Impossibility of coin flipping in generalized probabilistic theories via discretizations of semi-infinite programs

Jamie Sikora[1,*] and John H. Selby[2,1,†]
[1]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[2]*ICTQT, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland*

(Received 24 August 2019; revised 9 April 2020; accepted 23 June 2020; published 23 October 2020)

Coin flipping is a fundamental cryptographic task where spatially separated Alice and Bob wish to generate a fair coin flip over a communication channel. It is known that ideal coin flipping is impossible in both classical and quantum theory. In this work, we give a short proof that it is also impossible in generalized probabilistic theories under the generalized no-restriction hypothesis. Our proof relies crucially on a formulation of cheating strategies as semi-*infinite* programs, i.e., cone programs with infinitely many constraints. This introduces a formalism which may be of independent interest to the quantum community.

DOI: 10.1103/PhysRevResearch.2.043128

I. INTRODUCTION

In this paper, we consider the possibility of cryptography in theories more general than quantum or classical theory. One may ask why this is a worthwhile endeavour, and for this we give several reasons. The first reason is to future-proof current results, which is important in the context of cryptography. While developing quantum cryptography and computation, the community quickly came to realize that classical cryptography results need to be re-evaluated for the new quantum era. Since results in quantum cryptography typically rely on the validity of quantum mechanics being a faithful description of nature, these too all have to re-evaluated if quantum theory is one day superseded by a new theory, regardless of how minor or radical the departure from quantum mechanics is. Another reason is to gain a better understanding of results in quantum theory. For instance, it is insightful to sit back and think about what parts of quantum theory were needed to prove a result. Did we require entanglement? Were we just assuming these states are in superposition? Can we reprove this only assuming the no-signaling principle? By answering such questions, we gain a better understanding of quantum mechanics itself as well as the resources necessary for performing particular tasks.

In this and many other works in cryptography, optimization theory is a key ingredient in the analysis. On a high level, we want to maximize how much someone can "cheat" a protocol, whereby it is understood that the inability to cheat translates into security, and vice versa. The goal is often to design protocols which minimize cheating. We, however, take the opposite approach in this work and prove a limitation on designing *any* protocol for a particular task, namely coin flipping, discussed below.

II. COIN FLIPPING

Coin flipping is the cryptographic task where Alice and Bob generate a random bit b over a communication channel such that when Alice and Bob are honest, both output the same bit b and this bit is uniformly random [1]. Coin flipping is a primitive that is used mainly for building larger, more sophisticated cryptographic protocols in the two-party setting, and hence an understanding of its properties, along with its security limitations, is important. For example, coin flipping has been used in the creation of optimal oblivious transfer protocols [2], is related to bit commitment (see, for example, Refs. [3–5]), and variants have been studied such as weak coin flipping [6], unbalanced coin flipping [7], and die rolling [8]. Moreover, since secure oblivious transfer implies secure bit commitment [9] which in turn implies secure coin flipping,[1] proving the insecurity of coin flipping in a generalized probabilistic theory (GPT) setting (as we do in this paper) automatically implies the insecurity of these other tasks, as well as any others that imply secure coin flipping.

More formally, the coin flipping task is as follows. Suppose Alice has a set of strategies (basically, a description of how she interacts with Bob) given by the set $\mathcal{A}$ and Bob has a set of strategies given by the set $\mathcal{B}$. We do not just consider deterministic strategies but also those that occur as the result of some measurement procedure. We denote the probability of a pair of strategies occurring as $\mathsf{Prob}(A, B)$ which is between 0 and 1 for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

A coin-flipping protocol consists of the following:

*Present address: Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada and Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, United States; sikora@vt.edu
†john.h.selby@gmail.com

[1]This is obvious from the definitions.

(1) A triple of strategies for Alice $(A_0, A_1, A_{\text{abort}})$ which correspond to the measurement outcomes of some deterministic strategy A_{det} and

(2) A triple of strategies for Bob $(B_0, B_1, B_{\text{abort}})$ which correspond to the measurement outcomes of some deterministic strategy B_{det},

satisfying

$$\mathsf{Prob}(A_b, B_b) = 1/2 \ \text{ for } \ b \in \{0, 1\}. \tag{1}$$

The conditions above ensure that the protocol behaves as expected, that the bit b is uniform and shared between Alice and Bob. Ideally, we wish that neither Alice nor Bob can cheat by digressing from protocol and disturbing the conditions given by (1). However, this may not be the case, and as such, we need to measure this disturbance. The security measure in coin flipping is given by the amount a dishonest Alice or a dishonest Bob can bias the output distribution away from uniform. To make this formal, we define the following symbols:

(1) $P^*_{\text{Alice},b}$: The maximum probability that dishonest Alice can force honest Bob to accept the outcome b.

