Tight Bound on Randomness for Violating the CHSH Inequality

Yifeng Teng,\textsuperscript{1,a)} Siwei Wang,\textsuperscript{1,b)} Mingfei Zhao,\textsuperscript{1,c)} and Shenghao Yang\textsuperscript{2,d)}

\textsuperscript{1)} Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, P. R. China

\textsuperscript{2)} Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong, P. R. China

(Dated: 22 April 2015)

Free will has been studied to achieve loophole-free Bell’s inequality test and to provide device-independent quantum key distribution security proofs. The required free will (or randomness) such that a local hidden variable model (LHVM) can violate the Clauser-Horne-Shimony-Holt (CHSH) inequality has been studied, but a tight bound has not been proved for a practical case that i) the device settings of the two parties in the Bell test are independent; and ii) the device settings of each party can be correlated or biased across different runs. For a randomness measure related to min-entropy, we prove in this paper a tight bound of the required randomness for this case such that the CHSH inequality can be violated by certain LHVM. Our technique used in the proof is also of independent interest.

\textsuperscript{a)} tchief4685@gmail.com
\textsuperscript{b)} acmilan120791899@126.com
\textsuperscript{c)} mingfeizhao@hotmail.com
\textsuperscript{d)} shenghao.yang@gmail.com
I. INTRODUCTION

The Clauser-Horne-Shimony-Holt (CHSH) inequality\(^1\) is the most often used inequality\(^2\) in Bell test experiments. Experimental demonstrations of the violation of CHSH inequalities have been conducted since 1982\(^3\) (see also Giustina et al’s work\(^4\) and the references therein). These Bell tests, however, suffer from an inherent loophole that the settings of the participated devices may not be chosen totally randomly, called the randomness (free will) loophole. A small amount of correction between the device settings makes it possible that a local hidden variable model (LHVM) can reproduce predictions of quantum mechanics\(^5–9\). This loophole also weakens the Bell’s inequality based security proofs of device-independent quantum key distribution\(^10–12\) and randomness expansion\(^13–15\).

One of the essential questions in the randomness loophole is the bound of randomness such that the correctness of Bell tests can (or cannot) be guaranteed. Using a min-entropy type randomness measure, the bound of randomness required in a test can be formulated as an optimization problem\(^16–18\). In this paper, we study this optimization problem and obtain the asymptotic optimal value explicitly in the case that the two parties of the test have independent settings, but the setting of each party can be biased or correlated across different runs.

A. Problem Formulation

Let \(n\) be a positive integer, and \(X, Y\) be two random variables over \(\{0, 1\}^n\) with a joint distribution \(p_{XY}\). We may consider that \(X\) and \(Y\) are the device settings of the two parties in an \(n\)-run Bell test, respectively. Define

\[
P = \left( \max_{x,y \in \{0,1\}^n} p_{XY}(x,y) \right)^{1/n}.
\]

When \(X\) and \(Y\) are independent and uniformly distributed, \(P = 1/4\), which is the minimum value of \(P\) and corresponds to the case of complete randomness. When \(X\) and \(Y\) are deterministic, \(P = 1\), which corresponds to the case of zero randomness. Note that \(P\) is related to the min-entropy:

\[
H_\infty(X,Y) := -\log \max_{x,y \in \{0,1\}^n} p_{XY}(x,y) = -n \log P.
\]
Regard the vectors $\mathbf{x} \in \{0, 1\}^n$ as column vectors and denote by $\mathbf{x}^T$ the transpose of the $\mathbf{x}$. The optimization problem of our interest is

$$\min_{p_{XY}} P \quad \text{s.t.} \quad \frac{1}{n} \sum_{x,y} \mathbf{x}^T y p_{XY}(x,y) \leq \frac{4 - S_Q}{8},$$

where $S_Q = 2\sqrt{2}$ is a quantum constant. Readers may refer to $^{16-18}$ to see how this problem is obtained. For the completeness, we briefly derive this problem in Appendix A with the minimum physics context.

Various special cases of this problem has been solved. When $n = 1$, Hall$^7$ and Koh et al$^{16}$ showed that the optimal value of (1) is $(S_Q + 4)/24 \approx 0.285$. When $n \to \infty$, Pope and Kay$^{17}$ showed that the optimal value of (1) converges to $3 - S_Q - 4 h_b(\frac{1}{8}) \approx 0.258$, where

$$h_b(t) = -t \log_2 t - (1 - t) \log_2 (1 - t)$$

is the binary entropy function.

The case that $X$ and $Y$ are independent is of particular interest. Towards a loophole free Bell test, physicists have designed experiments with independent device settings$^{19}$. In quantum key distribution, the experimental devices of the two parties may be manufactured independently and separated spatially, reducing the potential correlation of the device settings generated by the adversary. For independent device settings, (1) becomes

$$\min_{p_X,p_Y} P \quad \text{s.t.} \quad \frac{1}{n} \sum_{x,y} x^T y p_X(x)p_Y(y) \leq \frac{4 - S_Q}{8}. \quad (2)$$

When $n = 1$, it was obtained by Koh et al$^{16}$ that the optimal value of (2) is $S_Q/8 \approx 0.354$.

When $n \to \infty$, let $P_Q$ be the limit of the optimal value of (2). Yuan, Cao and Ma$^{18}$ show numerically that $P_Q \lesssim 0.264$. The value of $P_Q$ has the following interpretation. For any independent device settings with randomness less than $P_Q$, it is not possible to have a LHVM that violates CHSH inequality. But for any value $P > P_Q$, there exists a LHVM that violates CHSH inequality where the device settings are independent, but have randomness less than or equal to $P$. 

3
TABLE I. Previous results.

|                | correlated devices | independent devices |
|----------------|---------------------|---------------------|
| $n = 1$        | $(S_Q + 4)/24 \approx 0.285$ | $S_Q/8 \approx 0.354$ |
| $n \to \infty$| $3^{-S_Q/8} 2^{-h_b\left(\frac{1-S_Q}{8}\right)} \approx 0.258$ | $\leq 0.264$ |

B. Our Contribution

In this paper, we close the unresolved case in Table I by proving that the optimal value of (2), as $n \to \infty$, converges exactly to

$$P_Q = 4^{-h_b(\sqrt{c_Q})} = 0.26428\ldots,$$

where $c_Q = \frac{4 - S_Q}{8} \approx 0.1464$. Our formula has a min-entropy interpretation: $-n \log_2 P_Q = 2nh_b(\sqrt{c_Q})$, i.e., each bit in $X$ and $Y$ has an average min-entropy $h_b(\sqrt{c_Q})$. To achieve $P_Q$, we use the uniform distribution over the set of sequences in $\{0, 1\}^n$ with at most $n\sqrt{c_Q}$ ones for $X$ and $Y$. This is the same technique as in Yuan, Cao and Ma, but we provide an analytical characterization using an information theoretic approach.

The major part of the paper is to show the converse that no distributions of $X$ and $Y$ with randomness less than $P_Q$ can be feasible for (2). To prove the converse, we introduce a concept profile to characterize a set of binary sequences, and study some properties of profiles. The technique of profile seems to be firstly used here and may of independent interest for other problems.

