A time-dependent Tsirelson's bound from limits on the rate of information gain in quantum systems

You may also be interested in:

Optimal signal states for quantum detectors
Ognyan Oreshkov, John Calsamiglia, Ramon Muñoz-Tapia et al.

Quantum state discrimination and its applications
Joonwoo Bae and Leong-Chuan Kwek

Entropy in general physical theories
Anthony J Short and Stephanie Wehner

A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations
Miguel Navascués, Stefano Pironio and Antonio Acín

Device-independent quantum key distribution secure against collective attacks
Stefano Pironio, Antonio Acín, Nicolas Brunner et al.

Rigidity of quantum steering and one-sided device-independent verifiable quantum computation
Alexandru Gheorghiu, Petros Wallden and Elham Kashefi

Non-adaptive measurement-based quantum computation and multi-party Bell inequalities
Matty J Hoban, Earl T Campbell, Klearchos Loukopoulos et al.

Simplified instantaneous non-local quantum computation with applications to position-based cryptography
Salman Beigi and Robert König

A monogamy-of-entanglement game with applications to device-independent quantum cryptography
Marco Tomamichel, Serge Fehr, Jdrzej Kaniewski et al.

A time-dependent Tsirelson's bound from limits on the rate of information gain in quantum systems

Andrew C Doherty[1] **and Stephanie Wehner**[2,3]

[1] Centre for Engineered Quantum Systems, School of Physics,
University of Sydney, Sydney, Australia
[2] Centre for Quantum Technologies, National University of Singapore,
2 Science Drive 3, 117543 Singapore, Singapore
E-mail: wehner@comp.nus.edu.sg

Abstract. We consider the problem of distinguishing between a set of arbitrary quantum states in a setting in which the time available to perform the measurement is limited. We provide simple upper bounds on how well we can perform state discrimination in a given time as a function of either the average energy or the range of energies available during the measurement. We exhibit a specific strategy that nearly attains this bound. Finally, we consider several applications of our result. Firstly, we obtain a time-dependent Tsirelson's bound that limits the extent of the Bell inequality violation that can be in principle be demonstrated in a given time t. Secondly, we obtain a Margolus–Levitin type bound when considering the special case of distinguishing orthogonal pure states.

[3] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Entropic measures tell us how much information a quantum register E contains about some classical register X *in principle*. But just how quickly does this information become available to us? In this little note, we derive bounds on the amount of information available after a given time t. As expected, our bounds depend on the resources we have available in the form of the *available energy*.

Throughout this paper, we will choose to measure information in terms of the *min-entropy*, which is the relevant quantity when we consider single-shot experiments and quantum cryptography. As we will explain in detail below, this measure is directly related [1] to the probability of success in state discrimination [2–6]. As a result, we focus on bounding the probability of success in distinguishing states $\{\rho_x\}_{x\in\mathcal{X}}$ where we are given ρ_x with probability p_x. Let $P_{\text{guess}}(X|E)_{H,t}$ denote this success probability after time t when using a particular Hamiltonian H in the measurement process. After providing a more careful discussion of the measurement process, we show the following results.

1.1. Results

1.1.1. A bound for two states. We first consider the case of only two input states ρ_0, ρ_1, for which it is easy to compute the optimal success probability if we have unlimited time (or resources) available [2]. We first provide a general bound in terms of the spectrum of the Hamiltonian (corollary B.2). For the special case of of two equiprobable states ($p_0 = p_1 = 1/2$), this bound simply reads

$$P_{\text{guess}}(X|E)_{H,t} \leqslant \frac{1}{2} + \frac{\gamma t \|H\|_\infty D(\rho_0, \rho_1)}{2\hbar}, \tag{1}$$

where $D(\rho_0, \rho_1)$ is the trace distance between the two states, $\|H\|_\infty$ is the largest eigenvalue of the Hamiltonian, and γ is a small constant. This bound is directly related to our ability to distinguish two input states given an unlimited amount of time, where the best measurements give us [2]

$$P_{\text{guess}}(X|E) = \frac{1}{2} + \frac{D(\rho_0, \rho_1)}{2}. \tag{2}$$

It is important to note that our time bounds depend on the energy *available* for the state discrimination process, and not how much energy could potentially be lost in any such process. We proceed to show that our bound is nearly tight up to a constant factor (theorem D.1) by providing an explicit measurement strategy. Finally, we prove a bound in terms of the average energies of the input states (theorem B.3). However, this bound does not compare as easily to the case of unlimited time.

1.1.2. A bound for many input states. When considering the case of an arbtirary number of input states $\rho_0, \ldots, \rho_{N-1}$ it is difficult to compute the maximum success probability, even in the case of unlimited time. In particular, no general closed form expression is known—only for the case of single qubit encodings does there exist a way to construct the optimal measurements geometrically [7]. In general, we can only approximate the optimal measurements numerically [5, 8–14], or resort to bounds on the success probability [4, 15–20]. As such, it becomes harder to relate $P_{\text{guess}}(X|E)_{H,t}$ to case of unlimited time. We hence provide a general bound in terms of the average energies alone. In particular, we show (theorem C.1) that

$$P_{\text{guess}}(X|E) \leqslant p_{x_{\max}} + \sum_{x=0}^{N-1} p_x \operatorname{tr}(H\rho_x), \tag{3}$$

where $x_{\max}$ is the smallest $x \in \{0, \ldots, N-1\}$ such that $p_{x_{\max}} \geqslant p_x$ for all x.

1.1.3. Applications. Finally, we discuss two applications of our bound. The first is to the study of Bell inequalities [21]. Typically, we care about determining the maximum quantum violation of such inequalities. In contrast, we ask what is the maximum violation that can be achieved in a fixed amount of time. When considering such inequalities as games between two players, Alice and Bob (see section 4.2.1), the 'amount' of quantum violation is determined by the probability p_{win} that the players win the game maximized over all states and measurements. For the Clauser–Horn–Shimony–Holt (CHSH) inequality [22], we have that classically

$$p_{\text{win}} \leqslant \tfrac{3}{4} \tag{4}$$

for any strategy of Alice and Bob. However, in quantum mechanics there exists a strategy that achieves

$$p_{\text{win}} = \frac{1}{2} + \frac{1}{2\sqrt{2}}. \tag{5}$$

This is optimal in the sense that there is no quantum strategy that achieves a higher winning probability. This upper bound on the winning probability is known as *Tsirelson's bound* [23] for the CHSH inequality. Here, we show (corollary F.2) that if we demand answers from Alice and Bob after time t

$$p_{\text{win}} \leqslant \frac{3}{4} + \frac{\gamma t \|H\|_{\infty}}{\hbar\sqrt{2}}, \tag{6}$$

where H is Bob's Hamiltonian involved in the measurement process, and γ is a small constant. We will also see that to achieve Tsirelson's bound, Alice and Bob need at least time

$$t \geqslant \frac{\hbar}{\gamma \|H\|_{\infty}}. \tag{7}$$

Our bounds tell us that there is indeed a fundamental time that is needed to establish non-local correlations of a certain strength. We will discuss these bounds in detail in section 4.2.

As a second application, we use our bound to obtain a form of the Margolus–Levitin theorem [24] that provides us with a lower bound on how much time it takes to transform a pure state into an orthogonal state. Since the Margolus–Levitin theorem provides a bound on the speed of evolution, it clearly provides a bound on the minimum amount of time that is required to obtain the optimal (time unlimited) success probability for state discrimination. However, note that we are interested in bounding $P_{\text{guess}}(X|E)_{H,t}$ even for shorter periods of time. We will discuss the relation of our work and the Margolus–Levitin theorem in detail in section 4.1.

1.2. Related work

Next to the Margolus–Levitin theorem [24], our work is related to several bounds [25, 26] on how fast information can be transmitted in principle given energy constraints (see [27] for a survey of results). These bounds generally consider the von Neumann entropy as a measure of information and are concerned with determining the capacity for sending information as a function of energy. That is, they consider how quickly we could convey information in the best possible way. In contrast, we consider the case of arbitrary encodings ρ_x, which may not be optimal to transmit classical information. In fact, even in the case of unlimited time the probability that we can reconstruct x from ρ_x could be very small. Our setting also differs in the sense that we focus solely on extracting classical information into a classical register in a sense that we will make precise below.

Our work is also related to several previous papers [28–30] that study the rate of change in entropies of a system that is in contact with an environment. Again, our work is of a somewhat different flavor since we are interested in extracting classical information, and our bounds furthermore involve average energies, rather than the largest energy $\|H\|_{\infty}$ of the (interaction) Hamiltonian H alone.

2. Gaining classical information

2.1. Quantifying information

Let us now consider more formally what we mean by gaining classical information encoded in a quantum system. Imagine that there is some finite set $\mathcal{X}$ of possible classical symbols to be encoded. For any symbol $x \in \mathcal{X}$, we thereby use $\rho_x \in \mathcal{B}(\mathcal{H}_{\mathrm{enc}})$ to denote its encoding into a quantum state on the system $\mathcal{H}_{\mathrm{enc}}$. We also refer to $\mathcal{H}_{\mathrm{enc}}$ as the *encoding space*. Our *a priori* ignorance about the classical information x is captured by the probability distribution p_x according to which the encoding space is prepared in the state ρ_x.

Throughout, we quantify how much information we have about x given access to the encoding space $\mathcal{H}_{\mathrm{enc}}$ in terms of the min-entropy [1]

$$\mathrm{H}_\infty(X|E) := -\log P_{\mathrm{guess}}(X|E), \tag{8}$$

where

$$P_{\mathrm{guess}}(X|E) := \sup_{\substack{\forall x\, M_x \geqslant 0 \\ \sum_{x \in \mathcal{X}} M_x = \mathbb{I}}} \sum_{x \in \mathcal{X}} p_x \operatorname{tr}(M_x \rho_x) \tag{9}$$

is the probability that we guess x, maximized over all possible measurements on the encoding space. Finding the optimal measurement is known as state discrimination and can be done using semidefinite programming [5, 8]. The min-entropy accurately measures information in a cryptographic setting [31], and for single shot experiments. This is in contrast to the von Neumann entropy which is concerned with the asymptotic case of a large number of identical experiments.

