Aristotle University of Thessaloniki –
Laboratoire d'Informatique de Paris 6 (Sorbonne University)

Tsintsilidas Dimitrios 9343

Master's Thesis

Certifiable Random Number Generation Using Nonlocal Games

Advisor: Alex Bredariol Grilo

Thessaloniki, July 2022

Abstract

Random number generation is one of the most important and hardest computational tasks because of its various applications, especially in cryptography. Nowadays, classical cryptography uses pseudorandom number generators, which cannot extend randomness and are only based on computational assumptions. Quantum mechanics has perfect randomness in its postulates, so they can provide a way to construct cryptographically secure protocols that can expand randomness and certify that it is generated by quantum phenomena. In order to achieve this, we can use a type of games, called nonlocal games, in which the players do not communicate and if they can use quantum properties, they have a greater probability to win than using any classical strategy. Checking the winning probability can verify the use of a quantum mechanical system and as a result, ensure that the output is truly random. We can use this simple idea to make protocols that expand an initial random seed exponentially and even use it for infinite randomness generation.

In this thesis, we will present the basics of quantum information theory and nonlocal games, so we can use them to build a protocol for exponential randomness expansion and after that a protocol for infinite randomness expansion. These topics have been studied in the literature before. The main contribution of this thesis is constructing a protocol for infinite randomness expansion that uses only 3 quantum devices instead of 4, which was the lowest bound known for this task until now.

Contents

Chapter 1

Introduction

1.1 Quantum Cryptography

Quantum mechanics is a fundamental theory in physics that was discovered in the 20th century, but has found plenty of applications since then. In the 1970s, some attempts were made to connect quantum mechanics with information theory, which is a theory that was relatively new then, developed by Shannon in the 1940s [1]. In the 1980s, several physicists, like Paul Benioff [2], Richard Feynman [3] and Yuri Manin [4] suggested that we can use the properties of quantum mechanics for computational tasks, such as simulations of quantum systems. However, the main result that made quantum computing look promising for the future of computing in general, is Shor's algorithm, which was developed by Peter Shor in 1994 [5]. This algorithm makes possible to find the prime factors of an integer in polynomial time, while every known classical algorithm requires at least sub-exponential time, which is not practical for real-world applications. Thus, this breakthrough suggested that one of the most used cryptosystems, the RSA would no longer be secure, if quantum computers can be constructed.

In parallel with the development of quantum computing, it was realised that an alternative to these cryptosystems, which could be broken by Shor's algorithms, can use the same properties of quantum mechanics, so the field of quantum cryptography was invented. This idea was first proposed in the 1970s by Stephen Wiesner [6], who introduced the concept of quantum money and quantum conjugate coding. Quantum conjugate coding referred to storing messages by encoding them in two "conjugate basis", so that only one of them can be decoded due to Heisenberg's Uncertainty Principle. This could be applied to make unforgeable bank notes, which is called quantum money. However, the main breakthrough came when in 1984 Charles H. Bennett and Gilles Brassard [7] used the idea of conjugate coding in transmitting information, which resulted in the first protocol for quantum key distribution, which is called the BB84 protocol. A further improvement in the security of this protocol was achieved when Artur Ekert proposed in 1991 [8] to use entanglement and Bell's inequalities.

Entanglement is a quantum phenomenon that proved to be essential for quantum cryptography, since it can produce correlations that cannot be described by classical systems. These correlations make it possible for a user that communicates with unreliable black-box servers (the user has no information about the system) to check if these servers uses quantum properties or not. This is called self-testing. The term was first introduced by Mayers and Yao [9], who used self-testing to reduce some assumptions needed for quantum cryptography, thus inventing device-independent quantum cryptography. Their idea is to make protocols that do not rely on trusting that the devices we use are honest about using quantum properties.

This has also potential in a real-life scenario; quantum computers, when they can be fully functional, will only do some very specific tasks, so any user will not have their own quantum computer, but would be able to communicate with a central server that provides the computational power of quantum computers. However, the user cannot be sure that this server is trustful, so there must be a way that the user can verify that the server is honest. Apart from verification, device-independent protocols are made for quantum key distribution, randomness amplification and randomness expansion.

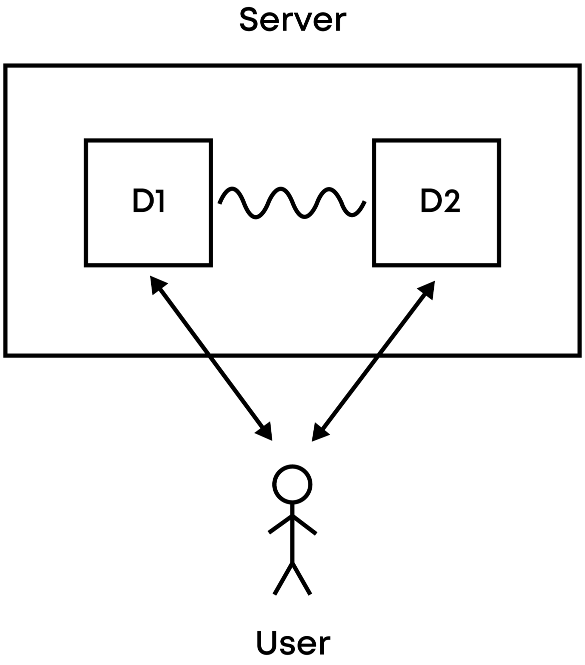


Figure 1.1: The setup of the protocols described in the thesis: There is a user that communicates with a server, which consists of different quantum devices that do not communicate, but they can share entanglement.

1.2 Randomness Generation

Random number generation is a process that generates a sequence of numbers that does not have some pattern and there is no way to be predicted with high probability. It is usually accomplished by devices or algorithms called random number generators (RNG) and has various applications in computer simulations, gambling, cryptography etc. Especially in cryptography, the use of randomness is crucial, because it ensures the security of our computational and communications infrastructure. Considering all the devices that are connected on the Internet right now always require more randomness to communicate, we can see that the amount of randomness needed is astronomical. On top of that, high quality randomness is essential for some cryptographic applications, like sharing a secret key.

We have two main methods to make RNGs: physical and computational. Physical RNGs, also called Hardware RNGs, generate randomness by means of a physical process, often based on microscopic phenomena, such as thermal noise. The computational method does not use some device, but some algorithm that uses a random input to expand it and generate a random output. However, since an algorithm does only deterministic operations, using the same input, which is known as seed, will give the same outcome. That is why we call these RNGs, Pseudorandom Number Generators (PRNGs).

There are some drawbacks in using PRNGs. The entropy of the output cannot be greater than the entropy of the input, so we cannot expand the randomness we already have. Also, the security of PRNGs is based on computational assumptions, such as the hardness of solving some problem, which we do not know definitely that it is true.

Moreover, there have been some attempts to certify randomness by only observing the sequence of the numbers. Many statistical tests have been invented and in the 1960s, Kolmogorov developed Kolmogorov complexity [10], which is an indicator of randomness contained in a certain string. However, there is a flaw in these tests, because some RNG can deterministically produce some string that passes the tests. So, in order to have true randomness we need an inherently random process.

Using the effects of quantum mechanics, we can design protocols for quantum RNGs that overcome all these problems. Since quantum mechanics is a theory with perfect randomness as one of its postulates, we can use it to produce and expand randomness, even using deterministic operations on the quantum system. We do not need to count on computational assumptions, but only on the correctness of quantum mechanics and using its properties, like Bell inequalities and self-testing, we can verify the random string without statistical tests, which are unreliable.

1.3 Motivation

In the literature, the study of certified randomness expansion protocols using quantum mechanics started in 2006, when Colbeck [11] suggested to use the violation of Bell's inequalities with 2 quantum devices to check if a system generates randomness. Classical security for these kind of protocols was proved in Pironio et al. [12], Fehr et al. [13], and Pironio and Massar [14], as well as in Coudron et al [15]. In 2012, Vazirani and Vidick [16] provided a protocol for exponential randomness expansion, which was secure against quantum adversaries, as well. In 2016, Coudron and Yuen using the Vazirani-Vidick protocol as subprotocol made the first adaptive protocol for infinite randomness expansion using 8 quantum devices [17]. In the same year, there were two other papers, one from Miller and Shi [18], who constructed a robust protocol for exponential expansion, improving the work of Vazirani and Vidick, and one from Chung, Shi and Wu [19], who proved the security for composition of protocols. These two results combined provided a procedure for infinite randomness expansion with 4 quantum devices. The goal of this thesis is to improve this result, so as to use 3 quantum devices for infinite randomness. For constructing such a protocol, we used the results from Chung, Shi and Wu [19], as well as a breakthrough result from Metger, Fawzi, Sutter and Renner [20], as building blocks of the proof.

1.4 Structure of this thesis

Chapter 2 We present the main concepts of quantum information theory that are essential for quantum computing and cryptography.

Chapter 3 This is an introduction to nonlocal games, a class of games, which are used in the proof of Bell's inequality and are a way to separate quantum from classical correlations. These games are the base for constructiong randomness expansion protocols.

Chapter 4 In this chapter, we present the main features used for randomness expansion protocol, a simple protocol for exponential expansion with proof of classical security and some upper bounds of non-adaptive protocols.

Chapter 5 This is the main contribution of this thesis. We present the Equivalence Lemma and Generalised Entropy Accumulation theorem and combine them to construct a protocol for infinite randomness expansion with 3 quantum devices.

Chapter 2

Quantum Information Basics

2.1 Qubit

To start with, we have to define the basic unit of quantum information, the qubit [21]. A qubit is analogous to a bit in classical information theory. Just as a bit can take two values, 0 and 1, so a qubit is a quantum-mechanical system which can be in two different states. However, in quantum mechanics these two states can co-exist in some combination of these two states, which is called superposition.

