Robust self-testing of (almost) all pure two-qubit states

MSc Thesis *(Afstudeerscriptie)*

written by

Tim Coopmans

under the supervision of **Dr Jędrzej Kaniewski** and **Dr Christian Schaffner**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense:	**Members of the Thesis Committee:**
April 25, 2017	Dr Serge Fehr
	Dr Jędrzej Kaniewski
	Dr Christian Schaffner
	Prof Sonja Smets
	Prof Ronald de Wolf (chair)

Institute for Logic, Language and Computation

Abstract

In a nonlocal scenario, physically isolated players each have a device that inputs and outputs classical information. Certain correlations between the joint input and output of the devices almost uniquely identify the quantum state that they share. This phenomenon is known as self-testing and has applications in quantum cryptography with untrusted devices. It was for example shown that for every pure two-qubit state, there exists a two-player Bell experiment whose correlations, or rather the Bell value that is computed from the correlations, can be used here to self-test that state; the Bell value that is used stems from a family of Bell inequalities called *tilted CHSH inequalities*. A special case is the regular *CHSH inequality*, which is used to self-test the singlet, a maximally entangled state of two qubits. For practical applications, estimation errors and the presence of external noise require self-testing statements to be robust to errors.

In this thesis, we extend previous work on self-testing of the singlet with the CHSH game. First, we use tilted CHSH inequalities to improve the robustness of previously found self-testing statements for (almost) all pure partially entangled states. Our result consists of the explicit construction of local quantum channels for the two players, from which we derive operator inequalities that we verify numerically. Using a recently developed method [Kan16], the improved bounds can be inferred. Furthermore, we construct a state that violates the CHSH inequality but for which there exist no local quantum channels that achieve greater fidelity than a trivial lower bound (i.e. achieve fidelity with the singlet greater than what is achievable using a separable state). This result implies that CHSH violation is not sufficient for the two players to 'extract' a singlet from their actual state by just local operations. Future research could focus on extending our results to different self-testable states such as the GHZ-states and on self-testing in the scenario where only one of the two players has a potentially untrusted device (quantum steering).

Keywords: self-testing, extractability, device-independence, quantum cryptography

Acknowledgments

First and foremost, I would like to thank my supervisor Jed Kaniewski, who spent a great deal of his time on me - someone said that it must have been 75% - to teach, help and pave the way for me to get a better grasp of the strange world of quantum information theory. Regardless of the nature of my questions (elementary, repetitive, or about academia in general), Jed made time to answer them, even early in the morning or late at night. Thank you for the good times in Copenhagen, Jed; I enjoyed working together and I hope that many students after me will be able to enjoy your enthusiastic, relaxed and especially very friendly character, not to mention your music lyrics jokes.

A lot of thanks also to Christian Schaffner, who in the first place gave me the opportunity to go and work under supervision of a fellow researcher. Even though I was abroad, he kept spending time to listen to what Jed and I were doing. Now and then he asked how things were going non-researchwise, which is indicative of his quality as a supervisor.

It has been a wonderful experience to stay in Copenhagen. I would like to thank the members of the QMath group for making me feel part of the group in the very first week (which started with climbing trees). Many thanks also to Matthias Christandl, the head of the group, for funding part of my stay.

I would also like to thank the thesis committee members, Serge Fehr, Sonja Smets, and Ronald de Wolf, for taking the time to read this thesis.

Many thanks to the international friends who kept my mind off research questions when needed, and to my friends at home for keeping in touch with me while I was abroad.

Special thanks to Marlies, for being there at exactly the right moments.

Contents

1 Introduction

The first ideas about 'quantum computers' were developed in the 1980s [Fey82, Ben82]; such computers could exploit properties that are exclusive to quantum systems, such as quantum entanglement - the property that generally, the components of quantum systems cannot be described independently. Since then, many applications of quantum computers have been found. The most promising of such applications are arguably the integer factorization algorithm by Shor [Sho94], which can break today's widely-used RSA encryption algorithm, the database search algorithm which was originally designed by Grover and later improved by Boyer et al. [Gro96, BBHT98], and the development of quantum-key-distribution (QKD) protocols, which started with the work of Bennett and Brassard [BB84] and Ekert [Eke91].

In the latter of these branches, the central question is how to establish a random bit string (the key) that is known only to the two parties who wish to communicate. QKD schemes heavily rely on the quantum-mechanical properties of the devices used for their execution, such as entanglement. Indeed, it has been shown that an eavesdropper can use imperfections in the devices in order to break security [SK14]. In practice, however, quantum entanglement is difficult to maintain due to interference of external noise with the entangled system. Additionally, the security proofs of several QKD protocols such as the protocol by Bennett and Brassard [BB84] only work when the dimension of the Hilbert space of the communicating parties is known [AGM06, MMMO06]. Finally, in the cryptographic scenario of malicious adversaries, one needs to take into account the possibility that the manufacturer of the devices has intentionally tampered with the devices in order to eavesdrop on the communicated information. Eavesdropping on communication devices is not just a hypothetical situation: recently, a series of documents were published that showed that for the Central Intelligence Agency (CIA) of the U.S.A., large-scale hacking of mobile phones and smart TVs is standard practice [Wik]. Reasons such as these led researchers to consider *device-independence*, the paradigm in which properties of the devices are no longer assumed to operate as specified. Let us first dive into the main tool for executing protocols in a device-independent fashion: *Bell nonlocality*.

1.1 Bell nonlocality

In a Bell experiment, several parties are physically isolated so that classical communication between the parties is impossible[1], and each party is regarded as a black box which takes input and gives output. After a long line of research that started with the criticism on quantum physics with the famous paper of Einstein, Podolsky and Rosen [EPR35], the first to correctly realize the implications of Bell experiments for this discussion was John Bell [Bel64]. We briefly go into the topic of Bell nonlocality. For a recent review on Bell tests and applications, see the work of Brunner et al. [BCP+14].

As an example of a Bell experiment, we consider the bipartite scenario with two players, usually called Alice and Bob (see Figure 1). Alice and Bob each have a device which takes input x, y and gives output a, b, respectively. In a real-world experiment, the devices could for example have a series of knobs (one for each possible input setting), each of which corresponds to a measurement on some physical state inside the device. The experiment consists of multiple runs; in each run, Alice and Bob each select an input at random and record the outcome. After a large number of runs, Alice and Bob come together and compute their joint statistics: the conditional probability distribution $\Pr(a, b|x, y)$. We assume that the rounds are identical and independently distributed (the 'i.i.d. assumption'), i.e. the

[1] In theory, this can be established by spatially separating the parties over a long distance and obtaining the required data from the Bell experiment quickly; the non-signaling condition is then implied by the impossibility of sending signals which travel faster than light.

probability distribution underlying the observed statistics is well-defined.

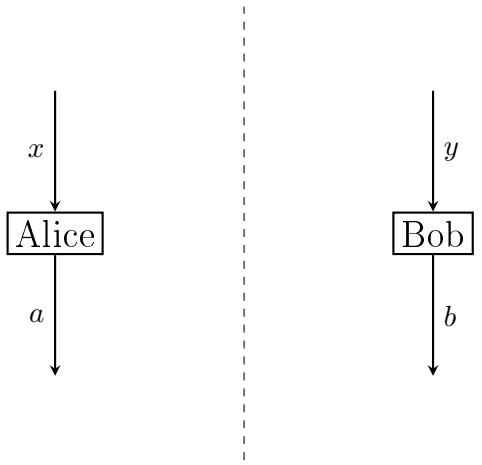

Figure 1: A Bell experiment with two players, Alice and Bob, who each have a device. In a single run of the Bell experiment, each of the two players' devices takes binary input, picked uniformly at random. Alice and Bob record the binary output of the device. After many runs, the two players come together and compute their joint probability distribution $\Pr(a, b|x, y)$. The two players are physically isolated, as indicated by the dashed vertical line.

If we assume that the devices are adequately described by classical physics, then the behaviour of the devices can be assumed to be *local*: that is, whatever operations the devices apply internally, the behaviour of a device can only be determined by anything that is spatially close to the device. In particular, since Alice and Bob are physically isolated, the behaviour of the two devices can only be correlated to each other if the devices were somehow connected *before* the experiment. In physics, such pre-established connections are referred to as *local variables*.

In the simplest case, the output of each device is a deterministic function of its input. For example, if Alice's device has two knobs, labeled 0 and 1, it could be that the device outputs the label of the knob that was pushed. In this case, we say that the behaviour of the device corresponds to a deterministic response function.

It has been shown [Fin82] that any local behaviour of the devices can be explained by choosing in each run a deterministic response function, where the choice of response function is made according to some probability distribution. However, if the knobs of the devices correspond to measurements on entangled quantum states, the devices' behaviour generally cannot be expressed by such a probabilistic mixture of deterministic functions. For this reason, Bell experiments have taken a prominent position in the debate on the validity of quantum mechanics. Throughout the past decades, many Bell experiments have been performed. It was only recently that the two main 'loopholes' in these experiments were closed simultaneously[2] [HBD+15, HKB+16, GVW+15, SMSC+15]. The experimentally observed correlations indeed could not be explained by local variables, which shows that, on the most fundamental level, the real world cannot be explained by any local-variable theory such as classical physics.

[2]First, there is the *locality loophole*, the prohibition of classical communication between the boxes. Second, there is the *detection loophole*: for any behaviour of the devices, Alice and Bob could change the observed statistics $\Pr(a, b|x, y)$ by not taking all runs of the Bell experiment into account in their calculations; this *post-selection* is therefore not allowed when dealing with the data obtained from Bell experiments.

In practice, it is more convenient to work with a single number that 'measures' the degree of nonlocality than with the entire conditional probability distribution $\Pr(a, b|x, y)$. Such a number is found in the *Bell value*, a linear combination of the probabilities $\Pr(a, b|x, y)$. In case the Bell value is greater than what can be achieved by distributions produced by local variable theories, we speak of a *Bell violation* of a *Bell inequality*.

Since entangled quantum states can yield a Bell violation while classical devices cannot, Alice and Bob can find out if their devices contain some kind of quantum state on the sole basis of the correlations. Although this fact is quite remarkable already, the connection between the behaviour of the devices and the observed statistics goes even further: if we assume quantum mechanics to be the underlying theory governing the behaviour of the devices, then certain statistics allow us to *uniquely* identify the quantum state and the measurements of the boxes (up to some well-understood equivalences). This was first explicitly shown by Mayers and Yao [MY98, MY04], but already implicit in earlier work [Cir80, PR92, SW88]. It is exactly this feature of Bell nonlocality that is useful in the context of cryptography.

1.2 From Bell experiments to device-independence and self-testing

The first ideas on the use of nonlocal correlations in order to test for a nosy eavesdropper in the context of key distribution were already implicit in the celebrated QKD scheme by Ekert [Eke91]. The first device-independent quantum key distribution protocol (DI-QKD) was developed by Barrett et al. [BHK05], who showed how two parties can generate a single random shared bit, secure against any post-quantum eavesdropper. Later work in DI-QKD mainly focused on developing similar protocols that generate more shared key bits with fewer device uses, and on proving their security [AMP06, MW06, AGM06, MPA11, VV14, DFR16, AFRV16]. For an overview of the field, we refer to two reviews [BCP$^+$14, ER14].

Related to DI-QKD is the field of device-independent randomness generation, where the central question is: how can we generate truly random bits with the use of potentially untrusted devices? This is a simpler task than device-independent key distribution, since one can always execute a DI-QKD protocol locally to generate a random key and subsequently use the key as a random bit string. The generation of randomness naturally points toward Bell inequalities, since it can be shown that Bell violation implies the presence of intrinsic randomness (for a comprehensive derivation, see the lecture notes by Scarani [Sca12]). The first work on device-independent randomness generation used Bell-violating devices [Col07]. Later work has mainly focused on finding protocols that are secure under relaxed assumptions and on proving their limitations [PAM$^+$10, CK11, SCA$^+$11, BPPP14, MS16]. For a more elaborate overview, we again refer to Brunner et al. [BCP$^+$14].

A third branch of the device-independent approach to quantum information theory is the field of *self-testing*: device-independent characterization of the quantum state and measurements. The central question in *self-testing of quantum states* is: given the observed correlations of several parties in a Bell experiment, what knowledge can be inferred about the quantum state that the parties possess? This knowledge is usually expressed as a self-testing statement: a bound on a measure that compares the optimal state and the actual state, where the bound is a function of a Bell value. A similar question can be asked for *self-testing of measurements*.

In self-testing, it is the connection between observed correlations on the one hand and the shared state and applied measurements on the other that is the object of research, rather than the use of

this connection for cryptographic purposes. Indeed, it is most remarkable that *classical* players (in our example: Alice and Bob) can determine which *quantum* state is shared by the devices, solely on the basis of correlations.

As already mentioned, Mayers and Yao were the first to explicitly note the usefulness of empirical correlations for practical applications [MY98, MY04]. In their work, the object under study was a maximally entangled state of two qubits, to which we will refer as the *singlet* in this thesis. It was already shown before for a particular Bell inequality, the CHSH inequality [CHSH69], that particular correlations are only reproduced by devices sharing a singlet (again, up to some well-understood equivalences) [PR92, BMR92]. By connecting the practical usefulness and the theoretical work, the singlet state was the first state to be shown to be self-testable. Since then, many other states and measurements have been shown to be 'identifiable' from certain Bell inequalities and observed correlations (where a precise meaning of 'identifiable' is given in Chapter 3) [McK14, MYS12, YN13, BP15, CGS16, McK16].

In practice, the observed statistics cannot be expected to yield the perfect correlations. The reason for this is twofold; first, since the correlations are estimated only from a finite numbers of runs of the Bell test, they are subject to statistical errors. The presence of external noise is another factor that contributes to the failure of Bell experiments to achieve perfect correlations in real-life experiments. It is for this reason that in order to relate observed correlations to the states and measurements that the devices could possibly share, we need self-testing statements that are *robust to errors*.

1.3 Our contributions

In this thesis, we improve the previously known robustness of self-testing statements for almost all pure partially entangled two-qubit states. Also, we prove that a violation of the CHSH inequality by a particular quantum state does not imply that a singlet can be 'extracted' from that state (see below).

Pure two-qubit states can be self-tested with a family of Bell inequalities which go by the name of *tilted CHSH inequalities*. That is, for every pure two-qubit state, there exists a tilted CHSH inequality for which particular correlations $\Pr(a, b|x, y)$ can only be obtained if the players share that particular state, up to some equivalences. Self-testing statements for pure two-qubit states are expressed in terms of bounds on the fidelity of the ideal two-qubit state and the actual state after Alice and Bob have applied quantum operations locally: the maximal fidelity that can be obtained in this way is referred to as the *extractability*. By generalizing the approach of Kaniewski [Kan16], we improve upon all previously known self-testing bounds for the extractability [YN13, BP15, BNS$^+$15].

In the method we used, the quantum operations that Alice and Bob apply locally are called *extraction channels*, since the two players attempt to 'extract' the optimal state from the state they actually possess. By constructing local extraction channels for each tilted CHSH inequality, we reduce the problem of proving a self-testing statement to proving a particular operator inequality. We numerically verified this operator inequality for (almost) every tilted CHSH inequality, which results in new improved self-testing statements for (almost) all pure two-qubit states.

Our second result is the construction of a state with the following two properties: (a) the state violates the CHSH inequality; (b) there exist no local extraction channels that achieve a fidelity with the maximally entangled two-qubit state that is strictly greater than the trivial lower bound. Here, 'trivial' refers to the fidelity which can be achieved for *any* state. In this sense, the singlet extractability of our constructed state is trivial.

A result of the same flavor was already known for *bound entangled states*. Entanglement distillation refers to the process of creating a maximally entangled state out of several copies of a less entangled state by LOCC (local operations and classical communication) [BBPS96]. Bound entangled states are entangled but undistillable (i.e. it is not possible to distill entanglement from them). Since classical communication and the use of several copies is not included in the definition of 'extractability', distillability is a stronger notion than extractability. Vértesi and Brunner constructed a bound entangled state which violates a Bell inequality that is different from the CHSH one [VB14]; in fact, their state does not violate the CHSH inequality. The state ρ_{TE} we constructed, however, does violate the CHSH inequality but it is not possible to extract even a "fraction" of a maximally violating state (the singlet) from ρ_{TE} using local operations.

This result is remarkable: although one needs entanglement for CHSH violation and maximal violation implies that the shared state is equivalent to a singlet, it is in general not possible to create more overlap with a maximally entangled two-qubit state from *any* entangled state using local operations only.

As a corollary to our result, there exist no self-testing statements, formulated in terms of the extractability, that yield nontrivial information about the actual state whenever the CHSH violation is very small.

The state we constructed is a classical-quantum state and can be written as a probabilistic mixture of, on each side, a classical three-outcome register and a qubit. Intuitively, one has to apply very different local extraction maps in order to maximize the singlet fidelity with a product state or a maximally entangled state. By tweaking the probabilistic weights of our state, we obtain a state which violates the CHSH inequality but still has trivial singlet extractability.

1.4 Organisation of the thesis

Chapter 2 gives a brief overview of the theory behind Bell operators and defines some general notions of quantum information theory. Also, the CHSH scenario will be explained in more detail. A large part of the chapter is devoted to auxiliary lemmas that are needed for our construction of the state that violates the CHSH inequality but nevertheless has trivial singlet extractability. Chapter 3 outlines previous work on robust self-testing of quantum states, with particular emphasis of the formalization of the self-testing problem, since this is a nontrivial problem on its own. Our main results are contained in Chapter 4 and 5. In Chapter 4, we construct a state that violates the CHSH inequality and has trivial singlet extractability. Chapter 5 contains the derivation of improved robustness for self-testings statements for (almost) all pure two-qubit states using the tilted CHSH scenario. Opportunities for future research are given in Chapter 6.

2 Preliminaries

In this section, we first review very briefly some notions from the quantum information formalism that are relevant to this thesis (Section 2.1). Then, in Section 2.2, we give an overview of Bell inequalities. In the last part of this chapter, Section 2.3, several lemmas are proven that are needed for our main result in Chapter 4.

For a more complete review of quantum information theory, we refer to the book by Nielsen and Chuang [NC00] and to John Preskill's lecture notes [Pre15]. A mathematically more rigorous introduction to quantum information theory can be found in the lecture notes by John Watrous [Wat16].

2.1 Quantum states and quantum operations

We review several notions from quantum information theory that are relevant to this thesis. All Hilbert spaces in this thesis are of finite dimension, and all operators are linear operators acting on Hilbert spaces of finite dimension.

2.1.1 Some properties and operations on matrices

A Hilbert space, a vector space over the complex numbers, will be denoted by $\mathbb{H}$. We will denote the standard Euclidean inner product on a Hilbert space $\mathbb{H}$ by $\langle \vec{v}|\vec{w}\rangle := \sum_i \overline{v_i} w_i$ with $\vec{v}, \vec{w} \in \mathbb{H}$. Any vector $\vec{v} \in \mathbb{H}$ has norm $||\vec{v}|| := \sqrt{\langle \vec{v}|\vec{v}\rangle}$.

The notation X and Y will denote linear operators (matrices) that act on vectors from a Hilbert space. A linear operator $X : \mathbb{H}_1 \to \mathbb{H}_2$ is said to have input dimension $\dim(\mathbb{H}_1)$ and output dimension $\dim(\mathbb{H}_2)$.

We start with some notation. The complex conjugate of a complex number z is denoted by $\overline{z}$. The transpose of the representation of an operator X in a particular basis is written as X^T and its conjugate transpose, defined as its transpose where every entry is replaced by its complex conjugate, is written as $X^\dagger$. If X satisfies $XX^\dagger = X^\dagger X$, then X is called normal. The tensor product of X and Y will be written as $X \otimes Y$. The identity matrix of an n-dimensional vector space will be written as $\mathbb{1}_n$, or simply $\mathbb{1}$ when the dimension of the space that it acts on is clear from the context.

A unitary matrix is a matrix U that satisfies $UU^\dagger = U^\dagger U = \mathbb{1}$. Equivalently, a unitary operator is a bijection that preserves inner products. Since a basis of a Hilbert space consists of pairwise orthonormal vectors, we infer that unitaries correspond to changes of basis. We refer to $UXU^\dagger$ as the unitary conjugation of a matrix X with U.

If $X = X^\dagger$, then we call X hermitian. Any hermitian matrix can be decomposed as $X = \sum_j \lambda_j |j\rangle\langle j|$ for real eigenvalues λ_j and orthogonal eigenstates $\{|j\rangle\}_j$. This decomposition is the 'eigendecomposition of X'. A function f on hermitian operators is defined as $f(X) = \sum_j f(\lambda_j) |j\rangle\langle j|$. For example, for hermitian matrices with nonnegative eigenvalues, the square root of X is given by $\sqrt{X} = \sum_j \sqrt{\lambda_j} |j\rangle\langle j|$. The modulus of a normal matrix X is defined as $|X| := \sqrt{X^\dagger X}$.

The set of 2×2 matrices is spanned by $\mathbb{1}_2$ together with the Pauli matrices, defined as

$$\mathsf{X} := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathsf{Y} := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathsf{Z} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where i is the imaginary unit.

For convenience, we write

- $\text{Herm}(\mathbb{H})$ for the set of hermitian operators $\mathbb{H} \to \mathbb{H}$;
- $\text{Pos}(\mathbb{H}) \subseteq \text{Herm}(\mathbb{H})$ for the set of hermitian operators $\mathbb{H} \to \mathbb{H}$ with nonnegative eigenvalues. Such operators are called positive semidefinite. Hermitian operators with strictly positive eigenvalues are called positive definite.

Let $\mathbb{H}$ be a Hilbert space of dimension d. The trace of a matrix $X \in \text{Herm}(\mathbb{H})$ is the sum of its eigenvalues and can be computed as $\text{Tr}(X) := \sum_{j=1}^{d} \langle j| X |j\rangle$ where $\{|j\rangle\}_{j=1}^{d}$ is an arbitrary basis of $\mathbb{H}$. The determinant of X, which is the product of its eigenvalues, is denoted by $\det(X)$. If a matrix X acts on composite system $\mathbb{H}_A \otimes \mathbb{H}_B$, then we denote its partial trace, where 'system A is traced out', as $\text{Tr}_A(X) := \sum_{k=1}^{d_A} (\langle k| \otimes \mathbb{1}_{\mathbb{H}_B}) X (|k\rangle \otimes \mathbb{1}_{\mathbb{H}_B})$, where d_A is the dimension of $\mathbb{H}_A$ and $\{|k\rangle\}_{k=1}^{d_A}$ is an arbitrary basis of $\mathbb{H}_A$.

Let $X, Y \in \text{Herm}(\mathbb{H})$. We say that the operator inequality $X \geq Y$ holds if $X - Y$ is positive semidefinite. Similarly, $X > Y$ holds when $X - Y$ is positive definite.

For any real number $p \geq 1$, the Schatten p-norm of $X : \mathbb{H}_1 \to \mathbb{H}_2$ is defined as

$$||X||_p := \left(\text{Tr}\left((X^\dagger X)^{\frac{p}{2}} \right) \right)^{\frac{1}{p}}.$$

We mention three commonly used special cases of the Schatten-p-norms:

- The case $p = 1$ corresponds to the *trace norm*: $||X||_{\text{tr}} := ||X||_1 = \text{Tr}\left(\sqrt{X^\dagger X}\right)$.
- The *Frobenius norm* is the special case for $p = 2$: $||X||_2 = \sqrt{\text{Tr}\left(X^\dagger X\right)}$.
- The limit of $p \to \infty$ yields the *spectral norm*:

$$||X||_\infty = \max\{||Xu|| \mid u \in \mathbb{H}_1, ||u|| = 1\}.$$

2.1.2 Quantum states

In quantum mechanics, a physical system is represented as a complex Hilbert space. The state of a physical system is represented by a vector of unit length in that Hilbert space and is denoted in Dirac's ket notation, e.g. $|\varphi\rangle$.

The state of a qubit is a vector in a complex Hilbert space of dimension 2; in general, the state of a qudit of dimension d is given by a vector in a complex Hilbert space of dimension d. The vector $(1, 0)$ is abbreviated as $|0\rangle$ and we write $(0, 1)$ as $|1\rangle$. We also define $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$.

A composite system made up out of m components, each represented by complex Hilbert space $\mathbb{H}_j$ for $1 \leq j \leq m$, is represented by the complex Hilbert space $\bigotimes_{j=1}^{m} \mathbb{H}_j$. If the state of component j is given by $|\varphi_j\rangle$, then the state of the composite system is $\bigotimes_{j=1}^{m} |\varphi_j\rangle \in \bigotimes_{j=1}^{m} \mathbb{H}_j$.

In reality, the precise state of a quantum system is usually not known. To deal with such partial knowledge, density matrices are a helpful tool. Let $\mathbb{H}$ be a Hilbert space that represents a certain quantum system. A density matrix is defined as $\rho \equiv \sum_{k=1}^{n} p_k |\varphi_k\rangle\langle\varphi_k|$, where $n \in \mathbb{N}_{\geq 1}$ is a strictly positive integer, p is an n-dimensional probability vector and the $|\varphi_k\rangle \in \mathbb{H}$ for all $1 \leq k \leq n$. This definition is equivalent to requiring that $\rho \geq 0$ while $\text{Tr}(\rho) = 1$. We denote the set of density matrices

by $\mathrm{D}(\mathbb{H}) := \{\rho \in \mathrm{Pos}(\mathbb{H}) \mid \mathrm{Tr}(\rho) = 1\}$. From now on, density matrices will be referred to as 'quantum states' or simply 'states'. States that have an eigenvalue equal to 1 are referred to as 'pure states', otherwise we call them 'mixed'. The density matrix that describes the state of a composite system of m components, where the j-th component is represented by the Hilbert space $\mathbb{H}_j$ for $1 \leq j \leq m$, is an element of $\mathrm{D}\big(\bigotimes_{j=1}^{m} \mathbb{H}_j\big)$.

For two linear operators $A : \mathbb{H}_1 \to \mathbb{H}_2$ and $B : \mathbb{H}_1 \to \mathbb{H}_2$, the Hilbert-Schmidt inner product is defined as

$$\langle A, B \rangle = \mathrm{Tr}\Big(A^\dagger B\Big) \tag{2.1}$$

For $\rho, \sigma \in \mathrm{D}(\mathbb{H})$, this inner product is the trace inner product $\langle \rho, \sigma \rangle := \mathrm{Tr}(\rho\sigma)$. Note that if $\rho = |\varphi\rangle\langle\varphi|$ and $\sigma = |\psi\rangle\langle\psi|$ are pure states, then $\mathrm{Tr}(\rho\sigma) = |\langle\varphi|\psi\rangle|^2$.

Other than the trace inner product, there are different measures of closeness on $\mathrm{D}(\mathbb{H})$, such as the 'fidelity'.

Definition 2.1. *The fidelity of two quantum states ρ, σ is given by $F(\rho, \sigma) = ||\sqrt{\rho}\sqrt{\sigma}||_{\mathrm{tr}}^2$.*

We state a few properties of the fidelity that will be useful in the formalization of the self-testing problem, as given in Chapter 3.

Lemma 2.2. *Let ρ, σ be density matrices. The fidelity $F(\rho, \sigma)$ has the following properties:*

1. *Symmetry: $F(\rho, \sigma) = F(\sigma, \rho)$;*
2. *$0 \leq F(\rho, \sigma) \leq 1$, and $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$;*
3. *If at least one of ρ, σ is pure, then $F(\rho, \sigma) = \mathrm{Tr}(\rho\sigma) = \langle \rho, \sigma \rangle$.*

For the proof of this lemma, we refer to John Watrous' lecture notes, proof of Proposition 3.12 [Wat16].

2.1.3 Entanglement

A product state is a state of the form $\rho \otimes \sigma$. A separable state is a state that can be written as

$$\sum_{j=1}^{n} p_j \ \rho_j \otimes \sigma_j$$

with p an n-dimensional probability vector. Any state that cannot be written as a separable state is called entangled. For Hilbert space dimension d, the 'maximally mixed state' is $\frac{\mathbb{1}_d}{d}$. A bipartite pure state $\rho_{AB} \in \mathrm{D}(\mathbb{H}_1 \otimes \mathbb{H}_2)$ is called 'maximally entangled' if $\mathrm{Tr}_A(\rho_{AB})$ and $\mathrm{Tr}_B(\rho_{AB})$ are both maximally mixed states.

In this thesis, we refer to a maximally-entangled state of two qubits as 'a singlet'. The particular maximally-entangled state $\frac{|01\rangle - |10\rangle}{\sqrt{2}}$ will be denoted by $|\phi\rangle$.