The min-entropy and the von Neumann entropy can be arbitrarily different, as is easily seen by considering the example where the encoding is trivial, that is, $\rho_x = \rho_{x'}$ for all x and x'. The strategy that maximizes the guessing probability $P_{\mathrm{guess}}(X|E)$ is then simply given by outputting the most likely symbol, i.e. $\mathrm{H}_\infty(X|E) = -\log \max_x p_x$,[4] and the conditional von Neumann entropy obeys $\mathrm{H}(X|E) = \mathrm{H}(X) = -\sum_x p_x \log p_x$. Consider now $\Sigma = \{0, 1\}^n$ to be the set of bitstrings of length n and suppose the all '0' string occurs with probability $p_{0^n} = 1/2$, and with probability $1/2$ any of the remaining strings occurs with equal probability. Clearly, we have $\mathrm{H}_\infty(X|E) = 1$, whereas $\mathrm{H}(X) \approx n/2$. That is, the von Neumann entropy can be very large, even if there is one symbol that occurs with extremely high probability. We will remark on the rate of information extraction from a quantum system in terms of the von Neumann entropy later on, but focus on the single shot case given by the min-entropy, or equivalently the probability of error in state discrimination.

2.2. Producing a classical output

To determine how quickly we can acquire classical information, we first need to specify what it means to output classical information from a measurement. Here, we model this process with the help of an additional 'classical' ancilla system $\mathcal{H}_{\mathrm{anc}}$ that contains the output. A classical system is associated with a fixed basis, which without loss of generality we take to be the computational basis. Preparation and measurement of a classical system can only be done on this basis, which intuitively corresponds to the idea of storing classical information: the ancilla can be prepared

---

[4] In analogy to the von Neumann entropy, $\mathrm{H}_\infty(X|E) = \mathrm{H}_\infty(X)$ since $\mathrm{H}_\infty(X) := -\log \max_x p_x$.

in any state of the fixed basis, and is subsequently measured on this basis after time t. The information contained in this register captures the notion of a classical probability distribution over the basis elements.

We model the process of state discrimination as follows. The problem is to discriminate between N states ρ_x on the encoding space $\mathcal{H}_{enc}$, where N is the number of possible classical symbols. At the beginning of the experiment the ancilla system is initialized to the symbol occurring with the largest probably $|x_{max}\rangle$, where

$$x_{max} := \text{argmax}_x \, p_x. \tag{10}$$

This initial condition captures the distinguisher's *a priori* knowledge: recall without access to the quantum register $H_\infty(X|E) = -\log p_{x_{max}}$. If there are multiple classical symbols with the same value $p_{x_{max}}$, we take the smallest one in lexicographic order. We will discuss the choice of initial state in detail below. The ancilla system has total dimension $d_{\mathcal{H}_{anc}} = N$ and the other directions correspond to the classical symbols x. The experimenter implements a unitary U on $\mathcal{H}_{enc} \otimes \mathcal{H}_{anc}$ during a specified time t. At this point the ancilla system is passed to a referee who will decide whether information has been gathered successfully by measuring $\mathcal{H}_{anc}$ in the computational basis, using measurement operators

$$P_x := |x\rangle\langle x|, \tag{11}$$

where the subscript x denotes the corresponding classical output. Hence the success probability of correctly identifying the state ρ_x using this procedure when the ancilla was initially in the state $|x_{max}\rangle$ is given by

$$\text{tr}((\mathbb{I}_{\mathcal{H}_{enc}} \otimes P_x) U (\rho_x \otimes |x_{max}\rangle\langle x_{max}||) U^\dagger). \tag{12}$$

See figure 1 for a schematic depiction of this process. Note that the ancilla is measured by the referee at no time cost. This is a natural assumption in our setting where we imagine that the final information is extracted by a referee who is not limited by any energy constraints. Such a referee naturally arises in, for example, the setting of Bell inequalities which we consider later. We will from now on assume that measurements producing classical outcomes are always performed this way.

To bound how much min-entropy we have after time t, our goal is to place bounds on the success probability in terms of the unitary

$$U = \exp\left(-\frac{iHt}{\hbar}\right), \tag{13}$$

that is, in terms of the interaction Hamiltonian

$$H = \sum_n E_n |E_n\rangle\langle E_n| \tag{14}$$

and the time t. Throughout, we will assume that $H \geqslant 0$ and that the lowest energy level is in fact $E_0 = 0$. Any other Hamiltonian differs from such an H by a term proportional to the identity, which does not contribute to the speed of information gain. We explicitly chose not to use the common convention $\hbar = 1$ to make it easier to draw comparisons to the Margolus–Levitin theorem [24] later.

Before turning to our actual bounds, let us first introduce some additional notation which we will refer to throughout the paper. We will use

$$\tilde{\rho}_x := \rho_x \otimes |x_{max}\rangle\langle x_{max}|, \tag{15}$$
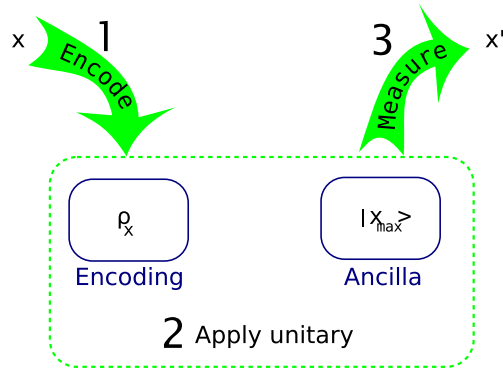
Figure 1. Our protocol for distinguishing quantum states in finite time. Firstly, the encoding register is placed into an encoding ρ_x of the classical symbol x chosen with probability p_x. The ancilla is initialized in the state $|x_{\max}\rangle$. Secondly, we can perform a unitary interaction $U = \exp(-iHt/\hbar)$ for time t between the encoding and the ancilla register. Note that the energy does not change in this step, but that the speed of this interaction depends on the energy already available. Finally, the ancilla register is measured by the referee in the computational basis to determine a guess x' for x. If $x' = x$, then we successfully recovered the classical information. In the setting of Bell inequalities considered later, the ancilla register is simply the message returned to the referee.

to denote the combined state consisting of the input state ρ_x on the encoding space, and the initial state of the ancilla $|x_{\max}\rangle\langle x_{\max}|$. We also write

$$R := U - I = \sum_n (\exp(-iE_n t/\hbar) - 1)|E_n\rangle\langle E_n|. \tag{16}$$

Furthermore, it will be convenient to rewrite the success probability (12) in terms of measurement operators

$$M_x := U^\dagger(\mathbb{I} \otimes P_x)U = \mathbb{I} \otimes P_x + W_x, \tag{17}$$

as $\mathrm{tr}(M_x\tilde{\rho}_x)$, where

$$W_x = W_x^1 + W_x^2 \tag{18}$$

and

$$W_x^1 := (\mathbb{I} \otimes P_x)R + R^\dagger(\mathbb{I} \otimes P_x), \tag{19}$$

$$W_x^2 := R^\dagger(\mathbb{I} \otimes P_x)R. \tag{20}$$

The average success probability for a particular Hamiltonian H and time t can now be written as

$$P_{\mathrm{guess}}(X|E)_{H,t} := \sum_{x \in \mathcal{X}} p_x \, \mathrm{tr}\,(M_x\tilde{\rho}_x). \tag{21}$$

3. Time versus information gain

We are now ready to derive our bounds. For simplicity, we will outline how this can be done for the case of two equiprobable states, and merely state our general result. Precise statements as well as a detailed derivation can be found in the appendix.

3.1. An upper bound to $P_{\text{guess}}(X|E)$

We now first derive an upper bound to the guessing probability. For the case of two equiprobable states (i.e. $N = 2$ and $p_x = 1/2$ for all $x \in \mathcal{X}$), such bounds are easy to obtain when we allow unlimited time (or energy). In particular, it is well known that in this case the success probability is given by [2]

$$P_{\text{guess}}(X|E) := \frac{1}{2} + \frac{D(\rho_0, \rho_1)}{2}, \tag{22}$$

where $D(\rho_0, \rho_1) = \frac{1}{2}\|\rho_0 - \rho_1\|_1$ is the trace distance of the two states. Let us now consider what happens in our time-limited scenario for a particular interaction Hamiltonian H. First of all, recall that for two equiprobable states, the ancilla is initialized to the smallest value

$$|x_{\text{max}}\rangle = |0\rangle. \tag{23}$$

For two states, the success probability P_{succ} averaged over the choice of input state, using the measurement given by operators M_1 and $M_0 = \mathbb{I} - M_1$ from (17), can now be expressed as

$$P_{\text{succ}}(X|E)_{H,t} = \frac{1}{2}\left[\text{tr}\,(M_0\tilde{\rho}_0) + \text{tr}\,(M_1\tilde{\rho}_1)\right] \tag{24}$$

$$= \frac{1}{2}\left[1 + \text{tr}\,(M_1(\tilde{\rho}_1 - \tilde{\rho}_0))\right] \tag{25}$$

$$= \frac{1}{2}\left[1 + \text{tr}\,(\rho_1 - \rho_0)\,\text{tr}\,(P_1|x_{\text{max}}\rangle\langle x_{\text{max}}|) + \text{tr}\,(W_1(\tilde{\rho}_1 - \tilde{\rho}_0))\right] \tag{26}$$

$$= \frac{1}{2} + \frac{\text{tr}\,(W_1(\tilde{\rho}_1 - \tilde{\rho}_0))}{2}, \tag{27}$$

where the fourth equality follows immediately from the fact that $P_1|x_{\text{max}}\rangle\langle x_{\text{max}}| = |x_{\text{max}}\rangle\langle x_{\text{max}}|P_1 = 0$. Let us now upper bound the term involving W_1. Again using that $P_1|x_{\text{max}}\rangle\langle x_{\text{max}}| = |x_{\text{max}}\rangle\langle x_{\text{max}}|P_1 = 0$, we have

$$\text{tr}\left(W_1^1(\tilde{\rho}_1 - \tilde{\rho}_0)\right) = 0. \tag{28}$$

Define $\tilde{A} := \tilde{\rho}_1 - \tilde{\rho}_0$, and consider its diagonalization $\tilde{A} = \sum_j \lambda_j |u_j\rangle\langle u_j|$. Let $\tilde{A}^+ := \sum_{j,\lambda_j \geqslant 0} \lambda_j |u_j\rangle\langle u_j|$ and $\tilde{A}^- := \tilde{A} - \tilde{A}^+$. Using the fact that $R \cdot R^\dagger$ is a positive map [32] and $0 \leqslant \mathbb{I} \otimes P_x \leqslant \mathbb{I}$, we can now bound the term

$$\text{tr}(W_1^2\tilde{A}) \leqslant \text{tr}(R\tilde{A}^+R^\dagger) \tag{29}$$

$$\leqslant 2\sum_n (1 - \cos(tE_n/\hbar))\langle E_n|\tilde{A}^+|E_n\rangle. \tag{30}$$

Substituting back into our original bound (27) gives us

$$P_{\text{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + \sum_n (1 - \cos(tE_n/\hbar))\langle E_n|\tilde{A}^+|E_n\rangle. \tag{31}$$

This is the basic inequality that we can use, along with some restriction on the allowed energies E_n, to bound the success probability for state discrimination in time t. In the rest of the paper we will apply this in two main settings, bounded maximum energy and bounded average energy.

