Three-input majority function as the unique optimal function for the bias amplification using nonlocal boxes

Ryuhei Mori*
School of Computing, Tokyo Institute of Technology, Tokyo 152-8552, Japan
(Received 21 April 2016; revised manuscript received 25 August 2016; published 28 November 2016)

Brassard *et al.* [Phys. Rev. Lett. **96**, 250401 (2006)] showed that shared nonlocal boxes with a CHSH (Clauser, Horne, Shimony, and Holt) probability greater than $\frac{3+\sqrt{6}}{6}$ yield trivial communication complexity. There still exists a gap with the maximum CHSH probability $\frac{2+\sqrt{2}}{4}$ achievable by quantum mechanics. It is an interesting open question to determine the exact threshold for the trivial communication complexity. Brassard *et al.*'s idea is based on recursive bias amplification by the three-input majority function. It was not obvious if another choice of function exhibits stronger bias amplification. We show that the three-input majority function is the unique optimal function, so that one cannot improve the threshold $\frac{3+\sqrt{6}}{6}$ by Brassard *et al.*'s bias amplification. In this work, protocols for computing the function used for the bias amplification are restricted to be nonadaptive protocols or a particular adaptive protocol inspired by Pawłowski *et al.*'s protocol for information causality [Nature (London) **461**, 1101 (2009)]. We first show an adaptive protocol inspired by Pawłowski *et al.*'s protocol, and then show that the adaptive protocol improves upon nonadaptive protocols. Finally, we show that the three-input majority function is the unique optimal function for the bias amplification if we apply the adaptive protocol to each step of the bias amplification.

DOI: 10.1103/PhysRevA.94.052130

I. INTRODUCTION

Bell showed that quantum mechanics allows correlations that cannot be generated by classical physics [1]. Clauser, Horne, Shimony, and Holt (CHSH) found simpler constraints on correlations which could be violated in quantum mechanics, but are always satisfied in classical physics [2], and which in fact characterize the set of correlations generated by classical physics on the binary setting [3]. Apart from the concrete mathematical description of quantum mechanics, we can only consider abstract statistical behavior realized by quantum mechanics. A nonlocal box is an abstract device which represents statistical behavior of separable measurements on a possibly entangled state in quantum mechanics and superquantum theory as well. A nonlocal box is assumed to be shared by two parties, Alice and Bob. A nonlocal box has input ports and output ports on both sides. A nonlocal box is specified by the conditional probability distribution $p(a,b \mid x,y)$ representing the probability of outputting a to Alice and b to Bob when Alice and Bob input x and y into the nonlocal box, respectively. Here, all of x, y, a, and b are assumed to be either of 0 or 1. They cannot communicate by using the nonlocal box since it satisfies the no-signaling condition

$$\sum_{b\in\{0,1\}} p(a,b \mid x,0) = \sum_{b\in\{0,1\}} p(a,b \mid x,1),$$

$$\sum_{a\in\{0,1\}} p(a,b \mid 0,y) = \sum_{a\in\{0,1\}} p(a,b \mid 1,y).$$

The CHSH probability P_{CHSH} is a measure of the nonlocality of the nonlocal box, defined by

$$P_{\mathrm{CHSH}} := \frac{1}{4}\sum_{x,y}\sum_{\substack{a,b\\ a\oplus b=x\wedge y}} p(a,b \mid x,y).$$

While the maximum CHSH probability given by classical physics is $P_{\mathrm{CHSH}} = 3/4$, that for quantum mechanics is $P_{\mathrm{CHSH}} = \frac{2+\sqrt{2}}{4}$ [4]. On the other hand, Popescu and Rohrich showed that there exists a nonlocal box, called the PR box, with $P_{\mathrm{CHSH}} = 1$ [5]. Hence, it is a natural question why quantum mechanics cannot achieve a CHSH probability greater than $\frac{2+\sqrt{2}}{4}$. Van Dam showed that if Alice and Bob share an unlimited number of PR boxes, they can compute an arbitrary function $f(x,y)$ only by sending 1 bit to each other where x and y are n bits owned by Alice and Bob, respectively [6]. This explains why nature does not allow $P_{\mathrm{CHSH}} = 1$, since we strongly believe that trivial communication complexity must not be allowed by nature. Furthermore, Brassard *et al.* showed that a nonlocal box with $P_{\mathrm{CHSH}} > \frac{3+\sqrt{6}}{6}$ yields trivial communication complexity in a probabilistic setting [7]. It has not been known whether or not the communication complexity is trivial when the CHSH probability is between $\frac{2+\sqrt{2}}{4}$ and $\frac{3+\sqrt{6}}{6}$. Later, Pawłowski *et al.* completely characterized the quantum CHSH probability $\frac{2+\sqrt{2}}{4}$ by using a principle called information causality [8]. However, it is still interesting to determine the exact threshold of P_{CHSH} for trivial communication complexity.

In this paper, we show that trivial communication complexity below $\frac{3+\sqrt{6}}{6}$ cannot be proved by Brassard *et al.*'s technique. Their technique is based on recursive bias amplification from exponentially small bias to constant bias by using the three-input majority function Maj_3. It was not obvious that Maj_3 is the best choice for the bias amplification. In this paper,

*mori@c.titech.ac.jp

we show that Maj_3 is the unique optimal function for the bias amplification.

Theorem 1. The three-input majority function is the unique optimal function for Brassard *et al.*'s technique of bias amplification using nonlocal boxes. Hence, one cannot obtain a threshold for trivial communication complexity smaller than $\frac{3+\sqrt{6}}{6}$ by Brassard *et al.*'s technique.

In Brassard *et al.*'s protocol, the three-input majority function Maj_3 is computed by a nonadaptive protocol; i.e., inputs for nonlocal boxes are independent of outputs of other nonlocal boxes. In this work, we introduce an adaptive protocol inspired by [8], and show that the adaptive protocol is no worse than an arbitrary noadaptive protocol. Then, we show Theorem 1 for generalizations of Brasssard *et al.*'s protocol in which an arbitrary Boolean function is used for the bias amplification in place of Maj_3, and is computed by the adaptive protocol. In this work, protocols for the computation of the function corresponding to Maj_3 are restricted to be nonadaptive protocols or the adaptive protocol inspired by [8]. For the proof of Theorem 1, we use the Fourier analysis of Boolean functions developed in theoretical computer science [9].

II. PRELIMINARIES

A. XOR protocol and nonlocal boxes

We introduce some notions and notations.

