Relaxed Bell inequality as a trade-off relation between measurement dependence and hiddenness

Gen Kimura [1,*] Yugo Susuki,[1] and Kei Morisue [2,†]
[1] College of Systems Engineering and Science, Shibaura Institute of Technology, Saitama 330-8570, Japan
[2] Department of Physics, Waseda University, Tokyo 169-8555, Japan

(Received 27 June 2022; revised 10 April 2023; accepted 17 July 2023; published 18 August 2023)

Quantum correlations that violate the Bell inequality cannot be explained by any local-hidden-variable theory that is measurement independent. However, this violation merely signifies the incompatibility of the underlying assumptions of reality, locality, and measurement independence, without providing a quantitative measure of the extent to which each assumption is violated. In contrast, Hall introduced a measure for each of the following assumptions: indeterminism, signaling, and measurement dependence, and Hall also generalized the Bell-Clauser-Horne-Shimony-Holt inequality, which provides a quantitative trade-off relationship between these assumptions. In this paper, we consider the introduction of hidden variables to be an essential assumption of Bell's theorem and introduce a quantification of hidden variables, which we term "hiddenness." We derive a trade-off relation between hiddenness and measurement dependence that applies to any local-hidden-variable theory.

DOI: 10.1103/PhysRevA.108.022214

I. INTRODUCTION

In 1964, Bell made a surprising discovery that quantum correlations cannot be explained by local realism [1,2]. He introduced an inequality satisfied by all local-hidden-variable models and showed that there are quantum correlations that violate the inequality when an appropriately chosen set of measurements is made. The inequality is now known as the Bell inequality, and the sequence of results is referred to as Bell's theorem [3]. Importantly, the quantities appearing in the inequality are made up of measurable values (expected values or probabilities). As a result, these inequalities can be directly verified through experiments. Since the discovery of Bell's theorem, many experiments have confirmed (along with the verification of various loopholes) that certain quantum correlations violate the inequality as predicted by quantum theory [4–10]. Consequently, it is widely believed that quantum theory is correct and, more importantly, that phenomena that cannot be explained by local realism indeed exist. Subsequently, these quantum correlations have been found to be useful resources for various areas of quantum information processing, such as quantum computers, quantum teleportation, and quantum cryptography [11,12]. Notably, the experimental fact of the violation of the Bell inequality is recognized as the key feature that makes some information processes completely secure, such as key distribution [13–15] and random-number generation [16,17].

In Bell's argument, however, the violation of the Bell inequality tells us just that the underlying assumptions are logically incompatible. In particular, nothing can be said about how much each assumption must be violated to explain the experimental facts. In addition, clarifying the assumptions behind Bell's theorem is not very obvious because some of them are often implicit and can also vary between studies. In Bell's original paper [1], he explicitly assumed (i) determinism and (ii) locality in hidden-variable models. There is also another essential assumption that is often "overlooked"; namely, (iii) the choice of measurement setting is independent of the hidden variable [18–21], usually reflecting the existence of "free will" or true randomness.[1] Hence, Hall proposed quantitative measures for these three assumptions, referring to them as indeterminism I, signaling S, and measurement dependence M, respectively. Moreover, he successfully generalized the Bell-Clauser-Horne-Shimony-Holt (Bell-CHSH) inequality [23] that provides a trade-off relation between the empirical value C (the CHSH value) and measures I, S, and M in the simplest Bell scenario (with binary measurement settings and binary outcomes) [24,25]. To the best of our knowledge, this was the first attempt to shed light on Bell's theorem in this way.

Subsequently, several works appeared to follow this line of research. Another relaxed Bell inequality was derived for models in which only one party can violate measurement independence [26] and was further generalized in [27] by the introduction of the measures for measurement dependence for each party. In [28], the authors discovered a relaxed Bell inequality in any local model, which includes the upper and lower bounds of the probability of measurement contexts. Their finding shows that an arbitrarily small amount of measurement independence is sufficient to manifest quantum

*gen@shibaura-it.ac.jp
†Present address: 160-0022 Tokyo Japan.

[1] Strictly speaking, we also assume that the measurement choice can be described by a probability theory. In contrast, free will that happens noncausally [22] regardless of whether it can be described by a probability theory has also been discussed.

nonlocality. Hall and Branciard observed a related relaxed Bell inequality and discussed quantum nonlocality within the context of both causal and retrocausal models [29]. Ghadimi considered a model that relaxed the signaling property (usually referred to as parameter dependence [3,30] or active nonlocality [31]) and derived another relaxed Bell inequality [32].

The significance of these studies extends to applications in informatics, where the underlying variables can be interpreted as potential resources for eavesdroppers. Koh *et al.* showed a trade-off relation between the CHSH value, the guessing probability, and the free-will parameter (related to the measurement dependence introduced above) with the application of secure randomness expansion [33].

However, the very assumption of introducing hidden variables has not been subject to quantification thus far. We claim that this assumption should also be quantitatively measured as one of the fundamental underlying assumptions in Bell's theorem. In this paper, we introduce such a measure by $H = \#(\Lambda) - 1$, where $\#(\Lambda)$ is the size (i.e., the cardinality) of the set of hidden variables Λ and H refers to the "hiddenness." If the variable is in the hands of the eavesdropper, the measure can be interpreted as the amount of information he or she can store. We derive a trade-off relation (Theorem 1) between H and measurement dependence M which is satisfied for all (measurement-dependent) local models [28]. In the case of $M = 0$ (measurement independent), the relation recovers the Bell-CHSH inequality. On the other hand, with the observed data that violate the Bell-CHSH inequality, M and H have a trade-off relation: To decrease M, H must increase and vice versa. Interestingly, the trade-off relationship saturates when $H \geqslant 3$ and coincides with Hall's inequality [25] in the case of local models. Therefore, this result is also a generalization of Hall's inequality for local models. We also show that our inequalities describe the necessary and sufficient condition of a local model in the sense that not only do all local models satisfy them but also there is a local model for which M and H satisfy the inequalities.

To demonstrate these results, we technically derive the optimal CHSH value $C_{\rm opt}$ [see (17)] that can be attained by some local model and establish the lower and upper bounds of $C_{\rm opt}$ with given measurement dependence M and hiddenness H (Propositions 1 and 2). While the upper bound gives the trade-off relation mentioned above, the lower bound also gives a quantitative estimate of the fact that the violation of the Bell-CHSH inequality can still be explained by a local-hidden-variable model if we relax the measurement-independence assumption.