3.1.1. A bound in terms of the maximum energy. From (31), we can immediately obtain a bound on the success probability for state discrimination in terms of the maximum energy $\|H\|_\infty$ of the coupling Hamiltonian H. ($\|H\|_\infty$ is just the largest eigenvalue of H.) This bound is attractive since it is simple to derive and has the appealing feature that it involves the trace distance between the two states, and is thus directly related to the probability that we distinguish the two states given an unlimited amount of time. However, there are many systems of physical interest where the maximum energy of system states is effectively unbounded. Even though we may without loss of generality assume that the spectrum is bounded for a particular set of input states (see appendix), this bound is nevertheless quite unsatisfying in these situations since it can be very weak. For this reason, we use the fundamental inequality (31) in the next section to derive a bound on the success probability that depends only on the average energy.

Note that since $\mathrm{tr}(\tilde{A}^+) = D(\tilde{\rho}_0, \tilde{\rho}_1) = D(\rho_0, \rho_1)$[5] we immediately obtain that the success probability obeys

$$P_{\mathrm{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + (1 - \cos(tC_{\max}/\hbar))D(\rho_0, \rho_1), \tag{32}$$

where $C_{\max} = \mathrm{argmax}_{E_n}(1 - \cos(tE_n/\hbar))$.[6] If $tE_n/\hbar \leqslant \pi$ for all n, then this upper bound simply reads

$$P_{\mathrm{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + (1 - \cos(t\|H\|_\infty/\hbar))D(\rho_0, \rho_1), \tag{33}$$

which will be useful for comparison below. However, for any values of $tE_n/\hbar$, there is a way upper bound (32) in terms of $\|H\|_1$. In particular, it is easy to verify that

$$P_{\mathrm{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + \frac{\gamma t\|H\|_\infty D(\rho_0, \rho_1)}{2\hbar}, \tag{34}$$

where

$$\gamma := \begin{cases} 5/\pi, & \text{if } 1 < tC_{\max}/\hbar < 4, \\ 3/\pi, & \text{otherwise.} \end{cases} \tag{35}$$

3.1.2. A bound in terms of the average energy. A sometimes more satisfying bound can be obtained in terms of the average energy. Note that we can upper bound (31) as

$$\frac{1}{2} + \frac{1}{2}\frac{\gamma t}{\hbar} \sum_n E_n \langle E_n|\tilde{A}^+|E_n\rangle, \tag{36}$$

---

[5] To see that this is an equality, recall that $D(\tilde{\rho}_0, \tilde{\rho}_1) := \frac{1}{2}\|A\|_1$. By the definition of the L1-norm, this means $D(\tilde{\rho}_0, \tilde{\rho}_1) = \frac{1}{2}\max_{-\mathbb{I} \leqslant P \leqslant \mathbb{I}} \mathrm{tr}(PA)$. The maximum is attained for $P = \Pi_{A^+} - \Pi_{A^-}$ where Π_{A^+} and Π_{A^-} are projectors onto the support of A^+ and A^-, respectively. This gives $D(\tilde{\rho}_0, \tilde{\rho}_0) = \frac{1}{2}\left(\mathrm{tr}(A^+) + \mathrm{tr}(A^-)\right)$. The claim now follows by noting that $\mathrm{tr}(A^+) = \mathrm{tr}(A^-)$ since $\mathrm{tr}(\rho_0) = \mathrm{tr}(\rho_1) = 1$.
[6] Recall that argmax refers to the element attaining the maximum value, and not the maximum itself, where here we follow convention and take it to be the lowest if multiple E_n give the same value.

and hence we may use the fact that

$$\tilde{A}^+ = \tfrac{1}{2}(\tilde{A}^+ - \tilde{A}^-) + \tfrac{1}{2}(\tilde{A}^+ + \tilde{A}^-) \tag{37}$$

$$= \tfrac{1}{2}(\tilde{\rho}_1 - \tilde{\rho}_0) + \tfrac{1}{2}|\tilde{\rho}_1 - \tilde{\rho}_0|, \tag{38}$$

to obtain

$$P_{\text{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + \frac{\gamma t \, (\text{tr}(H|\tilde{\rho}_1 - \tilde{\rho}_0|) + \text{tr}(H\tilde{\rho}_1) - \text{tr}(H\tilde{\rho}_0))}{4\hbar}. \tag{39}$$

Now, the asymmetry between the labels 0 and 1 is inessential. The bound is true if we swap the two state labels, as may be seen by repeating the above derivation swapping the role of the two state labels. Averaging these two bounds we find the following symmetric bound:

$$P_{\text{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + \frac{\gamma t \, \text{tr}(H|\tilde{\rho}_1 - \tilde{\rho}_0|)}{4\hbar}. \tag{40}$$

This bound should be compared with the bound (33) in terms of the maximum energy in which the trace distance appears. The quantity on the right hand side of (40) is loosely an energy-weighted trace distance. Whereas this bound is certainly stronger for a particular choice of H, it no longer bears an obvious quantitative relation to the Helstrom bound in terms of the trace distance. In deriving (40) we have made use of the knowledge of the optimal measurements for distinguishing a pair of states. This is no longer possible in more complicated cases, even where unlimited time is allowed [6]. We can weaken the bound somewhat, using the fact that $\rho, H \geqslant 0$ to obtain a bound explicitly in terms of the average energy as follows:

$$P_{\text{succ}}(X|E)_{H,t} \leqslant \frac{1}{2} + \frac{\gamma t \, \text{tr}[H(\tilde{\rho}_0 + \tilde{\rho}_0)]}{4\hbar}. \tag{41}$$

So we see that the average energy of the joint system and ancilla place a bound on the success probability of state discrimination, as claimed. This bound may be generalized easily to the case of more than two classical symbols and an arbitrary distribution $\{p_x\}_x$. We show in the appendix that

Theorem 3.1. *Suppose $H \geqslant 0$. Then the probability of distinguishing $\rho_0, \ldots, \rho_{N-1}$ given with probabilities $p_0, \ldots, p_{N-1}$ using the Hamiltonian H obeys*

$$P_{\text{succ}}(X|E)_{H,t} \leqslant p_{x_{\max}} + \frac{\hat{\gamma} t}{\hbar} \sum_{x=0}^{N-1} p_x \, \text{tr}(H\tilde{\rho}_x), \tag{42}$$

where

$$\hat{\gamma} := \begin{cases} 5/\pi, & \text{if } \forall E_n, 1 < t E_n/\hbar < 4, \\ 3/\pi, & \text{otherwise.} \end{cases} \tag{43}$$

Note that the term $\sum_x p_x \, \text{tr}(H\tilde{\rho}_x)$ is the energy of the encoding and ancilla register averaged over the choice of input symbols.

3.2. A lower bound on $P_{\text{guess}}(X|E)$

We now exhibit a specific measurement strategy for two equiprobable states, which attains our upper bound up to a constant factor. We again focus on the case of two possible input states,

as for the general setting there is no analytic procedure of obtaining the optimal measurements even in the setting of unlimited time. Our construction for two states will make explicit use of this optimal measurement.

Let $A = \rho_1 - \rho_0$. It is well known [2] that the optimal distinguishing measurement in the time unlimited case without the use of an ancilla is given by $\{\Pi_{A^+}, \Pi_{A^-}\}$, where Π_{A^+} and Π_{A^-} are projectors on the positive and negative eigenspace of A respectively. To construct our Hamiltonian H, let us diagonalize $A = \sum_j \lambda_j |u_j\rangle\langle u_j|$, and define $A^+ := \sum_{j, \lambda_j \geqslant 0} \lambda_j |u_j\rangle\langle u_j|$ and $A^- := A^+ - A$. Consider the operator

$$\hat{H} := \Pi_{A^-} \otimes \mathbb{I} + \Pi_{A^+} \otimes (|0\rangle\langle 1| + |1\rangle\langle 0|). \tag{44}$$

Clearly, $\hat{H}$ is Hermitian and unitary, and hence has eigenvalues ± 1. In fact, $\hat{H}$ is the unitary we would use to achieve the optimum distinguishing probability if we were unconcerned with time. We now define a Hamiltonian H

$$H := E_{\max}(\hat{H} + \mathbb{I})/2. \tag{45}$$

For comparison with our upper bound of (33) H obeys the condition $H \geqslant 0$ and has a largest eigenvalue equal to $E_{\max} = \|H\|_\infty$. A simple calculation provided in the appendix shows that for our choice of H we have

$$P_{\text{succ}}(X|E)_{H,t} = \frac{1}{2} + \frac{1}{4}(1 - \cos(t\|H\|_\infty/\hbar))D(\rho_0, \rho_1), \tag{46}$$

which gives a lower bound to $P_{\text{succ}}(X|E)$ maximized over all possible H in time t. This bound matches the upper bound of (33) up to a factor of $1/4$.

Note that $\hat{H}$ effectively implements a variant of the controlled-NOT (c-NOT) operation on the encoding space and the ancilla. For more than two input states, one could construct a similar $\hat{H}$ implementing a controlled addition mod N on the ancilla, as long as the optimum distinguishing measurement in the case of unlimited time is a projective measurement on the encoding space. This would give a similar relation between time and the original probability of distinguishing the given states. However, it is known that there do exist choices of encodings ρ_x such that the optimum measurement is not projective, and hence we omit this restricted form of generalization.

4. Applications

Let us now consider several applications of our simple bound.

4.1. Minimum distinguishing time and the Margolus–Levitin theorem

The first application we are interested in, is a return to our initial question: just how quickly can we acquire information? That is, what is the minimum time needed to extract classical information encoded in a quantum system? Note that with the Hamiltonian H in the lower bound for two equiprobable states, there does indeed exist a way to optimally distinguish the two states in time $t = \hbar\pi/\|H\|_\infty$. However, since there is a small gap to our upper bound it would be an open question as to whether it is possible to achieve the same in an even shorter amount of time.