Definition 2. For a Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, the XOR protocol with bias ϵ is a process of computations by Alice and Bob in which Alice and Bob compute a and b, respectively, by using nonlocal boxes and shared random bits but without any communication, such that $a \oplus b = f(x,y)$ with probability $(1+\epsilon)/2$.

There is a simple XOR protocol with bias 2^{-n} for an arbitrary function [7].

Lemma 3. There is an XOR protocol with bias 2^{-n} for arbitrary function $f(x,y)$ without using nonlocal boxes.

Proof. Let $r \in \{0,1\}^n$ be shared uniform random bits. Let $a = f(x,r)$. Let $b = 0$ if $r = y$ and $b = r'$ otherwise, where $r' \in \{0,1\}$ is Bob's private uniform random bit. Then, $a \oplus b = f(x,y)$ with probability $\frac{1}{2} + \frac{1}{2^{n+1}}$. ■

Definition 4. The nonlocal box is said to be isotropic if

$$\sum_{\substack{a,b \\ a\oplus b = x\wedge y}} p(a,b \mid x,y)$$

does not depend on x and y and if the marginal distributions for a and b are uniform for any x and y.

It was shown in Refs. [10,11] that the isotropic nonlocal box can be simulated by an arbitrary nonlocal box with the same CHSH probability.

Lemma 5. Using an arbitrary given nonlocal box, the isotropic nonlocal box with the same CHSH probability can be simulated.

From Lemma 5, in this study we assume that all nonlocal boxes are isotropic. Forster *et al.* showed that non-isotropic nonlocal boxes can be used for the nonlocality distillation, which is the amplification of the CHSH probability [12]. Brunner and Skrzypczyk showed that there exists a nonisotropic nonlocal box with $P_{\mathrm{CHSH}} = 3/4 + \epsilon$ for arbitrary small $\epsilon > 0$, which allows the simulation of a nonlocal box arbitrarily close to the PR box [13]. Of course, such a nonlocal box cannot be simulated in quantum mechanics even if the CHSH probability of the nonlocal box is achievable by quantum mechanics. In this study, we do not consider the nonlocality distillation, but consider the XOR protocol using isotropic nonlocal boxes.

B. Fourier analysis

Fourier analysis is the main mathematical tool in this work.

Definition 6. Any Boolean function $f : \{+1,-1\}^n \to \{+1,-1\}$ can be represented by a polynomial on $\mathbb{R}$ uniquely:

$$f(x) = \sum_{S\subseteq[n]} \widehat{f}(S) \prod_{i\in S} x_i$$

where $[n] := \{1,2,\ldots,n\}$. Here, $(\widehat{f}(S))_{S\subseteq[n]}$ are called the Fourier coefficients of f. When we consider the Fourier coefficients of Boolean function $f : \{0,1\}^n \to \{0,1\}$, we regard f as the function from $\{+1,-1\}^n$ to $\{+1,-1\}$. From Parseval's identity, the sum of squares of the Fourier coefficients is 1.

Let $\mathrm{supp}(\widehat{f}) := \{S \subseteq [n] \mid \widehat{f}(S) \neq 0\}$. For $S \subseteq [n]$, let 1_S be a vector on $\mathbb{F}_2$ of length n such that ith element of 1_S is 1 iff $i \in S$. Let $\dim(\widehat{f})$ be the Fourier dimension of f which is the dimension of linear space on $\mathbb{F}_2$ spanned by $\{1_S \mid S \in \mathrm{supp}(\widehat{f})\}$.

C. One-way communication complexities

We introduce notions on the one-way communication complexity of $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Let M_f be a $2^n \times 2^n$ matrix whose (x,y) element is $f(x,y)$. Let $D_{\to}(f)$ be the one-way communication complexity of f from Alice to Bob, which is the minimum m such that there exist functions $s : \{0,1\}^n \to \{0,1\}^m$ and $h : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}$ satisfying the identity $f(x,y) = h(s(x),y)$. Similarly, let $D_{\leftarrow}(f)$ be the one-way communication complexity of f from Bob to Alice. The one-way communication complexities can be characterized by the number of distinct rows and columns of M_f, i.e., $D_{\to}(f) = \lceil \log_2 n_{\mathrm{rows}}(M_f) \rceil$ and $D_{\leftarrow}(f) = \lceil \log_2 n_{\mathrm{cols}}(M_f) \rceil$, where $n_{\mathrm{rows}}(M_f)$ and $n_{\mathrm{cols}}(M_f)$ denote the number of distinct rows and the number of distinct columns of M_f, respectively. We also define

$$D^{\oplus}_{\to}(f) := \min_{A:\{0,1\}^n\to\{0,1\}} D_{\to}(f(x,y) \oplus A(x)),$$

$$D^{\oplus}_{\leftarrow}(f) := \min_{B:\{0,1\}^n\to\{0,1\}} D_{\leftarrow}(f(x,y) \oplus B(y)).$$

Here, $D^{\oplus}_{\to}(f)$ is the minimum number of bits Alice has to send to Bob such that Alice can compute a and Bob can compute b satisfying $a \oplus b = f(x,y)$.

D. Other notations

For odd n, let $\mathrm{Maj}_n : \{0,1\}^n \to \{0,1\}$ be the majority function on n variables. For even n, let Maj_n be the set of majority functions on n variables where the definitions for the tie cases are arbitrary. Since there are $\binom{n}{\frac{n}{2}}$ tie cases,

$|\mathrm{Maj}_n| = 2^{\binom{n}{\frac{n}{2}}}$ for even n. Note that a function $f : \{0,1\}^{2k} \to \{0,1\}$ which ignores one of the $2k$ input variables and outputs the majority of the other $2k-1$ variables is a member of Maj_{2k}. Finally, let $\delta := 2P_{\mathrm{CHSH}} - 1$, i.e., $P_{\mathrm{CHSH}} = \frac{1+\delta}{2}$. Here, we call δ the bias of the CHSH probability.