We can formulate this mathematically. Suppose that the two basic states are $|0\rangle$ and $|1\rangle$, which are some unit vectors orthogonal to each other on a vector space with dimension 2, consider $\mathbb{C}^2$. These two states obviously correspond to classical 0 and 1. The notation $|\cdot\rangle$ is called the "Dirac notation" and it is used to denote any quantum state. So, all the (linear) combinations of the two basic states are of the form

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \quad a, b \in \mathbb{C}$$

In order this to be a valid state in superposition, according to the laws of quantum mechanics, it must hold that $|a|^2 + |b|^2 = 1$. That means that the state $|\psi\rangle$ is also a unit vector in the vector space $\mathbb{C}^2$. Thus, the state space of a qubit is the set of all unit vectors in $\mathbb{C}^2$, which we denote as $S(\mathbb{C}^{\nvDash})$.

Since the vector space $\mathbb{C}^2$ is finite, we can consider that $|0\rangle$ and $|1\rangle$ constitute the standard basis, so

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and every state is of the form

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \qquad a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1$$

So, if a qubit can be in infinitely many states, why did we say that it is a two state quantum system? In fact, we can only observe two different states of the system using a measurement on the state $|\psi\rangle$: the state $|0\rangle$ with probability $|a|^2$ and the state $|1\rangle$ with probability $|b|^2$. For example if we measure the state

$$\frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle$$

we can get $|0\rangle$ or $|1\rangle$ with probability $(1/\sqrt{2})^2 = 1/2$.

After the measurement the state stays the same with the one we found with the measurement, so the initial state has lost its superposition. Then, we say that the state has collapsed to the specific outcome state.

However, we can measure a qubit not only using the vectors of the standard basis, but also any other orthonormal basis of the vector space $\mathbb{C}^2$. For instance, we can define a new basis as

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \qquad |-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

We will call the basis $(|+\rangle, |-\rangle)$ the Hadamard basis.

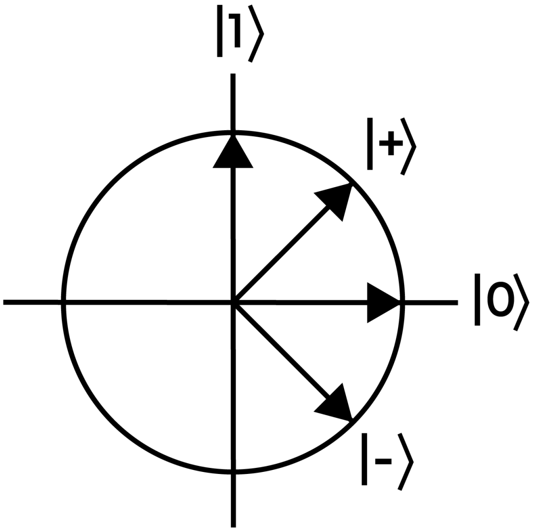


Figure 2.1: Standard basis $(|0\rangle, |1\rangle)$ and Hadamard basis $(|+\rangle, |-\rangle)$

When we measured in the standard basis, the probability of some outcome was the squared magnitude of the projection on the vector corresponding to this outcome. Hence we can generalise this for any basis using the inner product:

$$Pr(|\psi\rangle \ outcomes \ |+\rangle) = |(|\psi\rangle, |+\rangle)|^2 = |\,|\psi\rangle^* |+\rangle\,|^2 = |\,\langle\psi|+\rangle\,|^2$$

where $\langle\psi| = |\psi\rangle^* = (\overline{a} \ \ \overline{b})$ is the conjugate transpose of $|\psi\rangle$. The post-measurement state will again be the outcome state; this does not depend on the basis we use for the measurement.

It is clear that the procedure of measurement is probabilistic in nature, so it generates randomness. So, we can use this procedure to generate randomness for some applications, as we will see later.

2.2 Multiple qubits

Now let us suppose that we have two qubits (which are defined in two different vector spaces $\mathbb{C}^2$). If we want to describe them as a unified system, we have to take the tensor product of these two spaces, so as to keep the linear properties induced by quantum mechanics. Hence, we can define a two-qubit system in the space $\mathbb{C}^2 \otimes \mathbb{C}^2$ and the set of all two-qubit states as

$$S(\mathbb{C}^2 \otimes \mathbb{C}^2) = \{|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 : \||\psi\rangle\|_2^2 = 1\}$$

where $\|\cdot\|_2$ is the Euclidean norm.

The tensor product with vectors is defined as

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} = ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle$$

and we have $|ac|^2 + |ad|^2 + |bc|^2 + |bd|^2 = (|a|^2 + |b|^2)(|c|^2 + |d|^2) = 1$.

In these two-qubit systems, we can make measurements like in the case of qubit, but with 4 different outcomes, since any basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$ has cardinality 4. So, for example measuring

$$|\psi\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle$$

in the standard basis could get $|00\rangle$ with probability $|a|^2$ and after that the state would collapse to $|00\rangle$. The same holds respectively for the other basis states. However, it is possible to measure only one of the qubits. Measuring the first qubit would get $|0\rangle$ with probability $|a|^2 + |b|^2$ and $|1\rangle$ with probability $|c|^2 + |d|^2$. In the first case, the first qubit collapses to $|0\rangle$, so the possible states are $|00\rangle$ and $|01\rangle$. Thus, to find the post-measurement state we normalise the two possible states like this:

$$|\psi'\rangle = \frac{a\,|00\rangle + b\,|01\rangle}{|a|^2 + |b|^2}$$

We can generalise the definition of a two-qubit system to an n-qubit system, which is defined on the space $(\mathbb{C}^2)^{\otimes n}$. The measurement of an n-qubit system is defined similarly to that of the qubit with the inner product and the post-measurement state, as we have seen above. Obviously, every basis of $(\mathbb{C}^2)^{\otimes n}$ has cardinality 2^n, so by measuring an n-qubit system we can get 2^n different outcomes.

2.2.1 Entanglement

The properties of the tensor product allow for another remarkable phenomenon in quantum mechanics, apart from superposition, which is called entanglement. Since any state in $\mathbb{C}^2 \otimes \mathbb{C}^2$ with Euclidean norm equal to 1 is possible, we can consider the state:

$$|\psi\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

This state has the property that it cannot be written as the tensor product of two single qubits ($\nexists\ |\psi_1\rangle, |\psi_2\rangle \in \mathbb{C}^2$ such that $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$). If they existed, then assuming $|\psi_1\rangle = (a\ b)^T$ and $|\psi_2\rangle = (c\ d)^T$, we must have $ac = 1/\sqrt{2},\ ad = 0,\ bc = 0,\ bd = 1/\sqrt{2}$, which is impossible.

We will call all those states that cannot be decomposed to the tensor product of single qubit states, entangled. Entanglement can create some correlations between systems, which are not possible classically. We will study this in the next chapter.

2.3 Operators

2.3.1 Unitary evolution

After we have defined how information is represented in quantum systems, we need to determine the ways that these states change. So, we need some operation that maps one state (which is a vector) to another state. These maps from a vector space to another vector space are (linear) operators. In the case of finite vector spaces, which we need here, the operators are matrices. That means that an operator $\mathcal{O} : \mathbb{C}^n \to \mathbb{C}^m$ can be represented by an $m \times n$ matrix. We will see that the use of operators in quantum information is crucial.

The operators used for the evolution of a system must preserve the norm of the vectors, because the state vectors are all unit. These matrices, in the case of $n \times n$ operators, are called unitary matrices. A square matrix U is a unitary matrix if its conjugate transpose U^* is also its inverse, which means $UU^* = U^*U = \mathbb{I}$.

So, all the actions on a quantum state can be represented by a unitary matrix and reversely any unitary matrix represents an action on a state. If we have $|\psi\rangle$ as the initial state and $|\psi'\rangle$ as the final state (with the same number of qubits), then there is a unitary matrix U such that $|\psi'\rangle = U|\psi\rangle$.

In quantum computation, unitary matrices for small number of qubits are also called gates, because their action is analogous to logical gates in classical computation. The most famous gate is the Hadamard gate for one qubit:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

which maps $|0\rangle$ to $|+\rangle$, $|1\rangle$ to $|-\rangle$ and vice versa.

2.3.2 Measurements

Instead of measuring with different basis, we can also represent measurement using a collection of measurement operators $\{M_m\}$ where m is the index of the different outcomes. The rule we define by measuring with M_m is

$$p(m) = Pr(|\psi\rangle \text{ } outcomes \text{ } m) = \|M_m|\psi\rangle\|_2^2 = \langle M_m|\psi\rangle, M_m|\psi\rangle\rangle = \langle\psi| M_m^* M_m |\psi\rangle$$

and the post-measurement state is

$$\frac{M_m|\psi\rangle}{\langle\psi| M_m^* M_m |\psi\rangle}$$

Also, for completeness with probabilities, we have

$$\forall|\psi\rangle \quad 1 = \sum_m p(m) = \sum_m \langle\psi| M_m^* M_m |\psi\rangle \iff \sum_m M_m^* M_m = \mathbb{I}$$

If we define the operator $E = M^*M$, it is easy to see that E is a positive definite operator. That is why we call this type of measurement POVM (Positive Operator-Valued Measurement).

2.3.3 Projective Measurements and Observables

Measuring with a basis $\{|a_1\rangle, |a_2\rangle, ..., |a_n\rangle\}$ is equivalent with measuring with the operators $P_1 = |a_1\rangle \langle a_1|, P_2 = |a_2\rangle \langle a_2|$ and so on. These operators are Hermitian (equal with their conjugate transpose), because $P_i^* = (|a_i\rangle \langle a_i|)^* = \langle a_i|^* |a_i\rangle^* = |a_i\rangle \langle a_i| = P_i$. Hence

$$\langle \psi| P_i^* P_i |\psi\rangle = \langle \psi|a_i\rangle \langle a_i|a_i\rangle \langle a_i|\psi\rangle = (\langle a_i|\psi\rangle)^2$$

since $|a_i\rangle$ is a unit vector ($\langle a_i|a_i\rangle = 1$). So we have proved the equivalence. We also see that $P_i^2 = P_i$ and $P_i P_j = 0$. The operators with these properties are also called orthogonal projections.

Using orthogonal projections, we can define a more special kind of measurement than POVM, called projective measurement. This measurement is made by a family of orthogonal projections $\{P_i\}$ and is described by another operator, which is called observable.

