T | \ \forall \ x \neq x' \in \mathcal{X}, g_{T,R}(x) \neq g_{T,R}(x')\}.$$

That is, any map associates a unique state $\rho_{g_{T,R}(x)}$ with each string x. Second, we now construct maps $e : [\ell] \to [\ell]$ and $c : \mathcal{A} \times [\ell] \to \mathcal{A}$ that specify which measurement $E_j = M_{e(j)}$ we will use to extract a particular entry x_j of $x = x_1, \ldots, x_m$ from $\rho_{g_{T,R}(x)}$, for a potential relabeling of the outcomes as $M_{e(j)}^{c(a,j)} = E_j^a$ given by the map c. Let $\mathcal{D} = \{(e,c)\}$ denote the set of all such collections of maps. Finally, we may choose P_X to be any distribution over $\mathcal{X}$, where we will assign probability $P_X(x) = 0$ to any $x \notin R$. Note that this means P_X is effectively a distribution over R. If we take $P_X = P_{X_1} \times \cdots \times P_{X_m}$ to be a product distribution over all strings, then $e(j) = j$ is simply the identity, i.e., we will use measurement M_j to decode the jth element of the string.

We first of all show that Lemma 1 gives us a lower bound on the dimension of the quantum system for *any* distribution P_X, $T \subseteq [\ell]$, $R \subseteq \mathcal{X}$, and mapping $g_{T,R}$, c, and e. This will be important in Sec. III, where such mappings are fixed when considering a particular nonlocal game. We state both consequences of Lemma 1 and Corollary 1 explicitly.

[2] A dit is a unit of information analogous to a bit that can take $|\mathcal{A}| = d$ values.

Corollary 2. Let $\mathcal{S}=\{\rho_1,\ldots,\rho_\ell|\rho_j\in\mathcal{S}(\mathcal{H}^Q)\}$ be a set of states and let $\mathcal{M}=\{M_1,\ldots,M_m|M_j=(M_j^a)\in\mathcal{B}(\mathcal{H}^Q),a\in\mathcal{A}\}$ be a set of POVMs satisfying $p(a|j,r)=\mathrm{Tr}(M_j^a\rho_r)$ for some given set of probabilities $\{p(a|j,r)|a\in\mathcal{A},j\in[m],r\in[\ell]\}$. Then for any $T\subseteq[\ell]$, $R\subseteq\mathcal{X}$ with $|R|=|T|=\min(\ell,|\mathcal{X}|)$, and $g_{T,R}\in\mathcal{G}_{T,R}$, $(e,c)\in\mathcal{D}$, and any distribution P_X over $R\subseteq\mathcal{X}=\mathcal{A}^{\times m}$ giving ensemble $\{P_X(x),\rho_{g_{T,R}(x)}\}$ we must have

$$\dim(\mathcal{H}^Q)\geq 2^{H(X)-\Sigma_j H(X_j|Z_j)},$$

where $P_{Z_j|X_j}(z_j|x_j):=\mathrm{Tr}(M_{e(j)}^{c(x_j,j)}\rho_{g_{T,R}(x)})$. Furthermore,

$$\dim(\mathcal{H}^Q)\geq 2^{H(X)-\Sigma_j[h(p_j)+(1-p_j)\log_2(|\mathcal{A}|-1)]},$$

where $p_j=\Sigma_{x\in\mathcal{X}}P_X(x)\mathrm{Tr}(M_{e(j)}^{c(x_j,j)}\rho_{g_{T,R}(x)})$ is the average recovery probability of the jth entry of X when measuring $M_{e(j)}$ on $\mathcal{H}^Q$.

We are now ready to state our main result as an immediate consequence of Corollary 2.