III. BRASSARD *et al.*'s PROTOCOL

Brassard *et al.*'s basic idea is bias amplification by Maj_3. They showed that Maj_3 can be computed by using two PR boxes. Here, we give a simple argument showing that two PR boxes are sufficient to compute $\mathrm{Maj}_3(x \oplus y)$. The $\mathbb{F}_2$-polynomial representation of the three-input majority function is $\mathrm{Maj}_3(z_1,z_2,z_3) = z_1z_2 \oplus z_2z_3 \oplus z_3z_1$. Hence, one obtains the representation

$$\begin{aligned}&\mathrm{Maj}_3(x_1 \oplus y_1, x_2 \oplus y_2, x_3 \oplus y_3)\\ &\quad = (x_1 \oplus x_2)(y_2 \oplus y_3) \oplus (x_2 \oplus x_3)(y_1 \oplus y_2)\\ &\qquad \oplus (x_1x_2 \oplus x_2x_3 \oplus x_3x_1) \oplus (y_1y_2 \oplus y_2y_3 \oplus y_3y_1). \quad (1)\end{aligned}$$

The following is the protocol for computing a and b. Alice and Bob can compute their local terms $\mathrm{Maj}_3(x) := x_1x_2 \oplus x_2x_3 \oplus x_3x_1$ and $\mathrm{Maj}_3(y) := y_1y_2 \oplus y_2y_3 \oplus y_3y_1$ without communication, respectively. For the each of first two terms in Eq. (1), they use the PR box. For the first PR box, Alice and Bob input $x_1 \oplus x_2$ and $y_2 \oplus y_3$ and obtain a_1 and b_1, respectively. Similarly, for the second PR box, Alice and Bob input $x_2 \oplus x_3$ and $y_1 \oplus y_2$ and obtain a_2 and b_2, respectively. Then, Alice and Bob output $a := \mathrm{Maj}_3(x) \oplus a_1 \oplus a_2$ and $b := \mathrm{Maj}_3(y) \oplus b_1 \oplus b_2$, respectively. This is the XOR protocol without error using two PR boxes. Von Neumann showed that the probability of correctness of computations sufficiently close to $1/2$ is amplified by noisy Maj_3 iff the computation of Maj_3 succeeds with probability greater than $5/6$ [14]. Hence, the threshold of the above protocol is given by the condition $P_{\mathrm{CHSH}}^2 + (1 - P_{\mathrm{CHSH}})^2 > 5/6 \iff P_{\mathrm{CHSH}} > \frac{3+\sqrt{6}}{6}$. Under this condition, the iterative applications of Maj_3 to independent samples obtained by the protocol in Lemma 3 give a constant bias.

Brassard *et al.* invented the above elegant protocol, and showed that if $P_{\mathrm{CHSH}} > \frac{3+\sqrt{6}}{6}$, there exists an XOR protocol with constant bias for arbitrary function f. However, there is no reason why Maj_3 should be used for the bias amplification. We can use arbitrary functions, e.g., the majority function on five variables, in place of Maj_3. Of course, on a given number n of input variables, the majority functions Maj_n minimize the threshold value, corresponding to $5/6$ for Maj_3. However, nonmajority functions may require a smaller number of nonlocal boxes than the majority functions. Hence, nonmajority functions are also candidates for the generalization of Brassard *et al.*'s protocol. We have to generalize two quantities "2" and "5/6" in the case of Maj_3, which are the number of nonlocal boxes needed for the computation and the threshold for the probability of the correctness of computation of the function for the bias amplification, respectively. In this work, these two quantities are clearly characterized.

Although we can consider a general function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ in place of $\mathrm{Maj}_3(x \oplus y)$, in this study, we restrict f to be an XOR function, i.e., $f(x,y) = g^{\oplus}(x,y) := g(x \oplus y)$ for some $g : \{0,1\}^n \to \{0,1\}$. It seems to be a natural restriction since the inputs x and y have meaning only when their XOR is taken. Linden *et al.* showed that quantum mechanics has no advantage on the XOR protocol for computation of the XOR function when the input distribution is also an XOR function [15].

IV. NONADAPTIVE PR-CORRECT PROTOCOLS

Brassard *et al.* consider the protocol according to the $\mathbb{F}_2$-polynomial representation (1) for computing $\mathrm{Maj}_3^{\oplus}$. In this section, we show that for arbitrary given $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, this protocol is the best protocol for computing $f(x,y)$ among all protocols satisfying nonadaptivity and PR-correctness.

Definition 7. An XOR protocol is said to be nonadaptive if inputs for nonlocal boxes does not depend on outputs of other nonlocal boxes. An XOR protocol is said to be PR-correct if the protocol computes the target function $f(x,y)$ without error when the nonlocal boxes are PR boxes. An XOR protocol is said to be nonredundant if the inputs $(l_i(x),r_i(x))_{i=1,\dots,t}$ for the nonlocal box satisfy

$$\begin{aligned}A(x) \oplus B(y) \quad \oplus \quad \bigoplus_{i=1}^{t} (C_i \wedge l_i(x) \wedge r_i(y)) = 0\\ \iff (C_i)_{i=1,\dots,t} = 0, \quad A(x) = B(y). \quad (2)\end{aligned}$$

The following lemma was shown by Kaplan *et al.* [16]. Here, we give a short proof using Fourier analysis.

Lemma 8. The outputs of both players in nonadaptive PR-correct nonredundant protocol must be parity of the outputs of nonlocal boxes and a function of local inputs.

Proof. Let $l_1(x), \dots, l_t(x)$ and $r_1(y), \dots, r_t(y)$ be the inputs of nonlocal boxes from Alice and Bob, respectively. Let $a_1, \dots, a_t$ and $b_1, \dots, b_t$ be the outputs of the nonlocal boxes for Alice and Bob, respectively. From any protocol, one can obtain a modified protocol using $(l_i'(x) := l_i(x) \oplus l_i(0), r_i'(x) := r_i(y) \oplus r_i(0))_{i=1,\dots,t}$ as the inputs for nonlocal boxes since replacements of a_i and b_i by $a_i' \oplus l_i'(x)r_i(0) \oplus l_i(0)r_i(0)$ and $b_i' \oplus l_i(0)r_i'(y)$ for $i = 1, \dots, t$, respectively, simulate the original protocol, where $(a_i', b_i')_{i=1,\dots,t}$ are the outputs of the nonlocal boxes in the modified protocol. This transformation preserves nonadaptivity, PR-correctness, and nonredundancy. This transformation also preserves whether or not the outputs of both players are parity of the outputs of nonlocal boxes and a function of local inputs. Hence, without loss of generality, we can assume that $l_1(0) = \cdots = l_t(0) = r_1(0) = \cdots = r_t(0) = 0$. Assume that $a = u_x(a_1, \dots, a_t)$ and $b = v_y(b_1, \dots, b_t)$. Since the protocol is PR-correct, $a \oplus b = u_x(a_1, \dots, a_t) \oplus v_y(a_1 \oplus z_1(x,y), \dots, a_t \oplus z_t(x,y))$ must be constant for all $(a_1, \dots, a_t) \in \{0,1\}^t$, where $z_i(x,y) := l_i(x) \wedge r_i(y)$. By letting $x = 0$ ($y = 0$), we obtain that v_y (u_x) is equal to u_0 (v_0) or its negation for any y (x), respectively. Hence, there exist Boolean functions $F : \{0,1\}^t \to \{0,1\}$, $\varphi, \psi : \{0,1\}^n \to \{0,1\}$ such that $u_x(a_1, \dots, a_t) = \varphi(x) \oplus F(a_1, \dots, a_t)$ and $v_y(b_1, \dots, b_t) = \psi(y) \oplus F(b_1, \dots, b_t)$. On the other hand, it