An observable represents a physical quantity that can be measured in a quantum system, like energy, momentum, spin. It is associated with a Hermitian matrix A. The interpretation of this matrix is that its eigenvalues are the different outcomes of measuring this quantity and the eigenvectors are the states that the system can collapse to, if the corresponding eigenvalue is measured. Thus, if A has eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ (which are all real, because A is Hermitian) and eigenvectors $|a_1\rangle, |a_2\rangle, ..., |a_n\rangle$, we can write A as

$$A = \sum_{i=1}^{n} \lambda_i P_i$$

where $P_i = |a_i\rangle \langle a_i|$ is the projection on the eigenspace of $|a_i\rangle$. This is called the spectral decomposition.

Using the above relation, we can get

$$\langle \psi| A |\psi\rangle = \sum_{i=1}^{n} \lambda_i \langle \psi| P_i |\psi\rangle = \sum_{i=1}^{n} p(i)\lambda_i$$

which is the expected value of the outcome.

2.3.4 Pauli Matrices

Pauli matrices are 4 qubit observables that have some nice properties and are fundmental for the study of quantum information theory. They are defined by:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Below there are some useful properties:

1. They are unitary, Hermitian and their square equals the identity matrix. So, for all $A \in \{I, X, Y, Z\}$,

   - $AA^* = I$
   - $A = A^*$
   - $A^2 = I$

- $\lambda(A) = \pm 1$

2. X, Y, Z anti-commute pairwise

$$XY = -YX \quad YZ = -ZY \quad ZX = -XZ$$

3. $Y = iXZ, \quad X = iZY, \quad Z = iYX$

4. Each one of X, Y, Z corresponds to a projective measurement:

   - The observable Z is equivalent to measuring in the standard basis, because it is a projection onto $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ with eigenvalue 1 and onto $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ with eigenvalue -1.
   - The observable X is equivalent to measuring in the Hadamard basis, because it is a projection onto $|+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ with eigenvalue 1 and onto $|-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ with eigenvalue -1.
   - The observable Y is a projection onto $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$ with eigenvalue 1 and onto $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$ with eigenvalue -1.

2.4 Density Matrix

Thus far, we have seen how the quantum states can be represented as vectors and what actions we can do on these vectors. However, there is another equivalent representation for quantum states, which is often more helpful; it is the density matrix. Suppose that the state of the quantum system is not completely known, but can be in one of a number of states $|\psi_i\rangle$ with probability p_i $(i = 1, 2, ..., k)$. Then, we say that the system is in a mixed state $\{p_i, |\psi_i\rangle\}$. We can convert this mixed state into a matrix (or operator) which is called the density matrix (operator) and is defined by

$$\rho = \sum_{i=1}^{k} p_i |\psi_i\rangle \langle\psi_i|$$

It can be proved that all the properties of vector states also hold for the density matrices. We will see how we can reformulate them.

At first, if we have a quantum state and we want it to change using a unitary matrix U, then a state $|\psi\rangle$ will become $U|\psi\rangle$. Analogously, we can have

$$\rho = \sum_{i=1}^{k} p_i |\psi_i\rangle \langle\psi_i| \xrightarrow{U} \sum_{i=1}^{k} p_i U |\psi_i\rangle \langle\psi_i| U^* = U\rho U^*$$

We can generalise the operation $\rho \to U\rho U^*$ with maps from the space of density matrices with dimension n to the space of density matrices with dimension m. We call these maps quantum channels and they are the most general way to represent evolution of quantum systems. Quantum channels must preserve the properties of density matrices. That is why they must be completely positive and trace preserving, which is denoted as $CPTP(A, B)$, where A is the input quantum system and B the output quantum system.

We can also reformulate measurements. If we want to perform a measurement with measurement operators M_m, then $p(m)_i = \langle\psi_i| M_m^* M_m |\psi_i\rangle = Tr(M_m^* M_m |\psi_1\rangle\langle\psi_i|)$. To have the last equality, consider an orthonormal basis to the vector space that $|\psi\rangle$ belongs, which is $|0\rangle, |1\rangle, ..., |n\rangle$ with some i such that $|i\rangle = |\psi\rangle$. Then

$$Tr(A|\psi\rangle\langle\psi|) = \sum_i \langle i| A |\psi\rangle \langle\psi|i\rangle = \langle\psi| A |\psi\rangle$$

Therefore, for the mixed state $\{p_i, |\psi_i\rangle\}$, we have

$$p(m) = \sum_{i=1}^k p_i p(m|i) = \sum_{i=1}^k p_i p(m)_i = \sum_{i=1}^k p_i Tr(M_m^* M_m |\psi_1\rangle\langle\psi_i|) = Tr(M_m^* M_m \rho)$$

Similarly, we see that the post-measurement state, if we have m as output of the measurement, will be

$$\rho_m = \frac{M_m \rho M_m^*}{Tr(M_m^* M_m \rho)}$$

We can identify some properties of the density matrices.

1. It is easy to see that ρ is a square matrix, since $|\psi\rangle\langle\psi|$ is always square.

2. ρ is Hermitian, because

$$\rho^* = \left(\sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i|\right)^* = \sum_{i=1}^k p_i (|\psi_i\rangle\langle\psi_i|)^* = \sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i| = \rho$$

3. $Tr(\rho) = 1$. We can show that $Tr(|\psi\rangle\langle\psi|) = 1$ easily by taking the trace in respect to a basis which includes $|\psi\rangle$. So

$$Tr(\rho) = Tr\left(\sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_{i=1}^k p_i Tr(|\psi_i\rangle\langle\psi_i|) = \sum_{i=1}^k p_i = 1$$

4. All the eigenvalues of a density matrix are non-negative or equivalently ρ is a positive-semidefinite matrix. Suppose that for the eigenvalue λ, we have $\rho|v\rangle = \lambda|v\rangle \iff \lambda = \langle v|\rho|v\rangle$. Then

$$\lambda = \sum_{i=1}^k p_i \langle v|\psi_i\rangle\langle\psi_i|v\rangle = \sum_{i=1}^k p_i (\langle v|\psi_i\rangle)^2 \geq 0$$

5. We can combine different states with the tensor product. So, if we have the different quantum states $\rho_1, \rho_2, ..., \rho_n$, then their common state is $\rho = \rho_1 \otimes \rho_2 \otimes ... \otimes \rho_n$.

6. We can write the density matrix for a qubit (2×2) as a linear combination of Paulis:

$$\rho = \frac{1}{2}(I + r_x X + r_y Y + r_z Z)$$

where r_x, r_y, r_z are real numbers and are the coordinates of a point in the unit ball. We can generalise this and write any density matrix of a quantum system as a linear combination of tensor combinations of Paulis.

Conversely, suppose that ρ is a positive-semidefinite square matrix with trace 1. Then, all its eigenvalues are in $[0, 1]$ and by the spectral theorem

$$\rho = \sum_i \lambda_i \left|i\right\rangle \left\langle i\right|$$

which is the mixed state $\{\lambda_i, \left|i\right\rangle\}$, so ρ is a density matrix.

Finally, if we have a composite quantum system, we can use the density matrix to find the state of only some of the subsystems. This is called the reduced state operator. Suppose we have the systems A and B and they are in the state ρ^{AB}. Then the reduced state operator for system A is

$$\rho^A = Tr_B(\rho^{AB})$$

Tr_B is known as the partial trace over B and is defined by

$$Tr_B(\left|a_1\right\rangle \left\langle a_2\right| \otimes \left|b_1\right\rangle \left\langle b_2\right|) = \left|a_1\right\rangle \left\langle a_2\right| Tr(\left|b_1\right\rangle \left\langle b_2\right|) = \left\langle b_2|b_1\right\rangle \left|a_1\right\rangle \left\langle a_2\right|$$

where $\left|a_1\right\rangle, \left|a_2\right\rangle$ are possible states of the subsystem A and $\left|b_1\right\rangle, \left|b_2\right\rangle$ are possible states of the subsystem B. We can generalise this definition for any density matrix by using linearity, like in the common trace operator.

Chapter 3

Nonlocal Games

3.1 EPR Paradox and Nonlocal Games

In a 1935 paper titled "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?", the physicists Albert Einstein, Boris Podolsky and Nathan Rosen [22] proposed a thought experiment, with which they argued that quantum mechanics is an incomplete theory. Specifically, they argued there are some phenomena that cannot be predicted with quantum mechanics and there must be a theory that can contain them, too.

The thought experiment they described is as follows:

1. There are two participants, which we will call Alice and Bob, and they share a pair of qubits in an entangled state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. This is called an EPR pair. We consider that the two participants cannot communicate with each other during the experiment. For example, we can say that they are at planets light years away, so the communication is impossible during the experiment.

2. Alice can measure her qubit in the standard or Hadamard basis. If she chooses the standard basis, then she gets $|0\rangle$ or $|1\rangle$ with equal probability, and the overall state collapses to $|00\rangle$ or $|11\rangle$, respectively. In a similar manner, if she chooses the Hadamard basis, then she gets $|+\rangle$ or $|-\rangle$ with equal probability, and the overall state collapses to $|{+}{+}\rangle$ or $|{-}{-}\rangle$, respectively.

3. As a result, if Bob chooses to measure the state on the same basis with Alice, he will get the same outcome as her.

This thought experiment presented according to the authors of the paper a paradox, because the action on Alice's qubit could affect Bob's qubit. In other words, Bob could predict Alice's outcome, even if he was light years away in distance, which meant that information was transmitted faster than light and this is forbidden by the theory of relativity.

The EPR paradox showed that quantum mechanics is a nonlocal theory and suggested that this incompleteness can be solved by a hidden variable theory. However, in 1964, John Stewart Bell showed that there is no local theory, even with hidden variables, that can describe some correlations that arise from quantum mechanics [23], [24]. In fact, Bell's theorem states that the set of quantum correlations is strictly greater than the set of classical correlations (that arise from classical physics).

A simple way to prove Bell's theorem is by proposing a type of games which are called nonlocal games. In this type of games, there is a referee and several players, who can share entanglement and cannot communicate with each other (this is where nonlocality comes from). The referee sends randomly chosen questions to each players from a set of possible questions and each player can answer from a set of possible answers. At the end the referee uses a verification function, which has as parameters the questions they sent and the answers they received, and decides if they will accept, which means that the players win, or reject, which means that they lose.