In order to characterize all maximally entangled two-qubit states, we need the following lemma.

Lemma 2.3. (Schmidt decomposition)
Let $|\varphi\rangle$ be a pure state, shared between two players A and B. Then there exists a set of states $|j_A\rangle$,

orthonormal in the Hilbert space of player A, and a set of states $|j_B\rangle$, orthonormal in the Hilbert space of player B, such that

$$|\varphi\rangle = \sum_j \lambda_j |j_A\rangle \otimes |j_B\rangle$$

where the $\lambda_j \in \mathbb{R}_{\geq 0}$ satisfy $\sum_j \lambda_j^2 = 1$. The λ_j are referred to as the Schmidt coefficients of $|\varphi\rangle$.

Furthermore, the Schmidt coefficients are unique. If the Schmidt coefficient λ_j is nondegenerate for some j, then the states $|j_A\rangle$ and $|j_B\rangle$ are also unique.

For a concise proof of the lemma, we refer to the book by Nielsen and Chuang, Theorem 2.7 [NC00]. The argument for the uniqueness in given in Preskill's lecture notes, Section 2.4 [Pre15].

It is not hard to compute that every maximally entangled two-qubit state has two equal Schmidt coefficients $\frac{1}{\sqrt{2}}$. Using the Schmidt decomposition, we obtain a parametrization of such states.

Corollary 2.4. *All maximally entangled two-qubit states can be written as*

$$\frac{|a_0\rangle \otimes |b_0\rangle + |a_1\rangle \otimes |b_1\rangle}{\sqrt{2}}$$

where $\{|a_0\rangle, |a_1\rangle\}$ and $\{|b_0\rangle, |b_1\rangle\}$ are orthonormal bases for the single-qubit Hilbert space.

Since unitary matrices correspond to changes of bases, we see that we can obtain any maximally entangled two-qubit state from another by application of local unitaries.

2.1.4 Measurements and observables

By a POVM (Positive Operator-Value Measure) measurement on a quantum system, represented by Hilbert space $\mathbb{H}$, with possible outcomes $m \in M$ for M some finite set, we mean a set of operators $\{E_m\}_{m \in M}$ that satisfy $\sum_{m \in M} E_m = \mathbb{1}_{\mathbb{H}}$, where each E_m is a positive linear map $\mathbb{H} \to \mathbb{H}$. If the state of the system immediately before measurement is ρ, then the probability of measuring m is given by $p(m) = \mathrm{Tr}(E_m \rho)$.

POVM measurements are the most general type of measurements allowed in quantum mechanics. We will thus refer to a POVM measurement as simply 'measurement'.

A measurement with two outcomes, with measurement operators E_{+1} and E_{-1}, can be described by giving the *observable* $Q = E_{+1} - E_{-1}$. Note that given Q, we can compute $E_{\pm 1} = (\mathbb{1} \pm Q)/2$; hence the observable Q contains all the information needed to retrieve the measurement operators E_{+1} and E_{-1}.

Projective measurements form a special class of measurements, where the set of operators are projectors: all E_m satisfy $E_m^2 = E_m$. Any measurement can be made into a projective measurement by embedding the state that is measured into a larger Hilbert space.

Let $\alpha \in [0, 2\pi)$ and let $|\varphi\rangle \in \mathbb{H}$. Since the pure density matrices of $|\varphi\rangle$ and $e^{i\alpha}|\varphi\rangle$ are the same, multiplication with the *global complex phase* $e^{i\alpha}$ does not change the measurement statistics.

2.1.5 Bloch ball

The space of matrices of size 2×2 is spanned by $\mathbb{1}_2, \mathsf{X}, \mathsf{Y}$ and Z. Since density matrices have unit trace, any single-qubit density matrix ρ can be written as

$$\rho = \frac{1}{2}(\mathbb{1}_2 + x\mathsf{X} + y\mathsf{Y} + z\mathsf{Z})$$

with $x, y, z \in \mathbb{R}$ since ρ is hermitian. It is a straightforward exercise to prove that the eigenvalues of ρ are nonnegative if and only if $x^2 + y^2 + z^2 \leq 1$. Thus we can represent any single-qubit density matrix as an element (x, y, z) of a three-dimensional unit ball. This visualization is commonly referred to as the Bloch ball.

2.1.6 Unitaries and isometries

We already mentioned unitary operators, which are bijective inner-product preserving operators. More generally, an isometry is an inner-product preserving map which is not necessarily bijective. An isometry V acting on $\mathbb{H}$ satisfies $V^\dagger V = \mathbb{1}_{\dim(\mathbb{H})}$ since $\langle\varphi|\psi\rangle = \langle\varphi| V^\dagger V |\psi\rangle$ for all $|\varphi\rangle, |\psi\rangle \in \mathbb{H}$. If V is an isometry acting on $\mathbb{H}$, then the map $\rho \to V\rho V^\dagger$ is an isometry on $\mathrm{D}(\mathbb{H})$. Any unitary is an isometry; an example of an isometry that is not a unitary is e.g. adding an ancilla qubit in a pure state.

2.1.7 Quantum channels

The most general evolution of quantum states is described by quantum channels. By definition, a quantum channel $\Lambda : \mathrm{D}(\mathbb{H}_1) \to \mathrm{D}(\mathbb{H}_2)$ is a completely-positive trace-preserving map. There are several ways to argue that this definition captures all physically possible transformations that a quantum state could be subject to. Three of such derivations are given in the book of Nielsen and Chuang [NC00]; we present two of the resulting definitions here.

The first equivalent definition for Λ to be a quantum channel is as follows.

Lemma 2.5. *The map $\Lambda : \mathrm{D}(\mathbb{H}_1) \to \mathrm{D}(\mathbb{H}_2)$ is a quantum channel precisely if there exist operators $\Gamma_k : \mathbb{H}_1 \to \mathbb{H}_2$ for $1 \leq k \leq n$ for some number $n \in \mathbb{N}_{\geq 1}$ such that $\sum_{k=1}^n \big(\Gamma_k\big)^\dagger \Gamma_k = \mathbb{1}_{\mathbb{H}_1}$ and*

$$\Lambda(\rho) = \sum_{k=1}^m \Gamma_k \rho \big(\Gamma_k\big)^\dagger. \tag{2.2}$$

The operators Γ_k are denoted as the 'Kraus operators' of Λ. In general, the set of Kraus operators of a quantum channel is not uniquely defined.

The second definition has to do with another decomposition of quantum channels.

Lemma 2.6. *The map $\Lambda : \mathrm{D}(\mathbb{H}_1) \to \mathrm{D}(\mathbb{H}_2)$ is a quantum channel if and only if there exists a pure state $\sigma \in \mathrm{D}(\mathbb{H}_3)$ and a unitary U such that for every $\rho \in \mathrm{D}(\mathbb{H}_1)$, we have*

$$\Lambda(\rho) = \mathrm{Tr}_S \left(U(\rho \otimes \sigma)U^\dagger\right) \tag{2.3}$$

where S is a subsystem of $\mathbb{H}_1 \otimes \mathbb{H}_3$.

For any quantum channel $\Lambda : \mathrm{D}(\mathbb{H}_1) \to \mathrm{D}(\mathbb{H}_2)$, there exists a unique map $\Lambda^\dagger : \mathrm{D}(\mathbb{H}_2) \to \mathrm{D}(\mathbb{H}_1)$ satisfying $\langle\Lambda(\rho), \sigma\rangle = \langle\rho, \Lambda^\dagger(\sigma)\rangle$, where $\rho \in \mathrm{D}(\mathbb{H}_1)$ and $\sigma \in \mathrm{D}(\mathbb{H}_2)$. The map $\Lambda^\dagger$ is called the *dual*

channel of Λ. It is not hard to see that when Λ is decomposed as in Equation (2.2), then $\Lambda^\dagger$ can be written as

$$\Lambda^\dagger(\rho) = \sum_{k=1}^{m} (\Gamma_k)^\dagger \rho \Gamma_k.$$

In general, the dual of a quantum channel is not necessarily trace-preserving and thus need not be a quantum channel itself.

2.1.8 Dephasing channels and amplitude-damping channels

Two particular classes of qubit-to-qubit channels will be frequently used throughout the thesis: dephasing channels and amplitude-damping channels. We state the general form of these channels here.

A dephasing channel or phase-damping channel contracts the Bloch ball to an ellipsoid. Formally, we define the Kraus operators $E_0 := (1 - \frac{p}{2})\mathbb{1}_2$ and $E_1 := \frac{p}{2}U$, where U is some unitary and $0 \leq p \leq 1$ is the dephasing parameter. If we choose $U = \mathsf{X}$, for example, then the dephasing channel contracts the Bloch ball to an ellipsoid with as major axis the line through $\frac{\mathbb{1}}{2}$ and X. Setting $p = 1$ corresponds to full dephasing.

Let $|\varphi\rangle$ be a single-qubit state and let $|\psi\rangle$ be a state orthogonal to it. The *qubit-to-qubit amplitude-damping channel that damps towards* $|\varphi\rangle$ has Kraus operators $|\varphi\rangle\langle\varphi| + \sqrt{1-\gamma}\,|\psi\rangle\langle\psi|$ and $\sqrt{\gamma}\,|\varphi\rangle\langle\psi|$ where $0 \leq \gamma \leq 1$ is the damping parameter. By setting $\gamma = 0$, the amplitude-damping channel reduces to identity. Setting $\gamma = 1$ yields full damping: when applied to the Bloch ball, the entire Bloch ball is contracted to the pure state $|\varphi\rangle\langle\varphi|$, represented as a point on the boundary of the ball.

2.2 From Bell inequalities to self-testing

In this section, we review the mathematics behind Bell nonlocality as briefly introduced in Chapter 1. For a thorough review of Bell nonlocality, we refer to Brunner et al. [BCP+14]. For a deeper understanding of the connection between Bell nonlocality on the one hand and device-independence and self-testing on the other, see the lecture notes by Scarani [Sca12].

In a famous paper of Einstein, Podolsky and Rosen [EPR35], the authors describe a paradox that leads them to the conclusion that quantum mechanics cannot be a complete theory and that "hidden variables" must be added to the theory in order to predict all properties of a physical system with certainty. Einstein, Podolsky and Rosen describe two entangled systems[3] The authors reason that, depending on which measurement on the first of the two systems we apply, there are different wave functions that describe the state of the second system. Hence there are physical properties of the second system that we cannot predict with certainty. This observation leads the authors, who believe that two different properties cannot be realized in a same system simultaneously, to the conclusion that quantum mechanics is incomplete.

Bohm reformulated the thought experiment of Einstein, Podolsky and Rosen with two atoms, whose spin is the physical property under consideration [Boh51, p.614]. It is this formulation that led John Bell to prove [Bel64] that the quantum mechanical predictions of an entangled system of two spin-$\frac{1}{2}$ particles cannot be reproduced using any theory that assumes that the physical properties of the atoms are only influenced by anything spatially close; such a theory is called *local*. Alternatively, we say that the predictions of local theories can be explained by *local variables*. We usually refer to Bell's result

[3] It is this work from which the name *EPR-pair* to refer to a maximally entangled state of two qubits comes from.

that no local theory yields the same statistical predictions as quantum mechanics, as *Bell's theorem*. In the next section, we describe what it means for a theory to be local and outline the general framework of Bell inequalities, which can be used to experimentally falsify any local theory of physics.

2.2.1 Bell games and local variables

We give a review of Bell experiments or Bell games. The basic setup is the same as described in the introduction, Chapter 1, although in this case the questions to the players in the game are generated randomly by a third party, a referee, rather than by the players themselves. This setup is suited to the cryptographic setting where a customer (the referee) bought a source that produces entangled particles and attempts to verify this claim by having two friends (Alice and Bob) play a Bell game. In the setting as described in this section, the player-device pair is what is considered as a black box, rather than the device on its own.

In a Bell game, we have several players who receive input from a referee and return classical output. For simplicity, we describe the two-player setup here, with players Alice and Bob; the generalization to an arbitrary number of players is straightforward. Alice (Bob) receives input x (y) and returns output a (b). For both players, the referee chooses the input uniformly at random from a finite set; the sizes of Alice's input set and Bob's input set need not be equal. The same holds for their output sets. The players are physically isolated to prevent classical signaling between them; this could for example be implemented by separating them in space and requiring them to answer so quickly to the referee, that they would need superluminal signaling in order to receive information from the other players.

After playing several rounds of the game, the referee computes the joint statistics $\Pr(a, b|x, y)$. Since in real-life, the number of rounds is finite, the referee can only estimate the actual correlation using standard statistical methods; but in theory, by playing sufficiently many rounds, the real correlations can be approximated with arbitrary precision.

The goal of the referee is to describe the possible behaviors of the players using only the observed statistics $\Pr(a, b|x, y)$. The framework for obtaining information from the observed statistics is structured as follows. Consider the set of all explanations that we consider possible for the behaviour of Alice and Bob. Suppose that to each pair of such possible behaviors, we associate some number λ. At each run of the Bell game, a pair of behaviors λ is chosen according to some probability distribution $\rho(\lambda)$. The probability distribution that the referee in a Bell game obtains can now be written as

$$\Pr(a, b|x, y) = \int d\lambda \ \rho(\lambda|x, y) \ \Pr(a, b|x, y, \lambda). \tag{2.4}$$

For example, suppose that the input and output of all the players are single bits, and that moreover the output of an individual player can be computed by some deterministic function of the player's input. In that case, each λ corresponds to a pair of deterministic functions $\{0, 1\} \to \{0, 1\}$; there are 16 pairs in total.

Now let us impose some conditions on $\Pr(a, b|x, y)$ for the players' behaviour to be explainable by local variables. First, the practice of science is based upon the possibility to generate copies of a system, to which we then perform different measurements. This corresponds to the fact that the choice of λ in each run does not depend on the inputs x and y:

$$\rho(\lambda|x, y) = \rho(\lambda). \tag{2.5}$$

This requirement is called *measurement independence*.

Secondly, after fixing λ, the behaviour of each player cannot depend on the other's input since Alice and Bob are physically isolated. Formally, this becomes

$$\Pr(a, b|x, y, \lambda) = \Pr(a|x, \lambda)\Pr(b|y, \lambda) \tag{2.6}$$

Combining Equations (2.4), (2.5) and (2.6) yields the following description for a probability distribution to be explainable by a local theory:

$$\Pr_{\text{local}}(a, b|x, y) = \int d\lambda\ \rho(\lambda)\Pr(a|x, \lambda)\Pr(b|y, \lambda). \tag{2.7}$$

It was shown by Fine [Fin82] that a probability distribution $\Pr(a, b|x, y)$ can be explained by local variables if and only if it can be explained by *deterministic* local variables; that is, when a players' output is a deterministic function of λ and his/her input only. Note that even for local variables, the output of the parties may still be obtained via a stochastic process: the element of chance is then absorbed into the choice of λ in each run.

2.2.2 The local polytope and Bell inequalities

Let us fix the sizes of the input set and output set of each of the players. Using the formalization above, it is not hard to show the family of probability distributions of the form $\Pr(a, b|x, y)$ that are allowed by local variables is a convex set. Indeed, suppose that $\mathcal{P}_1$ and $\mathcal{P}_2$ are such probability distributions which can both be explained by a local theory; then $q\mathcal{P}_1 + (1 - q)\mathcal{P}_2$ for some $q \in [0, 1]$ is explained by assuming that before every round of the Bell game, a coin with bias q is flipped, and, depending on the outcome, a set of possible behaviors of the players of $\mathcal{P}_1$ or of $\mathcal{P}_2$ is picked.

The extremal points are given by the possible combinations of deterministic strategies that the players could apply. Because each number of possible inputs and outputs is finite (by definition of a Bell game), the number of extremal points is finite too. Hence the convex set is a polytope: this is the *local polytope* for given input and output set sizes.

A *facet* is one of the hyperplanes that delimits the local polytope. A hyperplane can be expressed as

$$\{\vec{n} \cdot \vec{\mathcal{P}} = c \mid \vec{\mathcal{P}}\} \tag{2.8}$$

where $\vec{n}$ is a normal vector orthogonal to the hyperplane and oriented to the outside of the polytope, c is some real constant and for a particular probability distribution $\Pr(a, b|x, y)$, the vector $\vec{\mathcal{P}}$ represents $\{\Pr(a, b|x, y) \mid a, b, x, y\}$. A facet of the local polytope is a hyperplane of dimension $d - 1$ (where d is the dimension of the local polytope) that satisfies the following two conditions. First, the intersection of the local polytope with the facet must be nonempty. Moreover, since the polytope lies on one side of the facet, all elements $\vec{\mathcal{P}}$ of the local polytope satisfy

$$\vec{n} \cdot \vec{\mathcal{P}} \leq c. \tag{2.9}$$

Some facets correspond to trivial constraints, such as the requirement that probabilities sum up to one. Other facets impose non-trivial constraints on the probability distribution, however: there exist quantum states Alice and Bob can share such that Equation (2.9) is violated. If a facet is not trivial, we refer to Equation (2.9) as its corresponding *Bell inequality*. The maximal constant c for which Equation (2.9) is satisfied for a quantum state and measurements yielding probability distribution $\vec{\mathcal{P}}$, is called the *Bell value* or *classical value* of the Bell inequality. Since there exist games for which these facets are violated by players who share particular quantum states, the *quantum set*, the set of probability

distributions that can be produced by quantum realizations, is strictly larger than the local polytope. If we drop the condition that Alice and Bob are restricted by quantum realizations, the resulting family of probability distributions is even larger than the quantum set; this is the *nonsignaling set*.

It is straightforward to compute that measurements on separable states yield a probability distribution that can be explained by local variables. By taking the contraposition of this statement, we see that entanglement is necessary for Bell violation.

We now zoom in on a particular Bell experiment, the *CHSH game* [CHSH69].

2.2.3 The CHSH game and self-testing of the singlet

In a single round of the CHSH game, two physically isolated players, Alice and Bob, are given single-bit questions x and y by a third party, the referee. The questions are picked uniformly at random. Alice and Bob share a bipartite state ρ_{AB} that they may use. After receiving their questions, they perform their local measurements and return single-bit answers, a and b, respectively. Alice and Bob are said to win the game if $a + b \pmod 2 \equiv x \wedge y$ and lose otherwise.

Denote the set of measurement operators that Alice applies to her part of the state by $\{P_0^x, P_1^x\}$. Similarly, Bob's measurement operators can be denoted by $\{Q_0^y, Q_1^y\}$. According to the quantum formalism, the conditional probabilities are

$$\Pr(a, b|x, y) = \mathrm{Tr}\big[(P_a^x \otimes Q_b^y)\rho_{AB}\big] \tag{2.10}$$

Now write

$$A_x := P_0^x - P_1^x \quad , \quad B_y := Q_0^y - Q_1^y$$

One could think of A_x and B_y as observables, but this is not needed for what follows; one could also think of these as definitions that are mathematically convenient.

By summing up the expressions as in Equation (2.10) for $\Pr(a, b|x, y)$ that refer to the winning conditions of the CHSH game, one can show that the probability that Alice and Bob win the game equals

$$\Pr(\text{win}) = \frac{1}{2} + \frac{\beta}{8}$$

where $\beta := \mathrm{Tr}[W\rho_{AB}]$ is the *Bell value* and W is the *CHSH operator*:

$$W := \sum_{x,y=0}^{1} (-1)^{x \cdot y} A_x \otimes B_y. \tag{2.11}$$

The referee collects statistics by letting Alice and Bob play multiple rounds. During each round they are physically isolated, but they are allowed to refresh their state and to communicate in between the rounds. This could for example be implemented by letting Alice and Bob set up their state through a quantum channel that is controlled by the referee.

It is not hard to show that if ρ_{AB} is classical or separable, then $\beta \leq 2$. For quantum states, it has been shown that the Bell value β cannot exceed $2\sqrt{2}$: this is known as Cirel'son's bound [Cir80]. This upper bound is achieved exactly when Alice and Bob share the singlet state $|\phi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$, up to local unitaries and auxiliary degrees of freedom [PR92, BMR92]. That is, there exist local measurement

operators such that the state ρ violates the CHSH inequality maximally, if and only if there exist local unitaries U_A, U_B and a state σ such that

$$\rho = (U_A \otimes U_B)(|\phi\rangle\langle\phi| \otimes \sigma)(U_A^\dagger \otimes U_B^\dagger) \tag{2.12}$$

where $|\phi\rangle$ is the singlet state. We will refer to a bipartite state ρ_{AB} as an *ideal state* if and only if there exist local measurement operators such that the CHSH inequality is violated maximally. The latter condition is the case precisely if ρ_{AB} can be written in the form of Equation (2.12).

Here, it is important to mention that the 'i.i.d assumption'. To be precise: in each round, a random variable underlies the observed outcomes given the incomes. These random variables are identical and independently distributed over the different rounds. With this assumption, we can estimate the actual probabilities with arbitrary precision by playing sufficiently many rounds. Without the i.i.d. assumption, Alice and Bob could prepare different states and apply different measurements during the rounds; thus the probabilities need not converge. In particular, without this assumption, Alice and Bob could be classical players and achieve CHSH violation in the limit of infinitely many rounds with nonzero probability. The i.i.d. assumption makes our model clear and reasonable.

It is crucial here that we do not allow for classical communication between Alice and Bob; doing so would enlarge the set of ideal states. To see this, consider the state $(|00\rangle + |11\rangle + |22\rangle)/\sqrt{3}$. It is not hard to see that this state does not violate CHSH maximally since it cannot be written in the form $(U_A \otimes U_B)(|\phi\rangle\langle\phi| \otimes \sigma)(U_A^\dagger \otimes U_B^\dagger)$, where U_A, U_B are unitaries, $|\phi\rangle$ is the singlet and σ an arbitrary state. However, such a transformation is possible under local operations with classical communication (see Theorem (1) in [Nie99]).

Consider the cryptographic setting where a customer bought two quantum-cryptographic devices from a vendor, who claims that his devices produce singlets, and that each device receives a qubit from each maximally-entangled pair when we connect the two devices with a quantum channel. Then the CHSH game can be used to verify this claim: the customer asks two friends, Alice and Bob, to come help testing the devices. The testing setup is evident: Alice and Bob connect the devices to let them share a singlet, then physically isolate themselves from each other and subsequently play a single round of the CHSH game, where the customer acts as a referee. They repeat the process to collect statistics. If the resulting Bell violation is $2\sqrt{2}$, the claim that the devices share the singlet state has been verified. Using the CHSH game, the customer can thus **self-test the singlet**. It is worthwhile to emphasize that the customer can be classical, and moreover does not need to possess any knowledge of quantum physics.

In reality, due to experimental noise and the finiteness of the number of rounds that can be played, the Bell violation cannot be expected to equal $2\sqrt{2}$ even when Alice and Bob do share a singlet state. This practical aspect requires self-testing statements to be **robust to errors**, which is the main topic of this thesis. We will get into further details on robust self-testing and the formalization of the self-testing problem in general in Chapter 3.

2.2.4 Jordan's lemma

In the device-independent scenario, there are no limitations on the dimension of the observables of Alice and Bob. The CHSH operator, which can be written in terms of the observables, is therefore difficult to analyze in general. Fortunately, the following lemma greatly simplifies the analysis.

Lemma 2.7. (Jordan's lemma)
Let M_0 and M_1 be Hermitian operators with eigenvalues 1 and -1. Then there exists a basis in which both operators are block-diagonal, where the blocks have dimensions 2×2 at most.

For the proof, we refer to Scarani's lecture notes [Sca12, p.32].

Jordan's lemma is useful in the context of obtaining self-testing statements from operator inequalities, as explained in Section 3.5.

2.2.5 Werner states

Many results have been proven on the relations between observed statistics, Bell inequalities and the quantum states and measurements achieving maximal violation, both for specific Bell inequalities as well as general results (we again refer to the work of Brunner et al. for an overview [BCP+14]). We wish to emphasize a specific result by Werner [Wer89], who proved that there exist bipartite mixed states (later called *Werner states*) that do not violate *any* Bell inequality, despite the fact that they are entangled. The Werner states where both parties have a qubit can be written as a convex combination of the singlet and the maximally mixed state:

$$\rho = p\,|\phi\rangle\langle\phi| + (1-p)\frac{\mathbb{1}_4}{4}$$

where $|\phi\rangle := (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state and $p \in [0,1]$ is the *visibility parameter*. It can be shown that for $p > \frac{1}{3}$, the Werner states are entangled. However, the correlation statistics that Alice and Bob can achieve with this state can be described by local variables for $p \leq \frac{5}{12}$ when any POVM measurements are allowed [Bar02] (for just projective measurements, the statistics can be explained by local variables for $p \leq \frac{1}{2}$ [Wer89]). Therefore all Werner states with $\frac{1}{3} < p \leq \frac{1}{2}$ are entangled, but nonetheless cannot violate any Bell inequality.

Thus although entanglement is needed to violate a Bell inequality, it is not sufficient. In Chapter 4, we show a result that has, in the context of self-testing, a similar flavor: there exists a bipartite state that is entangled, but there are no quantum operations that Alice and Bob could apply locally to achieve the fidelity of this state with the singlet to more than some trivial lower bound. For details, we refer to Chapter 4.

Classical communication is prohibited in Bell experiments. In order to relate our result from Chapter 4 to similar results in the LOCC setting (local operations *with* classical communication), we need the notions of entanglement distillation and bound entanglement.

2.2.6 Entanglement distillation and the partial transpose

Two separated parties can send quantum information to each other using just local operations and classical communication (LOCC) through a process called 'teleportation', which requires the presence of highly entangled bipartite states. In order to obtain such states, one uses 'entanglement distillation', which refers to the use of several copies of (slightly) entangled bipartite states in order to create a single maximally entangled state between two players using LOCC operations. A bipartite state is called 'undistillable' if it is not possible to distill entanglement from any number of copies of this state. A state is called 'bound entangled' if it is entangled but nevertheless undistillable. For a more thorough introduction to entanglement distillation, we refer to Horodecki et al. [HHHH09].

A useful tool to test for distillability is the partial transpose of a density matrix. Let $\rho_{AB} \in \mathrm{D}(\mathbb{H}_A \otimes \mathbb{H}_B)$ be a density matrix, which can be written as

$$\rho_{AB} = \sum_{i=1,j=1}^{d_A} \sum_{k=1,l=1}^{d_B} p_{kl}^{ij}\,|i\rangle\langle j| \otimes |k\rangle\langle l|$$

with coefficients $p_{kl}^{ij} \in \mathbb{C}$, where d_A (d_B) is the dimension of $\mathbb{H}_A$ ($\mathbb{H}_B$). The partial transpose of ρ_{AB} is defined as

$$\sum_{i=1,j=1}^{d_A} \sum_{k=1,l=1}^{d_B} p_{lk}^{ij} \, |i\rangle\langle j| \otimes |k\rangle\langle l| \, .$$

In order to test whether a bipartite state is separable, one could use the 'Peres-Horodecki criterion', also known as the 'PPT criterion' [Per96] ('PPT' stands for 'positive partial transpose'). The PPT criterion states that if ρ_{AB} is separable, then its partial transpose is positive semidefinite. The converse does not hold in general. However, PPT entangled states are "weakly" entangled in the sense that they are not distillable.

If a state is not PPT, it is called NPT (which stands for 'nonpositive partial transpose'). It was shown by Werner and Wolf that nonpositivity of the partial transpose is necessary for CHSH violation [WW01]. The only known criterion for proving a state to be undistillable, is to show that it is PPT. It is still an open question whether states exist which are both bound entangled and NPT.

2.3 Preliminaries to Chapter 4

This section contains a number of auxiliary lemmas for our proof that CHSH violation does not imply nontrivial singlet extractability (see Chapter 4).

All the lemmas in this section state properties of quantum states or quantum channels. First we show that applying the modulus to a hermitian operator can never decrease inner products with positive matrices.

Lemma 2.8. *Let $X \in \mathrm{Herm}(\mathbb{H})$ and let $Y \in \mathrm{Pos}(\mathbb{H})$. Then $\langle X, Y\rangle \leq \langle |X|, Y\rangle$, where $\langle ., .\rangle$ denotes the Hilbert-Schmidt inner product as defined in Equation* (2.1).

