ABSTRACT

Title of Dissertation: Improved Algorithms and Primitives for Quantum Cryptography

Nishant Rodrigues
Doctor of Philosophy, 2023

Dissertation Directed by: Dr. Brad Lackey
Department of Computer Science

In this dissertation we develop novel primitives and algorithms in quantum cryptography, specifically for quantum key distribution and random number generation. We show a device-independent quantum key distribution (DIQKD) algorithm that is based on the notion of synchronous correlations. Most algorithms in DIQKD literature rely on the well-known CHSH inequality which is neither symmetric nor synchronous. We propose a new synchronous Bell inequality that simplifies the QKD setting by being fully symmetric so that the roles of the two parties in the protocol, Alice and Bob, are completely interchangeable. This has implications for QKD hardware since an identical set of devices can be produced for both parties instead of separate devices for each. We also achieve key rates comparable to CHSH-based protocols.

This dissertation also focuses on closing the causality or locality loophole present in device-independent schemes. An assumption that is critical to device-independent protocols is that the two parties are acausally separated and cannot communicate with each other

once they receive their inputs. This is typically referred to as the *nonsignaling* condition. If the condition of nonsignaling is violated, then an attacker may simulate the entire protocol classically. This erases any certificate of quantumness produced by the violation of a Bell inequality. We pose a new security assumption with respect to the adversary's uncertainty about the two parties' measurement bases. We derive a bound for this uncertainty and show that if the uncertainty grows any larger than the threshold, there is no strategy any adversary can use to cheat in the protocol. This closes the causality loophole and makes the protocol easier to implement for practical use.

A primitive widely used in QKD and other cryptographic protocols is a random bit generator. We define *ideal* and *real* models of random bit generators and show their efficiency and security in the Constructive Cryptography framework. We specifically look at random bit generators based on process tomography of one-qubit channels. We consider ideal quantum random bit generators, then introduce some state preparation and bit-flip errors to define real quantum random bit generators. We show that the ideal and real generators are close to each other in statistical distance.

The third part of this dissertation presents some initial ideas for quantum lattice sieving algorithms. Lattices are very important objects in the effort to construct cryptographic primitives that are secure against quantum attacks. A central problem in the study of lattices is that of finding the shortest non-zero vector in the lattice. Asymptotically, *sieving* is the best known technique for solving the shortest vector problem, however, sieving requires memory exponential in the dimension of the lattice. This work tries to provide better memory complexity while also improving runtime. Our ideas are inspired by classical heuristic sieving algorithms and make an attempt to quantize those algorithms.

IMPROVED ALGORITHMS AND PRIMITIVES FOR QUANTUM CRYPTOGRAPHY

by

Nishant Rodrigues

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Andrew Childs, Chair
Dr. Brad Lackey, Advisor
Professor Dana Dachman-Soled
Professor David Mount
Professor Yi-Kai Liu

To my lovely grandma Isabella whom I miss dearly.

Acknowledgments

This dissertation would not have been possible without the constant support of my advisor Brad Lackey. When I first started graduate school, I had no idea what quantum cryptography was. I had interest in working on classical cryptography but I struggled to find an academic advisor working in classical cryptography. Brad was co-teaching an introductory course on quantum information along with Andrew Childs and Shelby Kimmel during my first semester as a graduate student. I was introduced to Brad through the course and later ended up reaching out to him about doing research on open problems. I am really grateful that I met Brad; he has been a coach and mentor to me all these years. During the first few years working with Brad he would prepare lecture-style notes and we would go over them during our regularly scheduled meetings. I have learned a lot and continue to learn from him. I appreciate his time, effort and guidance.

I am very grateful to Andrew Childs for agreeing to chair my dissertation committee, and for all the help along the way, not only while dealing with the administrative hurdles that came along but also through helpful chats and guidance over the years.

I'd like to thank the committee members Dave Mount, Dana Dachman-Soled, and Yi-Kai Liu for agreeing to be on my committee and for their helpful feedback with my research.

I owe a lot of gratitude to all the friends I made in graduate school. Shravan, Devesh, and Alejandro, thank you for being very supportive and for providing a safe space to

talk about graduate school struggles and mental health. I want to thank Ananth, Heba, Bhargavi, and Yusuf for their friendship and support throughout my time at UMD.

My brother Aniket and sister-in-law Vijaya, my parents George and Shubhada, and my aunt Amita have been my biggest role models and also my greatest supporters. My family, including my cousins Sanal and Sanika, and my two adorable nephews Kiaan and Shaarvin, have been a constant source of love and support throughout my PhD.

I cannot end this section without thanking my wonderful partner Mackenzie Rech. She has helped me get through difficult times throughout this journey by being a source of strength during the setbacks and a voice of reason and compassion in moments of doubt. She cherished even the tiniest of accomplishments and has cheered me on from start to finish. I couldn't have done this without her.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

In this dissertation, we explore new algorithms and frameworks for quantum cryptography. Charles Bennet and Gilles Brassard kick-started the field of quantum cryptography by proposing the first protocol for generating shared classical keys using quantum states. For a long time quantum cryptography was synonymous with quantum key distribution. However, quantum cryptography has now grown into a rich field of research and encompasses a variety of paradigms and protocols such as uncloneable cryptography, quantum randomness generation and expansion, quantum interactive protocols, proofs of quantumness, and quantum money among others. An adjacent area of research is that of post-quantum cryptography which has gained traction recently due to NIST's Post-Quantum Cryptography Standardization competition that aims to standardize cryptographic encryption and digital signature algorithms that are secure against attacks by quantum computers.

Constructive Cryptography, also referred to as Abstract Cryptography [1] is a framework for analyzing cryptographic protocols from the top-down, rather than the traditional bottom-up approach. Traditionally, analyzing cryptographic protocols first requires defining a model of computation (ex. Turing machines), writing an algorithm using this model, then defining a notion of complexity for this algorithm (ex. number of steps, or number of queries to an oracle), and then arguing about the efficiency of the algorithm based on

that notion of complexity. We finally then prove security of the algorithm based on the infeasibility of a game. The security proofs quite often also get very complicated.

Constructive Cryptography aims to simplify cryptographic proofs by defining *systems* at the highest level of abstraction, and then moving down to lower levels only when necessary. This allows us to prove properties of systems at every level without worrying about finer-grained details necessary at lower levels. Systems can be composed using simple algebraic rules. This framework allows us to compose different systems, and if each of the individual systems is secure, then the composed system also remains secure. An important paradigm that is used in constructive cryptography is that of defining *ideal* systems and *real* systems. An ideal system is a system that provides the desired functionality under ideal circumstances. A real system provides the same functionality, however considers the effect of noise in the system, which could be due to the presence of an adversary that may be trying to gain information about the system or trying to sabotage it. The goal is to show that if the ideal and real systems are "close", determined by some measure of differentiability, then a protocol that uses the ideal system remains secure (up to the same parameter of closeness) when the ideal system is replaced by the real system. In Section 3.5, we introduce some definitions and state some lemmas that are fundamental to the classical framework.

There has been work analyzing classical cryptographic primitives in the framework including private-key algorithms [2, 3], public-key algorithms [4, 5], digital signatures [6], key-encapsulation [7, 8], and other primitives such as one-time pads [9] and coin-tossing [10]. There is also some preliminary work on analyzing the security of quantum cryptographic protocols in this framework including device-dependent quantum key distribution

[11], quantum message transmission [12], and also in showing the impossibility of quantum and relativistic bit commitment [13]. While previous work addresses complete cryptographic protocols, the analysis of basic primitives such as random bit generators remains open. In section 1.1, we highlight our main contributions and progress towards the analysis of random bit generators in the constructive framework.

Quantum Key Distribution (QKD) is a major research area in quantum cryptography. The idea of QKD is to distribute a shared classical secret key to two parties by using quantum resources. There are two main ways of implementing QKD schemes:

1. *Prepare-and-send*, where Alice prepares a number of qubits in one of two incompatible bases and sends them to Bob. He then measures the states to obtain a binary string based on his measurement outcomes. They then use classical communication to select the bits that are part of the key (those where Alice and Bob selected the same basis), and then do other error correction and privacy amplification to make sure they end up with the same key, and that an adversary has minimal information about the shared key.

2. *Entanglement-based*, where Alice and Bob share a number of EPR pairs [14]. They also select which rounds in the protocol are *testing* rounds, and which ones are *key-generation* rounds. The key-generation rounds are used to generate the shared secret. The testing rounds are used to test if the value of a certain inequality, known as a Bell inequality, is equal to a specific value. If the value of the Bell inequality is what they expected, they use monogamy of entanglement to show that the adversary cannot have any correlation with the key bits that they both received.

The first QKD protocol was given in the seminal paper [15]. Most prepare-and-send protocols require the use of trusted devices for key generation. Proofs of security against quantum adversaries followed in [16, 17, 18, 19]. While theoretically secure, real world implementation challenges and imperfect devices gave rise to side channel attacks and cheating strategies [20, 21, 22, 23, 24]. Some of the practical challenges like photon-number-splitting attacks were overcome using decoy states [25, 26].

In [27], Mayers and Yao proposed the idea of 'device-independence', where the protocol should produce a secure key-bit even in the presence of untrusted and imperfect devices. Entanglement-based schemes form the basis of device-independent QKD. The first fully device-independent QKD protocol was given in [28], and was based on the well known CHSH inequality [29]. For a brief introduction and history of device-independent QKD, see section 3.2.4. The initial proofs of security for DIQKD [28, 30] were quite complicated and were subsequently simplified by the Entropy Accumulation Theorem [31]. The simplified proofs were given by [32]. The one limitation of existing literature is that most DIQKD protocols are based on the CHSH inequality [29]. The CHSH inequality is not symmetric, which means that the two parties in the protocol execute different strategies based on their respective inputs. DIQKD protocols are typically split into two kinds of rounds: (i) testing rounds and (ii) key-generation rounds. The adversary does not know which rounds will be testing versus key-generation and thus must come up with a strategy that helps her cheat in both cases. Proofs of security rely on the fact that the adversary cannot cheat in both types of rounds simultaneously and therefore checking for the violation of the CHSH inequality in the testing rounds helps bound the entropy of the outputs in the key-generation rounds. In CHSH based protocols, the two parties also need additional pre-shared randomness to

select which rounds of the protocol act as testing versus key-generation rounds.

Finally, a key assumption in device-independent protocols is that the two parties are not allowed to communicate with each other once the receive their inputs and before they produce their outputs. Most loophole-free experiments [33, 34, 35] separate Alice and Bob at a distance far enough such that the measurement signal of one party is outside the light cone of the other, thus preventing signaling by the fact that no communication is possible faster than the speed of light. However this method requires distant separation of the two parties which introduces other practical challenges like maintaining highly entangled quantum states over that distance. In section 1.1 we highlight how we solve some of the challenges and relax some of the requirements associated with CHSH based protocols.

1.1 Overview of contributions

There are three main contributions this dissertation makes: (i) introducing new primitives for random bit generators in the Constructive Cryptography framework, (ii) a fully device-independent quantum key distribution protocol, and (iii) a quantum lattice sieving algorithm that attempts to improve the memory and runtime complexity of lattice sieving algorithms.

1.1.1 Random bit generators

In a device-independent QKD protocol, Alice and Bob both use random bits to determine which measurement they should perform on their shared EPR pair. As a result, random bit generators play a very crucial role in the working of the protocol. In general, random

number generators are very basic, but rather very critical primitives in a lot of cryptographic protocols, which motivates their study in the constructive cryptography framework. To that effect, in Chapter 4 we define **ideal and real random bit generator models** along with ideal and real quantum random bit generators using process tomography of one-qubit quantum channels, and analyze their security in the constructive framework.

1.1.2 Synchronous device-independent quantum key distribution

We show a **fully device-independent quantum key distribution protocol** based on *synchronous* correlations in Chapter 6. The mathematical analysis of synchronous correlations first appeared in our work [36] and is also revisited in Chapter 5. The main contributions and advantages of our protocol over other CHSH based protocols are as follows.

1. The protocol is symmetric with respect to Alice and Bob, each performs exactly the same tasks.

2. Alice and Bob do not need to share randomness before starting the protocol, there is no need to pre-select which rounds of the protocol are testing rounds and which ones are key-generation rounds as these will be determined based on their independent selection of measurement bases.

3. Each round of the protocol is effective, by which we mean every testing round improves the estimate of the Bell violation (6.3) and every key generation round either improves the bound on any asynchronicity (6.7) or produces one bit of uniform shared secret.

4. Our protocol is well suited for circumventing the causality loophole common to device-independent protocols. By bounding the asynchronicity in the protocol, we can bound an adversary's maximum uncertainty about Alice and Bob's measurement selections.

The synchronicity condition in our protocol creates a loophole: can we achieve a maximal Bell violation using *asynchronous* correlations and no maximally entangled states? We answer this in the affirmative in Chapter 5 and show an example classical asynchronous strategy that achieves a maximal Bell violation. We also show how to close this asynchronicity loophole and extend our rigidity result to the almost-synchronous case in Section 6.4. Finally, we close the locality or causality loophole in device-independent schemes by introducing a new security assumption in Section 6.6. As a result, we have a protocol that is very well suited to real world implementation where the symmetry property means that the exact same devices can be manufactured for Alice and Bob, and the closing of the causality loophole eliminates the distance requirement present in prior experimental work.

1.1.3 Quantum lattice sieving

The third part, which has been included in Appendix A, gives a **quantum sieving algorithm** for finding the shortest vector in lattices. Lattices form the basis of many post-quantum cryptographic algorithms and this work is an important step towards the crypt-analysis of these algorithms.

1.2 A note about mathematical notation

Since the work in this dissertation involves topics in classical as well as quantum cryptogra-

phy, there is a very big overlap in the notation that is typically found in literature. In order to be consistent with existing literature, and also reduce confusion, we introduce notation in the sections as we go instead of redefining new global notation that deviates from what exists in literature.

Chapter 2: Preliminaries

2.1 Quantum Background

In this section we provide some background on basic concepts in quantum information theory relevant to our discussion, and introduce notation that we will use in this section. This section is not extensive, and we point the interested reader to [37, 38, 39, 40] for a more thorough introduction to quantum information and computing.

A qubit is a unit vector in a two-dimensional complex vector space also known as a Hilbert space, which we denote using $\mathfrak{H} = \mathbb{C}^2$. We write a qubit as $|\phi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$, where $\alpha_0, \alpha_1 \in \mathbb{C}$, and $|0\rangle = (1,0)^T$ and $|1\rangle = (0,1)^T$ are known as the computational basis states. Since $|\psi\rangle$ is a unit vector, $|\alpha_0|^2 + |\alpha_1|^2 = 1$. Quantum states in separate registers are represented using the tensor product. For example, if $|\phi_A\rangle \in \mathfrak{H}_A$ and $|\phi_B\rangle \in \mathfrak{H}_B$, then $|\phi_A\rangle \otimes |\phi_B\rangle \in \mathfrak{H}_A \otimes \mathfrak{H}_B$. To simplify notation, tensor product states are sometimes written without the $\otimes$ symbol, eg. $|0\rangle \otimes |0\rangle$ is often written as $|00\rangle$. A maximally entangled state on two qubits, also known as a Bell pair or EPR pair is denoted as $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Given a matrix $\mathbf{A} = \{a_{ij}\}$, we write $\mathrm{tr}(\mathbf{A}) = \sum_i a_{ii}$, which is the sum of the diagonal entries. We denote by $\mathbf{A}^\dagger = \{a^*_{ji}\}$ the adjoint or conjugate-transpose of $\mathbf{A}$. The three Pauli

matrices, which have special symbols, are given by

$$\sigma_X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \sigma_Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \text{ and } \sigma_Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Note that $\mathbf{P}^\dagger = \mathbf{P}$ for $\mathbf{P} \in \{\sigma_X, \sigma_Y, \sigma_Z\}$ and $\text{tr}(\sigma_X) = \text{tr}(\sigma_Y) = \text{tr}(\sigma_Z) = 0$. We denote by $\mathbb{1}_d$ the identity matrix of dimension d, and omit the dimension when it is clear from the context. The single-qubit Hadamard gate is given by $\mathbf{H} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. A matrix $\mathbf{A}$ is Hermitian if it is equal to its conjugate-transpose, i.e. $\mathbf{A} = \mathbf{A}^\dagger$. A matrix $\mathbf{U}$ is called unitary if $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. Unitary matrices (or operators) preserve inner products, i.e. if $|\psi\rangle = \mathbf{U}|\phi\rangle$, then $\langle\psi|\psi\rangle = \langle\phi|\mathbf{U}^\dagger\mathbf{U}|\phi\rangle = \langle\phi|\phi\rangle$, and therefore, quantum gates are represented using unitary matrices.

We also use an equivalent description of quantum states given by *density* operators, which are defined as:

Definition 1 (Density operator)**.** *A density operator, or density matrix, ρ is defined as a Hermitian operator that satisfies the following properties:*

- *It must be positive, which means $\langle\psi|\rho|\psi\rangle \geq 0$ for any $|\psi\rangle$.*
- *It must have trace* 1*, denoted as* $\text{tr}(\rho) = 1$.

Definition 2 (Quantum channel)**.** *[39, Definition 4.4.3] A quantum channel is a linear, completely positive, trace preserving map, corresponding to a quantum physical evolution*

A quantum channel $\mathcal{C}$ is a general way to describe the evolution of quantum systems. Suppose we have a system initially in state ρ_{in}, that interacts with the environment in some

state ρ_{env}, then the evolution of the system is unitary and the final state of our system is obtained by tracing out the environment.

$$\mathcal{C}(\rho_{in}) = \mathrm{tr}_{env}\left(U(\rho_{in} \otimes \rho_{env})U^\dagger\right) = \rho'.$$

A quantum channel acting on density operator ρ can be written in its operator sum representation:

$$\mathcal{C}(\rho) = \sum_k \mathbf{E}_k \rho \mathbf{E}_k^\dagger \tag{2.1}$$

where $\{\mathbf{E}_k\}$ are known as the Kraus operators for the channel, with $\sum_k \mathbf{E}_k^\dagger \mathbf{E}_k = \mathbb{1}$ to preserve traces [37, Section 8.2.3]. Kraus operators for a channel are generally not unique.

2.2 Process Tomography

We develop the theory needed to describe random bit generators based on process tomography. In general, process tomography is a method of characterizing what a quantum channel does by learning its process matrix. We use it to show how to get a bit with the maximum min-entropy by performing the right measurement on a given channel. Min-entropy is given by $H_{\min} = -\log \max_i p_i$ for a probability distribution $\{p_i\}$. We focus on min-entropy because it has uses in quantum key distribution where one of the classical post-processing steps requires a high min-entropy source. We defer the analysis of QKD protocols to Section 3.2.

In order to generate random bits, we want to start by preparing a qubit in the $|0\rangle$

state, perform a quantum operation on it, and then measure the state to obtain a bit with maximum min-entropy. However, in order to design the right measurement that optimizes our min-entropy, we need to characterize the quantum operation using process tomography. Once we have the full characterization of the quantum channel, we make the measurement and obtain the required random bit. An example of a channel one might use to obtain random bits is given in Example (3) below. In this section we develop the theory needed to characterize one-qubit channels.

We start with a brief discussion of quantum channels and their properties. Using eq. (2.1), a quantum channel $\mathcal{C}$ acting on a one-qubit state ρ can be written as

$$\begin{aligned}\mathcal{C}(\rho) &= \sum_k \mathbf{E}_k \rho \mathbf{E}_k^\dagger \\ &= \sum_k \sum_{\mu,\nu} e_{k\mu} e_{k\nu}^* \sigma_\mu \, \rho \, \sigma_\nu^\dagger \\ &= \sum_{\mu,\nu} \chi_{\mu\nu} \, \sigma_\mu \, \rho \, \sigma_\nu^\dagger \end{aligned} \tag{2.2}$$

where in the second equality we wrote $\mathbf{E}_k = \sum_\mu e_{k\mu}\sigma_\mu$ for some complex numbers $e_{k\mu}$ since the Pauli matrices form an operator basis for $\mathbb{C}^2$, and in the third equality we wrote $\chi_{\mu\nu} = \sum_k e_{k\mu} e_{k\nu}^*$. $\mathbf{X} = \{\chi_{\mu\nu}\}$ is known as the process matrix of the channel. Kraus operators for a channel are not unique in general as there may be unitary freedom in choosing the operator sum representation. However, the process matrix for a channel is always unique. Density operators are Hermitian, so taking the adjoint of eq. (2.2), we get

$$\mathcal{C}(\rho)^\dagger = \sum_{\mu,\nu} \overline{\chi_{\mu\nu}} \, \sigma_\nu \, \rho \, \sigma_\mu^\dagger = \sum_{\mu,\nu} \overline{\chi_{\nu\mu}} \, \sigma_\mu \, \rho \, \sigma_\nu^\dagger$$

and hence $\chi_{\mu\nu} = \overline{\chi_{\nu\mu}}$, which means that $\mathbf{X}$ is Hermitian. Since $\chi_{\mu\nu} \in \mathbb{C}$, we can write the entries of $\mathbf{X}$ in the form $a + bi$, where $i = \sqrt{-1}$.

$$\mathbf{X} = \begin{bmatrix} a_0 & a_1 + ib_1 & a_2 + ib_2 & a_3 + ib_3 \\ a_1 - ib_1 & a_4 & a_5 + ib_5 & a_6 + ib_6 \\ a_2 - ib_2 & a_5 - ib_5 & a_7 & a_8 + ib_8 \\ a_3 - ib_3 & a_6 - ib_6 & a_8 - ib_8 & a_9 \end{bmatrix}$$

We examine $\mathbf{X}$ and see that it satisfies certain key properties. Since the output of the channel is a density operator we know that it has trace one. Using $\rho = \frac{1}{2}\mathbb{1}$, we get:

$$1 = \text{tr}\left[\mathcal{C}\left(\frac{1}{2}\mathbb{1}\right)\right] = \frac{1}{2}\sum_{\mu,\nu} \chi_{\mu\nu}\text{tr}[\sigma_\mu\sigma_\nu]$$

For $\mu \neq \nu$, $\text{tr}[\sigma_\mu\sigma_\nu] = 0$ and for $\mu = \nu$, $\text{tr}[\sigma_\mu\sigma_\nu] = \text{tr}[\mathbb{1}] = 2$. Therefore,

$$\begin{aligned} 1 &= \frac{1}{2}\sum_{\mu} \chi_{\mu\mu}\text{tr}[\sigma_\mu\sigma_\mu] \\ &= \sum_{\mu} \chi_{\mu\mu} \end{aligned}$$

Similarly, using $\rho \in \{|0\rangle\langle 0|, |+\rangle\langle +|, |i\rangle\langle i|\}$ and $\sum_\mu \chi_{\mu\mu} = a_0 + a_4 + a_7 + a_9 = 1$, we get the following equalities:

$$\text{tr}[\mathcal{C}(|0\rangle\langle 0|)] = a_0 + 2a_3 + a_4 + a_7 + a_9 + 2b_5$$

$$1 = 1 + 2a_3 + 2b_5$$

$$\therefore b_5 = -a_3$$

$$\mathrm{tr}\left[\mathcal{C}(|i\rangle\langle i|)\right] = a_0 + 2a_2 + a_4 + a_7 + a_9 - 2b_6$$

$$\therefore b_6 = a_2$$

$$\mathrm{tr}\left[\mathcal{C}(|+\rangle\langle +|)\right] = a_0 + 2a_1 + a_4 + a_7 + a_9 + 2b_8$$

$$\therefore b_8 = -a_1$$

Let Π^0_σ (respectively Π^1_σ) be the projector onto the $+1$ (respectively -1) eigenspace of $\sigma \in \{\sigma_X, \sigma_Y, \sigma_Z\}$. We define probability variables $p_\rho^{\Pi^j_\sigma} \equiv \mathrm{tr}[\Pi^j_\sigma \mathcal{C}(\rho)]$. Using the equation,

$$p_\rho^{\Pi^j_\sigma} = \sum_{\mu,\nu} \chi_{\mu\nu} \mathrm{tr}[\Pi^j_\sigma \sigma_\mu \rho \sigma_\nu^\dagger],$$

we obtain the system of equations,

$$\begin{aligned}
p_{|0\rangle\langle 0|}^{\Pi^0_{\sigma_Z}} &= a_0 + 2a_3 + a_9 \\
p_{|0\rangle\langle 0|}^{\Pi^1_{\sigma_Z}} &= -2a_3 + a_4 + a_7 \\
p_{|1\rangle\langle 1|}^{\Pi^0_{\sigma_Z}} &= 2a_3 + a_4 + a_7 \\
p_{|+\rangle\langle +|}^{\Pi^0_{\sigma_Z}} &= \frac{1}{2}a_0 + 2a_3 + \frac{1}{2}a_4 + a_6 + \frac{1}{2}a_7 + \frac{1}{2}a_9 - b_2 \\
p_{|-\rangle\langle -|}^{\Pi^0_{\sigma_Z}} &= \frac{1}{2}a_0 + 2a_3 + \frac{1}{2}a_4 + \frac{1}{2}a_7 + a_8 + \frac{1}{2}a_9 + b_1 \\
p_{|0\rangle\langle 0|}^{\Pi^0_{\sigma_Y}} &= \frac{1}{2}a_0 + 2a_2 + \frac{1}{2}a_4 + \frac{1}{2}a_7 + a_8 + \frac{1}{2}a_9 - b_1 \\
p_{|1\rangle\langle 1|}^{\Pi^0_{\sigma_Y}} &= \frac{1}{2}a_0 + 2a_2 + \frac{1}{2}a_4 + \frac{1}{2}a_7 - a_8 + \frac{1}{2}a_9 + b_1 \\
p_{|+\rangle\langle +|}^{\Pi^0_{\sigma_Y}} &= \frac{1}{2}a_0 + 2a_2 + \frac{1}{2}a_4 + a_5 + \frac{1}{2}a_7 + \frac{1}{2}a_9 + b_3 \\
p_{|-\rangle\langle -|}^{\Pi^0_{\sigma_Y}} &= a_0 + 2a_2 + a_7 \\
p_{|0\rangle\langle 0|}^{\Pi^0_{\sigma_X}} &= \frac{1}{2}a_0 + 2a_1 + \frac{1}{2}a_4 + a_6 + \frac{1}{2}a_7 + \frac{1}{2}a_9 + b_2
\end{aligned}$$

$$
\begin{aligned}
p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} &= \frac{1}{2}a_0 + 2a_1 + \frac{1}{2}a_4 - a_6 + \frac{1}{2}a_7 + \frac{1}{2}a_9 - b_2 \\
p^{\Pi^0_{\sigma_X}}_{|+\rangle\langle +|} &= a_0 + 2a_1 + a_4 \\
p^{\Pi^0_{\sigma_X}}_{|-\rangle\langle -|} &= \frac{1}{2}a_0 + 2a_1 + \frac{1}{2}a_4 + a_5 + \frac{1}{2}a_7 + \frac{1}{2}a_9 - b_3
\end{aligned}
$$

Solving this system of linear equations, we determine all the a's and b's as follows:

$$
\begin{aligned}
a_0 &= -\frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_X}}_{|+\rangle\langle +|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} \\
&\qquad + \frac{1}{2}\,p^{\Pi^0_{\sigma_Y}}_{|-\rangle\langle -|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{4}\,p^{\Pi^1_{\sigma_Z}}_{|0\rangle\langle 0|} \\
a_1 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} - \frac{1}{4} \\
a_2 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} - \frac{1}{4} \\
a_3 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} - \frac{1}{4}\,p^{\Pi^1_{\sigma_Z}}_{|0\rangle\langle 0|} \\
a_4 &= -\frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_X}}_{|+\rangle\langle +|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} - \frac{1}{2}\,p^{\Pi^0_{\sigma_Y}}_{|-\rangle\langle -|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{4}\,p^{\Pi^1_{\sigma_Z}}_{|0\rangle\langle 0|} \\
a_5 &= -\frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_X}}_{|-\rangle\langle -|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Y}}_{|+\rangle\langle +|} \\
a_6 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Z}}_{|+\rangle\langle +|} \\
a_7 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} - \frac{1}{2}\,p^{\Pi^0_{\sigma_X}}_{|+\rangle\langle +|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Y}}_{|-\rangle\langle -|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{4}\,p^{\Pi^1_{\sigma_Z}}_{|0\rangle\langle 0|} \\
a_8 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Z}}_{|-\rangle\langle -|} \\
a_9 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} - \frac{1}{2}\,p^{\Pi^0_{\sigma_X}}_{|+\rangle\langle +|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} \\
&\qquad - \frac{1}{2}\,p^{\Pi^0_{\sigma_Y}}_{|-\rangle\langle -|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{4}\,p^{\Pi^1_{\sigma_Z}}_{|0\rangle\langle 0|} \\
b_1 &= -\frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Z}}_{|-\rangle\langle -|} \\
b_2 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_Z}}_{|1\rangle\langle 1|} - \frac{1}{2}\,p^{\Pi^0_{\sigma_Z}}_{|+\rangle\langle +|} \\
b_3 &= \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} + \frac{1}{4}\,p^{\Pi^0_{\sigma_X}}_{|1\rangle\langle 1|} - \frac{1}{2}\,p^{\Pi^0_{\sigma_X}}_{|-\rangle\langle -|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{4}\,p^{\Pi^0_{\sigma_Y}}_{|1\rangle\langle 1|} + \frac{1}{2}\,p^{\Pi^0_{\sigma_Y}}_{|+\rangle\langle +|}
\end{aligned}
$$

Consider a generic measurement $\mathcal{M} = \{\Pi^0 = |\psi\rangle\langle\psi|, \Pi^1 = |\psi^\perp\rangle\langle\psi^\perp|\}$ written in Bloch sphere coordinates, where $|\psi\rangle = \cos(\frac{\theta}{2})|0\rangle + e^{i\phi}\sin(\frac{\theta}{2})|1\rangle$. Our goal is to optimize for θ and ϕ, so that we obtain the maximum min-entropy after measuring state $\rho' = \mathcal{C}(\rho)$ Using $\rho = |0\rangle\langle 0|$ and our measurement basis, the probability of getting bit 0 is given by $\mathrm{tr}[\Pi^0\mathcal{C}(|0\rangle\langle 0|)]$. Writing the channel in terms of its process matrix form, we get,

$$\begin{aligned}
\Pr[0] &= \mathrm{tr}[\Pi^0\mathcal{C}(|0\rangle\langle 0|)] \\
&= \sum_{\mu\nu}\chi_{\mu\nu}\mathrm{tr}[\Pi^0\sigma_\mu|0\rangle\langle 0|\sigma_\nu^\dagger] \\
&= \sum_{\mu\nu}\chi_{\mu\nu}\langle 0|\sigma_\nu^\dagger\Pi^0\sigma_\mu|0\rangle \\
&= \sin(\theta)\cos(\phi)\left(p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \sin(\theta)\sin(\phi)\left(p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \cos(\theta)\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \frac{1}{2}
\end{aligned} \tag{2.3}$$

Recall that min-entropy is given by $H_\infty = \min(-\log p_i)$. In our case we only have two outcomes, and therefore the maximum min-entropy achievable is 1, when $\Pr[0] = \Pr[1] = \frac{1}{2}$. Therefore in order to maximize min-entropy we minimize the function $f(\theta,\phi) = |\Pr[0] - \frac{1}{2}|$. Working directly with eq. (2.3), we see that in order to get $\Pr[0] = \frac{1}{2}$, we want

$$\sin(\theta)\cos(\phi)\left(p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \sin(\theta)\sin(\phi)\left(p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \cos(\theta)\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) = 0$$

At this point, we can proceed in multiple ways. If either of $p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|}, p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|}, p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|}$ are very close to $\frac{1}{2}$, we can choose θ and ϕ accordingly to get min-entropy very close to 1.

- For example, if $p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|}$ is equal to (or very close) to $\frac{1}{2}$, we can choose $\theta = 0$ (computa-

tional basis) to get min-entropy equal to (very close) to 1.

- Similarly for $p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|}$ $\left(p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|}\right)$ equal to (close to) $\frac{1}{2}$, we choose $\theta = \frac{\pi}{2}, \phi = 0$ ($\theta = \frac{\pi}{2}, \phi = \frac{\pi}{2}$).