This paper is organized as follows: In Sec. II, we provide a brief introduction of hidden-variable theory and the measure for hiddenness. In Sec. III, we show a relaxed Bell inequality as a trade-off relation with the measurement dependence and hiddenness. In Sec. IV, we construct tight models that attain the equality of our trade-off relation. Finally, in Sec. V, we present our conclusion and discussion.

II. MEASURE FOR HIDDENNESS

Let us consider a bipartite physical system A and B (for Alice and Bob) in which the measurements are supposed to be performed in spacelike separated regions. The experimentally accessible probability is the set of joint probabilities $p(a, b|x, y)$, where x and y denote the measurements performed by Alice and Bob and a and b denote their respective outcomes. In the hidden-variable theory, we introduce a hidden variable $\lambda \in \Lambda$ so that the empirical joint probabilities are obtained by averaging over the hidden variable:

$$p(a, b|x, y) = \sum_{\lambda\in\Lambda} p(\lambda|x, y)p(a, b|x, y, \lambda), \tag{1}$$

where $p(\lambda|x, y)$ is the probability of λ given the values of (x, y) and $p(a, b|x, y, \lambda)$ represents the joint probability of the outcome (a, b) given the values of (x, y) and λ. Notice that one should replace the summation to the integral in the general situation. However, a nontrivial structure arises when we consider a model in which the size (i.e., cardinality) of the set Λ of hidden variables is finite. Bearing this in mind, we use a summation symbol for the hidden variables, except in Appendix B, where we provide a proof for (continuously) infinite models. In order to prove the Bell inequality, we need to assume both the locality condition and the measurement independence: The locality condition states that the outcomes a and b with fixed λ are statistically independent, i.e., $p(a, b|x, y, \lambda) = p(a|x, \lambda)p(b|y, \lambda)$. (It is worth noting that the locality condition is equivalent to assuming both parameter independence and outcome independence [3,30,31].) On the other hand, measurement independence asserts that the measurement context (x, y) and λ are independent: $p(\lambda|x, y) = p(\lambda)$. To summarize, in the context of local-hidden-variable theory with measurement independence, the following equation holds true:

$$p(a, b|x, y) = \sum_{\lambda\in\Lambda} p(\lambda)p(a|x, \lambda)p(b|y, \lambda). \tag{2}$$

In this paper, we relax the measurement independence and consider a (measurement-dependent) local model [28]:

$$p(a, b|x, y) = \sum_{\lambda\in\Lambda} p(\lambda|x, y)p(a|x, \lambda)p(b|y, \lambda). \tag{3}$$

Following Ref. [20], we shall use the same measure for the measurement dependence:

$$M := \sup_{x,y,x',y'} \sum_{\lambda} |p(\lambda|x, y) - p(\lambda|x', y')|. \tag{4}$$

Note that $0 \leqslant M \leqslant 2$ and $M = 0$ if and only if the model is measurement independent. However, it is useful to express M in terms of the total variation distance [34] (sometimes known as the trace distance [11] or the Kolmogorov distance [35,36]) as follows. The total variation distance between two probability measures P and Q is defined as

$$\delta(P, Q) := \sup_{E} |P(E) - Q(E)|, \tag{5}$$

where the supremum is taken over all the events $E \subset \Lambda$. In the discrete model $\Lambda = \{\lambda_1, \lambda_2, \dots\}$, it is easy to see the relation $\delta(P, Q) = \frac{1}{2}\sum_\lambda |p_\lambda - q_\lambda|$, where P and Q are given by the probability distributions $(p_\lambda)_{\lambda\in\Lambda}$ and $(q_\lambda)_{\lambda\in\Lambda}$, respectively. Therefore, we can express M as

$$M = 2 \sup_{x,y,x',y'} \delta(P_{xy}, P_{x'y'}), \tag{6}$$

where $P_{xy} = [p(\lambda|x,y)]_{\lambda\in\Lambda}$ denotes the probability distribution for λ with the measurement context (x, y).

As mentioned previously, the introduction of hidden variables is one of the essential assumptions underlying Bell's theorem. Therefore, we believe it is important to quantify this assumption in order to understand its meaning. In this paper, we introduce the following simple measure H, which we term *hiddenness*:

$$H := \#(\Lambda) - 1, \tag{7}$$

where $\#(\Lambda)$ is the cardinality of the set Λ. Obviously, $H \geqslant 0$ and takes only a discrete natural number (including infinity). This measure quantifies the degree to which we need to introduce hidden variables in order to explain the empirical statistics. Another interpretation is the memory size available to the eavesdropper if the hidden variable is in his or her possession.

The minimum case, where $H = 0$, corresponds to a trivial scenario where there is only one elementary event for the hidden variable. This essentially means there is no introduction of any hidden variable. The simplest, but nontrivial, case is $H = 1$, where there are two hidden elementary events, $\Lambda = \{\lambda_1, \lambda_2\}$. This scenario is logically possible because one can imagine a world with a "hidden coin" (e.g., $\lambda_1 =$ "tails" or $\lambda_2 =$ "heads"). Similarly, for $H = 2, 3, \ldots$, we need $3, 4, \ldots$ hidden elementary events for the hidden variable, respectively. The following section introduces a relaxed Bell inequality that establishes a trade-off relation between M and H, which holds in all local models.

III. RELAXED BELL INEQUALITY

In this section, we examine the simplest Bell scenario, often referred to as the CHSH setting, which involves binary measurement settings $x, y = 0, 1$ and binary measurement outcomes $a, b = \pm 1$. As an experimentally accessible quantity, we consider the *CHSH value* defined by

$$C := \langle 00\rangle + \langle 01\rangle + \langle 10\rangle - \langle 11\rangle, \tag{8}$$

where $\langle xy\rangle := \sum_{a,b=\pm 1} ab\ p(a,b|x,y)$ $(x, y = 0, 1)$ denotes the expectation value of the product of the measurement outcomes for joint measurement setting (x, y). It is well known [23] that for any measurement-independent local-hidden-variable model (2), the CHSH value is always bounded from above by 2, which is known as the Bell-CHSH inequality:

$$C \leqslant 2. \tag{9}$$

It should be emphasized that Bell-CHSH inequalities [the set of eight inequalities obtained by taking the absolute value and changing the position of the minus sign in (8)] provide not only a necessary condition but also a sufficient condition for the statistics to be explainable by measurement-independent local-hidden-variable models [37]. This fact causes Bell-CHSH inequalities to be of special interest.