holds on the $\{+1,-1\}$ domain that

$$ab = \left(\sum_{S\subseteq[t]} \widehat{u}_x(S)\prod_{i\in S} a_i\right)\left(\sum_{S\subseteq[t]} \widehat{v}_y(S)\prod_{i\in S}(a_i z_i(x,y))\right) = \sum_{S_1,S_2\subseteq[t]} \widehat{u}_x(S_1)\widehat{v}_y(S_2)\prod_{i\in S_2} z_i(x,y)\prod_{i\in(S_1\cup S_2)-(S_1\cap S_2)} a_i$$
$$= \sum_{S\subseteq[t]}\left(\sum_{\substack{S_1,S_2,\\ (S_1\cup S_2)-(S_1\cap S_2)=S}} \widehat{u}_x(S_1)\widehat{v}_y(S_2)\prod_{i\in S_2} z_i(x,y)\right)\prod_{i\in S} a_i. \tag{3}$$

This is the Fourier expansion of $u_x(a_1,\dots,a_t)\oplus v_y(a_1\oplus z_1(x,y),\dots,a_t\oplus z_t(x,y))$ as a function of $a_1,\dots,a_t$. Since the function must be constant, the Fourier coefficients for the empty set must be ±1, i.e.,

$$\sum_{S_1\subseteq[t]} \varphi(x)\psi(y)\widehat{F}(S_1)^2\prod_{i\in S_1} z_i(x,y)\in\{+1,-1\}$$

for any $x,y\in\{0,1\}^n$. Hence, for any $x,y\in\{0,1\}^n$, $\prod_{i\in S_1} z_i(x,y)$ must be common for all $S_1\in\mathrm{supp}(\widehat{F})$. The equality $\prod_{i\in S_1} z_i(x,y)=\prod_{i\in S_2} z_i(x,y)$ for $S_1\neq S_2$ implies $\prod_{i\in(S_1\cup S_2)-(S_1\cap S_2)} z_i(x,y)=1$, which shows the existence of a redundant nonlocal box. Hence, $\widehat{F}(S_1)\neq0$ for unique $S_1\subseteq[t]$. This implies that u_x $[v_y]$ is the parity of variables in S_1 and $\varphi(x)$ $[\psi(y)]$, respectively. ■

Naturally, we can ask whether or not the non-redundancy is restriction, i.e., whether or not we can reduce the error probability of the protocol by using the redundancy when the nonlocal boxes are not the PR boxes. The following lemma says that redundancy does not help to reduce the error probability of a nonadaptive PR-correct protocol.

Lemma 9. For an arbitrary given nonadaptive PR-correct protocol, there exists a nonadaptive PR-correct nonredundant protocol whose error probability is at most that of the original protocol for any bias δ of the CHSH probability.

Proof. As in the proof of Lemma 8, we can assume without loss of generality that $l_1(0)=\cdots=l_t(0)=r_1(0)=\cdots=r_t(0)=0$. Similarly to (3), when the nonlocal boxes are not necessarily the PR boxes, ab is equal to

$$\sum_{S\subseteq[t]}\left(\sum_{\substack{S_1,S_2,\\ (S_1\cup S_2)-(S_1\cap S_2)=S}} \widehat{u}_x(S_1)\widehat{v}_y(S_2)\prod_{i\in S_2}(e_i z_i(x,y))\right)\prod_{i\in S} a_i,$$

where e_i represents the error of the output of ith nonlocal box; i.e., $e_i=+1$ if the ith nonlocal box computes correctly and $e_i=-1$ otherwise. Recall that the bias of the CHSH probability is δ; i.e., the expectation of e_i is δ. Since the nonlocal boxes are isotropic, e_i is independent of any other variables x, y, $(a_j)_{j\in[t]}$, and $(e_j)_{j\in[t]\setminus\{i\}}$ for $i\in[t]$. Since the nonlocal boxes are isotropic, a_i is uniformly distributed for all $i\in[t]$. Hence, the expectation of ab (the bias of $a\oplus b$) is

$$\sigma(x,y)\varphi(x)\psi(y)\sum_{S_1\subseteq[t]}\widehat{F}(S_1)^2\delta^{|S_1|}$$
$$=:\sigma(x,y)\varphi(x)\psi(y)\,\mathrm{Stab}_\delta(F),$$

where $\sigma(x,y)$ denotes the common sign of $\prod_{i\in S_1} z_i(x,y)\in\{+1,-1\}$ for all $S_1\in\mathrm{supp}(\widehat{F})$. Since the protocol is PR-correct, $\sigma(x,y)\sigma(x)\psi(y)\in\{+1,-1\}$ must be equal to $f(x,y)$. Hence, the output of the protocol is correct with probability $[1+\mathrm{Stab}_\delta(F)]/2$. On the other hand, since $\prod_{i\in S} z_i(x,y)\in\{+1,-1\}$ is common for all $S\in\mathrm{supp}(\widehat{F})$, we can obtain a nonadaptive PR-correct protocol by replacing $u_x(a_1,\dots,a_t)$ and $v_y(b_1,\dots,b_t)$ by $\varphi(x)\oplus\bigoplus_{i\in S^*} a_i$ and $\psi(y)\oplus\bigoplus_{i\in S^*} b_i$ for $S^*:=\mathrm{argmin}_{S\in\mathrm{supp}(\widehat{F})}|S|$, respectively. In order to obtain a nonadaptive PR-correct nonredundant protocol, we shrink the set S^* to $T\subseteq S^*$ if S^* includes the redundancy [the local terms $\varphi(x)$ and $\psi(y)$ should also be modified according to the shrinkage]. The bias of the probability of correctness of the protocol is $\delta^{|T|}\geqslant\delta^{|S^*|}\geqslant\mathrm{Stab}_\delta(F)$. ■

Lemma 9 implies that, if we are interested in the minimization of the error probability among all nonadaptive PR-correct protocols, we only have to consider nonadaptive PR-correct nonredundant protocols.