Alternatively, if we want to achieve min-entropy exactly 1, we can pick a phase $\phi = \phi_0$ and compute θ as

$$\theta = \texttt{atan2}\left(-\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right), \left(\cos(\phi_0)\left(p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \sin(\phi_0)\left(p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{2}\right)\right)\right)$$

If $\phi_0 = 0$, $\theta = \texttt{atan2}\left(-\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right), p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{2}\right)$
If $\phi_0 = \frac{\pi}{2}$, $\theta = \texttt{atan2}\left(-\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right), p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} - \frac{1}{2}\right)$

Example 3. *Ideally, we want to prepare our initial qubit in the $|0\rangle\langle 0|$ state, apply a Hadamard gate and measure in the computational basis to get a random bit with min-entropy 1. However in practice, due to the environment and other factors, it is hard to prepare the exact state $|0\rangle\langle 0|$ and apply the exact Hadamard operator to it. We introduce an error term in the Hadamard operator and use a bit-flip error model in our state preparation of the initial state $|0\rangle\langle 0|$. A simple calculation shows that:*

$$e^{-i\mathbf{H}t} = \cos(t)\mathbb{1} - i\sin(t)\mathbf{H}$$

$$e^{-i\mathbf{H}\frac{\pi}{2}} = -i\mathbf{H}$$

Note that $i\mathbb{1} = e^{i\mathbb{1}\frac{\pi}{2}}$, and therefore we have

$$\mathbf{H} = e^{i\mathbb{1}\frac{\pi}{2}} e^{-i\mathbf{H}\frac{\pi}{2}} = e^{-i(\mathbf{H}-\mathbb{1})\frac{\pi}{2}}$$

Now that we have $\mathbf{H}$ *written out in this form, we allow for a small error in* t *that we denote* dt*. Thus we can rewrite* $\mathbf{H}$ *with over-rotation error as* $\mathcal{H}_{dt} = exp\left(-i\left(\mathbf{H} - \mathbb{1}\right)\left(\frac{\pi}{2} + dt\right)\right)$*. Using this as our Hadamard operator, we write the channel as follows:*

$$\mathcal{H}_{dt}(\rho) = \mathcal{H}_{dt}\rho\mathcal{H}_{dt}^{\dagger}$$

Substituting the values of $\mathcal{H}_{dt}$ *and* $\mathcal{H}_{dt}^{\dagger}$*, and simplifying the expression, we get*

$$\mathcal{H}_{dt}(\rho) = \left(\sin(dt)\mathbb{1} + i\cos(dt)\mathbf{H}\right)\rho\left(\sin(dt)\mathbb{1} - i\cos(dt)\mathbf{H}\right)$$

We use this description of the Hadamard in conjunction with the error channel: $\mathcal{E}(\rho) = (1-p)\rho + p\sigma_X\rho\sigma_X$*. Therefore our combined channel can be written as*

$$\mathcal{H}_{dt}(\mathcal{E}(\rho)) = (1-p)\mathcal{H}_{dt}\rho\mathcal{H}_{dt}^{\dagger} + p\mathcal{H}_{dt}\sigma_X\rho\sigma_X\mathcal{H}_{dt}^{\dagger}.$$

Using $|0\rangle\langle 0|$ *as our initial state the channel is written as*

$$\mathcal{C}(|0\rangle\langle 0|) = \mathcal{H}_{dt}(\mathcal{E}(|0\rangle\langle 0|)) = (1-p)\mathcal{H}_{dt}|0\rangle\langle 0|\mathcal{H}_{dt}^{\dagger} + p\mathcal{H}_{dt}|1\rangle\langle 1|\mathcal{H}_{dt}^{\dagger}$$

Using our method described, we compute $p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|}, p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|}, p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|}$ *for* $\mathcal{C}(|0\rangle\langle 0|)$ *and get,*

$$\begin{aligned}
p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} &= -\frac{1}{2}\left(2p-1\right)\cos^2\left(dt\right) + \frac{1}{2} \\
p^{\Pi^0_{\sigma_Y}}_{|0\rangle\langle 0|} &= -\frac{1}{2}\left(2\sqrt{2}p - \sqrt{2}\right)\cos\left(dt\right)\sin\left(dt\right) + \frac{1}{2} \\
p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} &= -\frac{1}{2}\left(2p-1\right)\sin^2\left(dt\right) + \frac{1}{2}
\end{aligned}$$

Using phase $\phi = 0$, we compute θ as

$$\theta = \texttt{atan2}\left(\frac{1}{2}(2p-1)\sin^2(dt), -\frac{1}{2}(2p-1)\cos^2(dt)\right)$$

This gives us $H_\infty = -\log(\max p_i) = -\log(\Pr[0]) = -\log\left(\frac{1}{2}\right) = 1$

Chapter 3: Literature Review

3.1 Nonlocal Games

In this section we review the literature on nonlocal games in the context of quantum key distribution. Two player nonlocal games are characterized by a tuple $(\mathcal{X}_A, \mathcal{X}_B, \mathcal{Y}_A, \mathcal{Y}_B, \mu, \mathcal{D})$ where the terms in the tuple is as follows:

- $\mathcal{X}_A, \mathcal{X}_B$: Finite sets of *input* (or *question*) alphabets for Alice and Bob respectively,
- $\mathcal{Y}_A, \mathcal{Y}_B$: Finite sets of *output* (or *answer*) alphabets for Alice and Bob respectively,
- μ: Probability distribution on $\mathcal{X}_A \times \mathcal{X}_B$,
- $D : \mathcal{X}_A \times \mathcal{X}_B \times \mathcal{Y}_A \times \mathcal{Y}_B \to \{0, 1\}$: The decision predicate that determines whether the players win or lose the game.

The two-player setting can easily be generalized to k-player games, where we have an input and output alphabet for each player involved, the probability distribution μ is defined over all input sets, and the decision predicate D takes as input all k questions and answers, and outputs a bit to determine whether the k players win or lose the game. For the rest of the section and thesis, we focus on two player nonlocal games. We think of nonlocal games in terms of the probability of producing specific outputs given specific inputs. We maximize

the value:

$$v(p) = \sum_{\substack{y_A, y_B \\ x_A, x_B}} D(x_A, x_B, y_A, y_B) p(y_A, y_B \mid x_A, x_B) \mu(x_A, x_B)$$

Definition 4 (Classical value)**.** *The classical value v_c^* of a game is obtained by maximizing $v(p)$ over probability distributions of the form:*

$$p(y_A, y_B \mid x_A, x_B) = \sum_{\omega \in \Omega} p_\Omega(\omega) p_A(y_A \mid x_A, \omega) p_B(y_B \mid x_B, \omega)$$

where Ω is a finite set.

Definition 5. *The quantum value v_q^* of a game is obtained by maximizing $v(p)$ over probability distributions of the form:*

$$p(y_A, y_B \mid x_A, x_B) = \mathrm{tr}((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})\rho)$$

where $\{\{E_y^x\}_{y \in \mathcal{Y}}\}_{x \in \mathcal{X}}$ and $\{\{F_y^x\}_{y \in \mathcal{Y}}\}_{x \in \mathcal{X}}$ are POVMs on $\mathfrak{H}_A$ and $\mathfrak{H}_B$ respectively, and ρ is a pure state on $\mathfrak{H}_A \otimes \mathfrak{H}_B$.

3.1.1 CHSH game

The CHSH game is named after its authors Clauser, Horne, Shimony, and Holt [29]. In this game the input consists of bits, i.e. $\mathcal{X}_A = \mathcal{X}_B = \mathcal{X} = \{0, 1\}$ and the output also consists of

bits $\mathcal{Y}_A = \mathcal{Y}_B = \mathcal{Y} = \{0, 1\}$. The decision predicate D is given by

$$D(x_A, x_B, y_A, y_B) = \begin{cases} 1 & \text{if } x_A \wedge x_B = y_A \oplus y_B \\ 0 & \text{otherwise} \end{cases}$$

The maximum winning probability using classical correlations is $\frac{3}{4}$ and is given by a simple strategy: Alice and Bob always return 0 (or 1) independent of their input bits.

We can win the win with probability higher than the classical winning probability by using quantum correlations. The maximum quantum winning probability is $\cos^2(\frac{\pi}{8}) \approx 0.85$, and is given by the quantum strategy: Alice and Bob share a maximally entangled quantum state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Upon receiving $x_A = 0$, Alice measures her qubit in the computational basis (Z basis), and outputs $y_A = 0$ for the $+1$ outcome and $y_A = 1$ for the -1 measurement outcome. If she receives $x_A = 1$, she measures her qubit in the Hadamard basis (X basis) and outputs $y_A = 0$ (or $y_A = 1$) corresponding to the $+1$ (-1) outcome. If Bob receives $x_B = 0$, he measures his qubit in the $\frac{X+Z}{\sqrt{2}}$ and returns y_B depending on the measurement outcome. If he receives $x_B = 1$, he measures in the $\frac{Z-X}{\sqrt{2}}$ basis and outputs y_B depending on the measurement outcome.

To further develop this notion of a *gap* between the classical and quantum value of a nonlocal game, let's look at the difference between the winning and losing probability of the game. Let $\{E_{y_A}^{x_A}\}$ and $\{F_{y_B}^{x_B}\}$ be Alice and Bob's projection valued measures. Given inputs $00, 01$, and 10, Alice and Bob need to produce the same outputs in order to win.

The probability for producing same outputs is given by:

$$\langle \Phi^+ | E_0^{x_A} \otimes F_0^{x_B} | \Phi^+ \rangle + \langle \Phi^+ | E_1^{x_A} \otimes F_1^{x_B} | \Phi^+ \rangle$$

The probability that they produce different outputs is given by:

$$\langle \Phi^+ | E_0^{x_A} \otimes F_1^{x_B} | \Phi^+ \rangle + \langle \Phi^+ | E_1^{x_A} \otimes F_0^{x_B} | \Phi^+ \rangle$$

In the event that they receive inputs 11, they win if they produce different outputs which happens with probability $\langle \Phi^+ | E_0^{x_A} \otimes F_1^{x_B} | \Phi^+ \rangle + \langle \Phi^+ | E_1^{x_A} \otimes F_0^{x_B} | \Phi^+ \rangle$, while they produce the same outputs with probability $\langle \Phi^+ | E_0^{x_A} \otimes F_0^{x_B} | \Phi^+ \rangle + \langle \Phi^+ | E_1^{x_A} \otimes F_1^{x_B} | \Phi^+ \rangle$

We define observables M_{x_A} and N_{x_B} as follows:

$$M_{x_A} := E_0^{x_A} - E_1^{x_A}$$

$$N_{x_B} := F_0^{x_B} - F_1^{x_B}$$

The difference between the winning and losing probability is given by:

$$\langle \Phi^+ | M_0 \otimes N_0 + M_0 \otimes N_1 + M_1 \otimes N_0 - M_1 \otimes N_1 | \Phi^+ \rangle \tag{3.1}$$

We write $C := M_0 \otimes N_0 + M_0 \otimes N_1 + M_1 \otimes N_0 - M_1 \otimes N_1$ to denote the CHSH operator. We discuss some bounds on the operator norm of C. Bell [41] showed that in the classical (or hidden variable theories) case, $|C| \leq 2$, and subsequently Tsirelson [42] showed that $|C| \leq 2\sqrt{2}$ in the quantum case. Hence, inqualities of this form are often referred to

as Bell inequalities or Tsirelson inequalities. A violation of the Bell inequality lets us certify quantum behavior, and a maximal violation of the inequality lets us certify specific entangled quantum states. This property is known as *self-testing* or *rigidity*. The CHSH game is rigid: a maximal violation can only be achieved using a state equivalent to a Bell pair, thus observing a maximal violation certifies that we must have a Bell pair. In the CHSH case, there is a direct relationship between the maximum winning probability of the game and the value of the Bell inequality and is given by $w = \frac{1}{8}\beta + \frac{1}{2}$, where w is the winning probability, and β is the value of the Bell term.

The CHSH game forms the basis of many device independent quantum key distribution protocols, which we explore in Section 3.2.

3.1.2 GHZ game

The Greenberger-Horne-Zeilinger (GHZ) game is a nonlocal game played between three players Alice, Bob and Carol. The referee provides inputs to the three parties $(x_A, x_B, x_C) \in \{000, 011, 101, 110\}$ and the three parties produce outputs $(y_A, y_B, y_C) \in \{0,1\}^3$. The decision predicate D is given by:

$$D(x_A, x_B, x_C, y_A, y_B, y_C) = \begin{cases} 1 & \text{if } x_A \vee x_B \vee x_C = y_A \oplus y_B \oplus y_C \\ 0 & \text{otherwise} \end{cases}$$

The optimal classical strategy wins with probability $\frac{3}{4}$. One example strategy that achieves this bound is: All three parties always output 1 no matter the input. We have $1 \oplus 1 \oplus 1 = 1$ and $x_A \vee x_B \vee x_C = 1$ for 3 out of the 4 input choices.

In contrast, it is possible to win the game perfectly (i.e. with probability 1) using a quantum strategy. The strategy is as follows: The three players share the GHZ state: $|\mathsf{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$. Each player uses the same measurement strategy. On input 0, the player measures in the Hadamard basis ($\{|+\rangle, |-\rangle\}$). On input 1, the player measures in the Y basis ($|+i\rangle, |-i\rangle$). This strategy wins the game with probability 1. The GHZ game is rigid, and thus winning the game with certainty certifies that the three players hold the GHZ state. This also means their measurement operators must be isomorphic (up to local unitaries) to the ideal winning strategy discussed above.

3.1.3 Magic Square game

The Magic Square game was studied independently by Mermin [43] and Peres [44]. In the game, Alice and Bob draw inputs $x_A, x_B \in \{0, 1, 2\}$. Alice produces output bits $m_0, m_1, m_2 = m_0 \oplus m_1$, while Bob produces bits $n_0, n_1, n_2 = 1 \oplus n_0 \oplus n_1$. They win the game if $m_{x_B} = n_{x_A}$.

One can think about this game as filling a 3×3 grid with 0s and 1s. Alice's role is to fill out the rows such that the parity of each row is even, while Bob's role is to fill out the columns such that the parity of each column is odd. They win the game if their outputs agree for a randomly chosen cell in the grid. Classically, there exist strategies where Alice and Bob can produce outputs that agree on 8 out of the 9 cells, and thus they can win the game with probability $\frac{8}{9}$. It turns out that this is the best they can hope to achieve classically. However in the quantum case, they can win the game with certainty by using

$Z \otimes \mathbb{1}$	$\mathbb{1} \otimes Z$	$Z \otimes Z$
$\mathbb{1} \otimes X$	$X \otimes \mathbb{1}$	$X \otimes X$
$-Z \otimes X$	$-X \otimes Z$	$Y \otimes Y$

Table 3.1: Alice and Bob's observables for the Magic Square game

the strategy: Alice and Bob share two copies of the EPR pair:

$$|\Phi^+\rangle \otimes |\Phi^+\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B\right) \otimes \frac{1}{\sqrt{2}} \left(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B\right)$$

and make measurements according to Table 3.1 depending on their inputs, where X, Y and Z are the Pauli matrices. Note that the observables in each row and each column commute with each other. As a result, observables in any given row (or column) are diagonalizable in the same basis, and thus can be measured simultaneously. This allows Alice and Bob to produce matching outputs every time and win the game with probability 1. A nice property of the strategy listed in Table 3.1 is that it is rigid. It can be used to self-test two copies of the EPR pair.

3.2 Quantum Key Distribution

3.2.1 Introduction

Quantum Key Distribution (QKD) aims to utilize quantum resources in order to establish a shared classical secret key between two or more players. The two primary objectives of QKD are 'correctness' and 'security'. By correctness we mean that the two parties must possess the same key at the end of the protocol, and by security we mean that any adversary in the protocol obtains minimal to no knowledge about the secret key at the

end of the protocol. These terms are also often extended to include noisy implementations and known as 'completeness' (correctness in an honest but possibly noisy execution of the protocol) and 'soundness' (security in the presence of an adversary) conditions. We will discuss completeness and soundness conditions in more detail later in Section 6.5 when we describe our quantum key distribution protocol.

QKD protocols can be constructed using two paradigms: (i) prepare-and-send and (ii) entanglement-based. We discuss these in more detail.

3.2.2 Prepare-and-send protocols

Bennett and Brassard [15] gave the first quantum key distribution protocol, and thus the protocol and its variants are known by the initials of the authors - BB84. The main idea is to transmit non-orthogonal quantum states chosen according to randomly drawn bit strings. If an adversary attempts to distinguish between the non-orthogonal states, she must do so at the expense of disturbing the quantum states. Alice and Bob can detect this interference and can abort if the error exceeds a pre-decided threshold. More formally, consider a setup with two parties Alice and Bob that want to establish a shared classical secret key. We may assume that an adversary Eve is present and is trying to get information about the shared secret key. Algorithm 1 shows the BB84 protocol. Once Alice and Bob have a subset n of data bits, they can use information reconciliation and privacy amplification to produce near-uniform key bits that are secure against a (quantum) Eve.

Algorithm 1: BB84 Protocol

Input:

$\delta \in \{0, 1\}$: Threshold for check bits

n: Expected number of raw key bits

m: Expected number of final key bits

1 Alice draws two bit strings a and b of length $(4 + \delta)n$ and uses them to produce quantum states: $|\psi_{a_k b_k}\rangle$ such that:

$$|\psi_{00}\rangle = |0\rangle$$
$$|\psi_{01}\rangle = |+\rangle$$
$$|\psi_{10}\rangle = |1\rangle$$
$$|\psi_{11}\rangle = |-\rangle.$$

2 Alice sends Bob $\bigotimes_{k=1}^{(4+\delta)n} |\psi_{a_k b_k}\rangle$.

3 Bob draws a random bit string b', and measures the $(4 + \delta)n$ states in the computational basis if $b'_k = 0$ or the Hadamard basis if $b'_k = 1$. He records his output as a'_k.

4 Alice and Bob publicly share their bitstrings b and b'.

5 They discard all a_k and a'_k where $b_k \neq b'_k$. If there are fewer than $2n$ bits left, they abort.

6 They choose a random subset n of the remaining outputs and use those as check bits. If more than δn bits disagree, they abort the protocol.

7 They use the remaining n bits to perform information reconciliation and privacy amplification to produce m final key bits.

3.2.3 Entanglement-based protocols

The E91 protocol is named after Artur Ekert who proposed the protocol in a 1991 paper [14]. Alice and Bob share maximally entangled quantum states and make measurements in either the computational or the Hadamard basis. If they measure in the same basis, they end up with identical outputs, and if they measure in different basis, the distribution of their outputs is random. After they make measurements, they share their measurement basis choices and only keep output bits where they made measurements in the same basis. One can bound Eve's information about the key bits by using the fidelity of the shared entangled

pairs. This idea laid the foundation for device-independent quantum key distribution based on nonlocal games and maximally entangled pairs.

3.2.4 Device-independent quantum key distribution

So far we have looked at protocols where Alice and Bob have full control over state preparation and measurement devices in the QKD protocol. However, in practice, Alice and Bob may wish to purchase QKD devices and equipment from a third-party and may not always trust those devices. Using such untrusted devices, can we still produce a shared secret key that is secure against quantum or more powerful adversaries? In order to answer that question, we first list some assumptions that we make about the protocol setup and the devices used in it:

1. Alice and Bob generate (local) inputs and feed them to their respective devices and receive outputs from them.

2. Alice and Bob are allowed to decide on a strategy before the execution of the protocol and may share resources like quantum states and classical randomness, but are not allowed to communicate with each other after they receive their inputs, and before they produce their outputs.

3. The devices used in the protocol may be noisy and may even be programmed by the adversary.

4. Alice, Bob, and the adversary Eve are all bound by the laws of quantum mechanics and that quantum mechanics is a complete theory.

The idea for device-independence was first introduced by Mayers and Yao in [27]. Some of the initial research in this direction gave efficient and noise-tolerant protocols [45, 46, 47, 48, 49, 50] but made some independence assumptions on the devices. In [51, 52] device-independent protocols were given that required perfect devices. The first fully device-independent QKD protocol was proposed by Vazirani and Vidick in [28]. Their protocol was based on a slightly modified version of the CHSH game, where Alice has an extra input labeled "2". The input pair $(x_A^i = 2, x_B^i = 1)$ is used to determine which rounds in the protocol are key generation rounds. Their modified game uses the rigidity results of the CHSH game to certify that Alice and Bob held maximally entangled quantum states. Their protocol is also robust to noise. The general idea for DIQKD protocols is this: the protocol is split into two types of rounds (i) testing rounds where we test for the violation of a Bell inequality and (ii) key rounds which are used to generate the raw key. In this protocol, the Bell inequality being tested is the CHSH inequality. We reiterate their CHSH condition since it is slightly modified for this protocol:

- If $x_A^i, x_B^i \in \{0, 1\}$ then $y_A^i \oplus y_B^i = x_A^i \wedge x_B^i$
- If $x_A^i, x_B^i = (2, 1)$ then $y_A^i \oplus y_B^i = 0$, i.e. $y_A^i = y_B^i$
- If $x_A^i, x_B^i = (2, 0)$ then all outputs are valid.

Let opt be the maximum probability with which two isolated devices produce outputs satisfying the CHSH condition under the laws of quantum mechanics. When the inputs are from the set $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$, the game is identical to the CHSH game and can be won with probability at most $\cos^2\left(\frac{\pi}{8}\right)$. When the input pair is from $\{(2, 0), (2, 1)\}$, the game can be won with certainty. Since all six input pairs are equally likely, we have

$\mathsf{opt} = \frac{2}{3}\cos^2(\frac{\pi}{8}) + \frac{1}{3}$. In order to achieve probability equal to opt, Alice and Bob share maximally entangled quantum states $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and measure in the eigenbases of the ± 1-observables:

$$\begin{array}{ll} \text{Alice:} & \text{Bob:} \\ x_A^i = 0 : Z & x_B^i = 0 : \dfrac{1}{\sqrt{2}}(Z+X) \\ x_A^i = 1 : X & x_B^i = 1 : \dfrac{1}{\sqrt{2}}(Z-X) \\ x_A^i = 2 : \dfrac{1}{\sqrt{2}}(Z-X) & \end{array} \tag{3.2}$$

They output $y^i = 0(1)$ corresponding to the $+1(-1)$ outcome for their respective observables. One can check using eq. (3.1) that this gives the correct winning probability.

For completeness we state the protocol in Algorithm 2. The proof of security uses multiple theoretical tools among which are the 'guessing game' [28, Lemma 11] and the 'quantum reconstruction paradigm' [28, Lemma 9] and are used to place a lower bound on the smooth min-entropy of Alice's and Bob's outputs conditioned on adversary Eve's side information. They show a secure key rate of $\approx 2.5\%$ and a raw key rate of $\approx 15\%$ as the noise $\eta \to 0$. In the other direction, they are able to achieve a positive key rate for a maximum noise level of $\eta \approx 1.2\%$. More recently, the Entropy Accumulation Theorem was proposed by [31], and improved in [53], which provides a more general framework for establishing upper and lower bounds on the conditional min-entropy. We reserve the discussion of the EAT for Section 3.3.

Algorithm 2: Vazirani-Vidick DIKQD protocol

Input:

η: Threshold for error rate

m: The number of rounds in the protocol

ϵ: The security parameter (used to bound statistical distance from uniform)

γ: Expected number of testing rounds

1 For $i \in [m]$:

2 Alice and Bob draw $x_A^i \in \{0,1,2\}$ and $x_B^i \in \{0,1\}$. They produce outputs y_A^i and $y_B^i \in \{0,1\}$ respectively according to eq. (3.2)

3 Alice chooses a subset $\mathbf{B} \subseteq [m]$ of size γm to use as testing rounds. Alice and Bob exchange their input and outputs in those rounds test the CHSH violation. If the CHSH condition is not satisfied in a fraction of rounds that is greater than the threshold η, they abort.

4 Alice and Bob publicly reveal the rest of their inputs. They check to make sure there are enough key rounds, i.e. rounds where their inputs were $(2,1)$. Let $\mathbf{C}$ be this set. If $\left||\mathbf{C}| - \frac{m}{6}\right| > 10\sqrt{m}$, they abort.

5 They perform standard information reconciliation and privacy amplification on their outputs in $\mathbf{C} - \mathbf{B}$.

3.2.5 Classical Post-Processing

Steps (3) and (4) in Algorithm 2 are referred to as the Parameter Estimation steps. This is when Alice and Bob check to see if their input and output pairs meet certain conditions, for example the violation of the CHSH inequality, and abort otherwise. After parameter estimation, Alice and Bob hold weakly correlated bit strings that they wish to turn into strongly correlated and strongly secure bit strings. The first step is to make sure that they hold identical bit strings except with negligible probability. This process is known as Information Reconciliation or Error Correction. Alice and Bob use a classical error correction protocol to correct any errors in their respective bit strings. An efficient way to do this is by using two-universal hashing. We present a formal definition.

Definition 6. *Let $\mathcal{F}$ be a family of functions from $\mathcal{X}$ to $\mathcal{Y}$, and μ be a probability distribution over $\mathcal{F}$. We call the pair $(\mathcal{F}, \mu)$ two-universal if for any $x \neq x' \in \mathcal{X}$ and f drawn*

according to μ,

$$\Pr_{f \in \mathcal{F}} [f(x) = f(x')] \leq \frac{1}{|\mathcal{Y}|}$$

We will often take μ to be the uniform distribution over $\mathcal{F}$. Such two-universal hash function families always exist for the case $\mathcal{X} = \{0,1\}^n$ and $\mathcal{Y} = \{0,1\}^m$ with $0 \leq m \leq n$, as shown in [54, 55]. Alice samples a function f according to the distribution μ and sends to Bob the pair $(f, f(k_A))$. Bob uses f to compute $f(k_B)$ and checks if $f(k_A) = f(k_B)$. If not, they abort the protocol. By choosing an appropriate size for the range of the hash function family i.e. $|\mathcal{Y}| = 2^{\lceil \log 1/\epsilon_{IR} \rceil}$, for an arbitrarily small ϵ_{IR}, we can guarantee that Alice and Bob's bit strings are equal except with probability at most ϵ_{IR}.

Once the information reconciliation step is complete, Alice and Bob use a *randomness extractor* to get their final keys secure against a quantum adversary, a process known as Privacy Amplification. At this point, Alice and Bob are assumed to have identical bit strings, so we only look at privacy amplification on Alice's side. Bob's analysis is identical. Since Eve is allowed to have quantum side information, Alice's and Eve's system can be described using a classical-quantum state $\rho_{AE} \in \mathfrak{H}_A \otimes \mathfrak{H}_E$ as follows:

$$\rho_{AE} = \sum_{y_A} p_{\mathcal{Y}}(y_A) |y_A\rangle\langle y_A| \otimes \rho_E^{y_A}$$

where $\{|y_A\rangle\}$ forms an orthonormal basis and $\rho_E^{y_A}$ is Eve's subnormalized state conditioned on obtaining the classical bit string y_A. The randomness extractor also takes as input a uniform seed $s \in \{0,1\}^d$. We formally define a quantum-proof randomness extractor.

Definition 7. $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is called a* (k, ϵ) *quantum-proof randomness*

extractor if for all classical quantum states $\rho_{YE} \in \mathfrak{H}_A \otimes \mathfrak{H}_E$ *with classical* Y *and min-entropy* $H_{\min}(Y \mid E) \geq k$*, we have*

$$\frac{1}{2} \left\| \rho_{\mathsf{Ext}(Y,S)SE} - \frac{\mathbb{1}}{2^m} \otimes \rho_S \otimes \rho_E \right\|_1 \leq \epsilon$$

Here the idea is that the uniform seed is independent from the output of the randomness extractor. An example of a quantum-proof randomness extractor is one that is based on two-universal hashing. Alice and Bob use the uniform seed to select a hash function f from the family of functions $\mathcal{F}$ described above and use the output of the function as their final key. For example Alice's final key would be $k_A = f(y_A)$, and similarly for Bob. We state [56, Corollary 5.6.1] which bounds the L_1-distance from uniform using the smooth min-entropy (see Definition 52).

Lemma 8. *Let* ρ_{XE} *be a density operator on* $\mathfrak{H}_X \otimes \mathfrak{H}_E$ *which is classical with respect to an orthonormal basis* $\{|x\rangle\}_{x \in \mathcal{X}}$ *of* $\mathfrak{H}_X$*. Let* $\mathcal{F}$ *be a two-universal family of hash functions from* $\mathcal{X}$ *to* $\mathcal{Y} = \{0,1\}^m$*, and let* $\epsilon \geq 0$*. Then*

$$\left\| \rho_{F(X)EF} - \rho_U \otimes \rho_{EF} \right\|_1 \leq 2\epsilon + 2^{-\frac{1}{2}(H_{\min}^{\epsilon}(\rho_{XE} \mid E) - m)}$$

where $\rho_U = \frac{1}{|\mathcal{Y}|} \sum_{y \in \mathcal{Y}} |y\rangle\langle y|$ *is the maximally mixed state on* $\mathfrak{H}_Y$*.*

3.3 Entropy Accumulation

In this section we review a theoretical tool known as the Entropy Accumulation Theorem (EAT) [31] that helps derive upper and lower bounds on the min-entropy of the outputs in

a device-independent quantum key distribution protocol. The main idea behind the EAT is to generalize the Quantum Asymptotic Equipartition Property to the case when our overall system does not consist of mutually independent and identically distributed (IID) sub-systems. In the IID case, the total entropy of the system is equal to the sum of its parts. This is not the case in DIQKD because the state of the system in round i may depend on the all previous rounds $0, \cdots, i-1$ (but not future rounds $i+1, \cdots$). The EAT says that given enough rounds, the total entropy generated in the protocol can be expressed as the number of rounds multiplied by the entropy generated in a single round along with a small correction term. Intuitively this says that as the number of rounds increases, we start to get closer and closer to the IID case (which asymptotically is exactly equal to the IID case).

We begin by introducing some notation and developing some of the key ideas necessary for the EAT. We will make use of EAT channels which are completely positive and trace-preserving maps that satisfy certain key properties. For a detailed explanation of EAT channels, see [32, Definition 2.5]. We denote EAT channels by $\mathcal{M}_i : R_{i-1} \mapsto R_i A_i B_i I_i C_i$, where A_i, B_i, C_i and I_i are classical systems with finite dimensions (typically denoting Alice and Bob's inputs and outputs and any other randomness they may share). R_i are arbitrary registers, typically denoting an adversary's side information and the noise from the environment.

Definition 9 (Tradeoff functions)**.** *[32, Definition 2.6] Let $\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_n$ be EAT channels, and let $\mathcal{C}$ denote the alphabet of $C_1, \cdots, C_n$. We call a function $f_{\min}$ a min-tradeoff*

function for $\{\mathcal{M}_i\}$ *if it is differentiable and convex, and satisfies*

$$f_{\min}(p) \leq \inf_{\substack{\sigma_{R_{i-1}R'}: \\ \mathcal{M}_i(\sigma)_{C_i}=p}} H(A_iB_i \mid I_iR')_{\mathcal{M}_i(\sigma)}$$

where the infimum is taken for all $i \in [n]$ *over states of* $\mathcal{M}_i$ *for which the marginal of* C_i *of the output state is the probability distribution* p*.*

We call a function $f_{\max}$ *a max-tradeoff function for* $\{\mathcal{M}_i\}$ *if it is differentiable and concave, and satisfies*

$$f_{\max}(p) \geq \sup_{\substack{\sigma_{R_{i-1}R'}: \\ \mathcal{M}_i(\sigma)_{C_i}=p}} H(A_iB_i \mid I_iR')_{\mathcal{M}_i(\sigma)}$$

where the supremum is taken for all $i \in [n]$ *over states of* $\mathcal{M}_i$ *for which the marginal of* C_i *of the output state is the probability distribution* p*.*

Definition 10. *Given a string* $\mathbf{s} = s_1, s_2, \cdots, s_n \in \Sigma_n$*, where* Σ *is a finite alphabet, we use* $\mathsf{freq}_{\mathbf{s}}(s')$ *to denote the probability distribution* $\frac{\sum_{i=1}^{n} \delta_{s_i,s'}}{n}$ *for* $s' \in \Sigma$*.*

We state the main theorem next.

Theorem 11 (EAT)**.** *[32, Theorem 2.7] Let* $\mathcal{M}_i : R_{i-1} \mapsto R_iA_iB_iI_iC_i$ *be EAT channels for* $i \in [n]$*, and let* $\rho_{ABICE} = (\mathrm{tr}_{R_n} \circ \mathcal{M}_n \circ \cdots \circ \mathcal{M}_1) \otimes \mathbb{1}_E \, \rho_{R_0E}$ *be the final state obtained after applying the EAT channels. Let* Ω *be an event defined over* $\mathcal{C}^n$ *and* p_Ω *the probability of* Ω *in* ρ*, and* $\rho_{\,|\,\Omega}$ *be the final state conditioned on* Ω*. We also let* $\epsilon_s \in (0, 1)$*.*

Given a min-tradeoff function $f_{\min}$ *for* $\{\mathcal{M}_i\}$ *such that* $f_{\min}(\mathsf{freq}_{\mathbf{c}}) \geq t_{\min}$ *for any*

$t_{\min} \in \mathbb{R}$ *and* $\mathbf{c} \in \mathcal{C}^n$ *for which* $\Pr[\mathbf{c}]\rho_{|\Omega} > 0$,

$$H_{\min}^{\epsilon_s}(AB \mid IE)_{\rho_{|\Omega}} > n \cdot t_{\min} - \sqrt{n}\left(2\left(\log(1+2d_{A_iB_i}) + \lceil \|\nabla f_{\min}\|_\infty \rceil\right)\sqrt{1-2\log(\epsilon_s \cdot p_\Omega)}\right)$$

where $d_{A_iB_i}$ *is the dimension of systems* A_iB_i.

Similarly, given a max-tradeoff function $f_{\max}$ *for* $\{\mathcal{M}_i\}$ *such that* $f_{\max}(\mathsf{freq}_{\mathbf{c}}) \leq t_{\max}$ *for any* $t_{\max} \in \mathbb{R}$ *and* $\mathbf{c} \in \mathcal{C}^n$ *for which* $\Pr[\mathbf{c}]\rho_{|\Omega} > 0$,

$$H_{\min}^{\epsilon_s}(AB \mid IE)_{\rho_{|\Omega}} < n \cdot t_{\max} + \sqrt{n}\left(2\left(\log(1+2d_{A_iB_i}) + \lceil \|\nabla f_{\max}\|_\infty \rceil\right)\sqrt{1-2\log(\epsilon_s \cdot p_\Omega)}\right).$$

The main challenge while applying the EAT to any key distribution protocol is to find an appropriate min-tradeoff function with the required properties. In Section 6.5, we derive a min-tradeoff function for our synchronous device independent quantum key distribution protocol.