Our main finding is the following.

Theorem 1. For any local model,

$$C \leqslant \min[H, 3]M + 2 \tag{10}$$

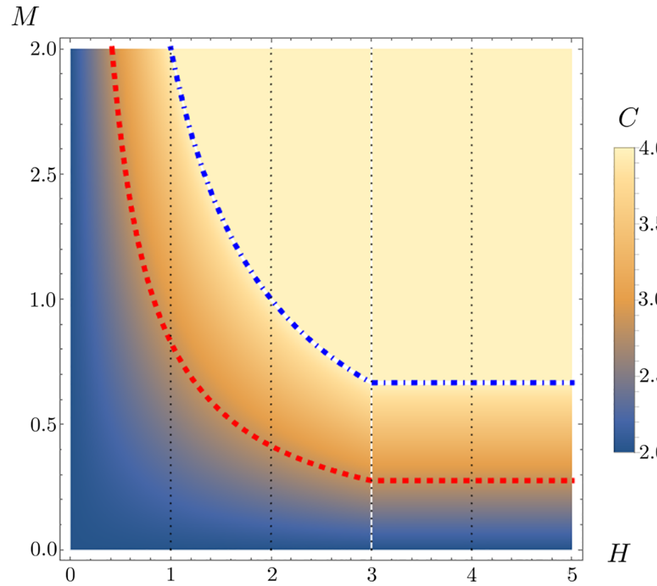

FIG. 1. The upper bound of the CHSH value C in (10) is plotted as a trade-off relation between H and M. The red dashed line corresponds to the maximal violation of the Bell-CHSH inequality in quantum systems: $C = 2\sqrt{2}$. The region above the blue dot-dashed line is a trivial violation: $C = 4$.

(in addition to the trivial bound $C \leqslant 4$).[2] The inequality is tight in the sense that there is a local model that can attain the equality in (10).

Inequality (10) is a generalization of Bell-CHSH inequality since it recovers the inequality by setting $M = 0$ (measurement independence). In general, it provides a trade-off relation between the CHSH value C, measurement dependence M, and hiddenness H. With a given violation of the Bell-CHSH inequality ($C > 2$), one can estimate the trade-off between M and H: The less hiddenness H there is, the more measurement dependence M is required and vice versa (see Fig. 1). Suppose, for instance, that we observe the maximum violation of the Bell-CHSH inequality in quantum theory, i.e., the Tsirelson bound $C = 2\sqrt{2} \simeq 2.8$. Then, for $H \geqslant 3$, one should give up measurement independence of at least $M = \frac{2}{3}(\sqrt{2} - 1) \simeq 0.276$. For $H = 2$, M should be greater than or equal to $\sqrt{2} - 1 \simeq 0.414$, and for $H = 1$, $M \geqslant 2(\sqrt{2} - 1) \simeq 0.828$.

An immediate corollary of Theorem 1 is that, for any local model,

$$C \leqslant 3M + 2. \tag{11}$$

This fact was previously observed by Hall [25], and therefore, the relation (10) generalizes his result for local models. Since this holds for any H, an ultimate lower bound for M exists:

$$\frac{C-2}{3} \leqslant M \ (\Leftrightarrow C \leqslant 3M + 2). \tag{12}$$

[2] Hence, the inequality can be written as $C \leqslant \min[\,\min[H, 3]M + 2, 4]$. We adopt the form (10) just to avoid this ugly expression.

Note that the case $H = 0$, which essentially corresponds to a local model without the introduction of hidden variables, also gives the Bell-CHSH inequality.

It would be interesting to consider the implications of our results for information security. For example, let us imagine that the hidden variable is in the possession of an eavesdropper and that H corresponds to the size of an exploitable information source. Our results suggest the following: In order to cheat legitimate users with an apparent violation of the Bell inequality based on a local model, the eavesdropper has to have great control over the measurements of the legitimate users if the memory size H is small. Conversely, if the eavesdropper has a large H, he or she does not need to worry much about controlling the measurements. Interestingly, however, there is a threshold ($H \geqslant 3$) beyond which the measurement dependence M cannot be reduced any further.

In the following, we provide a proof of Theorem 1 in two steps: First, we introduce the optimal CHSH value $C_{\rm opt}$ that can be achieved by a local model. Second, we establish the tight upper bound (as well as the tight lower bound) for $C_{\rm opt}$, which yields (10). Although the proof is done with a discrete model, the result is still valid for an uncountable model [$\#(\Lambda) = \infty$], which saturates to (12). However, since the proof requires a slightly different approach for the uncountable case, it is presented in Appendix B.

A. Optimal CHSH value for the local model

Using local expectation values $A_x := \sum_a ap(a|x,\lambda)$ ($x = 0, 1$) and $B_y := \sum_b bp(b|y,\lambda)$ ($y = 0, 1$), the CHSH value (8) for any local model (3) is given by

$$C = \sum_\lambda (z_1 A_0 B_0 + z_2 A_0 B_1 + z_3 A_1 B_0 - z_4 A_1 B_1), \tag{13}$$

where $z_i := p(\lambda|i)$ and $i = 1, 2, 3, 4$ corresponds to the measurement context $(x, y) = (0, 0), (0, 1), (1, 0), (1, 1)$, respectively. [In what follows, we often use the same labeling i for measurement contexts (x, y) for convenience.] Notice here that all λ dependences in A_x, B_y, and z_i are omitted. We have a tight inequality:

$$C \leqslant \sum_\lambda \left(\sum_{i=1}^4 z_i - 2 \min_i z_i \right) \tag{14}$$

from the following lemma.