Our techniques used to prove the main result are summarized in the next section, followed by the details in Section III.

II. OUTLINE OF THE PROOFS

As described in the previous section, we formulate an optimization problem as follows.

**Problem 1.** For any given $c \in (0, 1/4]$ and every positive integer $n$, consider the following programming

$$\min_{p_X, p_Y} \left(\max_x p_X(x) \max_y p_Y(y)\right)^{1/n},$$

s.t.

$$\frac{1}{n} \sum_{x, y \in \{0, 1\}^n} p_X(x)p_Y(y)x^Ty \leq c,$$
where \( p_X \) and \( p_Y \) are probability distributions over \( \{0, 1\}^n \). Let \( P_n \) be the optimal value of the above programming. We are interested in the limit of the sequence \( \{P_n\} \) when \( n \to \infty \).

Specifically we will need the case that \( c = c_Q \) for the physics problem of interests. Now we state the following theorem.

**Theorem 1.** For Problem 1 with \( c = c_Q \), \( \lim_{n \to \infty} P_n = 4^{-h_b(\sqrt{c_Q})} \), where

\[
h_b(t) = -t \log_2 t - (1 - t) \log_2 (1 - t)
\]

is the binary entropy function.

In the following of this section, we give an outline of the main techniques towards proving this theorem. We have the following bound for \( P_n \).

**Proposition 1.** For all sufficiently large \( n \), \( 1/4 \leq P_n < 1/2 \).

A. Simplified Problem

Let \( S_X \) and \( S_Y \) be the support of distributions \( p_X \) and \( p_Y \), respectively. Problem 1 can be simplified if we only consider distributions that are uniform over support. Suppose that

\[
p_X(x) = \frac{1}{|S_X|}, \quad \forall x \in S_X,
\]

\[
p_Y(y) = \frac{1}{|S_Y|}, \quad \forall y \in S_Y.
\]

Then we have

\[
\left( \max_x p_X(x) \max_y p_Y(y) \right)^{1/n} = \frac{1}{\sqrt{|S_X| \cdot |S_Y|}},
\]

and

\[
\frac{1}{n} \sum_{x,y \in \{0,1\}^n} p_X(x)p_Y(y)x^Ty = \frac{1}{n|S_X| \cdot |S_Y|} \sum_{x \in S_X, y \in S_Y} x^T y.
\]

Define a new problem as follows:

**Problem 2.** For any given \( c \in (0, 1/4] \) and every positive integer \( n \), consider the following programming

\[
\begin{aligned}
\min_{S_X,S_Y} & \quad \frac{1}{\sqrt{|S_X| \cdot |S_Y|}}, \\
\text{s.t.} & \quad \frac{1}{n|S_X| \cdot |S_Y|} \sum_{x \in S_X, y \in S_Y} x^T y \leq c,
\end{aligned}
\]
where $S_X$ and $S_Y$ are subsets of $\{0,1\}^n$. Let $P'_n$ be the optimal value of the above program-
ing. We are interested in the limit of the sequence $\{P'_n\}$ when $n \to \infty$.

It is obvious that $P_n \leq P'_n$ since only distributions that are uniform over support are
considered in Problem 2. The following theorem enables us to focus on $\lim_n P'_n$.

**Theorem 2.** $\lim_{n \to \infty} P'_n / P_n = 1$.

### B. Profiles

To study the properties of a set of binary vectors, we introduce the concept of profile. For
any positive integer $m$, we call vector $a = (a_1, a_2, \ldots, a_m) \in [0,1]^m$ a profile or an $m$-profile.
For each $S \subseteq \{0,1\}^n$, define the profile of set $S$ as

$$
\Gamma(S) = \begin{cases} 
1/|S| \sum_{s \in S} s, & |S| > 0; \\
(0,0, \ldots, 0), & |S| = 0.
\end{cases}
$$

We see that $\Gamma(S)$ is an $n$-profile.

Define the characteristic function of an $m$-profile $a$ as $f_a : [0, 1] \to [0, 1]$ such that

$$
f_a(t) = \begin{cases} 
a_1, & t = 0; \\
\lfloor atm \rfloor, & \forall 0 < t \leq 1.
\end{cases}
$$

For two profiles $a$ and $b$, we say $a \leq b$ if for any $0 \leq r \leq 1$, $f_a(r) \leq f_b(r)$, where $a$ and $b$ may not include the same number of components.

The following lemma tells us how to represent the constraint in Problem 2 in a simple way using profiles.

**Lemma 1.** In Problem 2, the left hand side of constraint (4) can be expressed as

$$
\frac{1}{n |S_X| \cdot |S_Y|} \sum_{x \in S_X, y \in S_Y} x^T y = \frac{1}{n} a^T b,
$$

where $a = \Gamma(S_X)$ and $b = \Gamma(S_Y)$. 


Proof. We can write
\[
\frac{1}{n} \cdot \frac{1}{|S_X|} \cdot \frac{1}{|S_Y|} \sum_{x \in S_X, y \in S_Y} x^T y = \frac{1}{n} \cdot \frac{1}{|S_X|} \cdot \frac{1}{|S_Y|} \left( \sum_{x \in S_X} x \right)^T \left( \sum_{y \in S_Y} y \right) \\
= \frac{1}{n} \cdot \frac{1}{|S_X|} \cdot \frac{1}{|S_Y|} (|S_X|a)^T(|S_Y|b)
\]
(5)
\[
= \frac{1}{n} a^T b,
\]
where (5) follows from the definition of the profile of a set of binary vectors.  

For a vector \(a\), we denote by \(a_i\) the \(i\)-th component of \(a\). The following theorem states that to get the value of \(P'_n\), we only need to consider \(S_X\) and \(S_Y\) with certain monotone property of their profiles.

**Theorem 3.** For all \(n\), there exist \(S_X, S_Y \subseteq \{0, 1\}^n\) that achieve \(P'_n\) in Problem 2 such that for \(a = \Gamma(S_X)\) and \(b = \Gamma(S_Y)\), \(0.5 \geq a_1 \geq a_2 \geq ... \geq a_n \geq 0\) and \(0 \leq b_1 \leq b_2 \leq ... \leq b_n \leq 0.5\).

By Theorem 3 it is sufficient for us to consider only profiles \(a \in [0, 0.5]^m\). For each \(m\)-profile \(a\), define its \(n\)-volume to be
\[
V_n(a) = \max \{|S| : S \subseteq \{0, 1\}^n, \Gamma(S) \leq a\}. \tag{6}
\]

**Lemma 2.** For any two profiles \(p\) and \(q\), if \(p \leq q\), we have \(V_n(p) \leq V_n(q)\) for every positive integer \(n\).

**Proof.** Notice that for any \(n\), any \(n\)-profile smaller than \(p\) is smaller than \(q\), then the lemma suffices.  

The following theorem gives an upper bound on the volume of a profile, which will be used in the proof of the lower bound on \(P'_n\).