Proof. Note that the operator inequality $X \leq |X|$ holds trivially. From this, we derive that, in particular, for every eigenvector $|j\rangle$ of Y, we have $\langle j|X|j\rangle \leq \langle j| \, |X| \, |j\rangle$. Now write Y in its eigenbasis: $Y = \sum_{j=1}^{d} \lambda_j \, |j\rangle\langle j|$, with $\lambda_j \geq 0$ for all $1 \leq j \leq d$, where d is the dimension of the Hilbert space $\mathbb{H}$. Then we compute

$$\langle |X|, Y\rangle - \langle X, Y\rangle = \langle |X| - X, Y\rangle = \mathrm{Tr}((|X| - X)Y) = \sum_{j=1}^{d} \lambda_j \big(\langle j| \, |X| \, |j\rangle - \langle j|X|j\rangle \big)$$

which is positive since $\lambda_j \geq 0$ and $\langle j| \, |X| \, |j\rangle \geq \langle j|X|j\rangle$ for all $1 \leq j \leq d$. □

Throughout our proofs in the thesis, we will use the fact that the trace inner product of two positive operators is nonnegative, as shown in the next lemma.

Lemma 2.9. *Let $X, Y \in \mathrm{Pos}(\mathbb{H})$. Then $\mathrm{Tr}(XY) \geq 0$.*

Proof. Write Y in its eigenbasis as $Y = \sum_{j=1}^{d} \lambda_j \, |j\rangle\langle j|$ with $1 \leq j \leq d = \dim(\mathbb{H})$. Since X is positive, we have $\langle j|X|j\rangle \geq 0$ for all j. Hence $\mathrm{Tr}(XY) = \sum_{j=1}^{d} \lambda_j \, \langle j|X|j\rangle \geq 0$. □

We also prove that the trace inner products of three arbitrary single-qubit states obey a restriction that is reminiscent of the triangle inequality for norms.

Lemma 2.10. *Let ρ, σ and τ be density matrices of single qubits. Suppose that*

$$\mathrm{Tr}(\rho\sigma) \geq 1 - \delta_1$$

$$\mathrm{Tr}(\sigma\tau) \geq 1 - \delta_2$$

for $0 \leq \delta_1, \delta_2 \leq \frac{1}{2}$.

Then $\mathrm{Tr}(\rho\tau) \geq (1 - 2\delta_1)(1 - 2\delta_2)$. This bound is tight for some single-qubit density matrices.

Proof. Since ρ is a density matrix of a single qubit, we can write $\rho = \frac{1}{2}(\mathbb{1}_2 + r_\mathsf{X}\mathsf{X} + r_\mathsf{Y}\mathsf{Y} + r_\mathsf{Z}\mathsf{Z})$ with $||\vec{r}||^2 = (r_\mathsf{X})^2 + (r_\mathsf{Y})^2 + (r_\mathsf{Z})^2 \leq 1$, for a three-vector of coefficients $\vec{r} = (r_\mathsf{X}, r_\mathsf{Y}, r_\mathsf{Z})$. Similarly for σ and τ, with vectors $\vec{s}$ and $\vec{t}$, respectively.

It is straightforward to calculate $\mathrm{Tr}(\rho\sigma) = \frac{1}{2}(1 + \vec{r} \cdot \vec{s})$, where $\vec{r} \cdot \vec{s} := r_\mathsf{X}s_\mathsf{X} + r_\mathsf{Y}s_\mathsf{Y} + r_\mathsf{Z}s_\mathsf{Z}$ is the standard inner product on Euclidean spaces. The expressions for $\mathrm{Tr}(\rho\tau)$ and $\mathrm{Tr}(\sigma\tau)$ follow by appropriate substitution of the direction vector. With these expression for the trace of the products of ρ,σ and τ, the lower bounds become $(\vec{r} \cdot \vec{s}) \geq 1 - 2\delta_1$ and $(\vec{s} \cdot \vec{t}) \geq 1 - 2\delta_2$.

We distinguish two cases:

- Case $\vec{s} = \vec{0}$. Then $\sigma = \frac{\mathbb{1}_2}{2}$, so $\mathrm{Tr}(\rho\sigma) = \mathrm{Tr}(\sigma\tau) = \frac{1}{2}$, hence $\delta_1 = \delta_2 = \frac{1}{2}$ (since we have assumed that $\delta_1, \delta_2 \leq \frac{1}{2}$). Hence the bound $\mathrm{Tr}(\rho\tau) \geq (1 - 2\delta_1)(1 - 2\delta_2)$ reduces to $\mathrm{Tr}(\rho\tau) \geq 0$, which indeed holds by Lemma 2.9.

- Case $\vec{s} \neq \vec{0}$. Dragomir proved an inequality between three inner products [Dra05, p.47]:

  $$\left|\frac{(\vec{r} \cdot \vec{s}) \cdot (\vec{s} \cdot \vec{t})}{||\vec{s}||^2} - \frac{(\vec{r} \cdot \vec{t})}{2}\right| \leq \frac{||\vec{r}||^2||\vec{t}||^2}{2} \quad \text{for } \vec{s} \neq \vec{0}.$$

  Since $||\vec{r}||, ||\vec{t}|| \leq 1$ for our problem, this reduces to

  $$\left|\frac{(\vec{r} \cdot \vec{s}) \cdot (\vec{s} \cdot \vec{t})}{||\vec{s}||^2} - \frac{(\vec{r} \cdot \vec{t})}{2}\right| \leq \frac{1}{2} \quad \text{for } \vec{s} \neq \vec{0}$$

  In particular, we can remove the modulus-symbols to obtain

  $$\frac{(\vec{r} \cdot \vec{s}) \cdot (\vec{s} \cdot \vec{t})}{||\vec{s}||^2} - \frac{(\vec{r} \cdot \vec{t})}{2} \leq \frac{1}{2} \quad \text{for } \vec{s} \neq \vec{0}$$

  Reordering results in

  $$\frac{(\vec{r} \cdot \vec{s}) \cdot (\vec{s} \cdot \vec{t})}{||\vec{s}||^2} \leq \frac{1 + (\vec{r} \cdot \vec{t})}{2} \quad \text{for } \vec{s} \neq \vec{0}$$

  Since $||\vec{s}|| \leq 1$ and $(\vec{r} \cdot \vec{s}) \geq 1 - 2\delta_1 \geq 0$ and $(\vec{s} \cdot \vec{t}) \geq 1 - 2\delta_2 \geq 0$, we get

  $$(\vec{r} \cdot \vec{s}) \cdot (\vec{s} \cdot \vec{t}) \leq \frac{(\vec{r} \cdot \vec{s}) \cdot (\vec{s} \cdot \vec{t})}{||\vec{s}||^2} \leq \frac{1}{2}(1 + (\vec{r} \cdot \vec{t})) = \mathrm{Tr}(\rho\tau) \quad \text{for } \vec{s} \neq \vec{0}$$

  Now substituting our bounds $(\vec{r} \cdot \vec{s}) \geq 1 - 2\delta_1$ and $(\vec{s} \cdot \vec{t}) \geq 1 - 2\delta_2$ yields the desired result.

Tightness. Let

$$\rho = \frac{\mathbb{1}_2 + \mathsf{X}}{2}, \quad \sigma = \frac{\mathbb{1}_2 + \frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}}{2}, \quad \tau = \frac{\mathbb{1}_2 + \mathsf{Z}}{2}$$

be pure density matrices. Set $\delta = \frac{1}{2} - \frac{1}{2\sqrt{2}}$. Then we verify that

$$\mathrm{Tr}(\rho\sigma) = \frac{1}{4}\,\mathrm{Tr}\left(\mathbb{1}_2 + \frac{1}{\sqrt{2}}\mathbb{1}_2\right) = \frac{1}{2} + \frac{1}{2\sqrt{2}} = 1 - \delta$$

and by symmetry, $\mathrm{Tr}(\sigma\tau) = 1 - \delta$ too. Also,

$$\mathrm{Tr}(\rho\tau) = \frac{1}{4}\,\mathrm{Tr}(\mathbb{1}_2) = \frac{1}{2} = \left(\frac{1}{\sqrt{2}}\right)^2 = \left(1 - \left[1 - \frac{1}{\sqrt{2}}\right]\right)^2 = \left(1 - 2\left[\frac{1}{2} - \frac{1}{2\sqrt{2}}\right]\right)^2 = (1-2\delta)^2$$

hence the bound is tight for ρ, σ, τ. □

The bipartite state that we constructed (see Chapter 4) is a 'classical-quantum state', which is defined as follows.

Definition 2.11. *Let $\mathbb{H}_c$ and $\mathbb{H}_q$ be Hilbert spaces of dimensions d_c and d_q, respectively. A state $\rho_{cq} \in \mathrm{D}(\mathbb{H}_c \otimes \mathbb{H}_q)$ is called **classical-quantum** if it can be written in the form*

$$\rho_{cq} = \sum_{j=1}^{d_c} p_j \,|j\rangle\langle j| \otimes \sigma_j$$

where the set $\{|j\rangle\}_{j=1}^{d_c}$ forms an orthonormal basis of $\mathbb{H}_c$, p is a probability vector and $\sigma_j \in \mathrm{D}(\mathbb{H}_q)$ for all $1 \leq j \leq d_c$.

Let $\mathcal{B}_c = \{|j\rangle\}_{j=1}^{d_c}$ be an orthonormal basis for $\mathbb{H}_c$. We refer to the set

$$S_{\mathcal{B}_c} \;:=\; \left\{ \sum_{j=1}^{d_c} p_j \,|j\rangle\langle j| \otimes \sigma_j \;\middle|\; \sigma_j \in \mathrm{D}(\mathbb{H}_q) \text{ and } p \text{ a probability vector of dimension } d_c \right\}$$

as the 'set of classical-quantum states on $\mathbb{H}_c \otimes \mathbb{H}_q$ given basis $\mathcal{B}_c$'.

Quantum channels acting on classical-quantum states allow for a special decomposition; we can think that the channel reads the value of the classical register and applies a channel from $\mathbb{H}_q$ which depends on this value.

Lemma 2.12. *Let $\mathbb{H}_c, \mathbb{H}_q$ and $\mathbb{H}_{\mathrm{out}}$ be Hilbert spaces of dimensions d_c, d_q and d_{out}, respectively. Let $\mathcal{B}_c = \{|j\rangle\}_{j=1}^{d_c}$ be an orthonormal basis for $\mathbb{H}_c$ and denote by $S_{\mathcal{B}_c}$ the set of classical-quantum states on $\mathbb{H}_c \otimes \mathbb{H}_q$ given basis $\mathcal{B}_c$.*

Let $\Lambda : \mathrm{D}(\mathbb{H}_c \otimes \mathbb{H}_q) \to \mathrm{D}(\mathbb{H}_{\mathrm{out}})$ be a quantum channel. Then there exists a set of d_c quantum channels $\{\Lambda_j : \mathrm{D}(\mathbb{H}_q) \to \mathrm{D}(\mathbb{H}_{\mathrm{out}})\}_{j=1}^{d_c}$ such that for every state $\rho_{CQ} := \sum_{j=1}^{d_c} p_j \,|j\rangle\langle j| \otimes \sigma_j \in S_{\mathcal{B}_c}$, we have

$$\Lambda(\rho_{CQ}) = \sum_{j=1}^{d_c} p_j \Lambda_j(\sigma_j).$$

Proof. Extend the basis $\mathcal{B}_c$ to a basis of $\mathbb{H}_c \otimes \mathbb{H}_q$, denoted by $\mathcal{B}_{cq} := \{|j\rangle \otimes |j'\rangle \mid 1 \leq j \leq d_c, 1 \leq j' \leq d_q\}$, where $\{|j'\rangle\}_{j'=1}^{d_q}$ is an arbitrary basis of $\mathbb{H}_q$.

Write Λ in terms of its Krauss operators:

$$\Lambda(\rho) := \sum_{i=1}^{N} \Gamma_i \rho (\Gamma_i)^\dagger \quad \text{with} \ \sum_i (\Gamma_i)^\dagger \Gamma_i = \mathbb{1}_{d_c \cdot d_q}.$$

Now write each Krauss operator as $\Gamma_i = \sum_{k=1}^{d_c} \langle k| \otimes \Gamma_{i,k}$, where $|k\rangle \in \mathcal{B}_c$ and $\Gamma_{i,k} : \mathbb{H}_q \to \mathbb{H}_{\text{out}}$. One can do this, for example, by writing each Γ_i, which can be represented as a matrix with d_{out} rows and $d_c \cdot d_q$ columns, in the basis $\mathcal{B}_{cq}$. Then partition the matrix into d_c blocks of d_{out} rows and d_q columns; each such a block is denoted by $\Gamma_{i,k}$, for $1 \leq k \leq d_c$.

Now we rewrite

$$\begin{aligned}
\Lambda(\rho_{CQ}) &= \sum_{j=1}^{d_c} p_j \sum_{i=1}^{N} \Gamma_i (|j\rangle\langle j| \otimes \sigma_j)(\Gamma_i)^\dagger \\
&= \sum_{j=1}^{d_c} p_j \sum_{i=1}^{N} \sum_{k=1}^{d_c} \sum_{k'=1}^{d_c} \big(\langle k| \otimes \Gamma_{i,k} \big) \big(|j\rangle\langle j| \otimes \sigma_j \big) \big(|k'\rangle \otimes (\Gamma_{i,k'})^\dagger \big) \\
&= \sum_{j=1}^{d_c} p_j \sum_{i=1}^{N} \sum_{k=1}^{d_c} \sum_{k'=1}^{d_c} \delta_{k,j} \delta_{j,k'} \Gamma_{i,k} \sigma_j (\Gamma_{i,k'})^\dagger \\
&= \sum_{j=1}^{d_c} p_j \sum_{i=1}^{N} \Gamma_{i,j} \sigma_j (\Gamma_{i,j})^\dagger
\end{aligned}$$

where δ denotes the Kronecker-delta-function. We complete the proof by showing that the maps $\Lambda_j : \mathrm{D}(\mathbb{H}_q) \to \mathrm{D}(\mathbb{H}_{\text{out}})$, $\sigma \mapsto \sum_{i=1}^{N} \Gamma_{i,j} \sigma (\Gamma_{i,j})^\dagger$ are quantum channels. For this, we only need to prove that $\sum_{i=1}^{N} (\Gamma_{i,j})^\dagger \Gamma_{i,j} = \mathbb{1}_{d_q}$ for all $j \in \{1, \ldots, d_c\}$.

Fix $j \in \{1, \ldots, d_c\}$. Let $|a\rangle, |b\rangle \in \mathbb{H}_q$ and let $|j\rangle$ be the j-th basis vector in B_{d_c}. Then we derive

$$\begin{aligned}
\langle a| \sum_{i=1}^{N} (\Gamma_{i,j})^\dagger \Gamma_{i,j} |b\rangle &= \sum_{k=1}^{d_c} \sum_{k'=1}^{d_c} \Big(\langle j| \otimes \langle a| \Big) \Big(\sum_{i=1}^{N} \big(|k\rangle \otimes (\Gamma_{i,k})^\dagger \big) \big(\langle k'| \otimes \Gamma_{i,k'} \big) \Big) \Big(|j\rangle \otimes |b\rangle \Big) \\
&= \sum_{i=1}^{N} \Big(\langle j| \otimes \langle a| \Big) \Big(\sum_{k=1}^{d_c} \sum_{k'=1}^{d_c} \big(|k\rangle \otimes (\Gamma_{i,k})^\dagger \big) \big(\langle k'| \otimes \Gamma_{i,k'} \big) \Big) \Big(|j\rangle \otimes |b\rangle \Big).
\end{aligned}$$

Now, using the fact that $\sum_{i=1}^{N} (\Gamma_i)^\dagger \Gamma_i = \mathbb{1}_m$ and $\Gamma_i = \sum_{k=1}^{d_c} \langle k| \otimes \Gamma_{i,k}$ for every $i \in \{1, \ldots, N\}$, we obtain

$$\begin{aligned}
\langle a| \sum_{i=1}^{N} (\Gamma_{i,j})^\dagger \Gamma_{i,j} |b\rangle &= (\langle j| \otimes \langle a|) \mathbb{1}_m (|j\rangle \otimes |b\rangle) \\
&= \langle j|j\rangle \cdot \langle a|b\rangle \\
&= \langle a|b\rangle
\end{aligned}$$

hence $\sum_{i=1}^{N} (\Gamma_{i,j})^\dagger \Gamma_{i,j} = \mathbb{1}_m$ indeed.

□

The following lemma captures the fact that if a qubit-to-qubit quantum channel moves the maximally mixed state $\frac{\mathbb{1}_2}{2}$ to a certain point close to the surface of the Bloch ball, then the entire Bloch ball is mapped to a small region close to this point.

Lemma 2.13. *Let $P = c_{\mathsf{X}}\mathsf{X} + c_{\mathsf{Y}}\mathsf{Y} + c_{\mathsf{Z}}\mathsf{Z}$, where $c_{\mathsf{X}}, c_{\mathsf{Y}}, c_{\mathsf{Z}}$ are real-valued coefficients obeying $c_{\mathsf{X}}^2 + c_{\mathsf{Y}}^2 + c_{\mathsf{Z}}^2 = 1$. Furthermore, let Λ be a quantum channel, such that the eigenvalues of $\Lambda(\frac{\mathbb{1}_2}{2})$ are δ and $1-\delta$, with $0 \leq \delta \leq \frac{1}{2}$.*
Then the matrix modulus of $\Lambda(P)$ can be upper bounded by $|\Lambda(P)| \leq 2\sqrt{\delta}\mathbb{1}_2$.

Proof. By positivity of quantum channels we have $\Lambda\left(\frac{\mathbb{1}_2 \pm P}{2}\right) \geq 0$. Denote $\omega := \Lambda(\frac{\mathbb{1}_2}{2})$, so that we obtain $-2\omega \leq \Lambda(P) \leq 2\omega$.

Every matrix in this proof of which the entries are written explicitly, is written in the eigenbasis of ω. This is straightforward for the matrix ω:

$$\omega = \begin{pmatrix} \delta & 0 \\ 0 & 1-\delta \end{pmatrix}.$$

In order to write $\Lambda(P)$ in the eigenbasis of ω, we note the following facts. Since Λ is a quantum channel, it sends hermitian operators to hermitian operators. Also, Λ is trace-preserving. Combining these two, we see that $\Lambda(P)$ can be written in the eigenbasis of ω as

$$\Lambda(P) = \begin{pmatrix} t & y \\ y^* & -t \end{pmatrix}$$

for some $t \in \mathbb{R}, y \in \mathbb{C}$.
Now $\Lambda(P) \geq -2\omega$ translates into

$$\begin{pmatrix} 2\delta + t & y \\ y^* & 2 - 2\delta - t \end{pmatrix} \geq 0$$

which implies that the determinant of $\Lambda(P) + 2\omega$ is greater than or equal to zero:

$$(2\delta + t) \cdot (2 - 2\delta - t) - |y|^2 \geq 0.$$

Similarly, from $\Lambda(P) \leq 2\omega$, we obtain

$$(2\delta - t) \cdot (2 - 2\delta + t) - |y|^2 \geq 0.$$

Adding the two equalities gives

$$(2\delta + t)(2 - 2\delta - t) + (2\delta - t)(2 - 2\delta + t) - 2|y|^2 \geq 0.$$

Expanding the left hand side yields

$$8\delta - 8\delta^2 - 2t^2 - 2|y|^2 \geq 0.$$

Dividing by 2 on each side yields

$$4\delta - 4\delta^2 - t^2 - |y|^2 \geq 0$$

from which we obtain

$$t^2 + |y|^2 \leq 4\delta - 4\delta^2 = 4\delta(1-\delta) \leq 4\delta \cdot 1 = 4\delta \tag{2.13}$$

where we used the fact that $\delta \geq 0$.

We use Equation (2.13) to show that $2\sqrt{\delta} - |\Lambda(P)| \geq 0$. We compute (in the eigenbasis of ω):

$$\begin{aligned} 2\sqrt{\delta}\mathbb{1}_2 - |\Lambda(P)| &= 2\sqrt{\delta}\mathbb{1}_2 - \sqrt{\Lambda(P)^2} \\ &= 2\sqrt{\delta}\mathbb{1}_2 - \sqrt{\begin{pmatrix} t^2+|y|^2 & 0 \\ 0 & t^2+|y|^2 \end{pmatrix}} \\ &= \begin{pmatrix} 2\sqrt{\delta} - \sqrt{t^2+|y|^2} & 0 \\ 0 & 2\sqrt{\delta} - \sqrt{t^2+|y|^2} \end{pmatrix}. \end{aligned}$$

Using Equation (2.13), we now conclude that $2\sqrt{\delta}\mathbb{1}_2 - |\Lambda(P)|$ is a positive matrix.

□

The fidelity of any separable bipartite state with another state is bounded as in the following lemma.

Lemma 2.14. *Let ρ be a separable bipartite state. The fidelity of ρ with a pure bipartite state φ is upper bounded as*

$$F(\rho, \varphi) \leq \lambda_{\max}(\varphi)^2$$

where $\lambda_{\max}(\varphi)$ denotes the largest Schmidt coefficient of the state φ. Furthermore, if $\lambda_{\max}$ is nondegenerate, then the bound is saturated if and only if ρ equals the unique pure product state that corresponds to $\lambda_{\max}(\varphi)$ in the Schmidt decomposition of φ.

Proof. We prove the case where ρ is a pure product state. Since a separable state can be written as convex combination of product states and any product state can be written as convex combination of pure product states, considering pure product states is sufficient for proving the lemma.

Write $\rho = \psi_A \otimes \psi_B$ with $\psi_A = |\psi_A\rangle\langle\psi_A|$ and $\psi_B = |\psi_B\rangle\langle\psi_B|$ pure states. Furthermore, let

$$|\varphi\rangle = \textstyle\sum_j \lambda_j |j_A\rangle \otimes |j_B\rangle$$

be a Schmidt decomposition of $\varphi = |\varphi\rangle\langle\varphi|$. Using Lemma 2.2 and the fact that φ is pure, we have

$$F(\rho, \varphi) = F(\psi_A \otimes \psi_B, \varphi) = \mathrm{Tr}((\psi_A \otimes \psi_B)\varphi)$$

hence

$$\begin{aligned} F(\rho, \varphi) &= \mathrm{Tr}((\psi_A \otimes \psi_B)\varphi) \\ &= \left|(\langle\psi_A| \otimes \langle\psi_B|)\,|\varphi\rangle\right|^2 \\ &= \Big|\sum_j \lambda_j (\langle\psi_A|j_A\rangle \cdot \langle\psi_B|j_B\rangle)\Big|^2 \\ &\leq \sum_j \lambda_j^2 |\langle\psi_A|j_A\rangle \cdot \langle\psi_B|j_B\rangle|^2 \\ &\leq \lambda_{j^{\max}}^2 \cdot 1 \end{aligned} \tag{2.14}$$

where $\lambda_{j^{\max}} = \lambda_{\max}(\varphi)$ denotes the largest Schmidt coefficient of φ.

If $\lambda_{j^{\max}}$ is nondegenerate, then Inequality (2.14) is saturated if and only if $|\varphi_A\rangle\otimes|\varphi_B\rangle = |j_A^{\max}\rangle\otimes|j_B^{\max}\rangle$. Conversely, if $|\varphi_A\rangle\otimes|\varphi_B\rangle = |j_A^{\max}\rangle\otimes|j_B^{\max}\rangle$, then indeed $F(\rho,\varphi) = \lambda_{j^{\max}}^2 = \lambda_{\max}(\varphi)^2$. We thus conclude that the upper bound is saturated if and only if $|\varphi_A\rangle \otimes |\varphi_B\rangle = |j_A^{\max}\rangle \otimes |j_B^{\max}\rangle$.

□

3 Self-testing: a brief overview

One of the most promising branches of quantum cryptography is arguably the field of Quantum Key Distribution (QKD), where the main task is to use quantum states and measurements to generate a random bit string (the key) known only to the two communicating parties and no-one else. Surprisingly, QKD protocols were constructed that remained secure even if they are executed using untrusted devices. Such protocols are now called 'device-independent' quantum-key-distribution (DI-QKD) protocols. Bell nonlocality is a key tool for DI-QKD, since observed statistics can allow us to rule out any classical correlations and moreover can in some extremal cases almost uniquely identify the state and measurements.

The development of the field of *self-testing*, which refers to device-independent certification of the state and measurements, cannot be seen independently from its applications in quantum cryptography. We first give a brief overview of the development of the device-independent perspective on QKD and randomness generation, before formalizing the self-testing problem. As already mentioned in Chapter 2, in the entire thesis the term 'singlet' will be used to refer to a maximally entangled pair of two qubits.

3.1 The emergence of the device-independence paradigm

In the celebrated QKD protocol by Ekert [Eke91], Alice and Bob receive qubits from a singlet source, so that each of them ends up with one qubit from each maximally entangled pair. The two parties apply measurements to their part of the states and subsequently use the CHSH inequality to test for eavesdropping: suppose that an eavesdropper had intercepted one or more qubits and measured them before sending them on. Then the resulting two-qubit state that the intercepted qubits were part of is not entangled. Hence the presence of an eavesdropper decreases the CHSH value β to at most the classical value 2. The case of an imperfect source that does not send maximally entangled states, thereby seriously compromising the security of this protocol, will also yield strictly smaller CHSH violation. Ekert was the first to show that Bell nonlocality can be useful for quantum cryptographic purposes.

Later, Bennett et al. [BBM92] proposed a simpler QKD scheme (called 'BBM' afterwards), based on maximally entangled pairs of qubits too but without the need to invoke Bell nonlocality. This protocol is an entanglement-based version of the BB84 scheme [BB84].

Such entanglement-based protocols were the initial motivation behind the research of Mayers and Yao, who were the first to explicitly notice the usefulness of certifying state and measurements using observed correlations [MY98, MY04] (although the ideas were implicit in earlier papers [Cir80, PR92, SW88]). They remark that imperfect sources can be constructed that undermine the security of the protocol while it requires thorough investigation to detect the imperfection of the source; therefore, a 'self-checking source' is needed, and we can 'self-test' the underlying state and measurements. At the time, it was already known that all states that violate the CHSH inequality maximally can be written in the form

$$(U_{AA'} \otimes U_{BB'})(|\varphi_{AB}\rangle\langle\varphi_{AB}| \otimes \tau_{A'B'})(U_{AA'}^\dagger \otimes U_{BB'}^\dagger) \tag{3.1}$$

does so too, where $|\varphi_{AB}\rangle$ is the singlet state, $U_{AA'}$ and $U_{BB'}$ are unitaries and $\tau_{A'B'}$ can be any state [PR92, BMR92]. However, although Mayers and Yao explicitly mention the relevance of Bell inequalities in Ekert's QKD protocol, they did not explicitly use Bell inequalities for self-testing themselves. Instead, they used the entire conditional probability distribution, obtained from a Bell experiment, and showed that particular observed correlations uniquely identify the underlying state and measurements, up to the same equivalences as defined in Equation (3.1). It is noteworthy to mention that the conditions that Mayers and Yao impose on the observed correlations clearly violate the CHSH inequality,

although not maximally; thus, use of the entire set of observed correlations rather than just a Bell value yields more information about the underlying state, as one would expect.

After the first paper by Mayers and Yao on self-testing, it took a while before the field of device-independence took off. Scarani, one of the pioneers in the field, attributes this in hindsight to the fact that quantum cryptography was still young [Sca12, Appendix B]. After quantum cryptography had outgrown its early stage, the first work to pave the way for a new field was the work of Barrett, Hardy and Kent [BHK05], who used Bell nonlocality to construct the first device-independent quantum cryptographic protocol which is secure against all non-signaling eavesdroppers. It is important to note here that such a protocol remains useful even if quantum mechanics turns out to be invalid, as long as the adversaries are non-signaling. With the proposed protocol, Alice and Bob can generate a single shared bit.

Not long after, Acín et al. [AGM06] published a paper in which several previous results were reconsidered. First, the authors noted that the BB84 protocol and its entanglement-based variant BBM are not secure in the context of device-independence. The realization that the security of BB84 and BBM are based upon the assumptions that qubits are being sent came from the perspective of Bell nonlocality: in both protocols, the joint statistics that Alice and Bob compute can always be reproduced using local variables (see [Sca12, p.12]), hence the correlations that are produced locally could in reality be using a larger, separable state. This makes both protocols unfit for use with untrusted devices. Others confirmed the compromised security of BB84 and BBM in a device-independent setting for different reasons [MMMO06, Appendix A]. Inspired by the Ekert protocol, Acín et al. use the CHSH inequality to construct a QKD protocol and call it 'CHSH-protocol'. In this work, the set of states that maximally violate the CHSH inequality, as given by Equation (3.1), is explicitly used.

At roughly the same time, Scarani et al. [SGB$^+$06] built upon the work of Barrett et al. to prove that noisy quantum states can also be used to achieve correlations that are 'sufficiently nonlocal' for key generation.

Later, Acín et al. zoom in on on quantum adversaries rather than the more general nonsignaling ones, in order to obtain better key rates [ABG$^+$07]. It is this paper also, in which the term 'device-independence' is first coined.