Lemma 1. For any positive tuple $\boldsymbol{z} := (z_1, z_2, z_3, z_4) \in \mathbb{R}^4$, we have

$$\max_{\substack{A,A' \in [-1,1] \\ B,B' \in [-1,1]}} (z_1 AB + z_2 AB' + z_3 A'B - z_4 A'B') \tag{15}$$

$$= g(\boldsymbol{z}) := \sum_{i=1}^4 z_i - 2 \min_i z_i. \tag{16}$$

See Appendix A for the proof. Now we introduce the optimal CHSH value for local models as

$$C_{\rm opt} := \max_{p(a|x,\lambda), p(b|y,\lambda)} C.$$

By utilizing the normalization condition, i.e., $\sum_j \sum_\lambda p(\lambda|j) = \sum_j 1 = 4$ and (14), we have an alternative form:

$$C_{\rm opt} = \sum_\lambda g(\boldsymbol{z}) = 4 - 2 \sum_\lambda \min_i p(\lambda|i). \tag{17}$$

We observe the trivial inequality $C_{\rm opt} \leqslant 4$ directly from this form. It also can be shown that the equality is attained by appropriate local probabilities $p(a|x,\lambda)$ and $p(b|y,\lambda)$ such that they attain the maximums in Lemma 1 for each λ.

In the subsequent sections, we derive the lower and upper bounds of $C_{\rm opt}$ with given H and M. The trivial case with $H = 0$ is described separately here. This case can be described by introducing a trivial set of hidden variable, i.e., a singleton set $\Lambda = \{\lambda_1\}$, so that

$$p(a, b|x, y) = p(a|x, \lambda_1) p(b|y, \lambda_1) \tag{18}$$

and $p(\lambda_1|x, y) = 1$ for all x, y. Since $M = 0$ is always satisfied in this case, the measurement dependence will not occur. We also have $C_{\rm opt} := 4 - 2 \min_i p(\lambda_1|i) = 2$. Hence, only the case

$$M = 0, \ C_{\rm opt} = 2 \tag{19}$$

is possible for the trivial case $H = 0$.

For nontrivial cases where $H = 1, 2, 3, \ldots$, measurement dependence M can take any value in the range [0,2]. However, we will see below that nontrivial lower and upper bounds of $C_{\rm opt}$ appear.

B. Lower bound of $C_{\rm opt}$ for the local model

Proposition 1. For any local model,

$$M + 2 \leqslant C_{\rm opt}. \tag{20}$$

The inequality is tight.

Proof. We use the same notation $z_i := p(\lambda|i)$ and the labeling $i = 1, 2, 3, 4$ for $(x, y) = (0, 0), (0, 1), (1, 0), (1, 1)$ as in the previous section. Recalling the definition (4), let $i_1 < i_2 \in \{1, 2, 3, 4\}$ such that

$$M = \sum_\lambda |z_{i_1} - z_{i_2}|.$$

Adding the normalization conditions $\sum_\lambda z_{i_3} = \sum_\lambda z_{i_4} = 1$, where $i_3 < i_4 \in \{1, 2, 3, 4\} \setminus \{i_1, i_2\}$, we observe $M + 2 = \sum_\lambda (|z_{i_1} - z_{i_2}| + z_{i_3} + z_{i_4}) \leqslant \sum_\lambda \max[(z_{i_1} - z_{i_2} + z_{i_3} + z_{i_4}), \ (-z_{i_1} + z_{i_2} + z_{i_3} + z_{i_4})]$. The last expression is bounded from above by $C_{\rm opt}$ since $C_{\rm opt} = \sum_\lambda g(\boldsymbol{z}) = \sum_\lambda \max[z_1 + z_2 + z_3 - z_4, z_1 + z_2 - z_3 + z_4, z_1 - z_2 + z_3 + z_4, -z_1 + z_2 + z_3 + z_4]$.

The tightness of the inequality will be shown in Sec. IV. ■

Proposition 1 shows that for any given measurement dependence $M \in [0, 2]$, a local hidden-variable model in which the CHSH value can reach $M + 2$ exists. This reflects the often overlooked fact that a Bell inequality can be violated even by a local-hidden-variable model if measurement independence is relaxed [18–21].

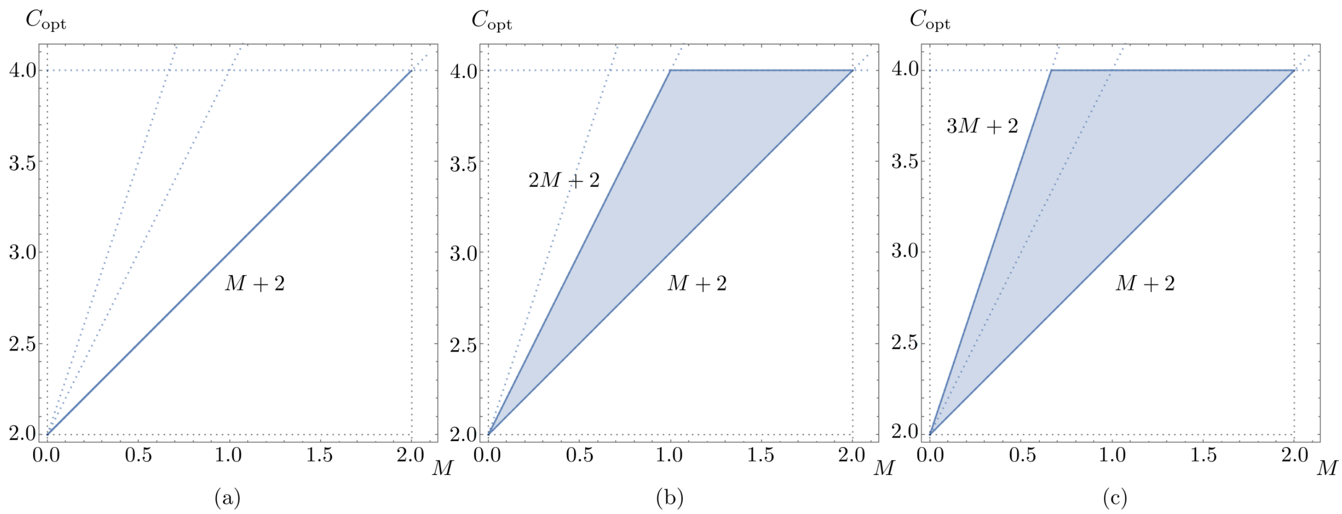

FIG. 2. Relation between the measurement dependence M and the optimal CHSH value $C_{\rm opt}$ for local models with (a) $H = 1$, (b) $H = 2$, and (c) $H \geqslant 3$. The blue shaded regions are feasible regions in local models.

C. Upper bound of $C_{\rm opt}$ for the local model

Proposition 2. For any local model,

$$C_{\rm opt} \leqslant \begin{cases} 2 & (H=0), \\ M+2 & (H=1), \\ 2M+2 & (H=2), \\ 3M+2 & (H\geqslant 3). \end{cases} \tag{21}$$

The inequalities are tight.