3.4 Quantum Randomness Generation and Expansion

In Chapter 4 we will see some methods for generating random bits using quantum states. While such methods provide 'true' randomness, one might want to verify that the quantum device in use is behaving according to specification. This takes us back in to the realm of 'device-independence' and a lot of the techniques used for device-independent quantum key distribution carry over to the randomness generation case. The general idea is to use non-local games in order to self-test the device by carefully selecting testing rounds to perform Bell tests while using the other rounds as randomness generation rounds. This allows us to

use a small amount of randomness for selecting the testing rounds while generating exponentially more randomness through the generation rounds. This is often also referred to as randomness amplification. Vazirani and Vidick showed the first proof of security against quantum adversaries for a randomness amplification protocol in [57]. Security against quantum adversaries is needed in order to compose the randomness amplification protocol with other protocols. Subsequent work by Miller and Shi [30] showed cryptographic security and combined with the work of Coudron and Yuen [58], and Chung, Shi, and Wu [59] showed unbounded randomness expansion using only four devices. This can be achieved by cross-feeding the output of one pair of devices into another pair and repeating this process an unbounded number of times to achieve unbounded expansion.

3.5 Constructive Cryptography

In this section we provide some of the main ideas used in this framework, while referring the reader to the original papers [1, 60] for some of the more technical definitions and proofs.

Definition 12. *A* system *is defined as an abstract object with* interfaces *that let it interact with other systems. The interfaces are labeled using a set, typically* $\mathcal{I} = \{1, \cdots, n\}$.

At the highest level, there are three main types of systems: *resources*, *converters*, and *distinguishers*. We define each one of these systems and describe their properties.

3.5.1 Resources

Let's start with *resources*. A resource system can be thought of as providing access to a resource used in a protocol. An example of a concrete resource might be a secret key, or an

authenticated channel. Resource systems are denoted using capital letters, ex. R, S, and are equipped with an interface label set. In a protocol, for example, with parties Alice, Bob, and Eve, we can think of resources having interfaces for each of the parties, i.e. an interface label set $\mathcal{I} = \{A, B, E\}$.

3.5.2 Converters

Next, we talk about *converters*. Intuitively a converter system can be thought of as converting one resource into another resource. Converters are denoted using Greek letters eg. α, β. Converters have two interfaces called the *inside* and *outside* interface respectively. The inside interface attaches to a resource, while the outside interface becomes the new interface for the combined system. For example, an interface $i \in \mathcal{I}$ of a resource R can connect to the inside interface of a converter α, and the outside interface of α is now the interface i of the new resource $\alpha^i R$.

For every interface $i \in \mathcal{I}$ that a resource R has, we can think of attaching a converter α_i at that interface. As a result, we can have $\boldsymbol{\alpha} = \{\alpha_1, \cdots, \alpha_n\}$ connect to the n interfaces of the resource R. When necessary, we simplify the notation $\alpha_i^i R$ and write $\alpha_i R$ instead to denote that the converter α_i is attached to the ith interface of the resource R. $\boldsymbol{\alpha} R := \alpha_1 \alpha_2 \cdots \alpha_n R$.

Let ϕ be a set of resources, and Σ denote a set of converters. A converter $\alpha \in \Sigma$ induces a function $\phi \to \phi : R \to \alpha^i R$, for a resource $R \in \phi$. The set Σ has the following properties:

- It is equipped with a composition operation $\circ$ such that for two converters $\alpha, \beta \in \Sigma$

$$(\beta \circ \alpha)^i R = \beta^i(\alpha^i R),$$

  and is closed under composition i.e. $\Sigma \circ \Sigma = \Sigma$.

- It contains an identity element $\mathrm{id} \in \Sigma$ such that

$$\mathrm{id} \circ \alpha = \alpha \circ \mathrm{id} = \alpha$$

3.5.3 Distinguishers

The next type of system we describe is called a *distinguisher* system. Intuitively, distinguishers are systems that distinguish one resource from the other. A distinguisher for an n-interface resource, is an $(n+1)$-interface system, where the inside n interfaces connect to the interfaces of the resource, and the outside interface outputs a bit. A distinguisher D connected to a resource R is denoted by DR. Since distinguishers output a bit, DR is a binary random variable. Depending on the notion of security we are interested in, we can define a class of feasible distinguishers which correspond to computational security, and a class of all (unbounded) distinguishers, which correspond to information-theoretic security. Given two resources the closeness between them is measured using the best distinguishing advantage a distinguisher has given a specific class of distinguishers. For example, we say that two resources R and S are ϵ-close if and only if the distinguishing advantage for distinguishers $D \in \mathcal{D}$ is bounded by ϵ. This can be written as $R \approx_\epsilon S \iff \Delta^{\mathcal{D}}(R, S) \leq \epsilon$.

When $\epsilon = 0$ we simplify notation and write $R \approx_0 S$ as $R \approx S$. The notation $\Delta^{\mathcal{D}}(R, S)$ is defined as

$$\Delta^{\mathcal{D}}(R, S) := \sup_{D \in \mathcal{D}} \Delta^{D}(R, S),$$

where $\Delta^{D}(R, S)$ is the statistical distance between the distributions of binary variables DR and DS.

3.5.4 Lemmas and definitions

In this section we state some definitions and lemmas without proof for brevity. The proofs are fairly straightforward and can be found in [1, 60].

Definition 13. *A* resource specification *is defined as a subset $\mathcal{R} \subseteq \phi$ of the set of resources that satisfy a given specification.*

Often times in cryptographic protocols, we want to construct one resource from other resources, for example, we might construct a secure channel using a secret key and an authenticated channel. We need a way to formally write that a resource specification $\mathcal{S}$ was *constructed* from another resource specification $\mathcal{R}$ using a converter $\pi \in \Sigma$.

Definition 14. $\mathcal{R} \xrightarrow{\pi} \mathcal{S} \iff \pi\mathcal{R} \subseteq \mathcal{S}$

Lemma 15. *Such a construction is composable, which means that*

$$(\mathcal{R} \xrightarrow{\pi} \mathcal{S}) \wedge (\mathcal{S} \xrightarrow{\pi'} \mathcal{T}) \implies \mathcal{R} \xrightarrow{\pi' \circ \pi} \mathcal{T}.$$

Definition 16. *The ϵ-ball around a specification $\mathcal{R}$, denoted $\mathcal{R}^\epsilon$ is defined as:*

$$\mathcal{R}^\epsilon = \{R' \mid \exists R \in \mathcal{R} : R \approx_\epsilon R'\}.$$

To capture the case when an adversary may apply an arbitrary converter to a resource specification, we write $\mathcal{R}^* := \mathcal{R}\Sigma = \{R\alpha \mid R \in \mathcal{R}, \alpha \in \Sigma\}$, where $*$ denotes an arbitrary converter.

Definition 17. *A* pseudo-metric *d on a set ϕ is a function $d : \phi \times \phi \to \mathbb{R}_{\geq 0}$ such that for $R, S, T \in \phi$, we have*

1. $d(R, R) = 0$
2. $d(R, S) = d(S, R)$ *(symmetry)*
3. $d(R, S) \leq d(R, T) + d(T, S)$ *(triangle inequality)*

Note that if $d(R, S) = 0 \iff R = S$, then d is a metric *on the set ϕ.*

A metric d on ϕ is called non-expanding *if $d(\alpha R, \alpha S), d(R\beta, S\beta) \leq d(R, S)$ for all $\alpha, \beta \in \Sigma$.*

Lemma 18. *For a non-expanding metric d, we have that*

$$\mathcal{R} \xrightarrow{\pi} \mathcal{S} \implies \mathcal{R}^\epsilon \xrightarrow{\pi} \mathcal{S}^\epsilon.$$

Lemma 19. $\mathcal{R} \xrightarrow{\pi} \mathcal{S} \implies \mathcal{R}^* \xrightarrow{\pi} \mathcal{S}^*$.

3.5.5 Modeling cryptographic protocols

In order to model a cryptographic protocol with three parties, Alice, Bob, and Eve (the adversary), and show indifferentiability between two resources R (real), and S (ideal), we can use the following two conditions. Honest parties connect to the left interfaces and the adversary connects to the right interface. We can denote the converters for the honest parties (π_A, π_B) as a single converter $\pi_{AB} = (\pi_A, \pi_B)$:

- $\pi_{AB} R \perp_E \approx_\epsilon S \perp_E$. This is a condition of availability. This means that if no adversary is present, indicated by a $\perp_E$ at Eve's interface, the two parties must still be able to carry out the protocol.

- $\pi_{AB} R \approx S \sigma_E$. This condition models security. This means that anything that Eve can do in the real setting, she can also do in the ideal setting. Stated another way, anything Eve cannot do in the ideal setting, she isn't able to do in the real setting. The converter σ_E is referred to as a *simulator*. It simulates the adversary's behavior in the ideal case.

In the next section we present some random bit generator models and prove some of their properties in the constructive cryptography framework.

Chapter 4: Random Bit Generators

This chapter focuses on introducing ideal and real primitives for random bit generators. Random number generators play an important role in quantum key distribution protocols as they are used by the two parties involved to locally select their measurement bases. Combining results in this chapter with the results from Section 6.6 is a major step towards showing a full proof of DIQKD in the constructive cryptography framework.

4.1 *Ideal* and *Real* random bit generators

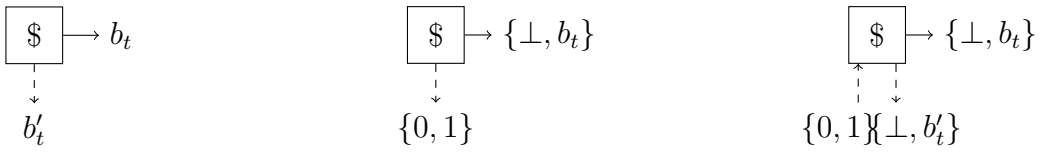
In this section we introduce three ideal bit generators. These bit generators have two interfaces, one for the honest party Alice, and one for an adversary Eve. We then describe real random bit generators where we make assumptions about how much information the adversary Eve has about the random bit delivered to Alice. We also show an ideal quantum random bit generator, and two concrete quantum random bit generators, and prove in the constructive framework that the real generators are close to the ideal generator, thus implying that we can construct the ideal generator from one of the real generators within a small error bound.

4.1.1 Preliminaries

The statistical distance between probability distributions over two random variables X and Y is given by $\Delta_{st} = \frac{1}{2}\sum_x |\Pr[X = x] - \Pr[Y = x]|$. In this section, we do not make any assumptions on the computational power of the adversary, and hence, the measure of distance and security we consider is based on statistical distance, which corresponds to information-theoretic security. We use $x \xleftarrow{\$} U$ to denote that x is being sampled from the distribution U. The probability distribution function for the *Beta* distribution is given as $f(x; \alpha, \beta) = Cx^{\alpha}(1-x)^{\beta}$, where C is a normalization constant. A beta-distributed random variable X with parameters α and β is denoted by $X \sim \texttt{Beta}(\alpha, \beta)$. Min-entropy is defined as $H_{\infty} = \min_x\{-\log \Pr[X = x]\}$.

We use the $\perp$ symbol to denote no output at an interface (either because there was no output produced or the adversary blocked it).

4.1.2 Definitions and descriptions of the models



(a) IDEAL1 (b) IDEAL2 (c) IDEAL3

Figure 4.1: Ideal random bit generators

Definition 20 (IDEAL1)**.** *At time t,* IDEAL1 *outputs a uniformly random bit b_t to the honest party Alice, and another uniformly random bit b'_t to dishonest party Eve. We have that* $\forall(t, t'), \Pr[b_t = x \text{ and } b'_t = x'] = \Pr[b_t = x]\Pr[b'_t = x']$. *This is illustrated in Figure(4.1a).*

Algorithm 3: Simulator σ

```
if input == 0 then
    output ⊥
else
    output b'_t ←$ U where U is the distribution that the random bit generator uses
```

Definition 21 (IDEAL2)**.** *At time t,* IDEAL2 *outputs either $\perp$ or b_t at Alice's interface. At Eve's interface, she sees* 0 *if and only if Alice sees $\perp$, and* 1 *if and only if Alice sees b_t at Alice's interface. This models the case where Eve learns whether Alice receives a bit. This model is illustrated in Figure(4.1b).*

Definition 22 (IDEAL3)**.** *Eve inputs* 0 *or* 1 *at her interface, and based on Eve's input at time t, Alice and Eve, at their respective interfaces, see $\perp$, $\perp$ if Eve's input was 0, and b_t, b'_t if Eve's input was 1. This models the case where Eve controls whether Alice receives a bit. This model is illustrated in Figure(4.1c).*

Lemma 23. *There exists a simulator σ such that* IDEAL1 $\approx$ IDEAL2 σ

Proof. We show that for any distinguisher D,

$$d(\mathsf{IDEAL1}, \mathsf{IDEAL2}\,\sigma) = |\Pr[D^{\mathsf{IDEAL1}} = 1] - \Pr[D^{\mathsf{IDEAL2}\,\sigma} = 1]| = 0$$

We define the simulator σ in Algorithm 3.

By construction of the simulator, b'_t is independently drawn from the distribution U whenever Eve's input to the simulator is 1 and thus, the mutual information between the distributions for b_t and b'_t is zero. □

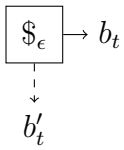


Figure 4.2: Real random generator 1 (REAL1)

Definition 24 (REAL1)**.** *At time t,* REAL1 *is a resource parameterized by ϵ that outputs a uniformly random bit b_t to the honest party Alice and b'_t to the dishonest party Eve. The dishonest party Eve has an advantage ϵ that we quantify as follows. Let $b_t(x) = \Pr[b_t = x]$ be Alice's probability distribution, and $b'_t(x') = \Pr[b'_t = x']$ be Eve's probability distribution. The joint distribution $\forall x, x' \in [n], (b_t(x), b'_t(x'))$ is given as*

$$(b_t(x), b'_t(x')) = \Pr[b_t = x \text{ and } b'_t = x'] = \Pr[b'_t = x' | b_t = x] \Pr[b_t = x].$$

We have that for a fixed x, x',

$$(b_t(x), b'_t(x')) = \begin{cases} \frac{1}{2}\left(\frac{1}{2} + \epsilon\right) & \text{if } x = x' \\ \frac{1}{2}\left(\frac{1}{2} - \epsilon\right) & \text{if } x \neq x' \end{cases}$$

Lemma 25. IDEAL1 $\approx_\epsilon$ REAL1*.*

Proof. We show that for the class of distinguishers $\mathcal{D}$ based on statistical distance,

$$\forall D \in \mathcal{D}, \Delta^D(\mathsf{IDEAL1}, \mathsf{REAL1}) = |\Pr[D^{\mathsf{IDEAL1}} = 1] - \Pr[D^{\mathsf{REAL1}} = 1]| \leq \epsilon.$$

We prove this by showing the statistical distance between the distributions for IDEAL1 and REAL1 is bounded by ϵ:

$$\begin{aligned}
\delta(\mathsf{IDEAL1}, \mathsf{REAL1}) &= \frac{1}{2} \sum_{x,x'} |[(b_t(x), b'_t(x'))]_{\mathsf{IDEAL1}} - [(b_t(x), b'_t(x'))]_{\mathsf{REAL1}}| \\
&= \frac{1}{2} \Big(\sum_{x=x'} |[(b_t(x), b'_t(x))]_{\mathsf{IDEAL1}} - [(b_t(x), b'_t(x))]_{\mathsf{REAL1}}| \\
&\qquad + \sum_{x \neq x'} |[(b_t(x), b'_t(x'))]_{\mathsf{IDEAL1}} - [(b_t(x), b'_t(x'))]_{\mathsf{REAL1}}| \Big) \\
&= \frac{1}{2} \left[2 \left(\left| \frac{1}{4} - \frac{1}{2} \left(\frac{1}{2} + \epsilon \right) \right| \right) + 2 \left(\left| \frac{1}{4} - \frac{1}{2} \left(\frac{1}{2} - \epsilon \right) \right| \right) \right] \\
&= \frac{1}{2} \left[\epsilon + \epsilon \right] \\
&= \epsilon
\end{aligned}$$

□

Next, we look at a class of random number generators based on process tomography of one-qubit channels, see Example (3). This involves making repeated measurements in the X and Z basis to compute measurement angles that give maximum min-entropy for the output random bit. In the ideal case, we have precise values for $p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|}$ and $p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|}$ and the min-entropy is 1. However in the real case, there is statistical error in the $p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|}$ and $p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|}$ values and thus the min-entropy value deviates very slightly from 1. We analyze these statistical errors in order to characterize the distinguishing advantage of a distinguisher connected to the real and ideal QRNG systems.

Let p_X and p_Z be the probability of getting outcome 0 on an X and Z measurement respectively. Let s_X, n_X and s_Z, n_Z be the number of successes and number of measurements

in the X and Z bases respectively. The p_X and p_Z can be modeled by Beta distributions given by $p_X \sim \texttt{Beta}(s_X + 1, n_X - s_X + 1)$ and $p_Z \sim \texttt{Beta}(s_Z + 1, n_Z - s_Z + 1)$ [61]. We use maximum likelihood estimators $\hat{p}_X = \frac{s_X}{n_X}$ and $\hat{p}_Z = \frac{s_Z}{n_Z}$ to estimate $\hat{\theta} = \tan^{-1}\left(-\frac{\hat{p}_Z - \frac{1}{2}}{\hat{p}_X - \frac{1}{2}}\right)$. Thus

$$\hat{\theta} = \tan^{-1}\left(-\frac{n_X}{n_Z}\left(\frac{2s_Z - n_Z}{2s_X - n_X}\right)\right)$$

Therefore $\Pr\{\theta \in [\hat{\theta} - d\theta, \hat{\theta} + d\theta]\}$ is given by

$$\begin{aligned}\Pr\{\theta \in [\hat{\theta} - d\theta, \hat{\theta} + d\theta]\} &= \int_0^1 \int_0^1 \left\{\tan^{-1}\left(-\frac{p_Z - \frac{1}{2}}{p_X - \frac{1}{2}}\right) \in [\hat{\theta} - d\theta, \hat{\theta} + d\theta]\right\} d\beta(p_X) d\beta(p_Z) \\ &= \int_0^1 \int_0^1 \left\{\left(-\frac{p_Z - \frac{1}{2}}{p_X - \frac{1}{2}}\right) \in [\tan(\hat{\theta} - d\theta), \tan(\hat{\theta} + d\theta)]\right\} d\beta(p_X) d\beta(p_Z)\end{aligned}$$

Let $l_1 = \frac{1}{2} + \left(p_X - \frac{1}{2}\right)\tan(\hat{\theta} - d\theta)$ and $l_2 = \frac{1}{2} + \left(p_X - \frac{1}{2}\right)\tan(\hat{\theta} + d\theta)$. Thus we have,

$$\begin{aligned}\Pr\{\theta \in [\hat{\theta} - d\theta, \hat{\theta} + d\theta]\} &\gtrsim \int_{\frac{1}{2}}^1 \int_0^1 p_Z \in [l_1, l_2]\, d\beta(p_Z) d\beta(p_X) \\ &= \int_{\frac{1}{2}}^1 \left\{\int_{l_1}^{l_2} d\beta(p_Z)\right\} d\beta(p_X) \\ &= \int_{\frac{1}{2}}^1 I_{\theta_+(p_X)}(s_Z + 1, n_Z - s_Z + 1) - I_{\theta_-(p_X)}(s_Z + 1, n_Z - s_Z + 1) d\beta(p_X) \qquad (4.1)\end{aligned}$$

We describe the two quantum random bit generators as follows:

Definition 26 ($\mathsf{IDEAL}_{\mathsf{qrng}-1}$)**.** *Prepare a qubit in the $|0\rangle\langle 0|$ state, apply the Hadamard operator to it, and measure in the computational basis. We have* $\Pr[0] = \Pr[1] = \frac{1}{2}$.

Definition 27 ($\mathsf{REAL}_{\mathsf{qrng}-1}$)**.** *Prepare n qubits in the $|0\rangle\langle 0|$ state with state preparation*

error p (modeled as a bit flip channel with probability p), apply a Hadamard operator which may have slight overrotation error dt. We estimate $p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|}$ and $p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|}$ by measuring $n-1$ of the prepared qubits. We then use the reserved qubit to generate an output bit with optimal min-entropy as given by Example (3). Statistical error in measurements introduces an error ϵ in the computation of the optimal measurement angle θ, characterized by Equation (4.1).

Lemma 28. $\mathsf{IDEAL}_{\mathsf{qrng}-1} \approx_\epsilon \mathsf{REAL}_{\mathsf{qrng}-1}$, *where ϵ is the statistical error in estimating the optimal angle in* $\mathsf{REAL}_{\mathsf{qrng}-1}$.

Proof. Let $\Pr[X=x]_{\mathsf{ideal}}$ and $\Pr[X=x]_{\mathsf{real}}$ be the probability distributions of the ideal and real generators respectively. We use Equation (4.1) to bound the probability that the computed value $\hat{\theta}$ is within ϵ of the actual value θ. Let p_ϵ be said bound. We can make ϵ arbitrarily small, eg. ≤ 0.05, and p_ϵ arbitrarily close to 1, eg. ≥ 0.99. Computing the statistical distance between the two probability distributions gives us:

$$
\begin{aligned}
\delta(\mathsf{IDEAL}_{\mathsf{qrng}-1},\mathsf{REAL}_{\mathsf{qrng}-1}) &= \frac{1}{2}\sum_x |\Pr[X=x]_{\mathsf{ideal}} - \Pr[X=x]_{\mathsf{real}}| \\
&= \frac{1}{2}\Big(|\Pr[X=0]_{\mathsf{ideal}} - \Pr[X=0]_{\mathsf{real}}| + |\Pr[X=1]_{\mathsf{ideal}} - \Pr[X=1]_{\mathsf{real}}|\Big) \\
&= \frac{1}{2}\Big(|\Pr[X=0]_{\mathsf{ideal}} - \Pr[X=0]_{\mathsf{real}}| \\
&\qquad + |1-\Pr[X=0]_{\mathsf{ideal}} - 1 + \Pr[X=0]_{\mathsf{real}}|\Big) \\
&= |\Pr[X=0]_{\mathsf{ideal}} - \Pr[X=0]_{\mathsf{real}}| \\
&= \left|\frac{1}{2} - \left(\sin(\theta\pm\epsilon)\left(p^{\Pi^0_{\sigma_X}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \cos(\theta\pm\epsilon)\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \frac{1}{2}\right)\right| \\
&= \Big|\sin(\theta\pm\epsilon)\Big(-\frac{1}{2}(2p-1)\cos^2(dt)\Big) \\
&\qquad + \cos(\theta\pm\epsilon)\Big(-\frac{1}{2}(2p-1)\sin^2(dt)\Big)\Big|
\end{aligned}
$$

For $p, dt \leq 0.05$, we have

$$\begin{aligned}\delta(\mathsf{IDEAL}_{\mathsf{qrng}-1}, \mathsf{REAL}_{\mathsf{qrng}-1}) &\leq \left|\pm\frac{1}{2}\sin(\epsilon)\right| \\ &\leq \frac{\epsilon}{2} \quad \text{for small } \epsilon.\end{aligned}$$

Therefore, $\delta(\mathsf{IDEAL}_{\mathsf{qrng}-1}, \mathsf{REAL}_{\mathsf{qrng}-1}) \leq \frac{\epsilon}{2}$ except with probability $1 - p_\epsilon$. □

Consider a variation of $\mathsf{REAL}_{\mathsf{qrng}-1}$, where instead of doing process tomography, we always make a measurement in the computational basis, i.e. $\theta = \phi = 0$.

Definition 29 ($\mathsf{REAL}_{\mathsf{qrng}-2}$)**.** *Prepare a qubit in the $|0\rangle\langle 0|$ state with state preparation error (bit flip error with probability p). Apply a Hadamard with overrotation dt, and measure in the Z basis.*

Lemma 30. $\mathsf{IDEAL}_{\mathsf{qrng}-1} \approx_{\epsilon^2} \mathsf{REAL}_{\mathsf{qrng}-2}$, *where $p, dt \leq \epsilon$*

Proof. Computing the statistical distance between the probability distributions gives us:

$$\begin{aligned}\delta(\mathsf{IDEAL}_{\mathsf{qrng}-1}, \mathsf{REAL}_{\mathsf{qrng}-2}) &= \frac{1}{2}\sum_x |\Pr[X = x]_{\mathsf{ideal}} - \Pr[X = x]_{\mathsf{real}}| \\ &= |\Pr[X = 0]_{\mathsf{ideal}} - \Pr[X = 0]_{\mathsf{real}}| \\ &= \left|\frac{1}{2} - \left(\left(p^{\Pi^0_{\sigma_Z}}_{|0\rangle\langle 0|} - \frac{1}{2}\right) + \frac{1}{2}\right)\right| \\ &= \left|-\frac{1}{2}(2p-1)\sin^2(dt)\right| \\ &\leq \left|\frac{1}{2}\sin^2(\epsilon)\right| \leq \frac{\epsilon^2}{2} \quad \text{for small } \epsilon\end{aligned}$$

□

Chapter 5: Synchronous Correlations

In this chapter we introduce the notion of synchronous correlations and discuss some results from [36] that can be found in greater detail in the paper. We show the existence of synchronous Bell inequalities in the case of three-input/two-output nonlocal synchronous games and show that they can be used to self-test an EPR pair. These Bell inequalities along with the self-testing property form the basis of the device-independent quantum key distribution protocol described in Chapter 6.

We study synchronous correlations in the context of nonlocal games, so we assume that there are two parties Alice and Bob that receive inputs $x_A, x_B \in \mathcal{X}$ and produce outputs $y_A, y_B \in \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are finite sets. We denote by $p(y_A, y_B \mid x_A, x_B)$ the conditional probability distribution or *correlation* of Alice and Bob's outputs given their inputs.

Definition 31. *A correlation is* synchronous *if* $p(y_A, y_B \mid x, x) = 0$ *whenever* $x \in \mathcal{X}$ *and* $y_A \neq y_B \in \mathcal{Y}$.

Definition 32. *A correlation is* symmetric *if* $p(y_A, y_B \mid x_A, x_B) = p(y_B, y_A \mid x_B, x_A)$

These correlations can be analyzed as polytopes of (column) stochastic matrices [62, 63]. For $|\mathcal{X}| = n$ and $|\mathcal{Y}| = m$, the correlation $p(y_A, y_B \mid x_A, x_B)$ produces an $m^2 \times n^2$ column stochastic matrix. The polytope that arises out of general correlations is $(m^2-1)n^2$

dimensional and has m^2n^2 vertices which correspond to deterministic strategies. Adding in the conditions for synchronicity slices this polytope and gives us an $(m^2-1)n^2 - n(m^2 - m)$-dimensional polytope with $m^2n^2 - m^2n + mn$ vertices. These vertices correspond to synchronous deterministic strategies.

We will focus on *nonsignalling correlations*, which means that the marginal distribution of Alice's output does not depend on Bob's input, and vice-versa for Bob. More formally,

Definition 33. *A correlation p is nonsignalling if*

1. *for all y_A, x_A, x_B, x'_B:*

$$\sum_{y_B} p(y_A, y_B \mid x_A, x_B) = \sum_{y_B} p(y_A, y_B \mid x_A, x'_B)$$

2. *for all y_A, x_A, x'_A, x_B:*

$$\sum_{y_A} p(y_A, y_B \mid x_A, x_B) = \sum_{y_A} p(y_A, y_B \mid x'_A, x_B)$$

Definition 34. *A local hidden variables strategy, or classical correlation is one that takes the form*

$$p(y_A, y_B \mid x_A, x_B) = \sum_{\omega \in \Omega} \mu(\omega)\, p_A(y_A \mid x_A, \omega)\, p_B(y_B \mid x_B, \omega)$$

for a finite set Ω and probability distribution μ over Ω.

Above p_A, p_B are local probability distributions Alice and Bob use to produce their respective outputs, and the (Ω, μ) can be viewed as their shared randomness. Clearly every

classical correlation is nonsignalling. We state some results about synchronous classical correlations.

Theorem 35. *The set of synchronous classical correlations with input $\mathcal{X}$ and output $\mathcal{Y}$ is bijective to the collection of probability distributions on the set of functions $\mathcal{X} \to \mathcal{Y}$. Given such a probability distribution, the associated strategy is: Alice and Bob sample a function $f : \mathcal{X} \to \mathcal{Y}$ according the specified distribution, and upon receiving x_A, x_B they output $y_A = f(x_A)$ and $y_B = f(x_B)$.*

Proof. Let p be a synchronous correlation. Thus we have that, for each $x_A = x_B = x \in \mathcal{X}$ and $y_A \neq y_B \in \mathcal{Y}$,

$$0 = \sum_{\omega \in \Omega} \mu(\omega) p_A(y_A \mid x, \omega) p_B(y_B \mid x, \omega) \tag{5.1}$$

Since we assume that $\mu(\omega) > 0$ for all $\omega \in \Omega$, we must have that for each $x \in \mathcal{X}$ and $\omega \in \Omega$: $p_A(y_A \mid x, \omega) p_B(y_B \mid x, \omega) = 0$ if $y_A \neq y_B$. Let us fix an $x \in \mathcal{X}$ and $\omega \in \Omega$. Since $\sum_{y_A} p_A(y_A \mid x, \omega) = 1$, $\exists y \in \mathcal{Y}$ such that $p_A(y \mid x, \omega) > 0$, and so from eq. (5.1) above, $p_B(y_B \mid x, \omega) = 0$ for $y_B \neq y$. This lets us conclude that

$$1 = \sum_{y_B} p_B(y_B \mid x, \omega) = p_B(y_0 \mid x, \omega)$$

Since $p_B(y \mid x, \omega) = 1$, the above argument carries through exactly with the roles of Alice and Bob reversed. This gives us $p_A(y|x, \omega) = 1$. Therefore for each $\omega \in \Omega$, we get a function $f_\omega : \mathcal{X} \to \mathcal{Y}$ given by $f_\omega(x) = y$ where y is the value with $p_A(y \mid x, \omega) = p_B(y \mid x, \omega) = 1$. Thus there is a natural mapping from Ω to the set of functions $\mathcal{X} \to \mathcal{Y}$. □

Corollary 36. *The extreme points of the synchronous hidden variables strategies from $\mathcal{X}$*

to $\mathcal{Y}$ can be canonically identified with the set of functions $\mathcal{X} \to \mathcal{Y}$.

Corollary 37. *Every synchronous classical strategy is symmetric.*

We focus now on quantum correlations and state some results that will be used in the device independent QKD protocol in Chapter 6. Specifically, the protocol is symmetric between Alice and Bob which means that their roles are interchangeable and need not be decided prior to executing the protocol.

Definition 38. *A correlation p is quantum if it takes the form*

$$p(y_A, y_B \,|\, x_A, x_B) = \mathrm{tr}\left(\rho \left(E_{y_A}^{x_A} \otimes F_{y_B}^{x_B}\right)\right)$$

where $\{\{E_y^x\}_{y\in\mathcal{Y}}\}_{x\in\mathcal{X}}$ and $\{\{F_y^x\}_{y\in\mathcal{Y}}\}_{x\in\mathcal{X}}$ are POVMs on $\mathfrak{H}_A$ and $\mathfrak{H}_B$ respectively, and ρ is a density operator on $\mathfrak{H}_A \otimes \mathfrak{H}_B$.

Quantum correlations are nonsignalling due to the fact that $\{E_y^x\}_{y\in\mathcal{Y}}$ and $\{F_y^x\}_{y\in\mathcal{Y}}$ are POVMs. We will take $\mathfrak{H}_A$ and $\mathfrak{H}_B$ to be finite-dimensional. For synchronous quantum correlations, the POVMs E_y^x and F_y^x are actually projection-valued. The proof for this can be found in [64, Proposition 1], but we state the result here.

Lemma 39. *Let $p(y_A, y_B \,|\, x_A, x_B) = \mathrm{tr}(\rho(E_{y_A}^{x_A} \otimes F_{y_B}^{x_B}))$ be a synchronous quantum correlation. Then the POVMs $\{E_y^x\}_{y\in Y}$ and $\{F_y^x\}_{y\in Y}$, for $x \in X$, are projection-valued measures. Moreover each E_y^x commutes with $\mathrm{tr}_B(\rho)$ and each F_y^x commutes with $\mathrm{tr}_A(\rho)$.*

Theorem 40. *Every synchronous quantum correlation can be expressed as the convex combination of synchronous quantum correlations with maximally entangled pure states. In*

particular, if a synchronous quantum correlation $\mathrm{tr}(\rho(E_{y_A}^{x_A} \otimes F_{y_B}^{x_B}))$ *is extremal then we may take* $\rho = |\psi\rangle\langle\psi|$ *with* $|\psi\rangle$ *maximally entangled.*

Theorem 41. *Let* $\mathcal{X}, \mathcal{Y}$ *be finite sets,* $\mathfrak{H}$ *a* d*-dimensional Hilbert space, and for each* $x \in \mathcal{X}$ *a projection-valued measure* $\{E_y^x\}_{y\in\mathcal{Y}}$ *on* $\mathfrak{H}$*. Then*

$$p(y_A, y_B \mid x_A, x_B) = \frac{1}{d}\mathrm{tr}(E_{y_A}^{x_A} E_{y_B}^{x_B})$$

defines a synchronous quantum correlation. Moreover every synchronous quantum correlation with maximally entangled pure state has this form.