V. THE NUMBER OF NONLOCAL BOXES

Lemma 8 implies that an arbitrary nonadaptive PR-correct nonredundant protocol corresponds to an $\mathbb{F}_2$-polynomial representation of $f(x,y)$:

$$f(x,y)=A(x)\oplus B(y)\oplus\bigoplus_{i=1}^{t} l_i(x)r_i(y). \tag{4}$$

Since the bias of the correctness of the corresponding protocol is δ^t, we define the following measure of the complexity.

Definition 10. The nonlocal box complexity $\mathrm{NLBC}(f)$ is the minimum t such that there exists a representation (4).

The nonlocal box complexity can be characterized by the rank of some matrix on $\mathbb{F}_2$. The following theorem slightly generalizes a theorem in Ref. [16].

Theorem 11. For any $f:\{0,1\}^n\times\{0,1\}^n\to\{0,1\}$,

$$\mathrm{NLBC}(f)=\mathrm{rank}_{\mathbb{F}_2}(M_{f'}),$$

where $f'(x,y) = f(x,y) \oplus f(x,0) \oplus f(0,y) \oplus f(0,0)$ and where $M_{f'}$ is a $2^n \times 2^n$ matrix on $\mathbb{F}_2$ such that its (x,y) element is equal to $f'(x,y)$.

Proof. First, we show $\mathrm{NLBC}(f) \leqslant \mathrm{rank}_{\mathbb{F}_2}(M_{f'})$. If $\mathrm{rank}_{\mathbb{F}_2}(M_{f'}) = r$, there is a matrix factorization $M_{f'} = UV$ for some $2^n \times r$ matrix U and $r \times 2^n$ matrix V. This implies that $f'(x,y) = \bigoplus_{i=1}^{r} a_i(x)b_i(y)$, where $a_i(x)$ denotes the (x,i) element of U and where $b_i(y)$ denotes the (i,y) element of V. Hence, it holds that $f(x,y) = (f(x,0) \oplus f(0,0)) \oplus f(0,y) \oplus \bigoplus_{i=1}^{r} a_i(x)b_i(y)$, and hence $\mathrm{NLBC}(f) \leqslant r$.

Conversely, if $\mathrm{NLBC}(f) = t$, there is a representation $f(x,y) = A(x) \oplus B(y) \oplus \bigoplus_{i=1}^{t} l_i(x)r_i(y)$. There also exists a representation $f'(x,y) = A'(x) \oplus B'(y) \oplus \bigoplus_{i=1}^{t} l_i(x)r_i(y)$. Since $f'(0,y) = f'(x,0) = 0$ for all x and y, by expanding constant terms in $l_i(x)$ and $r_i(x)$ we obtain a representation $f'(x,y) = \bigoplus_{i=1}^{t} l'_i(x)r'_i(y)$. This implies that there is a matrix factorization $M_{f'} = UV$ for $2^n \times r$ matrix U and $r \times 2^n$ matrix V where the (x,i) element of U is $l'_i(x)$ and the (i,y) element of V is $r'_i(y)$. Hence, $\mathrm{rank}_{\mathbb{F}_2}(M_{f'}) \leqslant t$. ■

Remark. If we restrict the decomposition to be symmetric, i.e., $l_i = r_i$ for all $i = 1, \ldots, t$, an extra one dimension is required for arbitrary XOR function $g^{\oplus}$ [17].

Lemma 12. For any $g : \{0,1\}^n \to \{0,1\}$, $\mathrm{NLBC}(g^{\oplus}) = 0$ only when g is a parity of some variables or its negation. Furthermore, $\mathrm{NLBC}(g^{\oplus})$ cannot be equal to 1.

Proof. From Theorem 11, $\mathrm{NLBC}(g^{\oplus}) = 0$ implies $g(x \oplus y) \oplus g(x) \oplus g(y) \oplus g(0) = 0$. Hence, it holds that $g(x \oplus y) \oplus g(0) = (g(x) \oplus g(0)) \oplus (g(y) \oplus g(0))$, so that $g(z) \oplus g(0)$ is linear, i.e., parity of some variables. Assume $\mathrm{NLBC}(g^{\oplus}) = 1$. From Theorem 11, $\mathrm{rank}_{\mathbb{F}_2}(M_{g^{\oplus'}})$ must be equal to 1. Since $M_{g^{\oplus'}}$ is a symmetric matrix, there is a decomposition $M_{g^{\oplus'}} = vv^t$, where v denotes a $\mathbb{F}_2$ vector of length 2^n. On the other hand, the diagonal elements of $M_{g^{\oplus'}}$ must be zero. That implies $v = 0$, and hence $\mathrm{NLBC}(g^{\oplus}) = 0$. This is a contradiction. ■

Example 13. The following table shows the nonlocal box complexity of Maj_n computed numerically by a computer:

n	3	5	7	9	11	13	15	17
$\mathrm{NLBC}(\mathrm{Maj}_n^{\oplus})$	2	14	26	254	494	1090	1818	65534

In Example 13, it is not easy to find any rule between n and the nonlocal box complexity, although $\mathrm{NLBC}(\mathrm{Maj}_n^{\oplus}) = 2^{n-1} - 2$ may happen frequently, e.g., $n = 3,5,9,17$. Generally, it is considered to be difficult to express $\mathrm{rank}_{\mathbb{F}_2}(M_f)$ in a simple form for arbitrary given f. Note that the rank on $\mathbb{R}$ is always at least the rank on $\mathbb{F}_2$. Since $\mathrm{rank}_{\mathbb{R}}(M_{g^{\oplus}})$ is equal to the number of nonzero Fourier coefficients of g [18], $2^{n-1} + 1$ is an upper bound of $\mathrm{NLBC}(\mathrm{Maj}_n^{\oplus})$ for odd n [an inequality $\mathrm{rank}_{\mathbb{F}_2}(M_f) - 2 \leqslant \mathrm{NLBC}(f) \leqslant \mathrm{rank}_{\mathbb{F}_2}(M_f)$ can be obtained in a way similar to that of Theorem 11]. Here, we introduce a lower bound of the nonlocal box complexity using the one-way communication complexity.

Lemma 14. For any $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$,

$$\mathrm{NLBC}(f) \geqslant \max\{D^{\oplus}_{\rightarrow}(f), D^{\oplus}_{\leftarrow}(f)\}.$$

Proof. Assume $f(x,y)$ has the form (4). Bob can compute $f \oplus A(x)$ from $(l_i(x))_{i=1,\ldots,\mathrm{NLBC}(f)}$. ■

It obviously holds that $D^{\oplus}_{\rightarrow}(f) \geqslant D_{\rightarrow}(f) - 1$. If g is an odd function, i.e., $g(\overline{z}) = \overline{g(z)}$ where $\overline{z}$ denotes the bit inversion of z, then $D^{\oplus}_{\rightarrow}(g^{\oplus}) = D_{\rightarrow}(g^{\oplus}) - 1$ since $g(x \oplus y) \oplus g(x) = g(\overline{x} \oplus y) \oplus g(\overline{x})$.