The players choose some strategy before the game begins. A deterministic strategy means that for some specific question, the player responds with a specific answer. Formally, we can represent this type of strategies with a family of functions $\{f_i(x_i)\}$ with i an index over the players, x_i the question on player i and f_i a function from the set of possible questions to the set of possible answers. Each function has only the input of the player as parameter because of the no communication condition. We can extend the deterministic strategies with the probabilistic ones, where the players can use randomness for their answers. We can suppose that the randomness is shared, because if it is private then the players' functions can use only some specific parts of the random string. So, in probabilistic strategies, we have the family of functions $\{f_i(x_i, r)\}$, where r is a shared random string. These strategies are classical. The quantum strategy is represented with a shared state $|\psi\rangle$ and for each player a set of observables $\mathcal{O}_{i,x_i}$. That means that a player depending on their input question x_i chooses to measure their own part of the shared state $|\psi\rangle$ with the corresponding observable. Each eigenvalue (or equivalently outcome) of the observable corresponds to a possible answer.

In this way, we can define two values, the classical value of the game, which is the maximum winning probability for the players if they can only use classical strategies, and the quantum value, which is the maximum winning probability for the players if they use a quantum strategy. The former is denoted as $\omega_c(G)$ and the latter $\omega_q(G)$ for some game G. Bell's theorem or Bell's inequality actually says that there exists a game G such that $\omega_q(G) > \omega_c(G)$. Next, we will see 2 such games, the CHSH game and the Magic Square game.

3.2 The CHSH Game

One instance of this type of games is called the CHSH game. It is named after John Clauser, Michael Horne, Abner Shimony and Richard Holt [25], who described it in a paper published in 1969. The game is described below:

1. The referee selects uniformly random bits $x, y \in \{0, 1\}$ and sends x to Alice and y to Bob.

2. Alice and Bob using a strategy that they have agreed before the game, answer to the referee by sending bits a and b respectively.

3. At the end, the referee accepts if and only if the XOR of a, b equals the AND of x, y ($a + b \equiv xy \ (mod\ 2)$).

We will see that if the players choose a classical strategy then the maximum winning probability 75% ($\omega_c(CHSH) = \frac{3}{4}$). On the other hand, there is a quantum strategy that uses entanglement and has winning probability around 85% ($\omega_q(CHSH) = \cos^2(\frac{\pi}{8})$).

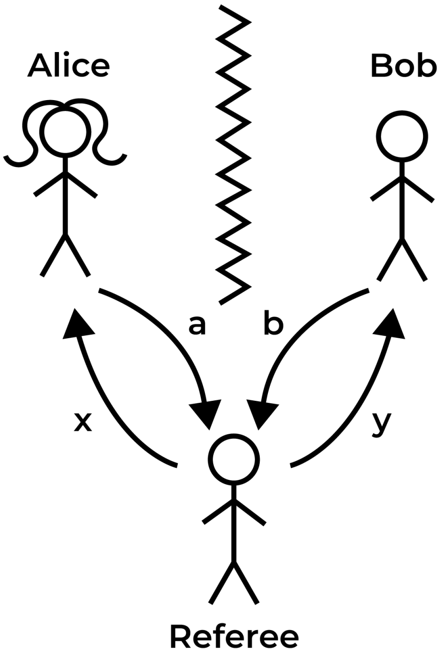


Figure 3.1: In the CHSH game, the referee sends the bit x to Alice and the bit y to Bob, and they must answer with a bit a and a bit b respectively without communicating. They win if $a + b \equiv xy\ (mod\ 2)$.

3.2.1 Classical Strategies for the CHSH Game

We start with the case of deterministic strategies, which means that the outputs a, b of Alice and Bob is a function of their respective inputs (since no communication is allowed). So, we have $a = f(x)$ and $b = g(y)$. We have four possible cases of inputs:

- when $x = 0$ and $y = 0$, then the players win if and only if $f(0) + g(0) \equiv 0\ (mod\ 2)$.
- when $x = 0$ and $y = 1$, then the players win if and only if $f(0) + g(1) \equiv 0\ (mod\ 2)$.
- when $x = 1$ and $y = 0$, then the players win if and only if $f(1) + g(0) \equiv 0\ (mod\ 2)$.
- when $x = 1$ and $y = 1$, then the players win if and only if $f(1) + g(1) \equiv 1\ (mod\ 2)$.

If we sum all the above relations, we get $0 \equiv 1\ (mod\ 2)$. So, we see that we cannot choose functions $f(x)$ and $g(y)$ such that all four of these equations hold, but only at most three of them. Thus, in this case the probability of winning is at most 75%.

However, Alice and Bob can use randomness in their strategy and this is still classical. So we have the case of probabilistic strategies. In this case, we consider that the players have access to a shared random variable r, which could be generated before the experiment. They can use this variable to choose their respective outputs, so we have that $a = f(x, r)$ and $b = g(y, r)$. The assumption of shared randomness also covers the case of private randomness, because the functions of Alice and Bob can only use some independent parts of the bits of r.

In this case the expected winning probability is

$$Pr[winning] = \sum_{r \in R} Pr(r) \cdot \omega(f(x,r), g(y,r))$$

where $\omega(f(x,r), g(y,r))$ denotes the probability of winning with the deterministic functions $f(x,r)$ and $g(y,r)$.

As a result, any probabilistic strategy is a linear combination of deterministic strategies, which means that the maximum winning probability is again 75%.

Thus, we conclude that for any classical strategy, the winning probability is at most 75%.

3.2.2 Quantum Strategy for the CHSH Game

Next, we will provide a quantum strategy with greater winning probability than the classical one. To begin with, we consider that before playing the game, Alice and Bob have shared the two qubits of an EPR pair. They have also agreed to measure their qubit with the following observables (denoted A_0, A_1, B_0, B_1):

Alice will measure in the standard basis, if she receives $x = 0$, which means $A_0 = Z$, and in the Hadamard basis, if she receives $x = 1$, which means $A_1 = X$.

Bob will measure in different basis. If he receives $y = 0$, we consider a rotation of $\pi/8$ to the standard basis, so the new basis vectors are

$$\cos\left(\frac{\pi}{8}\right)|0\rangle + \sin\left(\frac{\pi}{8}\right)|1\rangle, \quad \cos\left(\frac{5\pi}{8}\right)|0\rangle + \sin\left(\frac{5\pi}{8}\right)|1\rangle$$

The corresponding observable is $B_0 = \frac{1}{\sqrt{2}}(X + Z)$.

If he receives $y = 1$, then we have a rotation of $-\pi/8$ to the standard basis, so as the new basis vectors to be

$$\cos\left(-\frac{\pi}{8}\right)|0\rangle + \sin\left(-\frac{\pi}{8}\right)|1\rangle, \quad \cos\left(\frac{3\pi}{8}\right)|0\rangle + \sin\left(\frac{3\pi}{8}\right)|1\rangle$$

The corresponding observable is $B_0 = \frac{1}{\sqrt{2}}(Z - X)$.

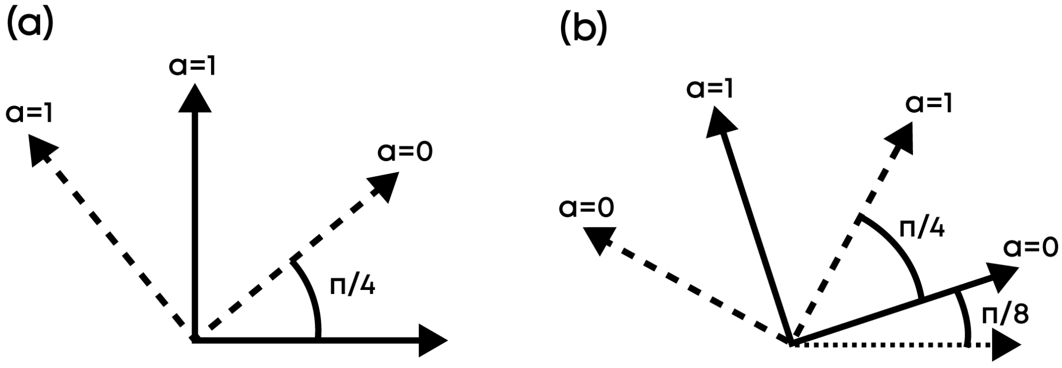


Figure 3.2: The different basis used for the quantum strategy of the CHSH game: (a) is for Alice and (b) is for Bob

(Note: When we measure in a binary basis with vectors $|a\rangle, |b\rangle$, then these vectors are the eigenvectors of the corresponding observable and the respective eigenvalues are 1 and -1. So, from the spectral theorem the observable $\mathcal{O}$ is

$$\mathcal{O} = 1 \cdot |a\rangle \langle a| - 1 \cdot |b\rangle \langle b|$$

Using this relation, it is easy to compute A_0, A_1, B_0, B_1.)

Now let us see the outcomes of the two players. We will use the eigenvalues, so outcome 1 means that the output bit is 0 and outcome -1 means that the output bit is 1.

- If the inputs are $x = y = 0$, then Alice measures in the standard basis and gets $|0\rangle$ or $|1\rangle$ with probability 1/2. If Alice gets $|0\rangle$ (outcome corresponding to the eigenvalue 1), then Bob's qubit collapses to $|0\rangle$, as well. So measuring with B_0, there is $\cos^2(\frac{\pi}{8})$ probability that he also gets outcome 1 ($\cos^2(\frac{\pi}{8})$ is the squared inner product of $|0\rangle$, which is the state of the system, and $\cos\left(\frac{\pi}{8}\right)|0\rangle + \sin\left(\frac{\pi}{8}\right)|1\rangle$, which is the eigenvector of 1). Respectively, if Alice gets $|1\rangle$ (eigenvalue -1), then Bob's state is also $|1\rangle$ and the probability of measuring -1 with B_0 is $\sin^2(\frac{5\pi}{8}) = \cos^2(\frac{\pi}{8})$. So, the overall success probability in this case is $\cos^2(\frac{\pi}{8})$.