4.1.1. Minimum time. However, note that our upper bounds on $P_{\text{succ}}(X|E)_{H,t}$ can also be understood as lower bounds on the time required to optimally distinguish the given states, retrieving the maximum amount of information from the encoding. Let us first consider our most general bound for large $\mathcal{X}$. We have that if we can distinguish optimally in time $t_{\text{distinguish}}$ our upper bound must be at least as large as the optimum $P_{\text{guess}}(X|E)$. That is,

$$p_{x_{\max}} + \frac{\gamma\, t_{\text{distinguish}}}{\hbar} \sum_{x=0}^{N-1} p_x \operatorname{tr}(H\tilde{\rho}_x) \geqslant P_{\text{guess}}(X|E), \tag{47}$$

and hence

$$t_{\text{distinguish}} \geqslant \frac{(P_{\text{guess}}(X|E) - p_{x_{\max}})\hbar}{\gamma \sum_{x=0}^{N-1} p_x \operatorname{tr}(H\tilde{\rho}_x)}. \tag{48}$$

4.1.2. Margolus–Levitin theorem. Let us now consider the special case where two equi-probable input encodings are perfectly distinguishable. That is, $\rho_0 = |0\rangle\langle 0|$ and $\rho_1 = |1\rangle\langle 1|$. Our task is now quite simple: we merely wish to turn the state $|1\rangle|0\rangle$ of the encoding and ancilla system to the state $|1\rangle|1\rangle$; that is, we wish to transform one vector into its orthogonal. Note that given unlimited time (or energy) we can succeed perfectly at this task and hence $P_{\text{guess}}(X|E) = 1$. From (48) we thus have

$$t_{\text{distinguish}} \geqslant \frac{\hbar}{2\gamma \operatorname{tr}(H\tilde{\rho}_1)}. \tag{49}$$

Our bound can hence also be understood as putting a limit on the time that it takes to turn a state vector to its orthogonal (on the ancilla), given some additional resource (the encoding register).

A bound on the minimum time that it takes to turn a vector into its orthogonal is indeed known as the Margolus–Levitin theorem [24]. In particular, their bound applied to our situation involving both the encoding and the ancilla register gives

$$t_{\text{ML}} \geqslant \frac{\hbar\pi}{2 \operatorname{tr}(H\tilde{\rho}_1)}. \tag{50}$$

Such a bound had previously only been derived from the time-energy uncertainty principle, where instead of the average energy we have the energy spread, i.e the difference in the largest and smallest eigenvalue of the Hamiltonian (see [33] for a review of history). The Margolus–Levitin theorem has been used to place bounds on the fundamental speed of computation [33], and was even slightly improved for some special cases [34]. Note, however, that for the Hamiltonian constructed in (45) we have $\operatorname{tr}(H\tilde{\rho}_1) = E_{\max}/2$ and hence the bound provided by Margolus–Levitin is in fact tight as we know that (45) lets us achieve the optimum success probability in time $t = \hbar\pi/E_{\max}$. This shows that it is our upper, rather than our lower, bound that can be improved.

Since we have $\gamma = 3/\pi$ or $\gamma = 5/\pi$ depending on the parameters, our bound is slightly worse than the Margolus–Levitin bound, which stems from our somewhat crude bound on $(1 - \cos(t E_n/\hbar))$. Note, however, that our bound considers a more specialized situation, namely turning the ancilla to its orthogonal given the encoding, but in turn applies to any kind of input states.

That we obtain a Margolus–Levitin-type theorem as a side effect of our analysis is not very surprising: clearly, the speed of dynamical evolution places a bound on how quickly we can transfer information from one system into the other. In turn, however, note that a

bound on how quickly transformation can be transferred does also translate into bounds on the speed of evolution and one can think of the speed of dynamical evolution when applied to a computation [33] as being limited by how quickly one can transfer the necessary information required for the subsequent stage of computation.

4.2. Time-dependent Tsirelson-bound

As another example on how our bound can be used we will derive a time-dependent Tsirelson's bound [23] for the Bell inequality [21] known as the CHSH inequality [22].

4.2.1. CHSH as a game.
We briefly describe the CHSH inequality in its more modern form as a game involving two distant players, Alice and Bob. A detailed account of this formulation and how it allows us to recover the original form of the CHSH inequality can, for example, be found in [35]. In the CHSH game, we imagine that we pose a question $y \in \{0, 1\}$ to Alice and a question $z \in \{0, 1\}$ to Bob, chosen uniformly at random, i.e. $p(y) = p(z) = 1/2$. These questions can be identified with the choice of measurement setting in the usual formulation. Alice and Bob now return answers $a \in \{0, 1\}$ and $b \in \{0, 1\}$, respectively, where we say that Alice and Bob *win* the game if and only if

$$y \cdot z = a + b \bmod 2. \tag{51}$$

Alice and Bob may thereby agree on any strategy beforehand, but they can no longer communicate once the game starts. In the quantum setting, this strategy corresponds to a choice of shared state and measurements, and in an experiment the no-signaling assumption is employed to enforce their inability to communicate. Clearly, one may write the probability that Alice and Bob win for a particular strategy as

$$p_{\text{win}} = \frac{1}{4} \sum_{y,z \in \{0,1\}} \sum_{\substack{a,b \\ a+b = y \cdot z}} \Pr[a, b | y, z], \tag{52}$$

where $\Pr[a, b | y, z]$ denotes the probability that Alice and Bob return answers a and b given questions y and z. For any classical strategy, $p_{\text{win}} \leqslant 3/4$ but quantumly there exists a strategy that achieves $p_{\text{win}} = 1/2 + 1/(2\sqrt{2}) \approx 0.853$. This is in fact optimal, since Tsirelson has shown [23, 36] that for any quantum strategy

$$p_{\text{win}} \leqslant \frac{1}{2} + \frac{1}{2\sqrt{2}}. \tag{53}$$

4.2.2. Strategies and state discrimination.
For our purposes, it will be convenient to employ a simple observation about what Bob has to do in order to produce the right answer in the game, which was described in more detail in [35]. Let $\rho_{y,a}$ denote the state of Bob's system conditioned on the fact that Alice received question y and has given answer a. Note that Bob's system will be placed in this state with probability $p(y, a) = p(a|y)/2$. For $z = 0$, (51) the rules of the game state that Alice and Bob win if and only if Bob returns the same answer as Alice, that is, $b = a$. In other words, Bob would like to determine, which of the following two states he is given

$$\sigma_0^{z=0} := \left(q_0^{z=0,0} \rho_{0,0} + q_1^{z=0,0} \rho_{1,0} \right), \tag{54}$$

$$\sigma_1^{z=0} := \left(q_0^{z=0,1} \rho_{0,1} + q_1^{z=0,1} \rho_{1,1} \right), \tag{55}$$

where

$$q_y^{z=0,0} = p(0|y)/(p(0|0) + p(0|1)), \tag{56}$$

$$q_y^{z=0,1} = p(1|y)/(p(1|0) + p(1|1)), \tag{57}$$

and the probability of $\sigma_x^{z=0}$ is given by $p_x^{z=0} = (p(x|0) + p(x|1))/2$. That is, Bob would simply try to extract classical information stored in quantum states, which is exactly the setting that our bound applies to. Producing a classical outcome on the ancilla system is very natural in this setting as we can imagine that when giving his answer Bob simply returns his ancilla to a referee who decides whether Alice and Bob win [7]. Similarly, if $z = 1$ Bob would like to determine which of the following two states he is given

$$\sigma_0^{z=1} := \left(q_0^{z=1,0} \rho_{0,0} + q_1^{z=1,0} \rho_{1,1} \right), \tag{58}$$

$$\sigma_1^{z=1} := \left(q_0^{z=1,1} \rho_{0,1} + q_1^{z=1,1} \rho_{1,0} \right), \tag{59}$$

where

$$q_{y=0}^{z=1,0} = p(0|0)/(p(0|0) + p(1|1)), \tag{60}$$

$$q_{y=1}^{z=1,0} = p(1|1)/(p(0|0) + p(1|1)), \tag{61}$$

$$q_{y=0}^{z=1,1} = p(1|0)/(p(1|0) + p(0|1)), \tag{62}$$

$$q_{y=1}^{z=1,1} = p(0|1)/(p(1|0) + p(0|1)), \tag{63}$$

the probability of $\sigma_0^{z=1}$ is $p_0^{z=1} = (p(0|0) + p(1|1))/2$, and the probability of $\sigma_1^{z=1}$ is $p_1^{z=1} = (p(1|0) + p(0|1))/2$. The probability that Alice and Bob win the game for a particular strategy can now be expressed as

$$p_{\text{win}} = \frac{1}{2} \sum_{z \in \{0,1\}} P_{\text{guess}}(X^z|E^z), \tag{64}$$

where we write $P_{\text{guess}}(X^z|E^z)$ for Bob's success probability in solving the state discrimination problems described above for $z \in \{0, 1\}$. From this perspective, Tsirelson's bound provides us with an upper bound on how well we can solve these two problems on average.

4.2.3. A time-limited game. In the usual setting of this game, Alice and Bob are essentially given an unlimited amount of time and energy to produce their answers. But how well can they do given only a limited amount of energy and time? Here, we consider a time-limited version of the CHSH game, in which Alice and Bob are given a fixed time t to produce their answers. If no answers are given at time t, we automatically rule that Alice and Bob lose. Our goal will be to derive a time-dependent version of (53). For simplicity, we will thereby assume that Alice has an essentially unlimited amount of energy at her disposal and only Bob will be restricted

---

[7] Note that our assumption that the referee is unrestricted contrasts with the view of computational complexity in which such games play a role in interactive proof systems. There, Alice and Bob are all-powerful, but the referee has limited time at his disposal to decide the outcome of the game. We would like to emphasize that our aim here is entirely different since we are merely interested in the strength of correlations between Alice and Bob that can be obtained within a certain time frame.

in some fashion. Given the perspective that Bob has to solve a state discrimination problem to produce the right answer as explained above, it is clear that we can use our general bound to address this setting. The use of an ancilla register is very natural, as we can view it as the message system holding Bob's answer that is returned to the referee.

In the usual scenario, Alice and Bob can choose which state to share at the start of the game as part of their strategy. Note, however, that we cannot allow arbitrary starting states to begin with, as we want to put a limit on the energy that Bob has at his disposal. For simplicity, however, we will make the sole assumption that Bob's Hamiltonian is bounded as $\|H\|_\infty$. In the appendix, we will derive a general time dependent Tsirelson bound from this assumption where we will need our generalization of the time bound for two input states to the case of non-uniform input distributions.

Here, we will focus on the essential idea that underlies this bound, which already becomes apparent if we consider a slightly simpler scenario in which Alice's marginal distributions are uniform ($p(a|y) = 1/2$ for all y). This scenario is well motivated if we imagine that there is a source supplying Alice and Bob with the maximally entangled state that lies outside of their control, and their strategy is restricted to their choice of two-outcome observables. In this case, Alice's outcome distribution will either be deterministic or uniform. In the deterministic case, Alice essentially plays a classical strategy. To obtain a quantum advantage in the case of unlimited time, Alice's outcome distributions will be uniform, and we will hence focus on this case.