(2) $P^*_{\text{Bob},b}$: The maximum probability that dishonest Bob can force honest Alice to accept the outcome b.

(3) ϵ: The bias of the coin-flipping protocol defined as

$$\epsilon := \max\{P^*_{\text{Alice},0}, P^*_{\text{Alice},1}, P^*_{\text{Bob},0}, P^*_{\text{Bob},1}\} - 1/2. \tag{2}$$

We wish to design protocols such as to minimize ϵ, with a perfect protocol having $\epsilon = 0$. In classical and quantum theory, this is known to be impossible [10,11]. In this work, we show that under some assumptions on $\mathcal{A}$ and $\mathcal{B}$, ϵ can be lower bounded by a positive constant, thus showing near-perfect coin flipping is impossible in any theory satisfying those assumptions.

To study the range of possible ϵ, we need to study the four quantities $P^*_{\text{Alice},0}$, $P^*_{\text{Alice},1}$, $P^*_{\text{Bob},0}$, and $P^*_{\text{Bob},1}$. Let us first consider $P^*_{\text{Bob},0}$. We can write this succinctly by the rudimentary optimization problem:

$$P^*_{\text{Bob},0} = \sup_{B \in \mathcal{B}} \{\mathsf{Prob}(A_0, B)\}. \tag{3}$$

This optimization problem exactly captures how much Bob can force Alice to output 0 maximized over all physical strategies he can perform. Before studying this problem using optimization theory, we require a mathematical structure on the quantities involved. We now discuss such a structure which is given by the study of generalized probabilistic theories.

III. GENERALIZED PROBABILISTIC THEORIES (GPTs)

To study (3) more generally than quantum and classical theory, we require a more general setting for physical theories. Here we work in the framework of generalized probabilistic theories which formalizes any physical theory with an operational description. There have been many approaches to GPTs; see, for example, Refs. [12–21] for introductions to these frameworks. GPTs have been successfully used for studying cryptography [13,22–27] and computation [28–36] in theories more general than quantum theory. We, however, do not actually need to introduce the full framework of GPTs for the purposes of this work. Instead, we just consider the structure that any such theory would impose on the sets of strategies for Alice and Bob.

As mentioned above, we do not just want to consider the strategies which occur deterministically, but those which may correspond to obtaining a particular outcome in some experiment. That is, given a strategy $A \in \mathcal{A}$ for Alice and a strategy $B \in \mathcal{B}$ for Bob, we obtain a probability $\mathsf{Prob}(A, B)$ that these two strategies jointly occur. In particular, there is always a "zero strategy" $0 \in \mathcal{A}$ such that $\mathsf{Prob}(0, B) = 0$ for all $B \in \mathcal{B}$. Conceptually, one can think of this as Alice aborting the protocol or simply not taking part in the first place.

First, we assume that these spaces of strategies are *convex* where we interpret convex combinations as *probabilistic mixtures*. That is, we assume that

$$pA_1 + (1 - p)A_2 \tag{4}$$

is in the set $\mathcal{A}$ and represents the strategy where with probability p Alice uses strategy A_1 and with probability $1 - p$ Alice uses strategy A_2. Given this understanding of the convex structure, the calculated probabilities must satisfy

$$\mathsf{Prob}\left(\sum_i p_i A_i, B\right) = \sum_i p_i \,\mathsf{Prob}(A_i, B), \tag{5}$$

where the set $\{p_i\}$ form a probability distribution. An equivalent equation holds for convex combinations of Bob's strategies. This means that a strategy for Alice induces a linear functional on the space of strategies for Bob (and vice versa).

Rather than working directly with the spaces of strategies $\mathcal{A}$ and $\mathcal{B}$, we work with operational equivalence classes of strategies. We say that two strategies A_1 and A_2 are operationally equivalent if

$$\mathsf{Prob}(A_1, B) = \mathsf{Prob}(A_2, B), \quad \forall B \in \mathcal{B} \tag{6}$$

and similarly for Bob's strategies. We denote these equivalence classes as $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$. (While we are working with the equivalence classes of strategies, our main result applies to the original strategies themselves; the equivalence classes just provide a convenient tool for our proof. We elaborate on this in the Appendix B.)

Note that our earlier assumptions imply that $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$ are both convex sets in some vector space V which are bounded and have nonempty interior. (For completeness, we prove that the sets are bounded in Appendix B). Moreover, we assume that the vector space V is finite-dimensional. This assumption is typically made in the study of GPTs for technical convenience. It can, however, be motivated by the idea that in a tomographic characterization of the strategies of Alice, one can only, in practice, perform a finite number of different experiments and therefore we must characterize the strategies by a finite number of probabilities.