- If the inputs are $x = 0, y = 1$, then Alice measures in the standard basis here as well. If she gets $|0\rangle$, then Bob measuring $|0\rangle$ with B_1 gets outcome 1 with probability $\cos^2(-\frac{\pi}{8}) = \cos^2(\frac{\pi}{8})$. Similarly, they both get outcomes -1 with probability $\sin^2(\frac{3\pi}{8}) = \cos^2(\frac{\pi}{8})$. So, again the success probability is $\cos^2(\frac{\pi}{8})$.

- If the inputs are $x = 1, y = 0$, then Alice measures in Hadamard basis, so she gets $|+\rangle$ or $|-\rangle$ with equal probability. If she gets $|+\rangle$ (outcome 1), Bob's state is also $|+\rangle$, so measuring with B_0 the probability of outcome 1 is

  $$\left(\frac{\cos\left(\frac{\pi}{8}\right)}{\sqrt{2}} + \frac{\sin\left(\frac{\pi}{8}\right)}{\sqrt{2}}\right)^2 = \frac{1+\sin\left(\frac{\pi}{4}\right)}{2} = \frac{1+\cos\left(\frac{\pi}{4}\right)}{2} = \cos^2\left(\frac{\pi}{8}\right)$$

  Similarly, if Alice gets $|-\rangle$ (outcome -1), then Bob's state is also $|-\rangle$ and measuring with B_0 gets outcome -1 with probability

  $$\left(\frac{\cos\left(\frac{5\pi}{8}\right)}{\sqrt{2}} - \frac{\sin\left(\frac{5\pi}{8}\right)}{\sqrt{2}}\right)^2 = \frac{1-\sin\left(\frac{5\pi}{4}\right)}{2} = \frac{1+\cos\left(\frac{\pi}{4}\right)}{2} = \cos^2\left(\frac{\pi}{8}\right)$$

  So, the success probability is again $\cos^2(\frac{\pi}{8})$.

- If the inputs are $x = 1, y = 1$, then Alice measures in Hadamard basis here as well. However, in this case the two players must have opposite outcomes. In the case of Alice getting $|+\rangle$ (outcome 1), then the probability of Bob getting outcome -1 by measuring $|+\rangle$ with B_1 is $\left(\frac{\cos\left(\frac{3\pi}{8}\right)}{\sqrt{2}} + \frac{\sin\left(\frac{3\pi}{8}\right)}{\sqrt{2}}\right)^2 = \cos^2(\frac{\pi}{8})$. Similarly, if Alice gets $|-\rangle$ (outcome -1), then Bob gets outcome 1 with probability $\left(\frac{\cos\left(-\frac{\pi}{8}\right)}{\sqrt{2}} - \frac{\sin\left(-\frac{\pi}{8}\right)}{\sqrt{2}}\right)^2 = \cos^2(\frac{\pi}{8})$.

To sum up, the success probability is $\cos^2(\frac{\pi}{8})$ in each case, so it is also generally the success probability for this strategy.

3.2.3 Optimality of $\cos^2(\frac{\pi}{8})$

It may seem that the value $\cos(\frac{\pi}{8})^2 = \frac{1}{2} + \frac{1}{2\sqrt{2}}$ is arbitrary, but it turns out that it is the maximal value we can achieve with a quantum strategy ($\omega_q(CHSH) = \cos^2(\frac{\pi}{8})$). This is called the Tsirelson's Inequality and it was proved in 1980 by Boris Tsirelson [26].

Proof. We can model a quantum strategy with the triple $S = (|\psi\rangle, A, B)$, where $|\psi\rangle \in \mathcal{C}^d \otimes \mathcal{C}^d$ is a quantum state shared between Alice and Bob with d being the dimension of their respective system, and A_x, B_y are the observables on the two systems, depending on the inputs x, y. The observables can have only two different outcomes, so their eigenvalues are 1 (for output 0) and -1 (for output 1) and we can decompose them like $A_x = A_{x,0} - A_{x,1}$ with $A_{x,0}, A_{x,1}$ being orthogonal projectors with sum the identity matrix ($B_x = B_{x,0} - B_{x,1}$ respectively).

Thus the probability of winning the CHSH game can be expressed as:

$$\omega_q(CHSH, S) = \sum_{x,y} Pr(x,y) \sum_{a,b:a\oplus b = x\wedge y} Pr(a,b|x,y) = \frac{1}{4} \sum_{a,b,x,y:a\oplus b=x\wedge y} \langle\psi| A_{x,a} \otimes B_{y,b} |\psi\rangle$$

Using the relations for the observables, we can reduce $\omega_q(CHSH, S) \leq \frac{1}{2} + \frac{1}{2\sqrt{2}}$ to:

$$\langle\psi| A_0 \otimes B_0 + A_1 \otimes B_0 + A_0 \otimes B_1 - A_1 \otimes B_1 |\psi\rangle \leq 2\sqrt{2}$$

To prove this, we define $C_0 = \frac{B_0+B_1}{\sqrt{2}}$ and $C_0 = \frac{B_0-B_1}{\sqrt{2}}$, so we can rewrite this as:

$$\langle\psi| A_0 \otimes C_0 + A_1 \otimes C_1 |\psi\rangle \leq 2$$

To simplify this relation, we can use the spectral norm of the operator T, which we denote as $\|T\|$ and is equal to the maximum singular value of T or equivalently the maximum value of $\langle\psi_1| T |\psi_2\rangle$ for unit vectors $|\psi_1\rangle, |\psi_2\rangle$. So, $\langle\psi| T |\psi\rangle \leq \|T\|$ and we need to prove that:

$$\|A_0 \otimes C_0 + A_1 \otimes C_1\| \leq 2$$

Then using $\|T\|^2 \leq \|T^2\|$, which is true for any operator, it suffices to prove:

$$\|(A_0 \otimes C_0 + A_1 \otimes C_1)^2\| \leq 4 \iff \|A_0^2 \otimes C_0^2 + A_1^2 \otimes C_1^2 + A_0A_1 \otimes C_0C_1 + A_1A_0 \otimes C_1C_0\| \leq 4$$

A_x and B_y are binary observables, so they square to identity ($A_x^2 = (A_{x,0} - A_{x,1})^2 = A_{x,0}^2 + A_{x,1}^2 - A_{x,0}A_{x,1} - A_{x,1}A_{x,0} = A_{x,0} + A_{x,1} = \mathbb{I}$, where we used that the square of a projector equals the projector and that $A_{x,0}, A_{x,1}$ are orthogonal and sum to identity).

Using this fact, we get

$$A_0^2 \otimes C_0^2 = \mathbb{I} \otimes \frac{1}{2}(B_0^2 + B_1^2 + B_0B_1 + B_1B_0) = \mathbb{I} \otimes (\mathbb{I} + \frac{1}{2}(B_0B_1 + B_1B_0))$$

$$A_1^2 \otimes C_1^2 = \mathbb{I} \otimes (\mathbb{I} - \frac{1}{2}(B_0B_1 + B_1B_0))$$

Hence,

$$\|(A_0 \otimes C_0 + A_1 \otimes C_1)^2\| = \|2\mathbb{I} \otimes \mathbb{I} + A_0A_1 \otimes C_0C_1 + A_1A_0 \otimes C_1C_0\|$$
$$\leq 2 + \|A_0A_1 \otimes C_0C_1\| + \|A_1A_0 \otimes C_1C_0\| = 2 + 2\|A_0A_1\| \cdot \|C_0C_1\|$$

where for the first inequality we used the triangle inequality and the fact that the norm of the identity is 1 and for the last equality $\|A \otimes B\| = \|A\| \cdot \|B\|$ and $\|AB\| = \|BA\|$.

We see that $\|A_0A_1\| \leq \|A_0\|\|A_1\| = \sqrt{\|A_0^2\|\|A_1^2\|} = 1$ and we also have

$$\|C_0C_1\| = \left\|\frac{1}{2}(B_0^2 - B_1^2 - B_0B_1 + B_1B_0\right\| = \frac{1}{2}\|B_1B_0 - B_0B_1\| \leq \frac{1}{2}(\|B_1B_0\| + \|B_0B_1\|) \leq 1$$

Using the last two inequalities we have the desired result. □

3.3 The Magic Square Game

In the previous section, we proved that the maximum winning probability for the CHSH game is around 85%. Someone may wonder if there is a game that the players cannot always win with classical strategies, but there is a quantum strategy so they have success probability 1. Well, there is! It is called the Magic Square Game or Mermin-Peres Magic Square [27], [28]. The setup is:

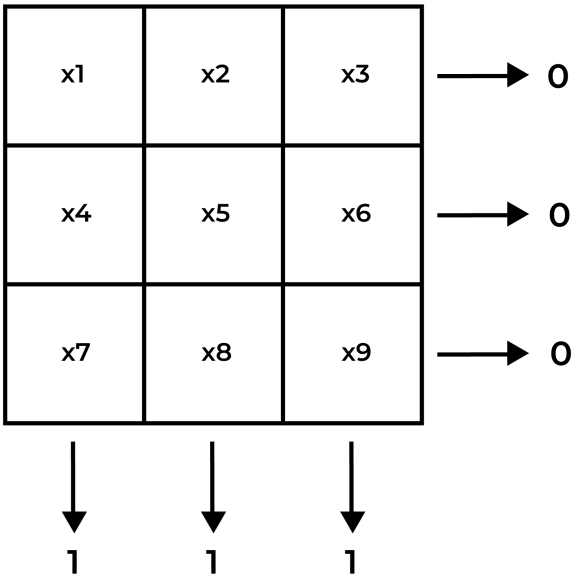
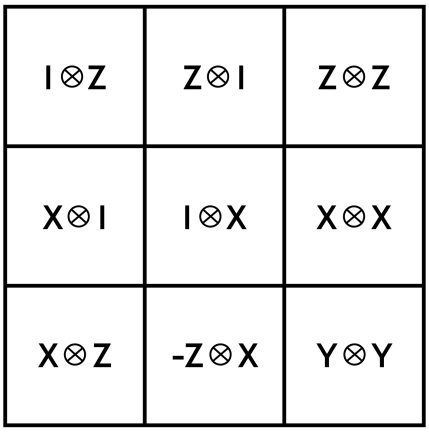


Figure 3.3: The Magic Square game setup: the sum of the binary variables at each row must be even (0 $(mod\ 2)$) and the sum of each column must be odd (1 $(mod\ 2)$).