Theorem 1. Let $\mathcal{S}=\{\rho_1,\ldots,\rho_\ell|\rho_j\in\mathcal{S}(\mathcal{H}^Q)\}$ be a set of states and let $\mathcal{M}=\{M_1,\ldots,M_m|M_j=(M_j^a)\in\mathcal{B}(\mathcal{H}^Q),a\in\mathcal{A}\}$ be a set of POVMs satisfying $p(a|j,r)=\mathrm{Tr}(M_j^a\rho_r)$ for some given set of probabilities $\{p(a|j,r)|a\in\mathcal{A},j\in[m],r\in[\ell]\}$. Then

$$\dim(\mathcal{H}^Q)\geq 2^C,$$

with

$$C:=\max_{T,R,g_{T,R}(e,c),P_X} H(X)-\sum_j H(X_j|Z_j),$$

where the maximization is taken over all subsets $T\subseteq[\ell]$, $R\subseteq\mathcal{X}$ with $|R|=|T|=\min(\ell,|\mathcal{X}|)$, and probability distributions P_X over R, and mappings $g_{T,R}\in G_{T,R}$, $(e,c)\in\mathcal{D}$, with $P_{Z_j|X_j}(z_j|x_j):=\mathrm{Tr}(M_{e(j)}^{c(x_j,j)}\rho_{g_{T,R}(x)})$.

Note that if we fix T, R, $g_{T,R}$, and (e,c) and furthermore restrict the maximization to product distributions $P_X=P_{X_1}\times\cdots\times P_{X_m}$ we have from Lemma 1 combined with Corollary 2 that

$$\dim(\mathcal{H}^Q)\geq 2^{\Sigma_j C_j},$$

where $C_j=\max_{P_{X_j}}I(X_j;Z_j)$ is the Shannon channel capacity, and $I(X_j;Z_j)=H(X_j)-H(X_j|Z_j)$ is the mutual information. Unfortunately, we do not know how hard it is to evaluate the quantity C in general when maximizing over all parameters. However, since for fixed $T,R,g_{T,R},(e,c)$ it is equivalent to computing the Shannon channel capacity it may not be an easy task for arbitrary distributions $p(a|j,r)$.

Let us look at a very simple example taken from [12,13] that illustrates our bound. The entries of the following table correspond to the probabilities $p(a|M_j,\rho_x)$, for the two possible states labeled using strings 00 and 11:

M	a	ρ_{00}	ρ_{11}
M_1	0	1	0
	1	0	1
M_2	0	1	1/2
	1	0	1/2

Note that in this example $\mathcal{X}=\{0,1\}^2$ consists of the possible strings of two bits, but only 00 and 11 occur with nonzero probability. For simplicity, suppose we are given these states with probability 1/2 each, and hence we have $H(X)=1$. Note that we can distinguish the two states perfectly using the first measurement, and hence both encoded bits can be recovered perfectly, $p_1=p_2=1$. By reference to Corollary 1 we see that at least a two-dimensional system is required to recover these statistics. If we can only perform projective measurements, then [12,13] say that we need more than one qubit. Note, however, that this is not the case for generalized measurements. To perform the second measurement M_2 we can perform the first measurement $M_1=\{M_1^a\}$, and output 0 for outcome, 0, but for outcome 1 we flip a coin that gives us 0 and 1 with probability 1/2 each. This corresponds to letting $M_2^0=M_1^0+M_1^1/2$ and $M_2^1=M_1^1/2$. Hence, our bound is tight for this trivial example. Below, we provide a second example that is inspired by the Clauser-Horne-Shimony-Holt (CHSH) inequality.

Our analysis shows that it is indeed possible to obtain bounds on the dimension in the quantum setting, partially answering an open question from [12,13] which asked to find such bounds for *projective* measurements. In particular, note that if we choose a uniform prior over $N=\min(|\mathcal{S}|,2^{\mathcal{M}})$ possible states, and consider only two outcome measurements we have by Corollary 1 that the dimension of the system must obey $\log_2 d\geq\Sigma_{j=1}^{\log_2 N}[1-h(p_j)]\geq\log_2 Nc$ with $c=\min_j[1-h(p_j)]$. This means that in the case where p_j is not arbitrarily close to 1/2, and N itself is very large, the dimension required is not significantly different from the one required by the ontological model [12,13]. It is worth considering the dependence on p_j which seems to be absent from this particular ontological model. If we merely want to represent the data classically in a way such that we can extract an arbitrary bit x_j alone with probability $p_j=p$ and the prior distribution over the strings is uniform, it is known that there do exist classical random access codes for which the dimension obeys $\log_2 d=\log_2 N[1-h(p)]+O(\log_2\log_2 N)$ [19]. Intuitively, the description of the ontological model includes much more information and hence has a larger size. When examining information processing within such an ontological model, it may however be worth considering whether it has a better representation for a particular task at hand.