Employing the Riesz representation theorem [37] in the case of linear functionals on finite-dimensional vector spaces, one can show that we can always compute the probabilities as

$$\mathsf{Prob}(A, B) = \langle \tilde{A}, \tilde{B} \rangle. \tag{7}$$

From now on we take $\tilde{\mathcal{A}}$ as the set of Alice's strategies (similarly $\tilde{\mathcal{B}}$ as the set of Bob's strategies) and hence drop the

tildes for convenience as the strategy representation should be clear from context.

We can now rewrite the optimization problem (3) in the form

$$P^*_{\mathrm{Bob},0} = \sup_{B\in\mathcal{B}}\{\langle A_0, B\rangle\}. \tag{8}$$

Because of the convex structure of the set $\mathcal{B}$, this is a convex optimization problem. However, since we want to prove general bounds on cheating, we require more structure on the sets $\mathcal{A}$ and $\mathcal{B}$ for our analysis.

IV. A PHYSICAL ASSUMPTION

Clearly some assumption on the sets $\mathcal{A}$ and $\mathcal{B}$ is required to prove anything meaningful. For example, consider any physical theory and restrict both Alice and Bob to a set of strategies that are ϵ-close to their honest strategies. This allows us to define a (rather boring) GPT in which ideal coin flipping is possible up to some small error. To avoid GPTs with these unnecessary restrictions, we make the assumption that any mathematically feasible strategy for Bob can be physically realized.

To formally define this lack of restriction for Bob, we start with defining two important quantities studied in convex analysis. The *polar set* of the set C is given as

$$C^o := \{W : \langle W, Z\rangle \leqslant 1,\ \forall Z \in C\} \tag{9}$$

and its *dual cone* is given as

$$C^* := \{W : \langle W, Z\rangle \geqslant 0,\ \forall Z \in C\}. \tag{10}$$

Notice we have $\mathcal{B} \subseteq \mathcal{A}^* \cap \mathcal{A}^o$ and $\mathcal{A} \subseteq \mathcal{B}^* \cap \mathcal{B}^o$ because *every* choice of strategies for Alice and Bob yields a proper probability.

We can now define our physical assumption.

Definition 1. The *generalized no-restriction hypothesis for Bob* states that $\mathcal{B} = \mathcal{A}^* \cap \mathcal{A}^o$.

To support this assumption, one can argue that if Alice knows that her set of strategies is given as $\mathcal{A}$ then to be able to guarantee security against Bob she should not make any assumptions about what Bob can do. In other words, we also maximize over all physical theories, which in this case translates to allowing Bob to have the largest set of strategies as possible.

This is closely related to the (standard) no-restriction hypothesis [19], which is a commonly used assumption in the study of GPTs that can be expressed as the idea that all mathematically possible measurements are physically allowed. Here, we generalize this idea to the level of arbitrary strategies.

One could equally well consider Bob's perspective and assume the generalized no-restriction hypothesis for Alice, i.e., $\mathcal{A} = \mathcal{B}^* \cap \mathcal{B}^o$. Surprisingly these two assumptions are not equivalent; see Fig. 1 for an example of this fact. However, for the purposes of this work, we need to only assume it for one party. We henceforth assume it for Bob, but by symmetry the following arguments can be adapted to the case where it is assumed instead for Alice.

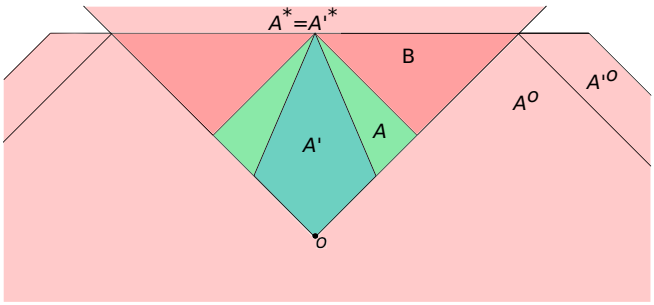

FIG. 1. Alice has two strategy sets A and A' corresponding to two different theories. We see that B is equal to both $A^* \cap A^o$ and $(A')^* \cap (A')^o$ and hence the generalized no-restriction hypothesis for Bob does not imply the same for Alice. We do have that $A = B^* \cap B^o$, so sometimes the assumption does hold for both Alice and Bob.