**Theorem 4.** Fix an integer \(m\) and let \(a \in [0, 0.5]^m\) be an \(m\)-profile. For any positive integer \(n\), the \(n\)-volume of profile \(a\) satisfies
\[
V_n(a) \leq 2^{\frac{3}{2}n(\sum_{i=1}^m h_b(a_i)+o(1))}, \tag{7}
\]
where \(h_b\) is the binary entropy function defined in (3) and \(o(1) \to 0\) as \(n \to \infty\).
C. Converse and Achievability

**Theorem 5.** For any sequence of $S_X, S_Y \subseteq \{0,1\}^n$ such that

$$\frac{1}{n} \sum_{x,y \in \{0,1\}^n} \frac{1}{|S_X|} \cdot \frac{1}{|S_Y|} x^T y \leq cQ,$$

we have

$$\liminf_{n \to \infty} \frac{1}{\sqrt{|S_X||S_Y|}} \geq 4^{-h_b(\sqrt{cQ})}.$$

We then give a construction of $S_X$ and $S_Y$ to show that the bound in Theorem 5 is tight.

**Theorem 6.** There exists a sequence of $S_X, S_Y \subseteq \{0,1\}^n$ such that

$$\frac{1}{n} \sum_{x,y \in \{0,1\}^n} \frac{1}{|S_X|} \cdot \frac{1}{|S_Y|} x^T y \leq cQ,$$

and

$$\lim_{n \to \infty} \frac{1}{\sqrt{|S_X||S_Y|}} = 4^{-h_b(\sqrt{cQ})}.$$

Now we are ready to prove Theorem 1.

**Proof of Theorem 1** Theorem 5 implies that $\liminf_{n \to \infty} P'_n \geq 4^{-h_b(\sqrt{cQ})}$ and Theorem 6 implies that $\limsup_{n \to \infty} P'_n \leq 4^{-h_b(\sqrt{cQ})}$. Thus $\lim_{n \to \infty} P'_n = 4^{-h_b(\sqrt{cQ})}$, which together with Theorem 2 proves Theorem 1. \qed

III. PROOFS

A. Proof of Proposition 1

The lower bound follows that $\max_x p_X(x) \geq 1/2^n$ for any distribution $p_X$ over $\{0,1\}^n$. To prove the upper bound, consider the following two distributions:

$$p_X(x) = \begin{cases} 1 - 2c, & x = 0 \\ 2c/(2^n - 1), & x \neq 0, \end{cases}$$

and $p_Y(y) = 1/2^n$ for all $y \in \{0,1\}^n$. We then have

$$\frac{1}{n} \sum_{x,y \in \{0,1\}^n} p_X(x)p_Y(y)x^T y = \frac{2c}{2^n(2^n - 1)} \cdot \frac{1}{n} \sum_{x,y \in \{0,1\}^n} x^T y = \frac{c}{2^{n-1}(2^n - 1)} 2^{(n-1)} \leq c,$$
and
\[
P_n \leq \left( \max_x p_X(x) \max_y p_Y(y) \right)^{1/n} = \frac{1}{2} \left( \max\{1 - 2c, 2c/(2^n - 1)\} \right)^{1/n} = \frac{1 - 2c}{2} < \frac{1}{2},
\]
where the second equality follows from \( c \leq 1/4 \) and the last inequality follows from \( c > 0 \).

B. Proof of Theorem 2

Suppose that \( p_X \) and \( p_Y \) on \( \{0, 1\}^n \) achieve the minimum objective value \( P_n \) in Problem 1. Write
\[
\sum_{x, y \in \{0, 1\}^n} p_X(x) p_Y(y) x^T y = \sum_x p_X(x) \theta_{p_Y}(x),
\]
where
\[
\theta_{p_Y}(x) = x^T \left( \sum_y p_Y(y) y \right).
\]

Let \( P_X = \max_x p_X(x) \). We know that \( P_X > 0 \). If \( P_X = 1 \), then there exists \( x_0 \) such that \( p_X(x_0) = 1 \). In this case, \( P_n = 1/2 \) since otherwise we may instead choose \( p_X \) such that \( p_X(0) = 1 \) and \( p_Y \) such that \( p_Y(y) = 1/2^n \) for all \( y \in \{0, 1\}^n \). Thus we have a contradiction to \( P_n < 1/2 \) (see Proposition 1). Therefore, \( 0 < P_X < 1 \).

Now consider the following programming:
\[
\min_{p_X} \sum_x p_X(x) \theta_{p_Y}(x),
\]
subject to \( p_X(x) \leq P_X, \forall x \in \{0, 1\}^n \).

Let \( p_X^* \) be an optimal distribution that minimizes the objective of (A9). Since \( \theta_{p_Y}(x) \geq 0 \) for all \( x \), in the distribution \( p_X^* \), there must be \( \lfloor \frac{1}{P_X} \rfloor \) sequences \( x \) with \( p_X^*(x) = P_X \) and one sequence \( z \) with \( p_X^*(z) = 1 - \lfloor \frac{1}{P_X} \rfloor P_X \). For any other sequence \( x \), we have \( p_X^*(x) = 0 \).

We then have
\[
\sum_x p_X^*(x) \theta_{p_Y}(x) \leq \sum_x p_X(x) \theta_{p_Y}(x) \leq nc,
\]
and
\[
\left( \max_x p_X^*(x) \max_y p_Y(y) \right)^{1/n} = \left( P_X \cdot \max_y p_Y(y) \right)^{1/n} = P_n.
\]
Therefore, \( p_X^* \) and \( p_Y \) also obtain the minimum objective value \( P_n \) in Problem 1.

Let \( S_X \) be the support of \( p_X^* \). We have \( |S_X| = \lfloor \frac{1}{P_X} \rfloor \), and for any \( x \in S_X \), \( \theta_{p_Y}(z) \geq \theta_{p_Y}(x) \).
Let \( \bar{p}_X \) be the uniform distribution over \( S_X \setminus \{z\} \). Notice for all \( x \in S_X \setminus \{z\} \),
\[
\bar{p}_X(x) \geq p_X^*(x),
\]

9
\[ \sum_{x \in S_X \setminus \{z\}} (\bar{p}_X(x) - p_X^*(x)) = p_X^*(z). \]

We have
\[
\sum_{y, x} \bar{p}_X(x) p_Y(y) x^T y - \sum_{x, y} p_X^*(x) p_Y(y) x^T y = \sum_{x \in S_X \setminus \{x^*\}} (\bar{p}_X(x) - p_X^*(x)) \theta_{p_Y}(x) - p_X^*(x) \theta_{p_Y}(z)
\leq \sum_{x \in S_X \{x^*\}} (\bar{p}_X(x) - p_X^*(x)) \theta_{p_Y}(x) - p_X^*(x) \theta_{p_Y}(z)
= 0.
\]

Thus
\[
\sum_{x, y} \bar{p}_X(x) p_Y(y) x^T y \leq \sum_{x, y} p_X^*(x) p_Y(y) x^T y \leq nc. \tag{9}
\]

Let \(P_n^\dagger = \min_{p_X, p_Y} \left(\max_x p_X(x) \max_y p_Y(y)\right)^{1/n}\) such that \(p_X\) and \(p_Y\) satisfy the constraint of Problem \(\text{I}\) and \(p_X\) is uniform over its support. We have
\[
P_n \leq P_n^\dagger \leq \left(\max_x \bar{p}_X(x) \max_y p_Y(y)\right)^{1/n}
= \left(\frac{1}{\lceil 1/P_X \rceil} \max_y p_Y(y)\right)^{1/n}
\leq \left(\frac{1}{\lceil 1/P_X \rceil - 1} \max_y p_Y(y)\right)^{1/n}
\leq \left(3P_X \max_y p_Y(y)\right)^{1/n}
= 3^{1/n} P_n,
\]

where the second inequality follows that \(\bar{p}_X\) and \(p_Y\) satisfy the constraint of Problem \(\text{I}\) (see \((A10))\), and the last inequality follows from \(0 < P_X < 1\) and Lemma \(3\) (to be proved later in this section). Therefore, \(\lim_{n \to \infty} P_n^\dagger/P_n = 1\).