A couple of years later, the first results on randomness generation with potentially untrusted devices and randomness amplification from Bell nonlocality were obtained [Col07, PAM$^+$10, CK11].

We can say that, based upon the numerous results in the field, in the meantime, device-independent QKD has clearly taken off. Most research so far has focused on creating more key with increasingly noisier states and fewer device uses and their proofs of security in a device-independent scenario [AMP06, MW06, MPA11, VV14, DFR16, AFRV16]. Randomness generation and amplification with potentially untrusted devices has also seen a lot of activity [SCA$^+$11, VV12, CVY13, BPPP14, MS16]. We refer to Ekert and Renner's review for a more complete set of references [ER14].

Whenever 'device-independence' is used nowadays, the term usually refers to the use of Bell nonlocality in applications such as QKD and randomness generation. While increasingly more results on this side were obtained, the fundamental interests in the connection between observed correlations and the underlying state and measurements itself also started to grow: this is the subject of self-testing, which deals with device-independent state and measurement certification. In self-testing scenarios, the fact that nonlocality can be shown using observed correlations is taken one step further (see also Figure 2): if we **assume that quantum physics gives a correct description of the physical system under**

consideration, then given observed correlations, what knowledge can be inferred about the state and the measurements? The first work of Mayers and Yao stated correlations that are only produced by the singlet (up to the equivalences from Equation (3.1)); an immediate follow-up question is: what if the correlations are not perfect? In practice, the real correlations are based upon a finite number of runs of a Bell game and can therefore only be estimated. Moreover, experiments are never free of noise. We thus need to be able to perform self-testing with **robustness against errors**. Bardyn et al. [BLM+09] was one of the first to consider several figures of merit to formalize the robustness problem, followed by McKague et al. [McK10, MYS12, McK14].

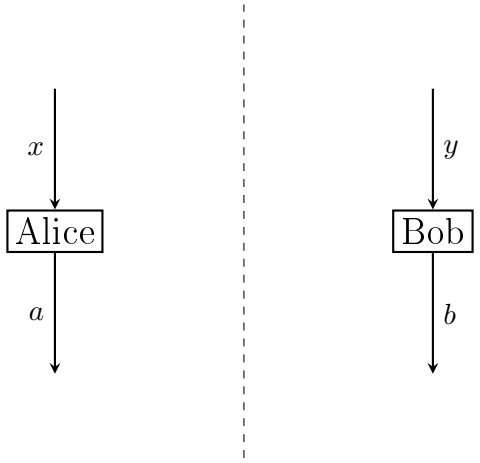

Figure 2: The setup for self-testing of bipartite quantum states. Two players, Alice and Bob, each have a device. Both devices are given classical input (x and y, respectively), which corresponds to the application of measurements inside the devices, and classical output (a and b). The devices are physically isolated, so that sending signals from one to the other is not possible. The central question is: given the conditional probabilities $\Pr(a, b|x, y)$, what information can be obtained about the state that the devices share?

This thesis presents our work on self-testing of quantum states. In this chapter, we give an overview of the states that have been shown to be identifiable with certain correlations. We discuss several figures of merit used to formalize the problem of robust self-testing and we show how previous self-testing proofs fit in in these formalizations. This chapter is ordered as follows. First, we define when a state can be self-tested and discuss several formalizations of **robust** self-testing, the setting in which the observed correlations are not identical to the correlations a reference state would produce. Then we give a brief overview of which states have been shown to be self-testable so far. After this, we turn our attention to an outline of the methods that were used so far to prove self-testing statement: we identified two main methods for finding robust self-testing statements and categorized the main results according to these two methods. This chapter aims to only give a brief overview of the main results in the field of self-testing. In the next chapters, more focus will be on self-testing of the singlet and of partially entangled two-qubit states, since our work extends the known results for the certification of these states.

3.2 Formalization of the self-testing problem of state certification

In self-testing of states, the main question is: given observed correlations, what can be inferred about the underlying state? Before giving a formalization of this question, let us focus on what we *cannot* infer. For simplicity, we focus on the CHSH inequality as an example.

We know that if the devices share the singlet state, then their correlations can maximally violate the CHSH inequality if particular measurements are applied inside the devices. However, the devices could as well contain states that are left untouched during the measurement. As more poetically formulated by McKague et al. [MYS12], isolated qubits do not exist in nature: for example, when a qubit is realized as the spin of an electron, then we need the entire electron, including properties such as angular momentum. Therefore, if we add degrees of freedom to the singlet state (that is, we tensor it with an arbitrary state upon which the measurements act trivially), then the devices will still maximally violate the CHSH inequality.

Moreover, Alice could change the local bases of her state and simultaneously change her measurement operators in the same way, which will not affect the correlation statistics when she plays the CHSH game. Likewise for Bob. Hence, our definitions should also allow for local changes of bases.

Neither of these features can be detected on the basis of the observed correlations only. Therefore, if ρ_{AB} maximally violates a particular Bell inequality, then

$$(U_{AA'} \otimes U_{BB'})(\rho_{AB} \otimes \tau_{A'B'})(U_{AA'}^\dagger \otimes U_{BB'}^\dagger) \tag{3.2}$$

does so too, where $U_{AA'}$ and $U_{BB'}$ are unitaries and $\tau_{A'B'}$ can be any state. Note that this expression is exactly the same as Equation (3.1); if we let ρ_{AB} denote the singlet state, then **all** states that violate the CHSH inequality can be written as in Equation (3.2). Thus, surprisingly, the two inherent limitations to self-testing turn out to be not only necessary, but also sufficient in some cases, such as in the case of the CHSH inequality.

For a formalization of the self-testing problem, we follow the general framework proposed by McKague et al. [MYS12], which builds upon the original work by Mayers and Yao and was already used in earlier work [McK14]. This formalization has been used by many others [McK14, MYS12, YVB$^+$14, BP15, BNS$^+$15, CGS16].

The framework is best understood when taking the 'necessary and sufficient' condition from Equation (3.2) as a starting point. Let us first consider the case in which all states are pure. In line with McKague et al., we define a pure state $|\psi\rangle$ to be self-testable using a particular Bell test if the following holds: if a pure state $|\psi\rangle$ has the property that there exist local measurements that yield the same correlations as $|\varphi\rangle$ in the Bell test, then $|\varphi\rangle$ can be obtained from $|\psi\rangle$ by adding pure ancilla states and applying local changes of bases (i.e. local unitaries). Since adding ancillas in a pure state, followed by an application of unitaries does not change inner products, these considerations bring us onto the use of local isometries. Formally, we require the existence of local isometries V_A and V_B and a "junk" state $|\text{junk}\rangle$ such that

$$\big(V_A \otimes V_B\big)(|\varphi\rangle) = |\psi\rangle \otimes |\text{junk}\rangle\,. \tag{3.3}$$

Another motivation for the use of local isometries is the fact that the correlations can be written as inner products $\Pr(a,b|x,y) = \operatorname{Tr}\big((P_a^x \otimes Q_b^y)\rho_{AB}\big)$. As

$$\operatorname{Tr}\big((P_a^x \otimes Q_b^y)\rho_{AB}\big) = \operatorname{Tr}\big((V_A P_a^x V_A^\dagger \otimes V_B Q_b^y V_B^\dagger)(V_A \otimes V_B)\rho_{AB}(V_A^\dagger \otimes V_B^\dagger)\big),$$

we see that that correlations are unaffected by local isometries.

The definition as given in Equation (3.3) extends in a straightforward manner to the case where $|\psi\rangle$ and the added ancillas are mixed.

Definition 3.1. *Let $\mathbb{H}_A, \mathbb{H}_B, \mathbb{H}'_A, \mathbb{H}'_B$ be Hilbert spaces. Let $\psi_{AB} \in \mathrm{D}(\mathbb{H}_A \otimes \mathbb{H}_B)$ be a pure bipartite state. Let $\rho_{A'B'} \in \mathrm{D}(\mathbb{H}'_A \otimes \mathbb{H}'_B)$ be a bipartite state.*

We say that the state $\rho_{A'B'}$ holds *the pure state ψ_{AB} if there exist Hilbert spaces $\mathbb{E}_A$ and $\mathbb{E}_B$, local isometries $V_{A'} : \mathbb{H}'_A \to \mathbb{H}_A \otimes \mathbb{E}_A$ and $V_{B'} : \mathbb{H}'_B \to \mathbb{H}_B \otimes \mathbb{E}_B$ and a 'junk state' $\sigma^{\mathrm{junk}}_{A''B''} \in \mathrm{D}(\mathbb{E}_A \otimes \mathbb{E}_B)$, such that*

$$\big(V_{A'} \otimes V_{B'}\big)\rho_{A'B'}\big(V^\dagger_{A'} \otimes V^\dagger_{B'}\big) = \psi_{AB} \otimes \sigma^{\mathrm{junk}}_{A''B''}.$$

We now say that the pure state ψ_{AB} is self-testable if we can infer from particular correlations in a Bell experiment that the underlying state holds ψ_{AB}.

Definition 3.2. ***(Self-testability)***
We say that the pure state ψ_{AB} is self-testable *if there exist correlations $\Pr(a, b|x, y)$, obtained from a Bell experiment as depicted in Figure 2, from which we can infer that the underlying state $\rho_{A'B'}$ in the Bell experiment holds ψ_{AB}.*

The state $\rho_{A'B'}$ is called 'input state' or 'physical state' and the pure state ψ_{AB} is the 'target state' or 'reference state'. The correlations $\Pr(a, b|x, y)$ as stated above are called 'perfect correlations' or 'perfect statistics'.

In order perform do both state and measurement certification, Definition 3.1 can be extended with the requirement that the measurements used in the Bell experiment act nontrivially on the target state only. Since we only treat state certification in this thesis, we omit this requirement. For the framework for simultaneous state and measurement certification, we refer to McKague et al. [MYS12].

It is not straightforward to extend this definition to the case where the correlations are different from the ideal case. In the next section, we will give an overview of several different formalization of state certification in the presence of errors. The focus will be on formulating *self-testing statements*, which convey the relation between the observed statistics and the 'distance' (in some measure) between the physical system and the reference system. In this thesis, we focus on not using all correlations, but just the Bell value.

3.3 Measures for robust state certification

In practice, it is not possible to achieve the ideal correlations, since any real-world experiment will be subject to noise and the experimental statistical fluctuation of the correlations will never vanish, since the number of runs of the Bell experiment is finite. For these reasons, we want to be able to infer information about the physical state in the case of imperfect statistics too.

In this section, we describe three formalizations of robustness bounds for self-testing statements, where the bound is a function of Bell violation. These three are the MYS-measure, the Mayers-Yao fidelity and the extractability. For each of the formalization, the setup is the same: we start with a Bell inequality $\mathcal{B}$ with quantum value β_Q and classical value $\beta_C < \beta_Q$, and a target state ψ_{AB} which maximally violates $\mathcal{B}$. Moreover, the Bell inequality $\mathcal{B}$ has the property that a state $\rho_{A'B'}$ achieves maximal violation if and only if $\rho_{A'B'}$ holds ψ_{AB}, in the sense of Definition 3.1. The three different approaches given in this section formalize the following question: if a state $\rho_{A'B'}$ achieves Bell value β on $\mathcal{B}$, then how 'close' are $\rho_{A'B'}$ and ψ_{AB}?

This 'closeness' is captured by the robustness 'measures' treated in this section: the MYS-measure F_{MYS}, the Mayers-Yao measure F_{MY} and the extractability Ξ. Each of these measures maps the tuple of input state and target state to a real number in the interval $[0, 1]$. Before treating these three notions in detail, let us give the general framework for self-testing statements using one of these measures.

Definition 3.3. (Self-testing statements in either MYS-measure/Mayers-Yao measure/extractability) *Suppose that the target state ψ_{AB} is self-testable using the perfect statistics that yield maximal violation β_Q to some Bell inequality $\mathcal{B}$. Write β_C for the classical value of $\mathcal{B}$.*

A robust self-testing statement using $\mathcal{B}$ and measure $f \in \{F_{\text{MYS}}, F_{\text{MY}}, \Xi\}$ *is given by a continuous function $g : [\beta_C, \beta_Q] \to [0, 1]$ with the following two properties: first, $\lim_{\beta \to \beta_Q} g(\beta) = 1$, and, moreover, for all states $\rho_{A'B'}$ that violate $\mathcal{B}$ with violation β_{obs}, we have*

$$f(\rho_{A'B'} \to \psi_{AB}) \geq g(\beta_{\text{obs}}).$$

We now treat the three robustness measures in detail. A brief remark on notation: the letter V will be used to refer to isometries and its index will denote the register it acts on. For example, the isometry $V_{AA'}$ acts upon the combined register AA'.

The first formalization, called the MYS-measure in this thesis, follows naturally from the definition of self-testing as given in Definitions 3.1 and 3.2. The name MYS-measure comes from the fact that this formalization was given before (for a pure input state) by McKague, Yang and Scarani [MYS12].

Definition 3.4. *Let ψ_{AB} be a self-testable state. We define the MYS-measure of target state ψ_{AB} from input state $\rho_{A'B'}$ as*

$$F_{\text{MYS}}(\rho_{A'B'} \to \psi_{AB}) := \sup_{V_{A'}, V_{B'}, \sigma^{\text{junk}}_{A''B''}} F\big((V_{A'} \otimes V_{B'})\rho_{A'B'}(V^\dagger_{A'} \otimes V^\dagger_{B'}), \psi_{AB} \otimes \sigma^{\text{junk}}_{A''B''}\big)$$

where the supremum is taken over a state $\sigma^{\text{junk}}_{A''B''}$ of some dimension and local isometries $V_{A'}$ and $V_{B'}$ of appropriate input/output dimensions.

Note that in the MYS-measure, observing perfect statistics (which yield $\beta_{\text{obs}} = \beta_Q$) implies that the input state holds the target state. To see this, note that the MYS-measure equals 1 if and only if there exist isometries $V_{A'}$ and $V_{B'}$ and a state $\sigma^{\text{junk}}_{A''B''}$ such that

$$(V_{A'} \otimes V_{B'})\rho_{A'B'}(V^\dagger_{A'} \otimes V^\dagger_{B'}) = \psi_{AB} \otimes \sigma^{\text{junk}}_{A''B''}.$$

When the statistics are not perfect, the Bell violation β_{obs} is strictly smaller than β_Q; intuitively, the isometry $\Phi = \Phi_A \otimes \Phi_B$ in the MYS-measure maps the input state to the 'closest' target state with auxiliary degrees of freedom.

The MYS-measure has a trivial lower bound. As an example, consider the case where the target state ψ_{AB} is the pure state $\frac{|01\rangle - |10\rangle}{\sqrt{2}}$ and the Bell test used is the CHSH game. Then we can define the local isometries

$$\begin{aligned} V_{A'} &: |\varphi\rangle \to |0\rangle \otimes |\varphi\rangle \\ V_{B'} &: |\varphi\rangle \to |1\rangle \otimes |\varphi\rangle \end{aligned}$$

so that for any input state $\rho_{A'B'}$, we have $(V_{A'} \otimes V_{B'})\rho_{A'B'}(V_{A'}^\dagger \otimes V_{B'}^\dagger) = |01\rangle \otimes \rho_{A'B'}$. By setting the state $\sigma^{\text{junk}}_{A''B''}$ equal to $\rho_{A'B'}$, we get

$$F\big((V_{A'} \otimes V_{B'})\rho_{A'B'}(V_{A'}^\dagger \otimes V_{B'}^\dagger), \psi_{AB} \otimes \rho_{A'B'}\big) = F\big(\,|01\rangle\langle 01| \otimes \rho_{A'B'}, \psi_{AB} \otimes \rho_{A'B'}\big) = \frac{1}{2}.$$

In general, the MYS-measure as given in Definition 3.4 can be trivially lower bounded by $\lambda^2_{\max}(\psi_{AB})$, where $\lambda_{\max}$ denotes the largest Schmidt coefficient.

Definition 3.4 has been used by many authors for self-testing statements: see the next section for a brief overview of the authors that work with this formalization. All proofs used with this definition so far rely on explicit construction of the isometry. The original definition proposed by McKague et al. [MYS12] includes measurement certification too; in fact, each of the proofs resulting in bounds on the MYS-measure first derive constraints on the measurement operators from the Bell value in order to obtain such bounds. When we only wish to perform state certification, finding bounds for certifying state and measurements simultaneously is restrictive: observing a Bell violation $\beta_{\text{obs}} < \beta_Q$ is either due to the fact that the underlying state is noisy, or misalignment of the measurements, or a mixture of the two. For this reason, we might hope to obtain better bounds if we consider state certification and measurement certification separately.

A second formalization of robust self-testing is formulated in terms of what Bardyn et al. proposed as the 'Mayers-Yao fidelity' [BLM+09].

Definition 3.5. *Let ψ_{AB} be a self-testable target state. The* Mayers-Yao fidelity *of ψ_{AB} from input state $\rho_{A'B'}$ is defined as*

$$F_{\text{MY}}(\rho_{A'B'} \to \psi_{AB}) := \sup_{\rho} F(\rho_{A'B'}, \rho) \tag{3.4}$$

where the supremum is taken over bipartite states of the form

$$\rho = (V_{AA''} \otimes V_{BB''})(\psi_{AB} \otimes \sigma^{\text{junk}}_{A''B''})(V_{AA''}^\dagger \otimes V_{BB''}^\dagger) \tag{3.5}$$

where $\sigma^{\text{junk}}_{A''B''}$ is a state of some dimension and $V_{AA''}$ and $V_{BB''}$ are isometries of appropriate input/output dimensions[4].

In the case of self-testing of the singlet, for example, the Mayers-Yao fidelity and the MYS-measure are suited for the black-box scenario in which we wish to find out how much the input state differs from a singlet (up to local unitaries and additional degrees of freedom). In the scenario where we wish to test the state, realized in devices which we bought from a potentially untrusted vendor, we need to take all possible quantum operations into account rather than just isometries. In line with the work of Bardyn et al. [BLM+09] and Kaniewski [Kan16], we define a third robustness measure, the extractability.

Definition 3.6. *Let ψ_{AB} be a self-testable state. The extractability of target state ψ_{AB} from input state $\rho_{A'B'}$ is defined as*

$$\Xi(\rho_{A'B'} \to \psi_{AB}) := \max_{\Lambda_A, \Lambda_B} F((\Lambda_A \otimes \Lambda_B)(\rho_{A'B'}), \psi_{AB}) \tag{3.6}$$

where the maximum is taken over quantum channels of the correct input/output dimensions (called ***extraction channels*** *in this context).*

[4] In the work of Bardyn et al., the supremum is taken over unitaries rather than isometries, which does not yield a fidelity that is defined on all input states. To see this, consider the singlet as target state ψ_{AB}. Then ρ in Equation (3.5) has even local dimension if the maps V_A and V_B were unitaries. Hence if the input state $\rho_{A'B'}$ is of odd local dimension, then the fidelity in Equation (3.4) is not defined.

Just like the MYS-measure, the extractability has the property that observing $\beta_{\text{obs}} = \beta_Q$ (which corresponds to an extractability of 1) implies that the input state $\rho_{A'B'}$ holds the target state ψ_{AB}. We show this.

By definition, we have that $\Xi(\rho_{A'B'} \to \psi_{AB}) = 1$ iff there exist local channels Λ_A, Λ_B for which $F((\Lambda_A \otimes \Lambda_B)(\rho_{A'B'}), \psi_{AB}) = 1$, which is equivalent to

$$(\Lambda_A \otimes \Lambda_B)(\rho_{A'B'}) = \psi_{AB}. \tag{3.7}$$

If $\rho_{A'B'}$ holds ψ_{AB} as in Definition 3.1, then Equation (3.7) holds by picking the isometries from Definition 3.1 as quantum channels. For the converse statement, assume that Equation (3.7) holds for some quantum channels Λ_A, Λ_B. Using Lemma 2.6, we can write

$$\psi_{AB} = \text{Tr}_{A'''} \left(\text{Tr}_{B'''} \left((U_{A'A''} \otimes U_{B'B''})(\rho_{A'B'} \otimes \sigma_{A''} \otimes \sigma_{B''})(U^\dagger_{A'A''} \otimes U^\dagger_{B'B''}) \right) \right) \tag{3.8}$$

where $U_{A'A''}, U_{B'B''}$ are unitaries, $\sigma_{A''}$ and $\sigma_{B''}$ are pure states, A''' is a subsystem of AA'' and B''' is a subsystem of BB''. Since ψ_{AB} is pure, we infer that the RHS of Equation (3.8) before tracing out can be written as

$$(U_{A'A''} \otimes U_{B'B''})(\rho_{A'B'} \otimes \sigma_{A''} \otimes \sigma_{B''})(U^\dagger_{A'A''} \otimes U^\dagger_{B'B''}) = \psi_{AB} \otimes \tau_{A'''B'''}.$$

Since $\sigma_{A''}$ is pure and $U_{A'A''}$ is a unitary, the map $\rho \mapsto U_{A'A''}(\rho \otimes \sigma_{A''})U^\dagger_{A'A''}$ is an isometry, and for similar reasons the map $\rho \mapsto U_{B'B''}(\rho \otimes \sigma_{B''})U^\dagger_{B'B''}$ is an isometry too. Therefore $\rho_{A'B'}$ holds ψ_{AB} in the sense of Definition 3.1. This concludes the proof that the extractability equals 1 if and only if $\rho_{A'B'}$ holds ψ_{AB}.

Bardyn et al. note that the extractability is also a relevant notion in practice when we want to establish how well we can improve a purchased source, by applying local operations before using the source.

The extractability has the same trivial lower bound as the MYS-measure. The most natural one for the extractability can be achieved when Alice and Bob discard their shares and replace their part by a fixed state. This corresponds to applying full amplitude damping channels locally. When Alice and Bob do so, the best amplitude damping channel they can apply yields $F((\Lambda_A \otimes \Lambda_B)(\rho_{A'B'}), \psi_{AB}) = \lambda^2_{\max}$, where $\lambda_{\max}$ is the maximal Schmidt coefficient of ψ_{AB}. So, just like for the MYS-measure, the extractability of the singlet is trivially lower bounded by $\frac{1}{2}$.

Since any isometry is a quantum channel, we see that the MYS-measure is related to the extractability as

$$F_{\text{MYS}}(\rho_{A'B'} \to \psi_{AB}) \leq \Xi(\rho_{A'B'} \to \psi_{AB})$$

for any input state $\rho_{A'B'}$ and target state ψ_{AB}. Thus the extractability is more "forgiving" than the MYS-measure.

We have already seen that the MYS-measure and the extractability share the same lower bound. We finish this section by noting two additional properties that both measures have in common: they are convex in the input state and moreover, their infimum over all input states with Bell value at least β can be upper bounded as a function of β.

Lemma 3.7. *Let $f(\rho_{A'B'} \to \psi_{AB})$ be either the extractability or MYS-measure of input state $\rho_{A'B'}$ and target state ψ_{AB}. Let the notation be as in the definition of a robust self-testing statement for ψ_{AB} using Bell inequality $\mathcal{B}$ (Definition 3.3). Then the following hold:*

1. *If* $\rho_{A'B'} = p\sigma_{A'B'} + (1-p)\tau_{A'B'}$ *for bipartite states* $\sigma_{A'B'}$ *and* $\tau_{A'B'}$ *and some* $p \in [0,1]$*, then*

$$f(\rho_{A'B'} \to \psi_{AB}) \leq p \cdot f(\sigma_{A'B'} \to \psi_{AB}) + (1-p) \cdot f(\tau_{A'B'} \to \psi_{AB}).$$

2. *For every* $\beta \in [\beta_C, \beta_Q]$*, all bipartite states* $\sigma_{A'B'}$ *which achieve Bell value at least* β *are upper bounded as*

$$f(\sigma_{A'B'} \to \psi_{AB}) \leq \lambda_{\max} + (1 - \lambda_{\max}) \cdot \frac{\beta - \beta_C}{\beta_Q - \beta_C}$$

where $\lambda_{\max}$ *denotes the largest Schmidt coefficient of* ψ_{AB}*.*

The first property is a direct consequence of the definition of extractability and subadditivity of the maximum. The second statement, where f is the extractability, is proven by considering a convex combination of the target state ψ_{AB} and a separable state (for details, we refer to Kaniewski [Kan16]). The proofs for the MYS-measure are analogous.

In the next section, we give an overview of the states that have been shown to be self-testable and which of the robustness measures as described in this section were used.

3.4 Self-testable states and methods to prove their self-testability

In the past decade, a couple of different methods have been applied to derive robust self-testing statements for a variety of states. We name some of them.

3.4.1 Self-testing statement from algebraic relations on the observables

Most of the self-testing statements cited in this chapter were derived by first proving some kind of algebraic relations on the observables.

Mayers and Yao studied the setup of two devices containing a bipartite state [MY04]. They considered the ideal case in which each device has a qubit, which are maximally entangled, and each device has three measurement settings: projective measurements at an angle $\{-\frac{\pi}{8}, 0, \frac{\pi}{8}\}$. This configuration yields nine (three settings on each side) probability distributions $\Pr(a, b|x, y)$ of settings x, y and outcomes a, b. Mayers and Yao showed that this setup is the *only* setup that gives rise to these probability distributions, up to auxiliary degrees of freedom and local changes of basis. As such, the observed correlations can be used to "self-check" the devices. The proof consists of a series of lemmas that are aimed at proving properties of the behaviour of (functions of) the observables when applied to the state. Mayers and Yao only consider the ideal case and do not provide any robustness bounds.

The first robust self-testing bound for the singlet with the CHSH inequality was given by McKague et al. [MYS12]. The approach is applied to self-testing the singlet, first with the CHSH inequality and then with the Mayers-Yao measure. McKague et al. provide state certification and measurement certification in one go, formulated in terms of the MYS-measure. Their proofs consist of an explicit construction of the isometry from Definition 3.4 as a function of the physical observables, and subsequently use commutation relations of the observables to prove their bounds in the MYS-measure.

Exactly the same kind of proof structure was used before to self-test graph states [McK14], and used in later work to prove the self-testing of many singlets in parallel [McK16]. Robust self-testing statements of all partially entangled two-qubit states with the use of the tilted CHSH inequality (the topic of Chapter 5) was proven using this method too: this was done by Bamps and Pironio [BP15], building upon the work of others [AMP12, YN13]. Building on the work of Yang and Navascues [YN13], it was

recently proven that all pure bipartite states can be self-tested [CGS16], again using the approach of explicit construction of the isometry. The latter result does not include robustness bounds.

What all these results have in common, is that the achieved robustness is very weak; the bounds become trivial for errors of the order of magnitude of $\beta_Q - \beta_{\text{obs}} \approx 10^{-4}$, where β_Q is the quantum value of the Bell inequality used in the proofs and β_{obs} is the observed Bell violation. This work focused on proving self-testability of states, but for robust state-certification that is relevant to real-world experiments, we will have to use a different approach. As argued before, the MYS-measure might be too strict for either state certification or measurement certification at the same time, whereas approaching these separately might yield better results. Indeed, a new method for measurement certification was recently proposed [Kan17].

3.4.2 Experimentally-relevant bounds

Bardyn et al. were the first to analytically prove a number of self-testing statements for the singlet and show bounds on both the Mayers-Yao fidelity as well as on the extractability [BLM$^+$09]. The bounds on the extractability were improved using a numerical method, the SWAP method, which was developed by Bancal et al. [BNS$^+$15]. Intuitively, the isometry in the definition of self-testability, Definition (3.2), locally *swaps* the right state, encoded in the input state, into ancilla qubits. Using this isometry, the SWAP method lower bounds the extractability by numerically optimizing over all correlations from a superset of the quantum set.

Bancal et al. apply the SWAP method to construct robust self-testing bounds for several states, among which the singlet using the Mayers-Yao correlations and any partially entangled two-qubit state using the tilted CHSH inequality [BNS$^+$15]. In later work, the SWAP method was applied for robust self-testing of particular partially-entangled qutrits [YVB$^+$14], the three-qubit W state [WCY$^+$14], the three- and four-qubit GHZ state and the four-qubit linear cluster state [PVN14] and the maximally entangled state of two qutrits [SAT$^+$16]. In all of these cases, the robustness is practically relevant.

The SWAP method is thus a versatile tool for finding self-testing statements of practically relevant robustness. However, since its computational cost grows fast when increasing the dimension of the target state, all applications of the method have so far been restricted to states of at most four qubits or two qutrits.

So far we have distinguished two main approaches for robust self-testing: one which is aimed at proving self-testing statements for families of states of arbitrarily large dimension. The corresponding proofs are all based on proving bounds of the behaviour of the local measurements and allow only for very small errors, which are too weak for use in practice. On the other hand, there are a few analytic results by Bardyn et al. with experimentally-relevant robustness, which were improved by Bancal et al. using the SWAP method, a numerical method that yields experimentally robust self-testing statements, but the size of the states that the method could handle so far is small (up to approximately four qubits) due to rapidly growing computational cost.