To obtain a time-dependent Tsirelson bound, we now employ our simple bound involving the original trace distance of the two states that we wish to discriminate (34). We have by Tsirelson's bound that

$$\frac{1}{2} \sum_{z \in \{0,1\}} P_{\text{guess}}(X^z | E^z) \leqslant \frac{1}{2} + \frac{1}{2\sqrt{2}}, \tag{65}$$

and hence by (22) and the fact that $p(a|y) = 1/2$

$$\frac{1}{2} \sum_{z \in \{0,1\}} D(\sigma_0^z, \sigma_1^z) \leqslant \frac{1}{\sqrt{2}}, \tag{66}$$

otherwise there would exist a better strategy for Alice and Bob at long times. So we have from (34) that

$$p_{\text{win}} \leqslant \frac{1}{2} + \frac{\gamma t \|H\|_\infty}{2\sqrt{2}\hbar}. \tag{67}$$

In particular, this means that if we allow only a limited amount of energy by Bob (e.g. by demanding that $\|H\|_\infty = 1$), then Bob needs time at least

$$t \geqslant \frac{\hbar}{\gamma \|H\|_\infty} \tag{68}$$

to achieve the optimum quantum violation of CHSH. Note that to achieve the optimum quantum violation, Alice's marginals will in fact be uniform, and hence this is indeed the minimum time required.

Clearly, for small time frames, it would be better for Alice and Bob to play a classical strategy in which Bob can just return the ancilla $|0\rangle$ 'as is' to the referee. The tradeoff between the classical and quantum strategies in our setting can be captured when considering arbitrary distributions, which we will address in the appendix. In particular, we will show that

Corollary 4.1. *Let Bob's Hamiltonian be scaled such that $H \geqslant 0$. Then the maximum success probability of winning the CHSH game for Alice and Bob in time t obeys*

$$p_{\text{win}}^t \leqslant \frac{3}{4} + \frac{\gamma t \|H\|_\infty}{\sqrt{2}\hbar}, \tag{69}$$

where

$$\gamma := \begin{cases} 5/\pi, & \text{if } 1 < t E_n/\hbar < 4, \\ 3/\pi, & \text{otherwise.} \end{cases} \tag{70}$$

We could also derive a more general bound in terms of Bob's average energy using (40). However, such a bound does not compare easily to the original Tsirelon's bound.

Of course, the minimum time (68) is extremely small, and irrelevant for any practical tests of CHSH. Indeed, it is not our intention to question the validity of present CHSH experiments or suggest any loopholes caused by an insufficient distance for Alice and Bob compared to the time it takes them to achieve Tsirelson's bound. Instead, we provided the present analysis as an illustrative example of how our bound applies.

We would like to point out that (69) tells us that the strength of non-local correlation is indeed a function of time. Furthermore, (68) tells us that there exists a fundamental time required to establish maximally strong quantum correlations. Finally, we note that one can also interpret (69) in another way: let us suppose that we were to fix a time t and observe that Alice and Bob tend to win the game with a probability of at least q. We can now rewrite (69) to obtain a lower bound on $\|H\|_\infty$. That is, we can conclude that Bob had a certain energy at his disposal, and the strength of non-local correlations in this setting provides us with a form of 'energy witness' for Bob. This also holds for the most general case discussed in the appendix.

5. Discussion

5.1. Choice of initial state

We obtained a series of simple bounds on how well we can recover classical information stored in a quantum system within a certain time frame. Let us now first consider what role the choice of initial state of the ancilla played in our bounds. During our discussions we assumed that the ancilla started out in the classical state corresponding to the most likely symbol x_{max}. This reflects the fact that the distinguisher *does* have full knowledge not only about the states ρ_x themselves, but also about the distribution p_x. In particular, this means that without touching the quantum register, he can always achieve a success probability of $p_{x_{\text{max}}}$ by outputting x_{max}. Clearly, we could have chosen any other classical symbol as our starting point, and our bounds can easily be adapted accordingly. This holds even for an arbitrary pure state of the ancilla. However, such a choice does not reflect the distinguisher's *a priori* knowledge.

Another option would be to let the ancilla start out in a special blank state, which intuitively corresponds to an outcome of 'don't know' and is orthogonal to any other outputs. It is straightforward to apply our methods to obtain a similar bound for this case. However, note that using a blank ancilla state is conceptually rather different since it means that we essentially neglect the *a priori* knowledge that a distinguisher has available.

5.2. Input size

Our bound is especially useful, if we are merely concerned with the probability of success that can be achieved within a certain time *t in principle*, using any physically allowed operation H. This is indeed interesting when we consider the problem of Bell inequalities where we wanted to obtain a bound on how well Alice and Bob can violate CHSH within a given time frame, when they can *choose* any Hamiltonian they like subject to energy constraints alone. In particular, we would like to emphasize that the time required to acquire classical information in our setting is not limited by the size of the alphabet $\mathcal{X}$, but merely by the choice of encodings. In practice, however, there are much more stringent constraints on how quickly information can be transferred that depend on the geometry of the ancilla, leading to additional constraints on the interaction Hamiltonian H. For example, it could be that H can consist of only two qubit interactions, and interactions between the encoding system and the ancilla are limited to their boundary. In this case, the size of the alphabet $\mathcal{X}$ clearly *does* matter, and stronger bounds therefore should depend strongly on the exact form of H. We note that some bounds on time scales for particular Hamiltonians H do follow from the decoherence and thermodynamics literature [28, 29] for pure state encodings, yet since such bounds typically involve $\|H_{\text{int}}\|_\infty$, where H_{int} is the interacting part of H, they offer little advantage in our setting. To observe how such bounds are related to ours is most easily seen when considering the conditional von Neumann entropy $\mathrm{H}(X|E)$. Note that if all ρ_x are pure the overall cqq-state ρ_{XEA}[8] is pure as well. Hence, $\mathrm{H}(X|E) = \mathrm{H}(XE) - \mathrm{H}(E) = \mathrm{H}(A) - \mathrm{H}(E)$. To determine how $\mathrm{H}(X|E)$ can change with time we would thus like to determine how the entropy of the reduced systems A and E evolves with time, which has been studied for the von Neumann entropy in the decoherence literature where an upper bound for the rate of change in entropy was obtained in terms of $\|H_{\text{int}}\|_\infty$ [28]. Similar considerations can be made for other entropies [37]. It is an interesting open question to obtain good bounds on such quantities for arbitrary H that take more of their structure into account.

5.3. Open questions

 Clearly, this is not the only interesting open question. Closely related is the question of how much time is required to demonstrate non-local correlations if Alice and Bob are yet more restricted. Again, this could take the form of physical constraints on the ancilla, or be considered in the framework of circuit complexity where one cares about the number of two qubit interactions, i.e. gates, that they have to apply. The example of CHSH is too small for such constraints to make a difference, but they do play an important role when considering more complicated inequalities.

Furthermore, it would be nice to see if the slight gap between our bound and the Margolus–Levitin theorem can be closed completely using a more stringent analysis for the case of orthogonal encodings ρ_x. In particular, this means that one would rederive the exact form of the Margolus–Levitin theorem from the rate of information transfer alone.

---

[8] A cqq-state is a classical-quantum-classical state, here classical on the source registers X, quantum on the encoding system E and, before the referee's measurement, quantum on the ancilla A.

Appendix A. Basic observations

In this appendix, we provide the technical details of our claims. To this end, we first establish two simple lemmas from which we later derive all our results. Since we will use these in everything that follows we consider the generalized problem where we wish to distinguish N states $\rho_0, \ldots, \rho_{N-1}$. The first lemma will be used to bound the success probabilities using measurement operators $M_x = \mathbb{I} \otimes P_x + W_x$ where the label $x \in \{0, \ldots, N-1\}$ corresponds to one of the N states we wish to identify.

Lemma A.1. *For any Hermitian operator $A \in \mathcal{B}(\mathcal{H}_{\text{in}})$ with diagonalization $A = \sum_j \lambda_j |u_j\rangle\langle u_j|$, and any $x \in \{0, \ldots, N-1\}$ the operator $\tilde{A} := A \otimes |x_{\max}\rangle\langle x_{\max}|$ satisfies*

$$\text{tr}(W_x \tilde{A}) \leqslant 2 \sum_n (1 - \cos(t E_n/\hbar))\langle E_n | \tilde{A}_x | E_n\rangle, \tag{A.1}$$

where

$$\tilde{A}_x := \begin{cases} |A| \otimes |x_{\max}\rangle\langle x_{\max}|, & \text{for } x = x_{\max}, \\ A^+ \otimes |x_{\max}\rangle\langle x_{\max}|, & \text{otherwise,} \end{cases} \tag{A.2}$$

$$A^+ := \sum_{j, \lambda_j \geqslant 0} \lambda_j |u_j\rangle\langle u_j|. \tag{A.3}$$

Proof. Using the definition of W_x from (18) we evaluate the terms involving W_x^1 and W_x^2 separately. Let us now first bound the term involving W_x^1. For $x \neq x_{\max}$, we have that

$$\text{tr}(W_x^1 \tilde{A}) = \text{tr}((\mathbb{I} \otimes P_x) R \tilde{A}) + \text{tr}(R^\dagger (\mathbb{I} \otimes P_x) \tilde{A}) \tag{A.4}$$

$$= 0, \tag{A.5}$$

where we used the linearity and cyclicity of the trace, as well as the fact that $P_x |x_{\max}\rangle\langle x_{\max}| = |x_{\max}\rangle\langle x_{\max}| P_x = 0$ for all $x \neq x_{\max}$. Let $A^- := \sum_{j, \lambda_j < 0} |u_j\rangle\langle u_j|$, and define $\tilde{A}^+ := A^+ \otimes |x_{\max}\rangle\langle x_{\max}|$ and $\tilde{A}^- := A^- \otimes |x_{\max}\rangle\langle x_{\max}|$. Note that $\tilde{A} = \tilde{A}^+ - \tilde{A}^-$. For $x = x_{\max}$ we can now use the fact that

$$R = U - I = \sum_n (\exp(-i t E_n/\hbar) - 1)|E_n\rangle\langle E_n|, \tag{A.6}$$

to write

$$\text{tr}(W_{x_{\max}}^1 \tilde{A}) = \text{tr}((\mathbb{I} \otimes P_{x_{\max}}) R \tilde{A}) + \text{tr}(R^\dagger (\mathbb{I} \otimes P_{x_{\max}}) \tilde{A}) \tag{A.7}$$

$$= \text{tr}((R + R^\dagger) \tilde{A}) \tag{A.8}$$

$$= \sum_{n=0}^{\infty} (\exp(-\mathrm{i}t E_n/\hbar) + \exp(\mathrm{i}t E_n/\hbar) - 2)\langle E_n|\tilde{A}|E_n\rangle \tag{A.9}$$

$$= 2\sum_{n=0}^{\infty} (\cos(t E_n/\hbar) - 1)\langle E_n|\tilde{A}|E_n\rangle \tag{A.10}$$

$$= 2\sum_{n=0}^{\infty} (\cos(t E_n/\hbar) - 1)\langle E_n|(\tilde{A}^+ - \tilde{A}^-)|E_n\rangle \tag{A.11}$$

$$\leqslant 2\sum_{n=0}^{\infty} (1 - \cos(t E_n/\hbar))\langle E_n|\tilde{A}^-|E_n\rangle, \tag{A.12}$$

where the fourth equality follows from Euler's formula, and the first inequality from the fact that $\cos(t E_n/\hbar) - 1 \leqslant 0$ and $\tilde{A}^+, \tilde{A}^- \geqslant 0$.