Corollary 42. *Every synchronous quantum correlation is symmetric.*

5.1 Synchronous Bell Inequalities

Bell inequalities are used a test of quantumness: violation of a Bell inequality certifies quantum behavior. We saw some Bell inequalities in Section 3.1 in the context of the CHSH game. They help characterize hidden variables correlations or classical correlations among general nonsignalling correlations. We focus on synchronous Bell inequalities in this section. We saw earlier that the set of general synchronous correlations forms a polytope, and the polytope of synchronous classical correlations is given by the convex hull of the set of functions from $\mathcal{X} \to \mathcal{Y}$. We will see that the facets of the classical synchronous polytope that are not facets of the nonsignalling synchronous polytope give us our synchronous Bell inequalities. We start with smallest case of two inputs and two outputs i.e. $|\mathcal{X}| = 2$ and $|\mathcal{Y}| = 2$. This case does not yield any synchronous Bell inequalities since every quantum

synchronous correlation is also classical. We state the results (without proof) more formally.

Proposition 43. *A synchronous nonsignaling correlation with $|\mathcal{X}| = 2$ is classical if and only if it is symmetric.*

Theorem 44. *Every synchronous quantum correlation with $|\mathcal{X}| = 2$ to any finite set is classical.*

As a consequence of the corollary above, there are no synchronous analogues of the CHSH inequality. Next we turn our attention to the three input, two output case i.e. $|\mathcal{Y}| = 2$ and $|\mathcal{X}| = 3$. First we introduce some notation that is typically used in literature as a coordinate system for the polytopes we are interested in.

$$\begin{aligned} a_{x_A} &= \sum_{y_A,y_B} (-1)^{(1,0).(y_A,y_B)} p(y_A, y_B|x_A, x_B) \\ b_{x_B} &= \sum_{y_A,y_B} (-1)^{(0,1).(y_A,y_B)} p(y_A, y_B|x_A, x_B) \\ c_{x_A,x_B} &= \sum_{y_A,y_B} (-1)^{(1,1).(y_A,y_B)} p(y_A, y_B|x_A, x_B). \end{aligned} \tag{5.2}$$

Note that the nonsignalling criteria show that a and b do not depend on x_B or x_A respectively. Moreover for any probability distribution

$$1 = \sum_{y_A,y_B} (-1)^{(0,0).(y_A,y_B)} p(y_A, y_B|x_A, x_B).$$

The inverse relationship is straightforward to compute:

$$\begin{aligned} p(0,0|x_A,x_B) &= \tfrac{1}{4}\left(1 + a_{x_A} + b_{x_B} + c_{x_A,x_B}\right), \\ p(0,1|x_A,x_B) &= \tfrac{1}{4}\left(1 + a_{x_A} - b_{x_B} - c_{x_A,x_B}\right), \\ p(1,0|x_A,x_B) &= \tfrac{1}{4}\left(1 - a_{x_A} + b_{x_B} - c_{x_A,x_B}\right), \\ p(1,1|x_A,x_B) &= \tfrac{1}{4}\left(1 - a_{x_A} - b_{x_B} + c_{x_A,x_B}\right). \end{aligned} \tag{5.3}$$

These must form a probability density function, and hence are nonnegative and sum to one. That they sum to one is clear from this expression; nonnegativity gives the inequalities

$$\begin{aligned} 1 + a_{x_A} + b_{x_B} + c_{x_A,x_B} &\geq 0, \\ 1 + a_{x_A} - b_{x_B} - c_{x_A,x_B} &\geq 0, \\ 1 - a_{x_A} + b_{x_B} - c_{x_A,x_B} &\geq 0, \\ 1 - a_{x_A} - b_{x_B} + c_{x_A,x_B} &\geq 0, \end{aligned} \tag{5.4}$$

which form the basis for our description of the polytope of nonsignaling correlations below.

Lemma 45. *A correlation p is symmetric and nonsignalling if and only if (i) $c_{x_A,x_B} = c_{x_B,x_A}$ and (ii) $a_x = b_x$.*

Proof. Suppose p is symmetric. Then

$$\begin{aligned} a_{x_A} &= \sum_{y_A,y_B} (-1)^{(1,0).(y_A,y_B)} p(y_A,y_B|x_A,x_B) \\ &= \sum_{y_A,y_B} (-1)^{(1,0).(y_A,y_B)} p(y_B,y_A|x_B,x_A) \\ &= \sum_{y_A,y_B} (-1)^{(1,0).(y_B,y_A)} p(y_A,y_B|x_B,x_A) \end{aligned}$$

$$= \sum_{y_A, y_B} (-1)^{(0,1).(y_A,y_B)} p(y_A, y_B | x_B, x_A) = b_{x_A}.$$

Similarly,

$$\begin{aligned} c_{x_A,x_B} &= \sum_{y_A, y_B} (-1)^{(1,1).(y_A,y_B)} p(y_A, y_B | x_A, x_B) \\ &= \sum_{y_A, y_B} (-1)^{(1,1).(y_A,y_B)} p(y_B, y_A | x_B, x_A) \\ &= \sum_{y_A, y_B} (-1)^{(1,1).(y_B,y_A)} p(y_A, y_B | x_B, x_A) \\ &= \sum_{y_A, y_B} (-1)^{(1,1).(y_A,y_B)} p(y_A, y_B | x_B, x_A) = c_{x_B,x_A}. \end{aligned}$$

The converse is clear from eq. (5.3). □

Lemma 46. *A correlation p is synchronous and nonsignalling if and only if for all $x \in X$ we have (i) $c_{x,x} = 1$ and (ii) $a_x = b_x$.*

Proof. Assume p is synchronous and nonsignalling. Thus $p(y_A, y_B|x, x) = 0$ if $y_A \neq y_B$. We always have

$$\begin{aligned} p(0,0|x,x) + p(0,1|x,x) + p(1,0|x,x) + p(1,1|x,x) \\ = p(0,0|x,x) + p(1,1|x,x) = 1. \end{aligned} \tag{5.5}$$

Thus from eq. (5.5) we have

$$c_{x,x} = \sum_{y_A, y_B} (-1)^{(1,1).(y_A,y_B)} p(y_A, y_B | x, x)$$

$$= p(0,0|x,x) - p(0,1|x,x) - p(1,0|x,x) + p(1,1|x,x)$$

$$= p(0,0|x,x) + p(1,1|x,x) = 1.$$

Similarly,

$$a_x = \sum_{y_A,y_B} (-1)^{(1,0)\cdot(y_A,y_B)} p(y_A,y_B|x,x)$$

$$= p(0,0|x,x) + p(0,1|x,x) - p(1,0|x,x) - p(1,1|x,x)$$

$$= p(0,0|x,x) - p(1,1|x,x)$$

$$= p(0,0|x,x) - p(0,1|x,x) + p(1,0|x,x) - p(1,1|x,x)$$

$$= \sum_{y_A,y_B} (-1)^{(0,1)\cdot(y_A,y_B)} p(y_A,y_B|x,x) = b_x$$

The converse is clear by setting $c_{x,x} = 1$ and $a_x = b_x$ in eq. (5.3). □

Specializing to the case of $|\mathcal{X}| = 3$, say $\mathcal{X} = \{0,1,2\}$, we can use Lemma 46 to reduce the set of inequalities in eq. (5.4). When $j = k$ we only obtain trivial inequalities $0 \geq 0$ and $1 \geq a_j \geq -1$. For $j \neq k$ we get the 24 inequalities

$$1 + a_j + a_k + c_{j,k} \geq 0, \qquad 1 - a_j - a_k + c_{j,k} \geq 0,$$

$$1 - a_j + a_k - c_{j,k} \geq 0, \qquad 1 + a_j - a_k - c_{j,k} \geq 0.$$

The synchronous hidden variables polytope is contained within the symmetric synchronous nonsignalling polytope, but has additional facets that describe it, which we denote

by Js, and are listed as follows:

$$\begin{aligned} J_0 &= \tfrac{1}{4}(1 + c_{0,1} - c_{0,2} - c_{1,2}) \;\geq\; 0 \\ J_1 &= \tfrac{1}{4}(1 - c_{0,1} + c_{0,2} - c_{1,2}) \;\geq\; 0 \\ J_2 &= \tfrac{1}{4}(1 - c_{0,1} - c_{0,2} + c_{1,2}) \;\geq\; 0 \\ J_3 &= \tfrac{1}{4}(1 + c_{0,1} + c_{0,2} + c_{1,2}) \;\geq\; 0. \end{aligned} \tag{5.6}$$

These four inequalities characterize when a synchronous nonsignaling strategy is classical.

5.2 Rigidity of quantum synchronous correlations

In this section we state some results about the four synchronous Bell inequalities. The first result shows a bound on the maximal violation of the Bell inequalities by a synchronous nonsignalling correlation while the second result shows a similar bound for synchronous quantum correlations. Without loss of generality we take $\mathcal{X} = \{0, 1, 2\}$ and $\mathcal{Y} = \{0, 1\}$.

Proposition 47. *Every synchronous nonsignaling correlation satisfies all of the inequalities* $J_0, J_1, J_2, J_3 \geq -\frac{1}{2}$. *However no individual correlation can violate more than one of the inequalities* $J_0, J_1, J_2, J_3 \geq 0$.

For synchronous quantum correlations, we give an analogue of Tsirl'son's bound [42].

Theorem 48. *Every synchronous quantum correlation satisfies all the inequalities* $J_i \geq -\frac{1}{8}$, *for all* $i \in \{0, 1, 2, 3\}$. *However no individual correlation can violate more than one of the inequalities* $J_i \geq 0$.

From the theorem above, we see that the violation of any of the J's can have magnitude at most $\frac{1}{8}$. The next theorem shows that there exists a unique quantum correlation that achieves this bound.

Theorem 49. *For each of the bounds of Theorem 48, there exists a unique synchronous quantum correlation from* $\{0,1,2\}$ *to* $\{0,1\}$ *that achieves it.*

Since we do not include a proof of the theorem, we state Alice and Bob's strategy for completeness. Let us begin by defining ± 1 observables $M_x = E_0^x - E_1^x$. From Theorem 41 and eq. (5.2), we have $a_x = \frac{1}{d}\mathrm{tr}\left(M_x\right)$ and $c_{x_A,x_B} = \frac{1}{d}\mathrm{tr}\left(M_{x_A}M_{x_B}\right)$. Then, Alice and Bob's strategy is given by:

$$[M_0] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, [M_1] \ = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}, \text{ and } [M_2] = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}. \tag{5.7}$$

The correlation matrix for this strategy is given by:

$$[p(y_A, y_B|x_A, x_B)] = \frac{1}{8}\left(\begin{array}{ccc|ccc|ccc} 4 & 1 & 1 & 1 & 4 & 1 & 1 & 1 & 4 \\ 0 & 3 & 3 & 3 & 0 & 3 & 3 & 3 & 0 \\ 0 & 3 & 3 & 3 & 0 & 3 & 3 & 3 & 0 \\ 4 & 1 & 1 & 1 & 4 & 1 & 1 & 1 & 4 \end{array}\right).$$

One can verify that $J_3 = -\frac{1}{8}$ for this correlation.

This consequence poses an issue for the synchronous device independent protocol we describe in the next chapter. The issue is that an adversary may use an asynchronous classical correlation and still produce a maximal violation of our Bell inequality. This

i	$f_i(0)$	$f_i(1)$	$f_i(2)$
0	0	0	0
1	0	0	1
2	0	1	0
3	0	1	1
4	1	0	0
5	1	0	1
6	1	1	0
7	1	1	1

Table 5.1: List of functions from $\{0, 1, 2\}$ to $\{0, 1\}$

renders the certificate produced by the violation of the Bell inequality useless. However, we are able to introduce measures to circumvent this so-called 'synchronicity loophole'.

5.3 Asymmetry and Asynchronicity

In this section, we briefly mention asymmetric and asynchronous correlations. For a more thorough treatment of the nonsignalling and hidden variables polytopes of these correlations, please refer to [36]. The Bell inequalites from the previous section only hold in the case that we have a synchronous and symmetric correlation. This gives rise to the question of whether one can violate the Bell inequalities using asynchronous and asymmetric correlations. We answer this question in the affirmative. In fact one can construct a simple strategy to produce an asynchronous hidden variables correlation that violates $J_3 \geq 0$. Recall that the hidden variables correlations were characterized by the set of functions from $\{0, 1, 2\} \to \{0, 1\}$. Thus we have eight functions f_i for $i \in \{0, \cdots, 7\}$ that Alice and Bob can each use to produce their outputs. We list the functions in Table 5.1.

The following example shows a strategy that achieves $J_3 = -\frac{1}{8}$ using an asynchronous classical strategy:

Example 50. *The strategy given by the convex sum*

$$\frac{1}{8}\left((f_0, f_7) + (f_7, f_0) + (f_1, f_1) + \cdots + (f_6, f_6)\right) \tag{5.8}$$

gives us the correlation matrix satisfying $J_3 = -\frac{1}{8}$*:*

$$\frac{1}{8}\left(\begin{array}{ccc|ccc|ccc} 3 & 1 & 1 & 1 & 3 & 1 & 1 & 1 & 3 \\ 1 & 3 & 3 & 3 & 1 & 3 & 3 & 3 & 1 \\ 1 & 3 & 3 & 3 & 1 & 3 & 3 & 3 & 1 \\ 3 & 1 & 1 & 1 & 3 & 1 & 1 & 1 & 3 \end{array}\right).$$

Alice and Bob use a shared classical random variable, say $\mu \xleftarrow{\$} \{0, ..., 7\}$ *drawn uniformly to determine their strategy in each round. Here, each value of* μ *corresponds to a function pair in eq.* (5.8)*.*

Chapter 6: Synchronous QKD

The bulk of this chapter appears in our paper [65].

6.1 Introduction

Quantum key distribution (QKD) allows two parties to establish a shared classical secret key using quantum resources. The two main requirements of QKD are (1) Correctness: the two parties, Alice and Bob, get the same key; and (2) Security: an adversary Eve gets negligible information about the key. Device-independent quantum key distribution (DI-QKD) is entanglement-based, and aims to prove security of QKD based solely on the correctness of quantum mechanics, separation of devices used by the two parties, and passing of statistical tests known as Bell violations [28, 30]. These protocols are usually specified by a non-local game, characterized by a conditional probability distribution or *correlation* $p(y_A, y_B \,|\, x_A, x_B)$. Intuitively, Alice and Bob obtain or generate random inputs x_A and x_B respectively, and the correlation describes the likelihood their entangled quantum devices return outputs y_A and y_B to each respectively. We will be interested in symmetric correlations and so will take $x_A, x_B \in X$ and $y_A, y_B \in Y$ where X and Y are finite sets; for our protocol specifically $X = \{0, 1, 2\}$ and $Y = \{0, 1\}$.

In general, security of a DI-QKD scheme relies on the monogamy of entanglement.

The key result is that maximally entangled quantum states are separable within any larger quantum system. In cryptographic terms, if Alice and Bob share a maximally entangled state then the results of measurements they make on this state will be uncorrelated to any other measurement results an adversary can perform. Hence, presuming the correctness of quantum mechanics, no adversary can have any information about key bits Alice and Bob may generate through this process. Generally, a DI-QKD protocol will involve two types of rounds: testing rounds where Alice and Bob (publicly) share their inputs and output results for performing statistics tests, and data rounds where they obtain shared secret bits. The goal of the testing rounds is to produce a certificate that Alice and Bob are operating on maximally entangled states.

Most current DI-QKD schemes are based on the CHSH inequality, a linear inequality in $p(y_A, y_B \mid x_A, x_B)$, which if satisfied characterizes classical statistics within a quantum system, see Section 3.2. Hence a violation of this inequality is a certificate of quantum behavior. This inequality exhibits "rigidity" in that the only quantum state that produces a maximal violation of the inequality is (up to natural equivalences) a Bell pair. Thus the goal of the testing rounds in a DI-QKD protocol is to statistically verify that the system produces a maximal violation of the CHSH inequality.

In a non-local game Alice and Bob may preshare an entangled resource in each round, but are not allowed any communication between receiving or generating their inputs x_A and x_B and measuring the system to obtain their outputs y_A and y_B. This is typically called a "nonsignaling" condition, leading to *nonsignaling correlations* which include all quantum strategies. If (even classical) communication between Alice and Bob is possible, then it is simple to classically simulate a correlation that produces a maximal violation of

the CHSH inequality, and hence any certificates of quantumness or entanglement are void [66]. This *locality* or *causality loophole* in the security proof is challenging to avoid; the only known means to close it is by having Alice and Bob acausally separated during each round: bounds on the speed of light prevent such communication [33, 34, 67].

A *synchronous correlation* is one such that $p(y_A, y_B \mid x, x) = 0$ whenever $y_A \neq y_B$ and $x \in X$. That is, whenever Alice and Bob input the same value they are guaranteed to receive the same outputs, although that value may be nondeterministic. These correlations have recently become popular owing to their use in the resolution of the Connes Embedding Conjecture and Tsirl'son's Problem [68], but have also been used to generalize combinatorial properties to the quantum setting [69, 70, 71].

We present a fully device-independent QKD protocol based on synchronous correlations. This protocol is symmetric, in that roles of Alice and Bob are completely interchangeable. This is an advantage over other DI-QKD protocols based on the CHSH inequality [28] (which is neither symmetric nor synchronous) as sender versus receiver roles do not need to be negotiated. Additionally, as Alice and Bob select their inputs independently they do not need pre-shared secret bits to decide upon testing versus data rounds.

The mathematical framework needed to prove device-independent security of this protocol was laid out in Chapter 5 and [36], where four analogues of the Bell/CHSH inequality for synchronous correlations were given. In this work we focus only on one of these, $J_3(p) \geq 0$ (see eq. (6.3) below). As well, bounds on quantum violations of these were given ($J_3(p) \geq -\frac{1}{8}$), and rigidity of correlations that achieve a maximal violation proven. The two critical analyses needed to complete a proof of security for our DI-QKD protocol are as follows. First, we must prove that if the system is observed to be close to the maximal

violation then it is close to the ideal system, which measures a Bell pair. Then, we provide an alternative security assumption that bypasses the causality loophole.

We tackle the first of these through two theorems. For context, Alice and Bob will select their inputs from $X = \{0, 1, 2\}$ and each measure a quantum system that produces a bit output from $Y = \{0, 1\}$. The ideal system, that produces $J_3(p) = -\frac{1}{8}$, involves measuring a Bell pair using three specific projection-valued measures $\{\hat{E}_y^x\}_{y=0,1}$ for $x = 0, 1, 2$ given in eq. (6.1) below. Any synchronous quantum correlation that achieves $J_3(p) = -\frac{1}{8}$ must have $E_y^x = \hat{E}_y^x \otimes \mathbb{1}$, and hence the measurements have no influence on the larger system. In Section 6.3 we show that if we take a *synchronous* quantum system that is close to achieving maximal J_3 violation, then it must be close to the ideal system in trace norm.

Unfortunately this introduces a "synchronicity" loophole: rigidity holds among synchronous correlations, but are there asynchronous correlations with $J_3 = -\frac{1}{8}$ that cannot certify maximal entanglement? In Section 6.4, we close this loophole using recent work on "almost synchronous" correlations [72]. This leads to our complete DI-QKD scheme given as Algorithm 4 below, where in addition to verifying a Bell violation one also bounds the total amount of asynchronicity of the correlation, S, as defined in eq. (6.7).

In Section 6.5 we use the Entropy Accumulation Theorem (EAT), see Section 3.3 and [31], to bound the the min-entropy of the outputs given an adversary's side-information. This allows us to derive the key rate of Algorithm 4.

Finally, in Section 6.6, we pose a new security assumption to close the *causality* or *locality* loophole: the adversary Eve may have unlimited communication and computational power, yet she has imperfect knowledge of Alice and Bob's inputs. Informally, given nonnegative values $\lambda \leq \frac{1}{8}$ and $\mu \leq 1$ there exists a bound ϵ_{max} such that if Eve's un-

certainty about Alice and Bob's inputs is greater than ϵ_{max} then there is no device she can create where Alice and Bob's expected Bell violation J_3 and asynchronicity S satisfy $-\frac{1}{8} \leq J_3 \leq -\frac{1}{8} + \lambda$ and $0 \leq S \leq \mu$.

6.2 Preliminaries

We present some definitions that will be used in the protocol later. Like other device-independent schemes, our protocol is expressed in terms of a nonlocal game, which is characterized by a conditional probability distribution (or correlation) $p(y_A, y_B|x_A, x_B)$ where $x_A, x_B \in X$ and $y_A, y_B \in Y$ are from finite sets X and Y. By a nonlocal game we mean the players Alice and Bob will receive inputs $x_A, x_B \in X$ from a referee and will produce outputs $y_A, y_B \in Y$. These are then adjudicated by the referee against some criterion, synchronicity in our case. Alice and Bob are not allowed to communicate once they receive their inputs, which is characterized by the famous nonsignaling conditions on the correlation [73, 74].

Unlike nonlocal games such as the CHSH or Magic Square games, or their generalizations [43, 44, 74, 75, 76], it is straightforward for Alice and Bob to create a perfect winning strategy for synchronicity. Prior to the games they agree on some function $f : X \to Y$, then regardless of how the referee selects $x_A, x_B \in X$, they output $y_A = f(x_A)$ and $y_B = f(x_B)$. Hence the "value" of any synchronous game (Alice's and Bob's expected success probability) is always 1, and so value plays no role in the following.

We denote the binary entropy function by $h(p) = -p\log(p) - (1-p)\log(1-p)$ for $p \in [0, 1]$. The von Neumann entropy of a quantum state ρ is given by $H(\rho) = -\mathrm{tr}(\rho\log(\rho))$.

Given two operators ρ_1 and ρ_2, we say $\rho_1 \geq \rho_2$ if $\rho_1 - \rho_2 \geq 0$.

Definition 51. *For a bipartite quantum state $\rho_{AB} \in \mathfrak{H}_A \otimes \mathfrak{H}_B$, the min-entropy of A conditioned on B is:*

$$H_{\min}(A \mid B)_{\rho_{AB}} = \max\{s \in \mathbb{R} : \exists \sigma_B \in \mathcal{D}(\mathfrak{H}_B) \text{ such that } 2^{-s}\mathrm{id}_A \otimes \sigma_B \geq \rho_{AB}\}$$

where $\mathcal{D}(\mathfrak{H}_B)$ is the set of density operators in $\mathfrak{H}_B$.

The ϵ-smooth version of the conditional min-entropy considers states that are ϵ-close to ρ_{AB}. The notion of closeness that is typically used is the purified distance $P(\rho, \sigma) = \sqrt{1 - F(\rho, \sigma)^2}$, where $F(\rho, \sigma)$ is the fidelity between states ρ and σ.

Definition 52. *For a bipartite quantum state $\rho_{AB} \in \mathfrak{H}$, the ϵ-smooth min-entropy of A conditioned on B is defined as:*

$$H^{\epsilon}_{\min}(A \mid B)_{\rho_{AB}} = \max_{\substack{\tilde{\rho}_{AB} \in \mathcal{S}(\mathfrak{H}) \\ P(\rho_{AB}, \tilde{\rho}_{AB}) \leq \epsilon}} H_{\min}(A \mid B)_{\tilde{\rho}_{AB}}$$

The quantum ϵ-smooth max-entropy is defined as:

$$H^{\epsilon}_{\max}(A \mid B)_{\rho_{AB}} = \log \inf_{\substack{\tilde{\rho}_{AB} \in \mathcal{S}(\mathfrak{H}) \\ P(\rho_{AB}, \tilde{\rho}_{AB}) \leq \epsilon}} \sup_{\sigma_B} \left\| \tilde{\rho}_{AB}^{\frac{1}{2}} \sigma_B^{-\frac{1}{2}} \right\|_1^2.$$

where $\mathcal{S}(\mathfrak{H})$ is the set of sub-normalized states in $\mathfrak{H}$ and $\|A\|_\alpha = \mathrm{tr}\left(\left(\sqrt{A^\dagger A}\right)^\alpha\right)^{\frac{1}{\alpha}}$.

6.3 A synchronous DIQKD protocol

We present a synchronous device-independent quantum key distribution protocol that is symmetric with respect to Alice and Bob, each party performing the same tasks.

Suppose Alice and Bob share an EPR pair. Each draws $x_A, x_B \in X = \{0, 1, 2\}$ respectively, and measures according to $\{\hat{E}_y^{x_A}\}_{y\in Y}$ and $\{\hat{E}_y^{x_B}\}_{y\in Y}$ to get outputs $y_A, y_B \in Y = \{0, 1\}$, where the projection-valued measures $\{\hat{E}_y^x\}_{y\in\{0,1\}}$ for $x \in \{0, 1, 2\}$ are:

$$\begin{aligned}
&\hat{E}_1^0 = |\phi_0\rangle\langle\phi_0|,\ \hat{E}_0^0 = \mathbb{1} - \hat{E}_1^0, \quad \text{where } |\phi_0\rangle = |1\rangle \\
&\hat{E}_1^1 = |\phi_1\rangle\langle\phi_1|,\ \hat{E}_0^1 = \mathbb{1} - \hat{E}_1^1, \quad \text{where } |\phi_1\rangle = \tfrac{\sqrt{3}}{2}|0\rangle + \tfrac{1}{2}|1\rangle \\
&\hat{E}_1^2 = |\phi_2\rangle\langle\phi_2|,\ \hat{E}_0^2 = \mathbb{1} - \hat{E}_1^2, \quad \text{where } |\phi_2\rangle = \tfrac{\sqrt{3}}{2}|0\rangle - \tfrac{1}{2}|1\rangle
\end{aligned} \tag{6.1}$$

The likelihood of Alice's and Bob's results are characterized by the correlation in Theorem 41:

$$p(y_A, y_B \mid x_A, x_B) = \frac{1}{2}\mathrm{tr}\left(\hat{E}_{y_A}^{x_A}\hat{E}_{y_B}^{x_B}\right).$$

In particular, this strategy produces a synchronous quantum correlation with correlation matrix:

$$[p(y_A, y_B|x_A, x_B)] =$$

$$\frac{1}{8}\begin{array}{c} \begin{matrix} (0,0) & (0,1) & (0,2) & (1,0) & (1,1) & (1,2) & (2,0) & (2,1) & (2,2) \end{matrix} \\ \begin{pmatrix} 4 & 1 & 1 & 1 & 4 & 1 & 1 & 1 & 4 \\ 0 & 3 & 3 & 3 & 0 & 3 & 3 & 3 & 0 \\ 0 & 3 & 3 & 3 & 0 & 3 & 3 & 3 & 0 \\ 4 & 1 & 1 & 1 & 4 & 1 & 1 & 1 & 4 \end{pmatrix} \begin{matrix} (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{matrix} \end{array} \tag{6.2}$$

One can verify this correlation yields a maximal violation of the Bell inequality, $J_3 = -\frac{1}{8}$, where

$$\begin{aligned} J_3 \;=\; & 1 - \tfrac{1}{4}\big(p(0,1 \mid 0,1) + p(1,0 \mid 0,1) + p(0,1 \mid 1,0) + p(1,0 \mid 1,0) \\ & + p(0,1 \mid 0,2) + p(1,0 \mid 0,2) + p(0,1 \mid 2,0) + p(1,0 \mid 2,0) \\ & + p(0,1 \mid 1,2) + p(1,0 \mid 1,2) + p(0,1 \mid 2,1) + p(1,0 \mid 2,1) \big). \end{aligned} \tag{6.3}$$

This correlation is rigid in that any synchronous quantum correlation that achieves $J_3 = -\frac{1}{8}$ must have implemented the strategy above. This follows from our Theorem 53 below. In particular, this maximal violation of J_3 is a self-test of the device to detect interference from adversary: Alice and Bob can certify that their devices hold maximally entangled pairs, and by monogamy of entanglement can establish that Eve doesn't have any information about their inputs.

Our protocol extends the above scenario to n rounds. It is important to note that the observable for our synchronous Bell inequality (6.3) only involves correlations where Alice and Bob use different inputs. Critically, neither Alice nor Bob must pre-select which rounds will used for testing versus key generation. Upon revealing their choices of bases,

testing rounds are given by those where they selected different bases and key generation rounds where they selected the same basis. In particular, they need not have any pre-shared randomness.

Of course no physical device adheres to a theoretical model perfectly, so in practice one still must perform standard information reconciliation and privacy amplification on the results.

Once the n rounds of the protocol are over, Alice and Bob communicate their basis selection over an authenticated classical channel. When they chose different bases (i.e. $x_A \neq x_B$), they exchange their measurement outcomes and use those to compute J_3. If the value of J_3 deviates too much from $-\frac{1}{8}$, they abort. The protocol is synchronous, therefore $y_A = y_B$ whenever $x_A = x_B$ and those can be used as the raw key bits for further standard privacy amplification and information reconciliation.

Our first main result is our technical rigidity statement that synchronous quantum correlations near $J_3 = -\frac{1}{8}$ have the desired security. Informally, after splitting off a space $\mathfrak{L}$ of small relative dimension, the correlation's projections are near (in trace norm) the ideal one, which separates Alice and Bob performing the perfect protocol on $\mathbb{C}^2$, and Eve and all other parties receiving no information having measurement outcomes from $\mathbb{1}_{\mathfrak{K}}$.

Theorem 53. *Let $p(y_A, y_B \mid x_A, x_B) = \frac{1}{d}\mathrm{tr}(E_{y_A}^{x_A} E_{y_B}^{x_B})$ be a synchronous quantum correlation with maximally entangled state, where $\{E_y^x\}$ is a projection-valued measure on a d-dimensional Hilbert space $\mathfrak{H}$. Suppose $J_3(p) \leq -\frac{1}{8} + \lambda$. Then on $\mathfrak{H} = \mathfrak{L} \oplus (\mathbb{C}^2 \otimes \mathfrak{K})$ there exists a projection-value measure $\{\tilde{E}_y^x\}$ where (1) $\tilde{E}_y^x = L_y^x + \hat{E}_y^x \otimes \mathbb{1}_{\mathfrak{K}}$, (2) $\frac{\dim \mathfrak{L}}{\dim \mathfrak{H}} \leq 8\lambda$, and*

(3) $\frac{1}{3}\sum_{x,y}\frac{1}{d}\mathrm{tr}\left(\left(E_y^x-\tilde{E}_y^x\right)^2\right)\leq 8\lambda$. *In particular, the expected statistical difference*

$$\frac{1}{3}\sum_{x,y}\left|p(y,y\mid x,x)-\frac{1}{2}\right|\leq\frac{1}{3}\left(\sqrt{8}\sqrt{\lambda}+32\lambda\right).$$

Proof. We begin by defining the ± 1-valued observables $M_x=E_0^x-E_1^x$, so $M_x^2=\mathbb{1}$, and following customary notation write

$$a_x=\frac{1}{d}\mathrm{tr}(M_x)\text{ and }c_{x_Ax_B}=\frac{1}{d}\mathrm{tr}(M_{x_A}M_{x_B}).$$

Similarly denote $\tilde{M}_x=\tilde{E}_0^x-\tilde{E}_1^x$. Notice $E_0^x=\frac{1}{2}(\mathbb{1}+M_x)$ and $E_1^x=\frac{1}{2}(\mathbb{1}-M_x)$ so

$$\frac{1}{3}\sum_{x,y}\frac{1}{d}\mathrm{tr}\left(\left(E_y^x-\tilde{E}_y^x\right)^2\right)=\frac{1}{6}\sum_x\frac{1}{d}\mathrm{tr}\left(\left(M_x-\tilde{M}_x\right)^2\right).$$

Now define $\Delta:=M_0+M_1+M_2$, and compute

$$\begin{aligned}\Delta^2&=M_0^2+M_1^2+M_2^2+M_0M_1+M_1M_0+M_0M_2+M_2M_0+M_1M_2+M_2M_1\\&=3\mathbb{1}+M_0M_1+M_1M_0+(M_0+M_1)M_2+M_2(M_0+M_1)\qquad(6.4)\\&=\mathbb{1}+M_0M_1+M_1M_0+(M_0+M_1+M_2)M_2+M_2(M_0+M_1+M_2)\\&=\mathbb{1}+M_0M_1+M_1M_0+\Delta M_2+M_2\Delta\qquad(6.5)\end{aligned}$$

We have Δ^2 relates to J_3, and hence we obtain the following bound:

$$\frac{1}{d}\mathrm{tr}(\Delta^2)=\frac{1}{d}\mathrm{tr}\left(M_0^2+M_1^2+M_2^2+2M_0M_1+2M_0M_2+2M_1M_2\right)$$

$$= \frac{3}{d}\mathrm{tr}\,(\mathbb{1}) + \frac{2}{d}\mathrm{tr}\,(M_0M_1 + M_0M_2 + M_1M_2)$$

$$= 3 + 2(c_{01} + c_{02} + c_{12}) = 1 + 2(1 + c_{01} + c_{02} + c_{12}) = 1 + 8J_3$$

$$\leq 1 + 8\left(-\frac{1}{8} + \lambda\right) = 8\lambda \tag{6.6}$$

Using two projections theory [77, 78, 79], we have a decomposition of the Hilbert space $\mathfrak{H}$

$$\mathfrak{H} = \mathfrak{L}_{00} \oplus \mathfrak{L}_{01} \oplus \mathfrak{L}_{10} \oplus \mathfrak{L}_{11} \oplus \bigoplus_{j=1}^{k} \mathfrak{H}_j,$$

where $\dim(\mathfrak{L}_{\alpha\beta}) = l_{\alpha\beta}$ for $\alpha, \beta \in \{0, 1\}$, and $\dim(\mathfrak{H}_j) = 2$, where the projections E_0^0 and E_0^1 take the form:

$$E_0^0 = 0_{l_{00}} \oplus 0_{l_{01}} \oplus \mathbb{1}_{l_{10}} \oplus \mathbb{1}_{l_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$E_0^1 = 0_{l_{00}} \oplus \mathbb{1}_{l_{01}} \oplus 0_{l_{10}} \oplus \mathbb{1}_{l_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} \cos^2\theta_j & \sin\theta_j\cos\theta_j \\ \sin\theta_j\cos\theta_j & \sin^2\theta_j \end{pmatrix}.$$

That is, we can express

$$M_0 = -\mathbb{1}_{\mathfrak{L}_{00}} \oplus -\mathbb{1}_{\mathfrak{L}_{01}} \oplus \mathbb{1}_{\mathfrak{L}_{10}} \oplus \mathbb{1}_{\mathfrak{L}_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$M_1 = -\mathbb{1}_{\mathfrak{L}_{00}} \oplus \mathbb{1}_{\mathfrak{L}_{01}} \oplus -\mathbb{1}_{\mathfrak{L}_{10}} \oplus \mathbb{1}_{\mathfrak{L}_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} \cos 2\theta_j & \sin 2\theta_j \\ \sin 2\theta_j & -\cos 2\theta_j \end{pmatrix}.$$

Now let us define $\tilde{M}_0, \tilde{M}_1, \tilde{M}_2$ as follows. Note that our ideal projections $\hat{E}_0^1, \hat{E}_1^1$ correspond

to angle $\hat{\theta} = \frac{2\pi}{3}$, and without loss of generality we can assume[1] $|\theta_j - \hat{\theta}| \leq \frac{\pi}{6}$.

$$\tilde{M}_0 = M_0 = -\mathbb{1}_{\mathfrak{L}_{00}} \oplus -\mathbb{1}_{\mathfrak{L}_{01}} \oplus \mathbb{1}_{\mathfrak{L}_{10}} \oplus \mathbb{1}_{\mathfrak{L}_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\tilde{M}_1 = -\mathbb{1}_{\mathfrak{L}_{00}} \oplus \mathbb{1}_{\mathfrak{L}_{01}} \oplus -\mathbb{1}_{\mathfrak{L}_{10}} \oplus \mathbb{1}_{\mathfrak{L}_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} \cos 2\hat{\theta} & \sin 2\hat{\theta} \\ \sin 2\hat{\theta} & -\cos 2\hat{\theta} \end{pmatrix},$$

$$\tilde{M}_2 = \mathbb{1}_{\mathfrak{L}_{00}} \oplus \mathbb{1}_{\mathfrak{L}_{01}} \oplus -\mathbb{1}_{\mathfrak{L}_{10}} \oplus -\mathbb{1}_{\mathfrak{L}_{11}} \oplus \bigoplus_{j=1}^{k} \begin{pmatrix} -1-\cos 2\hat{\theta} & -\sin 2\hat{\theta} \\ -\sin 2\hat{\theta} & 1+\cos 2\hat{\theta} \end{pmatrix}.$$

As desired, $\tilde{M}_x = (L_0^x - L_1^x) + \hat{M}_x \otimes \mathbb{1}_{\mathbb{C}^k}$, where the $\{L_y^x\}$ are the projection onto the summands $\mathfrak{L}_{\mu\nu}$.