III. NONLOCAL GAMES

We now show how our approach also leads to a lower bound on the dimension of the entangled state that two or more parties need to share in *any* Bell experiment, where we consider a bound for the CHSH inequality as a small example. In this case we can immediately compute the lower bound since all parameters P_X, T, R, $g_{T,R}$, and (e,c) are fixed. For the present purposes, it is convenient to view Bell experiments as a game between two, or more, distant players, who cooperate against a special party. We call this special party the *verifier*. In a two-player game with players Alice and Bob, the verifier picks two questions $s\in\mathcal{S}$ and $t\in\mathcal{T}$ and sends them to Alice and Bob, respectively. Alice and Bob, then return answers $a\in\mathcal{A}$ and $b\in\mathcal{B}$ to the verifier, who then

decides according to a fixed set of public rules whether Alice and Bob win by giving answers a and b to questions s and t. To win the game, Alice and Bob may agree on any strategy beforehand, but can no longer communicate once the game starts. Classically, such a strategy consists of shared randomness. In the quantum setting, they may choose any entangled state as part of their strategy and agree on any measurements to be performed on this state. Without loss of generality we can thus think of the questions as measurement settings and the answers as measurement outcomes.

More formally, the game is characterized by finite sets $\mathcal{S},\mathcal{T},\mathcal{A},\mathcal{B}$, a distribution $\pi:\mathcal{S}\times\mathcal{T}\rightarrow[0,1]$ according to which the verifier chooses his questions, and a predicate $V:\mathcal{A}\times\mathcal{B}\times\mathcal{S}\times\mathcal{T}\rightarrow\{0,1\}$, where $V(a,b|s,t)=1$ if and only if a and b are winning answers given questions s and t. Let π_A and π_B be the marginal probability distributions over $\mathcal{S}$ and $\mathcal{T}$, respectively. For simplicity, we also assume that we are dealing with a *unique* game, where V is defined in such a way that for each b,s,t there exists exactly one winning answer a for Alice. Our argument for the general case is analogous, and can be obtained by combining the correct answers into one, which effectively corresponds to performing a measurement with less outcomes. However, our proof just becomes much harder to read. For simplicity in our explanations, we will also assume that the possible answers are the same for each possible measurement setting.

Let $P[a|s]$ and $P[b|t]$ be the probabilities that Alice and Bob return answers a and b given questions s and t, respectively. Note that the no-signaling condition must hold and hence we may without loss of generality assume these probabilities to be independent of the other party's measurement setting. We now show how to use our approach from above to lower bound the dimension of the entangled state that Alice and Bob need to implement such a strategy. We are not concerned with the question whether there actually exists a strategy for Alice and Bob to obtain said distribution. This can be verified using the techniques of [2–4].

The simple trick is to realize that when Bob performs a measurement on his part of the state, he prepares a certain state on Alice's end. Let χ_b^t denote the state that is prepared for Alice if Bob has measurement setting $t\in\mathcal{T}$ and obtains outcome $b\in\mathcal{B}$. The probability that Alice holds the state χ_b^t is given by

$$P_X(t,b) := P[b|t]\pi_B(t),$$

where we combine t,b to index a string $x\in\mathcal{A}^{|\mathcal{S}|}$ as follows. Note that, since we are dealing with unique games, we can define a function $f:\mathcal{B}\times\mathcal{S}\times\mathcal{T}\rightarrow\mathcal{A}$ such that $f(b,s,t)=a$ for $V(a,b|s,t)=1$. We can label Alice's measurements with numbers from one up to $|\mathcal{S}|$ and hence without loss of generality we will take $\mathcal{S}=[|\mathcal{S}|]$ to represent the set of possible measurements for Alice. We define the string $x\in\mathcal{A}^{|\mathcal{S}|}$ as

$$x := f(b,1,t),\ldots,f(b,|\mathcal{S}|,t) \tag{1}$$

and let

$$\rho_x := \chi_b^t.$$

Since x is a function of b and t, we have

$$P_X(x) := P_X(t,b).$$

If Alice chooses measurement setting s she will try to give the correct answer a. Note that effectively she tries to retrieve the entry $x_s=f(b,s,t)$ from ρ_x, completing the analogy to quantum random access codes.