It remains to bound the term involving W_x^2. First of all, since $\Lambda(X) = RXR^\dagger$ is a positive map [32], and $\tilde{A}^+, \tilde{A}^- \geqslant 0$, we have that

$$R\tilde{A}^+ R^\dagger \geqslant 0, \tag{A.13}$$

$$R\tilde{A}^- R^\dagger \geqslant 0. \tag{A.14}$$

Note that for any $X, Z \geqslant 0$, we have $\mathrm{tr}(XZ) \geqslant 0$, and hence $\mathrm{tr}((\mathbb{I} \oplus P_x)R\tilde{A}^- R^\dagger) \geqslant 0$. Secondly, note that $RR^\dagger = R^\dagger R$ and we have

$$RR^\dagger = \sum_{n=0}^{\infty} (2 - \exp(\mathrm{i}t E_n/\hbar) - \exp(-\mathrm{i}t E_n/\hbar))\, |E_n\rangle\langle E_n| \tag{A.15}$$

$$= 2\sum_n (1 - \cos(t E_n/\hbar))|E_n\rangle\langle E_n|, \tag{A.16}$$

where the second equality follows by applying Euler's formula. We thus have

$$\mathrm{tr}(W_x^2\tilde{A}) = \mathrm{tr}((\mathbb{I} \otimes P_x)R\tilde{A}^+ R^\dagger) - \mathrm{tr}((\mathbb{I} \otimes P_x)R\tilde{A}^- R^\dagger) \tag{A.17}$$

$$\leqslant \mathrm{tr}(R^\dagger R\tilde{A}^+) \tag{A.18}$$

$$= 2\sum_n (1 - \cos(t E_n/\hbar))\langle E_n|\tilde{A}^+|E_n\rangle, \tag{A.19}$$

where the first inequality follows from (A.13), the fact that $0 \leqslant \mathbb{I} \otimes P_x \leqslant \mathbb{I}$ and the cyclicity of the trace, and the last equality from (A.15). Putting everything together, $\mathrm{tr}(W_x\tilde{A}) = \mathrm{tr}(W_x^1\tilde{A}) + \mathrm{tr}(W_x^2\tilde{A})$, we obtain the claimed result. $\qquad\square$

We will also make repeated use of the following bound. Note that whereas the bound applies to a very large range of values $E_n \geqslant 0$, we will later be particularly interested in the case of $t E_n/\hbar < 1$. Indeed the bound below is a great overestimate if $t E_n/\hbar > 2\pi$, as $2(1 - \cos(k)) \leqslant \gamma (k - 2\pi \lfloor k/(2\pi) \rfloor)$.

Lemma A.2. *Let $E_n \geqslant 0$. Then $2(1 - \cos(t E_n/\hbar)) \leqslant \gamma t E_n/\hbar$ where*

$$\gamma := \begin{cases} 5/\pi, & \text{if } 1 < t E_n/\hbar < 4, \\ 3/\pi, & \text{otherwise.} \end{cases} \tag{A.20}$$

Appendix B. A bound for two states

We now first consider the case where we are given just two states, ρ_0 and ρ_1. Here, we will consider the most general problem where p_0 and p_1 can be arbitrary.

B.1. A bound in terms of the trace distance

First of all, note that even for a general distribution $\{p_x\}_x$ the problem of distinguishing two states is easy to analyze [2]. In particular, we have that in the time-unlimited case for measurement operators acting directly on the encoding space

$$P_{\text{guess}}(X|E) = \max_{M_0, M_1} p_0 \, \text{tr}(M_0 \rho_0) + p_1 \, \text{tr}(M_1 \rho_1) \tag{B.1}$$

$$= p_0 + \max_{M_1} \text{tr}(M_1(p_1 \rho_1 - p_0 \rho_0)) \tag{B.2}$$

$$= p_0 + \Delta(p_1 \rho_1, p_0 \rho_0), \tag{B.3}$$

where $\Delta(p_1 \rho_1, p_0 \rho_0)$ is given by

$$\Delta(p_1 \rho_1, p_0 \rho_0) = \max_{0 \leqslant P \leqslant \mathbb{I}} \text{tr}(PA) \tag{B.4}$$

$$= \text{tr}(A^+), \tag{B.5}$$

where $A := p_1 \rho_1 - p_0 \rho_0$ with diagonalization $A = \sum_j \lambda_j |u_j\rangle\langle u_j|$ and $A^+ = \sum_{j, \lambda_j \geqslant 0} |u_j\rangle\langle u_j|$. (Note that Δ is not symmetric here and hence formally does not form a distance measure.) Similarly, we have

$$\Delta(p_0 \rho_0, p_1 \rho_1) = \text{tr}(A^-). \tag{B.6}$$

Note that for the time unlimited case, we could have equivalently expressed the success probability as

$$P_{\text{guess}}(X|E) = p_1 + \Delta(p_0 \rho_0, p_1 \rho_1). \tag{B.7}$$

It will also be useful to note that for $\tilde{\rho}_x = \rho_x \otimes |x_{\max}\rangle\langle x_{\max}|$,

$$\Delta(p_1 \rho_1, p_0 \rho_0) = \Delta(p_1 \tilde{\rho}_1, p_0 \tilde{\rho}_0). \tag{B.8}$$

Before stating our bound, let us introduce some additional notation. For two states, define

$$x_{\min} := 1 - x_{\max}. \tag{B.9}$$

We now first relate the problem of discriminating the two states in time t to the original success probability.

Lemma B.1. *The probability of distinguishing ρ_0 and ρ_1 given with probabilities p_0 and p_1 using the Hamiltonian $H = \sum_n E_n |E_n\rangle\langle E_n| \geqslant 0$ is bounded by*

$$P_{\text{succ}}(X|E)_{H,t} \leqslant p_{x_{\max}} + 2(1 - \cos(tC_{\max}/\hbar))\Delta(p_{x_{\min}}\rho_{x_{\min}}, p_{x_{\max}}\rho_{x_{\max}}), \tag{B.10}$$

where $C_{\max} = \text{argmax}_{E_n}(1 - \cos(tE_n/\hbar))$.

Proof. Using (B.1), (17) and the fact that $P_{x_{\min}}|x_{\max}\rangle = 0$ we may bound the success probability as

$$P_{\text{succ}}(X|E)_{H,t} \leqslant p_{x_{\max}} + \text{tr}(W_{x_{\min}}(p_{x_{\min}}\tilde{\rho}_{x_{\min}} - p_{x_{\max}}\tilde{\rho}_{x_{\max}})). \tag{B.11}$$

Applying lemma A.1 for $A = p_{x_{\min}}\rho_{x_{\min}} - p_{x_{\max}}\rho_{x_{\max}}$ we have that

$$\text{tr}(W_{x_{\min}}(A \otimes |x_{\max}\rangle\langle x_{\max}|)) \leqslant 2\sum_n(1 - \cos(tE_n/\hbar))\langle E_n|A^+ \otimes |x_{\max}\rangle\langle x_{\max}||E_n\rangle. \tag{B.12}$$

Hence from (B.5) and (B.8) we have

$$\text{tr}(W_{x_{\min}}(p_{x_{\min}}\tilde{\rho}_{x_{\min}} - p_{x_{\max}}\tilde{\rho}_{x_{\max}})) \leqslant 2(1 - \cos(tC_{\max}/\hbar))\Delta(p_{x_{\min}}\rho_{x_{\min}}, p_{x_{\max}}\rho_{x_{\max}}). \tag{B.13}$$

Our claim now follows by plugging this bound into (B.11). □

With the help of lemma A.2 for $E_n = C_{\max}$ one may now also use the fact that $\forall E_n$, $E_n \leqslant \|H\|_\infty$ to obtain a very simple bound in terms of the spectrum of the Hamiltonian.

Corollary B.2. *The probability of distinguishing ρ_0 and ρ_1 given with probabilities p_0 and p_1 using the Hamiltonian $H = \sum_n E_n|E_n\rangle\langle E_n| \geqslant 0$ is bounded by*

$$P_{\text{succ}}(X|E)_{H,t} \leqslant p_{x_{\max}} + \frac{\gamma t\|H\|_\infty \Delta(p_{x_{\min}}\rho_{x_{\min}}, p_{x_{\max}}\rho_{x_{\max}})}{\hbar}. \tag{B.14}$$

B.2. A bound in terms of the average energy

Inspecting the proof above with lemma A.2 in mind, it is indeed easy to see that we can also obtain a bound in terms of average energies. We first derive a somewhat stronger bound for two equiprobable states that actually depends on the 'average energy' of a function of both states.