First we bound the dimension of each $\mathfrak{L}_{\mu\nu}$. Consider (6.4) for Δ^2. If $|\psi_{01}\rangle \in \mathfrak{L}_{01}$, then

$$\begin{aligned}\langle\psi_{01}|\Delta^2|\psi_{01}\rangle &= \langle\psi_{01}|(3\mathbb{1} + M_0M_1 + M_1M_0 + (M_0+M_1)M_2 + M_2(M_0+M_1)|\psi_{01}\rangle \\ &= 3 - 1 - 1 + 0 + 0 = 1.\end{aligned}$$

The same equality holds for $|\psi_{10}\rangle \in \mathfrak{L}_{10}$, namely $\langle\psi_{10}|\Delta^2|\psi_{10}\rangle = 1$.

For a vector $|\psi_{00}\rangle$ in $\mathfrak{L}_{00}$ we again use (6.4) to get $\langle\psi_{00}|\Delta^2|\psi_{00}\rangle = 3 + 1 + 1 - 4\langle\psi_{00}|M_2|\psi_{00}\rangle$.

Now from Cauchy-Schwarz, and that $M_2^2 = \mathbb{1}$, we have

$$|\langle\psi_{00}|M_2|\psi_{00}\rangle| \leq |\langle\psi_{00}|\psi_{00}\rangle|^{\frac{1}{2}} |\langle\psi_{00}|M_2^2|\psi_{00}\rangle|^{\frac{1}{2}} = 1$$

[1] Direct examination of (6.1) reveals that any θ_j is within $\frac{\pi}{6}$ of the image of some E_y^x; the bound we prove is symmetric in x, y we may reorder the labeling in each $\mathfrak{H}_j$ so that θ_j is close to E_0^1 with $\hat{\theta} = \frac{2\pi}{3}$.

and thus $\langle\psi_{00}|\Delta^2|\psi_{00}\rangle \geq 1$. Similarly for $|\psi_{11}\rangle$ in $\mathfrak{L}_{11}$ we have

$$\langle\psi_{11}|\Delta^2|\psi_{11}\rangle = 5 + 4\langle\psi_{11}|M_2|\psi_{11}\rangle \geq 5 - 4|\langle\psi_{11}|M_2|\psi_{11}\rangle| \geq 1.$$

Putting everything together, since $\langle\psi_{\alpha\beta}|\Delta^2|\psi_{\alpha\beta}\rangle \geq 1$ on each $\mathfrak{L}_{\alpha\beta}$, for $\alpha, \beta \in \{0, 1\}$, summing over bases of the respective spaces

$$\frac{l}{d} = \frac{1}{d}(l_{00} + l_{01} + l_{10} + l_{11}) \leq \frac{1}{d}\sum_{j=1}^{l}\langle\psi_j|\Delta^2|\psi_j\rangle \leq \frac{1}{d}\mathrm{tr}(\Delta^2) \leq 8\lambda.$$

where the second-to-last inequality follows from Δ^2 being positive semidefinite.

This immediately provides the claimed bound on the statistical difference from uniform. We can explicitly bound the quantities $|a_0|$ and $|a_1|$ as follows:

$$|a_0| = \frac{1}{d}\left|\mathrm{tr}(M_0)\right| = \frac{1}{d}|-l_{00} - l_{01} + l_{10} + l_{11}| \leq \frac{l}{d} \leq 8\lambda$$
$$|a_1| = \frac{1}{d}\left|\mathrm{tr}(M_1)\right| = \frac{1}{d}|-l_{00} + l_{01} - l_{10} + l_{11}| \leq \frac{l}{d} \leq 8\lambda.$$

Using Cauchy-Schwarz, we bound $|a_2|$. As

$$a_0 + a_1 + a_2 = \frac{1}{d}\mathrm{tr}(\Delta) \leq \left(\frac{1}{d}\mathrm{tr}(\Delta^2)\right)^{\frac{1}{2}}\left(\frac{1}{d}\mathrm{tr}(\mathbb{1}^2)\right)^{\frac{1}{2}} \leq \sqrt{8\lambda},$$

we have $a_2 \leq \sqrt{8\lambda} - a_0 - a_1$ and therefore $|a_2| \leq \sqrt{8\lambda} + |a_0| + |a_1| \leq \sqrt{8}\sqrt{\lambda} + 16\lambda$.

Finally we bound each $\frac{1}{d}\mathrm{tr}\left(\left(M_x - \tilde{M}_x\right)^2\right)$. Note $M_0 - \tilde{M}_0 = 0$ by construction.

Then

$$\frac{1}{d}\mathrm{tr}\left(\left(M_1-\tilde{M}_1\right)^2\right)=\frac{1}{d}\sum_j \mathrm{tr}\left(\begin{pmatrix}\cos 2\theta_j-\cos 2\hat{\theta} & \sin 2\theta_j-\sin 2\hat{\theta}\\ \sin 2\theta_j-\sin 2\hat{\theta} & -\cos 2\theta_j+\cos 2\hat{\theta}\end{pmatrix}^2\right)$$
$$=\frac{1}{d}\sum_j(4-4\cos(2(\theta_j-\hat{\theta})))\;=\;\frac{8}{d}\sum_j\sin^2(\theta_j-\hat{\theta})$$

To bound this, we note that on any $\mathfrak{H}_j$:

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\cos 2\theta_j & \sin 2\theta_j\\ \sin 2\theta_j & -\cos 2\theta_j\end{pmatrix}+\begin{pmatrix}\cos 2\theta_j & \sin 2\theta_j\\ \sin 2\theta_j & -\cos 2\theta_j\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}=2\cos 2\theta_j\cdot\mathbb{1}_{\mathfrak{H}_j}.$$

From this we obtain

$$\left[\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}+\begin{pmatrix}\cos 2\theta_j & \sin 2\theta_j\\ \sin 2\theta_j & -\cos 2\theta_j\end{pmatrix}\right]^2=4\cos^2\theta_j\mathbb{1}_{\mathfrak{H}_j}.$$

Hence there exists a basis $\{|\psi_0\rangle,|\psi_1\rangle\}$ of $\mathfrak{H}_j$ such that

$$(M_0+M_1)|\psi_0\rangle=2\cos\theta_j|\psi_0\rangle \text{ and } (M_0+M_1)|\psi_1\rangle=-2\cos\theta_j|\psi_1\rangle.$$

Therefore again from (6.4) we have

$$\langle\psi_0|\Delta^2|\psi_0\rangle=3+2\cos 2\theta_j+4\cos\theta_j\langle\psi_0|M_2|\psi_0\rangle$$
$$\langle\psi_1|\Delta^2|\psi_1\rangle=3+2\cos 2\theta_j-4\cos\theta_j\langle\psi_1|M_2|\psi_1\rangle.$$

In particular, $\langle\psi_0|\Delta^2|\psi_0\rangle + \langle\psi_1|\Delta^2|\psi_1\rangle \geq 6 + 4\cos 2\theta_j - 8|\cos\theta_j|$. It is straightforward to show for $\theta \in \left[\frac{2\pi}{3} - \frac{\pi}{6}, \frac{2\pi}{3} + \frac{\pi}{6}\right]$ that $6 + 4\cos 2\theta - 8|\cos\theta| \geq 4\sin^2\left(\theta - \frac{2\pi}{3}\right)$. And hence we obtain the bound

$$\begin{aligned}\frac{1}{d}\mathrm{tr}(\Delta^2) &\geq \frac{1}{d}\sum_j (6 + 4\cos 2\theta_j - 8|\cos\theta_j|) \\ &\geq \frac{1}{d}\sum_j 4\sin^2(\theta_j - \hat{\theta}) \;=\; \frac{1}{2d}\mathrm{tr}\left(\left(M_1 - \tilde{M}_1\right)^2\right).\end{aligned}$$

In particular, $\frac{1}{d}\mathrm{tr}\left(\left(M_1 - \tilde{M}_1\right)^2\right) \leq 16\lambda$.

Finally, note $\tilde{M}_0 + \tilde{M}_1 + \tilde{M}_2 = -\mathbb{1}_{\mathfrak{L}_{00}} \oplus \mathbb{1}_{\mathfrak{L}_{01}} \oplus -\mathbb{1}_{\mathfrak{L}_{10}} \oplus \mathbb{1}_{\mathfrak{L}_{11}}$. By Jensen's inequality

$$\begin{aligned}\frac{1}{d}\mathrm{tr}\left(\left(M_2 - \tilde{M}_2\right)^2\right) &= \frac{1}{d}\mathrm{tr}\left(\left(\Delta - (-\mathbb{1}_{\mathfrak{L}_{00}} \oplus \mathbb{1}_{\mathfrak{L}_{01}} \oplus -\mathbb{1}_{\mathfrak{L}_{10}} \oplus \mathbb{1}_{\mathfrak{L}_{11}}) + (\tilde{M}_1 - M_1)\right)^2\right) \\ &\leq \frac{1}{d}\mathrm{tr}\left(\Delta^2\right) + \frac{1}{d}\mathrm{tr}\left(\mathbb{1}_{\mathfrak{L}}\right) + \frac{1}{d}\mathrm{tr}\left(\left(\tilde{M}_1 - M_1\right)^2\right) \leq 32\lambda.\end{aligned}$$

Therefore, $\frac{1}{3}\sum_{x,y}\frac{1}{d}\mathrm{tr}\left(\left(E_y^x - \tilde{E}_y^x\right)^2\right) \leq 8\lambda$ as desired.

It is straightforward to get a bound on the statistical difference to any synchronous quantum correlation close to $J_3 = -\frac{1}{8}$. Every synchronous quantum correlation is a convex sum of synchronous quantum correlations with maximally entangled states, and so we may write $p = \sum_j c_j p_j$ where p_j is as in the theorem above. Say $J_3(p_j) \leq -\frac{1}{8} + \lambda_j$, and so

$$J_3(p) = \sum_j c_j J_3(p_j) \leq -\frac{1}{8} + \sum_j c_j\lambda_j = -\frac{1}{8} + \lambda$$

where we define $\lambda = \sum_j c_j \lambda_j$. With two uses of Jensen's inequality,

$$\begin{aligned}\frac{1}{3}\sum_{x,y}\left|p(y,y\mid x,x)-\frac{1}{2}\right| &\leq \frac{1}{3}\sum_{j,x,y} c_j \left|p_j(y,y\mid x,x)-\frac{1}{2}\right| \\ &\leq \sum_j c_j(C\sqrt{\lambda_j}+C'\lambda_j) \leq C\sqrt{\lambda}+C'\lambda.\end{aligned}$$

□

Unfortunately, this does not yet produce a fully device-independent protocol as we still suffer from a "synchronicity" loophole. We discuss this loophole and close the loophole in the next section.

6.4 Measure of asynchronicity

That $J_3 = -\frac{1}{8}$ can be achieved by a unique synchronous quantum correlation, which necessarily can only be realized through a maximally entangled state, provides the device-independent security of the above QKD scheme. However this opens a "synchronicity" security loophole: can a (asynchronous) quantum device simulate $J_3 = -\frac{1}{8}$ without using maximally entangled states (and hence potentially leak information about the derived shared keys)? Fortunately a recent work shows that the same results apply to "almost" synchronous correlations [72]. This allows us to close this synchronicity loophole by also bounding the asynchronicity of the observed correlation.

Definition 54. *The* asynchronicity *with respect to a basis choice $x \in X$ and set of measurement outcomes Y is $S_x(p) = \sum_{y_A \neq y_B} p(y_A, y_B|x,x)$. The* total *(or* expected*) asynchronicity*

is

$$S(p) = \frac{1}{|X|} \sum_{x \in X} S_x(p) \tag{6.7}$$

In [72], this measure is called the "default to synchronicity" and denoted δ_{sync}. While the expected asynchronicity is the average likelihood of an asynchronous result where the inputs are sampled uniformly at random, all results here and in [72], apply to the case where the expectation is computed over inputs sampled with respect to some other fixed distribution. To bound the asynchronicity, we modify the scheme in Section 6.3 so that for some data rounds where Alice and Bob have selected the same inputs they still reveal their output, stated as Algorithm 4 below.

Here we state the main result [72, Theorem 3.1] in the notation used above. Note that this theorem refers to symmetric (albeit asynchronous) correlations, which is the natural setting as every synchronous quantum correlation is symmetric. This implies a special form for the projections in the correlation, involving the transpose with respect to the natural basis given by the Schmidt-decomposition of the entangled state used in the correlation.

Theorem 55 (Vidick)**.** *There are universal constants $c, C > 0$ such that the following holds. Let X and Y be finite sets and p a symmetric quantum correlation with input set X, measurement results Y, and asynchronicity $S = S(p)$. Write $p(y_A, y_B \mid x_A, x_B) = \langle \psi | E_{y_A}^{x_A} \otimes (E_{y_B}^{x_B})^T | \psi \rangle$ where $\{E_y^x\}_{y \in Y}$ is a POVM on a finite-dimensional Hilbert space $\mathfrak{H}$ and $|\psi\rangle$ a state on $\mathfrak{H} \otimes \mathfrak{H}$. Let $|\psi\rangle = \sum_{j=1}^{r} \sqrt{\sigma_j} \sum_{m=1}^{d_j} |\phi_{j,m}^A\rangle \otimes |\phi_{j,m}^B\rangle$ be the Schmidt decomposition, and write $|\psi_j\rangle = \frac{1}{\sqrt{d_j}} \sum_{m=1}^{d_j} |\phi_{j,m}^A\rangle \otimes |\phi_{j,m}^B\rangle$. Then*

1. *$\mathfrak{H} = \bigoplus_{j=1}^{r} \mathfrak{H}_j$ with $|\psi_j\rangle$ being maximally entangled on $\mathfrak{H}_j \otimes \mathfrak{H}_j$;*

2. *there is a projective measurement* $\{E_y^{j,x}\}_{y\in Y}$ *on each* $\mathfrak{H}_j$ *so that*

$$p_j(y_A, y_B \mid x_A, x_B) = \langle \psi_j | E_{y_A}^{j,x_A} \otimes (E_{y_B}^{j,x_B})^T | \psi_j \rangle = \frac{1}{d_j} \mathrm{tr}(E_{y_A}^{j,x_A} E_{y_B}^{j,x_B})$$

is a synchronous quantum correlation and $p \approx \sum_{j=1}^r d_j \sigma_j p_j$ *in that:*

$$\frac{1}{|X|} \sum_{x\in X} \sum_{y\in Y} \sum_{j=1}^r \frac{1}{d_j} \sum_{m=1}^{d_j} \langle \phi_{j,m}^A | \left(E_y^x - E_y^{j,x}\right)^2 | \phi_{j,m}^A \rangle \leq CS^c.$$

As indicated in [72, §4.1], this result can be used to transfer rigidity from synchronous to almost synchronous correlations. As $\sum_j d_j \sigma_j = 1$, we also transfer the bound on the statistical difference from uniform to convex sums in this theorem exactly as in the previous section. As for the full correlation we rephrase Lemma 2.10 of [72] in the context of Theorem 55 as follows.

Corollary 56. *Let* $p(y_A, y_B \mid x_A, x_B) = \langle \psi | E_{y_A}^{x_A} \otimes (E_{y_B}^{x_B})^T | \psi \rangle$ *be a quantum correlation with asynchronocity* S *as in Theorem 55, and let* $\bar{p} = \sum_{j=1}^r d_j \sigma_j p_j$ *with*

$$\frac{1}{|X|} \sum_{x\in X} \sum_{y\in Y} \sum_{j=1}^r \frac{1}{d_j} \sum_{m=1}^{d_j} \langle \phi_{j,m}^A | \left(E_y^x - E_y^{j,x}\right)^2 | \phi_{j,m}^A \rangle = \gamma$$

as given in Theorem 55. Then

$$\frac{1}{|X|^2} \sum_{x_A, x_B, y_A, y_B} |p(y_A, y_B \mid x_A, x_B) - \bar{p}(y_A, y_B \mid x_A, x_B)| \leq 3S + 4\sqrt{\gamma}.$$

Note that this bound on the statistical difference directly bounds $J_3(p)$ in terms of the convex sum of the analogous $J_3(p_j)$. Note that J_3, as seen in (6.3), is an affine function so

$J_3(\bar{p}) = \sum_{j=1}^{r} \sigma_j d_j J_3(p_j)$ using the notation of Theorem 55 above. Then immediately from Corollary 56, $|J_3(p) - J_3(\bar{p})| \leq \frac{27}{4} S + 9\sqrt{\gamma}$. In turn from Theorem 55 we have $\gamma \leq CS^c$, and so there are different universal constants C', c' so that

$$|J_3(p) - J_3(\bar{p})| \leq C' S^{c'}. \tag{6.8}$$

Corollary 57. *Let $p(y_A, y_B \mid x_A, x_B) = \langle \psi | E_{y_A}^{x_A} \otimes (E_{y_B}^{x_B})^T | \psi \rangle$ be a quantum correlation as in Theorem 55 and suppose $J_3(p) = -\frac{1}{8} + \lambda$. Then the Hilbert space decomposes as $\mathfrak{H} = \bigoplus_{j=1}^{r} \mathfrak{H}_j = \bigoplus_{j=1}^{r} (\mathfrak{L}_j \oplus (\mathbb{C}^2 \otimes \mathfrak{K}_j))$ where $\frac{\dim \mathfrak{L}_j}{\dim \mathfrak{H}_j} \leq 8\lambda_j$. On each summand we have projection-valued measures $\{\tilde{E}_y^{j,x}\}$ such that $\tilde{E}_y^{j,x} = L_y^{j,x} + \hat{E}_y^x \otimes \mathbb{1}_{\mathfrak{K}_j}$ and*

$$\frac{1}{3} \sum_{x,y} \sum_{j=1}^{r} \sigma_j d_j \left(\frac{1}{d_j} \sum_{m=1}^{d_j} \langle \phi_{j,m}^A | (E_y^x - \tilde{E}_y^{j,x})^2 | \phi_{j,m}^A \rangle \right) \leq C_1 S^c + C_2 \lambda$$

for universal constants c, C_1, C_2.

Proof. Given $\{E_y^x\}$ as above, we obtain projections $\{E_y^{j,x}\}$ defining synchronous correlations p_j from Theorem 55. Write $J_3(p_j) = -\frac{1}{8} + \lambda_j$. From Theorem 53, we obtain the given decomposition of the Hilbert space and projection-valued measures $\{\tilde{E}_y^{j,x}\}$ where

1. $\tilde{E}_y^{j,x} = L_y^{j,x} + \hat{E}_y^x \otimes \mathbb{1}_{\mathfrak{K}_j}$,
2. $\frac{\dim \mathfrak{L}_j}{\dim \mathfrak{H}_j} \leq 8\lambda_j$, and
3. $\frac{1}{3} \sum_{x,y} \frac{1}{d_j} \sum_{m=1}^{d_j} \langle \phi_{j,m}^A | (E_y^{j,x} - \tilde{E}_y^{j,x})^2 | \phi_{j,m}^A \rangle \leq C_2 \lambda_j$.

Then using the notation and (6.8) above $|J_3(p) - J_3(\bar{p})| = \left| \lambda - \sum_{j=1}^{r} \sigma_j d_j \lambda_j \right| \leq C' S^{c'}$ and

thus

$$\frac{1}{3}\sum_{x,y}\sum_{j=1}^{r}\sigma_j d_j\left(\frac{1}{d_j}\sum_{m=1}^{d_j}\langle\phi_{j,m}^A|(E_y^{j,x}-\tilde{E}_y^{j,x})^2|\phi_{j,m}^A\rangle\right)\leq C_2\sum_{j=1}^{r}\sigma_j d_j\lambda_j = C_2\lambda + C_2C'S^{c'}.$$

On the other hand,

$$\begin{aligned}&\frac{1}{3}\sum_{x,y}\sum_{j=1}^{r}\sigma_j d_j\left(\frac{1}{d_j}\sum_{m=1}^{d_j}\langle\phi_{j,m}^A|(E_y^{x}-E_y^{j,x})^2|\phi_{j,m}^A\rangle\right)\\ &\qquad\leq\frac{1}{3}\sum_{x,y}\sum_{j=1}^{r}\left(\frac{1}{d_j}\sum_{m=1}^{d_j}\langle\phi_{j,m}^A|(E_y^{x}-E_y^{j,x})^2|\phi_{j,m}^A\rangle\right)\leq C''S^{c''}\end{aligned}$$

directly from Theorem 55. So by Jensen's inequality

$$\begin{aligned}&\frac{1}{3}\sum_{x,y}\sum_{j=1}^{r}\sigma_j d_j\left(\frac{1}{d_j}\sum_{m=1}^{d_j}\langle\phi_{j,m}^A|(E_y^{x}-\tilde{E}_y^{j,x})^2|\phi_{j,m}^A\rangle\right)\\ &\qquad\leq\frac{2}{3}\sum_{x,y}\sum_{j=1}^{r}\sigma_j d_j\left(\frac{1}{d_j}\sum_{m=1}^{d_j}\langle\phi_{j,m}^A|(E_y^{x}-E_y^{j,x})^2|\phi_{j,m}^A\rangle\right)\\ &\qquad\quad+\frac{2}{3}\sum_{x,y}\sum_{j=1}^{r}\sigma_j d_j\left(\frac{1}{d_j}\sum_{m=1}^{d_j}\langle\phi_{j,m}^A|(E_y^{j,x}-\tilde{E}_y^{j,x})^2|\phi_{j,m}^A\rangle\right)\\ &\qquad\leq 2C_1S^c+2C_2\lambda\end{aligned}$$

for some universal constant C_1. □

6.5 Security and key-rate analysis

Our synchronous fully device-independent quantum key distribution protocol is stated in Algorithm 4. For an honest, but possibly noisy implementation of the protocol, we assume

Algorithm 4: Synchronous QKD Protocol

Input:

$\lambda \in [0, \frac{1}{8})$: Allowed error in J_3 violation

$\mu \in [0, \mu_0]$: Allowed error in asynchronicity S; μ_0 is a pre-decided threshold

$n \in \mathbb{N}$: Total number of rounds

$m \in \mathbb{N}$: Parameter for choosing asynchronicity check rounds. $\kappa := \frac{1}{m}$

$\gamma \in (0, 1]$: Expected fraction of test rounds

$\delta_{est}^{J_3} \in (0, 1)$: Width of statistical interval for the J_3 test

$\delta_{est}^{S} \in (0, 1)$: Width of statistical interval for the S test

EC: Error Correction protocol

PA: Privacy Amplification protocol

For $i \in [n]$:

Alice and Bob draw $x_A^i \leftarrow X$, $x_B^i \leftarrow X$ according to Equation (6.9)

They produce outputs y_A^i and y_B^i using $\{E_y^{x_A^i}\}$ and $\{E_y^{x_B^i}\}$ respectively

They share their inputs x_A^i and x_B^i.

Error Correction: Alice and Bob use error correction protocol EC to obtain outputs $\tilde{Y}_A$ and $\tilde{Y}_B$. They abort if the error correction protocol aborts.

Parameter Estimation:

Bob estimates the J_3 violation in rounds where $x_A^i \neq x_B^i$, i.e. he sets $R_i = 1$ if $\tilde{y}_A^i \neq y_B^i$ else 0. He aborts if $\sum_i R_i < \left[\gamma\left(\frac{3}{4} - \frac{2}{3}\lambda\right) - \delta_{est}^{J_3}\right] \cdot n$

He also estimates the asynchronicity S in rounds where $x_A^i = x_B^i$ and $i \pmod m = 0$, i.e. he sets $Q_i = 1$ if $\tilde{y}_A^i \neq y_B^i$ else 0 in those rounds. He aborts if $\sum_i Q_i < \left[\kappa(1-\gamma)\mu - \delta_{est}^{S}\right] \cdot n$

Privacy Amplification: Alice and Bob use privacy amplification protocol PA to create final keys K_A and K_B using $\tilde{y}_A^i$ and $\tilde{y}_B^i$ where $x_A^i = x_B^i$ and $i \pmod m \neq 0$.

that Alice and Bob perform measurements $E_{y_A}^{x_A^i} \otimes E_{y_B}^{x_B^i}$ on the state ρ_{AB}. We assume a depolarization channel and take ρ_{AB} to be the state $(1-\nu)|\Phi^+\rangle\langle\Phi^+| + \frac{\nu}{4}\mathbb{1}$, where $\nu \in [0, 1]$ is the depolarization noise and $|\Phi^+\rangle$ is the EPR pair. Using measurements according to eq. (6.1), we get $J_3 = -\frac{1}{8} + \frac{3}{8}\nu$, and $S = \frac{\nu}{2}$. A general framework for analyzing device-independent protocols was laid out in [32], which we use to show completeness and soundness of our protocol.

Lemma 58 (Completeness)**.** *Let ϵ_{EC}^c be the completeness error of the EC protocol, and ϵ_{EC} be the probability that the EC protocol does not abort but Alice and Bob hold dif-*

ferent outputs post error correction. Then, Protocol 4 has completeness error $\epsilon^c_{QKD} \leq exp\left(-2n\left((\delta^S_{est})^2\right) + (\delta^{J_3}_{est})^2\right)\right) + \epsilon^c_{EC} + \epsilon_{EC}$.

Proof. The protocol either aborts in the error correction step or the parameter estimation step. The probability of aborting during the J_3 and S tests can be bounded using Hoeffding's inequality as follows:

$$\Pr\left(\sum_j R_j > \left[\gamma\left(\frac{3}{4} - \frac{2}{3}\lambda\right) - \delta^{J_3}_{est}\right]\cdot n \bigwedge \sum_j Q_j > \left[\kappa(1-\gamma)\mu - \delta^S_{est}\right]\cdot n\right) \leq exp\left(-2n\left((\delta^S_{est})^2\right) + (\delta^{J_3}_{est})^2\right)\right).$$

The rest of the proof follows analogously to [32, Lemma 5.2 and Eq. 5.2] □

We use the Entropy Accumulation Theorem (EAT) [31], to bound the min-entropy of Alice and Bob's outputs with respect to an adversary Eve's side information. To that effect, we define Ω as the event that Alice and Bob do not abort the protocol in the parameter estimation step. The EAT yields a bound on the min-entropy, given we find an appropriate *min-tradeoff* function.

We state the min-entropy bound in the following theorem.

Theorem 59. *Let* $\rho_{Y_A Y_B X_A X_B T E}$ *be the joint state of Alice, Bob and Eve's system along with the register* T *for indicating testing versus data rounds, and let* Ω *be the event that the protocol does not abort during parameter estimation. We write* $\rho_{|\Omega}$ *for the state of the system conditioned on* Ω*. Let* $\epsilon_{EA}, \epsilon_s \in (0,1)$*. Then either*

1. *The protocol aborts with probability greater than* $1 - \epsilon_{EA}$*, or*

2. $H_{\min}^{\epsilon_s}(Y_AY_B|X_AX_BTE)_{\rho_{|\Omega}} > n \cdot \mathsf{OPT}(\epsilon_s, \epsilon_{EA})$*, where* OPT *is defined as follows:*

$$g(p) = \begin{cases} 1 - h\left(3 - 4\frac{p(1)}{\gamma}\right) & \frac{p(1)}{\gamma} \in \left[\frac{2}{3}, \frac{3}{4}\right] \\ 1 & \frac{p(1)}{\gamma} \in \left[\frac{3}{4}, 1\right], \end{cases}$$

$$f_{\min}^{J_3}(p, p_t) = \begin{cases} g(p) & \text{if } p(1) \leq p_t(1) \\ \frac{d}{dp(1)} g(p)\Big|_{p_t} \cdot p(1) + g(p_t) - \frac{d}{dp(1)} g(p)\Big|_{p_t} \cdot p_t(1) & \text{if } p(1) > p_t(1), \end{cases}$$

$$f_{EAT} = n f_{\min}^{J_3}(p, p_t) - \frac{2}{\sqrt{n}} \left(\log 9 + \left\lceil \frac{d}{dp(1)} g(p) \right\rceil\right) \sqrt{1 - 2\log(\epsilon_s \cdot \epsilon_{EA})},$$

$$\mathsf{OPT}(\epsilon_s, \epsilon_{EA}) = \max_{\frac{2}{3} < \frac{p_t(1)}{\gamma} < \frac{3}{4}} f_{EAT}(p, p_t, \epsilon_s, \epsilon_{EA}).$$

Before we state the proof, we develop some key ideas and prove some lemmas that will be used in the proof. In round $i \in [n]$, Alice and Bob draw from a local biased distribution with $p_0, p_1, p_2 \in [0, 1]$:

$$x_i = \begin{cases} i \pmod 3 & \text{with probability } p_0, \\ i + 1 \pmod 3 & \text{with probability } p_1, \\ i + 2 \pmod 3 & \text{with probability } p_2. \end{cases} \tag{6.9}$$

Without loss of generality we may assume that the total number of rounds is a multiple of 3, i.e. $n = 3N$ for some N. There are two cases in which they perform a testing round – first for testing the violation of the Bell inequality J_3, and second to test the asynchronicity of the protocol. Let γ be the probability of performing a J_3 test. Thus

we have $\gamma = p(x_A \neq x_B)$.

$$\gamma = p(x_A \neq x_B) = \frac{1}{3}\sum_{i=0}^{2} p(x_A^i \neq x_B^i) = 2(p_0p_1 + p_0p_2 + p_1p_2).$$

For the J_3 test we define a random variable R_i as follows:

$$R_i = \begin{cases} 1 & \text{if } y_A^i \neq y_B^i \text{ and } x_A^i \neq x_B^i, \\ 0 & \text{if } y_A^i = y_B^i \text{ and } x_A^i \neq x_B^i, \\ \perp & \text{if } x_A^i = x_B^i. \end{cases}$$

The probability that $R_i = 1$ is given by

$$\begin{aligned} p(R_i = 1) = p(y_A^i \neq y_B^i \wedge x_A^i \neq x_B^i) &= \sum_{i}^{3N} \sum_{\substack{y_A^i \neq y_B^i \\ x_A^i \neq x_B^i}} p(y_A^i, y_B^i \mid x_A^i, x_B^i) \cdot p(x_A^i, x_B^i) \cdot \frac{1}{3N} \\ &= \frac{1}{3}\sum_{i=0}^{2} \sum_{\substack{y_A^i \neq y_B^i \\ x_A^i \neq x_B^i}} p(y_A^i, y_B^i \mid x_A^i, x_B^i) \cdot p(x_A^i, x_B^i) \\ &= \frac{1}{3}(p_0p_1 + p_0p_2 + p_1p_2) \sum_{\substack{y_A \neq y_B \\ x_A \neq x_B}} p(y_A, y_B \mid x_A, x_B) \\ &= \frac{1}{3}(p_0p_1 + p_0p_2 + p_1p_2)(4 - 4J_3) = \gamma\left(\frac{2}{3} - \frac{2}{3}J_3\right). \end{aligned}$$

Similarly, we define a random variable Q_i corresponding to the asyncronicity. We reserve every mth key generation round to perform an asynchronicity check i.e. if $i = 0 \pmod m$ for i such that $x_A^i = x_B^i$. We denote by $\kappa = 1/m$ the fraction of asynchronicity

check rounds. We have

$$Q_i = \begin{cases} 1 & \text{if } y_A^i \neq y_B^i \text{ and } x_A^i = x_B^i \text{ and } i = 0 \pmod m, \\ 0 & \text{if } y_A^i = y_B^i \text{ and } x_A^i = x_B^i, \\ \perp & \text{if } x_A^i \neq x_B^i. \end{cases}$$

The probability that $Q_i = 1$ is given by

$$\begin{aligned} p(Q_i = 1) &= p(y_A^i \neq y_B^i \wedge x_A^i = x_B^i \wedge i = 0 \pmod m) \\ &= \frac{1}{m} \sum_i^{3N} \sum_{x_A^i = x_B^i} \sum_{y_A^i \neq y_B^i} p(y_A^i, y_B^i \mid x_A^i, x_B^i) \cdot p(x_A^i, x_B^i) \cdot \frac{1}{3N} \\ &= \frac{\kappa}{3} \sum_{i=0}^{2} \sum_{\substack{y_A^i \neq y_B^i \\ x_A^i = x_B^i}} p(y_A^i, y_B^i \mid x_A^i, x_B^i) \cdot p(x_A^i, x_B^i) \\ &= \frac{\kappa}{3}(p_0^2 + p_1^2 + p_2^2) \sum_{\substack{y_A \neq y_B \\ x_A = x_B}} p(y_A, y_B \mid x_A, x_B) = \frac{\kappa}{3}(1-\gamma) \cdot 3S = \kappa(1-\gamma)S. \end{aligned}$$

Thus if $p(x_A \neq x_B) = \gamma$, then the probability that we are in a testing round (J_3 or S), i.e. $T_i = 1$ is given by $\gamma + \kappa(1-\gamma)$. We can tune γ arbitrarily by choosing p_0, p_1 and p_2 appropriately.