We have a 3×3 grid, which has to be filled with 0's and 1's, such that the parity of every row is 0 and of every column is 1. This is impossible, since the parity of all the bits must be 0 and 1 at the same time, but we can convert it to a game. So, the referee randomly selects one row or column and sends it to Alice ($x \in \{r_1, r_2, r_3, c_1, c_2, c_3\}$) and then selects a random cell $y \in x$ and sends it to Bob. (Note: there is a variation of the game where the referee sends a random row to Alice and a random column to Bob, but the results are the same.) The players win, if Alice responds with three bits (a_1, a_2, a_3) that correspond to her input row or column and Bob responds with a bit b that corresponds to his input cell, such that the parity of a_1, a_2, a_3 is 0 if the input is row and 1 if it is column, and the bit b matches with Alice's bit in the corresponding cell.

We can use a similar argument with the CHSH game to see that Alice and Bob will fail in at least one of the 18 possible questions that can be asked, so the classical value of the game is 17/18. However, there is a quantum strategy that makes them succeed with probability 1.

The key observation for finding a perfect quantum strategy is that even though there is no solution for the system in $\mathbb{Z}_2$, there is an operator solution.



I⊗Z	Z⊗I	Z⊗Z
X⊗I	I⊗X	X⊗X
X⊗Z	-Z⊗X	Y⊗Y

Figure 3.4: The operator solution of the Magic Square game: each row has product $\mathbb{I} \otimes \mathbb{I}$ and each column has product $-\mathbb{I} \otimes \mathbb{I}$. The players use these observables to obtain their answer for a specific position on the grid.

Here, X, Y, Z are the Pauli matrices, for which we have $P^2 = \mathbb{I}$ with P any Pauli matrix, $Y = iXZ$ and $XZ = -ZX$. Using these properties, we see that the product of the observables in each row is the identity matrix $\mathbb{I}$ and the product in each column is $-\mathbb{I}$. Also, these observables have eigenvalues 1 and -1, which means that their output is binary, as we desire.

So, the two players can use two EPR pairs for their strategy, with each having one part of them. This means that the state they will use is:

$$|\psi\rangle = \left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\right)^{\otimes 2}$$

If both measure their respective part of the EPR pair with the same observable, we have already seen that they will get the same result. Thus if the players' measurements are consistent with the matrix above, then the outputs of Alice and Bob will be consistent, as well.

Moreover, the observables in each row and column commute, hence it does not matter which order Alice will choose to measure with her 3 observables and it will not affect the outcomes. Finally, using the product of the observables we see that the product of the outcomes will be 1 for each row and -1 for each column, which is the same as the parity constraint we have.

To sum up, Alice's answer will always satisfy the constraints because of the product of the operators she will use, and Bob's answer will be consistent with Alice's, because of the properties of the EPR pair. Thus, the players will win the game with probability 1 ($\omega_q(MS) = 1$).

Chapter 4

Randomness Expansion

4.1 Introduction

The task of generating independent random bits is very significant in modern-day computing, since it is used for various applications, such as cryptography, algorithms and physical simulations. The quality of randomness, which means how close the output distribution is to the desired distribution, is also very important for some applications, especially in cryptography, where the quality of randomness affects the security of a protocol. Thus, it is very important to construct physical devices that can produce reliably random bits.

However, it is impossible for classical computers to produce randomness. Since they use deterministic operations on their input, the entropy of the input will always be greater than the entropy of the output, which is called the data processing inequality. The best we can hope for is pseudorandomness generators, which produce a distribution that is relatively close to the desired one. There have also been some attempts to produce randomness from other physical sources, but the output was of low quality and it was very hard to test them [29].

Fortunately, quantum mechanics provides a way to both generate truly random bits and also test them. For the first part, we count on the fact that quantum mechanics has a probabilistic nature. For the second part, the main idea to verify that the source of randomness behaves quantumly is to use Bell's inequalities or nonlocal games. For example, if we have two non-communicating devices and play the CHSH game with them, then we can check their success probability and determine if they have used quantum correlations. So, in order to take this statistical test, the idea, which was first observed by Colbeck in his PhD thesis [11], was to play many times sequentially the CHSH game [16].

This method would require 2 random bits for the input and could result to 2 bits that contain less randomness (since there must be some correlation). However, using some rounds of the protocol with a predetermined input, we can expand the initial entropy of the input. In fact, we can expand exponentially the randomness of the seed (input) we use, which was first achieved by Vazirani-Vidick [16].

4.2 Preliminaries

4.2.1 Entropy

At first, we will need some definitions to gain a better perspective on randomness. Let X be a discrete random variable. We denote $supp(X)$ for the support set of X, which means all the values v such that $Pr(X = v) > 0$.

The Shannon entropy of X is denoted $H(X)$ and is defined as

$$H(X) = \sum_{x \in supp(X)} -Pr(X = x) \log(Pr(X = x))$$

We also define the max-entropy $H_0(X) = \log(|supp(X)|)$ and the min-entropy:

$$H_\infty(X) = -\log \max_{x \in supp(X)} Pr(X = x)$$

The Shannon entropy is a way to measure how much information is contained in a random variable, while the max-entropy is an upper bound for it (if all values have equal probabilities) and min-entropy shows how close the distribution of the random variable is to the uniform distribution. That is why we prefer to use the min-entropy as a measure of randomness.

The conditional min-entropy is defined as

$$H_\infty(X|Y) = -\log \left(\sum_y Pr(Y = y) 2^{-H_\infty(X|Y=y)} \right)$$

Moreover, for two discrete random variables X, Y, their statistical distance is

$$\|X - Y\|_1 = \sum_{u \in supp(Y) \cup supp(X)} \frac{1}{2} \Big| Pr_Y(u) - Pr_X(u) \Big|$$

With this we can define the smooth min-entropy of X for some $\epsilon > 0$ as

$$H_\infty^\epsilon(X) = \sup_{X' : \|X - X\|_1 \leq \epsilon} H_\infty(X')$$

and the smooth conditional min-entropy as

$$H_\infty^\epsilon(X|Y) = \sup_{X', Y' : \|(X,Y) - (X',Y')\| \leq \epsilon} H_\infty(X'|Y')$$

As we can see the smoothness is used to measure the maximum min-entropy that is approximately close to the initial distribution.

We will use the smooth min-entropy as the measure of randomness produced. The smoothness is used, because it is sufficient to be ϵ-close to the desired probabilities (it is cryptographically secure for small enough ϵ) and the min-entropy generally shows the number of uniformly random bits that can be extracted from the distribution of outputs. This extraction of random bits is achieved with a procedure called extractor (which is described in [30]). Extractor is a deterministic algorithm that takes an input with min-entropy n and outputs an n-bit string that is (close to) uniformly random. The extractor requires an extra seed of uniformly random bits (apart from the seed used in the protocol as described below), but we will not take this into account. There are some standard extractors that can be used, like Trevisan's extractor [31].

4.2.2 Concentration Inequalities

Lemma 4.2.1. (Chernoff Bound)

Let $X_1, X_2, ..., X_n$ be independent Bernoulli (taking values only 0 or 1) random variables with expected value μ. Then for any $\delta > 0$, it holds that

$$Pr\left(\sum_{i=1}^{n} X_i \leq (1-\delta)\mu n\right) \leq e^{-\delta^2 \mu n/2}$$

$$Pr\left(\sum_{i=1}^{n} X_i \geq (1+\delta)\mu n\right) \leq e^{-\delta^2 \mu n/(2+\delta)}$$

Combining both of them, we have for any $\delta \in [0, 1]$:

$$Pr\left(\left|\sum_{i=1}^{n} X_i - \mu n\right| \geq \delta \mu n\right) \leq 2e^{-\delta^2 \mu n/3}$$

Lemma 4.2.2. (Hoeffding's Inequality)

Let $X_1, X_2, ..., X_n$ be independent random variables, such that $Pr(X_i \in [a_i, b_i]) = 1$. Let μ be the expected value of their sum:

$$\mu = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right]$$

Then, for any $t > 0$, we have

$$Pr\left(\left|\sum_{i=1}^{n} X_i - \mu\right| \geq t\right) \leq 2e^{-2t^2/\sum_i (b_i - a_i)^2}$$

Lemma 4.2.3. (Azuma's inequality)

Consider a sequence of random variables $X_1, X_2, X_3...$ such that $\mathbb{E}[X_n] < \infty$ and

$$\mathbb{E}[X_{n+1}|X_1, ..., X_n] = X_n$$

for any $n \in \mathbb{N}$. Then this sequence is called a martingale.

If we consider another sequence of random variables $Y_1, Y_2, Y_3...$ and it holds that $\mathbb{E}[X_n] < \infty$ and $\mathbb{E}[X_{n+1}|Y_1, ..., Y_n] = X_n$ for any $n \in \mathbb{N}$, then the sequence $\{X_k\}$ is a martingale with respect to the sequence $\{Y_k\}$.

Suppose that the sequence $X_1, X_2, X_3...$ is a martingale with $Pr(|X_k - X_{k-1}| \leq c_k) = 1$. Then for any $n \in \mathbb{N}$ and $\epsilon > 0$, it holds that

$$Pr(X_n - X_1 \geq \epsilon) \leq e^{-\epsilon^2/2\sum_{i=1}^{n} c_i^2}$$

$$Pr(X_n - X_1 \leq -\epsilon) \leq e^{-\epsilon^2/2\sum_{i=1}^{n} c_i^2}$$

Combining both of them, we get

$$Pr(|X_n - X_1| \geq \epsilon) \leq 2e^{-\epsilon^2/2\sum_{i=1}^{n} c_i^2}$$

4.2.3 Randomness Expansion Protocols

We will define the properties of a randomness expansion protocol. The randomness expansion protocols we refer to here will use some 2-player nonlocal game G with input alphabets $\mathcal{X}, \mathcal{Y}$, output alphabets $\mathcal{A}, \mathcal{B}$ and $\omega_q(G) > \omega_c(G)$. In order to carry it out, the referee/user must communicate with two (or more) players/devices (we use both definitions below), which cannot communicate with each other.