Theorem B.3. *The probability of distinguishing ρ_0 and ρ_1 given with probabilities p_0 and p_1 using the Hamiltonian $H = \sum_n E_n|E_n\rangle\langle E_n| \geqslant 0$ in time t is bounded by*

$$P_{\text{succ}}(X|E)_{H,t} \leqslant p_{x_{\max}} + \frac{\gamma t}{2\hbar}\left[\text{tr}(H|p_{x_{\min}}\tilde{\rho}_{x_{\min}} - p_{x_{\max}}\tilde{\rho}_{x_{\max}}|) + p_{x_{\min}}\text{tr}(H\tilde{\rho}_{x_{\min}}) - p_{x_{\max}}\text{tr}(H\tilde{\rho}_{x_{\max}})\right], \tag{B.15}$$

where

$$\gamma := \begin{cases} 5/\pi, & \text{if } \exists E_n, 1 < tE_n/\hbar < 4, \\ 3/\pi, & \text{otherwise.} \end{cases} \tag{B.16}$$

Proof. Recall that by applying lemma A.1 with $A = p_{x_{\min}}\rho_{x_{\min}} - p_{x_{\max}}\rho_{x_{\max}}$ we have that

$$\text{tr}(W_{x_{\min}}(p_{x_{\min}}\tilde{\rho}_{x_{\min}} - p_{x_{\max}}\tilde{\rho}_{x_{\max}})) \leqslant 2\sum_n(1 - \cos(tE_n/\hbar))\langle E_n|A^+ \otimes |x_{\max}\rangle\langle x_{\max}||E_n\rangle. \tag{B.17}$$

We may now use lemma A.2 to obtain

$$\text{tr}(W_{x_{\min}}(p_{x_{\min}}\tilde{\rho}_{x_{\min}} - p_{x_{\max}}\tilde{\rho}_{x_{\max}})) \tag{B.18}$$

$$\leqslant \frac{\gamma t}{\hbar}\sum_n E_n\langle E_n|A^+ \otimes |x_{\max}\rangle\langle x_{\max}||E_n\rangle \tag{B.19}$$

$$= \frac{\gamma t \, \mathrm{tr}(H\tilde{A}^+)}{\hbar}. \tag{B.20}$$

Our claim now follows by noting that

$$\tilde{A}^+ = \tfrac{1}{2}(\tilde{A}^+ + \tilde{A}^-) + \tfrac{1}{2}(\tilde{A}^+ - \tilde{A}^-) \tag{B.21}$$

$$= \tfrac{1}{2} \left| p_{x_{\min}} \tilde{\rho}_{x_{\min}} - p_{x_{\max}} \tilde{\rho}_{x_{\max}} \right| + \tfrac{1}{2} \left(p_{x_{\min}} \tilde{\rho}_{x_{\min}} - p_{x_{\max}} \tilde{\rho}_{x_{\max}} \right). \tag{B.22}$$

□

Appendix C. A bound for many input states

Finally, we derive a bound for the most general case of distinguishing states $\rho_0, \ldots, \rho_{N-1}$ where we are given ρ_x with probability p_x.

Theorem C.1. *Suppose $H \geqslant 0$. Then the probability of distinguishing $\rho_0, \ldots, \rho_{N-1}$ given with probabilities $p_0, \ldots, p_{N-1}$ obeys*

$$P_{\mathrm{succ}}(X|E)_{H,t} \leqslant p_{x_{\max}} + \frac{\hat{\gamma} t}{\hbar} \sum_{x=0}^{N-1} p_x \, \mathrm{tr}(H\tilde{\rho}_x), \tag{C.1}$$

where

$$\hat{\gamma} := \begin{cases} 5/\pi, & \text{if } \exists E_n, \, 1 < t E_n/\hbar < 4, \\ 3/\pi, & \text{otherwise.} \end{cases} \tag{C.2}$$

Proof. Note that the success probability for a particular interaction H is now given by

$$P_{\mathrm{succ}}(X|E)_{H,t} = \sum_{x=0}^{N-1} p_x \, \mathrm{tr}\left(M_x \tilde{\rho}_x\right), \tag{C.3}$$

where

$$\mathrm{tr}(M_x \tilde{\rho}_x) = \mathrm{tr}((\mathbb{I} \otimes P_x)\tilde{\rho}_x) + \mathrm{tr}(W_x \tilde{\rho}_x). \tag{C.4}$$

Let us now first consider the case of $x = x_{\max}$. We have that

$$\mathrm{tr}((\mathbb{I} \otimes P_x)\tilde{\rho}_x) = \mathrm{tr}(\rho_x) = 1. \tag{C.5}$$

Using lemma A.1 with $A = \rho_x$ we hence have that

$$\mathrm{tr}(M_x \tilde{\rho}_x) \leqslant 1 + 2 \sum_n (1 - \cos(t E_n/\hbar)) \langle E_n | \tilde{\rho}_x | E_n \rangle. \tag{C.6}$$

We now turn to the case of $x \neq x_{\max}$. Since $P_x |x_{\max}\rangle\langle x_{\max}| = |x_{\max}\rangle\langle x_{\max}| P_x = 0$ and $\tilde{\rho}_x = \rho_x \otimes |x_{\max}\rangle\langle x_{\max}|$ we have $(\mathbb{I} \otimes P_x)\tilde{\rho}_x = 0$ for all $x \neq x_{\max}$. Again by applying A.1 with $A = \rho_x$ we obtain from $\rho_x \geqslant 0$ that

$$\mathrm{tr}(W_x \tilde{\rho}_x) \leqslant 2 \sum_n (1 - \cos(t E_n/\hbar)) \langle E_n | \tilde{\rho}_x | E_n \rangle. \tag{C.7}$$

Our claim now follows by using lemma A.2 to obtain $2(1 - \cos(t E_n/\hbar)) \leqslant \gamma t E_n/\hbar$ for $E_n \geqslant 0$. □

Appendix D. Attaining the bound

We now exhibit a Hamiltonian that achieves our upper bound for the success probability of distinguishing two states given with *a priori* equal probability.

Theorem D.1. *Suppose we are given ρ_0 and ρ_1 with* a priori *equal probability. Let $E_{\max} \geqslant 0$. Then there exist a Hamiltonian H with $\|H\|_\infty = E_{\max}$ that in time t achieves success probability*

$$P_{\text{succ}}(X|E)_{H,t} = \frac{1}{2} + \frac{1}{4}(1 - \cos(t\|H\|_\infty/\hbar))D(\rho_0, \rho_1). \tag{D.1}$$

In particular, we can distinguish the two states perfectly in time $t = \hbar\pi/E_{\max}$.

Proof. Let $A = \rho_1 - \rho_0$. We can diagonalize $A = \sum_j \lambda_j |u_j\rangle\langle u_j|$, and define $A^+ := \sum_{j,\lambda_j \geqslant 0} \lambda_j |u_j\rangle\langle u_j|$ and $A^- := A - A^+$. Consider the operator

$$\hat{H} := \Pi_{A^-} \otimes \mathbb{I} + \Pi_{A^+} \otimes (|0\rangle\langle 1| + |1\rangle\langle 0|), \tag{D.2}$$

where Π_{A^+} and Π_{A^-} are projectors on the support of A^+ and A^-, respectively. Clearly, $\hat{H}$ is Hermitian and unitary, and hence has eigenvalues ± 1. We now define the Hamiltonian H

$$H := E_{\max}(\hat{H} + \mathbb{I})/2. \tag{D.3}$$

Since any term in the Hamiltonian proportional to the identity does not affect the dynamics, we may replace the time evolution operator $\exp(-itH/\hbar)$ with

$$U = \exp(-it\|H\|_\infty\hat{H}/2\hbar) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{-it\|H\|_\infty\hat{H}}{2\hbar}\right)^n. \tag{D.4}$$

For our choice of $\hat{H}$ we have that

$$(\hat{H})^n = \begin{cases} \mathbb{I} & n \text{ even,} \\ \hat{H} & n \text{ odd.} \end{cases} \tag{D.5}$$

and as a result the Taylor expansion for U gives

$$U = (\cos(t\|H\|_\infty/2\hbar)\mathbb{I} - i\sin(t\|H\|_\infty/2\hbar)\hat{H}). \tag{D.6}$$

Using this unitary in our state discrimination problem, we obtain

$$P_{\text{succ}}(X|E)_{H,t} = \frac{1}{2} + \frac{1}{2}\,\text{tr}\left(U^\dagger(\mathbb{I}\otimes P_1)U(\tilde{\rho}_1 - \tilde{\rho}_0)\right) \tag{D.7}$$

$$= \frac{1}{2} + \frac{\sin^2(t\|H\|_\infty/2\hbar)}{2}\,\text{tr}(\hat{H}(\mathbb{I}\otimes P_1)\hat{H}(\tilde{\rho}_1 - \tilde{\rho}_0)). \tag{D.8}$$

It remains to evaluate the last term. First of all, note that

$$\hat{H}(\mathbb{I}\otimes P_1)\hat{H} = (\Pi_{A^+} \otimes |0\rangle\langle 0| + \Pi_{A^-} \otimes |1\rangle\langle 1|). \tag{D.9}$$

Since $\tilde{\rho}_1 - \tilde{\rho}_0 = (A^+ - A^-) \otimes |0\rangle\langle 0|$ and $\Pi_{A^+}\Pi_{A^-} = \Pi_{A^-}\Pi_{A^+} = 0$, we thus have

$$\text{tr}(\hat{H}(\mathbb{I}\otimes P_1)\hat{H}(\tilde{\rho}_1 - \tilde{\rho}_0)) = \text{tr}(A^+) \tag{D.10}$$

$$= D(\rho_0, \rho_1). \tag{D.11}$$

The claim follows by an application of the double angle formula. □

Appendix E. Constraining the eigenvalues of *H*

For completeness, we now remind ourselves why in many settings it is not unreasonable to assume that $\|H\|_\infty$ is indeed bounded. Note that when dealing with fixed input states $\rho_0, \ldots, \rho_{N-1}$, we can without loss of generality assume that the Hamiltonian H that leads to the optimal success probability possible within a certain time t is limited to the energy eigenspace sufficient to contain the support of the inputs states and the standard ancilla state. This holds even in an approximate sense. To see this, consider how the state on the encoding register and the ancilla evolve in time

$$\tilde{\rho}_0(t) := U(t)\tilde{\rho}_0 U(t)^\dagger, \tag{E.1}$$

$$\tilde{\rho}_1(t) := U(t)\tilde{\rho}_1 U(t)^\dagger. \tag{E.2}$$

Choose an error parameter ε and define Π_δ^ε to be the lowest rank operator such that $[\Pi_\delta^\varepsilon, H] = 0$ and

$$\tfrac{1}{2}\|\rho_0(0) - \Pi_\delta^\varepsilon \rho_0(0)\Pi_\delta^\varepsilon\|_1 \leqslant \varepsilon, \tag{E.3}$$

$$\tfrac{1}{2}\|\rho_1(0) - \Pi_\delta^\varepsilon \rho_1(0)\Pi_\delta^\varepsilon\|_1 \leqslant \varepsilon. \tag{E.4}$$

Since $[\Pi_\delta^\varepsilon, H] = 0$ and the L1-norm is unitarily invariant, we can immediately conclude that (E.3) still holds when we replace $\rho_0(0)$ and $\rho_1(0)$ with any subsequent states $\rho_0(t)$ and $\rho_1(t)$. However, this tells us that we can approximate U with a unitary $\hat{U}$ as

$$\hat{H} := \Pi_\delta^\varepsilon H \Pi_\delta^\varepsilon, \tag{E.5}$$

$$\hat{U}(t) := \exp(-\mathrm{i}\hat{H}t/\hbar), \tag{E.6}$$

without affecting any of the output states, except with a chosen error ε. By the definition of the L1-norm we have for any Hermitian operator A that

$$\tfrac{1}{2}\|A\|_1 = \sup_{-\mathbb{I} \leqslant P \leqslant \mathbb{I}} \mathrm{tr}(PA), \tag{E.7}$$

and hence (E.3) implies that using the unitary $\hat{H}$ in place of any original H leads to a change in success probability in the state discrimination problem of at most 2ε.