One way to get rid of the latter problem, is to find a new analytic method. This was done recently by Kaniewski, who derived analytic self-testing bounds whose robustness is also experimentally relevant [Kan16]. The next section is devoted to this method.

3.5 Self-testing from operator inequalities

The results in Chapter 5 build upon previous work by Kaniewski [Kan16], who showed that bounds on the extractability can be derived from certain operator inequalities. This section is devoted to this derivation. First, we show how to obtain self-testing statements from operator inequalities. Then, we show how such operator inequalities can be proven.

3.5.1 From operator inequality to self-testing statement

Let ψ_{AB} be a pure target state and $\mathcal{B}$ a Bell inequality. For simplicity, we treat the bipartite case here rather than the general multipartite case. We denote the Bell operator of $\mathcal{B}$ by W; the Bell operator can be written as (see Section 2.2):

$$W = \sum_{x,y,a,b} c_{x,y,a,b} P_a^x \otimes Q_b^y \tag{3.9}$$

where the $c_{x,y,a,b}$ are real constants and the P_a^x (Q_b^y) are the measurement operators for Alice (Bob) on input x (y).

We will derive an operator inequality that yields a lower bound on the extractability. For completeness, we state the extractability from Definition 3.6 again: the extractability of ψ_{AB} from $\rho_{A'B'}$ is defined as

$$\Xi(\rho_{A'B'} \to \psi_{AB}) = \max_{\Lambda_A, \Lambda_B} F\big((\Lambda_A \otimes \Lambda_B)(\rho_{A'B'}), \psi_{AB}\big)$$

where Λ_A and Λ_B are the extraction channels.

Now let us start the derivation with a particular operator inequality. Let Λ_A, Λ_B be local extraction channels (see Definition 3.6), that **only depend** on Alice's and Bob's local measurement operators P_a^x and Q_b^y, respectively (but **not** on the input state). Furthermore, define

$$K := (\Lambda_A^\dagger \otimes \Lambda_B^\dagger)(\psi_{AB}). \tag{3.10}$$

where $\Lambda^\dagger$ refers to the dual channel of quantum channel Λ. Now presume that for some fixed real parameters s and μ the operator inequality

$$K \geq sW + \mu \mathbb{1} \tag{3.11}$$

holds *for all possible measurements* (recall that W and the extraction channels are functions of the measurement operators P_a^x for Alice and Q_b^y for Bob). Then we can take the trace with the input state $\rho_{A'B'}$ on both sides of Inequality (3.11). Expanding the left hand side yields

$$\mathrm{Tr}\big(K\rho'_{AB}\big) = \langle (\Lambda_A^\dagger \otimes \Lambda_B^\dagger)(\rho_{AB}), \rho'_{AB} \rangle = \langle \rho_{AB}, \big(\Lambda_A \otimes \Lambda_B\big)(\rho'_{AB}) \rangle \tag{3.12}$$

and computing the right hand side of Inequality (3.11) boils down to

$$\mathrm{Tr}\big(sW\rho'_{AB} + \mu \mathbb{1}\rho'_{AB}\big) = s\beta + \mu \tag{3.13}$$

By combining Equations (3.12) and (3.13), we get

$$F\big((\Lambda_A \otimes \Lambda_B)(\rho_{A'B'}), \psi_{AB}\big) \geq s\beta + \mu$$

which holds for all possible measurement operators, since the operator inequality that we started with, Equation (3.11), did so too.

Using the definition of extractability, Definition 3.6, we obtain

$$\Xi(\rho_{A'B'} \to \psi_{AB}) \geq s\beta + \mu$$

which expresses that the extractability of ψ_{AB} from $\rho_{A'B'}$ is lower bounded by a linear function of the observed Bell violation: this is precisely a self-testing statement.

Proving the operator inequality given in Equation (3.11) is hard in general, since we made no assumptions about the dimension of the Hilbert space that the measurement operators act upon. Fortunately, the operator inequality can be proven in some cases, which can be seen in the next section.

3.5.2 How to prove such operator inequalities

We show that if the inputs and outputs of both Alice and Bob in the Bell experiment corresponding to $\mathcal{B}$ are binary, then proving the operator inequality from Equation (3.11) becomes tractable. As in Section 2.2.3, denote the observables of Alice and Bob by $A_x := P_0^x - P_1^x$ and $B_y := Q_0^y - Q_1^y$ (for $x, y \in \{0, 1\}$), which are applied to the shared state $\rho_{A'B'}$.

First, note that if the measurements that Alice and Bob apply are not projective, then Alice and Bob could add ancillas to make them so as first part of the extraction procedure.

Second, we can use Jordan's lemma (see Lemma 2.7) to further ease the analysis. By applying Jordan's lemma to Alice's observables A_0 and A_1, we see that

$$A_x = \bigoplus_m A_x^m \tag{3.14}$$

with $x \in \{0, 1\}$ and A_x^m has dimensions 2×2 at most. As part of the extraction procedure, Alice can embed her part of the state ρ_{AB} into a larger Hilbert space, in order to make all blocks of size 2×2.

A similar expression can be given for Bob's observables:

$$B_y = \bigoplus_n B_y^n \tag{3.15}$$

for $y \in \{0, 1\}$, where, for the same reason as given above, we can assume that the block B_y^n are of size 2×2.

Since the outcomes of the players are binary, the observables A_x and B_y obtain all the information needed to retrieve the measurement operators (see also Section 2.1.4). In particular, we can write the Bell operator W from Equation (3.9) in terms of the observables. Then, using Equations (3.14) and (3.15), we see that we can write W as

$$W = \bigoplus_{m,n} W^{mn}$$

where W^{mn} is of size 4×4. One can think of W^{mn} as a Bell operator where the underlying state consists of two qubits. As an example, suppose that W is the CHSH operator $W = \sum_{x=0,y=0}^{1} (-1)^{x \cdot y} A_x \otimes B_y$, then $W^{mn} = \sum_{x=0,y=0}^{1} (-1)^{x \cdot y} A_x^m \otimes B_y^n$.

A similar reasoning can be given for the operator K from Equation (3.10). We can now write

$$K - sW - \mu \mathbb{1} = \bigoplus_{mn} (K^{mn} - sW^{mn} - \mu \mathbb{1}_4)$$

where the matrices K^{mn} and W^{mn} have sizes 4×4. Now if for every block ,the operator inequality

$$K_{mn} \geq sW_{mn} + \mu \mathbb{1}_4 \tag{3.16}$$

holds, then the 'big' operator inequality (3.11) holds too.

Note that the inequality (3.16) holds for every block if we can prove the inequality *for all possible qubit observables*. We are free to choose the basis in which we prove the operator inequality (3.16); since the only property of a pair of qubit-observables A_0, A_1 that is invariant under unitary conjugation is the angle between them, we can write

$$A_r := \cos(a)\mathsf{X} + (-1)^r \sin(a)\mathsf{Z} \tag{3.17}$$

for $a \in [0, \frac{\pi}{2}]$. A similar expression holds for Bob's observables.

In short, we have seen in this section that in a Bell experiment where the players' input and output are both binary, we can derive robust self-testing bounds on the extractability from the operator inequality (3.11).

Using this method, Kaniewski showed that given CHSH violation β, the singlet extractability can be lower bounded by $\frac{1}{2} + \frac{1}{2} \cdot \frac{\beta - \beta^*}{2\sqrt{2} - \beta^*}$, where the *threshold violation* $\beta^* = \frac{16+14\sqrt{2}}{17} \approx 2.11$ is the smallest violation at which the bound becomes non-trivial. At the moment of writing, this is the best lower bound that has been established for self-testing the singlet with the CHSH inequality.

The local extraction channels that were used for this bound are dephasing channels, where the dephasing parameter depends on the angle between the observables. In order to derive new self-testing bounds for (almost) all pure partially entangled two-qubit states (see Chapter 5), we use the same type of dephasing channels, with slight adaptations.

4 CHSH violation does not imply nontrivial singlet extractability

Let us note a few properties of two sets of states, namely all separable states and all states that are equivalent to a singlet (i.e. are a singlet up to auxiliary degrees of freedom and local unitaries), in relation to the CHSH inequality and self-testing. We have seen that for these two sets, singlet extractability and CHSH violation go hand in hand: first, separable states have trivial singlet extractability and do not violate the CHSH inequality. States that are equivalent to a singlet, on the other hand, are the only states with the following two properties: they have singlet extractability equal to 1 (the maximal possible value) and, moreover, they violate the CHSH inequality maximally for the right choice of measurements.

Robust self-testing statements obtained from operator inequalities, as described in Section 3.5, were shown before to yield a threshold violation of $\beta^* \approx 2.11$ [Kan16]. That is, we can extract a nontrivial singlet from any state that achieves CHSH violation greater than β^*. By improving the local extraction channels that yielded these robustness bounds, we attempted to 'close the gap', in order to obtain a threshold violation of 2 (since separable states yield a CHSH violation of 2, obtaining $\beta^* < 2$ is impossible). Unfortunately, numerous attempts failed - in fact, all channels we considered performed considerably worse than the dephasing channels used to obtain $\beta^* \approx 2.11$.

The fact that none of our attempts came close to the performance of the dephasing channels, led us to conjecture that the dephasing channels are optimal (although we have not found a proof). Secondly, it brought us to reconsider whether it was possible to 'close the threshold gap' at all. It felt natural to ask about the intermediate regime: does **every** Bell violation $\beta > 2$ imply nontrivial singlet extractability? In this section, we answer this question in the negative.

In what follows, we construct a state with the following two properties: (a) the state violates the CHSH inequality; (b) the singlet extractability of the state, as defined in Definition 3.6, is trivial. The latter property means that for any pair of local extraction channels Λ_A and Λ_B, we have $F\big((\Lambda_A \otimes \Lambda_B)(\rho_{\mathrm{TE}}), \Phi_+\big) \leq \frac{1}{2}$, where ρ_{TE} is the state we constructed (with $\underline{\text{t}}$rivial $\underline{\text{e}}$xtractability) and Φ_+ is a maximally entangled two-qubit state.

The existence of such a state is surprising. Entanglement is necessary for Bell violation in general, and for CHSH violation in particular, a state even needs to be NPT entangled (which could be thought of as a "strong" form of entanglement, see also Section 2.2.6). On top of this, maximal CHSH violation is only possible with a state that 'contains' a maximally entangled two-qubit state. It is therefore intuitive that a state that violates the CHSH inequality should also be 'close' to a singlet, and this closeness should be detectable with a measure such as the extractability. Especially since the extractability involves quantum channels, we allow for the most general local (i.e. without communication) processing of the state. In spite of all these considerations, this intuition remarkably turns out to be incorrect: it is a surprising fact that there exist entangled states that violate the CHSH inequality but although Alice and Bob can apply any quantum operation, the fidelity of their state with the singlet will be at most as good as discarding their parts and replacing it by a fixed separable state.

Although entanglement is necessary for CHSH violation, the converse does not hold: Werner states (see Section 2.2.5) are entangled but do not violate any Bell inequality. Our result shows a similar property, in the relation between entanglement and singlet extractability rather than CHSH violation and entanglement: although nontrivial singlet extractability implies entanglement (see Lemma 2.14), the converse is proven incorrect by our counterexample.

This section is ordered as follows. We will first set up the conventions; that is, the specific maximally entangled two-qubit state that we use and the observables, two on each side, we select. Then we define the state $\rho_{\mathrm{TE}} \equiv \rho_{\mathrm{TE}}(\nu)$ as a function of a single parameter $\nu \geq 0$. We show that, with the observables as chosen before, the state $\rho_{\mathrm{TE}}(\nu)$ violates the CHSH inequality by construction if $\nu > 0$. The rest of the chapter proves that there exists a $\nu > 0$ indeed for which the fidelity of $(\Lambda_A \otimes \Lambda_B)(\rho_{\mathrm{TE}}(\nu))$ with the singlet is at most $\frac{1}{2}$ for any choice of extraction channels Λ_A and Λ_B. This concludes the proof that there exists a state that violates the CHSH inequality but nevertheless has trivial singlet extractability.

4.1 The target state

The CHSH inequality is maximally violated by all maximally entangled states of two qubits. Since all such states differ only by local unitary conjugation (see Section 2.1.3), Alice and Bob could always first apply their local unitaries before applying their local extraction channels. Hence the singlet extractability of any input state is unaffected by the choice of maximally entangled two-qubit state used as target state. For our proof, it will turn out to be convenient to work with a particular maximally entangled two-qubit state, which we define as follows.

Let Φ be the pure density matrix of $\frac{|00\rangle+|11\rangle}{\sqrt{2}} = (\mathbb{1}_2 \otimes (\mathsf{ZX}))\Big(\frac{|01\rangle-|10\rangle}{\sqrt{2}}\Big)$. Define the unitary U by

$$U := \cos\Big(\frac{3\pi}{8}\Big)\mathsf{Z} + \sin\Big(\frac{3\pi}{8}\Big)\mathsf{X}. \tag{4.1}$$

Using this unitary, we compute

$$\Phi_+ := (\mathbb{1} \otimes U)\Phi(\mathbb{1} \otimes U^\dagger) = \frac{1}{4}\Big(\mathbb{1}_2 \otimes \mathbb{1}_2 + \mathsf{Y} \otimes \mathsf{Y} + \mathsf{X} \otimes \frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}} + \mathsf{Z} \otimes \frac{\mathsf{X}-\mathsf{Z}}{\sqrt{2}}\Big]\Big) \tag{4.2}$$

Note that both $\mathbb{1}_2 \otimes (\mathsf{ZX})$ and $\mathbb{1}_2 \otimes U$ are unitaries. Hence, since Φ is a maximally entangled state and Φ_+ only differs from Φ by local unitary conjugation, so is Φ_+. In what follows, the state Φ_+ is the one we wish to extract.

By defining

$$V := \cos\Big(\frac{3\pi}{16}\Big)\mathbb{1} - i\sin\Big(\frac{3\pi}{16}\Big)\mathsf{Y} \tag{4.3}$$

it is straightforward to compute that

$$\Phi_+ \text{ is the pure density matrix of } \frac{1}{\sqrt{2}}(V \otimes V)(|00\rangle - |11\rangle).$$

Hence the symmetric Schmidt decomposition of Φ_+ is given by $\frac{1}{\sqrt{2}}\big(V|0\rangle \otimes V|0\rangle - V|1\rangle \otimes V|1\rangle\big)$.

4.2 The observables

The observables of Alice and Bob are operators acting on $\mathbb{C}^3 \times \mathbb{C}^2$. We define Alice's and Bob's observables as follows:

$$\begin{aligned} A_0 &= B_0 &&= \Big(|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|\Big) \otimes \mathsf{X} \\ A_1 &= B_1 &&= |0\rangle\langle 0| \otimes \mathsf{X} + |1\rangle\langle 1| \otimes \mathsf{Z} + |2\rangle\langle 2| \otimes (-\mathsf{X}). \end{aligned} \tag{4.4}$$

The CHSH operator W, which depends on the observables of Alice and Bob, has been defined in Equation (2.11). We repeat it here for completeness:

$$W = \sum_{j,k=0}^{1} (-1)^{j \cdot k} A_j \otimes B_k$$

Note that W can be written as

$$W = \sum_{x=0,y=0}^{2} |x\rangle\langle x|_A \otimes |y\rangle\langle y|_B \otimes W_{xy} \tag{4.5}$$

where the W_{xy} can be ordered in a diagram, as depicted in Table (4.6).

Bob	2	$2\mathsf{X}\otimes\mathsf{X}$	$2\mathsf{Z}\otimes\mathsf{X}$	$-2\mathsf{X}\otimes\mathsf{X}$
	1	$2\mathsf{X}\otimes\mathsf{X}$	$\mathsf{X}\otimes(\mathsf{X}+\mathsf{Z})+\mathsf{Z}\otimes(\mathsf{X}-\mathsf{Z})$	$2\mathsf{X}\otimes\mathsf{Z}$
	0	$2\mathsf{X}\otimes\mathsf{X}$	$2\mathsf{X}\otimes\mathsf{X}$	$2\mathsf{X}\otimes\mathsf{X}$
		0	1	2
			Alice	

(4.6)

Throughout the proof, we will frequently refer to the different choices of x and y. It will be convenient to use the visual representation from Table (4.6) for these choices. We will refer to the point $x = y = 1$ as 'the center', and to the set of the remaining (x, y) as 'the frame'. We denote the center by $\mathcal{C} := \{(1, 1)\}$ and the frame by $\mathcal{F} := \{0, 1, 2\}^2 - \mathcal{C}$.

4.3 The input state ρ_{TE}

We define the input state ρ_{TE} as

$$\rho_{\mathrm{TE}} := \sum_{x,y=0}^{2} p_{xy} \, |x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{xy} \tag{4.7}$$

where the states ρ_{xy} are given by

$$\rho_{xy} := \begin{cases} \Phi_+ & \text{if } x = y = 1 \text{ ("the center")} \\ \frac{1}{4}(\mathbb{1}\otimes\mathbb{1} + \frac{W_{xy}}{2}) & \text{otherwise ("the frame")} \end{cases}$$

where Φ_+ has been defined as in Equation (4.2). We denote $\nu := p_{11}$, and the remaining weights p_{xy} for $(x, y) \in \mathcal{F}$ are given by

$$p_{xy} = \frac{c_{xy}}{41} \cdot (1 - \nu)$$

where the c_{xy}'s are defined as

y	2	$c_{02} = 8$	$c_{12} = 0$	$c_{22} = 8$
	1	$c_{01} = 3$		$c_{21} = 0$
	0	$c_{00} = 11$	$c_{10} = 3$	$c_{20} = 8$
		0	1	2
			x	

(4.8)

Note that for $(x, y) \in \mathcal{F}$, the states ρ_{xy} are separable and ρ_{11} is a singlet, up to local unitary conjugation. We therefore have $\mathrm{Tr}(W_{xy}\rho_{xy}) = 2$ for $(x, y) \in \mathcal{F}$ with the W_{xy} defined as in Table (4.6), while for $x = y = 1$, we obtain $\mathrm{Tr}(W_{11}\rho_{11}) = 2\sqrt{2}$. By linearity, the Bell violation of ρ_{TE} is $2 + (2\sqrt{2} - 2)\nu$. Therefore the state ρ_{TE} violates the CHSH inequality if and only if $\nu > 0$. At the end of this chapter, we will set the weight ν to some strictly positive value to ensure trivial extractability.

4.4 Amplitude-damping channel

We already noted that the singlet extractability is trivially lower bounded by $\frac{1}{2}$, since Alice and Bob can always choose to ignore their shared state and locally replace it by a fixed state. Replacing a state corresponds to completely amplitude-damping channels. Since the Schmidt decomposition of any maximally-entangled state of two qubits is not unique because it has Schmidt coefficient $\frac{1}{2}$ with degeneracy two, there any many choices of local bases to write the state in the form of Corollary 2.4: if we write the state in the form

$$\frac{|a_0\rangle \otimes |b_0\rangle + |a_1\rangle \otimes |b_1\rangle}{\sqrt{2}}$$

where $\{|a_0\rangle, |a_1\rangle\}$ and $\{|b_0\rangle, |b_1\rangle\}$ are bases for Alice's and Bob's qubit, respectively, then a natural completely amplitude-damping channel that achieves singlet fidelity $\frac{1}{2}$ is damping to $|a_0\rangle$ ($|a_1\rangle$) for Alice and damping to $|b_0\rangle$ ($|b_1\rangle$) for Bob. In the rest of this chapter, it will be convenient to be able to refer to the amplitude-damping channels that achieve singlet fidelity $\frac{1}{2}$. For this reason, we choose one of such channels for our target state from Equation (4.2) and describe it here explicitly.

If Alice and Bob share the state $\frac{|00\rangle - |11\rangle}{\sqrt{2}}$, then they can obtain the target state Φ_+ as given in Equation (4.2) by both applying the unitary V from Equation (4.3) to their part of the state. Now write $|\varphi\rangle := V|0\rangle$ and $|\psi\rangle := V|1\rangle$, so that we can write the target state from Equation (4.2) in its symmetric Schmidt decomposition as

$$\Phi_+ = \frac{1}{2}\big(|\varphi\rangle|\varphi\rangle - |\psi\rangle|\psi\rangle\big)\big(\langle\varphi|\langle\varphi| - \langle\psi|\langle\psi|\big).$$

Then define the amplitude-damping channel that shrinks the entire Bloch sphere to the state $|\varphi\rangle\langle\varphi|$, given by

$$\Lambda_{\mathrm{AD},\varphi} : \rho \mapsto E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \tag{4.9}$$

where the Kraus operators E_0, E_1 are given by

$$E_0 := |\varphi\rangle\langle\varphi| \quad , \quad E_1 := |\varphi\rangle\langle\psi|.$$

It is straightforward to compute that $\Lambda_{\mathrm{AD},\varphi}(\rho) = |\varphi\rangle\langle\varphi|$ for **any** single-qubit state ρ. Hence for any two-qubit state ρ_{AB} that Alice and Bob share, we have $F\big((\Lambda_{\mathrm{AD},\varphi} \otimes \Lambda_{\mathrm{AD},\varphi})(\rho_{AB}), \Phi_+\big) = \frac{1}{2}$.

4.5 Main result and proof outline

For completeness, we state the trivial extractability of ρ_{TE} as a theorem.

Theorem 4.1. *There exists a $\nu > 0$ such that the state ρ_{TE}, as defined in Equation (4.7) violates the CHSH inequality, but nevertheless has trivial singlet extractability. To be precise: let W be the CHSH operator as given in Equation (2.11) where the observables are stated in Equation* (4.4)*, let Φ_+ be the maximally entangled state from Equation (4.2), and let Λ_A, Λ_B be quantum channels. Then there exists a $\nu > 0$ such that the state $\rho_{\mathrm{TE}} \equiv \rho_{\mathrm{TE}}(\nu)$ satisfies both*

$$\mathrm{Tr}(W\rho_{\mathrm{TE}}) > 2 \quad \textit{and} \quad \max_{\Lambda_A,\Lambda_B} F((\Lambda_A \otimes \Lambda_B)(\rho_{\mathrm{TE}}), \Phi_+) = \frac{1}{2}.$$

Our goal is to upper bound the fidelity $F((\Lambda_A \otimes \Lambda_B)(\rho), \Phi_+)$ by $\frac{1}{2}$, for any choice of extraction channels Λ_A,Λ_B. Using Lemma (2.12), this fidelity can be expressed in terms of a set of six qubit-to-qubit channels: three channels Λ_A^x for Alice, for each choice of $x \in \{0,1,2\}$, and similarly three channels Λ_B^y for Bob, with $y \in \{0,1,2\}$. To see this, recall that $F(\sigma_1, \sigma_2) = \langle \sigma_1, \sigma_2 \rangle$ when at least one of σ_1, σ_2 is pure, where $\langle X, Y \rangle := \mathrm{Tr}\big(Y^\dagger X\big)$. Since Φ_+ is a pure state, we can apply Lemma (2.12) to obtain six qubit-to-qubit quantum channels Λ_A^x, Λ_B^y for $x, y \in \{0,1,2\}$, such that

$$\begin{aligned} F((\Lambda_A \otimes \Lambda_B)(\rho_{\mathrm{TE}}), \Phi_+) &= \langle (\Lambda_A \otimes \Lambda_B)(\rho), \Phi_+ \rangle \\ &= \langle (\Lambda_A \otimes \Lambda_B)(\sum_{x,y=0}^{2} p_{xy} \, |x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{xy}), \Phi_+ \rangle \\ &= \sum_{x,y=0}^{2} p_{xy} \langle (\Lambda_A^x \otimes \Lambda_B^y)(\rho_{xy}), \Phi_+ \rangle. \end{aligned} \tag{4.10}$$

Since the states on the frame are separable, their singlet extractability is trivial. Hence for (x, y) on the frame, the values for $\langle (\Lambda_A^x \otimes \Lambda_B^y)(\rho_{xy}), \Phi_+ \rangle$ are upper bounded by $\frac{1}{2}$. We have seen previously that the amplitude-damping channels described in Section (4.4) achieve this upper bound. We will show that amplitude-damping channels are in fact *the only* channels achieving this bound, and moreover, that this also holds approximately. In particular, we show that there is a *trade-off* between the fidelity with the singlet that the extraction channels can achieve on the frame on the one hand, and in the center on the other. That is, if Alice and Bob apply extraction channels that achieve almost-optimal performance on all points on the frame, then the fidelity in the center cannot be much greater than $\frac{1}{2}$. Using Equation (4.10), we then pick the weight ν such that the fidelity with the singlet is upper bounded by $\frac{1}{2}$, for any choice of local extraction channels.

Our proof is divided into several steps. We first show (in Lemma 4.2) that if the local extraction channels Λ_A^x and Λ_B^y for (x, y) on the frame achieve fidelity close to $\frac{1}{2}$, then the extraction channels in the center cannot differ too much from the amplitude-damping channels defined in Section (4.4).

If the extraction channels in the center are approximately amplitude-damping channels that damp to a pure product state, they must, by linearity, shrink the entire Bloch sphere to a tiny volume close to the boundary. This point is made rigorous in Lemma 4.3. Note that this tiny volume cannot be too far off from the product state that the amplitude-damping channels from Section (4.4) damp towards. This observation leads us to Lemma 4.4: we show that the fidelity of any state in this tiny volume with singlet state from Equation (4.2) cannot be much greater than $\frac{1}{2}$. We then obtain an upper bound on the singlet extractability with the use of Equation (4.10). This upper bound allows us to pick the weight ν in the center such that the singlet extractability of our input state ρ becomes trivial.

4.6 Notation

Before starting the proof, we introduce some extra notation. Define

$$\sigma_x^{\pm} := \Lambda_A^x(|\pm\rangle\langle\pm|)$$

$$\tau_y^{\pm} := [U^\dagger(\Lambda_B^y(|\pm\rangle\langle\pm|))U]^T$$

where the unitary U has been defined in Equation (4.1). Now we use the fact that $\mathrm{Tr}(A \otimes B\Phi) = \frac{1}{2}\mathrm{Tr}\big(AB^T\big)$ for any 2×2 linear operators A and B, where Φ is the pure density matrix of $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$, as defined in Section 4.1. Using the relation between the states Φ and Φ_+ from Equation (4.2), we now

obtain

$$\langle(\Lambda_A^x \otimes \Lambda_B^y)(|{+}{+}\rangle\langle{+}{+}|), \Phi_+\rangle = \langle\sigma_x^+ \otimes (U(\tau_y^+)^T U^\dagger), \Phi_+\rangle = \langle\sigma_x^+ \otimes (\tau_y^+)^T, \Phi\rangle = \frac{1}{2}\operatorname{Tr}\big(\sigma_x^+\tau_y^+\big) \quad (4.11)$$

and, similarly,

$$\langle(\Lambda_A^x \otimes \Lambda_B^y)(|{+}{-}\rangle\langle{+}{-}|), \Phi_+\rangle = \frac{1}{2}\operatorname{Tr}\big(\sigma_x^+\tau_y^-\big), \quad (4.12)$$

$$\langle(\Lambda_A^x \otimes \Lambda_B^y)(|{-}{+}\rangle\langle{-}{+}|), \Phi_+\rangle = \frac{1}{2}\operatorname{Tr}\big(\sigma_x^-\tau_y^+\big), \quad (4.13)$$

$$\langle(\Lambda_A^x \otimes \Lambda_B^y)(|{-}{-}\rangle\langle{-}{-}|), \Phi_+\rangle = \frac{1}{2}\operatorname{Tr}\big(\sigma_x^-\tau_y^-\big). \quad (4.14)$$

The states ρ_{xy} on the frame are separable, hence their singlet extractability is at most $\frac{1}{2}$. We now define the differences from this upper bound as

$$\varepsilon_{xy} := \frac{1}{2} - \langle(\Lambda_A^x \otimes \Lambda_B^y)(\rho_{xy}), \Phi_+\rangle \quad \text{for } (x,y) \in \mathcal{F}. \quad (4.15)$$

In the remainder of the chapter, we show that the singlet extractability of the state in the center is upper bounded by $\frac{1}{2} + \mathcal{O}(\varepsilon_{\text{wav}})$, where ε_{wav} is a particular weighted average of the ε_{xy} on the frame, defined as

$$\varepsilon_{\text{wav}} := \frac{1}{41}\sum_{(x,y)\in\mathcal{F}} c_{xy}\varepsilon_{xy} = \frac{1}{1-\nu}\sum_{(x,y)\in\mathcal{F}} p_{xy}\varepsilon_{xy} \quad (4.16)$$

where the coefficients c_{xy} are given in Table (4.8).