Since there always exists a local model in which the CHSH value reaches $C_{\rm opt}$, this result proves Theorem 1.

In particular, for $H = 1$ [$\#(\Lambda) = 2$] together with Proposition 1, we have

$$C_{\rm opt} = M + 2. \tag{22}$$

Namely, in this case, the optimal CHSH value and M have a one-to-one relation [see Fig. 2(a)].

Proof for $H = 0$. We have already shown (19), so the relation trivially holds.

For $H = 1$, we provide a direct proof of (22) [i.e., (20) and (21) simultaneously].

Proof for $H = 1$. Let $\Lambda = \{\lambda_1, \lambda_2\}$. We denote for each $i = 1, 2, 3, 4$, $z_i := p(\lambda_1|i)$, so by the normalization condition, $1 - z_i = p(\lambda_2|i)$. Without loss of generality, we can assume $z_1 \geqslant z_2 \geqslant z_3 \geqslant z_4 \geqslant 0$. Noting that $|z_1 - z_4| + |(1 - z_1) - (1 - z_4)| = 2(z_1 - z_4)$, etc., it is easy to see that $M = 2(z_1 - z_4)$. Letting $w_i = \sum_{j=1}^4 z_j - 2z_i$ $(i = 1, 2, 3, 4)$ and invoking (17), we have $C_{\rm opt} = 4 - 2[\min_i z_i + \min_i (1 - z_i)] = 2 - 2z_4 + 2z_1$. Therefore, we have shown the equality $C_{\rm opt} = 2 + M$. ■

In what follows, we let $n = H - 1 = \#(\Lambda)$ $(n \geqslant 3)$ and label a hidden variable as $\Lambda = \{1, 2, \ldots, n\}$ instead of writing $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$.

Proof for $H = 2, 3$. To prove the cases $H = 2, 3$ $(n = 3, 4)$, we use the following lemma.

Lemma 2. For $n = 3, 4$, there exist $i_*, j_* = 1, 2, 3, 4$ and $\lambda_* = 1, 2, 3$ such that

$$\sum_{\lambda=1}^{n} p(\lambda|i_\lambda) + (n-1)|p(\lambda_*|i_*) - p(\lambda_*|j_*)| \geqslant 1, \tag{23}$$

where $i_\lambda \in \{1, 2, 3, 4\}$ denotes an index which attains the minimum of $p(\lambda|i)$:

$$p(\lambda|i_\lambda) = \min_i p(\lambda|i). \tag{24}$$

The proof of this lemma is given in Appendix A.

By using (6), (17), and the notation (24), inequality (21) for $H = 2, 3$ $(n = 3, 4)$ can be rewritten as

$$\sum_{\lambda=1}^{n} p(\lambda|i_\lambda) + (n-1)\max_{i,j} \delta(P_i, P_j) \geqslant 1, \tag{25}$$

where $P_i = [p(\lambda|i)]_{\lambda\in\Lambda}$. Let i_*, j_*, λ_* be a set with which (23) is satisfied in Lemma 2. Then, we have

$$\begin{aligned} &\sum_{\lambda=1}^{n} p(\lambda|i_\lambda) + (n-1)\max_{i,j} \delta(P_i, P_j) \\ &\quad\geqslant \sum_{\lambda=1}^{n} p(\lambda|i_\lambda) + (n-1)\delta(P_{i_*}, P_{j_*}) \\ &\quad\geqslant \sum_{\lambda=1}^{n} p(\lambda|i_\lambda) + (n-1)|p(\lambda_*|i_*) - p(\lambda_*|j_*)| \geqslant 1. \end{aligned} \tag{26}$$

In particular, the second inequality follows by invoking the definition of the total variation distance in the form (5) and applying the singleton event $E = \{\lambda_*\}$. ■

The cases with $H \geqslant 4$ will be proved by reducing the problems to the case with $H = 3$ in the following manner.

Proof for $H \geqslant 4$. In a similar manner, what we have to show is that, for $n \geqslant 5$,

$$\sum_{\lambda=1}^{n} p(\lambda|i_\lambda) + 3\max_{i,j} \delta(P_i, P_j) \geqslant 1. \tag{27}$$

TABLE I. Tight model for $H = 1$ ($p \in [0, 1]$).

λ	P_{00}	P_{01}	P_{10}	P_{11}
λ_1	0	p	p	p
λ_2	1	$1-p$	$1-p$	$1-p$

We define the partition of $\Lambda = \{1, 2, \dots, n\}$ by

$$E_\gamma = \{\lambda \in \Lambda \mid i_\lambda = \gamma\}\ (\gamma = 1, \dots, 4).$$

[Recall the notation (24).] Clearly, we have $\cup_\gamma E_\gamma = \Lambda$, $E_\gamma \cap E_{\gamma'} = \emptyset$ $(\gamma \neq \gamma')$. Some of E_γ could be the empty set. Now, we introduce a new set of hidden variables $\Gamma = \{\gamma\}_{\gamma=1,2,3,4}$ with coarse-grained probabilities $\tilde{P}_i = [\tilde{p}(\gamma|i)]_\gamma$ for each measurement context $i = 1, \dots, 4$, where $\tilde{p}(\gamma|i) := P_i(E_\gamma) = \sum_{\lambda\in E_\gamma} p_i(\lambda|i)$. Following the notation (24), we denote $\tilde{i}_\gamma \in \{1, 2, 3, 4\}$ $(\gamma \in \Gamma)$, with which it holds that $\tilde{p}(\gamma|\tilde{i}_\gamma) = \min_{i=1,2,3,4} \tilde{p}(\gamma|i)$. However, by the definition of E_γ, we can assume $\tilde{i}_\gamma = \gamma$ for all γ. To see this, we need to show $\tilde{p}(\gamma|\gamma) \leqslant \tilde{p}(\gamma|i)$ for any γ and i. But $\tilde{p}(\gamma|\gamma) = P_\gamma(E_\gamma) = \sum_{\lambda\in E_\gamma} p(\lambda|\gamma)$. Since $i_\lambda = \gamma$ for any $\lambda \in E_\gamma$, we have $\sum_{\lambda\in E_\gamma} p(\lambda|\gamma) = \sum_{\lambda\in E_\gamma} p(\lambda|i_\lambda) \leqslant \sum_{\lambda\in E_\gamma} p(\lambda|i) = P_i(E_\gamma) = \tilde{p}(\gamma|i)$. Applying Lemma 2 with $n = 4$ for Γ [noting $\#(\Gamma) = 4$], $\tilde{p}$, and $\tilde{i}_\gamma$, there exist $i_*, j_*, \gamma_* = 1, \dots, 4$ such that

$$\sum_{\gamma=1}^{4} \tilde{p}(\gamma|\tilde{i}_\gamma) + 3|\tilde{p}(\gamma_*|i_*) - \tilde{p}(\gamma_*|j_*)| \geqslant 1. \tag{28}$$