Appendix F. A general time-dependent Tsirelson's bound

Let us now consider a more general version of our time-dependent Tsirelson's bound in which we drop the assumption that the source emits a particular state, and that Alice makes a two-outcome projective measurement. The only assumption we will make now is that Bob's Hamiltonian is bounded $\|H\|_\infty = E_{\max}$.

For our proof, we will need the more general version of the two state discrimination problem in which the two states are not necessarily given with equal probabilities (see corollary B.2). Again, let us first briefly consider the time-unlimited case, where M_0 and M_1 are just measurements on a single system. Recall that we could express the success probability of distinguishing ρ_0 and ρ_1 given with probabilities p_0 and p_1, respectively, as

$$P_{\text{guess}}(X|E) = p_0 + \Delta(p_1\rho_1, p_0\rho_0). \tag{F.1}$$

At first glance, this expression appears a bit asymmetric—after all, what should be so special about p_0? Note, however, that by replacing $M_1 = \mathbb{I} - M_0$ in (B.1) we could also have expressed the success probability as

$$P_{\text{guess}}(X|E) = p_1 + \Delta(p_0\rho_0, p_1\rho_1).$$

In particular, it will be convenient to note that we could have also written the success probability as the average of these two terms

$$P_{\text{guess}}(X|E) = \frac{1}{2}\left(1 + \sum_{x\in\{0,1\}} \Delta(p_{\bar{x}}\rho_{\bar{x}}, p_x\rho_x)\right). \tag{F.2}$$

Let us now return to the time limited case, involving an interaction of the encoding and ancilla system, followed by a measurement on the ancilla. Recall that we have from corollary B.2 that

$$P_{\text{guess}}(X|E)_{H,t} \leqslant p_{x_{\max}} + \left(\frac{t\gamma\|H\|_\infty}{\hbar}\right)\Delta(p_{x_{\min}}\rho_{x_{\min}}, p_{x_{\max}}\rho_{x_{\max}}). \tag{F.3}$$

Note that in the time limited case we cannot simply average—the proof of corollary B.2 yields a different bound had we placed $p_{x_{\min}}$ in front (a small calculation shows that it will again single out $p_{x_{\max}}$). We are now ready to show our general bound, where we will use the notation developed in section 4.2.2.

Lemma F.1. *Let Bob's Hamiltonian be scaled such that $H \geqslant 0$. Then the maximum success probability of winning the CHSH game for Alice and Bob in time t obeys*

$$p_{\text{win}}^t \leqslant \frac{1}{2}\left(\sum_{z\in\{0,1\}} p_{x_{\max}}^z\right) + \frac{\gamma t\|H\|_\infty}{\sqrt{2}\hbar}, \tag{F.4}$$

where

$$\gamma := \begin{cases} 5/\pi, & \text{if } 1 < tC_{\max}/\hbar < 4, \\ 3/\pi, & \text{otherwise}, \end{cases} \tag{F.5}$$

with $C_{\max} = \text{argmax}_{E_n}(1 - \cos(tE_n/\hbar))$.

Proof. We have from Tsirelson's bound [23] that for any strategy of Alice and Bob no matter how much time or energy they may have available

$$p_{\text{win}} = \frac{1}{2} \sum_{z\in\{0,1\}} P_{\text{guess}}(X^z|E^z) \tag{F.6}$$

$$= \frac{1}{4} \sum_z \left(1 + \sum_{x\in\{0,1\}} \Delta(p_{\bar{x}}^z\sigma_{\bar{x}}^z, p_x^z\sigma_x^z)\right) \tag{F.7}$$

$$\leqslant \frac{1}{2} + \frac{1}{2\sqrt{2}}. \tag{F.8}$$

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Rearranging terms gives us

$$\frac{1}{4} \sum_{z,x \in \{0,1\}} \Delta(p_{\bar{x}}^z \sigma_{\bar{x}}^z, p_x^z \sigma_x^z) \leqslant \frac{1}{2} + \frac{1}{2\sqrt{2}} - \frac{1}{2} \tag{F.9}$$

$$= \frac{1}{2\sqrt{2}}. \tag{F.10}$$

Since $\Delta(\cdot, \cdot) \geqslant 0$, this means that

$$\frac{1}{2} \sum_z \Delta(p_{x_{\min}}^z \sigma_{x_{\min}}^z, p_{x_{\max}}^z \sigma_{x_{\max}}^z) \leqslant \frac{1}{\sqrt{2}}. \tag{F.11}$$

Our claim for the time limited case now follows by plugging this bound into (F.3)

$$p_{\text{win}}^t = \frac{1}{2} \sum_{z \in \{0,1\}} P_{\text{guess}}(X^z|E^z)_{H,t} \tag{F.12}$$

$$\leqslant \frac{1}{2} \left(\sum_{z \in \{0,1\}} p_{x_{\max}}^z \right) + \frac{C}{2} \sum_{z,x \in \{0,1\}} \Delta(p_{x_{\min}}^z \sigma_{x_{\min}}^z, p_{x_{\max}}^z \sigma_{x_{\max}}^z), \tag{F.13}$$

$$\leqslant \frac{1}{2} \left(\sum_{z \in \{0,1\}} p_{x_{\max}}^z \right) + \frac{C}{\sqrt{2}}, \tag{F.14}$$

where we have used the shorthand

$$C := \frac{t\gamma \|H\|_\infty}{\hbar}. \tag{F.15}$$

$\square$

At first glance, this bound may seem somewhat strange as it involves a potentially unknown term $\sum_x p_{x_{\max}}^z$. Note, however, that this term is determined by the distributions over the states that Bob should distinguish. It is this distribution, that determines the classical bound for CHSH and hence

$$\frac{1}{2} \sum_z p_{x_{\max}}^z \leqslant \frac{3}{4}. \tag{F.16}$$

Note that since the ancilla is initialized to the most likely x in each case, there exists a strategy for Alice and Bob with which they can play optimally classically in no time at all. However, since there is generally an interplay between the choice of distributions and the state Alice can create we derived the bound in its more general form above.

Corollary F.2. *Let Bob's Hamiltonian be scaled such that $H \geqslant 0$. Then the maximum success probability of winning the CHSH game for Alice and Bob in time t obeys*

$$p_{\text{win}}^t \leqslant \frac{3}{4} + \frac{\gamma t \|H\|_\infty}{\sqrt{2}\hbar}, \tag{F.17}$$

where

$$\gamma := \begin{cases} 5/\pi, & \text{if } 1 < tC_{\max}/\hbar < 4, \\ 3/\pi, & \text{otherwise,} \end{cases} \tag{F.18}$$

with $C_{\max} = \text{argmax}_{E_n}(1 - \cos(t E_n/\hbar))$.

IOP Institute of Physics 🔶 DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

References

[1] König R, Renner R and Schaffner C 2009 *IEEE Trans. Inf. Theory* **55** 4337
[2] Helstrom C W 1967 *Inf. Control* **10** 254
[3] Holevo A S 1973 *Probl. Pereda. Inf.* **9** 3
    Holevo A S 1973 *Prob. Inf. Transm.* **9** 177–83 (Engl. Transl.)
[4] Belavkin V P 1975 *Stochtics* **1** 315
[5] Yuen H P, Kennedy R S and Lax M 1975 *IEEE Trans. Inf. Theory* **21** 125
[6] Barnett S M and Croke S 2009 *Adv. Opt. Photonics* **1** 238
[7] Deconick M and Terhal B 2010 *Phys. Rev.* A **81** 062304
[8] Eldar Y, Megretski A and Verghese G 2003 *IEEE Trans. Inf. Theory* **49** 1007
[9] Eldar Y 2003 *IEEE Trans. Inf. Theory* **49** 446
[10] Reimpell M and Werner R F 2005 *Phys. Rev. Lett.* **94** 080501
[11] Jezek M, Rehacek J and Fiurasek J 2002 *Phys. Rev.* A **65** 060301
[12] Jezek M, Fiurasek J and Hradil Z 2003 *Phys. Rev.* A **68** 012305
[13] Tyson J 2009 arXiv:0902.0395
[14] Tyson J 2010 *J. Math. Phys.* **51** 092204
[15] Belavkin V P 1975 *Radio Eng. Electron. Phys.* **20** 1177
[16] Mochon C 2007 *Phys. Rev.* A **75** 042313
[17] Hausladen P and Wootters W 1994 *J. Mod. Opt.* **41** 2385
[18] Tyson J 2009*a* *Phys. Rev.* A **79** 032343
[19] Tyson J 2009*b* *J. Math. Phys.* **50** 032106
[20] Gopal D and Wehner S 2010 *Phys. Rev.* A **82** 022326
[21] Bell J S 1964 *Physics* **1** 195
[22] Clauser J, Horne M, Shimony A and Holt R 1969 *Phys. Rev. Lett.* **23** 880
[23] Tsirelson B 1980 *Lett. Math. Phys.* **4** 93
[24] Margolus N and Levitin L B 1998 *Physica* D **120** 188
[25] Lebedev D B and Levitin L B 1966 *IRE Inf. Control* **9** 1
[26] Pendry J B 1983 *J. Phys. A: Math. Gen.* **16** 2161
[27] Bekenstein J D and Schiffer M 1990 *Int. J. Mod. Phys.* C **1** 355
[28] Rodriguez-Rosario C A, Kimura G, Imai H and Aspuru-Guzik A 2011 *Phys. Rev. Lett.* **106** 050403
[29] Gogolin C 2010 *PhD Thesis* arXiv:1003.5058
[30] Kimura G, Ohno H and Hayashi H 2007 *Phys. Rev.* A **76** 042123
[31] Renner R 2005 *PhD Thesis* ETH Zurich, Germany, arXiv:quant-ph/0512258
[32] Bhatia R 2007 *Positive Definite Matrices* (Princeton, NJ: Princeton University Press)
[33] Lloyd S 2000 *Nature* **406** 1047
[34] Zielinski B and Zych M 2006 *Phys. Rev.* A **75** 034301
[35] Wehner S 2008 *PhD Thesis* University of Amsterdam arXiv:0806.3483
[36] Tsirelson B 1987 *J. Sov. Math.* **36** 557
[37] Hutter A and Wehner S 2011 in preparation