4.7 Upper bounding the extractability in the center as a function of ε_{wav}

The first step is to show that, when the ε_{xy} are small, the channels of Alice and Bob for $x = y = 1$ (denoted by Λ_A^1 and Λ_B^1) are 'close' to the amplitude-damping channels from Section (4.4). To do so, we show that $\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)$ and $\operatorname{Tr}\big(\tau_1^+\tau_1^-\big)$ cannot be much smaller than 1.

Lemma 4.2. *Both* $\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)$ *and* $\operatorname{Tr}\big(\tau_1^+\tau_1^-\big)$ *are lower bounded by* $\max(0, 1 - 328\varepsilon_{\text{wav}})$*, where* ε_{wav} *is defined in Equation (4.16).*

Proof. The fact that $\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)$ is lower bounded by zero, follows directly from the fact that $\sigma_1^+ = \Lambda_A^1(|+\rangle\langle+|)$ and $\sigma_1^- = \Lambda_A^1(|-\rangle\langle-|)$ are positive semidefinite operators (see Lemma 2.9). For $\operatorname{Tr}\big(\tau_1^+\tau_1^-\big)$, first realize that

$$\operatorname{Tr}\big(\tau_1^+\tau_1^-\big) = \operatorname{Tr}\left((\tau_1^+\tau_1^-)^T\right) = \operatorname{Tr}\left((\tau_1^-)^T(\tau_1^+)^T\right) = \operatorname{Tr}\left((\tau_1^+)^T(\tau_1^-)^T\right).$$

Then we note that $(\tau_1^\pm)^T = U^\dagger\Lambda_B^1(|\pm\rangle\langle\pm|)U$ is a single-qubit state, so the fact that $\operatorname{Tr}\big(\tau_1^+\tau_1^-\big) \geq 0$ follows with Lemma 2.9. Showing the other lower bound, $1 - 328\varepsilon_{\text{wav}}$, is less straightforward.

We show that both $\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)$ and $\operatorname{Tr}\big(\tau_1^+\tau_1^-\big)$ are lower bounded by $1 - 328\varepsilon_{\text{wav}} = 1 - 8(11\varepsilon_{00} + 8\varepsilon_{02} + 8\varepsilon_{20} + 8\varepsilon_{22} + 3(\varepsilon_{10} + \varepsilon_{01}))$.

Note that when $x = 0$ or $y = 0$, we have $\rho_{xy} = \frac{\mathbb{1}\otimes\mathbb{1} + X\otimes X}{4} = \frac{|{+}{+}\rangle\langle{+}{+}| + |{-}{-}\rangle\langle{-}{-}|}{2}$.
Using Equations (4.11)-(4.14), we define

$$\varepsilon_{xy}^{++} := \frac{1}{2} - \langle(\Lambda_A^x \otimes \Lambda_B^y)(|{+}{+}\rangle\langle{+}{+}|), \Phi_+\rangle = \frac{1}{2} - \frac{1}{2}\operatorname{Tr}\big(\sigma_x^+\tau_y^+\big)$$
$$\text{for } (x,y) \text{ with } x = 0 \text{ or } y = 0$$

and also

$$\begin{aligned}\varepsilon_{xy}^{--} &:= \frac{1}{2} - \langle(\Lambda_A^x \otimes \Lambda_B^y)(|--\rangle\langle--|), \Phi_+\rangle = \frac{1}{2} - \frac{1}{2}\operatorname{Tr}\left(\sigma_x^- \tau_y^-\right) \\ &\quad \text{for } (x,y) \text{ with } x = 0 \text{ or } y = 0\end{aligned}$$

Since product states have fidelity at most $\frac{1}{2}$ with any two-qubit maximally entangled state, we see that ε_{xy}^{++} and ε_{xy}^{--} are both nonnegative.

In this way, we obtain the following 10 equalities (two equalities for each point (x, y) with $x = 0$ or $y = 0$):

$$\begin{aligned}\operatorname{Tr}\left(\sigma_x^+ \tau_y^+\right) &= 1 - 2\varepsilon_{xy}^{++} \\ &\quad \text{and} \\ \operatorname{Tr}\left(\sigma_x^- \tau_y^-\right) &= 1 - 2\varepsilon_{xy}^{--} \\ &\quad \text{for } (x,y) \text{ with } x = 0 \text{ or } y = 0\end{aligned} \tag{4.17}$$

Also, note that

$$\begin{gathered}\varepsilon_{xy}^{++} + \varepsilon_{xy}^{--} = 2\varepsilon_{xy} \\ \text{for } (x,y) \text{ with } x = 0 \text{ or } y = 0\end{gathered} \tag{4.18}$$

where ε_{xy} has been defined in Equation (4.15).

We can define similar quantities when $x = y = 2$, for which we have $\rho_{22} = \frac{\mathbb{1}\otimes\mathbb{1} - X\otimes X}{4} = \frac{|+-\rangle\langle+-| + |-+\rangle\langle-+|}{2}$. If we define

$$\begin{aligned}\varepsilon_{22}^{+-} &:= \frac{1}{2} - \langle(\Lambda_A^2 \otimes \Lambda_B^2)(|+-\rangle\langle+-|), \Phi_+\rangle = \frac{1}{2} - \frac{1}{2}\operatorname{Tr}\left(\sigma_2^+ \tau_2^-\right) \\ \varepsilon_{22}^{-+} &:= \frac{1}{2} - \langle(\Lambda_A^2 \otimes \Lambda_B^2)(|-+\rangle\langle-+|), \Phi_+\rangle = \frac{1}{2} - \frac{1}{2}\operatorname{Tr}\left(\sigma_2^- \tau_2^+\right)\end{aligned}$$

then we get

$$\begin{aligned}\operatorname{Tr}\left(\sigma_2^+ \tau_2^-\right) &= 1 - 2\varepsilon_{22}^{+-} \\ \operatorname{Tr}\left(\sigma_2^- \tau_2^+\right) &= 1 - 2\varepsilon_{22}^{-+}\end{aligned} \tag{4.19}$$

Also,

$$\varepsilon_{22}^{+-} + \varepsilon_{22}^{-+} = 2\varepsilon_{22} \tag{4.20}$$

Now we repeatedly apply lemma 2.10 to equations (4.17) and (4.19) to obtain the desired bounds. For the lower bound for $\operatorname{Tr}\left(\sigma_1^+ \sigma_1^-\right)$, a chain of inequalities is depicted in Figure 3, resulting in the bound

$$\operatorname{Tr}\left(\sigma_1^+ \sigma_1^-\right) \geq 1 - 8(6\varepsilon_{00}^{++} + \varepsilon_{00}^{--} + \varepsilon_{10}^{--} + 2\varepsilon_{10}^{++} + 4\varepsilon_{20}^{++} + 4\varepsilon_{02}^{++} + 4\varepsilon_{22}^{+-}) \tag{4.21}$$

Equation (4.18) allows us to merge some of the terms, which results in

$$\operatorname{Tr}\left(\sigma_1^+ \sigma_1^-\right) \geq 1 - 8(5\varepsilon_{00}^{++} + \varepsilon_{00} + \varepsilon_{10} + \varepsilon_{10}^{++} + 4\varepsilon_{20}^{++} + 4\varepsilon_{02}^{++} + 4\varepsilon_{22}^{+-}) \tag{4.22}$$

From Equation (4.18) we obtain $\varepsilon_{xy}^{++} \leq 2\varepsilon_{xy}$ and $\varepsilon_{xy}^{--} \leq 2\varepsilon_{xy}$, and, similarly, Equation (4.20) yields $\varepsilon_{22}^{+-} \leq 2\varepsilon_{22}$. Substituting these in the inequality (4.22) results in

$$\mathrm{Tr}\left(\sigma_1^+\sigma_1^-\right) \geq 1 - 8(11\varepsilon_{00} + 3\varepsilon_{10} + 8\varepsilon_{20} + 8\varepsilon_{02} + 8\varepsilon_{22}) \tag{4.23}$$

By subtracting the nonnegative term $8 \cdot 3\varepsilon_{01}$, we obtain the desired lower bound

$$\mathrm{Tr}\left(\sigma_1^+\sigma_1^-\right) \geq 1 - 8(11\varepsilon_{00} + 3(\varepsilon_{10} + \varepsilon_{01}) + 8\varepsilon_{20} + 8\varepsilon_{02} + 8\varepsilon_{22}) = 1 - 328\varepsilon_{\mathrm{wav}} \tag{4.24}$$

The case for $\mathrm{Tr}\left(\tau_1^+\tau_1^-\right)$ is analogous, by replacing the σ's in Figure 3 by τ's and vice versa. We can do so because the derivation in Figure 3 only depends on Equations (4.17) and (4.19) and Lemma 2.10, each of which is invariant under swapping Alice and Bob. The resulting bound can be obtained by swapping the indices in Equation (4.23):

$$\mathrm{Tr}\left(\tau_1^+\tau_1^-\right) \geq 1 - 8(11\varepsilon_{00} + 3\varepsilon_{01} + 8\varepsilon_{20} + 8\varepsilon_{02} + 8\varepsilon_{22}) \tag{4.25}$$

from which we subtract the nonnegative term $8 \cdot 3\varepsilon_{10}$ to obtain the desired result.

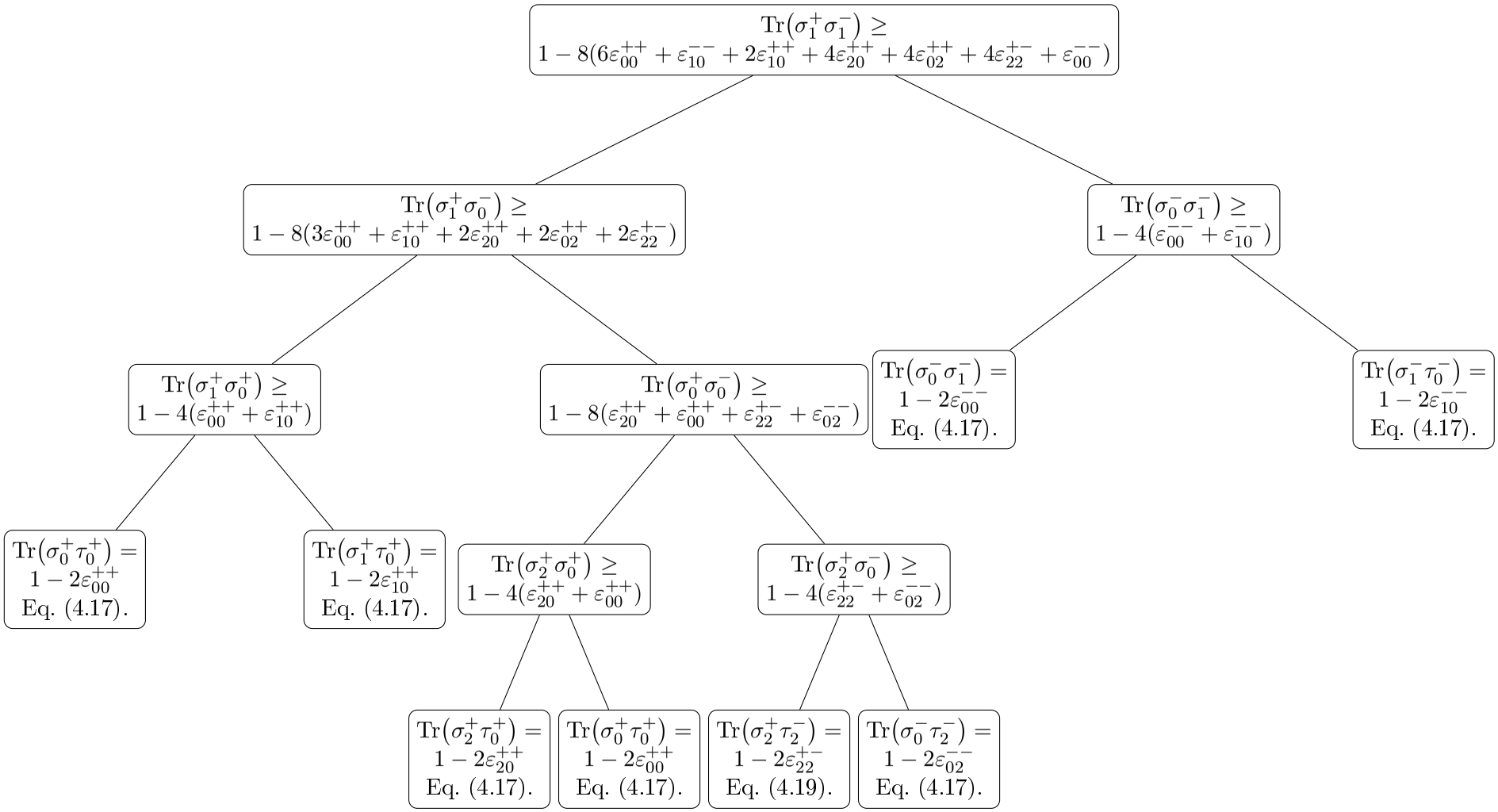

Figure 3: Starting from Equations (4.17) and (4.19), which are indicated in the leaf nodes, we repeatedly apply Lemma 2.10 to obtain Equation (4.21), which is depicted in the root node.
If a parent node depicts a lower bound on $\mathrm{Tr}(AB)$ using its children nodes with bounds on $\mathrm{Tr}(AC) \geq 1 - a$ and $\mathrm{Tr}(CB) \geq 1 - b$ where $a, b \geq 0$, then the derivation underlying this relation is $\mathrm{Tr}(AB) \overset{\text{Lemma 2.10}}{\geq} (1-2a)(1-2b) \geq 1 - 2(a+b)$.

□

As explained above, the previous lemma indicates that Λ_A^1 and Λ_B^1 are close to the full amplitude-damping channels from Section (4.4). Note that an completely-amplitude-damping channel to a pure state, shrinks the Bloch sphere to a point on its boundary. Hence any channel close to an amplitude-damping channel, must shrink the Bloch ball to some tiny volume, not far away from the boundary of the Bloch ball.

From this we infer that the center of the Bloch sphere, the maximally mixed state $\frac{\mathbb{1}_2}{2}$, is sent to a state that is close to the boundary. This idea is captured by the next lemma.

Lemma 4.3. *The smaller eigenvalue of $\Lambda_A^1(\frac{\mathbb{1}_2}{2})$ is upper bounded by $328\varepsilon_{\text{wav}}$, and so is the smaller eigenvalue of $\Lambda_B^1(\frac{\mathbb{1}_2}{2})$.*

Proof. Note that we can write $\Lambda_A^1(\frac{\mathbb{1}_2}{2}) = \Lambda_A^1(\frac{|+\rangle\langle+|+|-\rangle\langle-|}{2}) = \frac{\sigma_1^+ + \sigma_1^-}{2}$. Denote the eigenvalues of $\frac{\sigma_1^+ + \sigma_1^-}{2}$ by δ_A and $1-\delta_A$, with $0 \leq \delta_A \leq \frac{1}{2}$ (hence δ_A is the smaller eigenvalue). Using the fact that for any matrix M of rank 2, we have $2\det(M) = \big(\operatorname{Tr}(M)\big)^2 - \operatorname{Tr}\big(M^2\big)$, we obtain

$$\begin{aligned}
\frac{1}{2}\delta_A &\leq \delta_A(1-\delta_A) \\
&= \frac{1}{2}\det\Big(\frac{\sigma_1^+ + \sigma_1^-}{2}\Big) \\
&= \frac{1}{2}\Big[\Big(\operatorname{Tr}\Big[\frac{\sigma_1^+ + \sigma_1^-}{2}\Big]\Big)^2 - \operatorname{Tr}\Big(\Big[\frac{\sigma_1^+ + \sigma_1^-}{2}\Big]^2\Big)\Big] \\
&= \frac{1}{2}\cdot\frac{1}{4}\Big[\Big(\operatorname{Tr}\big(\sigma_1^+\big)\Big)^2 + \Big(\operatorname{Tr}\big(\sigma_1^-\big)\Big)^2 + 2\operatorname{Tr}\big(\sigma_1^+\big)\operatorname{Tr}\big(\sigma_1^-\big) - \operatorname{Tr}\Big((\sigma_1^-)^2\Big) - \operatorname{Tr}\Big((\sigma_1^-)^2\Big) - 2\operatorname{Tr}\big(\sigma_1^-\sigma_1^-\big)\Big] \\
&= \frac{1}{2}\cdot\frac{1}{4}\Big[4 - \operatorname{Tr}\big((\sigma_1^-)^2\big) - \operatorname{Tr}\big((\sigma_1^-)^2\big) - 2\operatorname{Tr}\big(\sigma_1^-\sigma_1^-\big)\Big] \\
&= \frac{1}{2}\Big[1 - \frac{1}{4}\Big[\operatorname{Tr}\big((\sigma_1^+)^2\big) + \operatorname{Tr}\big((\sigma_1^-)^2\big) + 2\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)\Big]\Big]
\end{aligned}$$

By the Cauchy-Schwarz inequality, we have

$$\big[\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)\big]^2 = \big[\langle\sigma_1^+, \sigma_1^-\rangle\big]^2 \leq \operatorname{Tr}\big((\sigma_1^+)^2\big)\cdot\operatorname{Tr}\big((\sigma_1^-)^2\big) \leq \operatorname{Tr}\big((\sigma_1^+)^2\big)\cdot 1 = \operatorname{Tr}\big((\sigma_1^+)^2\big)$$

and, similarly, $\big[\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big)\big]^2 \leq \operatorname{Tr}\big((\sigma_1^-)^2\big)$. From Lemma (2.10) we had $\operatorname{Tr}\big(\sigma_1^+\sigma_1^-\big) \geq \max(0, 1 - 328\varepsilon_{\text{wav}})$. We can therefore continue the series of inequalities as

$$\begin{aligned}
\frac{1}{2}\delta_A &\leq \frac{1}{2}\Big[1 - \frac{1}{4}\Big(2\max(0, 1-328\varepsilon_{\text{wav}})^2 + 2\max(0, 1-328\varepsilon_{\text{wav}})\Big)\Big] \\
&\leq \min\Big(\frac{1}{2}\Big[1 - \frac{1}{4}\Big(2(1-328\varepsilon_{\text{wav}})^2 + 2(1-328\varepsilon_{\text{wav}})\Big)\Big], \frac{1}{2}[1 - \frac{1}{4}(2\cdot 0 + 2\cdot 0)\Big) \\
&\leq \min\Big(\frac{1}{2}\Big[1 - \frac{1}{4}\Big(2(1-328\varepsilon_{\text{wav}}) + 2(1-328\varepsilon_{\text{wav}})\Big)\Big], \frac{1}{2}\Big) \\
&= \min\Big(\frac{1}{2}\Big[1 - (1-328\varepsilon_{\text{wav}})\Big], \frac{1}{2}\Big) \\
&\leq \frac{1}{2}\cdot 328\varepsilon_{\text{wav}}
\end{aligned}$$

hence $\delta_A \leq 328\varepsilon_{\text{wav}}$.

With an analogous argument, one can show that the smaller eigenvalue of $\frac{\tau_1^+ + \tau_1^-}{2}$ is upper bounded by $328\varepsilon_{\text{wav}}$ too. □

Now if a channel maps the center of the Bloch ball, the maximally mixed state $\frac{\mathbb{1}_2}{2}$, to a state close to the boundary, then by linearity the entire Bloch ball will be mapped to a tiny volume not far away from the boundary either. Consequently, the state $(\Lambda_A^1 \otimes \Lambda_B^1)(\Phi_+)$ is close to a product state, where Φ_+ is the target state, as defined in Equation (4.2). Since two-qubit product states have fidelity $\frac{1}{2}$ with any maximally entangled two-qubit state, we thus conclude that the fidelity $\langle (\Lambda_A^1 \otimes \Lambda_B^1)(\Phi_+), \Phi_+ \rangle$ cannot exceed $\frac{1}{2}$ by much. This idea is made rigorous in the next lemma.

Lemma 4.4. *The fidelity of Φ_+ with its image under $\Lambda_A^1 \otimes \Lambda_B^1$ is bounded as*

$$\langle (\Lambda_A^1 \otimes \Lambda_B^1)(\Phi_+), \Phi_+ \rangle \leq \frac{1}{2} + 656\varepsilon_{\text{wav}}.$$

Proof. Note that we have the following two bounds.

- We have
$$\left\langle (\Lambda_A^1 \otimes \Lambda_B^1) \frac{(\mathbb{1} \otimes \mathbb{1} + \sigma_{\mathsf{Y}} \otimes \sigma_{\mathsf{Y}})}{4}, \Phi_+ \right\rangle \leq \frac{1}{2} \tag{4.26}$$
since the first argument is a separable state, and the fidelity of a separable state with a two-qubit maximally entangled state is at most $\frac{1}{2}$ (see Corollary 2.14).

- Write P and P' such that $\frac{\mathbb{1}_2 + P}{2}$ and $\frac{\mathbb{1}_2 + P'}{2}$ are arbitrary vectors on the Bloch ball:
$$P := c_{\mathsf{X}}\mathsf{X} + c_{\mathsf{Y}}\mathsf{Y} + c_{\mathsf{Z}}\mathsf{Z} \text{ and } P' := c'_{\mathsf{X}}\mathsf{X} + c'_{\mathsf{Y}}\mathsf{Y} + c'_{\mathsf{Z}}\mathsf{Z}$$
where $c_{\mathsf{X}}, c_{\mathsf{Y}}, c_{\mathsf{Z}}, c'_{\mathsf{X}}, c'_{\mathsf{Y}}, c'_{\mathsf{Z}} \in \mathbb{R}$ and $c_{\mathsf{X}}^2 + c_{\mathsf{Y}}^2 + c_{\mathsf{Z}}^2 = (c'_{\mathsf{X}})^2 + (c'_{\mathsf{Y}})^2 + (c'_{\mathsf{Z}})^2 = 1$. In line with the notation of the proof of Lemma (4.3), denote the minimal eigenvalue of $\Lambda_A^1\left(\frac{\mathbb{1}_2}{2}\right)$ by δ_A, and write δ_B for the smaller eigenvalue of $\Lambda_B^1\left(\frac{\mathbb{1}_2}{2}\right)$. Then using Lemmas 2.13 and 2.8, we obtain

$$\begin{aligned}
\langle (\Lambda_A^1 \otimes \Lambda_B^1)(P \otimes P'), \Phi_+ \rangle &\overset{\text{Lemma 2.8}}{\leq} \langle |(\Lambda_A^1 \otimes \Lambda_B^1)(P \otimes P')|, \Phi_+ \rangle \\
&= \langle |(\Lambda_A^1)(P)| \otimes |(\Lambda_B^1)(P')|, \Phi_+ \rangle \\
&\overset{\text{Lemma 2.13}}{\leq} \langle 2\sqrt{\delta_A}\mathbb{1}_2 \otimes 2\sqrt{\delta_B}\mathbb{1}_2, \Phi_+ \rangle \\
&\leq 4\sqrt{\delta_A \delta_B}\left\langle \mathbb{1}_2 \otimes \mathbb{1}_2, \Phi_+ \right\rangle \\
&= 4\sqrt{\delta_A \delta_B} \\
&\overset{\text{Lemma 4.3}}{\leq} 4\sqrt{328\varepsilon_{\text{wav}} \cdot 328\varepsilon_{\text{wav}}} \\
&\leq 4 \cdot 328\varepsilon_{\text{wav}} \qquad (4.27) \\
&= 1312\varepsilon_{\text{wav}}. \qquad (4.28)
\end{aligned}$$

As defined in Equation (4.2), the target state Φ_+ is given by

$$\Phi_+ = \frac{1}{4}\left(\mathbb{1} \otimes \mathbb{1} + \mathsf{Y} \otimes \mathsf{Y} + \mathsf{X} \otimes \frac{\mathsf{X} + \mathsf{Z}}{\sqrt{2}} + \mathsf{Z} \otimes \frac{\mathsf{X} - \mathsf{Z}}{\sqrt{2}}\right).$$

Using Equations (4.26) and (4.28), we obtain

$$\begin{aligned}
\langle(\Lambda_A^1 \otimes \Lambda_B^1)(\Phi_+), \Phi_+\rangle &= \Big\langle(\Lambda_A^1 \otimes \Lambda_B^1)\Big(\frac{\mathbb{1} \otimes \mathbb{1} + \mathsf{Y} \otimes \mathsf{Y}}{4}\Big), \Phi_+\Big\rangle \\
&+ \frac{1}{4}\Big\langle(\Lambda_A^1 \otimes \Lambda_B^1)\Big(\mathsf{X} \otimes \frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}\Big), \Phi_+\Big\rangle \\
&+ \frac{1}{4}\Big\langle(\Lambda_A^1 \otimes \Lambda_B^1)\Big(\mathsf{Z} \otimes \frac{\mathsf{X}-\mathsf{Z}}{\sqrt{2}}\Big), \Phi_+\Big\rangle \\
&\leq \frac{1}{2} + \frac{1}{4} \cdot 1312\varepsilon_{\text{wav}} + \frac{1}{4} \cdot 1312\varepsilon_{\text{wav}} \\
&= \frac{1}{2} + 656\varepsilon_{\text{wav}}.
\end{aligned}$$

□

4.8 Upper bounding the singlet extractability of ρ_{TE}

To conclude the proof, we show that the state ρ_{TE}, after application of extraction channels, has trivial fidelity with the singlet for any choice of channels. To do so, we need to unfold our definitions and use lemma 4.4.

First, write $F\Big((\Lambda_A \otimes \Lambda_B)(\rho_{\text{TE}}), \Phi_+\Big)$ as a sum of inner product, as in Equation (4.10):

$$\begin{aligned}
F\Big((\Lambda_A \otimes \Lambda_B)(\rho_{\text{TE}}), \Phi_+\Big) &= \sum_{x,y=0}^{2} p_{xy}\langle(\Lambda_A^x \otimes \Lambda_B^y)(\rho_{xy}), \Phi_+\rangle \\
&= \nu \cdot \langle(\Lambda_A^1 \otimes \Lambda_B^1)(\rho_{11}), \Phi_+\rangle + \sum_{x,y\in\mathcal{F}} p_{xy}\langle(\Lambda_A^x \otimes \Lambda_B^y)(\rho_{xy}), \Phi_+\rangle
\end{aligned}$$

where, as before, we have denoted $\nu := p_{11}$. Substituting the definition of the ε_{xy} as in Equation (4.15), we get:

$$F\Big((\Lambda_A \otimes \Lambda_B)(\rho_{\text{TE}}), \Phi_+\Big) = \nu \cdot \langle(\Lambda_A^1 \otimes \Lambda_B^1)(\rho_{11}), \Phi_+\rangle + \sum_{x,y\in\mathcal{F}} p_{xy}(\frac{1}{2} - \varepsilon_{xy}).$$

Now by using Lemma (4.4), we obtain

$$F\Big((\Lambda_A \otimes \Lambda_B)(\rho_{\text{TE}}), \Phi_+\Big) \leq \nu \cdot \Big(\frac{1}{2} + 656\varepsilon_{\text{wav}}\Big) + \frac{\sum_{x,y\in\mathcal{F}} p_{xy}}{2} - \Big(\sum_{x,y\in\mathcal{F}} p_{xy}\varepsilon_{xy}\Big)$$

By substituting the definition of ε_{wav} as in Equation (4.16), we can continue the series of inequalities as

$$\begin{aligned}
F\Big((\Lambda_A \otimes \Lambda_B)(\rho_{\text{TE}}), \Phi_+\Big) &\leq \nu \cdot \Big(\frac{1}{2} + 656\varepsilon_{\text{wav}}\Big) + \frac{1-\nu}{2} - (1-\nu)\varepsilon_{\text{wav}} \\
&= \nu \cdot (656\varepsilon_{\text{wav}} + \varepsilon_{\text{wav}}) - \varepsilon_{\text{wav}} + \frac{1}{2} \\
&= \varepsilon_{\text{wav}} \cdot (657\nu - 1) + \frac{1}{2}.
\end{aligned}$$

Setting $\nu = \frac{1}{657} \approx 0.0015$ ensures that the first term vanishes, which results in $F\Big((\Lambda_A \otimes \Lambda_B)(\rho_{\mathrm{TE}}), \Phi_+\Big) \leq \frac{1}{2}$. With $\nu = \frac{1}{657}$, the Bell value of ρ_{TE} becomes $2 + (2\sqrt{2} - 2)\nu \approx 2.0013$. This concludes the proof of the main theorem of this chapter, Theorem 4.1.