Before proving Theorem 59, we first show a bound on the mutual information between Alice's output and Eve's system. Following the outline in [80], we assume that Eve provides Alice and Bob a Bell diagonal state with eigenvalues $\lambda_{\Phi^+}, \lambda_{\Phi^-}, \lambda_{\Psi^+}, \lambda_{\Psi^-}$ corresponding to

the Bell states

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \qquad |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),$$
$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \qquad |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

We may write the Bell diagonal state as

$$\rho_\lambda = \begin{pmatrix} \lambda_{\Phi^+} & & & \\ & \lambda_{\Psi^-} & & \\ & & \lambda_{\Phi^-} & \\ & & & \lambda_{\Psi^-} \end{pmatrix} \tag{6.10}$$

The following lemma provides a bound on the mutual information between Alice's output and Eve's system. This bound is then used in the proof of the theorem in bounding the min-entropy of Alice and Bob's outputs in the protocol conditioned on Eve's side information.

Lemma 60. *Let Y_A^i be Alice's output in round $i \in [n]$, and E be Eve's register. If Eve provides Alice and Bob the Bell diagonal state ρ_λ in eq. (6.10), with eigenvalues ordered as $\lambda_{\Phi^+} \geq \lambda_{\Psi^-}$ and $\lambda_{\Phi^-} \geq \lambda_{\Psi^+}$, we have*

$$\chi(Y_A^i : E) \leq h(\lambda_{\Phi^-}).$$

Proof. Given a state $\rho \in \mathfrak{H}_A \otimes \mathfrak{H}_B$ and measurement operators $E_{y_A}^{x_A}$ and $F_{y_B}^{x_B}$ corresponding to Alice and Bob respectively, the probability of getting outputs (y_A, y_B) from inputs (x_A, x_B) is given by the Born rule, $p(y_A, y_B \,|\, x_A, x_B) = \mathrm{tr}\left(\left(E_{y_A}^{x_A} \otimes F_{y_B}^{x_B}\right)\rho\right)$. For the Bell

diagonal state ρ_λ, this probability may be expanded as follows:

$$
\begin{aligned}
p(y_A, y_B \mid x_A, x_B) &= \mathrm{tr}\left((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})\rho_\lambda\right) \\
&= \lambda_{\Phi^+}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})|\Phi^+\rangle\langle\Phi^+|\right) + \lambda_{\Phi^-}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})|\Phi^-\rangle\langle\Phi^-|\right) \\
&\quad + \lambda_{\Psi^+}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})|\Psi^+\rangle\langle\Psi^+|\right) + \lambda_{\Psi^-}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})|\Psi^-\rangle\langle\Psi^-|\right) \\
&= \lambda_{\Phi^+}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes F_{y_B}^{x_B})|\Phi^+\rangle\langle\Phi^+|\right) + \lambda_{\Phi^-}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes ZF_{y_B}^{x_B}Z)|\Phi^+\rangle\langle\Phi^+|\right) \\
&\quad + \lambda_{\Psi^+}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes XF_{y_B}^{x_B}X)|\Phi^+\rangle\langle\Phi^+|\right) + \lambda_{\Psi^-}\mathrm{tr}\left((E_{y_A}^{x_A} \otimes YF_{y_B}^{x_B}Y)|\Phi^+\rangle\langle\Phi^+|\right) \\
&= \frac{\lambda_{\Phi^+}}{2}\mathrm{tr}\left(E_{y_A}^{x_A}\overline{F_{y_A}^{x_A}}\right) + \frac{\lambda_{\Phi^-}}{2}\mathrm{tr}\left(E_{y_A}^{x_A}\overline{ZF_{y_A}^{x_A}Z}\right) \\
&\quad + \frac{\lambda_{\Psi^+}}{2}\mathrm{tr}\left(E_{y_A}^{x_A}\overline{XF_{y_A}^{x_A}X}\right) + \frac{\lambda_{\Psi^-}}{2}\mathrm{tr}\left(E_{y_A}^{x_A}\overline{YF_{y_A}^{x_A}Y}\right)
\end{aligned}
$$

Using this probability, we can compute the values of J_3 and S. One can show that choosing $E_{y_A}^{x_A} = \overline{F_{y_B}^{x_B}}$ is the optimal choice for minimizing J_3 and S simultaneously, but we skip the proof here. We define projection operators using variables $\theta_1, \theta_2, \gamma_1$ and γ_2 which we later optimize:

$$
\begin{aligned}
E_0^0 &= |\phi_0\rangle\langle\phi_0| \quad \text{with } |\phi_0\rangle = |0\rangle \\
E_0^1 &= |\phi_1\rangle\langle\phi_1| \quad \text{with } |\phi_1\rangle = \cos\theta_1|0\rangle + e^{i\gamma_1}\sin\theta_1|1\rangle \\
E_0^2 &= |\phi_2\rangle\langle\phi_2| \quad \text{with } |\phi_2\rangle = \cos\theta_2|0\rangle + e^{i\gamma_2}\sin\theta_2|1\rangle
\end{aligned}
$$

and where the corresponding $E_1^x = \mathbb{1} - E_0^x$ for $x \in \{0, 1, 2\}$. Computing the asynchronicity

S directly according to eq. (6.7) we get

$$S = \frac{\lambda_{\Phi^-}}{3}\left[\sin^2(2\theta_1) + \sin^2(2\theta_2)\right] + \frac{\lambda_{\Psi^+}}{3}\left[3 - (\sin^2(2\theta_1) + \sin^2(2\theta_2)\right] + \lambda_{\Psi^-}$$

The λ_{Ψ^-} term doesn't depend on θ_1 and θ_2, so we may take $\lambda_{\Psi^-} = 0$ since we want to minimize S. Further, since $\sin^2(2\theta_1) + \sin^2(2\theta_2) \geq 0$ and $\lambda_{\Phi^-} \geq \lambda_{\Psi^+}$, we may take $\lambda_{\Psi^+} = 0$. Next we define δ_1 and δ_2 to be the deviation in angles from the angles in the optimal strategy defined in eq. (6.1) (the optimal angles are given by $\theta_1 = \frac{\pi}{3}$ and $\theta_2 = -\frac{\pi}{3}$). Thus the equations we obtain for J_3 and S using $\theta_1 = \frac{\pi}{3} + \delta_1$ and $\theta_2 = -\frac{\pi}{3} + \delta_2$ are:

$$\begin{aligned} J_3 = &-(2\lambda_{\Phi^-} - 1)\cos\left(\frac{\pi}{3} + \delta_1\right)\cos\left(-\frac{\pi}{3} + \delta_2\right)\sin\left(\frac{\pi}{3} + \delta_1\right)\sin\left(-\frac{\pi}{3} + \delta_2\right) \\ &+ \cos^2\left(\frac{\pi}{3} + \delta_1\right)\cos^2\left(-\frac{\pi}{3} + \delta_2\right) \end{aligned}$$

Since we want to minimize J_3, we minimize the term independent of the factor λ_{Φ^-}. We call this term c_{J_3} and find that this term is

$$\begin{aligned} c_{J_3} &= \cos\left(\frac{\pi}{3} + \delta_1\right)\cos\left(-\frac{\pi}{3} + \delta_2\right)\sin\left(\frac{\pi}{3} + \delta_1\right)\sin\left(-\frac{\pi}{3} + \delta_2\right) \\ &\quad + \cos^2\left(\frac{\pi}{3} + \delta_1\right)\cos^2\left(-\frac{\pi}{3} + \delta_2\right) \\ &= \cos\left(\frac{2\pi}{3} + \delta_1 - \delta_2\right)\cos\left(\frac{\pi}{3} + \delta_1\right)\cos\left(\frac{\pi}{3} - \delta_2\right) \end{aligned}$$

Minimizing c_{J_3} for δ_1 and δ_2 we find that $\delta_1 = \frac{\delta_2}{2}$, and $\delta_2 \in \{0, \frac{2\pi}{3}, \frac{4\pi}{3}\}$. The solutions $\delta_2 = \frac{2\pi}{3}$ and $\delta_2 = \frac{4\pi}{3}$ are equivalent to $\delta_2 = 0$, so we only consider the latter solution. This suggests that in order for Eve to minimize J_3, her strategy must match the ideal strategy

developed in eq. (6.1). Using $\delta_1 = \delta_2 = 0$, we get

$$\begin{aligned} J_3 &= -\frac{1}{8} + \frac{3}{8}\lambda_{\Phi^-} \\ S &= \frac{1}{2}\lambda_{\Phi^-} \end{aligned} \tag{6.11}$$

From [80, Lemma 5], we have

$$\begin{aligned} \chi(Y_A^i : E) &\leq H([\lambda_{\Phi^+}, \lambda_{\Phi^-}, \lambda_{\Psi^+}, \lambda_{\Psi^-}]) - h(\lambda_{\Phi^+} + \lambda_{\Phi^-}) \\ &= h(\lambda_{\Phi^-}) = \begin{cases} h(\frac{1}{3} + \frac{8}{3}J_3) \\ \\ h(2S) \end{cases} \end{aligned}$$

Where the second to last equality follows because $\lambda_{\Psi^+} = \lambda_{\Psi^-} = 0$, thus $H([\lambda_{\Phi^+}, \lambda_{\Phi^-}]) = h(\lambda_{\Phi^-})$, and $h(\lambda_{\Phi^+} + \lambda_{\Phi^-}) = h(1) = 0$ □

Proof of Theorem 59. In similar fashion to [32, Theorem 4.1], we need to find a min-tradeoff function in order to apply the EAT. From Lemma 60, we have $\chi(Y_A^i : E|X_A^i = 0) \leq h\left(\frac{1}{3} + \frac{8}{3}J_3\right)$. Thus

$$H(Y_A^i|X_A^i X_B^i E) \geq 1 - h\left(\frac{1}{3} + \frac{8}{3}J_3\right) \tag{6.12}$$

Inserting this back into eq. (6.12), we get

$$H(Y_A^i|X_A^i X_B^i E) \geq 1 - h\left(\frac{1}{3} + \frac{8}{3}\left(1 - \frac{3}{2}\frac{p(R_i = 1)}{\gamma}\right)\right) = 1 - h\left(3 - 4\frac{p(R_i = 1)}{\gamma}\right)$$

For $\frac{p(1)}{\gamma} \in \left[\frac{2}{3}, 1\right]$, let

$$g(p) = \begin{cases} 1 - h\left(3 - 4\frac{p(1)}{\gamma}\right) & \frac{p(1)}{\gamma} \in \left[\frac{2}{3}, \frac{3}{4}\right] \\ 1 & \frac{p(1)}{\gamma} \in \left[\frac{3}{4}, 1\right] \end{cases}$$

We note that we only define $g(p)$ in the regime $\frac{p(1)}{\gamma} \in \left[\frac{2}{3}, 1\right]$ since that range is operationally relevant. The function can be extended to values of $\frac{p(1)}{\gamma} \in \left[0, \frac{2}{3}\right]$ for completeness but is not necessary for the purposes of the proof. The function $g(p)$ has unbounded gradient at $\frac{p(1)}{\gamma} = \frac{3}{4}$, and therefore needs to be modified using the 'cutting-and-gluing' trick of [32] in order to define a min-tradeoff function that can be used in the EAT. To that effect, we define two functions l_1 and l_2 over a point p_t that can be later optimized:

$$l_1(p_t) = \left\lceil \frac{d}{dp(1)} g(p) \Big|_{p_t} \right\rceil, \qquad l_2(p_t) = g(p_t) - l_1(p_t) \cdot p_t(1)$$

and define $f_{\min}^{J_3}$ as follows:

$$f_{\min}^{J_3}(p, p_t) = \begin{cases} g(p) & \text{if } p(1) \leq p_t(1) \\ l_1(p_t) \cdot p(1) + l_2(p_t) & \text{if } p(1) > p_t(1) \end{cases}$$

Applying the EAT with min-tradeoff function $f_{\min}^{J_3}(p, p_t)$ for any p_t such that $\frac{2}{3} < \frac{p_t(1)}{\gamma} < \frac{3}{4}$, and plugging in $\frac{p(1)}{\gamma} = \frac{p(R_i=1)}{\gamma} = \frac{2}{3} - \frac{2}{3}J_3$, we get the bound on the smooth min-entropy $H_{\min}^{\epsilon_s}(Y_A Y_B \mid X_A X_B E)_{\rho_{|\Omega}}$ □

The soundness proof for the protocol follows identically to [32, Lemmas 5.3 and 5.4].

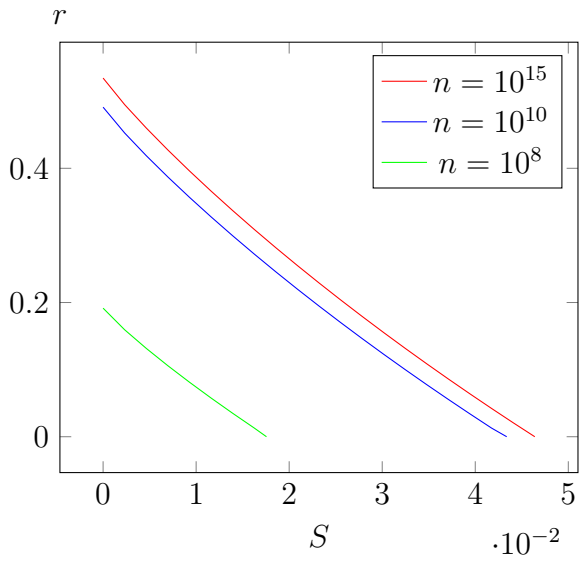


Figure 6.1: Values of $r = l/n$ against S

The key length l generated at the end of Protocol 4 is derived analogously to [32, Theorem 5.1 and Eq 5.4] which for completeness we state here:

$$l = n \cdot \mathsf{OPT}(\epsilon_s/4, \epsilon_{EA} + \epsilon_{EC}) - \mathsf{leak}_{\mathsf{EC}} - 3\log\left(1 - \sqrt{1 - (\epsilon_s/4)^2)}\right)$$
$$- \gamma \cdot n - \sqrt{n} 2\log 7\sqrt{1 - 2\log(\epsilon_s/4 \cdot (\epsilon_{EA} + \epsilon_{EC}))} - 2\log(1/\epsilon_{PA}) \tag{6.13}$$

where $\mathsf{leak}_{\mathsf{EC}}$ is discussed in detail in [32, §5.5.1 and Eq 5.9].

Based on Theorem 59 and [32, Theorem 5.1], we plot the key rate, defined as $r = \frac{l}{n}$. In Figure 6.1, we plot the key rate against the asynchronicity (referred to as the bit-error rate in [32]), and in Figure 6.2 we plot the key-rate against the total number of rounds while keeping asynchronicity constant. For large n, we are able to tolerate asynchronicity of up to 4.6% before the key-rate goes to 0. We use the values $\epsilon_{EC} = 10^{-10}, \epsilon_{EA} = \epsilon^s_{QKD} = 10^{-5}, \epsilon^c_{QKD} = 10^{-2}, p_0 = 0.97, p_1 = p_2 = 0.015$ and $\delta^{J_3}_{est} = 10^{-3}$ to plot the key rate curves in Figures 6.1 and 6.2

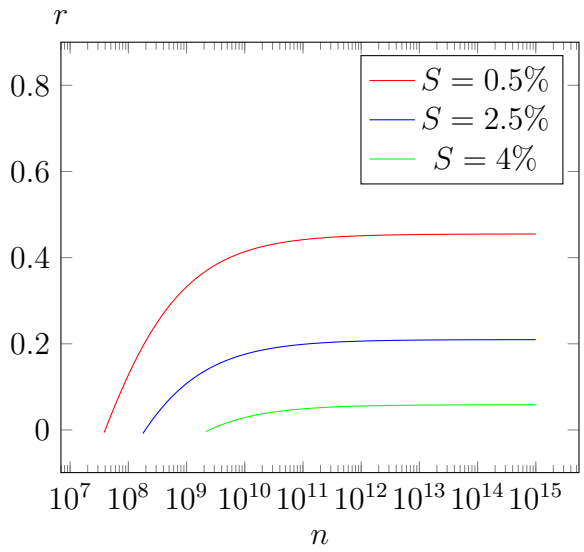


Figure 6.2: Values of $r = l/n$ against n

6.6 Causality Loophole

In this section we describe what is called the *causality* or *locality* loophole common to device independent quantum key distribution protocols that use non-local games, and propose a solution to the loophole using a new security assumption.

As seen in Section 6.4, the bound for the Bell inequality $J_3 \geq -\frac{1}{8}$ is sharp and rigid only among synchronous quantum correlations. There exist more powerful synchronous nonsignaling strategies that violate those bounds. Furthermore, if classical communication is allowed between the parties in the protocol, even greater violations can be achieved. This is the *causality loophole*: unless Alice and Bob are acausally separated, then the statistics for the synchronous Bell inequalities can simply be simulated using classical communication.

Example 61. *In order to simulate the correlation in eq. (6.3) that achieves the maximal violation $J_3 = -\frac{1}{8}$ among synchronous quantum correlations, Eve does the following:*

1. *She records Alice's input x_A and delivers a uniformly random bit $y_A = y \in \{0, 1\}$ to*

Alice;

2. *She checks if Bob's input x_B is equal to x_A,*

   - *if $x_B = x_A$ then she delivers the same output bit $y_B = y = y_A$ to Bob, or*
   - *if $x_B \neq x_A$ then she picks $y_B = y$ with probability $\frac{1}{4}$, and $y_B = 1 - y$ with probability $\frac{3}{4}$, and delivers this output bit to Bob.*

Note that Eve needs to communicate (x_A, y_A) to "Bob's side" of the protocol to compute his output.

In order to resolve the causality loophole in our protocol we pose a new security assumption: instead of limiting Eve's computational power or limiting the communication she can perform, we assume that she has imperfect knowledge of the basis Alice and Bob use in the protocol. We state this more formally as follows.

Let ϵ be Eve's uncertainty about Alice and Bob's inputs. Without loss of generality, we assume that this is symmetric across all basis selections. For $x', x \in \{0, 1, 2\}$ we have

$$\Pr\{\text{Eve guesses basis } x' \mid \text{Alice (or Bob) selects basis } x\} = \begin{cases} 1 - \epsilon & \text{when } x' = x \\ \\ \frac{\epsilon}{2} & \text{when } x' \neq x. \end{cases} \tag{6.14}$$

In greater generality, we model the basis selection that Alice and Bob use for their inputs as a classical-quantum state on $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathfrak{H}_E$, corresponding to Alice, Bob, and Eve respectively. Alice and Bob's states are classical while Eve can have quantum side infor-

mation which she may use to produce a correlation for a cheating strategy. We denote this state by ρ_{ABE}. For inputs $x_A, x_B \in \{0,1,2\}$ for Alice and Bob respectively, we have $\rho_{ABE} = |x_A\rangle\langle x_A| \otimes |x_B\rangle\langle x_B| \otimes \rho_E^{x_A,x_B}$, where $\rho_E^{x_A,x_B}$ quantifies Eve's side information. Based on (6.14) above we further decompose

$$\rho_E^{x_A,x_B} = \Big((1-\epsilon)^2\sigma_{x_A,x_B} + (1-\epsilon)\frac{\epsilon}{2}(\sigma_{x_A,x_B\oplus 1} + \sigma_{x_A,x_B\oplus 2} + \sigma_{x_A\oplus 1,x_B} + \sigma_{x_A\oplus 2,x_B}) + \frac{\epsilon^2}{4}(\sigma_{x_A\oplus 1,x_B\oplus 1} + \sigma_{x_A\oplus 1,x_B\oplus 2} + \sigma_{x_A\oplus 2,x_B\oplus 1} + \sigma_{x_A\oplus 2,x_B\oplus 2})\Big),$$

where we denote $x_A \oplus i \coloneqq x_A + i \pmod 3$, and similarly for x_B. Writing Eve's guess for Alice's input by z_A and for Bob's input by z_B, the σ_{z_A,z_B} for $z_A, z_B \in \{0,1,2\}$ are densities containing Eve's side information depending on her guess for x_A and x_B respectively. With these, we also allow Eve to have unlimited computational power and communication to produce outputs (y_A, y_B). We denote the resulting conditional probability distribution as $\Pr\{(y_A, y_B \mid z_A, z_B)\}_{\sigma_{z_A,z_B}}$. As this is also a correlation, Eve has her own Bell term which we denote by $\tilde{J}_3$ and her own asynchronicity term which we denote by $\tilde{S}$.

Eve's goal is to program Alice and Bob's devices such that the device outputs pass statistical tests for estimating Bell violation and asynchronicity. The following theorem shows that Eve's uncertainty ϵ is upper-bounded by a function of the allowed errors in Alice and Bob's Bell and asynchronicity terms. If Eve's uncertainty exceeds a certain threshold then there does not exist a distribution $\Pr\{(y_A, y_B \mid z_A, z_B)\}_{\sigma_{z_A,z_B}}$ she can use to provide outputs to Alice and Bob that still pass their Bell and asynchronicity checks. We state the theorem formally as follows.

Theorem 62. *Let $0 \leq \lambda < \frac{1}{8}$ be the allowed error in Alice and Bob's J_3 term, and $0 \leq \mu$*

be their asynchronicity bound. On Eve's side, let $\tilde{J}_3$ and $\tilde{S}$ be analogous Bell inequality and asynchronicity terms for her correlation. Let ϵ be Eve's uncertainty about Alice and Bob's inputs as given in eq. (6.14), and δ be such that $0 \leq \delta$. If $\epsilon > \epsilon^{\delta}_{max}$, where

$$\epsilon^{\delta}_{max} = \frac{2}{3} - \frac{2}{3}\left(\frac{\sqrt{144(\delta-1)\lambda + 64\lambda^2 + 6(36\delta + 8\lambda - 9)\mu - 72\mu^2 - 162\delta + 81}}{6\mu - 18\delta - 8\lambda + 9}\right),$$

then Eve's asynchronicity $\tilde{S} < \delta$.

Proof. For inputs $x_A, x_B \in \{0, 1, 2\}$, the correlation that Alice and Bob use to compute key bits and self-test their devices is then given by:

$$p(y_A, y_B \mid x_A, x_B) = \tag{6.15}$$
$$\sum_{z_A, z_B} \Pr\{y_A, y_B \mid z_A, z_B\}_{\sigma_{z_A, z_B}} \cdot \left\{ \begin{array}{ll} 1-\epsilon & \text{for } z_A = x_A \\ \\ \frac{\epsilon}{2} & \text{otherwise} \end{array} \right\} \cdot \left\{ \begin{array}{ll} 1-\epsilon & \text{for } z_B = x_B \\ \\ \frac{\epsilon}{2} & \text{otherwise} \end{array} \right\}$$

We begin by deriving expressions for the expected values of J_3 and S.

$$\begin{aligned} \langle 1 - J_3 \rangle = \frac{1}{4}\big(& p(0,1 \mid 0,1) + p(1,0 \mid 0,1) + p(0,1 \mid 1,0) + p(1,0 \mid 1,0) \\ & + \; p(0,1 \mid 0,2) + p(1,0 \mid 0,2) + p(0,1 \mid 2,0) + p(1,0 \mid 2,0) \\ & + \; p(0,1 \mid 1,2) + p(1,0 \mid 1,2) + p(0,1 \mid 2,1) + p(1,0 \mid 2,1)\big) \\ = & \left(1 - \epsilon + \tfrac{3}{4}\epsilon^2\right)(1 - \tilde{J}_3) + \left(\tfrac{3}{2}\epsilon - \tfrac{9}{8}\epsilon^2\right)\tilde{S} \end{aligned} \tag{6.16}$$

A similar computation for S gives us:

$$\begin{aligned}\langle S\rangle &= \frac{1}{3}\big(p(0,1\mid 0,0)+p(1,0\mid 0,0)+p(0,1\mid 1,1)\\ &\qquad +p(1,0\mid 1,1)+p(0,1\mid 2,2)+p(1,0\mid 2,2)\big)\\ &= \left(1-2\epsilon+\tfrac{3}{2}\epsilon^2\right)\tilde{S}+\left(\tfrac{4}{3}\epsilon-\epsilon^2\right)(1-\tilde{J}_3) \qquad (6.17)\end{aligned}$$

Using eqs. (6.16) and (6.17), we can solve for $\tilde{J}_3$ and $\tilde{S}$ as:

$$\begin{bmatrix} 1-\tilde{J}_3 \\ \tilde{S}\end{bmatrix} = \begin{bmatrix} 1-\epsilon+\frac{3}{4}\epsilon^2 & \frac{3}{2}\epsilon-\frac{9}{8}\epsilon^2 \\ \frac{4}{3}\epsilon-\epsilon^2 & 1-2\epsilon+\frac{3}{2}\epsilon^2\end{bmatrix}^{-1}\begin{bmatrix}\frac{9}{8}-\lambda \\ \mu\end{bmatrix}$$

We get solutions:

$$\tilde{J}_3=\frac{(3\epsilon^2-4\epsilon)(3-6\mu+8\lambda)+16\lambda-2}{4\,(3\epsilon-2)^2} \qquad \tilde{S}=\frac{(3\epsilon^2-4\epsilon)(6\mu-8\lambda+9)+24\mu}{6\,(3\epsilon-2)^2}. \qquad (6.18)$$

Plugging $\tilde{S}=\delta$ in eq. (6.18), and solving for ϵ gives us:

$$\epsilon^{\delta}_{max}=\frac{2}{3}-\frac{2}{3}\left(\frac{\sqrt{144(\delta-1)\lambda+64\lambda^2+6(36\delta+8\lambda-9)\mu-72\mu^2-162\delta+81}}{6\mu-18\delta-8\lambda+9}\right) \qquad (6.19)$$

□

Corollary 63. *For $\epsilon > \epsilon^0_{max}$, there is no correlation Eve can use to produce a cheating strategy against Alice and Bob.*

Proof. Plugging in $\delta = 0$ in eq. (6.19),

$$\epsilon_{max} \coloneqq \epsilon^0_{max} = \frac{2}{3} - \frac{2}{3}\left(\frac{\sqrt{64\lambda^2 + 6(8\lambda - 9)\mu - 72\mu^2 - 144\lambda + 81}}{6\mu - 8\lambda + 9}\right).$$

If Eve's uncertainty $\epsilon > \epsilon_{max}$, then $\tilde{S} < 0$, and since no correlation can have negative asynchronicity, no such $\Pr\{(y_A, y_B \mid z_A, z_B)\}_{\sigma_{z_A,z_B}}$ exists. □

By the corollary above, we conclude that Eve's uncertainty cannot grow too much before her asynchronicity becomes negative, therefore resulting in an infeasible strategy. Fixing a reasonable threshold for the error allowed in the Bell term, say $\lambda = 0.05$, we plot values of ϵ_{max} against varying values of Alice and Bob's allowed asynchronicity μ in Figure 6.3. The plot shows that even for allowed asynchronicity $\mu = 5\%$, Eve must have close to perfect certainty $\approx 97\%$ about Alice and Bob's inputs. Thus even with unlimited computational and communication power, when $\epsilon > \epsilon_{max}$, no correlation exists to perfectly simulate statistics that pass Alice and Bob's Bell and asynchronicity checks.

We further examine the regime where Eve's uncertainty $\epsilon > \epsilon_{max}$. In this case the best Eve can do in order to provide Alice and Bob an expected asynchronicity value $\langle S \rangle$ close to μ, is to use a synchronous correlation herself, i.e. $\tilde{S} = 0$. Fixing $\tilde{S} = 0$, we plot $\langle J_3 \rangle$ as Eve's uncertainty exceeds ϵ_{max}. Let $\gamma_\epsilon \coloneqq \epsilon - \epsilon_{max}$ denote how much Eve's uncertainty is above the maximum. Figure 6.4 shows that even with a lot of uncertainty, Eve can make $\langle J_3 \rangle$ as close to $-\frac{1}{8}$ as she likes. Since Eve is not restricted to quantum strategies, she can in fact violate the $-\frac{1}{8}$ bound. However, providing a $\langle J_3 \rangle$ value smaller than $-\frac{1}{8}$ is not in her best interest since Alice and Bob check if their estimated J_3 is in $[-\frac{1}{8}, -\frac{1}{8} + \lambda]$.

As a result, detecting Eve's interference depends *only* on the asynchronicity check.

Since Eve's $\tilde{S} = 0$, she has to provide a value for Alice and Bob's $\langle S \rangle = \tilde{\mu}$ that is strictly larger than their decided error threshold μ. We use eq. (6.17) to plot the effect of increasing ϵ past ϵ_{max} on $\langle S \rangle = \tilde{\mu}$ for a fixed λ and μ. Figure 6.5 shows the comparison between γ_ϵ and $\tilde{\mu}$ for $\mu = 0.05$ and $\lambda = 0.05$. In our analysis the choice of 0.05 for both λ and μ is arbitrary, and is made to demonstrate the effect of increasing Eve's uncertainty ϵ on the expected value $\langle S \rangle$. Alice and Bob may pick any reasonable error values for their J_3 and S terms without affecting the following calculations. From Figure 6.5, we see that $\tilde{\mu}$ increases sharply as γ_ϵ increases, which in turn implies that Alice and Bob's asynchronicity test always fails except with negligible probability. We show this using a straightforward Chernoff argument and bounding the probability that Alice and Bob's output is asynchronous in fewer than a μ fraction of the asynchronicity check rounds. Formally, let's assume Alice and Bob have m asynchronicity check rounds. Let A_i be a $\{0, 1\}$ random variable denoting whether their output is asynchronous in round $i \in [m]$. Since Eve provides an asynchronous output with probability $\tilde{\mu}$, we have

$$A_i = \begin{cases} 1 & \text{with probability } \tilde{\mu}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $A_S = \sum_i A_i$. Therefore $\langle A_S \rangle = \sum_i \langle A_i \rangle = m\tilde{\mu}$. Using a Chernoff bound we get

$$\Pr(A_S \leq m\mu) \leq \exp\left(-\frac{(\tilde{\mu} - \mu)^2 k}{2\tilde{\mu}}\right).$$

Alice and Bob can thus make this probability arbitrarily small by picking an appropriate value m for the number of asynchronicity check rounds they perform.