Two of the main parameters of these protocols are the seed length m and the number of rounds n. The number of rounds is how many times the players play the nonlocal game and the seed length is the length of the input bit-string to the referee, who produces the inputs to these repetitive games with a deterministic procedure from this specific seed. We require the seed to be (close to) uniformly random, so the size of m is approximately the entropy of the input. In order to have randomness expansion, we must have output randomness (entropy), which is greater than the size of m (which is not the case in pseudorandomness generators).

Moreover, for every protocol there is a test T that the referee uses to decide if the players succeed in the protocol (which is the event we will denote as WIN) or fail. The most usual test is the product test, in which the referee checks every round of the game separately and then uses these statistics with another function g to determine the event WIN. For the protocol that we will define later, we will use this type of tests, but for the upper bound we will consider the general case.

The completeness c of the protocol is a real number in $(0, 1]$, such that if the two players play with the ideal strategy for the game, then $Pr(WIN) \geq c$. We say that the completeness holds with quantum devices, if this ideal strategy can be implemented with quantum devices. We usually want the completeness to be exponentially close to 1 in n.

We also define the soundness s as the real number in $(0, 1]$, such that if playing with whichever strategy resulting $Pr(WIN) \geq s$, guarantees that $g \leq H^{\epsilon}_{\infty}(A, B|X, Y, WIN)$, where g is called the expansion and it is the lower bound on the output randomness. The parameter ϵ is the smoothness of the protocol and is defined above. We say that the soundness holds against quantum devices, if the strategy can be any quantum (or classical) strategy. We usually want the soundness and smoothness to be exponentially small in the seed length m.

Another important property of these protocols is robustness. Since ideal execution of any protocol without errors and noise is not possible in the real world (especially with current quantum technology), the referee's test must accept the output of the devices with high probability, even if there are some small deviations from the ideal strategy. So, if we consider the distribution produced by the ideal strategy with inputs x, y, $S_{ideal}(x, y)$, then another strategy $S(x, y)$ is η-close to the ideal if their statistical distance is at most η ($\|S_{ideal}(x, y) - S(x, y)\|_1 \leq \eta$). A protocol is η-robust, if it accepts with high probability any strategy which is η-close to the ideal one. In statistical tests, we usually have robustness as an error tolerance parameter to the expected value of the test's output.

Finally, we say that a protocol is non-adaptive, if the user cannot use the output of previous rounds to determine the inputs to future ones. In this chapter, we will see an exponential lower bound and two doubly exponential upper bounds for non-adaptive protocols. In the next chapter, we will present an adaptive protocol that achieves unbounded randomness with an extra device.

4.3 Exponential Randomness Expansion Protocol

The Vazirani-Vidick protocol used the CHSH game to expand randomness exponentially [16]. Subsequent works from Miller and Shi [18],[32],[33] provided more robust and cryptographically secure results. Also, in the paper from Arnon-Friedman, Renner and Vidick [34], a special theorem called the entropy accumulation theorem [35] was used to prove the security of this kind of protocols. These results also cover the hard case of a quantum adversary, which means that a quantum device that may be entangled with the devices used tries to extract information about the random string used. Our proof here will only cover the case of a classical adversary, but we will prove a stronger result in the next chapter, which uses the generalised version of the entropy accumulation theorem. The protocol used here is from Coudron, Vidick and Yuen [15].

All these protocols have similar intuition. There are some rounds called "test rounds" and some that are called "generation rounds". At the test rounds, the inputs to the two players-devices has the same distribution as in the definition of the nonlocal game. These rounds are used by the referee to check if the players play honestly (using entanglement), so that certifies that the randomness is generated by quantum properties. At the generation rounds, the input used is predetermined and the outputs are used for the extracted randomness. The predetermined input is used for economy of the random bits provided by the seed.

4.3.1 Protocol

Protocol arguments

- G : two-player non-local game, specified by a question set $\mathcal{X} \times \mathcal{Y}$, a probability distribution q on $\mathcal{X} \times \mathcal{Y}$, an answer set $\mathcal{A} \times \mathcal{B}$, and a winning condition $\omega : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \rightarrow \{0, 1\}$
- $x^* \in \mathcal{X}$, $y^* \in \mathcal{Y}$: fixed inputs used for generation rounds
- D_1, D_2 : the two untrusted devices that will play the game G repeatedly
- $n \in \mathbb{N}$: number of rounds
- $\gamma \in (0, 1]$: expected fraction of test rounds
- ω_{exp} : expected winning probability in G
- δ : error tolerance

Protocol steps

For rounds $i = 1, ..., n$, the referee performs the following steps:

1. They choose $T_i \in \{0, 1\}$ with $Pr[T_i = 1] = \gamma$. If $T_i = 1$, the referee chooses $X_i, Y_i \in \mathcal{X} \times \mathcal{Y}$ according to the question distribution q. If $T_i = 0$, the referee chooses $X_i = x^*$, $Y_i = y^*$.
2. They send input X_i to Alice and Y_i to Bob. They receive answers A_i and B_i, respectively.
3. If $T_i = 0$, the referee sets $C_i = \perp$. If $T_i = 1$, they set $C_i = \omega(X_i, Y_i, A_i, B_i)$.

At the end of the protocol, the referee aborts if $|\{i \ \ s.t. \ \ C_i = 0\}| > (1 - \omega_{exp} + \delta) \cdot \gamma n$.

4.3.2 Proof

The whole analysis below is due to Coudron, Vidick and Yuen [15].

Definition 4.3.1. A two player nonlocal game will be characterized as $(p_0, \eta, 1-\xi)$-randomness generating against quantum players, if there is a specific input $x_0 \in \mathcal{X}$, such that the probability that it is input, is $q_X(x_0) = \sum_{y \in \mathcal{Y}} q(x_0, y) \geq p_0$ and for any quantum strategy with success probability at least $\omega_q - \eta$, it holds that

$$\max_{a \in \mathcal{A}} p(A = a | X = x_0) \leq 1 - \xi.$$

That means that for this specific input to Alice, Alice's output cannot be fixed - there must be some randomness in her answer. It is clear that we want to find the smallest possible ξ, because then the distribution of the outputs would be closer to uniformly random and farther from a deterministic response. It is proved that the CHSH game is $(1/2, \eta, 1/2 + \sqrt{3\eta})$-randomness generating, if we choose input $x_0 = 0$ and the Magic Square is $(1/9, \eta, 12/13 + \eta)$-randomness generating. Also it is clearly necessary that $\eta < \omega_q(G) - \omega_c(G)$.

Theorem 4.3.2. *Let G be a $(p_0, \eta, 1-\xi)$-randomness generating game against quantum players with input distribution q and m_q the number of bits required to sample from the distribution q. Also, let m be the seed length of the protocol, n the number of rounds, $\delta = p_0\eta/8$ the error tolerance and ϵ and s the smoothness and the soundness of the protocol respectively. Also, we assume that $\epsilon \leq s$ and $s\epsilon > e^{-C\min(\eta^2, p_0\xi^2)\gamma n}$, where C is a universal constant. Then the protocol defined in the previous section is a randomness expansion protocol with completeness $c \geq 1 - e^{-O(\delta^2 n)}$, soundness s, smoothness ϵ and expansion $g \geq \xi n/8$ and is robust for any constant less than δ.*

We can use this theorem to have exponential randomness expansion. The number of random bits needed to determine which rounds to be test rounds are $O(n \cdot h(\gamma))$, where $h(\cdot)$ is the binary entropy function. Also, considering $|\mathcal{A} \times \mathcal{B}|$ as constant, the input for the test rounds needs $O(\gamma n)$ random bits. Therefore the seed must have length $O(n \cdot h(\gamma) + \gamma n)$. If we choose some small γ, for instance $\gamma = O(1/n)$, then the seed has length $O(\log n)$ and the expansion is $\Omega(n)$, so we get exponential randomness expansion.

4.3.2.1 Completeness

To prove the completeness of the protocol, we use the Chernoff bound (Lemma 4.2.1). We consider players that play every round independently, so defining Z_i as a binary random variable such that $Z_i = 1$ if the players succeed in the i-th test round, then when the players play honestly (expected value of winning is $\omega_q(G)$):

$$Pr\left(\left|\sum_i Z_i - \omega_q(G)\gamma n\right| \geq \delta\gamma n\right) \leq 2e^{-\delta^2\gamma n/3\omega_q(G)}$$

4.3.2.2 Soundness

Theorem 4.3.3. *Suppose for the constants ζ with $1/2 \geq \zeta \geq 2\gamma$, $\eta > 0$ and $s \geq \epsilon > 0$ it holds that*

$$\frac{\log(16/(\epsilon^2 s))}{n} < \frac{\min(p_0\zeta^2, \eta^2)\gamma}{30}$$

and also we have

$$H^{\epsilon}_{\infty}(A, B|X, Y, WIN) \leq \zeta n$$

and

$$Pr(WIN) \geq s$$

where WIN is the event that the two players pass the test of the referee.

Then, there must exist two quantum devices that can play a single round of G, and an output $a_0 \in \mathcal{A}$ *such that when the two devices play the game G, then*

$$Pr(WIN) \geq \omega_q(G) - 8\delta/p_0$$

and

$$Pr(A = a_0|X = x_0) \geq 1 - 8\zeta.$$

In other words, it holds that the response of these two devices is close to being deterministic. If we set $\zeta = \xi/8$ and using $\delta \leq p_0\eta/8$, we see that we have a contradiction with the game G being $(p_0, \eta, 1 - \xi)$-randomness generating. Therefore, we have the desired result.

Proof. We define the probability space for the execution of the protocol as $\Omega = \{(x, y, a, b, u) \in (\{0, 1\}^5)^n\}$, where (x, y) are the strings of inputs from the referee to the devices, (a, b) are the strings of outputs of the devices and u is a string with $u_i = 1$ for test rounds and $u_i = 0$ for generation rounds. We will mark the corresponding random variables with capital letters, so for example U_i is the random variable that equals 1 if and only if $u_i = 1$. For one successful execution of the protocol, we also define $W = \sum_i U_i$ and from the referee's test, since we have error tolerance δ, we get

$$\frac{1}{W} \sum_{i:U_i=1} \mathbf{1}[WIN_i] \geq \omega_q(G) - \delta$$

where $\mathbf{1}[WIN_i] = 1$, if the players succeed in the ith round and 0, if they fail.