The CHSH violation of the state we constructed greatly depends on the constants appearing in the several lemmas in this chapter. We note a few possibilities for optimization of these constants, thereby improving the CHSH violation of our state. First, the main lemma to the proof, Lemma 4.2, only considers Alice's and Bob's channels separately. However, we have already seen that the singlet extractability of our state equals $\frac{1}{2}$ if and only if the extraction channels of Alice and Bob damp towards the same term in the symmetric Schmidt decomposition of the singlet (see Section 4.4). Adjusting the proof while keeping this relation between Alice's and Bob's channels in mind might yield a greater CHSH violation than $\beta \approx 2.0013$. Second, Lemmas 4.4 and 4.3 involve the smallest eigenvalue of the maximally mixed state after application of the extraction channels. This smallest eigenvalue yields bounds on the fidelity with the singlet state as given in Lemma 2.13. However, upon close inspection, we see that these lemmas use the product of the eigenvalues (i.e. the determinant), rather than the smallest eigenvalue, which upper bounds the product. Improvement on the final CHSH violation could therefore be made by improving this lemma too.

5 The tilted CHSH inequality

5.1 Self-testing using the tilted CHSH inequality

The following family of Bell operators, which form a generalization of the CHSH operator, was introduced by Acín et al. [AMP12]:

$$B = \alpha A_0 \otimes \mathbb{1} + \sum_{j,k\in\{0,1\}} (-1)^{j\cdot k} A_j \otimes B_k \tag{5.1}$$

where A_0, A_1, B_0, B_1 are observables and $0 \leq \alpha < 2$ is a parameter. The classical value of the Bell operator is $2 + \alpha$ whereas quantum states can yield a violation up to $\sqrt{8 + 2\alpha^2}$. Consider the set of partially entangled two-qubit states, which can, up to local unitaries, be written in the form

$$\cos(\theta)\,|00\rangle + \sin(\theta)\,|11\rangle \tag{5.2}$$

where $\theta \in]0, \frac{\pi}{4}]$. Acín et al. showed that these states maximally violate the tilted CHSH inequality with the parameter $\alpha = 2/\sqrt{1 + 2\tan^2(2\theta)}$.

Building upon the work of Yang and Navascués [YN13], it was shown by Bamps and Pironio [BP15] that the states from Equation (5.2) are the *only* states that maximally violate the tilted CHSH inequality, up to additional degrees of freedom and local unitaries (these two features are necessary for self-testing, see Section 3.2). The proof by Bamps and Pironio also includes analytic self-testing bounds for these states. Their approach followed the line of work of McKague et al. as given in Definition 3.4; as such, they explicitly constructed the isometries as required by the definition of self-testability, Definition 3.2. The analytic bounds thus obtained are of order $O(\sqrt{\varepsilon})$, where ε is the difference between the observed violation and the maximal violation. These bounds are rather weak; for $\alpha = 0$ for example (which corresponds to self-testing of the singlet), the bound becomes trivial for $\varepsilon \gtrapprox 2.3 \cdot 10^{-3}$. For $\alpha = 1$, the bounds are trivial for $\varepsilon \gtrapprox 1.5 \cdot 10^{-4}$.

These bounds were significantly improved by Bancal et al. by application of the numerical SWAP method (see Section 3.4.2) [BNS+15]. Their results are depicted in Figure 4.

5.1.1 Our contribution

Numerically, we found improved self-testing bounds using the tilted CHSH inequality for all pure two-qubit states which can be parametrized, up to local unitaries, as in Equation (5.2) with $\theta \in [0.14, \frac{\pi}{4}]$. Further research is needed to verify that the bounds also hold for pure two-qubit states with $\theta \in]0, 0.14]$. We first state the final bounds.

Result 5.1. (Numerical result)
Denote by Φ_α a pure two-qubit state with Schmidt coefficients $\cos(\theta)$ and $\sin(\theta)$ with $\theta \in [0.14, \frac{\pi}{4}]$, where $\alpha = 2/\sqrt{1 + 2\tan^2(2\theta)}$. Let ρ be a state that achieves a violation β of the tilted CHSH inequality as given in Equation (5.1). Then

$$\max_{\Lambda_A,\Lambda_B} F\big((\Lambda_A \otimes \Lambda_B)(\rho), \Phi_\alpha\big) \geq s_\alpha \cdot \beta + \mu_\alpha \tag{5.3}$$

where s_α and μ_α are given by:

$$s_\alpha = \frac{1 - \frac{1}{4}\Big(1 + \sqrt{\frac{4-\alpha^2}{8+2\alpha^2}} + \sqrt{\frac{2\alpha^2}{8+2\alpha^2}}\Big)}{\sqrt{8 + 2\alpha^2} - (2 + \alpha)}, \tag{5.4}$$

$$\mu_\alpha = 1 - s_\alpha \cdot \sqrt{8 + 2\alpha^2}. \tag{5.5}$$

The remainder of this chapter is devoted to the derivation of this result, which we show using the method described in Section 3.5. Building upon the work of Acín et al., we first write down the state and measurement operators that maximally violate the tilted CHSH inequality. Then we explicitly extend the local extraction channels for the CHSH scenario [Kan16] to the scenario of the tilted CHSH inequality. Finally, for every $\alpha \in [0, 1.85]$ (which corresponds to $\theta \in [0.14, \frac{\pi}{4}]$) we compute an operator inequality and parameters s_α and μ_α that imply Equation (5.3). We verified these operator inequalities numerically; for details about the numerics, we refer to the end of this chapter (Section 5.2.7).

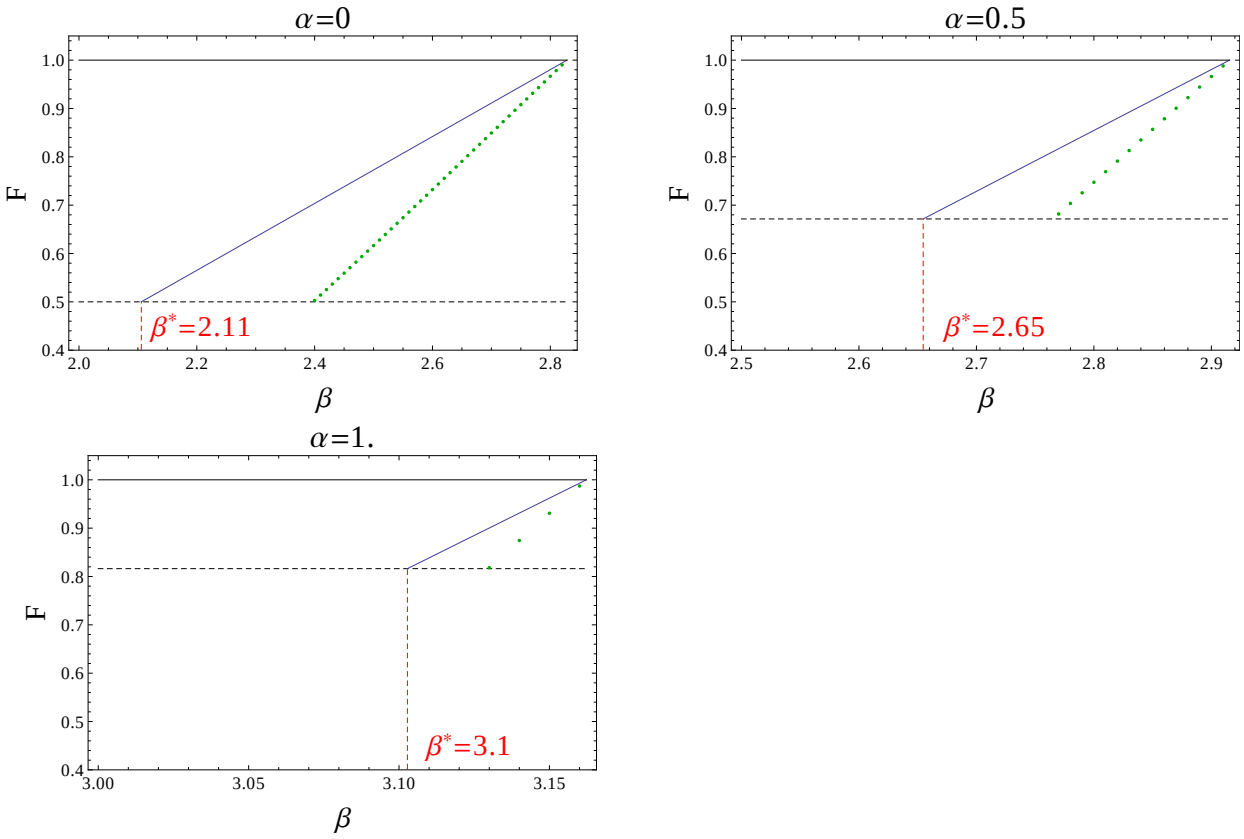

Figure 4: Lower bounds (solid line) on the extractability (see Definition 3.6) for partially entangled pure states $\cos(\theta)|00\rangle + \sin(\theta)|11\rangle$ as a function of the tilted CHSH violation β, where $\alpha = 2/\sqrt{1 + 2\tan^2(2\theta)}$. The horizontal axes range from the classical value to the quantum value. The trivial lower bound is indicated by a dashed horizontal line. Vertical dashed lines indicate the threshold violation β^*. The dotted lines are the previous best bounds as found by Bancal et al. [BNS+15]. The case $\alpha = 0$ corresponds to self-testing bounds for the maximally entangled two-qubit state as found before [Kan16].

5.2 Self-testing bounds from operator inequalities for (almost) all pure two-qubit states

5.2.1 The observables

Recall that, as a consequence of Jordan's lemma, we only need to consider qubit observables (see the text just before Equation (3.17) in Section 3.5). Similar to Equation (3.17), we define Alice's and Bob's

observables as

$$\begin{aligned} A_r(a) &:= \cos(a)\mathsf{X} + (-1)^r \sin(a)\mathsf{Z} \\ B_r(b) &:= \cos(b)\mathsf{X} + (-1)^r \sin(b)\mathsf{Z} \quad (r = 0, 1) \end{aligned}$$

where $a, b \in [0, \frac{\pi}{2}]$ are Alice's and Bob's angles, respectively. By the same reasoning as in Section 3.5, this formulation covers all possible choices of observables.

Using this observable parametrization, the tilted CHSH operator from Equation (5.1) becomes

$$\begin{aligned} B(a,b) &= 2\Big[(\cos(a)\mathsf{X} + \sin(a)\mathsf{Z}) \otimes (\cos(b)\mathsf{X} + \alpha\mathbb{1}) \\ &\quad + (\cos(a)\mathsf{X} - \sin(a)\mathsf{Z}) \otimes \sin(b)\mathsf{Z}\Big] \end{aligned} \tag{5.6}$$

5.2.2 The target states: partially entangled two-qubit states

The tilted CHSH inequality is maximally violated by pure states of the form $|\varphi_\theta\rangle := \cos(\theta)|00\rangle + \sin(\theta)|11\rangle$ for $\theta \in]0, \frac{\pi}{4}]$. Bamps and Pironio proved that the observables that achieve maximal violation with this state are [BP15]

$$\begin{array}{l|l} A_0 = \mathsf{Z} & B_0 = \cos(b^*)\mathsf{Z} + \sin(b^*)\mathsf{X} \\ A_1 = \mathsf{X} & B_1 = \cos(b^*)\mathsf{Z} - \sin(b^*)\mathsf{X} \end{array} \tag{5.7}$$

where $\tan(b^*) = \sin(2\theta)$. For the case $\alpha = 0$ (corresponding to $\theta = \frac{\pi}{4}$), the tilted CHSH operator reduces to the regular CHSH operator, and the observables that yield maximal violation are

$$\begin{array}{l|l} A_0 = \mathsf{Z} & B_0 = (\mathsf{Z} + \mathsf{X})/\sqrt{2} \\ A_1 = \mathsf{X} & B_1 = (\mathsf{Z} - \mathsf{X})/\sqrt{2}. \end{array} \tag{5.8}$$

For $\alpha = 0$, however, our observable parametrization yields maximal CHSH violation when

$$A_0 = B_0 = \frac{\mathsf{X} + \mathsf{Z}}{\sqrt{2}}, \quad A_1 = B_1 = \frac{\mathsf{X} - \mathsf{Z}}{\sqrt{2}}. \tag{5.9}$$

In order to obtain the state that maximally violates the tilted CHSH inequality for given α, we therefore need to change the local bases of the target state $|\varphi_\theta\rangle = \cos(\theta)\,|00\rangle + \sin(\theta)\,|11\rangle$. Let us first compute the pure density matrix $|\varphi_\theta\rangle\langle\varphi_\theta|$:

$$\begin{aligned} |\varphi_\theta\rangle\langle\varphi_\theta| &= \cos^2(\theta)\Big[\frac{\mathsf{Z}+\mathbb{1}}{2} \otimes \frac{\mathsf{Z}+\mathbb{1}}{2}\Big] + \sin^2(\theta)\Big[\frac{\mathsf{Z}-\mathbb{1}}{2} \otimes \frac{\mathsf{Z}-\mathbb{1}}{2}\Big] + \frac{1}{2}\cos(\theta)\sin(\theta)\Big[\mathsf{X}\otimes\mathsf{X} - Y \otimes Y\Big] \\ &= \frac{1}{4}\Big(\mathbb{1}\otimes\mathbb{1} + \sin(2\theta)\Big[\mathsf{X}\otimes\mathsf{X} - Y\otimes Y\Big] + \mathsf{Z}\otimes\mathsf{Z} + \cos(2\theta)\Big[\mathsf{Z}\otimes\mathbb{1} + \mathbb{1}\otimes\mathsf{Z}\Big]\Big). \end{aligned} \tag{5.10}$$

By comparing Equations (5.8) and (5.9), we see that we need to apply the following two local unitary transformations to $|\varphi_\theta\rangle\langle\varphi_\theta|$:

- for Alice: change X into $\frac{\mathsf{X}-\mathsf{Z}}{\sqrt{2}}$ and Z into $\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}$. This corresponds to a rotation of the Bloch sphere over an angle $\frac{\pi}{4}$ about the Y-axis;
- for Bob: change X into Z and vice versa, and change Y by $-\mathsf{Y}$. This corresponds to a rotation about the axis through $\frac{\mathbb{1}}{2}$ and $\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}$. During this process, X and Z are swapped and Y picks up a minus sign.

Applying these two transformations to $|\varphi_\theta\rangle\langle\varphi_\theta|$ from Equation (5.10), we obtain the target state that we will use:

$$\begin{aligned}\Phi_\alpha \;:=\; & \frac{1}{4}\Big(\mathbb{1}\otimes\mathbb{1}+\sin(2\theta)\Big[\frac{\mathsf{X}-\mathsf{Z}}{\sqrt{2}}\otimes\mathsf{Z}+\mathsf{Y}\otimes\mathsf{Y}\Big]+\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}\otimes\mathsf{X} \\ & +\cos(2\theta)\Big[\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}\otimes\mathbb{1}+\mathbb{1}\otimes\mathsf{X}\Big]\Big) \qquad (5.11) \\ = \; & \frac{1}{4}\Big[\mathbb{1}\otimes\mathbb{1}+\frac{1}{\sqrt{4+\alpha^2}}\Big(\alpha\sqrt{2}\Big[\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}\otimes\mathbb{1}+\mathbb{1}\otimes\mathsf{X}\Big] \\ & +\sqrt{4-\alpha^2}\mathsf{Y}\otimes\mathsf{Y}+\sqrt{4+\alpha^2}\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}\otimes\mathsf{X}+\sqrt{4-\alpha^2}\frac{\mathsf{X}-\mathsf{Z}}{\sqrt{2}}\otimes\mathsf{Z}\Big)\Big] \qquad (5.12)\end{aligned}$$

where we used the relation $\alpha = 2/\sqrt{1+2\tan^2(2\theta)}$ to rewrite the target state as a function of α instead of θ.

5.2.3 Trivial lower bound to extractability

The target state Φ_α is a pure state with Schmidt coefficients $\cos(\theta)$ and $\sin(\theta)$, where $\theta \in]0, \frac{\pi}{4}]$. The trivial lower bound to the extractability $\max_{\Lambda_A,\Lambda_B} F\big((\Lambda_A\otimes\Lambda_B)(\rho),\Phi_\alpha\big)$ is given by the square of the larger of the two Schmidt coefficients of Φ_α (see Section 3.5); hence the trivial lower bound is given by $\cos(\theta)^2 = \frac{1}{2}\big(\sqrt{\frac{2\alpha^2}{4+\alpha^2}}+1\big)$.

5.2.4 Optimal observables

We obtain the observable angles at which maximal violation occurs in our parametrization by applying the same transformations to the optimal observables as given in Equation (5.7). Doing so, we get the following optimal observables:

$$\begin{array}{l|l} A_0 = \frac{X+Z}{\sqrt{2}} & B_0 = \cos(b^*)X+\sin(b^*)Z \\ A_1 = \frac{X-Z}{\sqrt{2}} & B_1 = \cos(b^*)X-\sin(b^*)Z \end{array} \qquad (5.13)$$

where we had defined the angle b^* as $\tan(b^*) = \sin(2\theta) = \sqrt{\frac{4-\alpha^2}{4+\alpha^2}}$. Hence the optimal violation occurs at the angles $a = \frac{\pi}{4}$ and $b = b^*$ in our parametrization.

5.2.5 Extraction channels

Kaniewski used the following extraction channels for the CHSH inequality.

Extraction channels for CHSH [Kan16]
Alice and Bob have the same extraction channel Λ^{CHSH} as a function of their angle a and b, respectively, defined as:

$$[\Lambda^{\mathrm{CHSH}}(x)](\rho) := \frac{1+g(x)}{2}\rho + \frac{1-g(x)}{2}\Gamma(x)\rho\Gamma(x) \tag{5.14}$$

where

$$\Gamma(x) := \begin{cases} \mathsf{X} & \text{if } x \in [0, \pi/4] \\ \mathsf{Z} & \text{for } x \in (\pi/4, \pi/2] \end{cases}$$

and

$$g(x) := (1+\sqrt{2})(\sin(x) + \cos(x) - 1)$$

where x denotes the angle.

For the optimal angles for CHSH, $a = b = \frac{\pi}{4}$, we have $g(\frac{\pi}{4}) = 1$, hence the extraction channels are identity channels. For angles $x = 0$ or $x = \frac{\pi}{2}$, which correspond to commuting observables, the channels are full dephasing channels in the X-axis or Z-axis, respectively.

In order to find self-testing bounds using the tilted CHSH inequalities, we use dephasing channels again. When finding these extracting channels, we wish to keep two following properties of the channels: full dephasing for commuting observables and identity channels for the optimal angles. To do so, we need to modify the function g into functions g_α^A and g_α^B, such that the following two conditions hold:

$$\text{(a)} \qquad g_\alpha^A(0) = g_\alpha^A(\frac{\pi}{2}) = g_\alpha^B(0) = g_\alpha^B(\frac{\pi}{2}) = 0$$
$$\text{(b)} \qquad g_\alpha^A(\frac{\pi}{4}) = 1 \text{ and } g_\alpha^B(b^*) = 1.$$

(Recall that $b^* \equiv b^*(\alpha)$ is the optimal angle for Bob). Moreover, for the extraction channels from Equation (5.14) to be valid channels, the functions also need to satisfy $g_\alpha^A(x), g_\alpha^B(x) \in [-1, 1]$ for all $x \in [0, \frac{\pi}{2}]$.

For Alice, the optimal angle is independent of α, so we can set $g_\alpha^A(x) := g(x)$. For Bob's function g_α^B, we set out to find a function of the form $g_\alpha^B(x) = f_\alpha(x)g(x)$ for some function f to be determined later. The reason for scaling the original function g by multiplication is that, since g vanishes at $x = 0$ or $x = \frac{\pi}{2}$, the new function g_α^B will do so as well. We therefore only need to find a function f that satisfies $f_\alpha(b^*)g(b^*) = 1$ and $-1 \leq f_\alpha(x)g(x) \leq 1$.

Assuming that $f_\alpha(x)g(x)$ is maximal for $x = b^*$, we have

$$\frac{\partial}{\partial x} f_\alpha(x)g(x)\Big|_{x=b^*} = 0, \qquad \text{hence} \qquad f'_\alpha(b^*) = f_\alpha(b^*) \cdot \Big(- \frac{g'(b^*)}{g(b^*)} \Big)$$

A solution to this single-point equation is given by $f_\alpha(x) = e^{-\lambda(\alpha)\cdot x}$, where

$$\lambda(\alpha) := \frac{g'(b^*)}{g(b^*)} = \frac{\cos(b^*) - \sin(b^*)}{\cos(b^*) + \sin(b^*) - 1} = \frac{\sqrt{4+\alpha^2} - \sqrt{4-\alpha^2}}{\sqrt{4+\alpha^2} + \sqrt{4-\alpha^2} - 2\sqrt{2}}$$

where in the last equation, we used the relations $\tan(b^*) = \sin(2\theta)$ and $\alpha = 2/\sqrt{1 + 2\tan^2(2\theta)}$.

The second condition, $-1 \leq f_\alpha(x)g(x) \leq 1$ is satisfied by correctly normalizing $f_\alpha(x)$.

The functions Γ which are used in the extraction channel for the CHSH case need modification too. In particular, we need to change the angle at which the dephasing axis transition between the X-axis and the Z-axis. For the regular CHSH, this transition angle is the optimal angle $a = b = \frac{\pi}{4}$. It is natural to let this transition occur at the optimal angle for $\alpha \neq 0$ too.

Collecting all the modifications to the extraction channels as explained above, we obtain the following channels for the tilted CHSH scenario.

Extraction channels for the tilted CHSH scenario
Consider the extraction channel $\Lambda(x)$ as function of the angle x, defined as:

$$[\Lambda(x)](\rho) := \frac{1+g(x)}{2}\rho + \frac{1-g(x)}{2}\Gamma(x)\rho\Gamma(x)$$

where

$$\Gamma(x) := \begin{cases} \mathsf{X} & \text{if } x \in [0, x^*] \\ \mathsf{Z} & \text{for } x \in (x^*, \pi/2] \end{cases}$$

with x^* the transition angle and g a dephasing parameter function.

For Alice, the transition angle is $a^* = \frac{\pi}{4}$ and the function

$$g^A(a) := (1+\sqrt{2})(\cos(a) + \sin(a) - 1).$$

For Bob, we have $b^* = b^*(\alpha) = \arctan\left(\sqrt{\frac{4-\alpha^2}{4+\alpha^2}}\right)$ and the function g becomes

$$g^B_\alpha(x) := h(\alpha, x)/h(\alpha, b^*),$$

where

$$h(\alpha, x) := g^A(x) \cdot e^{-\lambda(\alpha)x}$$

and

$$\lambda(\alpha) := \frac{\sqrt{4+\alpha^2} - \sqrt{4-\alpha^2}}{\sqrt{4+\alpha^2} + \sqrt{4-\alpha^2} - 2\sqrt{2}}$$

5.2.6 Operator inequality

We compute $K_{\text{tilted}}(\alpha, a, b) := \left(\Lambda_A(\alpha, a)^\dagger \otimes \Lambda_B(\alpha, b)^\dagger\right)(\Phi_\alpha)$, where Λ_A and Λ_B are the extraction channels derived in the previous section and Φ_α is the partially entangled state as defined in Equation (5.12).

We introduce some notation first. If x^* denotes the optimal angle, write

$$\rho(x) = \begin{Bmatrix} \rho_{x \leq x^*} \\ \rho_{x > x^*} \end{Bmatrix}$$

when $\rho(x) = \rho_{x \leq x^*}$ if $x \leq x^*$ and $\rho(x) = \rho_{x > x^*}$ otherwise. Using this notation, we can write

$$\Lambda_A(a)(\mathsf{X}) = \mathsf{X} \cdot \begin{Bmatrix} 1 \\ g^A(a) \end{Bmatrix} \quad \text{and} \quad \Lambda_A(a)(\mathsf{Z}) = \mathsf{Z} \cdot \begin{Bmatrix} g^A(a) \\ 1 \end{Bmatrix}$$

and similarly,

$$\Lambda_B(b)(\mathsf{X}) = \mathsf{X}\cdot\begin{Bmatrix}1\\ g_\alpha^B(b)\end{Bmatrix} \quad \text{and} \quad \Lambda_B(b)(\mathsf{Z}) = \mathsf{Z}\cdot\begin{Bmatrix}g_\alpha^B(a)\\ 1\end{Bmatrix}.$$

It is now straightforward to compute

$$\begin{aligned}
K_{\text{tilted}}(\alpha,a,b) = \frac{1}{4}\Big[\quad & \mathbb{1}\otimes\mathbb{1} \\
+ \quad & \frac{1}{\sqrt{2}}\begin{Bmatrix}\mathsf{X}+g^A(a)\mathsf{Z}\\ g^A(a)\mathsf{X}+\mathsf{Z}\end{Bmatrix}\otimes\begin{Bmatrix}\mathsf{X}\\ g_\alpha^B(b)\mathsf{X}\end{Bmatrix} \\
+ \quad & \frac{\sqrt{4-\alpha^2}}{\sqrt{4+\alpha^2}}\Big(g^A(a)g_\alpha^B(b)\mathsf{Y}\otimes\mathsf{Y} + \frac{1}{\sqrt{2}}\begin{Bmatrix}\mathsf{X}-g^A(a)\mathsf{Z}\\ g^A(a)\mathsf{X}-\mathsf{Z}\end{Bmatrix}\otimes\begin{Bmatrix}g_\alpha^B(b)\mathsf{Z}\\ \mathsf{Z}\end{Bmatrix}\Big) \\
+ \quad & \frac{\alpha}{\sqrt{4+\alpha^2}}\Big(\begin{Bmatrix}\mathsf{X}+g^A(a)\mathsf{Z}\\ g^A(a)\mathsf{X}+\mathsf{Z}\end{Bmatrix}\otimes\mathbb{1} + \sqrt{2}\mathbb{1}\otimes\begin{Bmatrix}\mathsf{X}\\ g_\alpha^B(b)\mathsf{X}\end{Bmatrix}\Big) \\
\Big] &
\end{aligned}$$

Following the method to obtain linear self-testing bounds on the extractability from operator inequalities as described in Section 3.5, we need to find real-valued parameters s_α and μ_α, such that the operator inequality

$$T_{\alpha,s,\mu}(a,b) := K_{\text{tilted}}(\alpha,a,b) - s_\alpha B(\alpha,a,b) - \mu_\alpha \mathbb{1}_4 \geq 0 \tag{5.15}$$

holds, for all $\alpha\in[0,1.85]$ and $a,b\in[0,\frac{\pi}{2}]$, where $B(\alpha,a,b)$ is the tilted CHSH operator as defined in Equation (5.6).

By noting that $T_{\alpha,s}(a,b) = (\mathrm{H}\otimes\mathsf{X})T_{\alpha,s}(\frac{\pi}{2}-a,b)(H\otimes\mathsf{X})$ for all α and s, where $\mathrm{H}=(\mathsf{X}+\mathsf{Z})/\sqrt{2}$ is the Hadamard gate, we see that it is sufficient to only consider $a\in[0,\frac{\pi}{4}]$. We thus obtain

$$\begin{aligned}
T_{\alpha,s,\mu}(a,b) \quad = \quad & \Big(\frac{1}{4}-\mu\Big)\mathbb{1}\otimes\mathbb{1} \\
+ \quad & \frac{1}{4}\sqrt{\frac{4-\alpha^2}{4+\alpha^2}}g^A(a)g_\alpha^B(b)\mathsf{Y}\otimes\mathsf{Y} \\
+ \quad & \Big(\frac{1}{4\sqrt{2}}\begin{Bmatrix}1\\ g_\alpha^B(b)\end{Bmatrix} - 2s\cos(a)\cos(b)\Big)\mathsf{X}\otimes\mathsf{X} \\
+ \quad & \Big(\frac{1}{4\sqrt{2}}g^A(a)\begin{Bmatrix}1\\ g_\alpha^B(b)\end{Bmatrix} - 2s\sin(a)\cos(b)\Big)\mathsf{Z}\otimes\mathsf{X} \\
+ \quad & \Big(\frac{1}{4}\sqrt{\frac{4-\alpha^2}{8+2\alpha^2}}\begin{Bmatrix}g_\alpha^B(b)\\ 1\end{Bmatrix} - 2s\cos(a)\sin(b)\Big)\mathsf{X}\otimes\mathsf{Z} \\
- \quad & \Big(\frac{1}{4}\sqrt{\frac{4-\alpha^2}{8+2\alpha^2}}g^A(a)\begin{Bmatrix}g_\alpha^B(b)\\ 1\end{Bmatrix} - 2s\sin(a)\sin(b)\Big)\mathsf{Z}\otimes\mathsf{Z} \\
+ \quad & \frac{\alpha}{4\sqrt{4+\alpha^2}}\Big((\mathsf{X}+g^A(a)\mathsf{Z})\otimes\mathbb{1} + \sqrt{2}\begin{Bmatrix}1\\ g_\alpha^B(b)\end{Bmatrix}\mathbb{1}\otimes\mathsf{X}\Big) \\
- \quad & \alpha s(\cos(a)\mathsf{X}+\sin(a)\mathsf{Z})\otimes\mathbb{1}
\end{aligned} \tag{5.16}$$

for $\alpha\in[0,2[$ and $a\in[0,\frac{\pi}{4}]$ and $b\in[0,\frac{\pi}{2}]$, where the expressions in between curly brackets are dependent on Bob's transition angle b^*.