To apply Corollary 1, let $p(a|M_s,\chi_b^t)$ be the probability that Alice outputs a for measurement setting s and prepared state χ_b^t.

Corollary 3. In any nonlocal game where Alice obtains the correct outcome $a\in\mathcal{A}_s$ for measurement setting $s\in\mathcal{S}$ with probability p_s, the dimension of her Hilbert space $\mathcal{H}^A$ obeys

$$\dim(\mathcal{H}^A)\geqslant 2^{H(X)-\Sigma_s[h(p_s)+(1-p_s)\log_2(|\mathcal{A}_s|-1)]},$$

where X is the random variable corresponding to the choice of string as defined in Eq. (1).

Evidently, an analogous statement can be made for Bob. If we are considering more than two players, it is straightforward to extend our argument to bound the Hilbert space dimension of each individual player by grouping the remaining players together as one.

Let us look at a small example which illustrates the proof. Consider the CHSH game. Here, $\mathcal{A}=\mathcal{B}=\mathcal{S}=\mathcal{T}=\{0,1\}$ and π is the uniform distribution. Alice's goal is to obtain an outcome a such that $st=a+b \mod 2$. Letting $x=g(b,t)=f(b,0,t),f(b,1,t)$ we obtain an encoding of a two-bit string $x\in\{0,1\}^2$ as $g(0,0)=0,0$, $g(1,0)=1,1$, $g(0,1)=1,0$, and $g(1,1)=0,1$. How many qubits does Alice need to use if she always wants to give the correct answer with probability $\gamma=1/2+1/(2\sqrt{2})$? With analogy to the table of our previous example, we have probabilities $p(a|M_s,\rho_x)$ given by

M	a	ρ_{00}	ρ_{01}	ρ_{10}	ρ_{11}
M_0	0	γ	γ	$(1-\gamma)$	$(1-\gamma)$
	1	$(1-\gamma)$	$(1-\gamma)$	γ	γ
M_1	0	γ	$(1-\gamma)$	γ	$(1-\gamma)$
	1	$(1-\gamma)$	γ	$(1-\gamma)$	γ

We have for all $t,b\in\{0,1\}P[b|t]=1/2$ and hence $P_X(x)=P_X(t,b)=1/4$. Since everything is uniform we immediately obtain from Corollary 1 that $\log_2[\dim(\mathcal{H}^A)]\geqslant[1-H(p)]2\approx 0.8$. Hence Alice needs at least one qubit to no great surprise. We do not need to know a specific strategy, however, for the well-known CHSH state and measurements we would have an encoding of $\rho_{00}=|0\rangle\langle 0|$, $\rho_{01}=|-\rangle\langle -|$, $\rho_{10}=|+\rangle\langle +|$, and $\rho_{11}=|1\rangle\langle 1|$ which actually coincides with the best known quantum random access code for a two-bit string.

Bounds for other games for which we are given a distribution over the measurement outcomes can be shown in an analogous way. In general, if we are given the full distribution over all settings and outcomes we can apply the first part of Corollary 2 to obtain a slightly better bound, depending on the distribution.

IV. MIN-ENTROPY

Our task of lower bounding the dimension of the Hilbert space can be used to give a partial answer to an open prob-

lem in the analysis of cryptographic protocols in the bounded-quantum-storage [21,22], and noisy-quantum-storage models [14–16]. In particular, the example discussed in the previous section can be modified to give a simple counterexample that shows that an additivity property of the smooth min-entropy that has been shown to hold for independent quantum states [16] is not true in general. Note that a modified version may still hold with additional loss in the parameters. The same counterexample can also be used to show that exact min-entropy splitting with respect to quantum knowledge as it holds for independent states [16] is not possible in general without imposing further assumptions. We defer the details of this construction to the Appendix.

V. CONCLUSION

We have given a simple lower bound that places a fundamental limit on how large the dimension of the state has to be to implement certain measurement strategies. Our result shows that in the limit of a large number of measurement settings and states, the dimension of this state cannot generally be significantly smaller than the amount of classical information [e.g., in the form of (hidden) variables in an ontological model [12,13]] necessary to produce the desired statistics.