Example 15. It obviously holds that $D_{\rightarrow}(\mathrm{Maj}_n^{\oplus}) = n$. Since Maj_n is an odd function, it holds that $D^{\oplus}_{\rightarrow}(\mathrm{Maj}_n^{\oplus}) = n - 1$. From Example 13, this lower bound is tight for $n = 3$, but becomes looser as n increases. This lower bound seems not to be asymptotically tight.

In fact, the adaptive protocol introduced in the next section has bias $\delta^{D^{\oplus}_{\rightarrow}(g^{\oplus})}$ for arbitrary XOR function $g^{\oplus}$.

VI. ADAPTIVE PROTOCOL

A. Pawłowski *et al.*'s protocol

In this section, we show an adaptive protocol which is inspired by the adaptive protocol invented in Ref. [8]. Let the address function Addr_n be

$$\mathrm{Addr}_n(x_0, \ldots, x_{2^n-1}, y_1, \ldots, y_n) := x_y,$$

where $y = \sum_{i=1}^{n} y_i 2^{i-1}$. In Ref. [8], Pawłowski *et al.* characterized the quantum limit $\frac{2+\sqrt{2}}{4}$ of the CHSH probability by using a principle called information causality. What they essentially showed in Ref. [8] is following.

Lemma 16. There is a PR-correct protocol computing the address function Addr_n with bias δ^n.

Proof. The lemma is shown by the induction. There is n representation

$$\mathrm{Addr}_1(x_0,x_1,y_1) = x_0 \oplus y_1(x_0 \oplus x_1).$$

Hence, there exists a nonadaptive protocol computing Addr_1 with bias δ, so that the lemma holds for $n = 1$. For $n \geqslant 2$, there is a recursive formula

$$\mathrm{Addr}_n(x_0, \ldots, x_{2^n-1}, y_1, \ldots, y_n) = \mathrm{Addr}_1(x'_0, x'_1, y_n),$$

where

$$x'_0 := \mathrm{Addr}_{n-1}(x_0, \ldots, x_{2^{n-1}-1}, y_1, \ldots, y_{n-1}),$$
$$x'_1 := \mathrm{Addr}_{n-1}(x_{2^{n-1}}, \ldots, x_{2^n-1}, y_1, \ldots, y_{n-1}).$$

From the hypothesis of the induction, there is a PR-correct protocol computing x'_0 and x'_1 with bias δ^{n-1}. Let a_0 and b_0 (a_1 and b_1) be random variables corresponding to the outputs of the protocol computing x'_0 (x'_1), respectively. Then, if $\delta = 1$, one obtains

$$\begin{aligned}
&\mathrm{Addr}_n(x_0, \ldots, x_{2^n-1}, y_1, \ldots, y_n) \\
&\quad = \mathrm{Addr}_1(a_0 \oplus b_0, a_1 \oplus b_1, y_n) \\
&\quad = \mathrm{Addr}_1(a_0, a_1, y_n) \oplus \mathrm{Addr}_1(b_0, b_1, y_n) \\
&\quad = a_0 \oplus y_n(a_0 \oplus a_1) \oplus b_{y_n}.
\end{aligned}$$

From this observation, we recursively define the protocol for Addr_n in the following way. (P1) Compute a_0 and a_1 at Alice's side, and b_{y_n} at Bob's side using the protocol for Addr_{n-1}. (P2) Input $a_0 \oplus a_1$ and y_n into the common nonlocal box, and obtain a' and b'. (P3) Output $a := a_0 \oplus a'$ at Alice's side and $b := b' \oplus b_{y_n}$ at Bob's side. This protocol is obviously PR-correct. Since at each step, the error of bias δ is XORed, this protocol has bias δ^n. ■

B. The adaptive protocol

In the following, we show an adaptive protocol computing arbitrary given function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ using Pawłowski *et al.*'s protocol.

Theorem 17. For arbitrary function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, there is a PR-correct protocol computing f with bias $\delta^{\min\{D^{\oplus}_{\to}(f), D^{\oplus}_{\leftarrow}(f)\}}$.

Proof. Arbitrary function f can be represented by

$$f(x,y) = \mathrm{Addr}_n(f(x,0,\dots,0), f(x,0\dots,0,1), \dots, f(x,1,\dots,1), y_1, \dots, y_n).$$

From Lemma 16, there is an adaptive protocol computing f with bias δ^n.

We can consider compression of Bob's input since we do not have to distinguish y's belonging to equivalent columns of M_f. By applying the compression, we obtain the protocol with bias $\delta^{D_{\leftarrow}(f)}$. Furthermore, if we have an XOR protocol for $f(x,y) \oplus B(y)$, we also obtain an XOR protocol for $f(x,y)$ by replacing Bob's output b with $b \oplus B(y)$. Hence, we obtain the protocol with bias $\delta^{D^{\oplus}_{\leftarrow}(f)}$. In the same way, we also obtain the protocol with bias $\delta^{D^{\oplus}_{\to}(f)}$. ■

From Lemma 14 and Theorem 17, we obtain the following corollary.

Corollary 18: The adaptive PR-correct protocol in Theorem 17 is no worse than any nonadaptive PR-correct protocol.

VII. BIAS AMPLIFICATION

We now consider the bias amplification by general XOR function $g^{\oplus}$ in Brassard *et al.*'s protocol, where $g^{\oplus}$ is computed by the adaptive PR-correct protocol introduced in Theorem 17. If z is a random variable taking $+1$ with probability $\frac{1+\epsilon}{2}$ and -1 with probability $\frac{1-\epsilon}{2}$, its expectation is ϵ. The expectation ϵ is called the bias of random variable z. If the inputs for g is independently and identically distributed and have bias ϵ, the bias of output of g is given in the following formula.

Definition 19. For any $g : \{0,1\}^n \to \{0,1\}$, we define

$$\mathrm{Bias}_\epsilon(g) := \sum_{S \subseteq [n]} \widehat{g}(S)\epsilon^{|S|}.$$

Example 20. Since $\mathrm{Maj}_3(z_1,z_2,z_3) = (1/2)(z_1 + z_2 + z_3 - z_1z_2z_3)$, one obtains $\mathrm{Bias}_\epsilon(\mathrm{Maj}_3) = (3/2)\epsilon - (1/2)\epsilon^3$. Roughly speaking, the input bias ϵ is amplified to $(3/2)\epsilon$ for small ϵ.