5.2.7 Computing the optimal s and μ

Numerical evidence indicates that the operator $T_{\alpha,s,\mu}(a,b)$ has strictly positive eigenvalues for all $\alpha \in [0, 1.85]$ and $s \geq 0$ and for all $(a,b) \in [0,\frac{\pi}{4}] \times [0,\frac{\pi}{2}] - \{(\frac{\pi}{4}, b^*), (0,\frac{\pi}{2})\}$. We first compute the minimal eigenvalues of $T_{\alpha,s,\mu}$ at the two 'critical' points for which $T_{\alpha,s,\mu}(a,b)$ is singular:

$$\text{Minimal eigenvalues:} \tag{5.17}$$
$$\begin{aligned} T_{\alpha,s,\mu}(\frac{\pi}{4}, b^*) &: 1 - s\sqrt{8+2\alpha^2} \\ T_{\alpha,s,\mu}(0,\frac{\pi}{2}) &: \frac{1}{4}\Big(1 + \sqrt{\frac{4-\alpha^2}{8+2\alpha^2}} + \sqrt{\frac{2\alpha^2}{8+2\alpha^2}}\Big) - (2+\alpha)s \end{aligned}$$

Our goal is to find $s = s_\alpha$ and $\mu = \mu_\alpha$ that satisfy (a) the operator inequality as given in Equation (5.15) and (b) for which the threshold violation β^* is minimal (see Section 3.5 for more details on the threshold violation). Since the trivial lower bound to the extractability of the input state with the target state Φ_α is given by $\frac{1}{2}\big(\sqrt{\frac{2\alpha^2}{4+\alpha^2}} + 1\big)$ (see Section 5.2.3), we obtain the threshold violation as a function of α, s and μ:

$$\beta^* = \frac{1}{s}\left[\frac{1}{2}\Big(\sqrt{\frac{2\alpha^2}{4+\alpha^2}} + 1\Big) - \mu\right]$$

By substituting the minimal eigenvalues of $T_{\alpha,s,\mu}(a,b)$ at the two critical points as given in Equation (5.17), we find that the threshold violation is minimized at the intersection of these two eigenvalues, as a function of s. Setting these two eigenvalues equal and solving for s yields

$$s_\alpha := \frac{1 - \frac{1}{4}\Big(1 + \sqrt{\frac{4-\alpha^2}{8+2\alpha^2}} + \sqrt{\frac{2\alpha^2}{8+2\alpha^2}}\Big)}{\sqrt{8+2\alpha^2} - (2+\alpha)}.$$

The two lines intersect at the value

$$\mu_\alpha := 1 - s_\alpha \cdot \sqrt{8+2\alpha^2}.$$

Setting $\alpha = 0$ recovers the parameters for the self-testing lower bounds for the singlet as found by Kaniewski [Kan16].

5.2.8 Numerical verification for positivity of $T_{\alpha,s,\mu}(a,b)$

We used the built-in numerical function `Eigenvalues` of the computer program *Mathematica* (version 8.0) to compute the eigenvalues of $T_{\alpha,s_\alpha,\mu_\alpha}(a,b)$ as given in Equation (5.16), with s_α and μ_α as given in the previous section. The variables α,a and b were given the following values:

- $\alpha \in [0, 1.85]$, with step size $d\alpha = 0.05$;
- $a \in [0, 22 \cdot \frac{\varepsilon}{\sqrt{2}}] \approx [0, \frac{\pi}{4}]$ with step size $da = \frac{\varepsilon}{\sqrt{2}}$, where $\varepsilon = 0.05$;
- $b \in [0, 44 \cdot \frac{\varepsilon}{\sqrt{2}}] \approx [0, \frac{\pi}{2}]$ with step size $db = \frac{\varepsilon}{\sqrt{2}}$, where $\varepsilon = 0.05$.

The value $\alpha = 1.85$ corresponds to $\theta \approx 0.14$. For given α, the distance between $(a,b) \in [0,\frac{\pi}{4}] \times [0,\frac{\pi}{2}]$ and the closed grid point (a,b) for which the eigenvalues of $T_{\alpha,s_\alpha,\mu_\alpha}$ were computed is at most ε.

For all of these values, the minimal eigenvalue of $T_{\alpha,s_\alpha,\mu_\alpha}(a,b)$ was nonnegative, with a precision of 10^{-12}. We thus verified that the operator $T_{\alpha,s_\alpha,\mu_\alpha}(a,b)$ is positive semidefinite for $0 \leq \alpha \leq 1.85$.

Following the method as outlined in Section 3.5, we arrive at our main result, Result 5.1.

For $\alpha > 1.85$, the minimal eigenvalue of $T_{\alpha,s_\alpha,\mu_\alpha}(a,b)$ ranged from the order of 10^{-6} to 10^{-3}. More research will have to show if these are numerical artifacts or that a slight adaption of the dephasing channels will yield nonnegative eigenvalues for $T_{\alpha,s_\alpha,\mu_\alpha}(a,b)$.

6 Discussion and conclusion

Self-testing bounds as a function of Bell violation can be formulated in terms of the extractability (see Definition 3.6), which is the fidelity of the target state with the input state after the players of the Bell game are allowed to apply local quantum channels called extraction channels. Using this formulation, we improved upon previously known robustness bounds for self-testing of (almost) all pure partially entangled two-qubit states using tilted CHSH inequalities. The proof for our new robustness bounds is based upon the construction of extraction channels for maximally entangled two-qubit states by Kaniewski [Kan16]. We extended these channels to the case of partially entangled two-qubit states, thereby finding new robustness bounds which improve upon all previously known results.

Furthermore, we constructed a state that violates the CHSH inequality but has trivial singlet extractability. Here 'trivial' indicates that there are no channels that achieve strictly greater fidelity with the singlet than when Alice and Bob ignore their shares and replace it by a fixed state. For the singlet, the trivial extractability is $\frac{1}{2}$. The state we constructed is a classical-quantum state and can be written as a probabilistic mixture of, on each side, a classical three-outcome classical register and a qubit. We showed that no pair of extraction channels can perform well at both the separable two-qubit states and the maximally-entangled two-qubit state at the same time; as a consequence, the overall fidelity is at most $\frac{1}{2}$.

The existence of such a state as described here shows not only that the CHSH violation does not imply nontrivial singlet extractability, but also that entanglement is not sufficient for nontrivial singlet extractability. Our result is, in the context of self-testing, of the same flavour as the existence of Werner states [Wer89], which are states that are genuinely entangled but nonetheless admit a local hidden variable model and thus do not violate any Bell inequality (see also Section 2.2.5). In the case of two qubits, Werner states are singlets mixed with uniform noise; by linearity of the singlet fidelity, one directly computes that the Werner state with visibility parameter p has singlet extractability at least $\frac{1+3p}{4}$, which is nontrivial for $p > \frac{1}{3}$. Since all Werner states with $p \leq \frac{1}{2}$ cannot violate any Bell inequality, we infer that nontrivial singlet extractability does not imply CHSH violation. The state of implications between entanglement, CHSH violation and singlet extractability is depicted in Figure 5.

Rather than considering the relation between singlet extractability and entanglement, it is a natural step to ask about *bound entanglement*. Bound entangled states are entangled but undistillable. Vértesi and Brunner constructed a bound entangled state that exhibits Bell nonlocality [VB14]. As classical communication is allowed in the process of entanglement distillation, an undistillable state has trivial singlet extractability. Vértesi and Brunner proved undistillability of the state they proposed by showing positivity of its partial transpose and provided an explicit Bell inequality, different from the CHSH inequality, that the state violates. Since nonpositivity of the partial transpose is necessary for CHSH violation, the state Vértesi and Brunner constructed does not violate the CHSH inequality. Our result is different from the work of Vértesi and Brunner since the state that we constructed violates the CHSH inequality.

6.1 Future research

The state we constructed has CHSH violation $\beta \approx 2.0013$. As a corollary to our result, the observation of a CHSH violation $\beta \lesssim 2.0013$ does not yield any information about the singlet extractability of the underlying state. On the other hand, nontrivial singlet extractability can be inferred if $\beta > 2.11$. It is up to future research to close this gap and to examine what information can be inferred about the state if it yields a CHSH violation $\beta \in (2.0013, 2.11]$: what is the real threshold violation in this regime?

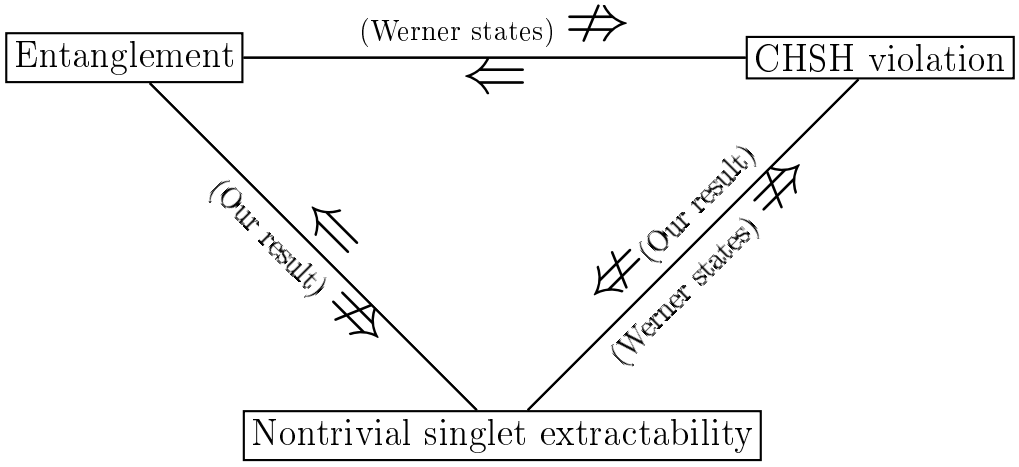

Figure 5: The relations between three properties that a quantum state can possess.

At the end of Chapter 4, we noted a few opportunities to improve the bounds on the CHSH violation of our constructed state. By optimizing our proof, we might be able to make the gap $\beta \in (2.0013, 2.11]$ smaller.

We have successfully obtained self-testing statements for two-qubit states from operator inequalities. A next step is to apply this method to different entangled states. In particular, future research could focus on the n-partite genuinely entangled states

$$\frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}}$$

which go by the name of *GHZ-states* [GHZ89]. Possibly, we could use operator inequalities to prove robust self-testing statements for the GHZ-states using the family of Mermin-Ardehali-Belinskii-Klyshko operators [Mer90, Ard92, BK93] (see also [Kan17] for a concise definition of this family of operators). Recently, provably tight self-testing bounds were constructed for the case $n = 3$ [Kan16]. Similar to our extension of self-testing statements of the maximally entangled two-qubit state to partially entangled two-qubit states, research could focus on extending the tight bounds for $n = 3$ to general $n > 3$. This task is of a different flavour than our results, however: since the operator inequalities that correspond to self-testing bounds grow with dimensionality, proving self-testing for n-partite GHZ states requires proving infinitely many operator inequalities whose dimensions grow exponentially with n.

Finally, one could examine the use of operator inequalities for a one-sided variant of self-testing called *quantum steering*. A steering setup is a variation to a Bell experiment where only Alice considers her device as a black box, while Bob has full control over his measurements (for a general review on steering, we refer to Cavalcanti and Skrzypczyk [CS16]). Such a setup is also called a *one-sided device-independent scenario*, since only Alice need not trust her device. The central question is: given Bob's measurements and his outcome, and Alice's input and output, what can be inferred about the shared state? Robustness bounds for steering have already been found [ŠASA16, GWK17]; future research could attempt to improve these with operator inequalities, using the method described in Section 3.5.

Bibliography

[ABG+07] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. Device-Independent Security of Quantum Cryptography against Collective Attacks. Phys. Rev. Lett., 98: 230501, 2007.
DOI: 10.1103/PhysRevLett.98.230501.

[AFRV16] R. Arnon-Friedman, R. Renner, and T. Vidick. Simple and tight device-independent security proofs. arXiv:1607.01797, 2016.
arXiv: 1607.01797.

[AGM06] A. Acín, N. Gisin, and L. Masanes. From Bell's Theorem to Secure Quantum Key Distribution. Phys. Rev. Lett., 97: 120405, 2006.
DOI: 10.1103/PhysRevLett.97.120405.

[AMP06] A. Acín, S. Massar, and S. Pironio. Efficient quantum key distribution secure against no-signalling eavesdroppers. New Journal of Physics, 8(8): 126–126, 2006.
DOI: 10.1088/1367-2630/8/8/126.

[AMP12] A. Acín, S. Massar, and S. Pironio. Randomness versus Nonlocality and Entanglement. Phys. Rev. Lett., 108: 100402, 2012.
DOI: 10.1103/PhysRevLett.108.100402.

[Ard92] M. Ardehali. Bell inequalities with a magnitude of violation that grows exponentially with the number of particles. Phys. Rev. A, 46: 5375–5378, 1992.
DOI: 10.1103/PhysRevA.46.5375.

[Bar02] J. Barrett. Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality. Phys. Rev. A, 65: 042302, 2002.
DOI: 10.1103/PhysRevA.65.042302.

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, 175, 1984.
DOI: 10.1016/j.tcs.2014.05.025.

[BBHT98] M. Boyer, G. Brassard, P. Høyer, and A. Tapp. Tight Bounds on Quantum Searching. Fortschritte der Physik, 46(4-5): 493–505, 1998.
DOI: 10.1002/(SICI)1521-3978(199806)46:4/5<493::AID-PROP493>3.0.CO;2-P.

[BBM92] C. H. Bennett, G. Brassard, and N. D. Mermin. Quantum cryptography without Bell's theorem. Phys. Rev. Lett., 68: 557–559, 1992.
DOI: 10.1103/PhysRevLett.68.557.

[BBPS96] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. Phys. Rev. A, 53: 2046–2052, 1996.
DOI: 10.1103/PhysRevA.53.2046.

[BCP+14] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. Reviews of Modern Physics, 86(2): 419–478, 2014.
DOI: 10.1103/revmodphys.86.419.

[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. Physics, 1964.

[Ben82] P. Benioff. Quantum mechanical hamiltonian models of turing machines. Journal of Statistical Physics, 29(3): 515–546, 1982.
DOI: 10.1007/bf01342185.

[BHK05] J. Barrett, L. Hardy, and A. Kent. No Signaling and Quantum Key Distribution. Phys. Rev. Lett., 95: 010503, 2005.
DOI: 10.1103/PhysRevLett.95.010503.

[BK93] A. V. Belinskii and D. N. Klyshko. Interference of light and Bell's theorem. Physics-Uspekhi, 36(8): 653, 1993.
DOI: 10.1070/PU1993v036n08ABEH002299.

[BLM+09] C.-E. Bardyn, T. C. H. Liew, S. Massar, M. McKague, and V. Scarani. Device-independent state estimation based on Bell's inequalities. Phys. Rev. A, 80: 062327, 2009.
DOI: 10.1103/PhysRevA.80.062327.

[BMR92] S. L. Braunstein, A. Mann, and M. Revzen. Maximal violation of Bell inequalities for mixed states. Phys. Rev. Lett., 68: 3259–3261, 1992.
DOI: 10.1103/PhysRevLett.68.3259.

[BNS+15] J.-D. Bancal, M. Navascués, V. Scarani, T. Vértesi, and T. H. Yang. Physical characterization of quantum devices from nonlocal correlations. Phys. Rev. A, 91: 022115, 2015.
DOI: 10.1103/PhysRevA.91.022115.

[Boh51] D. Bohm. Quantum theory. Prentice-Hall physics series, 1951.

[BP15] C. Bamps and S. Pironio. Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing. Phys. Rev. A, 91: 052111, 2015.
DOI: 10.1103/PhysRevA.91.052111.

[BPPP14] J. Bouda, M. Pawłowski, M. Pivoluska, and M. Plesch. Device-independent randomness extraction from an arbitrarily weak min-entropy source. Phys. Rev. A, 90: 032313, 2014.
DOI: 10.1103/PhysRevA.90.032313.

[CGS16] A. Coladangelo, K. T. Goh, and V. Scarani. All Pure Bipartite Entangled States can be Self-Tested. arXiv:1611.08062, 2016.
arXiv: 1611.08062.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. Phys. Rev. Lett., 23: 880–884, 1969.
DOI: 10.1103/PhysRevLett.23.880.

[Cir80] B. S. Cirel'son. Quantum generalizations of Bell's inequality. Letters in Mathematical Physics, 4(2): 93–100, 1980.
DOI: 10.1007/BF00417500.

[CK11] R. Colbeck and A. Kent. Private randomness expansion with untrusted devices. Journal of Physics A: Mathematical and Theoretical, 44(9): 095305, 2011.
Online: http://stacks.iop.org/1751-8121/44/i=9/a=095305.

[Col07] R. Colbeck. Quantum And Relativistic Protocols For Secure Multi-Party Computation (Phd thesis). arxiv:0911.3814, 2007.
arXiv: 0911.3814.

[CS16] D. Cavalcanti and P. Skrzypczyk. Quantum steering: a review with focus on semidefinite programming. Reports on Progress in Physics, 80(2): 024001, 2016.
DOI: 10.1088/1361-6633/80/2/024001.

[CVY13] M. Coudron, T. Vidick, and H. Yuen. Robust Randomness Amplifiers: Upper and Lower Bounds, pages 468–483. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
DOI: 10.1007/978-3-642-40328-6_33.

[DFR16] F. Dupuis, O. Fawzi, and R. Renner. Entropy accumulation. arXiv:1607.01796, 2016.
arXiv: 1607.01796.

[Dra05] S. S. Dragomir. Advances in Inequalities of the Schwarz, Grüss, and Bessel Type in Inner Product Spaces. Nova Publishers, 2005.
arXiv: math/0503059.

[Eke91] A. K. Ekert. Quantum cryptography based on Bell's theorem. Phys. Rev. Lett., 67: 661–663, 1991.
DOI: 10.1103/PhysRevLett.67.661.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? Phys. Rev., 47: 777–780, 1935.
DOI: 10.1103/PhysRev.47.777.

[ER14] A. Ekert and R. Renner. The ultimate physical limits of privacy. Nature, 507(7493): 443–447, 2014.
DOI: 10.1038/nature13132.

[Fey82] R. P. Feynman. Simulating physics with computers. International Journal of Theoretical Physics, 21(6-7): 467–488, 1982.
DOI: 10.1007/bf02650179.

[Fin82] A. Fine. Hidden Variables, Joint Probability, and the Bell Inequalities. Phys. Rev. Lett., 48: 291–295, 1982.
DOI: 10.1103/PhysRevLett.48.291.

[GHZ89] D. M. Greenberger, M. A. Horne, and A. Zeilinger. Going Beyond Bell's Theorem, pages 69–72. Springer Netherlands, Dordrecht, 1989.
DOI: 10.1007/978-94-017-0849-4_10.

[Gro96] L. K. Grover. A fast quantum mechanical algorithm for database search. In Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96. Association for Computing Machinery (ACM), 1996.
DOI: 10.1145/237814.237866.

[GVW+15] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger. Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons. Phys. Rev. Lett., 115: 250401, 2015.
DOI: 10.1103/PhysRevLett.115.250401.

[GWK17] A. Gheorghiu, P. Wallden, and E. Kashefi. Rigidity of quantum steering and one-sided device-independent verifiable quantum computation. New Journal of Physics, 19(2): 023043, 2017.
DOI: 10.1088/1367-2630/aa5cff.

[HBD+15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. Nature, 526(7575): 682–686, 2015.
DOI: 10.1038/nature15759.

[HHHH09] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. Rev. Mod. Phys., 81: 865–942, 2009.
DOI: 10.1103/RevModPhys.81.865.

[HKB+16] B. Hensen, N. Kalb, M. S. Blok, A. E. Dréau, A. Reiserer, R. F. L. Vermeulen, R. N. Schouten, M. Markham, D. J. Twitchen, K. Goodenough, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell test using electron spins in diamond: second experiment and additional analysis. Scientific Reports, 6(1), 2016.
DOI: 10.1038/srep30289.

[Kan16] J. Kaniewski. Analytic and Nearly Optimal Self-Testing Bounds for the Clauser-Horne-Shimony-Holt and Mermin Inequalities. Phys. Rev. Lett., 117: 070402, 2016.
DOI: 10.1103/PhysRevLett.117.070402.

[Kan17] J. Kaniewski. Self-testing of binary observables based on commutation. arXiv:1702.06845, 2017.
arXiv: 1702.06485.

[McK10] McKague, Matthew. Quantum Information Processing with Adversarial Devices. PhD thesis, 2010.
Online: http://hdl.handle.net/10012/5259.

[McK14] M. McKague. Self-Testing Graph States, pages 104–120. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
DOI: 10.1007/978-3-642-54429-3_7.

[McK16] M. McKague. Self-testing in parallel. New Journal of Physics, 18(4): 045013, 2016.
DOI: 10.1088/1367-2630/18/4/045013.

[Mer90] N. D. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. Phys. Rev. Lett., 65: 1838–1840, 1990.
DOI: 10.1103/PhysRevLett.65.1838.

[MMMO06] F. Magniez, D. Mayers, M. Mosca, and H. Ollivier. Self-testing of Quantum Circuits. In Automata, Languages and Programming, pages 72–83. Springer Nature, 2006.
DOI: 10.1007/11786986_8.

[MPA11] L. Masanes, S. Pironio, and A. Acín. Secure device-independent quantum key distribution with causally independent measurement devices. Nature Communications, 2: 238, 2011.
DOI: 10.1038/ncomms1244.

[MS16] C. A. Miller and Y. Shi. Robust Protocols for Securely Expanding Randomness and Distributing Keys Using Untrusted Quantum Devices. J. ACM, 63(4): 33:1–33:63, 2016.
DOI: 10.1145/2885493.

[MW06] L. Masanes and A. Winter. Unconditional security of key distribution from causality constraints. arXiv:quant-ph/0606049, 2006.
arXiv: quant-ph/0606049v1.

[MY98] D. Mayers and A. Yao. Quantum cryptography with imperfect apparatus. In Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280), pages 503–509, 1998.
DOI: 10.1109/SFCS.1998.743501.

[MY04] D. Mayers and A. Yao. Self Testing Quantum Apparatus. Quantum Info. Comput., 4(4): 273–286, 2004.
arXiv: quant-ph/0307205.

[MYS12] M. McKague, T. H. Yang, and V. Scarani. Robust self-testing of the singlet. Journal of Physics A: Mathematical and Theoretical, 45(45): 455304, 2012.
Online: http://stacks.iop.org/1751-8121/45/i=45/a=455304.

[NC00] M. A. Nielsen and I. L. Chuang. Quantum information and quantum computation. Cambridge: Cambridge University Press, 2(8): 23, 2000.

[Nie99] M. A. Nielsen. Conditions for a Class of Entanglement Transformations. Phys. Rev. Lett., 83: 436–439, 1999.
DOI: 10.1103/PhysRevLett.83.436.

[PAM+10] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by Bell's theorem. Nature, 464(7291): 1021–1024, 2010.
DOI: 10.1038/nature09008.

[Per96] A. Peres. Separability Criterion for Density Matrices. Phys. Rev. Lett., 77: 1413–1415, 1996.
DOI: 10.1103/PhysRevLett.77.1413.

[PR92] S. Popescu and D. Rohrlich. Which states violate Bell's inequality maximally? Physics Letters A, 169(6): 411 – 414, 1992.
DOI: 10.1016/0375-9601(92)90819-8.

[Pre15] J. Preskill. Lecture Notes: Quantum Computation. 2015.
Online: http://www.theory.caltech.edu/people/preskill/ph229/.

[PVN14] K. F. Pál, T. Vértesi, and M. Navascués. Device-independent tomography of multipartite quantum states. Phys. Rev. A, 90: 042340, 2014.
DOI: 10.1103/PhysRevA.90.042340.

[ŠASA16] I. Šupić, R. Augusiak, A. Salavrakos, and A. Acín. Self-testing protocols based on the chained bell inequalities. New Journal of Physics, 18(3): 035013, 2016.
Online: http://stacks.iop.org/1367-2630/18/i=3/a=035013.

[SAT+16] A. Salavrakos, R. Augusiak, J. Tura, P. Wittek, A. Acín, and S. Pironio. Bell inequalities for maximally entangled states. arXiv:1607.04578, 2016.
arXiv: 1607.04578.

[SCA+11] J. Silman, A. Chailloux, N. Aharon, I. Kerenidis, S. Pironio, and S. Massar. Fully Distrustful Quantum Bit Commitment and Coin Flipping. Phys. Rev. Lett., 106: 220501, 2011.
DOI: 10.1103/PhysRevLett.106.220501.

[Sca12] V. Scarani. The device-independent outlook on quantum physics. Acta Physica Slovaca, 62(4): 347–409, 2012.

[SGB+06] V. Scarani, N. Gisin, N. Brunner, L. Masanes, S. Pino, and A. Acín. Secrecy extraction from no-signaling correlations. Phys. Rev. A, 74: 042339, 2006.
DOI: 10.1103/PhysRevA.74.042339.

[Sho94] P. W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on, pages 124–134. IEEE, 1994.
DOI: 10.1109/sfcs.1994.365700.

[SK14] V. Scarani and C. Kurtsiefer. The black paper of quantum cryptography: Real implementation problems. Theoretical Computer Science, 560, Part 1: 27 – 32, 2014.
DOI: 10.1016/j.tcs.2014.09.015.

[SMSC+15] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam. Strong Loophole-Free Test of Local Realism. Phys. Rev. Lett., 115: 250402, 2015.
DOI: 10.1103/PhysRevLett.115.250402.

[SW88] S. J. Summers and R. Werner. Maximal violation of Bell's inequalities for algebras of observables in tangent spacetime regions. In Annales de l'IHP Physique théorique, volume 49, pages 215–243, 1988.
Online: http://www.numdam.org/article/AIHPA_1988__49_2_215_0.pdf.

[VB14] T. Vértesi and N. Brunner. Disproving the Peres conjecture by showing Bell nonlocality from bound entanglement. Nature Communications, 5: 5297, 2014.
DOI: 10.1038/ncomms6297.

[VV12] U. Vazirani and T. Vidick. Certifiable quantum dice. 2012.
DOI: 10.1145/2213977.2213984.

[VV14] U. Vazirani and T. Vidick. Fully Device-Independent Quantum Key Distribution. Phys. Rev. Lett., 113: 140501, 2014.
DOI: 10.1103/PhysRevLett.113.140501.

[Wat16] J. Watrous. Book: Theory of Quantum Information. 2016.
Online: https://cs.uwaterloo.ca/~watrous/TQI/.

[WCY+14] X. Wu, Y. Cai, T. H. Yang, H. N. Le, J.-D. Bancal, and V. Scarani. Robust self-testing of the three-qubit W state. Phys. Rev. A, 90: 042339, 2014.
DOI: 10.1103/PhysRevA.90.042339.

[Wer89] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. Phys. Rev. A, 40: 4277–4281, 1989.
DOI: 10.1103/PhysRevA.40.4277.

[Wik] WikiLeaks. Cia malware targets iphone, android, smart tvs.
Online: https://wikileaks.org/ciav7p1/#ANALYSIS. Accessed: 2017-04-04.

[WW01] R. F. Werner and M. M. Wolf. All-multipartite Bell-correlation inequalities for two dichotomic observables per site. Phys. Rev. A, 64: 032112, 2001.
DOI: 10.1103/PhysRevA.64.032112.

[YN13] T. H. Yang and M. Navascués. Robust self-testing of unknown quantum systems into any entangled two-qubit states. Phys. Rev. A, 87: 050102, 2013.
DOI: 10.1103/PhysRevA.87.050102.

[YVB+14] T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, and M. Navascués. Robust and Versatile Black-Box Certification of Quantum Devices. Phys. Rev. Lett., 113: 040401, 2014.
DOI: 10.1103/PhysRevLett.113.040401.