Our approach also gives a weak bound on the dimension of the entangled state needed to implement nonlocal strategies for any multiplayer nonlocal game. Note, however, that our bound will be quite weak if the probability of outputting the correct outcome is close to $1/2$, or the number of measurement outcomes is large. Furthermore, note that our bound also works for the case where the choice of Alice's measurement settings is uniform which may not be the case for a particular game, leaving the possibility of better bounds. Yet, our approach is a first direction to find bounds for general games. It is an interesting question whether the present idea of viewing the game as an encoding procedure leads to new upper bounds as well.

ACKNOWLEDGMENTS

We are indebted to the referee for helpful comments to improve the presentation of the paper. A.C.D. is supported by the Australian Research Council. S.W. is supported by NSF Grant No. PHY-04056720. S.W. thanks the University of Queensland for the generous travel support to attend QIP '07, and Oscar Dahlsten and Renato Renner for the invitation to the workshop "Information Primitives and Laws of Nature" at ETH Zurich.

APPENDIX: MIN-ENTROPY

In this appendix, we describe the counterexample mentioned in the text showing that the additivity property that was proved recently for the smooth min-entropy of independent quantum states does not hold in general. The same example can also be used to show that min-entropy splitting with respect to quantum knowledge as was shown for such states does not hold in general, without imposing additional constraints.

1. Definitions

To state the additivity lemma, we will need the following quantities introduced by Renner [23], reproduced here for convenience: Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$ and let $\sigma_B \in \mathcal{S}(\mathcal{H}^B)$. Then the *min-entropy of* ρ_{AB} *relative to* σ_B is given by

$$H_\infty(\rho_{AB}|\sigma_B) := -\log_2 \lambda,$$

where λ is the smallest real number such that $\lambda \mathbb{I}_A \otimes \sigma_B \geqslant \rho_{AB}$. We need a related quantity, where in addition we optimize over states σ_B defined as

$$H_\infty(\rho_{AB}|B) := \sup_{\sigma_B \in \mathcal{S}(\mathcal{H}^B)} H_\infty(\rho_{AB}|\sigma_B).$$

For a CQ state ρ_{XE}, we also use the shorthand

$$H_\infty(X|E) := \sup_{\sigma_E \in \mathcal{S}(\mathcal{H}^E)} H_\infty(\rho_{XE}|\sigma_E)$$

for the *conditional min-entropy of X given E*. It is difficult to get an intuitive understanding from this formal definition of conditional min-entropy, but one can show using semidefinite programming duality [24] that

$$H_\infty(X|E) = -\log_2 P_g(X|E), \tag{A1}$$

where $P_g(X|E)$ is defined as the maximum success probability of guessing X by measuring the E register of ρ_{XE}. Formally, for any (not necessarily normalized) cq-state ρ_{XE}, the guessing probability is defined as

$$P_g(X|E) := \sup_{\{M_x\}} \sum_x P_X(x) \mathrm{Tr}(M_x \rho_E^x),$$

where the supremum ranges over all positive operator valued measurements with measurement elements $\{M_x\}_{x \in \mathcal{X}}$, i.e., $M_x \geqslant 0$ and $\Sigma_x M_x = \mathbb{I}_E$. If all side information is classical, we recover the fact that the classical min-entropy is the negative logarithm of the maximum probability.

We will also refer to smooth versions of these quantities. Intuitively, we no longer consider the min-entropy of a fixed state ρ_{AB}, but allow us to move to some $\hat{\rho}_{AB}$ which is close to ρ_{AB}, but may have considerably larger min-entropy. These smooth quantities are often needed since they have some nicer properties than the conventional min-entropy. For $\varepsilon \geqslant 0$, the *ε-smooth min-entropy of* ρ_{AB} *relative to* σ_B is given by

$$H_\infty^\varepsilon(\rho_{AB}|\sigma_B) := \sup_{\hat{\rho}_{AB} \in \mathcal{K}^\varepsilon(\rho_{AB})} H_\infty(\hat{\rho}_{AB}|\sigma_B),$$

where $\mathcal{K}^\varepsilon(\rho_{AB}) := \{\hat{\rho}_{AB} \in \mathcal{P}(\mathcal{H}^A \otimes \mathcal{H}^B) \,|\, \|\rho_{AB} - \hat{\rho}_{AB}\|_1 \leqslant \mathrm{Tr}(\rho_{AB})\varepsilon$ and $\mathrm{Tr}(\hat{\rho}_{AB}) \leqslant \mathrm{Tr}(\rho_{AB})\}$. Finally, we need the related quantity of the *ε-smooth min-entropy of* ρ_{AB} *relative to* B defined by Renner, where we now again maximize over all states $\sigma_B \in \mathcal{S}(\mathcal{H}^B)$:

$$H_\infty^\varepsilon(\rho_{AB}|B) := \sup_{\sigma_B} H_\infty^\varepsilon(\rho_{AB}|\sigma_B).$$