V. OPTIMIZATION ANALYSIS

Under this assumption, we can now clean up the optimization problem for Bob (8) as

$$P^*_{\mathrm{Bob},0} = \sup_{B\in\mathcal{A}^*\cap\mathcal{A}^o}\{\langle A_0, B\rangle\} \tag{11}$$

$$= \sup_{B\in\mathcal{A}^*}\{\langle A_0, B\rangle : \langle B, A\rangle \leqslant 1, \forall A \in \mathcal{A}\}. \tag{12}$$

This type of optimization problem is called a *semi-infinite program* since the variable B is finite-dimensional but there are infinitely many constraints. (Note that this class is not the same as the more popular class of optimization problems called semidefinite programs.) Semi-infinite programming has a rich theory (see, for example, Ref. [38]), although it has yet to be used to study quantum theory or its generalizations, as far as we are aware.

For our needs, it suffices to look at relaxations of $P^*_{\mathrm{Bob},0}$ where we optimize instead using a discretization of the infinite set $\mathcal{A}$. To this end, we define a mesh, denoted here as $\mathcal{A}_\delta$, parameterized by a fineness measure $\delta > 0$, such that it has the following properties:

(1) $\mathcal{A}_\delta$ is finite, contains a basis for V, and is contained in $\mathcal{A}$, and

(2) $\forall A \in \mathcal{A}$, $\exists X \in \mathcal{A}_\delta$ such that $\|X - A\|_2 \leqslant \delta$.

Note that such a discretization always exists since $\mathcal{A}$ is bounded.

We now consider the discretized version of this optimization problem defined to optimize using $\mathcal{A}_\delta$ instead, as shown below:

$$P^\delta_{\mathrm{Bob},0} = \sup_{B\in\mathcal{A}^*}\{\langle A_0, B\rangle : \langle B, X\rangle \leqslant 1, \forall X \in \mathcal{A}_\delta\}. \tag{13}$$

First, note that we have $P^*_{\mathrm{Bob},0} \leqslant P^\delta_{\mathrm{Bob},0}$ since it relaxes (12) as $\mathcal{A}_\delta \subseteq \mathcal{A}$. Furthermore, since there are finitely many constraints, this is a (traditional) cone program, making it easier to analyze. Recently, there have been several applications of cone programming to the study of GPTs [22–24,39–41] and to quantum theory [5,42–45].

As expected, as one decreases δ (the fineness measure of the mesh), we have that $\mathcal{A}_\delta$ becomes a better approximation of the set $\mathcal{A}$. In particular, we have the lemma below.

Lemma 2. $\lim_{\delta\to 0^+} P^\delta_{\mathrm{Bob},0} = P^*_{\mathrm{Bob},0}$.

Proof. We first show that the feasible region of (13) is bounded. To this end, we define the function

$$f(Y) = \max_{X \in \mathcal{A}_\delta} \{|\langle X, Y\rangle|\}, \tag{14}$$

which is finite since $\mathcal{A}_\delta$ is finite. It can be easily checked that this is a norm (since $\mathcal{A}_\delta$ contains a basis) and is bounded for all B satisfying the constraints of (13). Since all norms are equivalent in finite-dimensional vector spaces, we know there exists a $\tau > 0$ such that $\|B\|_2 \leqslant \tau$ for all B feasible in (13).

Fix B feasible in (13) and $A \in \mathcal{A}$. We now wish to scale B by some constant $c > 0$ to ensure $\langle A, cB\rangle \leqslant 1$ [and thus cB is feasible in (12)]. Then, for $X \in \mathcal{A}_\delta$ δ-close to A, we have

$$\langle B, A\rangle = \langle B, X\rangle + \langle B, A - X\rangle \tag{15}$$

$$\leqslant \langle B, X\rangle + \|B\|_2 \|A - X\|_2 \tag{16}$$

$$\leqslant 1 + \tau\delta. \tag{17}$$

Thus, $\dfrac{1}{1+\tau\delta} B$ is feasible in (12). This implies that

$$P^*_{\text{Bob},0} \leqslant P^\delta_{\text{Bob},0} \leqslant (1 + \tau\delta)\, P^*_{\text{Bob},0}. \tag{18}$$

Taking limits finishes the proof. ∎

We now prove a lower bound on the product of Alice's cheating probability and the relaxation of Bob's cheating probability. This is the key step in proving our main result, which takes advantage of the simplified structure of the relaxed problem.

Lemma 3. $P^*_{\text{Alice},0} \cdot P^\delta_{\text{Bob},0} \geqslant 1/2$, for all $\delta > 0$.