Similar technique can be used to show that \(\lim_{n \to \infty} P_n^\prime/P_n = 1\), which completes the proof of this theorem. Specifically, suppose that \(p_X, p_Y\) on \(\{0, 1\}^n\) achieve \(P_n^\dagger\) where \(p_X\) is uniform on its support. Define \(P_Y = \max_y p_Y(y)\) and \(P_X = \max_x p_X(x)\). Similar to the above argument, there exists distribution \(p_Y^*\) such that

1. for \(\lfloor 1/P_Y \rfloor\) sequences \(y\), \(p_Y^*(y) = P_Y\), for another one sequence \(y_0\), \(p_Y^*(y_0) = 1 - \lfloor 1/P_Y \rfloor P_Y\), and for all other sequences \(y\), \(p_Y^*(y) = 0\);
2. $\sum_{x,y} p_X(x)p_Y^*(y)x^Ty \leq \sum_{x,y} p_X(x)p_Y(y)x^Ty \leq nc$; and

3. $(\max_x p_X(x) \max_y p_Y^*(y))^{1/n} = (P_X P_Y)^{1/n}$.

Let the support set of distributions $p_Y^*$ be $S_Y$, and let $\bar{p}_Y$ be the uniform distribution over $S_Y \setminus \{y_0\}$. Similar to the reasoning of (A10), we have

$$\sum_{x,y} p_X(x)\bar{p}_Y(y)x^Ty \leq \sum_{x,y} p_X(x)p_Y^*(y)x^Ty \leq nc.$$ 

Again, according to Lemma 3,

$$P_n^\dagger \leq P'_n \leq \left( P_X \max_y \bar{p}_Y(y) \right)^{1/n} = \left( P_X \frac{1}{[1/P_Y]} \right)^{1/n} \leq \left( P_X \frac{1}{[1/P_Y] - 1} \right)^{1/n} \leq (3P_X P_Y)^{1/n} = \sqrt{3} P_n^\dagger,$$

and hence $\lim_{n \to \infty} P'_n / P_n^\dagger = 1$.

**Lemma 3.** For every $x \in (0, 1)$,

$$x([1/x] - 1) \geq \frac{1}{3}.$$ 

**Proof.** If $x \geq \frac{1}{3}$, then

$$x([1/x] - 1) \geq x \geq \frac{1}{3}.$$ 

If $x < \frac{1}{3}$, then

$$x([1/x] - 1) \geq x(1/x - 2) \geq 1 - 2x > \frac{1}{3}.$$

\[ \square \]

**C. Proof of Theorem 3**

We first show that we only need to consider $S_X$ and $S_Y$ with profiles $a, b \in [0, 0.5]^n$. Suppose that for some $i$ we have $a_i > \frac{1}{2}$. We obtain a new set $S'_X$ by flipping the $i$-th bit of all vectors in $S_X$. Let $a' = \Gamma(S'_X)$. We have $a'_k = a_k$ for $k \neq i$ and $a'_i = 1 - a_i$. We know from
Lemma 1 that for the constraint (4) still holds with $S_X'$ in place of $S_X$ since $a'_i < 0.5 < a_i$.

While the objective function of Problem 2 with $S_X'$ in place of $S_X$ does not change since $|S_X'| = |S_X|$. Similarly we can modify $S_Y$ such that all $b_i \leq \frac{1}{2}$.

Without the loss of generality, we assume $a_1 \geq a_2 \geq \cdots \geq a_n$. Otherwise we just change the order of the bit in the string. Now we put $b_1, ..., b_n$ in a non-decreasing reordering as: $b'_1 \leq \cdots \leq b'_n$. There must exist set $S_Y' \subseteq \{0, 1\}^n$ such that $\Gamma(S_Y') = (b'_1, ..., b'_n)^T$ by changing the order of the bits for each string in set $S_Y$. Then we have

$$\frac{1}{n|S_X||S_Y'|} \sum_{x \in S_X', y \in S_Y'} x^T y = \sum_{i=1}^{n} a_i b'_i \leq \sum_{i=1}^{n} a_i b_i \leq c.$$ (10)

The proof is completed by $|S_X||S_Y'| = |S_X||S_Y|$.

D. Proof of Theorem 4

The logarithm in this proof has base 2. Consider subset $S \subset \{0, 1\}^n$ with $\Gamma(S) \leq a$. Define a random variable $X$ over $\{0, 1\}^n$ with support $S$ and $\Pr\{X = x\} = \frac{1}{|S|}$ for each $x \in S$. Recall that the $i$-th component of $x \in \{0, 1\}^n$ is denoted by $x_i$. Let $l_k = \lfloor \frac{kn}{m} \rfloor$ for $k = 0, 1, \ldots, m$. Since $E[X] = \Gamma(S) \leq a$, we have for $k = 1, \ldots, m$ and $i = 1, \ldots, l_k - l_{k-1}$, $E[X_{l_{k-1}+i}] \leq a_k$. Note that $X_i$ is a binary random variable. Hence the entropy $H(X_{l_{k-1}+i}) \leq h_b(a_k)$ for $k = 1, \ldots, m$ and $i = 1, \ldots, l_k - l_{k-1}$. Therefore,

$$\log |S| = H(X) \leq \sum_{k=1}^{m} \sum_{i=1}^{l_k-l_{k-1}} H(X_{l_{k-1}+i})$$

$$\leq \sum_{k=1}^{m} (l_k - l_{k-1}) h_b(a_k)$$

$$\leq \frac{n}{m} \left( \sum_i h_b(a_i) + o(1) \right),$$

where the last inequality follows that $l_k - l_{k-1} \leq \frac{n}{m} + 1$ and $o(1)$ tends to zero as $n$ tends to $\infty$. Since the above inequality holds for all subset $S \subset \{0, 1\}^n$ with $\Gamma(S) \leq a$, we have

$$V_n(a) \leq 2^{\frac{n}{m} \left( \sum_i h_b(a_i) + o(1) \right)}.$$
E. Proof of Theorem $\ref{thm:proof}$

Let $a = \Gamma(S_X)$, $b = \Gamma(S_Y)$. By Theorem $\ref{thm:gamma}$ it is sufficient for us to consider $S_X$ and $S_Y$ such that $0.5 \geq a_1 \geq \ldots \geq a_n \geq 0$ and $0 \leq b_1 \leq \ldots \leq b_n \leq 0.5$. Hence $f_a$ is decreasing on $[0, 1]$, and $f_b$ is increasing on $[0, 1]$.