When a Boolean function g is computed correctly with probability $\frac{1+\rho}{2}$, the output bias of g is $\rho\,\mathrm{Bias}_\epsilon(g)$. We say that the bias is amplified by g if the absolute value of bias of the output of g is larger than that of the input and if the sign of bias is preserved. The bias is amplified by the noisy g for sufficiently small input bias iff $\mathrm{Bias}_0(g) = 0$ and $\rho \frac{d\,\mathrm{Bias}_\epsilon(g)}{d\epsilon}|_{\epsilon=0} > 1$. Hence, we obtain the following theorem.

Theorem 21. Assume that $g : \{0,1\}^n \to \{0,1\}$ can be computed correctly with probability $\frac{1+\rho}{2}$. Then, the bias is amplified by the noisy g when the input bias is sufficiently small iff $\widehat{g}(\varnothing) = 0$ and $\rho > \rho_{\mathrm{B}}(g)$, where

$$\rho_{\mathrm{B}}(g) := \frac{1}{\max\left\{1, \sum_{i=1}^n \widehat{g}(\{i\})\right\}}.$$

The majority functions minimize $\rho_{\mathrm{B}}(g)$.

Lemma 22. For $g : \{0,1\}^n \to \{0,1\}$,

$$\rho_{\mathrm{B}}(g) \geqslant \frac{2^{n-1}}{n\binom{n-1}{\frac{n-1}{2}}} \quad \text{if } n \text{ is odd},$$

$$\rho_{\mathrm{B}}(g) \geqslant \frac{2^n}{n\binom{n}{\frac{n}{2}}} \quad \text{if } n \text{ is even}.$$

The equality is achieved by and only by the majority functions on n variables. Asymptotically, it holds that $\rho_{\mathrm{B}}(g) \geqslant \sqrt{\pi/(2n)}[1 + O(n^{-1/2})]$.

Proof. One obtains $\sum_{i\in[n]} \widehat{g}(\{i\}) = \mathbb{E}[g(x)(x_1 + \cdots + x_n)] \leqslant \mathbb{E}[|x_1 + \cdots + x_n|]$ where the equality holds only when g is Maj_n [9]. Hence, only the majority functions Maj_n maximize $\sum_{i\in[n]} \widehat{g}(\{i\})$. It is easy to complete the rest of the proof [9]. ■

Note that the lower bound for even n is equal to the lower bound for $n-1$. The condition on δ for the bias amplification by Brassard *et al.*'s protocol is $\delta^{D^{\oplus}_{\to}(g^{\oplus})} > \rho_{\mathrm{B}}(g)$.

Definition 23. For any $g : \{0,1\}^n \to \{0,1\}$,

$$\delta_{\mathrm{B}}(g) := \begin{cases} \rho_{\mathrm{B}}(g)^{\frac{1}{D^{\oplus}_{\to}(g^{\oplus})}} & \text{if } \widehat{g}(\varnothing) = 0 \text{ and } \rho_{\mathrm{B}}(g) < 1, \\ 1 & \text{otherwise.} \end{cases}$$

If $\delta > \delta_{\mathrm{B}}(g)$ for some $g : \{0,1\}^n \to \{0,1\}$, there exists an XOR protocol with constant bias.

Example 24. One obtains $\delta_{\mathrm{B}}(\mathrm{Maj}_3) = \sqrt{2/3}$, which means that the threshold for the CHSH probability is $\frac{1+\sqrt{2/3}}{2} = \frac{3+\sqrt{6}}{6}$ [7].

We can now rephrase Theorem 1 in the following form.

Theorem 25.

$$\inf_{g:\{0,1\}^n \to \{0,1\}, n\in\mathbb{N}} \delta_{\mathrm{B}}(g) = \sqrt{\frac{2}{3}}.$$

Furthermore, $\delta_{\mathrm{B}}(g) = \sqrt{2/3}$ iff g is essentially equivalent to Maj_3.

Here, we say that g is essentially equivalent to Maj_3 if g is the majority of some fixed three-input variables and ignores the other $n-3$ input variables. The following lemma was shown in Ref. [19].

Lemma 26. For any $g : \{0,1\}^n \to \{0,1\}$,

$$D_{\to}(g^{\oplus}) = \dim(\widehat{g}).$$

Since $D^{\oplus}_{\to}(f) \geqslant D_{\to}(f) - 1$, it holds that $D^{\oplus}_{\to}(g^{\oplus}) \geqslant \dim(\widehat{g}) - 1$.

Remark. If $A(x)$ in the definition of $D^{\oplus}_{\to}(f)$ is restricted to be linear, $D^{\oplus}_{\to}(g^{\oplus})$ is equal to the affine dimension of $\widehat{g}$, which is the minimum dimension of affine space on $\mathbb{F}_2$ including $\{1_S \mid S \in \mathrm{supp}(\widehat{g})\}$. Hence, the affine dimension of $\widehat{g}$ is an upper bound of $D^{\oplus}_{\to}(g^{\oplus})$.

First, we show that Theorem 25 holds for $n \leqslant 4$.

Lemma 27. It holds that $\delta_{\mathrm{B}}(g) \geqslant \sqrt{2/3}$ for all Boolean functions g on at most four variables. Furthermore, for $n \leqslant 4$, only functions essentially equivalent to Maj_3 satisfy $\delta_{\mathrm{B}}(g) = \sqrt{2/3}$.

Proof. Assume $D^{\oplus}_{\rightarrow}(g^{\oplus}) \leqslant 1$. Then, the protocol is nonadaptive. From Lemma 12, g must be linear, and hence $\rho_{\mathrm{B}}(g) = 1$. Assume $D^{\oplus}_{\rightarrow}(g^{\oplus}) \geqslant 2$. From Lemma 22, $\rho_{\mathrm{B}}(g) \geqslant 2/3$ for $n \leqslant 4$, and hence $\delta_{\mathrm{B}}(g) \geqslant \sqrt{2/3}$. From Example 24, it is achieved by Maj_3.