Proof. Let $B \in \text{int}(\mathcal{B}) = \text{int}(\mathcal{A}^* \cap \mathcal{A}^o) \subseteq \text{int}(\mathcal{A}^*)$ which exists since $\mathcal{B}$ has nonempty interior by construction. Then $B' := \frac{1}{2}B$ satisfies $B' \in \text{int}(\mathcal{A}^*)$ and $\langle B', X\rangle < 1$ for all $X \in \mathcal{A}_\delta$. This is known as a *strictly feasible* solution. Since $P^\delta_{\text{Bob},0}$ is bounded from above by Eq. (18), the strong duality theorem for cone programming (see, for example, Ref. [46]) states that $P^\delta_{\text{Bob},0}$ is equal to

$$\min_{y_X \geqslant 0} \left\{ \sum_{X \in \mathcal{A}_\delta} y_X : \sum_{X \in \mathcal{A}_\delta} y_X X - A_0 \in (\mathcal{A}^*)^* \right\} \tag{19}$$

and this problem attains[2] an optimal solution $\{y'_X\}$. Thus, we have $P^\delta_{\text{Bob},0} = \sum_{X \in \mathcal{A}_\delta} y'_X$. Define

$$A := \frac{1}{P^\delta_{\text{Bob},0}} \sum_{X \in \mathcal{A}_\delta} y'_X X = \sum_{X \in \mathcal{A}_\delta} \left(\frac{y'_X}{\sum_{\tilde{X} \in \mathcal{A}_\delta} y'_{\tilde{X}}} \right) X. \tag{20}$$

Notice that $A \in \mathcal{A}$ by convexity and $A - \frac{1}{P^\delta_{\text{Bob},0}} A_0 \in (\mathcal{A}^*)^*$ by the constraints in (19). Suppose Alice uses A as her strategy to force Bob to accept outcome 0. Then we have

$$P^*_{\text{Alice},0} \geqslant \langle A, B_0\rangle \geqslant \frac{1}{P^\delta_{\text{Bob},0}} \langle A_0, B_0\rangle = \frac{1}{2P^\delta_{\text{Bob},0}} \tag{21}$$

since $B_0 \in \mathcal{B} \subseteq \mathcal{A}^*$ and $\langle A_0, B_0\rangle = 1/2$ from Eq. (1). ∎

[2]Note that attainment of an optimal *dual* solution is not always stated explicitly in the proofs of strong duality, but it is indeed the case.

By combining the two lemmas, we have that $P^*_{\text{Alice},0} \cdot P^*_{\text{Bob},0} \geqslant 1/2$, and therefore the maximum of the two probabilities is at least $1/\sqrt{2}$. This gives the same lower bound on the bias Kitaev gave for the case of quantum theory [11], which was later reproved by Gutoski and Watrous using a representation of quantum strategies [47].

Theorem 4. Any coin-flipping protocol in a GPT satisfying the *generalized no-restriction hypothesis for Bob* (and/or Alice) satisfies $\epsilon \geqslant 1/\sqrt{2} - 1/2 \approx 0.207$. In particular, either Alice or Bob can force an outcome with probability at least $1/\sqrt{2}$.

Since quantum theory satisfies the generalized no-restriction hypothesis for both Alice and Bob [47], we have another proof that coin flipping is impossible in quantum theory.

VI. DISCUSSION

What is perhaps unusual about our main result is that we have found a numerical lower bound that holds for any GPT satisfying the generalized no-restriction hypothesis for Alice and/or Bob. Typically results in the study of GPTs either show something is possible or impossible, or consider a specific GPT (whose structure can be exploited). This is relevant for cryptographic purposes as well. If our result was simply saying that perfect coin flipping is impossible, then this does not rule out the existence of protocols with small bias, which would be enough for all intents and purposes. Theorem 4 says that near-perfect protocols cannot exist either. Moreover, the constant lower bound shows that the security of coin-flipping protocols cannot be boosted in the sense that a protocol with bias $\epsilon < 1/2$ cannot be used in a composition to reduce the bias arbitrarily close to 0.

The main technique in this work is our treatment of semi-infinite programs, in particular, how we discretized them into cone programs. We hope that our use of semi-infinite programs will raise awareness of this formalism for future uses in quantum theory and physics by breaking roadblocks when formulating difficult problems as optimization problems.

A. Future work

This bound on coin flipping is (asymptotically) achievable in quantum theory using a protocol which is classical apart from quantum subroutines [7]. This quantum subroutine is a black-box implementation of quantum *weak* coin flipping—a similarly defined task but with less stringent security requirements. The history of finding the best quantum weak coin-flipping protocol culminated in the work of Mochon [6]. This unpublished paper is 80 pages long and, even though it has been simplified [48] (see also Ref. [4]), is still not well understood. (Recent progress has been made, however, in Ref. [49].) Mochon's work relies on point games (developed by Kitaev), a notion which is dual, in a sense, to protocols [specified in this work as the pair of triples $((A_0, A_1, A_{\text{abort}}), (B_0, B_1, B_{\text{abort}}))$]. Even though point games are mysterious in the context of quantum theory, perhaps our generalization to the framework of GPTs will shed light. In fact, there is one immediate similarity to this work. A major step in Mochon's proof is the reduction from time-dependent point games to

time-independent point games. This, in a nutshell, strips away all the "time-dependent" information of the protocol. Our framework and proof, on the other hand, completely strips away all notion of time as it does not explicitly rely on the round-to-round strategy descriptions, and thus might make this point game reduction simpler or even trivial. It might even expose a GPT in which perfectly secure weak coin flipping is attainable in a finite number of rounds of communication. It would be interesting to find such a GPT, if one exists, and compare it to quantum theory where it is known that perfectly secure weak coin flipping only exists in the limit of infinite rounds of communication [6,50].