We also use the shorthand

$$H_\infty^\varepsilon(A|B) := H_\infty^\varepsilon(\rho_{AB}|B).$$

2. Additivity

In [16], Lemma 2.2, it was shown that for two independent quantum states $\rho_{X_1E_1}$ and $\rho_{X_2E_2}$ we have

$$H_\infty^\varepsilon(X_1|E_1) + H_\infty^\varepsilon(X_2|E_2) \geq H_\infty^{\varepsilon^4}(X_1X_2|E),$$

where $E=E_1E_2$ and E_1 and E_2 are independent. Hence, one might hope that something similar holds for a general ccq-state; in particular, that we have

$$H_\infty^\varepsilon(X_1|E) + H_\infty^\varepsilon(X_2|E) \geq H_\infty^{\varepsilon^4}(X_1X_2|E). \tag{A2}$$

However, we now show that there exists a cq-state

$$\rho_{X_1X_2E} = \sum_{x_1x_2 \in \{0,1\}} p_{x_1x_2} |x_1x_2\rangle\langle x_1x_2| \otimes \rho_x^E$$

that violates this statement for small ε.

From the chain rule for the smooth min-entropy, and the data-processing inequality [23], Theorem 3.2.12, we have

$$H_\infty^\varepsilon(X_1X_2|E) \geq H_\infty^\varepsilon(X_1X_2E) - H_0(E) \geq H_\infty^\varepsilon(X_1X_2) - H_0(E).$$

Using that $H_0(E)=\log_2 \text{rank}\ \rho_E$ we thus have

$$\log_2 \dim(\mathcal{H}^E) \geq H_\infty^\varepsilon(X_1X_2) - H_\infty^\varepsilon(X_1X_2|E).$$

Now consider the CHSH example given above. Let $p_{x_1x_2}$ be the uniform distribution, and again let $\rho_{00}=|0\rangle\langle 0|$, $\rho_{01}=|-\rangle\langle -|$, $\rho_{10}=|+\rangle\langle +|$, and $\rho_{11}=|1\rangle\langle 1|$. The random variables X_1 and X_2 here correspond to the choice of the first and second bit, respectively.

First, consider the case of $\varepsilon=0$. And suppose by contradiction that Eq. (A2) holds. Note that for our simple example we have for any $D \in \{1,2\}$ that $H_\infty(X_D|E)=-\log_2 t$ with $t=1/2+1/(2\sqrt{2})$, since the min-entropy directly relates to the guessing probability as outlined in Eq. (A1). Hence, we would have

$$\log_2 \dim(\mathcal{H}^E) \geq H_\infty(X_1X_2) - H_\infty(X_1X_2|E) \geq 2 + 2\log_2 t \approx 1.54.$$

However, we know that one qubit, i.e., $\log_2 \dim(\mathcal{H}^E)=1$, is sufficient for this encoding. For small ε, we can make a similar argument by virtue of the fact that $-\log_2[P_g(X_j|E) - \varepsilon] \geq H_\infty^\varepsilon(X_j|E)$ [25] and $H_\infty^\varepsilon(X_1X_2) \geq H_\infty(X_1X_2)$.