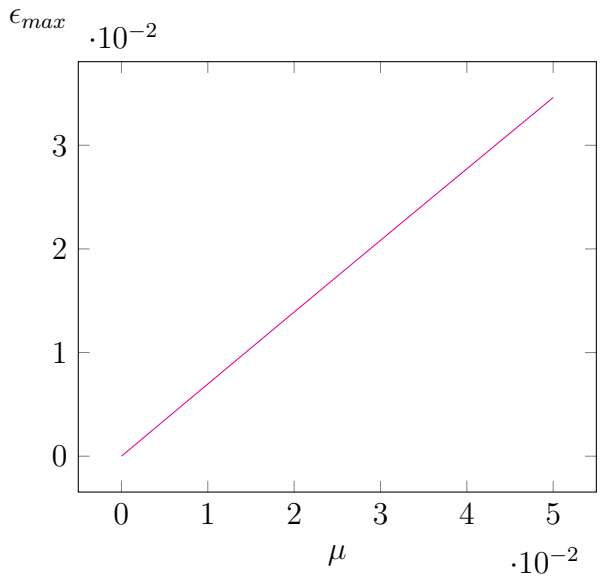


Figure 6.3: Values of μ vs. ϵ_{max} for which $\tilde{S}$ is non-negative

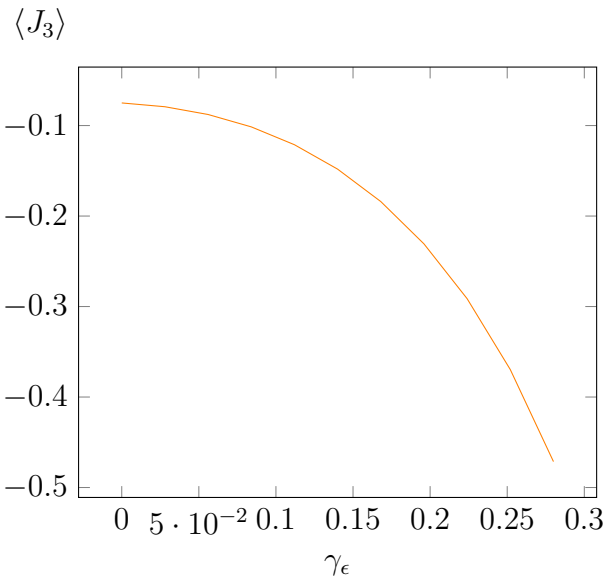


Figure 6.4: Values of $\langle J_3 \rangle$ vs. γ_ϵ for $\mu = 0.05, \lambda = 0.05$

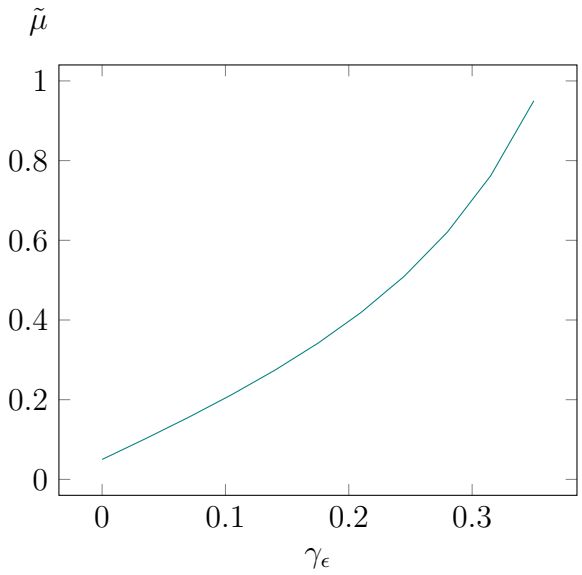


Figure 6.5: Values of $\langle S \rangle = \tilde{\mu}$ vs. γ_ϵ for $\mu = 0.05, \lambda = 0.05$

Chapter 7: Conclusion and Future Directions

7.1 Conclusions

In this dissertation we focused on three main topics. The first topic introduced novel primitives for ideal and real random bit generators in the Constructive Cryptography framework. We showed examples of quantum random bit generators based on process tomography and analyzed their efficiency in the constructive framework. The second topic focused on developing a fully device-independent quantum key distribution protocol that is based on the notion of synchronous correlations. This protocol had a few loopholes which we studied carefully and showed how to exploit. We also closed those same loopholes by introducing certain checks and assumptions that Alice and Bob can verify in order to guarantee that their protocol remains secure. Finally in the third topic we proposed a preliminary version of a quantum algorithm for lattice sieving. This algorithm continues to be a work in progress and has the potential to significantly reduce the memory complexity and runtime of lattice sieving algorithms, if the analysis carries out.

7.2 Future directions

We mention some directions and ideas for future research.

1. We analyzed random bit generators based on process tomography of one-qubit channels. This was the right step towards showing the efficiency of those random bit generators in the Constructive Cryptography framework. However in order to show rigorous security, one might want to consider device independent quantum random bit generators. In our work we closed the causality loophole by introducing a new security assumption. The next step would be to show a constructive proof that the composed protocol where our DIQKD protocol acts as a random bit generator to another instance of a DIQKD protocol remains secure even if the condition of nonlocality is violated, as long as the adversary has some uncertainty about the two parties' inputs.

2. Can we study other basic primitives in the Constructive Cryptography framework? One example is a *self-test*. What does an *ideal* self-test look like versus a *real* one?

3. Our protocol is based on synchronous correlations with three inputs and two outputs. An interesting direction would be to explore synchronous correlations with different number of inputs and outputs. Can we say anything about the synchronous Bell inequalities that occur in those cases?

4. Removing the assumption we made in the quantum lattice sieving algorithm is crucial to showing its claimed memory and runtime complexity. If the assumption holds, the result will have a major impact on the security parameters that organizations like NIST will set for the standardized post-quantum cryptographic algorithms.

Appendix A: Quantum Lattice Sieving

In this chapter, we provide a writeup for some ongoing work on developing quantum lattice sieving algorithms. This work was previously uploaded as a preprint, however we received feedback from the post-quantum cryptography community about an oversight in the amplitude amplification step. While we are hopeful that there is a work-around, we introduce an assumption in the meanwhile (given as Assumption 69) in order to present the rest of our results. We note that the assumption only impacts the claimed runtime of the algorithms. If the assumption is false, the algorithms will remain correct but will no longer have the claimed runtime.

A.1 Introduction

Lattices have become attractive tools in the design of post-quantum cryptographic protocols. This is evident from NIST's recent Post Quantum Cryptography Standardization competition, where three out of the four Round 3 finalists in the Public Key Encryption category, and two out of the three for Digital Signatures are lattice based. Certain lattice problems like that of finding the shortest vector in a given lattice are believed to be hard to solve in the worst case, even using quantum computers. As a result, cryptanalysis of existing lattice-based cryptosystems is necessary as a means of understanding the complex-

ity of solving hard lattice problems, and to set relevant security parameters used in these cryptosystems.

In this work we focus on the shortest vector problem (SVP), which is defined as follows:

Definition 64 (Shortest Vector Problem - SVP)**.** *Given a lattice $\mathcal{L}$, find a non-zero $\mathbf{s} \in \mathcal{L}$ such that $|\mathbf{s}| = \lambda_1(\mathcal{L})$.*

There are two main techniques for solving the SVP, *sieving* and *enumeration*. *Sieving* refers to the class of algorithms that process a list of lattice vectors at each sieve step and produce shorter vectors for the next sieve step. Sieving gives exponential time algorithms for solving the exact Shortest Vector Problem. This is in contrast to *enumeration* which gives super-exponential algorithms for SVP. We do not discuss enumeration techniques, but provide references for the interested reader [81, 82, 83, 84, 85, 86].

We provide a brief discussion on sieving algorithms starting with the algorithm by Micciancio and Voulgaris in [87]. The algorithm starts with an empty list, and adds new vectors to it at each sieve iteration. At each step the algorithm reduces the new vector with the existing vectors in the list. As the sieve proceeds, the list contains smaller and smaller vectors, and eventually outputs the shortest one. By reducing the new vector with the other vectors in the list, they were able to prove a bound on the size of the list of vectors, which in turn, determines the runtime $\tilde{\mathcal{O}}(2^{3.199d})$ and memory complexity $\tilde{\mathcal{O}}(2^{1.325d})$ of the algorithm. Notice that the constant in the exponent is still very high for this algorithm to be practical for lattices of large dimension. A more recent algorithm by Aggarwal et al. uses Discrete Gaussian Sampling to solve SVP in time and space $2^{d+o(d)}$ [88].

In recent years, there has been work on developing practical sieving algorithms that make heuristic assumptions, and show bounds on time and space required for the algorithm based on the heuristic assumption. Nguyen and Vidick gave a heuristic algorithm for solving SVP in [89] that runs in time $\mathcal{O}(2^{0.415d})$ and space $\mathcal{O}(2^{0.2075d})$. This practical variant makes the heuristic assumption that at each level the sieve vectors lie on the surface of the unit sphere, and are distributed uniformly and independently on it. This assumption, which we state in Heuristic (65), helps bound the total number of vectors needed in the initial step of the sieve. The NV-Sieve starts with a large initial set of vectors from the lattice. At each step of the sieve, the algorithm iterates over every vector in this set and checks to see if the vector has norm less than or equal to a specified quantity. If it does, then the vector is promoted to the next step of the sieve. If the norm is greater than the specified quantity, the algorithm searches in a list of *centers* to see if a *center* reduces with the current vector. If it does, the reduced vector is promoted to the next step of the sieve. If no center is found that reduces with the current vector, the current vector is added to the list of *centers* to use for reducing against future vectors. We can see how the structure of this algorithm resembles a real world sieve. At every sieve iteration, only shorter vectors go to the next level, and therefore at the end of the sieve we are left with very short vectors.

Several other heuristic algorithms have been proposed, including the GaussSieve from [87] which is based on ListSieve but without any theoretical guarantees on runtime. Laarhoven proposed algorithms based on locality sensitive hashing in [90] and [91]. The current best known classical algorithm for SVP is due to [92] and is based on solving a related problem known as Bounded Distance Decoding. This algorithm runs in time $2^{1.741d+o(d)}$ and space $2^{0.5d+o(d)}$. Concurrently, there has been work on developing quantum speedups

for these classical algorithms, usually by replacing the 'search' step in these algorithms with quantum amplitude amplification that provides a quadratic speedup.

A.1.1 Algorithm overview

We give an overview of our quantum sieving algorithm that is based on the classical Double Sieve algorithm due to Bai et al. [93]. The Double Sieve is not the best known classical algorithm, however, it is very simple to describe. This algorithm lends itself nicely to a quantum sieving algorithm, where our idea is to start with a superposition over lattice vectors, and search for shorter vectors using amplitude amplification at each step.

First we present the classical Double Sieve along with the heuristic assumptions made in the original paper, and then show how our quantum algorithm works. This algorithm uses memory of the order $\mathcal{O}(2^{0.2075d})$ and has time complexity $\mathcal{O}(2^{0.415d})$. Technical details of the algorithm can be found in the original paper, but we restate the heuristic and show some computations to derive bounds on the number of vectors needed in the initial step of the sieve, as this analysis will carry over to the quantum case. The algorithm starts by sampling an initial set of vectors S from the given lattice $\mathcal{L}$. The heuristic assumption made about this sample of vectors is as follows:

Heuristic 65. *The vectors $\frac{\mathbf{v}}{|\mathbf{v}|}$ for $\mathbf{v} \in S$ are distributed independently and uniformly on the surface of the unit sphere.*

Using this heuristic, we can compute the number of lattice vectors needed in the initial set S. Given a fixed vector $\mathbf{v}$ on the surface of the unit sphere, the density of vectors on the sphere that make an angle of θ with $\mathbf{v}$ is given by approximately $\sin\theta$. Given two vectors of

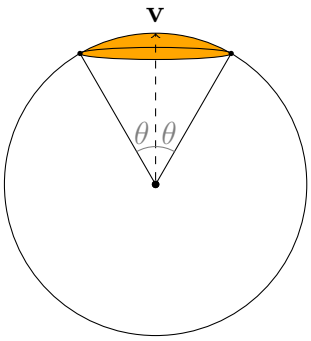


Figure A.1: The portion of the unit sphere that $\mathbf{v}$ covers is given by the volume of the spherical cap.

nearly equal length, $\mathbf{v}_1$ and $\mathbf{v}_2$, they reduce with each other, i.e. $|\mathbf{v}_2 - \mathbf{v}_1| \leq \min\{|\mathbf{v}_1|, |\mathbf{v}_2|\}$ only if the angle θ between them is less than $\frac{\pi}{3}$. Thus given a fixed vector on the unit sphere, the probability that another vector also on the sphere reduces with it, is given by $p = \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}$. We can think of this probability as the portion of the unit sphere that one vector covers, which is shown in Figure(A.1). In order for this probability to be close to 1, we need to cover the whole sphere, and so we need roughly $N \propto \frac{1}{p} = \left(\frac{4}{3}\right)^{\frac{d}{2}} = 2^{0.2075d}$ vectors. This gives us the number of vectors that we need to sample initially when the sieve starts.

The sieve algorithm proceeds by looking at the sum (and difference) of pairs of vectors and promotes the respective sum (or difference) which has norm less than the quantity γR to the next step of the sieve, where γ is usually set to $1 - \frac{1}{d}$, and R to the maximum norm of vectors at the current step of the sieve. Each step of the sieve uses the heuristic assumption. After a $poly(d)$ number of iterations of the sieve, we end up with sufficiently small vectors, which give a good approximation of the shortest vector. The algorithm is presented below.

The main loop iterates over pairs of vectors, and hence the runtime is quadratic in N,

Algorithm 5: Double Sieve [93]

Input:
S: The initial set of sampled vectors of size N
γ: Norm reduction factor
R: Bound on the maximum norm of vectors in S

Set $S' = \{\}$
for $\mathbf{u}, \mathbf{v} \in S$ **do**
  if $|\mathbf{u} \pm \mathbf{v}| \leq \gamma R$ **then**
    $S' = S' \cup \{\mathbf{u} \pm \mathbf{v}\}$
return S'

yielding a running time of $\mathcal{O}(2^{0.415d})$. The sieve starts with, and maintains a list of vectors of size N at each iteration, and so the memory is $\mathcal{O}(2^{0.2075d})$.

In the quantum sieve, we also sample $N = \mathcal{O}(2^{0.2075d})$ vectors, but we do that in superposition instead of storing a classical list. Assume that we have the ability to sample a uniform superposition over $\mathcal{O}(2^{0.2075d})$ lattice vectors. We tensor product two of these superpositions to give us a superposition over pairs of vectors in the lattice. At the initial stage, we fix R to be the maximum norm of the vectors in the lattice, as in the classical double sieve. Thus we have a superposition over pairs of vectors in the lattice whose norm is bounded by R.

$$\frac{1}{\sqrt{N}} \sum_{\substack{\mathbf{u}_1 \in S \\ |\mathbf{u}_1| \leq R}} |\mathbf{u}_1\rangle \otimes \frac{1}{\sqrt{N}} \sum_{\substack{\mathbf{u}_2 \in S \\ |\mathbf{u}_2| \leq R}} |\mathbf{u}_2\rangle = \frac{1}{N} \sum_{\mathbf{u}_1, \mathbf{u}_2 \in S} |\mathbf{u}_1, \mathbf{u}_2\rangle$$

This is analogous to iterating over pairs of vectors in the classical double sieve. Next we compute the norm of the difference $\mathbf{u}_2 - \mathbf{u}_1$ in an auxiliary register:

$$\frac{1}{N} \sum_{\mathbf{u}_1, \mathbf{u}_2 \in S} |\mathbf{u}_1, \mathbf{u}_2\rangle |0\rangle_{\text{aux}} \mapsto \frac{1}{N} \sum_{\mathbf{u}_1, \mathbf{u}_2 \in S} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle$$

We then perform amplitude amplification (A.A.) to find vectors such that $|\mathbf{u}_2 - \mathbf{u}_1| \leq \gamma R$. This is analogous to steps 3 and 4 in Algorithm 5 stated above. As a result of the amplitude amplification step, we now have a uniform superposition over vectors in the lattice whose norm is bounded by γR, with high probability.

$$\frac{1}{N} \sum_{\mathbf{u}_1, \mathbf{u}_2 \in S} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle \xmapsto{\text{A.A.}} \frac{1}{\sqrt{N}} \sum_{\substack{\mathbf{u}_1, \mathbf{u}_2 \in S \\ |\mathbf{u}_2 - \mathbf{u}_1| \leq \gamma R}} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle + |\phi_{\text{junk}}\rangle$$

To simplify notation, we only show the first two registers, measure to discard the junk state, compute $\mathbf{u}_2 - \mathbf{u}_1$ in the first register using a unitary U_{diff}, and rename $\mathbf{u}_2 - \mathbf{u}_1$ to $\mathbf{v}_1$:

$$\frac{1}{\sqrt{N}} \sum_{\substack{\mathbf{u}_1, \mathbf{u}_2 \in S \\ |\mathbf{u}_2 - \mathbf{u}_1| \leq \gamma R}} U_{\text{diff}} |\mathbf{u}_1, \mathbf{u}_2\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{\substack{\mathbf{u}_1, \mathbf{u}_2 \in S \\ |\mathbf{u}_2 - \mathbf{u}_1| \leq \gamma R}} |\mathbf{u}_2 - \mathbf{u}_1, \mathbf{u}_2\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{v}_1 = \mathbf{u}_2 - \mathbf{u}_1} |\mathbf{v}_1, \mathbf{u}_2\rangle$$

At this point we have successfully run one iteration of the sieve. We started with vectors with norm bounded by R, and ended up with vectors with norm bounded by γR, thus effectively reducing the norm of the vectors in our sieve by a factor of γ. In order to keep the sieve going, we need a second superposition over vectors with norm bounded by γR. We combine two such superpositions, run amplitude amplification, and obtain another superposition with vectors of norm bounded by $\gamma^2 R$. This process produces a binary tree of superpositions, such that to run iteration i of the sieve, we have to start with 2^i superpositions at the initial step. Notice that after each level i, we end up with a

superposition state that looks like

$$\frac{1}{\sqrt{N}} \sum_{|\mathbf{w}_1| \leq \gamma^i R} |\mathbf{w}_1, \text{history}_{2^i-1}\rangle \tag{A.1}$$

where $|\text{history}_{2^i-1}\rangle$ is a state on $2^i - 1$ registers which contains the history of the lattice vectors that produced the vector $\mathbf{w}_1$. In our presentation we do not explicitly write the history state, but we note that it exists, and is important to the memory and runtime analysis of the algorithm. We run $t = poly(d)$ iterations of the sieve, just as in the classical double sieve, which uses 2^t superpositions. At the end we are left with a uniform superposition over very small vectors in the lattice, and measuring the state, we get a good approximation to the shortest vector in the lattice. The runtime and memory analysis along with the full algorithm is presented in Section (A.3).

A.2 Preliminaries

A.2.1 Notation

The norm of a vector $\mathbf{v} = (v_1, \cdots, v_n)^\top$ is denoted $|\mathbf{v}|$, and we take this norm to be the Euclidean norm, $|\mathbf{v}| = \sqrt{\sum_i v_i^2}$. For a matrix $\mathbf{B}$, we denote its ith column by $\mathbf{b}_i$, and its norm by $|\mathbf{B}| = \max_i |\mathbf{b}_i|$. We denote by $B_n(\mathbf{v}, r)$ the ball of radius r around the vector $\mathbf{v}$. Thus $B_n(\mathbf{v}, r) = \{\mathbf{y} \in \mathbb{R}^n \mid |\mathbf{y} - \mathbf{v}| \leq r\}$. We simplify notation and write $B_n(\mathbf{0}, r) = B_n(r)$, and $B_n(\mathbf{0}, 1) = B_n$. We denote by $\mathcal{S}$ the d-dimensional unit sphere.

A.2.2 Lattices

Given d linearly independent vectors $\{\mathbf{b}_1, \cdots, \mathbf{b}_d\}$ with each $\mathbf{b}_i \in \mathbb{R}^n$, the *lattice* $\mathcal{L}$ generated by them is denoted by the set of all integer linear combinations of the *basis* vectors, and written as $\mathcal{L} = \{\sum_{i=1}^{d} x_i \mathbf{b}_i \mid x_i \in \mathbb{Z}\}$. We will restrict our discussion to lattices of full rank i.e. $n = d$, and denote using d the dimension of the lattice. Alternatively, the basis can be written as a matrix $\mathbf{B}$ with the vectors forming the columns of $\mathbf{B}$ which gives us $\mathcal{L} = \{\mathbf{Bx} \mid \mathbf{x} \in \mathbb{Z}^d\}$. The volume of the lattice, denoted $\text{vol}(\mathcal{L})$ is given by $|\det(\mathbf{B})|$, and is an invariant of the lattice. The norm of the shortest vector in the lattice is denoted $\lambda_1(\mathcal{L})$ and is known as the *first minimum* of the lattice. One of the central problems in the study of lattices is that of finding the shortest vector which we described in Section (A.1).

Several cryptographic primitives use other related average-case hard lattice problems like Small Integer Solutions (SIS) [94], Learning with Errors (LWE) [95], and their variants as the computational hardness assumption when defining the security of the cryptosystem. The SIS problem is stated as follows:

Definition 66 (Small Integer Solutions - SIS)**.** *Given n, m, q, ν with $\nu < q$, a matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, and lattice $\mathcal{L} = \{\mathbf{x} \in \mathbb{Z}^m : \mathbf{Ax} \equiv \mathbf{0} \mod q\}$, find $\mathbf{y} \in \mathcal{L}$ s.t. $|\mathbf{y}| \leq \nu$*

The Learning with Errors problem is stated as follows:

Definition 67 (Learning with Errors - LWE)**.** *Let $\mathbf{s} \in \mathbb{Z}_q^n$, for a given n and modulus q. Let χ be a probability distribution on $\mathbb{Z}_q$. We define a probability distribution $A_{\mathbf{s},\chi}$ with sampling as follows: draw $\mathbf{a} \in \mathbb{Z}_q^n$ at uniform, and $e \in \mathbb{Z}_q$ according to χ. Then return $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e) \mod q$. The problem is then defined as:*

Given n, q, χ and any number of independent samples from $A_{\mathbf{s},\chi}$, determine $\mathbf{s}$.

A.2.3 Gram-Schmidt orthogonalization

The Gram-Schmidt basis is an orthogonal set of vectors $\hat{\mathbf{B}} = \{\hat{\mathbf{b}}_1, \cdots, \hat{\mathbf{b}}_d\}$, where each $\hat{\mathbf{b}}_i$ is orthogonal to $span(\mathbf{b}_1, \cdots \mathbf{b}_{i-1})$. Given a basis $\mathbf{B}$, the Gram-Schmidt basis can be computed as follows: set $\hat{\mathbf{b}}_1 = \mathbf{b}_1$, and compute $\hat{\mathbf{b}}_i = \mathbf{b}_i - \sum_{j=1}^{i-1} \frac{\langle \hat{\mathbf{b}}_j, \mathbf{b}_i \rangle}{\langle \hat{\mathbf{b}}_j, \hat{\mathbf{b}}_j \rangle} \hat{\mathbf{b}}_j$, for $i = 2, \cdots, d$.

A.2.4 Quantum background

We denote quantum states as $\sum_x \alpha_x |x\rangle$, where $\alpha_x \in \mathbb{C}$, and $\sum_x |\alpha_x|^2 = 1$. We refer the reader to [37] or [38] for an introduction to quantum information processing. We present some details about the method of amplitude amplification that is relevant to the analysis of our algorithm.

Definition 68 (Amplitude Amplification [96, 97])**.** *Given a function f, consider the problem of searching for m marked elements in a set of N elements with $0 < m < N$, such that $f(x) = 1$ iff x is marked, and $f(x) = 0$ otherwise. Let $\mathcal{A}$ be a quantum algorithm that makes no measurements and produces the superposition $|\Psi\rangle = \mathcal{A}|0 \cdots 0\rangle = \sqrt{p_{good}}|\Psi_{good}\rangle + \sqrt{p_{bad}}|\Psi_{bad}\rangle$, where $|\Psi_{good}\rangle$ is a superposition over the m marked 'good' elements. Let U_f be the unitary that flips the phase of the good states, i.e. $U_f|\Psi_{good}\rangle = -|\Psi_{good}\rangle$, and leaves the bad states unchanged. Let $U_{0^\perp}$ be the unitary that flips the phase of the all-zeros state i.e. $U_{0^\perp}|0 \cdots 0\rangle = -|0 \cdots 0\rangle$ and leaves all other states unchanged. Let the iterate $Q = \mathcal{A}U_{0^\perp}\mathcal{A}^{-1}U_f$. Applying the iterate k times to the state $|\Psi\rangle$, i.e. $Q^k|\Psi\rangle$ results in the state $|\tilde{\Psi}\rangle = \sin(2k+1)\theta|\Psi_{good}\rangle + \cos(2k+1)\theta|\Psi_{bad}\rangle$, where $\sin^2\theta = p_{good}$. Setting*

$k = \mathcal{O}\left(\sqrt{\frac{N}{m}}\right)$ *and measuring* $|\tilde{\Psi}\rangle$ *produces the state* $|\Psi_{good}\rangle$ *with probability close to* 1.

We note that often times the goal of amplitude amplification is to search for an element in a given set of elements. However in our algorithm, we use it as a tool to take a uniform superposition over all elements in our set, and produce a uniform superposition over just the marked elements in our set with high probability.

A.2.5 Discrete Gaussians

Let $s > 0$ and $\rho_s(\mathbf{v}) = e^{\frac{-\pi|\mathbf{v}|^2}{s^2}}$ for $\mathbf{v} \in \mathbb{R}^d$, and $\rho_s(\mathcal{L}) = \sum_{\mathbf{v}\in\mathcal{L}} \rho_s(\mathbf{v})$ for a given lattice $\mathcal{L}$. The Discrete Gaussian over the lattice $\mathcal{L}$, denoted $D_{\mathcal{L},s}$ is a distribution where the probability of a vector $\mathbf{x} \in \mathcal{L}$ is given by

$$\Pr_{X\sim D_{\mathcal{L},s}}[X = \mathbf{x}] = \frac{\rho_s(\mathbf{x})}{\rho_s(\mathcal{L})}$$

The smoothing parameter for a lattice $\mathcal{L}$, denoted $\eta(\mathcal{L})$ is defined as the smallest real $s > 0$ such that $\rho_{\frac{1}{s}}(\mathcal{L}^* \setminus \{\mathbf{0}\}) \leq \epsilon$ for any positive real $\epsilon > 0$, and where $\mathcal{L}^* = \{\mathbf{v} \in \mathbb{R}^m \mid \langle \mathbf{v}, \mathbf{u}\rangle \in \mathbb{Z}, \forall \mathbf{u} \in \mathcal{L}\}$ is called the dual lattice to $\mathcal{L}$. In [98], the authors show how to efficiently sample from a Discrete Gaussian distribution, given a width greater than the smoothing parameter of the lattice. Combining the result above with a result from Grover et al. in [99], we assume we are able to create a superposition over lattice points of the form $\frac{1}{\sqrt{N}} \sum_{\mathbf{v}\in\mathcal{L}} |\mathbf{v}\rangle$ where N is the number of lattice points we want to sample. We call this subroutine $\mathsf{Sample}(\mathcal{L})$ that produces such a superposition for use in our algorithms.

A.3 Quantum Double Sieve

A.3.1 Binary quantum double sieve

The quantum double sieve is a quantization of the classical double sieve algorithm of Bai et al. [93]. The idea is to reduce pairs of vectors, the difference being that we do this in quantum superposition. We sample in superposition from a discrete Gaussian distribution of width greater than the smoothing parameter η for our lattice. Following the analysis in Section (A.1.1), we need $\mathcal{O}(2^{0.2075d})$ vectors in our initial sample in order for any fixed vector in our set to reduce against any other vector in the set with probability close to 1. We combine two such superpositions in order to get a superposition over pairs of vectors in the lattice. We then use amplitude amplification to search over pairs of vectors where the norm of the difference between the pair is less than γR. This is analogous to running one iteration of the classical double sieve. We do this iteratively for t steps of the sieve, starting with an initial set of 2^t superpositions as shown in Figure (A.2). The sieve runs for t iterations, thus immediately yielding the memory requirement for the algorithm. The runtime of this algorithm is determined by the number of amplitude amplification steps that we need to run. Algorithm 6 presents the binary quantum double sieve algorithm.

We explain the amplitude amplification step of the algorithm above. At level i of the protocol, $|\phi_1\rangle = \frac{1}{\sqrt{N}} \sum_{|\mathbf{u}_1| \leq \gamma^{i-1} R} |\mathbf{u}_1\rangle$ and $|\phi_2\rangle = \frac{1}{\sqrt{N}} \sum_{|\mathbf{u}_2| \leq \gamma^{i-1} R} |\mathbf{u}_2\rangle$, each with a history state that we do not write. At this point we tensor those two states to get a superposition over pair of vectors $\frac{1}{N} \sum_{\mathbf{u}_1, \mathbf{u}_2} |\mathbf{u}_1, \mathbf{u}_2\rangle$. Let $f_i(x) : \mathbb{R} \mapsto \{0, 1\}$ be defined as follows:

Algorithm 6: Binary Quantum Sieve

Input:

$\mathcal{L}$: The lattice
$\mathsf{Sample}(\mathcal{L})$: Subroutine to sample vectors from $\mathcal{L}$
γ: Norm reduction factor
R: Upper bound on norm of vectors sampled
C: Upper bound on the length of the shortest vector
N: Number of vectors to be sampled

1 Let t be such that $2^t R \leq C$.
2 Create 2^t superpositions of the form $|\phi_j\rangle = \frac{1}{\sqrt{N}} \sum_{|\mathbf{u}_j| \leq R} |\mathbf{u}_j\rangle$ for $j \in \{1, \cdots, 2^t\}$ using $\mathsf{Sample}(\mathcal{L})$.
3 **for** $i \in [t]$
4 Run amplitude amplification according to eq. (A.2) on pairs of vectors $|\phi_1, \phi_2\rangle, \cdots, |\phi_{2^{t-i+1}-1}, \phi_{2^{t-i+1}}\rangle$ to get 2^{t-i} states of the form $|\phi_j\rangle = \frac{1}{\sqrt{N}} \sum_{|\mathbf{u}_j| \leq \gamma^i R} |\mathbf{u}_j\rangle$ for $j \in \{1, \cdots, 2^{t-i}\}$.
5 Measure the first register of the final state $|\phi_1\rangle = \frac{1}{\sqrt{N}} \sum_{|\mathbf{u}_1| \leq \gamma^t R} |\mathbf{u}_1\rangle$ to output vector $\mathbf{u}_1$ of desired norm.

$$f_i(x) = \begin{cases} 1 & \text{if } x \leq \gamma^i R \\ \\ 0 & \text{otherwise} \end{cases} \tag{A.2}$$

We can define a unitary U_{f_i} that maps a state $|x\rangle$ to $(-1)^{f_i(x)}|x\rangle$, where x will be $|\mathbf{u}_2 - \mathbf{u}_1|$ in our case. Then we use standard amplitude amplification:

$$\frac{1}{N} \sum_{\mathbf{u}_1, \mathbf{u}_2} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle \xmapsto{AA} \frac{1}{\sqrt{N}} \sum_{|\mathbf{u}_2 - \mathbf{u}_1| \leq \gamma^i R} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle + |\text{junk}\rangle$$

We then measure to discard the junk state, compute $\mathbf{u}_2 - \mathbf{u}_1$ in the first register which becomes state $|\phi_1\rangle$ for the next iteration $i + 1$.

Assumption 69. *There exists an efficient quantum algorithm $\mathcal{A}$, as given in Definition 68, that uses Heuristic 65 and the history state from eq. (A.1) such that $\mathcal{A}|0 \cdots 0\rangle$ produces a state negligibly close in trace distance to $\frac{1}{\sqrt{N}} \sum_{|\mathbf{w}| \leq \gamma^i R} |\mathbf{w}\rangle$, where γ and R are the parameters*

used in Algorithms 5 and 6.