In short, if one were to develop GPT weak coin-flipping protocols with small bias, then the lower bound presented in this work might be achievable by imitating the quantum protocol. It would be interesting to see which GPTs allow for secure weak coin flipping, whether it is proved using point games, semi-infinite programming, or another yet-to-be-discovered method.

ACKNOWLEDGMENTS

We thank M. Plávala, G. Chiribella, and H. Barnum for helpful discussions. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation, and Science. This research was also supported in part by the Foundation for Polish Science through IRAP project cofinanced by EU within Smart Growth Operational Programme (Contract No. 2018/MAB/5).

APPENDIX A: PROOF THAT OUR MAIN RESULT APPLIES TO THE ORIGINAL STRATEGY CONTEXT (BEFORE MODDING OUT BY THE EQUIVALENCE RELATIONS)

Recall that the set of strategies for Alice (resp. Bob) is denoted by $\mathcal{A}$ (resp. $\mathcal{B}$) and we use $\tilde{\mathcal{A}}$ (resp. $\tilde{\mathcal{B}}$) to denote the same set after modding out by the equivalence relation:

$$A_1 \sim A_2 \iff \text{Prob}(A_1, B) = \text{Prob}(A_2, B) \ \text{ for all } B \in \mathcal{B} \tag{A1}$$

(and the analogous equivalence relation for Bob). Recall that we have an inner product such that $\text{Prob}(A, B) = \langle \tilde{A}, \tilde{B} \rangle$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

A protocol is defined by a triple $(A_0, A_1, A_{\text{abort}}) \subset \mathcal{A}$ for Alice and a triple $(B_0, B_1, B_{\text{abort}}) \subset \mathcal{B}$ for Bob satisfying certain properties that are not needed for the following discussion. Let $(\tilde{A}_0, \tilde{A}_1, \tilde{A}_{\text{abort}})$ and $(\tilde{B}_0, \tilde{B}_1, \tilde{B}_{\text{abort}})$ be the triples of equivalence classes for Alice and Bob, respectively.

We have shown in the paper that, under the generalized no-restriction hypothesis on $\tilde{\mathcal{A}}$ or $\tilde{\mathcal{B}}$, we have at least one of the two following conditions holding:

(1) There exists $\tilde{B} \in \tilde{\mathcal{B}}$ such that $\langle \tilde{A}_0, \tilde{B} \rangle \geqslant 1/\sqrt{2}$; or
(2) There exists $\tilde{A} \in \tilde{\mathcal{A}}$ such that $\langle \tilde{B}_0, \tilde{A} \rangle \geqslant 1/\sqrt{2}$.

Suppose the first condition holds. Then take any $B \in \mathcal{B}$ in the same equivalence class as $\tilde{B}$. Then we have

$$\text{Prob}(A_0, B) = \langle \tilde{A}_0, \tilde{B} \rangle \geqslant 1/\sqrt{2}. \tag{A2}$$

A similar argument exists if the second condition holds. Thus, our main result follows in the original context of the strategies as well, namely,

(1) There exists $B \in \mathcal{B}$ such that $\text{Prob}(A_0, B) \geqslant 1/\sqrt{2}$; or
(2) There exists $A \in \mathcal{A}$ such that $\text{Prob}(B_0, A) \geqslant 1/\sqrt{2}$.

APPENDIX B: PROOF THAT ALICE AND BOB'S STRATEGIES ARE BOUNDED IF THEY HAVE NONEMPTY INTERIORS

Suppose that $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$ are the convex sets representing Alice and Bob's strategies, respectively. [Note that we are using the set of strategies after we have modded out by the equivalence relation (A1).] We now show that they are each bounded if they have nonempty interiors. To this end, it suffices to prove this for *any* norm since we are working in a finite-dimensional real vector space. Since $\tilde{\mathcal{B}}$ has nonempty interior, we can define a basis

$$\mathcal{I} = \{\tilde{B}_1, \ldots, \tilde{B}_n\}, \tag{B1}$$

where $\tilde{B}_1, \ldots, \tilde{B}_n \in \tilde{\mathcal{B}}$. Define the function

$$f(X) = \sup_{\tilde{B} \in \mathcal{I}} |\langle X, \tilde{B} \rangle|, \tag{B2}$$

which is clearly non-negative and finite for all X.