Additivity of the smooth min-entropy was required as a tool to show a so-called min-entropy splitting lemma for independent quantum states [16]. Intuitively, the technique of min-entropy splitting, first introduced by Wullschleger [26] for classical min-entropy, states that if the min-entropy of two (or more) random variables X_1X_2 is high, then the min-entropy of either X_1 or X_2 must be greater than half the joint min-entropy. Here, we are interested in the min-entropy of X_1X_2 *conditioned* on quantum information. In particular, it was shown in [16], Lemma 2.7, that for $\varepsilon \geq 0$ and two independent states $\rho_{X_1E_1}$ and $\rho_{X_2E_2}$, satisfying

$$H_\infty^{\varepsilon^4}(X_1X_2|E_1E_2) \geq \alpha,$$

there exists a random variable $D \in \{1,2\}$ such that

$$H_\infty^\varepsilon(X_D|E) \geq \alpha/2,$$

with $E=E_1E_2$. It was an open problem in [16] whether this statement is also true for arbitrary ccq-states $\rho_{X_1X_2E}$. Since additivity falls, it is no longer clear whether this would be true in general. By the same argument as above, one can also see that for X_1 and X_2 the random variables corresponding to the encoding of the first or second bit, respectively, we cannot have that $H_\infty(X_1|E) \geq H_\infty(X_1X_2|E)/2$ or $H_\infty(X_2|E) \geq H_\infty(X_1X_2|E)/2$.

This small example shows that we must be very careful when trying to perform min-entropy splitting with respect to quantum information, and indeed one can also use the present example to disprove min-entropy splitting for non-independent states. However, it does not rule out that such a statement is still true with a significant loss in the smoothing parameter ε or by adding an additional fudge factor. Indeed, such statements involving additional factors are known if the number of random variables $X_1,\ldots,X_n$ is small compared to the size of the set $\mathcal{X}$ over which the variables $X_1,\ldots,X_n$ are distributed [25]. Unfortunately though, they do not give nice bounds in our setting.

[1] R. Blume-Kohout, C. Caves, and I. Deutsch, Found. Phys. **32**, 1641 (2002).
[2] M. Navascues, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).
[3] M. Navascués, S. Pironio, and A. Acin, e-print arXiv:0803.4290.
[4] A. Doherty, Y.-C. Liang, B. Toner, and S. Wehner, in *Proceedings of the 23rd IEEE Conference on Computational Complexity, 2008*, (IEEE Computer Society, Washington D.C., 2008), pp. 199–210.
[5] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).
[6] B. Tsirelson, Hadronic J. Suppl. **8**, 329 (1993).
[7] K. F. Pál and T. Vértesi, Phys. Rev. A **77**, 042105 (2008).
[8] T. Vértesi and K. F. Pál, Phys. Rev. A **77**, 042106 (2008).
[9] S. Wehner and A. Doherty (unpublished).
[10] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).
[11] J. Briët, and B. Toner (unpublished).
[12] N. Harrigan, T. Rudolph, and S. Aaronson, e-print arXiv:0709.1149.
[13] T. Rudolph (unpublished).
[14] S. Wehner, Ph.D. thesis, University of Amsterdam, 2008, e-print arXiv:0806.3483.
[15] S. Wehner, C. Schaffner, and B. M. Terhal, Phys. Rev. Lett. **100**, 220502 (2008).
[16] C. Schaffner, B. Terhal, and S. Wehner, e-print arXiv:0807.1333.
[17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).

[18] A. Nayak, in *Proceedings of the 40th IEEE FOCS, 1999*, (IEEE Computer Society, Washington D.C., 1999), pp. 369–376.

[19] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, in *Proceedings of the 31st ACM STOC, 1999*, (ACM, New York 1999), pp. 376–383.

[20] I. Kerenidis and R. de Wolf, J. Comput. Syst. Sci. **69**, 395 (2004).

[21] I. Damgård, S. Fehr, L. Salvail, and C. Schaffner, in *Proceedings of 46th IEEE FOCS, 2005*, (IEEE Computer Society, Washington D.C. 2005), pp. 449–458.

[22] I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner, in *Advances in Cryptology—CRYPTO '07*, Lecture Notes in Computer Science Vol. 4622 (Springer-Verlag, Berlin, 2007), pp. 360–378.

[23] R. Renner, Ph.D. thesis, ETH Zurich, 2005, e-print arXiv:quant-ph/0512258.

[24] R. König, R. Renner, and C. Schaffner, e-print arXiv:0807.1338.

[25] R. König and R. Renner e-print arXiv:0712.4291.

[26] J. Wullschleger, in *Advances in Cryptology—EUROCRYPT '07* (Ref. [22]). Lecture Notes in Computer Science.