Since $\tilde{i}_\gamma = \gamma$ and $i_\lambda = \gamma$ for $\lambda \in E_\gamma$, we have

$$\sum_{\gamma=1}^{4} \tilde{p}(\gamma|\tilde{i}_\gamma) = \sum_{\gamma=1}^{4}\sum_{\lambda\in E_\gamma} p(\lambda|\gamma) = \sum_{\lambda=1}^{n} p(\lambda|i_\lambda). \tag{29}$$

By applying the events E_γ in the original definition of the total variation distance (5), we have

$$\max_{i,j} \delta(P_i, P_j) \geqslant |P_{i_*}(E_{\gamma_*}) - P_{j_*}(E_{\gamma_*})|$$
$$= |\tilde{p}(\gamma_*|i_*) - \tilde{p}(\gamma_*|j_*)|. \tag{30}$$

The combination of (28), (29), and (30) implies (27).

The tightnesses for all of the cases above will be shown in the next section. ■

IV. TIGHT MODELS

In this section, we demonstrate the tightnesses of inequalities in both Propositions 1 and 2 by constructing explicit models of $p(\lambda|x, y)$ that achieve equalities. Since there is a local model with some $p(a|x, \lambda)$ and $p(b|y, \lambda)$ that achieves $C_{\rm opt}$, this demonstrates the tightness of Theorem 1.

Note that the case with $H = 0$ trivially attains the bounds shown in (19). In the following, we show the tightnesses for the upper bounds in Proposition 2 for the cases with $H = 1, 2$ and $H \geqslant 3$.

Case with $H = 1$. Let $\Lambda = \{\lambda_1, \lambda_2\}$. Let $P_{xy} = [p(\lambda|x, y)]_{\lambda\in\Lambda}$ be given by Table I with a one-parameter $p \in [0, 1]$. For this model, we have $C_{\rm opt} = 2p + 2$ and $M = 2p$. Hence, $C_{\rm opt} = M + 2$, where M runs over [0,2] for $p \in [0, 1]$.

TABLE II. Tight model for $H = 2$.

λ	P_{00}	P_{01}	P_{10}	P_{11}
		$p \in [0, 1/2]$		
λ_1	0	p	p	p
λ_2	p	0	p	p
λ_3	$1-p$	$1-p$	$1-2p$	$1-2p$
		$p \in [1/2, 1]$		
λ_1	0	$1-p$	$1-p$	$2p-1$
λ_2	p	0	p	$1-p$
λ_3	$1-p$	p	0	$1-p$

Case with $H = 2$. Let $\Lambda = \{\lambda_1, \lambda_2, \lambda_3\}$. Let $P_{xy} = [p(\lambda|x, y)]_{\lambda\in\Lambda}$ be given by Table II with a one-parameter $p \in [0, 1]$. For this model, we have $C_{\rm opt} = 4p + 2$ and $M = 2p$ for $p \in [0, 1/2]$; hence, $C_{\rm opt} = 2M + 2$, where M runs over [0,1]. And we have $C_{\rm opt} = 4$ and $M = 2p$ for $p \in [1/2, 1]$; hence, $C_{\rm opt} = 4$, where M runs over [1,2].

Case with $H \geqslant 3$. Let $\Lambda = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \dots\}$. Let $P_{xy} = [p(\lambda|x, y)]_{\lambda\in\Lambda}$ be given by Table III with a one-parameter $p \in [0, 1]$. For this model, we have $C_{\rm opt} = 6p + 2$ and $M = 2p$ for $p \in [0, 1/3]$; hence, $C_{\rm opt} = 3M + 2$, where M runs over [0, 2/3]. And we have $C_{\rm opt} = 4$ and $M = 2p$ for $p \in [1/3, 1]$; hence $C_{\rm opt} = 4$, where M runs over [1,2].

Next, the lower bound in Proposition 1 can be attained with the model given in Table I with the trivial addition of $p(\lambda|x, y) = 0$ for λ_i $(i \geqslant 3)$.

Moreover, we can show that the regions between the lower and upper bounds of $C_{\rm opt}$ are feasible for any M and H. Technically, this is far from trivial since both M and $C_{\rm opt}$ are not affine functions of $p(\lambda|x, y)$ in general. However, for the probabilities given in Table I (with a trivial extension for any $H \geqslant 0$) for the lower bound and the ones given in Tables II and III for the upper bound, we can easily show that both M and $C_{\rm opt}$ are affine for their convex combination. Hence, with these special choices of lower and upper bounds, their convex combinations fill the sandwiched regions. In Fig. 2,

TABLE III. Tight model for $H \geqslant 3$.

λ	P_{00}	P_{01}	P_{10}	P_{11}
		$p \in [0, 1/3]$		
λ_1	0	p	p	p
λ_2	p	0	p	p
λ_3	p	p	0	p
λ_4	$1-2p$	$1-2p$	$1-2p$	$1-3p$
λ_5	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		$p \in [1/3, 1]$		
λ_1	0	$\frac{1-p}{2}$	$\frac{1-p}{2}$	p
λ_2	p	0	$\frac{1-p}{2}$	$\frac{1-p}{2}$
λ_3	$\frac{1-p}{2}$	p	0	$\frac{1-p}{2}$
λ_4	$\frac{1-p}{2}$	$\frac{1-p}{2}$	p	0
λ_5	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

the feasible regions for M and $C_{\rm opt}$ given H are shown by blue shaded regions.