Moreover, since $\langle \tilde{A}, \tilde{B} \rangle = \text{Prob}(A, B)$ for all $\tilde{A} \in \tilde{\mathcal{A}}$ and $\tilde{B} \in \tilde{\mathcal{B}}$, we have $f(\tilde{A}) \leqslant 1$ for all $\tilde{A} \in \tilde{\mathcal{A}}$. Thus, all that remains is to prove that f is a valid norm.

We now show the triangle inequality. For any X and Y, we have

$$f(X+Y) = \sup_{\tilde{B} \in \mathcal{I}} |\langle X+Y, \tilde{B} \rangle| \tag{B3}$$

$$\leqslant \sup_{\tilde{B} \in \mathcal{I}} (|\langle X, \tilde{B} \rangle| + |\langle Y, \tilde{B} \rangle|) \tag{B4}$$

$$\leqslant \sup_{\tilde{B} \in \mathcal{I}} |\langle X, \tilde{B} \rangle| + \sup_{\tilde{B}' \in \mathcal{I}} |\langle Y, \tilde{B}' \rangle| \tag{B5}$$

$$= f(X) + f(Y). \tag{B6}$$

Thus, the triangle inequality holds. For any scalar $\alpha \in \mathbb{R}$, we have

$$f(\alpha X) = \sup_{\tilde{B} \in \mathcal{I}} |\langle \alpha X, \tilde{B} \rangle| \tag{B7}$$

$$= \sup_{\tilde{B} \in \mathcal{I}} (|\alpha| |\langle X, \tilde{B} \rangle|) \tag{B8}$$

$$= |\alpha| \sup_{\tilde{B} \in \mathcal{I}} |\langle X, \tilde{B} \rangle| \tag{B9}$$

$$= |\alpha| f(X). \tag{B10}$$

The last property to show is when $f(X) = 0$, we must have $X = 0$. Let X be such that $f(X) = 0$. Then clearly we must have $\langle X, \tilde{B} \rangle = 0$ for all $\tilde{B} \in \mathcal{I}$. Since $\mathcal{I}$ is a basis we have $\langle X, Y \rangle = 0$ for all vectors Y. By setting $Y = X$, we have that $\|X\| = 0$ implying that $X = 0$, as desired.

[1] M. Blum, Coin flipping by telephone a protocol for solving impossible problems, ACM SIGACT News **15**, 23 (1983).

[2] A. Chailloux, G. Gutoski, and J. Sikora, Optimal bounds for semi-honest quantum oblivious transfer, Chicago J. Theor. Comp. Sci. **13**, 1 (2016).

[3] A. Nayak and P. W. Shor, Bit-commitment based quantum coin flipping, Phys. Rev. A **67**, 012304 (2003).

[4] A. Nayak, J. Sikora, and L. Tunçel, Quantum and classical coin-flipping protocols based on bit-commitment and their point games, arXiv:1504.04217.

[5] A. Nayak, J. Sikora, and L. Tunçel, A search for quantum coin-flipping protocols using optimization techniques, Math. Program. **156**, 581 (2016).

[6] C. Mochon, Quantum weak coin flipping with arbitrarily small bias, arXiv:0711.4114.

[7] A. Chailloux and I. Kerenidis, Optimal quantum strong coin flipping, in *Proceedings of 50th IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, 2009), pp. 527–533.

[8] J. Sikora, Simple, near-optimal quantum protocols for die-rolling, Cryptography **1**, 11 (2017).

[9] J. Kilian, Founding cryptography on oblivious transfer, In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, 1988), pp. 20–31.

[10] H.-K. Lo and H. F. Chau, Why quantum bit commitment and ideal quantum coin tossing are impossible, Phys. D (Amsterdam, Neth.) **120**, 177 (1998).

[11] A. Kitaev, Quantum coin-flipping, talk at the 6th annual workshop on Quantum Information Processing, 2002 (unpublished).

[12] L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012.

[13] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

[14] G. Ludwig, *An Axiomatic Basis of Quantum Mechanics. 1. Derivation of Hilbert Space* (Springer-Verlag, Berlin, 1985).

[15] E. Brian Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. **17**, 239 (1970).

[16] C. H. Randall and D. J. Foulis, An approach to empirical logic, Am. Math. Mon. **77**, 363 (1970).

[17] C. Piron, Axiomatique quantique, Helv. Phys. Acta **37**, 439 (1964).

[18] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, New York, 1963).

[19] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Probabilistic theories with purification, Phys. Rev. A **81**, 062348 (2010).

[20] L. Hardy, Reformulating and reconstructing quantum theory, arXiv:1104.2066.