Define two $m$-profiles $\bar{a}$ and $a$ such that for $1 \leq i \leq m$,

$$\bar{a}_i = \left\lceil mf_a \left(\frac{i-1}{m}\right) \right\rceil, \quad a_i = \left\lfloor mf_a \left(\frac{i}{m}\right) \right\rfloor.$$

We have $f_{\bar{a}}$ and $f_a$ are decreasing on $[0, 1]$.

**Lemma 4.** $a \leq a \leq \bar{a}$. 

**Proof.** Notice that $f_a$ is a decreasing function. For every $0 \leq r \leq 1$,

$$f_a(r) = \bar{a}_{\lceil rm \rceil} \geq f_a \left(\frac{\lceil rm \rceil - 1}{m}\right) \geq f_a(r),$$

and similarly,

$$f_a(r) = a_{\lceil rm \rceil} \leq f_a \left(\frac{\lceil rm \rceil}{m}\right) \leq f_a(r).$$

Thus $a \leq a \leq \bar{a}$. \hfill $\square$

Define two $m$-profiles $\bar{b}$ and $b$ such that for $1 \leq i \leq m$,

$$\bar{b}_i = \left\lceil mf_b \left(\frac{i}{m}\right) \right\rceil, \quad b_i = \left\lfloor mf_b \left(\frac{i}{m}\right) \right\rfloor.$$

We have $f_b$ and $f_b$ are increasing on $[0, 1]$, and similar to Lemma 4, we have the following lemma.

**Lemma 5.** $b \leq b \leq \bar{b}$. 

Now we can prove the following lemma.

**Lemma 6.** For $m \geq 2$,

$$\frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \frac{1}{m} \sum_{i=1}^{n} a_i b_i < \frac{2}{m},$$

(11)
Proof. Observe that
\[
\frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \frac{1}{n} \sum_{i=1}^{n} a_i b_i = \frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \int_0^1 f_a(t) f_b(t) dt \\
\leq \frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \int_0^1 f_a(t) f_b(t) dt \\
= \frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \frac{1}{m} \sum_{i=1}^{m} a_i b_i.
\]

By definition, we have for \(1 \leq i \leq m-1\), \(ma_i \geq m\bar{a}_{i+1} - 1\) and \(mb_i \geq m\bar{b}_i - 1\). Hence
\[
\frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \frac{1}{m} \sum_{i=1}^{m} a_i b_i 
\leq \frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i - \frac{1}{m} \sum_{i=1}^{m} \left( \bar{a}_{i+1} - \frac{1}{m} \right) \left( \bar{b}_{i-1} - \frac{1}{m} \right) \\
= \frac{1}{m} \left( \bar{a}_1 \bar{b}_1 + \bar{a}_2 \bar{b}_2 + \sum_{i=3}^{m} \bar{a}_i (\bar{b}_i - \bar{b}_{i-2}) + \sum_{i=2}^{m-1} \left( \frac{\bar{a}_{i+1}}{m} + \frac{\bar{b}_{i-1}}{m} - \frac{1}{m^2} \right) \right) \\
\leq \frac{1}{m} \left( 0.25 + 0.25 + \sum_{i=3}^{m} 0.5(\bar{b}_i - \bar{b}_{i-2}) + \sum_{i=2}^{m-1} \left( \frac{0.5}{m} + \frac{0.5}{m} \right) \right) \\
= \frac{1}{m} \left( 1.5 - \frac{2}{m} + 0.5\bar{b}_m + 0.5\bar{b}_{m-1} - 0.5\bar{b}_2 - 0.5\bar{b}_1 \right) \\
\leq \frac{2}{m},
\]
where we use the fact that \(\bar{a}_i, \bar{b}_i \leq 0.5\).

By Lemma 6 and the condition of the theorem (using the form given in Lemma 1), we have
\[
\frac{1}{m} \sum_{i=1}^{m} \bar{a}_i \bar{b}_i \leq \frac{1}{n} \sum_{i=1}^{n} a_i b_i + \frac{2}{m} \leq c_Q + \frac{2}{m}. \tag{12}
\]

From Lemma 2 and Theorem 4 we know that
\[|S_X||S_Y| = V_n(a)V_n(b) \leq V_n(\bar{a})V_n(\bar{b}) \leq 2^{h_a} (\sum_{i=1}^{m} (h_a(a_i) + h_b(b_i)) + o(1)),\]

where \(o(1) \to 0\) as \(n \to \infty\). For \(0 \leq t \leq 0.25\), define
\[f(t) = \max_{2t \leq x \leq \frac{1}{2}} \left( h_b(x) + h_b \left( \frac{t}{x} \right) \right). \tag{13}\]

Some properties of the above function are given in Appendix B (see Lemma 7–9). Applying Lemma 9 and (12),
\[
\frac{1}{m} \sum_{i=1}^{m} (h_b(\bar{a}_i) + h_b(\bar{b}_i)) \leq \frac{1}{m} \sum_{i=1}^{m} f(\bar{a}_i, \bar{b}_i) \leq f \left( c_Q + \frac{2}{m} \right). \tag{14}
\]
Thus for any sufficiently large \( m \),
\[
\liminf_{n \to \infty} \frac{1}{\sqrt{|S_X||S_Y|}} \geq 2^{-f(c_Q + \frac{2}{m})}.
\]

Take \( m \to \infty \) we have
\[
\liminf_{n \to \infty} \frac{1}{\sqrt{|S_X||S_Y|}} \geq 2^{-f(c_Q)} = 4^{-h_b(\sqrt{c_Q})},
\]
where the last equality is implied by Lemma 7.

F. Proof of Theorem 6

For every \( n \), let \( S_X = S_Y = \{ x \in \{0, 1\}^n : x \text{ includes at most } n\sqrt{c_Q} \text{ 1s} \} \). Then
\[
|S_X| = |S_Y| = \sum_{i=0}^{\lfloor n\sqrt{c_Q} \rfloor} \binom{n}{i} = 2^{n(h_b(\sqrt{c_Q}) + o(1))},
\]
where \( o(1) \to 0 \) as \( n \to \infty \). Thus
\[
\lim_{n \to \infty} \frac{1}{\sqrt{|S_X||S_Y|}} = \frac{1}{2^{2h_b(\sqrt{c_Q})}} = 4^{-h_b(\sqrt{c_Q})}.
\]

From the constructions of \( S_X \) and \( S_Y \), we know that
\[
\Gamma(S_X) = \Gamma(S_Y) \leq (\sqrt{c_Q}, \sqrt{c_Q}, \cdots, \sqrt{c_Q}).
\]

Therefore
\[
\frac{1}{n} \sum_{x \in S_X, y \in S_Y} \frac{1}{|S_X||S_Y|} x^T y = \frac{1}{n} (\Gamma(S_X))^T \Gamma(S_Y) \leq \frac{1}{n} \sum_{i=1}^{n} (\sqrt{c_Q})^2 = c_Q.
\]

Thus \( S_X \) and \( S_Y \) satisfies constraints in Theorem 2.