Next, we show the uniqueness. From the above argument, it holds that $\delta_{\mathrm{B}}(g) = \sqrt{2/3}$ only when $\rho_{\mathrm{B}}(g) = 2/3$ and $D^{\oplus}_{\rightarrow}(g^{\oplus}) = 2$. From Lemma 22, $\rho_{\mathrm{B}}(g) = 2/3$ only when g is one of the 64 majority functions on four variables. In the following, we show that for $g \in \mathrm{Maj}_4$, $D^{\oplus}_{\rightarrow}(g^{\oplus}) = 2$ only when g is essentially equivalent to Maj_3. From Lemma 26, $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| \leqslant \dim(\widehat{g}) \leqslant 3$. If $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| \leqslant 2$, it holds that $\sum_{i\in[n]} \widehat{g}(\{i\}) \leqslant \sqrt{2} < 3/2$ from the Cauchy-Schwartz inequality. If $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| = 3$, g depends only on three variables since $\dim(\widehat{g}) \leqslant 3$. Hence, g is essentially equivalent to Maj_3. ■

From the following lemma, only Boolean functions with small Fourier dimension may outperform Maj_3.

Lemma 28. For any $g : \{0,1\}^n \to \{0,1\}$,

$$\delta_{\mathrm{B}}(g) \geqslant \left(\frac{1}{\dim(\widehat{g})}\right)^{\frac{1}{2[\dim(\widehat{g})-1]}}.$$

In particular, if $\dim(\widehat{g}) \geqslant 5$, it holds that $\delta_{\mathrm{B}}(g) > \sqrt{2/3}$.

Proof. One obtains

$$\begin{aligned}\dim(\widehat{g}) &\geqslant |\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| \\ &\geqslant \frac{\left(\sum_{i\in[n]} \widehat{g}(\{i\})\right)^2}{\sum_{i\in[n]} \widehat{g}(\{i\})^2} \geqslant \left(\sum_{i\in[n]} \widehat{g}(\{i\})\right)^2.\end{aligned}$$

In the above, the first inequality is trivial. The second inequality is the Cauchy-Schwartz inequality. The third inequality holds since the sum of squares of all of the Fourier coefficients is 1. Hence, $\rho_{\mathrm{B}}(g) \geqslant \dim(\widehat{g})^{-1/2}$. From Lemma 26, we obtain this theorem. ■

Lemmas 27 and 28 give the complete proof of Theorem 25.

Proof of Theorem 25. From Lemma 28, we only have to show that if $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| \leqslant \dim(\widehat{g}) \leqslant 4$, $\delta_{\mathrm{B}}(g) \leqslant \sqrt{2/3}$ only for g essentially equivalent to Maj_3. Assume $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| = 4$. Then, the Boolean function g depends only on four input variables since $\dim(\widehat{g}) \leqslant 4$. From Lemma 27, there is no function on four variables satisfying $\delta_{\mathrm{B}}(g) \leqslant \sqrt{2/3}$ except for functions essentially equivalent to Maj_3. Next, we assume $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| = 3$. In this case, $\sum_{i\in[n]} \widehat{g}(\{i\}) \leqslant \sqrt{3}$. Since $(1/\sqrt{3})^{1/3} > \sqrt{2/3}$, we can assume $D^{\oplus}_{\rightarrow}(g^{\oplus}) \leqslant 2$. Then, the Boolean function g depends only on three input variables since $\dim(\widehat{g}) \leqslant D^{\oplus}_{\rightarrow}(g^{\oplus}) + 1 \leqslant 3$. From Lemma 27, there is no function on three variables satisfying $\delta_{\mathrm{B}}(g) \leqslant \sqrt{2/3}$ except for Maj_3. Next, we assume $|\{i \in [n] \mid \widehat{g}(\{i\}) \neq 0\}| \leqslant 2$. In this case, $\sum_{i\in[n]} \widehat{g}(\{i\}) \leqslant \sqrt{2}$. Since $(1/\sqrt{2})^{1/2} > \sqrt{2/3}$, we can assume $D^{\oplus}_{\rightarrow}(g^{\oplus}) \leqslant 1$. From Lemma 12, it holds that $\delta_{\mathrm{B}}(g) = 1$. We conclude that there is no function satisfying $\delta_{\mathrm{B}}(g) \leqslant \sqrt{2/3}$ except for functions essentially equivalent to Maj_3. ■

VIII. CONCLUSION

In this paper, we show that the three-input majority function is the unique optimal function for Brassard *et al.*'s bias amplification on some conditions. This paper also develops a mathematical framework using Fourier analysis for problems on XOR protocols with nonlocal boxes. On the other hand, in this paper, functions $g^{\oplus}$ for the bias amplification are restricted to be XOR function, although it seems to be a natural restriction. Furthermore, protocols for computing functions $g^{\oplus}$, in this paper, are restricted to be a particular adaptive PR-correct protocol, which is improves upon arbitrary nonadaptive PR-correct protocols. General adaptive protocols may allow more reliable computation than these protocols [13]. Hence, the result of this paper does not show the limitation of the idea of the bias amplification, but shows only the limitation of the idea of bias amplification by an XOR function computed by the particular adaptive PR-correct protocol. The bias amplification by general adaptive computation of a non-XOR function would be an interesting direction of research.

ACKNOWLEDGMENT

This work was supported by MEXT KAKENHI Grant No. 24106008.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[3] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[4] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).
[5] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[6] W. van Dam, Nat. Comput. **12**, 9 (2013).
[7] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Phys. Rev. Lett. **96**, 250401 (2006).
[8] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Nature (London) **461**, 1101 (2009).
[9] R. O'Donnell, *Analysis of Boolean Functions* (Cambridge University Press, Cambridge, UK, 2014).
[10] L. Masanes, A. Acin, and N. Gisin, Phys. Rev. A **73**, 012112 (2006).
[11] A. J. Short, Phys. Rev. Lett. **102**, 180502 (2009).
[12] M. Forster, S. Winkler, and S. Wolf, Phys. Rev. Lett. **102**, 120401 (2009).
[13] N. Brunner and P. Skrzypczyk, Phys. Rev. Lett. **102**, 160403 (2009).
[14] J. von Neumann, *in Automata Studies*, edited by C. E. Shannon and J. McCarthy, Annals of Mathematics Studies 34 (Princeton University Press, Princeton, 1956), p. 43.

[15] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. Lett. **99**, 180502 (2007).
[16] M. Kaplan, S. Laplante, I. Kerenidis, and J. Roland, Quantum Inf. Comput. **11**, 40 (2011).
[17] A. Lempel, SIAM J. Comput. **4**, 175 (1975).
[18] A. Bernasconi and B. Codenotti, IEEE Trans. Comput. **48**, 345 (1999).
[19] A. Montanaro and T. Osborne, arXiv:0909.3392.