[21] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, A structure theorem for generalized-noncontextual ontological models, arXiv:2005.07161.

[22] J. Sikora and J. Selby, Simple proof of the impossibility of bit commitment in generalized probabilistic theories using cone programming, Phys. Rev. A **97**, 042302 (2018).

[23] J. H. Selby and J. Sikora, How to make unforgeable money in generalised probabilistic theories, Quantum **2**, 103 (2018).

[24] L. Lami, C. Palazuelos, and A. Winter, Ultimate data hiding in quantum mechanics and beyond, Commun. Math. Phys. **361**, 661 (2018).

[25] H. Barnum and A. Wilce, Information processing in convex operational theories, Elect. Notes Theor. Comp. Sci. **270**, 3 (2011).

[26] H. Barnum, Oscar C. O. Dahlsten, M. Leifer, and B. Toner, Nonclassicality without entanglement enables bit commitment, in *Proceedings of the IEEE Information Theory Workshop* (IEEE, Piscataway, NJ, 2008), pp. 386–390.

[27] J. Barrett, L. Hardy, and A. Kent, No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005).

[28] M. Krumm and M. P. Müeller, Quantum computation is the unique reversible circuit model for which bits are balls, npj Quantum Inf. **5**, 7 (2019).

[29] H. Barnum, C. M. Lee, and J. H. Selby, Oracles and query lower bounds in generalised probabilistic theories, Found. Phys. **48**, 954 (2018).

[30] A. J. P. Garner, Interferometric computation beyond quantum theory, Found. Phys. **48**, 886 (2018).

[31] J. Barrett, N. de Beaudrap, M. J. Hoban, and C. M. Lee, The computational landscape of general physical theories, npj Quantum Inf. **5**, 41 (2019).

[32] C. M. Lee and J. H. Selby, Deriving Grover's lower bound from simple physical principles, New J. Phys. **18**, 093047 (2016).

[33] C. M. Lee and M. J. Hoban, Bounds on the power of proofs and advice in general physical theories, Proc. R. Soc. London A **472**, 20160076 (2016).

[34] C. M. Lee and J. H. Selby, Generalised phase kick-back: The structure of computational algorithms from physical principles, New J. Phys. **18**, 033023 (2016).

[35] C. M. Lee and J. Barrett, Computation in generalised probabilisitic theories, New J. Phys. **17**, 083001 (2015).

[36] C. M. Lee and J. H. Selby, Higher-order interference in extensions of quantum theory, Found. Phys. **47**, 89 (2017).

[37] F. Riesz, Sur les opérations fonctionnelles linéaires, Compt. Rend. Acad. Sci. **149**, 974 (1909).

[38] A. Shapiro, Semi-infinite programming, duality, discretization and optimality conditions, Optimization **58**, 133 (2009).

[39] S. Fiorini, S. Massar, M. K. Patra, and H. R. Tiwary, Generalized probabilistic theories and conic extensions of polytopes, J. Phys. A: Math. Theor. **48**, 025302 (2014).

[40] A. Jenčová and M. Plávala, Conditions on the existence of maximally incompatible two-outcome measurements in general probabilistic theory, Phys. Rev. A **96**, 022113 (2017).

[41] J. Bae, D.-G. Kim, and L.-C. Kwek, Structure of optimal state discrimination in generalized probabilistic theories, Entropy **18**, 39 (2016).

[42] S. Gharibian, J. Sikora, and S. Upadhyay, QMA variants with polynomially many provers, Quantum Inf. Comput. **13**, 0135 (2013).

[43] S. Bandyopadhyay, A. Cosentino, N. Johnston, V. Russo, J. Watrous, and N. Yu, Limitations on separable measurements by convex optimization, IEEE Trans. Inf. Theory **61**, 3593 (2015).

[44] M. Laurent and T. Piovesan, Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone, Siam J. Optim. **25**, 2461 (2015).

[45] J. Sikora and A. Varvitsiotis, Linear conic formulations for two-party correlations and values of nonlocal games, Math. Program. **162**, 431 (2017).

[46] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2004).

[47] G. Gutoski and J. Watrous, Toward a general theory of quantum games, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, 2007), pp. 565–574.

[48] D. Aharonov, A. Chailloux, M. Ganz, I. Kerenidis, and L. Magnin, A simpler proof of existence of quantum weak coin flipping with arbitrarily small bias, SIAM J. Comp. **45**, 633 (2016).

[49] A. S. Arora, J. Roland, and S. Weis, Quantum weak coin flipping, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (ACM, New York, NY, 2019), pp. 205–216.

[50] A. Ambainis, A new protocol and lower bounds for quantum coin flipping, J. Comp. Syst. Sci. **68**, 398 (2004).