IV. CONCLUDING REMARKS

In this paper, we determine for Problem 1 that when \( c = c_Q \)
\[
\lim_{n \to \infty} P_n = 4^{-h_b(\sqrt{c})},
\]
which is of particular interest for quantum information. Note that our technique also shows that \( (17) \) holds for \( c_Q \leq c < 1/4 \). However, the existing technique in this paper does not imply \( (17) \) for \( c < c_Q \), which holds if we can show that \( f(t) \) (defined in \( (13) \)) is concave in \( [0, 0.25] \). But we can only show the concavity of \( f(t) \) for the range \( [0.0625, 0.25] \) (see Appendix B). Whether \( f(t) \) is concave in \( [0, 0.25] \) is of certain mathematical interest.
ACKNOWLEDGMENTS

We thank Xiongfeng Ma, Xiao Yuan and Zhu Cao for introducing us this problem and providing insightful comments to our work. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61471215. This work was partially funded by a grant from the University Grants Committee of the Hong Kong Special Administrative Region (Project No. AoE/E-02/08).

Appendix A: Derivation of the Optimization Problem

A Bell test experiment has two spatially separated parties, Alice and Bob, who can randomly choose their devices settings $x$ and $y$ from set $\{0, 1\}$ and generate output bits $a$ and $b$, respectively. The Clauser-Horne-Shimony-Holt (CHSH) inequality is that

$$S := \sum_{a,b,x,y \in \{0,1\}} (-1)^{a \oplus b + xy} q(a, b|x, y) \leq 2,$$

where $\oplus$ denotes the exclusive-or of two bits, and $q(a, b|x, y)$ is the probability that outputs $a$ and $b$ are generated when the device settings are $x$ and $y$. The theory of quantum mechanics predicts a maximum value for $S$ of $S_Q = 2\sqrt{2}$.

In a local hidden variable model (LHVM), assume that an adversary Eve controls a variable $\lambda$ taking discrete values so that

$$q(a, b|x, y) = \sum_\lambda q(a|x, \lambda) q(b|y, \lambda) q(\lambda|x, y),$$

where $q(a|x, \lambda)$ (resp. $q(b|y, \lambda)$) is the probability that $a$ is output when the setting of Alice (resp. Bob) is $x$ (resp. $y$), and $q(\lambda|x, y)$ is the conditional probability distribution of the variable $\lambda$ given $x$ and $y$. Free will is assumed in the derivation of the CHSH inequality, i.e.,

$$q(\lambda|x, y) = q(\lambda).$$

With this assumption, the inequality (A1) holds for any LHVM.

We consider the case that the device settings may not be chosen freely, i.e., (A2) may not hold. By the Bayes’ law,

$$q(\lambda|x, y) = \frac{q(x, y|\lambda)q(\lambda)}{q(x, y)} = 4q(x, y|\lambda)q(\lambda),$$

16
where \( q(x, y) \) is assumed to be \( 1/4 \) so that Alice and Bob cannot detect the existence of adversary Eve. In this case,

\[
S = \sum_{\lambda} S_\lambda q(\lambda), \quad (A3)
\]

where

\[
S_\lambda = 4 \sum_{a, b, x, y \in \{0, 1\}} (-1)^{a \oplus b + xy} q(a|x, \lambda)q(b|y, \lambda)q(x, y|\lambda).
\]

The adversary can pick probabilities \( q(\lambda), q(x, y|\lambda), q(a|x, \lambda) \) and \( q(b|y, \lambda) \) to fake the violation of a Bell’s inequality.

The following randomness measure are used in literature\textsuperscript{16–18}

\[
P = \max_{x, y, \lambda} q(x, y|\lambda).
\]

Note that \( P \) takes values from \( 1/4 \) to 1. When \( P = 1/4 \), all the device settings are uniformly picked independent of \( \lambda \). When \( P = 1 \), for at least one value of \( \lambda \), the device settings are deterministic.

We are interested in the minimum value of \( P \) such that \( S \geq S_Q \) for certain LHVMs. In other words, we want to solve the following problem

\[
\begin{align*}
\min & \quad \max_{x, y, \lambda} q(x, y|\lambda) \\
\text{s.t.} & \quad \sum_{\lambda} S_\lambda q(\lambda) \geq S_Q, \\
& \quad \sum_{\lambda} q(x, y|\lambda)q(\lambda) = \frac{1}{4},
\end{align*}
\]

(A4)

where \( S \) is defined in (A3) and the minimization is over all the possible (conditional) distributions \( q(\lambda), q(x, y|\lambda), q(a|x, \lambda) \) and \( q(b|y, \lambda) \).

Due to the convexity of the constraints with respect to \( q(a|x, \lambda) \) and \( q(b|y, \lambda) \), we can consider only deterministic distributions \( q(a|x, \lambda) \) and \( q(b|y, \lambda) \) without changing the optimal value of (A4). Let \( a = a(x, \lambda) \) and \( b = b(y, \lambda) \). Rewrite

\[
S_\lambda = 4 \sum_{x, y \in \{0, 1\}} (-1)^{a(x, \lambda) \oplus b(y, \lambda) + xy} q(x, y|\lambda). \quad (A5)
\]

For a give value of \( \lambda \), there are totally 16 different pairs of the output functions \((a, b)\). Table II lists the eight possible output functions with \( a(0, \lambda) = 0 \). We do not need to consider the eight possible output functions with \( a(0, \lambda) = 1 \) since they give the same set of \( S_\lambda \) as listed in the last column in Table II. Note that for the output functions with index 1, 2, 5, 7, the corresponding \( S_\lambda \) has only one negative term, so they are better than the other
TABLE II. Output function assignment. Here $\lambda$ in $a(x, \lambda)$, $b(y, \lambda)$ and $q(x, y|\lambda)$ is omitted.

|   | $a(0)$ | $a(1)$ | $b(0)$ | $b(1)$ | $S_{\lambda}/4$ |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | $q(0,0) + q(0,1) + q(1,0) - q(1,1)$ |
| 2 | 0 | 0 | 0 | 1 | $q(0,0) - q(0,1) + q(1,0) + q(1,1)$ |
| 3 | 0 | 0 | 1 | 0 | $-q(0,0) + q(0,1) - q(1,0) - q(1,1)$ |
| 4 | 0 | 0 | 1 | 1 | $-q(0,0) - q(0,1) - q(1,0) + q(1,1)$ |
| 5 | 0 | 1 | 0 | 0 | $q(0,0) + q(0,1) - q(1,0) + q(1,1)$ |
| 6 | 0 | 1 | 0 | 1 | $q(0,0) - q(0,1) - q(1,0) - q(1,1)$ |
| 7 | 0 | 1 | 1 | 0 | $-q(0,0) + q(0,1) + q(1,0) + q(1,1)$ |
| 8 | 0 | 1 | 1 | 1 | $-q(0,0) - q(0,1) + q(1,0) - q(1,1)$ |

TABLE III. Output function assignment.

| $\lambda$ | $a(0, \lambda)$ | $a(1, \lambda)$ | $b(0, \lambda)$ | $b(1, \lambda)$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 0 | 0 | 0 | 1 |
| 4 | 0 | 1 | 1 | 0 |

Moreover, due to symmetry, we can consider one of these four assignments. Here we use $a(x, \lambda) = b(y, \lambda) = 0$.