V. CONCLUSION AND DISCUSSION

In this paper, we have introduced the measure of hiddenness H and investigated a trade-off relation between H and the measurement dependence M for any local-hidden-variable models. In the CHSH setting, we derived a relaxed Bell inequality (10) that generalizes the Bell-CHSH inequality. Note that the introduction of hiddenness can generalize all known relaxed Bell's inequalities to date, especially in terms of the tightness of these inequalities. For instance, while inequality (11) was already known to give the tight bound within local models [25], our results show that the smallest cardinality of the set of hidden variables to attain the bound must be 4 (or, equivalently, $H = 3$). Given its potential impact, this aspect would hold particular importance for applications in computer science and also in cryptography. Interestingly, the structure of the trade-off changes between $H \leqslant 2$ and $H \geqslant 3$: While the trade-off reduces to Hall's inequality (11) for $H \geqslant 3$, a nontrivial dependence for H appears when $H \leqslant 2$. Moreover, the trade-off relation completely characterizes the range of measurement-dependent local models.

In the present paper, hiddenness H was introduced simply by the cardinality of the set of hidden variables, making H a discrete quantity. In addition, this measure does not reflect the statistics of the hidden variables. In an upcoming paper [38], we overcome this disadvantage by introducing another measure of hiddenness that uses the max entropy. This measure will provide a better reflection of the hidden-variable statistics. Furthermore, it would be interesting to generalize the results obtained in this paper by relaxing the condition of locality. To do this, we need to introduce measures for both parameter and outcome dependence.

ACKNOWLEDGMENTS

We are grateful to M. Hall for his valuable comments and discussion. We also thank R. Takakura and Y. Kitajima for their fruitful comments. G.K. is supported in part by JSPS KAKENHI Grant No. 17K18107.

APPENDIX A: PROOFS OF LEMMAS

Proof of Lemma 1. Since the objective function $f_z(A, A', B, B') := z_1AB + z_2AB' + z_3A'B - z_4A'B'$ in (15) is affine for all variable A, A', B, B', the maximum is attained by the extreme points $A, A', B, B' = \pm 1$. There are four cases to consider: (I) $A = A', B = B'$, (II) $A = A', B = -B'$, (III) $A = -A', B = B'$, and (IV) $A = -A', B = -B'$. Direct computations show that in case I $f = \pm(z_1 + z_2 + z_3 - z_4) = \pm(z - 2z_4)$, in case II $f = \pm(z - 2z_2)$, in case III $f = \pm(z - 2z_3)$, and in case IV $f = \pm(z - 2z_1)$, where $z := \sum_{j=1}^{4} z_j$. This shows that

$$\max_{A,A',B,B'\in[-1,1]} f_z(A, A', B, B') = \max_i |z - 2z_i|. \tag{A1}$$

However, as is easily shown, the maximum is always attained by a positive $z - 2z_i$. ■

Proof of Lemma 2. We provide the proof for $n = 3$. The case for $n = 4$ can be shown in parallel. Assume for all $i, j \in \{1, 2, 3, 4\}$ and $\lambda \in \{1, 2, 3\}$ that

$$\sum_{\lambda'=1}^{3} p(\lambda'|i_{\lambda'}) + 2|p(\lambda|i) - p(\lambda|j)| < 1. \tag{A2}$$

There always exists an $s \in \{1, 2, 3, 4\}$ which is different from all i_1, i_2, and i_3. Moreover, for each λ, one can choose $j_\lambda, k_\lambda \in \{1, 2, 3, 4\}$ such that i_λ, j_λ, k_λ, and s are all different from each other; that is, for any λ, $\{i_\lambda, j_\lambda, k_\lambda, s\} = \{1, 2, 3, 4\}$.

Applying the case with $\lambda = 1$, $i = i_1, j = j_1$ to (A2), one has $\sum_{\lambda'=1}^{3} p(\lambda'|i_{\lambda'}) + 2|p(1|i_1) - p(1|j_1)| = \sum_{\lambda'=1}^{3} p(\lambda'|i_{\lambda'}) - 2[p(1|i_1) - p(1|j_1)] = -p(1|i_1) + p(2|i_2) + p(3|i_3) + 2p(1|j_1)$. Thus, one has

$$-p(1|i_1) + p(2|i_2) + p(3|i_3) + 2p(1|j_1) < 1.$$

Also, for the case with $\lambda = 1$, $i = i_1$, $j = k_1$, one has

$$-p(1|i_1) + p(2|i_2) + p(3|i_3) + 2p(1|k_1) < 1.$$

Similarly, for $\lambda = 2, 3$,

$$\begin{aligned} p(1|i_1) - p(2|i_2) + p(3|i_3) + 2p(2|j_2) &< 1,\\ p(1|i_1) - p(2|i_2) + p(3|i_3) + 2p(2|k_2) &< 1,\\ p(1|i_1) + p(2|i_2) - p(3|i_3) + 2p(3|j_3) &< 1,\\ p(1|i_1) + p(2|i_2) - p(3|i_3) + 2p(3|k_3) &< 1. \end{aligned}$$

Summing all six inequalities above and dividing by 2, one gets

$$\begin{aligned} &p(1|i_1) + p(1|j_1) + p(1|k_1) + p(2|i_2) + p(2|j_2) + p(2|k_2)\\ &\quad + p(3|i_3) + p(3|j_3) + p(3|k_3) < 3. \end{aligned}$$

Since $i_\lambda \neq j_\lambda \neq k_\lambda \neq s$ for all $\lambda \in \{1, 2, 3\}$, the left-hand side can be grouped as

$$\sum_{i\neq s} \sum_{\lambda=1}^{3} p(\lambda|i) = 3,$$

which is contradictory and leads to inequality (A2). ■

APPENDIX B: PROOF FOR INFINITE MODELS

In this Appendix, we provide a proof of Theorem 1 for the case of an uncountable local-hidden-variable model with $\#(\Lambda) = \infty$, where inequality (10) reduces to inequality (11). We need to replace (3) and (4) with

$$p(a, b|x, y) = \int d\lambda\, p(\lambda|x, y)p(a|x, \lambda)p(b|y, \lambda), \tag{B1}$$

$$M := \sup_{x,y,x',y'} \int d\lambda \left|p(\lambda|x, y) - p(\lambda|x', y')\right|, \tag{B2}$$

where $p(\lambda|x, y)$ is the probability density of λ conditioned on the measurement context (x, y).