Theorem 70. *Given a lattice $\mathcal{L}$ of dimension d, $N = \mathcal{O}(2^{0.2075d})$, smoothing parameter η, norm reduction factor γ, and $R \geq \eta$, Algorithm (6) solves the SVP for $\mathcal{L}$ in t iterations, taking time $2^t\sqrt{N} = \mathcal{O}(R^c\sqrt{N})$ for some constant c, and memory $2^t = poly(R)$, if Assumption 69 holds.*

Proof. Algorithm (6) gives a binary tree of superpositions as shown in Figure (A.2). The amount of memory used by the algorithm is then given by 2^t where t is the height of the tree. At each level of the tree, we use amplitude amplification between pairs of superpositions, so the total number of amplitude amplification steps is also 2^t. Assume C is a small constant that is a good enough approximation for the length of the shortest vector. In order to compute the quantity 2^t, we have that:

$$\begin{aligned} \gamma^t R &\leq C \\ \therefore \log(\gamma^t R) &\leq \log(C) \\ \therefore t &\leq \frac{1}{\log(\gamma)} \left(\log(C) - \log(R)\right) \end{aligned}$$

Plugging in this value of t in 2^t we get:

$$\begin{aligned} 2^t &\leq 2^{\frac{1}{\log(\gamma)}(\log(C)-\log(R))} \\ &= 2^{\log(C^{\frac{1}{\log(\gamma)}})-\log(R^{\frac{1}{\log(\gamma)}})} \\ &= C^{\frac{1}{\log(\gamma)}} R^{-\frac{1}{\log(\gamma)}} = poly(R) \end{aligned}$$

Next we analyze the time complexity of the amplitude amplification step. At level i of

the algorithm, we perform amplitude amplification on a superposition of vectors $|\phi_1\rangle \otimes |\phi_2\rangle = \frac{1}{N}\sum_{|\mathbf{u}_1|,|\mathbf{u}_2|\leq\gamma^{i-1}R} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle$ which can be split up as a superposition over good and bad states. We note that we have roughly N marked good elements such that $|\mathbf{u}_2 - \mathbf{u}_1| \leq \gamma^i R$. We rewrite $|\phi_1\rangle \otimes |\phi_2\rangle$ as

$$\begin{aligned}
|\phi_1\rangle \otimes |\phi_2\rangle &= \frac{1}{N} \sum_{|\mathbf{u}_1|,|\mathbf{u}_2|\leq\gamma^{i-1}R} |\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle \\
&= \frac{1}{\sqrt{N}}|\psi_{\text{good}}\rangle + \sqrt{1 - \frac{1}{N}}|\psi_{\text{bad}}\rangle \\
\text{where } |\psi_{\text{good}}\rangle &= \sum_{|\mathbf{u}_2 - \mathbf{u}_1|\leq\gamma^i R} \frac{1}{\sqrt{N}}|\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle, \text{ and} \\
|\psi_{\text{bad}}\rangle &= \sum_{|\mathbf{u}_2 - \mathbf{u}_1|>\gamma^i R} \frac{1}{\sqrt{N^2 - N}}|\mathbf{u}_1, \mathbf{u}_2, |\mathbf{u}_2 - \mathbf{u}_1|\rangle
\end{aligned}$$

Next we perform standard amplitude amplification and get the state $|\psi_{\text{good}}\rangle$ with probability close to 1 using $\mathcal{O}(\sqrt{\frac{N^2}{m}})$ iterations of the search iterate, where m is the number of marked elements. In our case, we have $m = N$ marked elements, so amplitude amplification takes time $\mathcal{O}(\sqrt{\frac{N^2}{N}}) = \mathcal{O}(\sqrt{N})$. We perform a total of 2^t amplitude amplification steps, so the runtime of the algorithm is given by $\mathcal{O}(2^t\sqrt{N}) = \mathcal{O}(R^c\sqrt{N})$ for $c = \frac{1}{\log\frac{1}{\gamma}}$. □

Corollary 71. *Given a lattice $\mathcal{L}$ of dimension d, $N = \mathcal{O}(2^{0.2075d})$, smoothing parameter η, norm reduction factor γ, and $\eta \leq R = \mathcal{O}(2^{\alpha d})$ for $\alpha > 0$, Algorithm (6) solves SVP for $\mathcal{L}$ in time $\mathcal{O}(2^{(\alpha c+0.1038)d})$ and memory $\mathcal{O}(2^{\alpha cd})$ for some constant c, if Assumption 69 holds.*

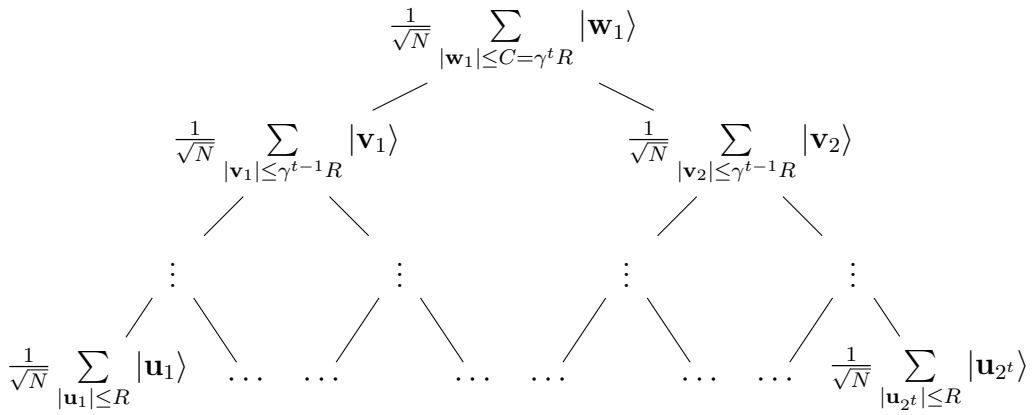


Figure A.2: Each level in the tree is one iteration of the sieve. In order to run the sieve for t iterations we need 2^t initial superpositions over vectors in our lattice

A.3.2 r-ary Quantum Sieve

We show a variation of the binary quantum double sieve technique to improve the memory bound of the sieve. At each step of the binary quantum sieve, we combined vectors of roughly the same norm, and searched for differences of them that gave us shorter vectors. We can provide a slight improvement in memory by combining two superpositions where the first one is over vectors of norm bounded by a quantity R, while the second is over vectors of norm bounded by a smaller quantity, say γR. We describe the process step by step.

Let us begin by sampling two superpositions where the norms of the vectors are bounded by R in both superpositions. Let these superpositions be over M vectors, for M that will be determined later.

$$\frac{1}{\sqrt{M}} \sum_{|\mathbf{v}_1| \leq R} |\mathbf{v}_1\rangle \otimes \frac{1}{\sqrt{M}} \sum_{|\mathbf{v}_2| \leq R} |\mathbf{v}_2\rangle = \frac{1}{M} \sum_{\mathbf{v}_1, \mathbf{v}_2} |\mathbf{v}_1, \mathbf{v}_2\rangle$$

We compute $|\mathbf{v}_2 - \mathbf{v}_1|$ in an auxiliary register, and using a unitary U_f defined in the same

way as Equation (A.2), we search for pairs of vectors where the norm of the difference is less than γR. Amplitude amplification gives us

$$\frac{1}{M}\sum_{\mathbf{v}_1,\mathbf{v}_2}|\mathbf{v}_1,\mathbf{v}_2,|\mathbf{v}_2-\mathbf{v}_1|\rangle \xmapsto{A.A.} \frac{1}{\sqrt{M}}\sum_{|\mathbf{v}_2-\mathbf{v}_1|\leq\gamma R}|\mathbf{v}_1,\mathbf{v}_2,|\mathbf{v}_2-\mathbf{v}_1|\rangle$$
$$\longmapsto \frac{1}{\sqrt{M}}\sum_{\substack{|\mathbf{w}_1|\leq\gamma R\\ \mathbf{w}_1=\mathbf{v}_2-\mathbf{v}_1}}|\mathbf{w}_1\rangle|\text{history}\rangle$$

Notice that this first step is the same as in the Binary Quantum Sieve. In the next step, however, we combine vectors with slightly larger norm, bounded by R, with the output from our amplitude amplification step. We run amplitude amplification on this superposition to find even shorter vectors, namely, those whose norm is bounded by $\gamma^2 R$.

$$\frac{1}{\sqrt{M}}\sum_{|\mathbf{w}_1|\leq\gamma R}|\mathbf{w}_1\rangle\otimes\frac{1}{\sqrt{M}}\sum_{|\mathbf{v}_3|\leq R}|\mathbf{v}_3\rangle \xmapsto{A.A.} \frac{1}{\sqrt{M}}\sum_{|\mathbf{w}_2|\leq\gamma^2 R}|\mathbf{w}_2\rangle$$

We continue combining shorter vectors with slightly longer vectors and search over the superposition for even shorter vectors. This process for t' steps is illustrated in Figure (A.3).

At the i^{th} (for $i < t'$) iteration, we combine vectors $\mathbf{w}_i$ of length bounded by $\gamma^i R$ with vectors $\mathbf{v}_{i+2}$ of length bounded by R. For our analysis, we normalize the lengths of $\mathbf{w}_i$ and $\mathbf{v}_{i+2}$ by R, so $|\mathbf{w}_i| \approx \gamma^i$ and $|\mathbf{v}_{i+2}| \approx 1$. Note that we require the angle $\theta_{\mathbf{w}_i,\mathbf{v}_{i+2}}$ between $\mathbf{w}_i$ and $\mathbf{v}_{i+2}$ to be $\leq \frac{\pi}{3}$, and $\gamma^i > \frac{1}{2}$ since no feasible solutions exist for $|\mathbf{v}_{i+2} - \mathbf{w}_i| \leq |\mathbf{w}_i|$ for $|\mathbf{w}_i| \leq \frac{1}{2}$ and $|\mathbf{v}_{i+2}| = 1$. Thus for $\gamma^i > \frac{1}{2}$, in order to get reductions between $\mathbf{w}_i$ and $\mathbf{v}_{i+1}$ we want vectors $\mathbf{w}_i \in B(\mathbf{0}, \gamma^i) \cap B(\mathbf{v}_{i+2}, \gamma.\gamma^i)$.

We restate a lemma from [93] that will help us compute the quantity $B(\mathbf{0}, \gamma^i) \cap$

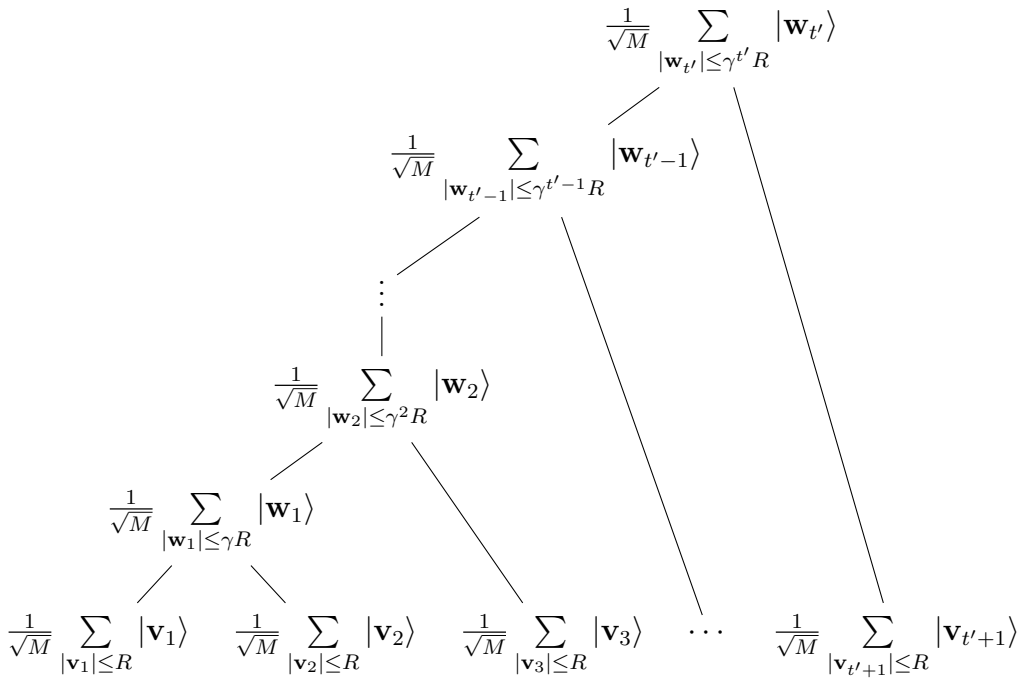


Figure A.3: Each level of the tree represents one step of the sieve and vectors on that level have norm shorter by a factor of γ. This tree shows that we get linear growth in the number of superpositions we need for up to t' iterations of the sieve, as opposed to an exponential number of superpositions in the binary sieve.

$B(\mathbf{v}_{i+2}, \gamma.\gamma^i)$

Lemma 72. *Given* $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^d$*, and* $r_1, r_2 \in \mathbb{R}$*, and* $|\mathbf{u}_1 - \mathbf{u}_2| = e$*, such that* $r_1, r_2 < e < r_1 + r_2$*,*

$$|B_d(\mathbf{u}_1, r_1) \cap B_d(\mathbf{u}_2, r_2)| \propto \left(\frac{-e^4 + 2e^2(r_1^2 + r_2^2) - (r_1^2 - r_2^2)^2)}{4e^2}\right)^{\frac{d}{2}} |B_d|$$

Setting $\gamma^i = s$ for simplicity, and using Lemma (72), we get

$$\begin{aligned} |B_d(\mathbf{0}, s) \cap B_d(\mathbf{v}_{i+2}, \gamma s)| &\propto \left(\frac{-1 + 2(s^2 + s^2\gamma^2) - (s^2 - s^2\gamma^2)^2)}{4}\right)^{\frac{d}{2}} |B_d| \\ &= \left(\frac{-1 + 2s^2(1 + \gamma^2) - s^4(1 - \gamma^2)^2)}{4}\right)^{\frac{d}{2}} |B_d| \end{aligned}$$

$$= \left(-\frac{1}{4s^2} + \frac{(1+\gamma^2)}{2} - \frac{s^2(1-\gamma^2)^2)}{4}\right)^{\frac{d}{2}} |B_d(s)| \tag{A.3}$$

This volume gives us the portion of the sphere of radius s that one fixed vector $\mathbf{w}_i$ covers. In order for any given vector in our set to reduce with a larger vector with probability close to 1, we need to cover the whole sphere. Therefore we need to sample $M \propto \frac{1}{p}$, where $p = \left(-\frac{1}{4\gamma^{2i}} + \frac{(1+\gamma^2)}{2} - \frac{\gamma^{2i}(1-\gamma^2)^2)}{4}\right)^{\frac{d}{2}}$.

Now that we have a bound on the number of samples we need, we can state the complete algorithm for solving SVP, shown in Algorithm (7).

Theorem 73. *Given a lattice $\mathcal{L}$ of dimension d, a norm reduction factor γ, and a value $M = \mathcal{O}(2^{\delta d})$ where $\delta = \frac{1}{2}\log(1/(-\frac{1}{4\gamma^{2t'}} + \frac{(1+\gamma^2)}{2} - \frac{\gamma^{2t'}(1-\gamma^2)^2)}{4}))$ and t' is the number of iterations of the sieve with $\gamma^{t'} > \frac{1}{2}$. Let η be the smoothing parameter and R such that $\eta \leq R = \mathcal{O}(2^{\alpha d})$ for $\alpha > 0$. Then Algorithm (7) solves SVP for $\mathcal{L}$ in memory $\mathcal{O}\left(2^{\left(\alpha c' + \frac{\delta}{2}\right)d}\right)$ and time $\mathcal{O}(R^{c'}\sqrt{M})$ for some constant c', if Assumption 69 holds.*

Proof. The appropriate value for M given γ and the number of iterations is given by the analysis of Eq (A.3). We can compute the exponent δ by taking the appropriate logarithm, and we get $\delta = \frac{1}{2}\log(1/(-\frac{1}{4\gamma^{2t'}} + \frac{(1+\gamma^2)}{2} - \frac{\gamma^{2t'}(1-\gamma^2)^2)}{4}))$ since we have t' iterations. Let $t' + 1 = r$. Let x be such that $\gamma^{t'x}R \leq C$ where C is a bound on the shortest vector (for example C can be taken to be the Minkowski bound). We need $(t'+1)^x = r^x$ vectors. We have:

$$\gamma^{t'x}R = \gamma^{(r-1)x}R \leq C$$

$$\therefore \log_r(\gamma^{(r-1)x}R) \leq \log_r(C)$$

Algorithm 7: r-ary Quantum Sieve

Input:

$\mathcal{L}$: The lattice
$\mathsf{Sample}(\mathcal{L})$: Subroutine to sample vectors from $\mathcal{L}$
γ: Norm reduction factor
R: Upper bound on norm of vectors sampled
C: Upper bound on the length of the shortest vector
t': r-ary tree depth limit
M: Number of vectors to sample

1 Let x be such that $2^{xt'}R \leq C$.
2 Set $r = t' + 1$.
3 Create r^x states of the form $|\phi_j\rangle = \frac{1}{\sqrt{M}} \sum_{|\mathbf{u}_j| \leq R} |\mathbf{u}_j\rangle$ for $j \in \{1, \cdots, r^x\}$ using $\mathsf{Sample}(\mathcal{L})$.
4 **for** $c \in [x]$
5 Run Algorithm 8 on every block of r superpositions $(|\phi_1\rangle, \cdots, |\phi_r\rangle), \cdots, (|\phi_{r^{x-c+1}-r}\rangle, \cdots, |\phi_{r^{x-c+1}}\rangle)$ to obtain r^{x-c} states of the form $|\phi_j\rangle = \frac{1}{\sqrt{M}} \sum_{|\mathbf{u}_j| \leq \gamma^{ct'}R} |\mathbf{u}_j\rangle$ for $j \in \{1, \cdots, r^{x-c}\}$.
6 Measure the first register of the final state $|\phi_1\rangle = \frac{1}{\sqrt{M}} \sum_{|\mathbf{u}_1| \leq \gamma^{xt'}R} |\mathbf{u}_1\rangle$ to output vector $\mathbf{u}_1$ of desired norm.

$$\therefore x \leq \log_r(C^{\frac{1}{(r-1)\log_r(\gamma)}}) - \log_r(R^{\frac{1}{(r-1)\log_r(\gamma)}})$$

$$\therefore r^x \leq C^{\frac{\log(t'+1)}{t'\log(\gamma)}} R^{\frac{\log(t'+1)}{t'\log(\frac{1}{\gamma})}} = \mathcal{O}(R^{c'})$$

The time complexity is identical to the analysis in Theorem (70), and hence we get that the runtime is $\mathcal{O}(R^{c'}\sqrt{M})$ where $c' = \frac{\log(t'+1)}{t'\log(\frac{1}{\gamma})}$. □

We note that the parameters δ, t', and γ are tunable, and can be optimized based on the value of α to get the lowest possible value for the constant c'.

A.4 Preprocessing superpositions

As communicated to us [100], we can clarify the results of Theorem 70 (with an additional suitable heuristic) by preprocessing our initial superposition. By choosing suitable block

Algorithm 8: Auxiliary algorithm to Algorithm 7 for running amplitude amplification on vectors of differing norms

Input:

γ, R, t': Parameters from Algorithm 7

c: Iteration level in step (4) of Algorithm 7

$|\phi_1\rangle, \cdots, |\phi_r\rangle$: Superpositions over vectors of norm bounded by $\gamma^{(c-1)t'}R$

1 **for** $i \in 2, \cdots, r$ **do**

2 Run amplitude amplification on states $|\phi_1\rangle \otimes |\phi_i\rangle$ using $f_{(i-1)+(c-1)t'}$ as per eq. (A.2). Relabel the output state as $|\phi_1\rangle$.

3 Return $|\phi_1\rangle = \frac{1}{\sqrt{M}} \sum_{|\mathbf{u}_1| \leq \gamma^{ct'}R} |\mathbf{u}_1\rangle$.

sizes, one can use BKZ with $2^{\tilde{\mathcal{O}}(d^{1-\lambda})}$ work to obtain vectors of length $R = 2^{\tilde{\mathcal{O}}(d^{\lambda})}$.

A.5 Future Work

We intend to find a way to turn the assumption we make (Assumption 69) about the amplitude amplification operator into a lemma. This work is a first step toward quantizing classical algorithms by doing more than just a replacement of the classical search step with quantum search. In our case, the Double Sieve was a very good candidate for using the 'sieving-in-superposition' technique. It might be interesting to see if this technique extends to other classical sieving algorithms like Laarhoven's SphereSieve[90] and HashSieve [91] to yield even better exponents in the runtime and memory complexities.

Bibliography

[1] Ueli Maurer and Renato Renner. Abstract cryptography. In *Innovations in Computer Science*. Citeseer, 2011.

[2] Ueli Maurer and Björn Tackmann. On the soundness of authenticate-then-encrypt: formalizing the malleability of symmetric encryption. In *Proceedings of the 17th ACM conference on Computer and communications security*, pages 505–515, 2010.

[3] Ueli Maurer, Andreas Rüedlinger, and Björn Tackmann. Confidentiality and integrity: A constructive perspective. In *Theory of Cryptography: 9th Theory of Cryptography Conference, TCC 2012, Taormina, Sicily, Italy, March 19-21, 2012. Proceedings 9*, pages 209–229. Springer, 2012.

[4] Christian Badertscher, Fabio Banfi, and Ueli Maurer. A constructive perspective on signcryption security. In *Security and Cryptography for Networks: 11th International Conference, SCN 2018, Amalfi, Italy, September 5–7, 2018, Proceedings 11*, pages 102–120. Springer, 2018.

[5] Dennis Hofheinz, Christian Matt, and Ueli Maurer. Idealizing identity-based encryption. In *Advances in Cryptology–ASIACRYPT 2015: 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29–December 3, 2015, Proceedings, Part I 21*, pages 495–520. Springer, 2015.

[6] Christian Badertscher, Ueli Maurer, and Björn Tackmann. On composable security for digital signatures. In *Public-Key Cryptography–PKC 2018: 21st IACR International Conference on Practice and Theory of Public-Key Cryptography, Rio de Janeiro, Brazil, March 25-29, 2018, Proceedings, Part I 21*, pages 494–523. Springer, 2018.

[7] Christian Badertscher. Key exchange security in constructive cryptography. Master's thesis, Eidgenössische Technische Hochschule Zürich, Department of Computer Science, 2012.

[8] Sandro Coretti, Ueli Maurer, and Björn Tackmann. A constructive perspective on key encapsulation. *Number Theory and Cryptography: Papers in Honor of Johannes Buchmann on the Occasion of His 60th Birthday*, pages 226–239, 2013.

[9] Christian Matt and Ueli Maurer. The one-time pad revisited. In *2013 IEEE International Symposium on Information Theory*, pages 2706–2710. IEEE, 2013.

[10] Grégory Demay and Ueli Maurer. Unfair coin tossing. In *2013 IEEE International Symposium on Information Theory*, pages 1556–1560. IEEE, 2013.

[11] Christopher Portmann and Renato Renner. Cryptographic security of quantum key distribution. *arXiv preprint arXiv:1409.3525*, 2014.

[12] Fabio Banfi, Ueli Maurer, Christopher Portmann, and Jiamin Zhu. Composable and finite computational security of quantum message transmission. In *Theory of Cryptography Conference*, pages 282–311. Springer, 2019.

[13] Venkatesh Vilasini, Christopher Portmann, and Lídia Del Rio. Composable security in relativistic quantum cryptography. *New Journal of Physics*, 21(4):043057, 2019.

[14] Artur K Ekert. Quantum cryptography based on Bell's theorem. *Physical review letters*, 67(6):661, 1991.

[15] Charles Bennett and Gilles Brassard. Quantum crytography: Public key distribution and coin tossing. In *Proc. IEEE Int. Conf. Computers, Systems, and Signal Processing, Bangalore, India, 1984*, pages 175–179, 1984.

[16] Hoi-Kwong Lo and Hoi Fung Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *science*, 283(5410):2050–2056, 1999.

[17] Eli Biham, Michel Boyer, P Oscar Boykin, Tal Mor, and Vwani Roychowdhury. A proof of the security of quantum key distribution. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 715–724, 2000.

[18] Peter W Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Physical review letters*, 85(2):441, 2000.

[19] Dominic Mayers. Unconditional security in quantum cryptography. *Journal of the ACM (JACM)*, 48(3):351–406, 2001.

[20] Christopher A Fuchs, Nicolas Gisin, Robert B Griffiths, Chi-Sheng Niu, and Asher Peres. Optimal eavesdropping in quantum cryptography. i. information bound and optimal strategy. *Physical Review A*, 56(2):1163, 1997.

[21] Juan-Ignacio Cirac and Nicolas Gisin. Coherent eavesdropping strategies for the four state quantum cryptography protocol. *Physics Letters A*, 229(1):1–7, 1997.

[22] Dagmar Bruß. Optimal eavesdropping in quantum cryptography with six states. *Physical Review Letters*, 81(14):3018, 1998.

[23] Helle Bechmann-Pasquinucci and Nicolas Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Physical Review A*, 59(6):4238, 1999.

[24] Valerio Scarani and Christian Kurtsiefer. The black paper of quantum cryptography: real implementation problems. *Theoretical Computer Science*, 560:27–32, 2014.

[25] Xiang-Bin Wang. Beating the photon-number-splitting attack in practical quantum cryptography. *Physical review letters*, 94(23):230503, 2005.

[26] Hoi-Kwong Lo, Xiongfeng Ma, and Kai Chen. Decoy state quantum key distribution. *Physical review letters*, 94(23):230504, 2005.

[27] Dominic Mayers and Andrew Yao. Quantum cryptography with imperfect apparatus. In *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No. 98CB36280)*, pages 503–509. IEEE, 1998.

[28] Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Physical Review Letters*, 113(14):Art–No, 2014.

[29] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880, 1969.

[30] Carl A Miller and Yaoyun Shi. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. *Journal of the ACM (JACM)*, 63(4):1–63, 2016.

[31] Frederic Dupuis, Omar Fawzi, and Renato Renner. Entropy accumulation. *Communications in Mathematical Physics*, 379(3):867–913, 2020.

[32] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. Simple and tight device-independent security proofs. *SIAM Journal on Computing*, 48(1):181–225, 2019.

[33] Bas Hensen, Hannes Bernien, Anaïs E Dréau, Andreas Reiserer, Norbert Kalb, Machiel S Blok, Just Ruitenberg, Raymond FL Vermeulen, Raymond N Schouten, Carlos Abellán, et al. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[34] Marissa Giustina, Marijn AM Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Åke Larsson, Carlos Abellán, et al. Significant-loophole-free test of bell's theorem with entangled photons. *Physical review letters*, 115(25):250401, 2015.

[35] Ming-Han Li, Cheng Wu, Yanbao Zhang, Wen-Zhao Liu, Bing Bai, Yang Liu, Weijun Zhang, Qi Zhao, Hao Li, Zhen Wang, et al. Test of local realism into the past without detection and locality loopholes. *Physical review letters*, 121(8):080404, 2018.

[36] Nishant Rodrigues and Brad Lackey. Nonlocal games, synchronous correlations, and Bell inequalities. *arXiv preprint arXiv:1707.06200v4*, 2020.

[37] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[38] Phillip Kaye, Raymond Laflamme, Michele Mosca, et al. *An introduction to quantum computing*. Oxford University Press on Demand, 2007.

[39] Mark M Wilde. *Quantum information theory*. Cambridge University Press, 2013.

[40] John Watrous. *The theory of quantum information*. Cambridge university press, 2018.

[41] John Stewart Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, 1964.

[42] Boris S Tsirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.

[43] N David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical review letters*, 65(27):3373, 1990.

[44] Asher Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3-4):107–108, 1990.

[45] Antonio Acin, Nicolas Gisin, and Lluis Masanes. From bell's theorem to secure quantum key distribution. *Physical review letters*, 97(12):120405, 2006.

[46] Antonio Acin, Serge Massar, and Stefano Pironio. Efficient quantum key distribution secure against no-signalling eavesdroppers. *New Journal of Physics*, 8(8):126, 2006.

[47] L Masanes, R Renner, M Christandl, A Winter, and J Barrett. Unconditional security of key distribution from causality constraints. *arXiv preprint quantph/0606049*, 2006.

[48] Lluis Masanes, Stefano Pironio, and Antonio Acín. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature communications*, 2(1):238, 2011.

[49] Esther Hänggi, Renato Renner, and Stefan Wolf. Efficient device-independent quantum key distribution. In *Advances in Cryptology–EUROCRYPT 2010: 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, French Riviera, May 30–June 3, 2010. Proceedings 29*, pages 216–234. Springer, 2010.

[50] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. Device-independent quantum key distribution secure against collective attacks. *New Journal of Physics*, 11(4):045021, 2009.

[51] Jonathan Barrett, Roger Colbeck, and Adrian Kent. Unconditionally secure device-independent quantum key distribution with only two devices. *Physical Review A*, 86(6):062326, 2012.

[52] Ben W Reichardt, Falk Unger, and Umesh Vazirani. A classical leash for a quantum system: Command of quantum systems via rigidity of chsh games. *arXiv preprint arXiv:1209.0448*, 2012.

[53] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner. Generalised entropy accumulation. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 844–850. IEEE, 2022.

[54] J Lawrence Carter and Mark N Wegman. Universal classes of hash functions. In *Proceedings of the ninth annual ACM symposium on Theory of computing*, pages 106–112, 1977.

[55] Mark N Wegman and J Lawrence Carter. New hash functions and their use in authentication and set equality. *Journal of computer and system sciences*, 22(3):265–279, 1981.

[56] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6(01):1–127, 2008.

[57] Umesh Vazirani and Thomas Vidick. Certifiable quantum dice: or, true random number generation secure against quantum adversaries. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 61–76, 2012.

[58] Matthew Coudron and Henry Yuen. Infinite randomness expansion with a constant number of devices. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 427–436, 2014.

[59] Kai-Min Chung, Yaoyun Shi, and Xiaodi Wu. Physical randomness extractors: generating random numbers with minimal assumptions. *arXiv preprint arXiv:1402.4797*, 2014.

[60] Ueli Maurer and Renato Renner. From indifferentiability to constructive cryptography (and back). In *Theory of Cryptography Conference*, pages 3–24. Springer, 2016.

[61] Wikipedia contributors. Conjugate prior — Wikipedia, the free encyclopedia, 2023. [Online; accessed 8-November-2023].

[62] Boris S Tsirel'son. Some results and problems on quantum Bell-type inequalities. *Fundamental questions in quantum physics and relativity*, pages 32–48, 1993.

[63] Wim van Dam. Implausible consequences of superstrong nonlocality. *Natural Computing*, 12(1):9–12, 2013.

[64] Peter J Cameron, Ashley Montanaro, Michael W Newman, Simone Severini, and Andreas Winter. On the quantum chromatic number of a graph. *Electron. J. Combin*, 14(1), 2007.

[65] Nishant Rodrigues and Brad Lackey. Fully Device-Independent Quantum Key Distribution Using Synchronous Correlations. In Omar Fawzi and Michael Walter, editors, *18th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2023)*, volume 266 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 8:1–8:22, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[66] Ben F Toner and Dave Bacon. Communication cost of simulating Bell correlations. *Physical Review Letters*, 91(18):187904, 2003.

[67] Lynden K Shalm, Evan Meyer-Scott, Bradley G Christensen, Peter Bierhorst, Michael A Wayne, Martin J Stevens, Thomas Gerrits, Scott Glancy, Deny R Hamel, Michael S Allman, et al. Strong loophole-free test of local realism. *Physical review letters*, 115(25):250402, 2015.

[68] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP* = RE. *arXiv preprint arXiv:2001.04383*, 2020.

[69] Laura Mancinska and David Roberson. Graph homomorphisms for quantum players. In *9th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2014)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2014.

[70] Vern I Paulsen, Simone Severini, Daniel Stahlke, Ivan G Todorov, and Andreas Winter. Estimating quantum chromatic numbers. *Journal of Functional Analysis*, 270(6):2188–2222, 2016.

[71] Se-Jin Kim, Vern Paulsen, and Christopher Schafhauser. A synchronous game for binary constraint systems. *Journal of Mathematical Physics*, 59(3):032201, 2018.

[72] Thomas Vidick. Almost synchronous quantum correlations. *Journal of mathematical physics*, 63(2):022201, 2022.

[73] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[74] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Computational Complexity, 2004. Proceedings. 19th IEEE Annual Conference on*, pages 236–249. IEEE, 2004.

[75] Alex Arkhipov. Extending and characterizing quantum magic games. *arXiv preprint arXiv:1209.3819*, 2012.

[76] Andrea Coladangelo and Jalex Stark. Robust self-testing for linear constraint system games. *arXiv preprint arXiv:1709.09267*, 2017.

[77] Werner Oskar Amrein and Kalyan B Sinha. On pairs of projections in a Hilbert space. *Linear algebra and its applications*, 208:425–435, 1994.

[78] Paul R Halmos. Two subspaces. *Transactions of the American Mathematical Society*, 144:381–389, 1969.

[79] Albrecht Böttcher and Ilya M Spitkovsky. A gentle guide to the basics of two projections theory. *Linear Algebra and its Applications*, 432(6):1412–1459, 2010.

[80] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Physical Review Letters*, 98(23):230501, 2007.

[81] Michael Pohst. On the computation of lattice vectors of minimal length, successive minima and reduced bases with applications. *ACM Sigsam Bulletin*, 15(1):37–44, 1981.

[82] Ravi Kannan. Improved algorithms for integer programming and related lattice problems. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 193–206, 1983.

[83] Ulrich Fincke and Michael Pohst. Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. *Mathematics of computation*, 44(170):463–471, 1985.

[84] Claus-Peter Schnorr and Horst Helmut Hörner. Attacking the Chor-Rivest cryptosystem by improved lattice reduction. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 1–12. Springer, 1995.

[85] Nicolas Gama, Phong Q Nguyen, and Oded Regev. Lattice enumeration using extreme pruning. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 257–278. Springer, 2010.

[86] Claus Peter Schnorr. Average time fast SVP and CVP algorithms for low density lattices and the factorization of integers. *Technical Report*, 2010.

[87] Daniele Micciancio and Panagiotis Voulgaris. Faster exponential time algorithms for the shortest vector problem. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1468–1480. SIAM, 2010.

[88] Divesh Aggarwal, Daniel Dadush, Oded Regev, and Noah Stephens-Davidowitz. Solving the shortest vector problem in 2^n time using discrete gaussian sampling. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 733–742, 2015.

[89] Phong Q Nguyen and Thomas Vidick. Sieve algorithms for the shortest vector problem are practical. *Journal of Mathematical Cryptology*, 2(2):181–207, 2008.

[90] Thijs Laarhoven. Sieving for shortest vectors in lattices using angular locality-sensitive hashing. In *Annual Cryptology Conference*, pages 3–22. Springer, 2015.

[91] Thijs Laarhoven and Benne de Weger. Faster sieving for shortest lattice vectors using spherical locality-sensitive hashing. In *International Conference on Cryptology and Information Security in Latin America*, pages 101–118. Springer, 2015.

[92] Divesh Aggarwal, Yanlin Chen, Rajendra Kumar, and Yixin Shen. Improved (provable) algorithms for the shortest vector problem via bounded distance decoding. *arXiv preprint arXiv:2002.07955*, 2020.

[93] Shi Bai, Thijs Laarhoven, and Damien Stehlé. Tuple lattice sieving. *LMS Journal of Computation and Mathematics*, 19(A):146–162, 2016.

[94] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM Journal on Computing*, 37(1):267–302, 2007.

[95] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 56(6):1–40, 2009.

[96] Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 305:53–74, 2002.

[97] Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219, 1996.

[98] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 197–206, 2008.

[99] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. *arXiv preprint quant-ph/0208112*, 2002.

[100] Thijs Laarhoven. Private communication.