Moreover, the constraint $\sum_{\lambda} q(x, y|\lambda)q(\lambda) = \frac{1}{4}$ is redundant. To illustrate the idea, consider a LHVM with $\lambda = 0$ a constant and $a(x, 0) = b(y, 0) = 0$. We use $q(x, y|0)$ to denote the device setting distribution related to this LHVM. Define a new LHVM with $\lambda = 1, 2, 3, 4$ and $q(\lambda) = 1/4$: The output functions are assigned according to Table III and the device setting distributions are assigned according to Table IV. It can be verified that the value of $P$ and $S$ are the same for both LHVMs and $q(x, y) = 1/4$.

With these simplifications, the above minimization problem becomes

$$\min \max_{x,y,\lambda} q(x, y|\lambda)$$

s.t. $4 \sum_{\lambda}(1 - 2q(1, 1|\lambda))q(\lambda) \geq S_Q$, \hspace{1cm} (A6)

Last, it is sufficient to consider a constant $\lambda$ for the above optimization. For any distribution $q$ with multiple $\lambda$’s achieving the optimal value of (A6), let $q'$ be a distribution with
a single \( \lambda \) such that \( q'(x, y) = q(x, y) \) for any \( x, y \). Observe that \( \max_{x, y} \max_{\lambda} q(x, y|\lambda) \geq \max_{x, y} q(x, y) = \max_{x, y} q'(x, y) \), we have \( q' \) optimizes the minimization problem with constant \( \lambda \). The above optimization now becomes

\[
\min \max_{x, y} q(x, y) \\
s.t. \quad 4 - 8q(1, 1) \geq S_Q,
\]  

(A7)

In the above formulations, only a single run of the test is performed. It is more realistic to consider that the device settings in different runs are correlated, which is referred to as the multiple-run scenario, where the device settings \( \mathbf{x} = (x_1, \ldots, x_n)^T \) and \( \mathbf{y} = (y_1, \ldots, y_n)^T \) in \( n \) runs of the tests follow a joint distribution \( q(\mathbf{x}, \mathbf{y}|\lambda) \). Similar to the discussion of the single-run scenario, for multiple runs, we have \( S = \sum_{\lambda} S_\lambda q(\lambda) \) with

\[
S_\lambda = \frac{4}{n} \sum_{\mathbf{x}, \mathbf{y} \in \{0, 1\}^n} q(\mathbf{x}, \mathbf{y}|\lambda) \sum_{i=1}^{n} (-1)^{a(x_i, \lambda) \oplus b(y_i, \lambda) + x_i y_i} \\
= 4 \sum_{\mathbf{x}, \mathbf{y} \in \{0, 1\}^n} q(\mathbf{x}, \mathbf{y}|\lambda) \left[ \pi(0, 0|\mathbf{x}, \mathbf{y})(-1)^{a(0, \lambda) \oplus b(0, \lambda)} + \pi(0, 1|\mathbf{x}, \mathbf{y})(-1)^{a(0, \lambda) \oplus b(1, \lambda)} \\
+ \pi(1, 0|\mathbf{x}, \mathbf{y})(-1)^{a(1, \lambda) \oplus b(0, \lambda)} + \pi(1, 1|\mathbf{x}, \mathbf{y})(-1)^{a(1, \lambda) \oplus b(1, \lambda) + 1} \right], \\
= 4 \sum_{x, y \in \{0, 1\}} (-1)^{a(x, \lambda) \oplus b(y, \lambda) + xy} \pi(x, y|\lambda)
\]  

(A8)

where \( \pi(x, y|\mathbf{x}, \mathbf{y}) \) is the fraction of \((x, y)\) pairs among the pairs \((x_k, y_k), k = 1, \ldots, n\), and \( \pi(x, y|\lambda) = \sum_{\mathbf{x}, \mathbf{y} \in \{0, 1\}^n} q(\mathbf{x}, \mathbf{y}|\lambda) \pi(x, y|\mathbf{x}, \mathbf{y}) \). Note that (A8) shares the same form as (A5).

Define the measure of measurement dependence for multiple runs as

\[
P = \left( \max_{\mathbf{x}, \mathbf{y}, \lambda} q(\mathbf{x}, \mathbf{y}|\lambda) \right)^{1/n}.
\]
The problem of interest now becomes

\[
\begin{align*}
&\min \left(\max_{x,y,\lambda} q(x, y | \lambda)\right)^{1/n} \\
&\text{s.t. } \sum_{\lambda} S_\lambda q(\lambda) \geq S_Q \\
&\quad \sum_{\lambda} q(x, y | \lambda) q(\lambda) = \frac{1}{n},
\end{align*}
\]

(A9)

where \(S_\lambda\) is defined in (A8). Note that when \(n = 1\), (A9) becomes (A4). Similar to the reasoning in single-run case, we only need to consider an LHVM with one deterministic strategy \(\lambda\) where \(a(x, \lambda) = b(y, \lambda) = 0\), and simplify problem (A9) to

\[
\begin{align*}
&\min \left(\max_{x,y} q(x, y)\right)^{1/n} \\
&\text{s.t. } \frac{1}{n} \sum_{x,y\in\{0,1\}^n} q(x, y)x^T y \leq \frac{4 - S_Q}{8}.
\end{align*}
\]

(A10)

Appendix B: Properties of a Function

We study some properties of the function \(f(t)\) defined in (13). Recall that

\[f(t) = \max_{2t \leq x \leq \frac{1}{4}} \left(h_b(x) + h_b\left(\frac{t}{x}\right)\right), \quad 0 \leq t \leq 0.25.\]

The next lemma implies that \(f(t) = 2h_b(\sqrt{t})\) for \(0.0625 \leq t \leq 0.25\).

**Lemma 7.** For \(0.0625 \leq t \leq 0.25\), \(2t \leq x \leq 0.5\), we have

\[h_b(x) + h_b\left(\frac{t}{x}\right) \leq 2h_b(\sqrt{t}),\]

where the equality holds for \(x = \sqrt{t}\).