The idea of the proof is to transform the uncountable hidden-variable model with $\#(\Lambda) = \infty$ into another hidden-variable model with $\#(\tilde{\Lambda}) = 2^4$:

$$\tilde{\Lambda} = \{\tilde{\lambda} = (a_0, a_1, b_0, b_1) \mid a_0, a_1, b_0, b_1 = 0, 1\}. \tag{B3}$$

We introduce the probability distributions on $\tilde{\Lambda}$ as

$$q(a_0, a_1, b_0, b_1|x, y) := \int d\lambda p(\lambda|x, y)p(a_0|x = 0, \lambda) p(a_1|x = 1, \lambda)p(b_0|y = 0, \lambda)p(b_1|y = 1, \lambda) \tag{B4}$$

and a deterministic model given by

$$q(a, b|x, y) := \sum_{a_0,a_1,b_0,b_1=0,1} q(a_0, a_1, b_0, b_1|x, y)\delta_{aa_x}\delta_{bb_y}.$$

It can easily be shown that $p(a, b|x, y) = q(a, b|x, y)$, and hence, the CHSH value does not change:

$$C = \tilde{C}.$$

Let $\tilde{M}$ be the measurement dependence for this new hidden-variable model on $\tilde{\Lambda}$:

$$\tilde{M} := \sup_{x,y,x',y'} \sum_{a_0,a_1,b_0,b_1=0,1} |q(a_0, a_1, b_0, b_1|x, y) - q(a_0, a_1, b_0, b_1|x', y')|. \tag{B5}$$

By substituting (B4) into (B5), we have

$$\begin{aligned}
\tilde{M} &= \sup_{x,y,x',y'} \sum_{a_0,a_1,b_0,b_1=0,1} \left| \int d\lambda[p(\lambda|x, y) - p(\lambda|x', y')]p(a_0|x = 0, \lambda)p(a_1|x = 1, \lambda)p(b_0|y = 0, \lambda)p(b_1|y = 1, \lambda)\right| \\
&\leqslant \sup_{x,y,x',y'} \sum_{a_0,a_1,b_0,b_1=0,1} \int d\lambda \left|p(\lambda|x, y) - p(\lambda|x', y')\right| p(a_0|x = 0, \lambda)p(a_1|x = 1, \lambda)p(b_0|y = 0, \lambda)p(b_1|y = 1, \lambda) \\
&= \sup_{X,Y,X',Y'} \int d\lambda \left|p(\lambda|x, y) - p(\lambda|x', y')\right| \\
&= M,
\end{aligned}$$

where we have used the triangle inequality for the integral and the normalization conditions for $p(a|x, \lambda)$ and $p(b|y, \lambda)$.

Since the model on $\tilde{\Lambda}$ is finite, we have already shown that

$$C \leqslant 3\tilde{M} + 2.$$

However, as shown above, we have $\tilde{M} \leqslant M$; this completes the proof of (11) in the uncountable model. ■

[1] J. S. Bell, Phys. Phys. Fiz. **1**, 195 (1964).
[2] J. S. Bell and A. Aspect, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 2004).
[3] W. Myrvold, M. Genovese, and A. Shimony, in *The Stanford Encyclopedia of Philosophy*, fall 2021 ed., edited by E. N. Zalta (Metaphysics Research Lab, Stanford University, Stanford, CA, 2021).
[4] S. J. Freedman and J. F. Clauser, Phys. Rev. Lett. **28**, 938 (1972).
[5] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
[6] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, Phys. Rev. Lett. **81**, 3563 (1998).
[7] B. Hensen, N. Kalb, M. S. Blok, A. E. Dréau, A. Reiserer, R. F. L. Vermeulen, R. N. Schouten, M. Markham, D. J. Twitchen, K. Goodenough, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Sci. Rep. **6**, 30289 (2016).
[8] L. K. Shalm *et al.*, Phys. Rev. Lett. **115**, 250402 (2015).
[9] M. Giustina *et al.*, Phys. Rev. Lett. **115**, 250401 (2015).
[10] N. Gisin, *Quantum Chance: Nonlocality, Teleportation and Other Quantum Marvels* (Springer, Cham, 2014).
[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, 2000).
[12] M. Hayashi, S. Ishizaka, A. Kawachi, G. Kimura, and T. Ogawa, *Introduction to Quantum Information Science* (Springer, Berlin, 2015).
[13] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[14] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[15] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).
[16] 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, Nature (London) **464**, 1021 (2010).
[17] R. Colbeck and A. Kent, J. Phys. A **44**, 095305 (2011).
[18] J. Bell, A. Shimony, M. Horne, and J. Clauser, Dialectica **39**, 85 (1985).
[19] C. H. Brans, Int. J. Theor. Phys. **27**, 219 (1988).
[20] M. J. W. Hall, Phys. Rev. Lett. **105**, 250404 (2010).
[21] J. Barrett and N. Gisin, Phys. Rev. Lett. **106**, 100406 (2011).
[22] J. Conway and S. Kochen, Found. Phys. **36**, 1441 (2006).
[23] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[24] M. J. W. Hall, Phys. Rev. A **82**, 062117 (2010).

[25] M. J. W. Hall, Phys. Rev. A **84**, 022102 (2011).
[26] M. Banik, M. R. Gazi, S. Das, A. Rai, and S. Kunkri, J. Phys. A **45**, 205301 (2012).
[27] A. S. Friedman, A. H. Guth, M. J. W. Hall, D. I. Kaiser, and J. Gallicchio, Phys. Rev. A **99**, 012121 (2019).
[28] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, and N. Gisin, Phys. Rev. Lett. **113**, 190402 (2014).
[29] M. J. W. Hall and C. Branciard, Phys. Rev. A **102**, 052228 (2020).
[30] J. P. Jarrett, Nous **18**, 569 (1984).
[31] E. Nelson, in *Stochastic Processes in Classical and Quantum Systems*, edited by S. Albeverio, G. Casati, and D. Merlini (Springer, Berlin, 1986), pp. 438–469.
[32] M. Ghadimi, Phys. Rev. A **104**, 032205 (2021).
[33] D. E. Koh, M. J. W. Hall, Setiawan, J. E. Pope, C. Marletto, A. Kay, V. Scarani, and A. Ekert, Phys. Rev. Lett. **109**, 160404 (2012).
[34] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times* (American Mathematical Society, Providence, 2017).
[35] C. Fuchs and J. van de Graaf, IEEE Trans. Inf. Theory **45**, 1216 (1999).
[36] G. Kimura, K. Nuida, and H. Imai, Rep. Math. Phys. **66**, 175 (2010).
[37] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[38] R. Takakura, K. Morisue, I. Watanabe, and G. Kimura, arXiv:2208.13634.