**Proof.** Fix \(t\). Let \(u(x) = h_b(x) + h_b\left(\frac{t}{x}\right)\). Observe that \(u(x) = u\left(\frac{t}{x}\right)\). Thus it suffices to show \(u(x) \leq 2h_b(\sqrt{t})\) for \(2t \leq x \leq \sqrt{t}\). Taking derivative on \(u\) we have

\[u'(x) = -\log x + \log(1-x) + \frac{t}{x^2} \log\left(\frac{t}{x}\right) - \frac{t}{x^2} \log\left(1 - \frac{t}{x}\right)\]

Let \(v(x) = -x \log x + x \log(1-x)\), we have

\[xu'(x) = v(x) - v\left(\frac{t}{x}\right),\]

(B1)
From $t \geq \frac{1}{16}$ we have

$$\frac{t}{x} \geq \frac{1}{2} - x \geq \frac{1}{4}. \quad (B2)$$

We may verify that $v$ is decreasing on $[0.25, 0.5]$. If $x \geq 0.25$, then $xu'(x) \geq 0$ since $x \leq \frac{t}{x}$. Otherwise, we may verify $v(x) \geq v(0.5 - x)$ for $x \leq 0.25$. Then apply (B2) to (B1) we have

$$vu'(x) = v(x) - v\left(\frac{t}{x}\right) \geq v(x) - v(0.5 - x) \geq 0 \quad (B3)$$

Therefore $u$ is an increasing function on $[2t, \sqrt{t}]$, which implies $u(x) \leq 2h_b(\sqrt{t})$. \hfill \square

**Lemma 8.** Function $f(t)$ is increasing on $[0, 0.25]$.

**Proof.** To show that $f$ is increasing, fix any $0 \leq t_1 < t_2 \leq 0.25$. We write $f(t_1) = h_b(x_1) + h_b(y_1)$ where $x_1$ maximizes $h_b(x) + h_b\left(\frac{t_1}{x}\right)$ for $x \in [2t_1, 0.5]$ and $x_1y_1 = t_1$. We know that $0 \leq x_1, y_1 \leq 0.5$. Find $x_2$ and $y_2$ such that $x_1 \leq x_2 \leq \frac{1}{2}, y_1 \leq y_2 \leq \frac{1}{2}$ such that $x_2y_2 = t_2$. Therefore

$$f(t_1) = h_b(x_1) + h_b(y_1) \leq h_b(x_2) + h_b(y_2) \leq f(t_2). \quad \square$$

**Lemma 9.** For any $c' \geq c_Q = \frac{2 - \sqrt{2}}{4} \approx 0.1464$, if $k$ real numbers $t_1, t_2, \ldots, t_k \in [0, 0.25]$ such that $\frac{1}{k} \sum_{i=1}^{k} t_i \leq c'$, we have

$$\frac{1}{k} \sum_{i=1}^{k} f(t_i) \leq f(c').$$

**Proof.** Let $f_0(t) = 2h_b\left(\sqrt{t}\right), 0 \leq t \leq 0.25$. From Lemma 7 $f(t) = f_0(t)$ for $t \geq 0.0625$. Let $f_1$ be the tangent line of $f_0$ on $(0.14, f_0(0.14))$. Notice that $h_b(x)$ and $\sqrt{x}$ are both concave on their domains. We see that $f_0(t)$ is also concave on $[0, 0.25]$. Observe that $f_0$ is concave and increasing on $[0, \frac{1}{4}]$, we have $f_1$ is an increasing function, while for every $t \in [0, 0.25]$, $f_0(t) \leq f_1(t)$.

Let $g(t)$ be a function defined on $[0, 0.25]$ such that

$$g(t) = \begin{cases} f_1(t) & 0 \leq t \leq 0.14; \\ f_0(t) & 0.14 < t \leq 0.25. \end{cases}$$

Observe that $g$ is linear on $[0, 0.14]$ and concave on $[0.14, 0.25]$, thus $g$ is concave on $[0, 0.25]$. For $0 \leq t < 0.0625$,

$$f(t) \leq f(0.0625) = f_0(0.0625) = 1.623 < 1.630 = g(0) \leq g(t).$$
For $0.0625 \leq t \leq 0.25$,

$$f(t) = f_0(t) \leq g(t).$$

Thus $g$ is always not smaller than $f$. Take $t'_1, t'_2, \ldots, t'_k \leq 0.25$ such that $t_i \leq t'_i$ for all $1 \leq i \leq k$, while $\frac{1}{k} \sum_{i=1}^{k} t'_i = c'$. Applying Jensen’s inequality we have

$$\frac{1}{k} \sum_{i=1}^{k} f(t_i) \leq \frac{1}{k} \sum_{i=1}^{k} f(t'_i) \leq \frac{1}{k} \sum_{i=1}^{k} g(t'_i) \leq g \left( \frac{1}{k} \sum_{i=1}^{k} t'_i \right) = g(c') = f(c'),$$

where the first inequality holds since $f$ is increasing, the second inequality holds since $g$ is always no less than $f$, and the last equality follows that $c' \geq c_Q > 0.14$.

REFERENCES

1. J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden-variable theories,” Phys. Rev. Lett. 23, 880–884 (1969).
2. J. S. Bell, “On the einstein-podolsky-rosen paradox,” Physics 1, 195–200 (1964).
3. A. Aspect, P. Grangier, and G. Roger, “Experimental realization of einstein-podolsky-rosen-bohm Gedankenexperiment: A new violation of Bell’s inequalities,” Phys. Rev. Lett. 49, 91–94 (1982).
4. M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger, “Bell violation using entangled photons without the fair-sampling assumption,” Nature 497, 227–230 (2013).
5. M. Feldmann, “New loophole for the einstein-podolsky-rosen paradox,” Found. Phys. Lett. 8, 41–53 (1995).
6. J. Kofler, T. Paterek, and i. c. v. Brukner, “Experimenter’s freedom in Bell’s theorem and quantum cryptography,” Phys. Rev. A 73, 022104 (2006).
7. M. J. W. Hall, “Local deterministic model of singlet state correlations based on relaxing measurement independence,” Phys. Rev. Lett. 105, 250404 (2010).
8. J. Barrett and N. Gisin, “How much measurement independence is needed to demonstrate nonlocality?” Phys. Rev. Lett. 106, 100406 (2011).
9. M. J. W. Hall, “Relaxed Bell inequalities and kochen-specker theorems,” Phys. Rev. A 84, 022102 (2011).
D. Mayers and A. Yao, “Quantum cryptography with imperfect apparatus,” in Proceedings of the 39th Annual Symposium on Foundations of Computer Science, FOCS ’98 (IEEE Computer Society, Washington, DC, USA, 1998) pp. 503–509.

A. Acín, N. Gisin, and L. Masanes, “From bell’s theorem to secure quantum key distribution,” Phys. Rev. Lett. 97, 120405 (2006).

U. Vazirani and T. Vidick, “Fully device-independent quantum key distribution,” Phys. Rev. Lett. 113, 140501 (2014).

S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, “Random numbers certified by Bell’s theorem,” Nature 464, 1021–1024 (2010).

R. Colbeck and R. Renner, “Free randomness can be amplified,” Nat Phys 8, 450–453 (2012).

C. Dhara, G. de la Torre, and A. Acín, “Can observed randomness be certified to be fully intrinsic?” Phys. Rev. Lett. 112, 100402 (2014).

D. E. Koh, M. J. W. Hall, Setiawan, J. E. Pope, C. Marletto, A. Kay, V. Scarani, and A. Ekert, “Effects of reduced measurement independence on Bell-based randomness expansion,” Phys. Rev. Lett. 109, 160404 (2012).

J. E. Pope and A. Kay, “Limited measurement dependence in multiple runs of a Bell test,” Phys. Rev. A 88, 032110 (2013).

X. Yuan, Z. Cao, and X. Ma, “Randomness requirement on CHSH Bell test in the multiple run scenario,” arXiv preprint arXiv:1409.7875 (2014).

J. Gallicchio, A. S. Friedman, and D. I. Kaiser, “Testing bell’s inequality with cosmic photons: Closing the setting-independence loophole,” Phys. Rev. Lett. 112, 110405 (2014).