We want to study only the test rounds with input x_0 (so we can use the property of randomness generating games). We denote with W' the number of rounds with $u_i = 1$ and $x_i = x_0$. Thus, we can eliminate some extreme cases using the Chernoff bound (Lemma 4.2.1):

$$Pr\left(|W - \gamma n| \geq \frac{\gamma n}{3}\right) \leq e^{-\gamma n/27} \leq \epsilon^2/4$$

and

$$Pr\left(|W' - p_0\gamma n| \geq \frac{p_0\gamma n}{3}\right) \leq e^{-p_0\gamma n/27} \leq \epsilon^2/4$$

where the last inequalities result from the condition on the parameters.

Next, we define the new event WIN', in which the players have passed the referee's test (event WIN) and also

$$|W - \gamma n| \leq \frac{\gamma n}{3}, \quad |W' - p_0\gamma n| \leq \frac{p_0\gamma n}{3}$$

Using the union bound and $\epsilon^2 \leq \epsilon/2$, we have that

$$Pr\left(|W - \gamma n| \leq \frac{\gamma n}{3} \wedge |W' - p_0\gamma n| \leq \frac{p_0\gamma n}{3}\right) \geq 1 - \epsilon^2/2 - \epsilon^2/2 \geq 1 - \epsilon/4 \tag{4.1}$$

Claim 4.3.4. *If X is a random variable and T an event such that* $Pr(T) \geq 1 - \beta$, *then we have that* $H^{\epsilon-2\beta}_{\infty}(X|T) \leq H^{\epsilon}_{\infty}(X)$.

Proof. We define the random variable Y that has the same distribution as X conditioned on T. Then there must be another random variable $\tilde{Y}$ on the probability space conditioned on T, such that $H_\infty(\tilde{Y}) = H_\infty^{\epsilon-2\beta}(Y) = H_\infty^{\epsilon-2\beta}(X|T)$ and $\left\|\tilde{Y} - Y\right\|_1 \leq \epsilon - 2\beta$. We can extend $\tilde{Y}$ to a new variable $\tilde{X}$ in an arbitrary way, under the condition $H_\infty(\tilde{X}) \geq H_\infty(\tilde{Y})$. Then, since $Pr(T) \geq 1-\beta$, we have $\left\|X - \tilde{X}\right\|_1 \leq (\epsilon - 2\beta)/(1-\beta) + \beta \leq \epsilon$ and we get the desired result. □

Using the above claim, the initial assumption that $H_\infty^\epsilon(A, B|X, Y, WIN) < \zeta n$ and (4.1), we get $H_\infty^{\epsilon/2}(A, B|X, Y, WIN') < \zeta n$. This means by the definition of smooth min-entropy that for any distribution q with $\|q - p\|_1 \leq \epsilon/2$, $H_\infty(A, B|WIN', X, Y)_q < \zeta n$. Here q and p are distribution on the probability space Ω' which is Ω conditioned on the event WIN'. So, if we define a set $S \subseteq \Omega'$ which contains all (x, y, a, b, u) such that $Pr((A, B) = (a, b)|(X, Y) = (x, y)) > 2^{-\zeta n}$ (we can derive this relation from the definition of min-entropy), then from $\|q - p\|_1 \leq \epsilon/2$, we get $Pr(S|WIN') \geq \epsilon/2$ (this is the measure on probability space Ω').

In order to find $Pr(S)$, we need to calculate $Pr(WIN')$ first. Since $Pr(WIN) \geq s$ and $s \geq \epsilon$, using the union bound we have

$$Pr(\overline{WIN'}) \leq 1 - s + \epsilon^2/4 + \epsilon^2/4 \leq 1 - s + \epsilon/2 \leq 1 - s/2 \iff Pr(WIN') \geq s/2$$

As a result, we get

$$Pr(S) = Pr(S|WIN')Pr(WIN') \geq \frac{\epsilon}{2} \cdot \frac{s}{2} = \frac{\epsilon s}{4}$$

We continue by proving the two claims below (4.3.5 and 4.3.6) that show some properties of the sequences $(x, y, a, b, u) \in S$.

Claim 4.3.5. *For all $(x, y, a, b, u) \in S$ except for a fraction of at most ϵ of them, it holds that*

$$\frac{1}{w'} \sum_{i \in [n], u_i = 1, x_i = x_0} Pr(A_i = a_i \mid X_i = x_0, (A, B, X, Y)_{<i} = (a, b, x, y)_{<i}) \geq 1 - 4\zeta$$

Proof. Since $(x, y, a, b, u) \in S$, $Pr((A, B) = (a, b)|(X, Y) = (x, y)) > 2^{-\zeta n}$. Thus, applying Bayes' Rule, we get:

$$Pr((A, B) = (a, b) \mid (X, Y) = (x, y)) > 2^{-\zeta n} \iff$$

$$\prod_{i=1}^{n} Pr(A_i = a_i \mid (A, B)_{<i} = (a, b)_{<i}, (X, Y) = (x, y)) > 2^{-\zeta n} \iff$$

$$\sum_{i=1}^{n} -\log Pr(A_i = a_i \mid (A, B)_{<i} = (a, b)_{<i}, (X, Y) = (x, y)) < \zeta n$$

where we ignored B_i, because it does not change the lower bound. Furthermore, $(X, Y)_{>i}$ and Y_i are independent from A_i, so we can reduce

$$\sum_{i=1}^{n} -\log Pr(A_i = a_i \mid X_i = x_i, (A, B, X, Y)_{<i} = (a, b, x, y)_{<i}) < \zeta n$$

Using concavity of logarithm and Jensen's inequality:

$$\frac{1}{n}\sum_{i=1}^{n} -\log Pr(A_i = a_i \mid X_i = x_i, (A, B, X, Y)_{<i} = (a, b, x, y)_{<i}) \geq$$

$$-\log\left(\frac{1}{n}\sum_{i=1}^{n} Pr(A_i = a_i \mid X_i = x_i, (A, B, X, Y)_{<i} = (a, b, x, y)_{<i})\right)$$

Hence we get:

$$\frac{1}{n}\sum_{i=1}^{n} Pr(A_i = a_i \mid X_i = x_i, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) > 2^{-\zeta} > 1-\zeta$$

where the final inequality is because $1/2 \geq \zeta$.

Since there are at most $4\gamma n/3$ test rounds:

$$\begin{aligned}
&\sum_{i=1}^{n} Pr(A_i = a_i \mid X_i = x_i, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \\
= &\sum_{i\in[R], X_i = x_0} Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \\
+ &\sum_{i\in[n], X_i \neq x_0} Pr(A_i = a_i \mid X_i \neq x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \\
\leq &\sum_{i\in[n], X_i = x_0} Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) + \frac{4\gamma n}{3}
\end{aligned}$$

Using the fact above:

$$\frac{1}{n}\sum_{i\in[n], X_i = x_0} Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) > 1-\zeta-\frac{4\gamma}{3} \geq 1-2\zeta.$$

Furthermore, we see that conditioned on $X_i = x_0$, a round is chosen as a test round independently. So, the variables $Pr(A_i = a_i \mid X_i = x_0, U_i = 1, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i})$ are independent and have expected value greater than $1-2\zeta$. Using Hoeffding's inequality (Lemma 4.2.2) we get:

$$Pr\left(\left|W'(1-2\zeta) - \sum_{i\in[n], U_i=1} Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i})\right| \geq W'\cdot 2\zeta\right)$$

$$\leq 2e^{-8\zeta^2 W'} \leq s\epsilon^2/4 \leq \epsilon Pr(S)$$

where the second inequality is from the choice of parameters.

We also have

$$Pr\left(\frac{1}{W'}Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \leq 1-4\zeta\right)$$

$$\leq Pr\left(\left|W'(1-2\zeta) - \sum_{i\in[n], U_i=1} Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i})\right| \geq W'\cdot 2\zeta\right)$$

$$\leq \epsilon Pr(S)$$

Thus, we get the desired result. □

Claim 4.3.6. *For all $(x,y,a,b,u) \in S$ except for a fraction of at most ϵ of them, it holds that*

$$\frac{1}{w}\sum_{i\in[n], u_i=1} Pr(WIN_i \mid (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \geq \omega_q(G) - 2\delta$$

Proof. We define the random variable $Z_i \in \{0,1\}$ for any $i = 1,..,W$ such that $Z_i = 1$ if the players succeed at the i-th test round. By the condition of winning we have that

$$\sum_i Z_i \geq W(\omega_q(G) - \delta) \tag{4.2}$$

We also define the sequence of random variables (V_k), such that

$$V_k = \sum_{i=1}^{k} Z_i - \mathbb{E}[Z_j|Z_{j-1},...,Z_1,U]$$

This sequence is a martingale with respect to the sequence $(W,Z_1),(W,Z_1,Z_2),...,(W,Z_1,..,Z_W)$, so by Azuma's inequality (Lemma 4.2.3) we get (since $V_1 = 0$)

$$Pr\left(\Big|\sum_i Z_i - \sum_i \mathbb{E}[Z_j|Z_{j-1},...,Z_1,W]\Big| \geq W\delta\right) \leq 2e^{-W\delta^2/2} \leq \epsilon^2 s/4$$

where we used the bound for W in S and the selection of parameters for the last inequality. Using (4.2), we can get

$$Pr\left(\sum_i \mathbb{E}[Z_j|Z_{j-1},...,Z_1,U]| \leq W(\omega_q(G) - 2\delta)\right)$$

$$\leq Pr\left(\sum_i \mathbb{E}[Z_j|Z_{j-1},...,Z_1,U]| \leq \sum_{i=1}^{k} Z_i - W\delta\right) \leq \epsilon^2 s/4$$

However, the probabilities here are over Ω', since we used (4.2). In order to remove the condition on WIN', we can multiply by $s/2 \leq Pr(WIN')$, so the bound becomes $\epsilon^2 s^2/8 \leq \epsilon^2 s/4 \leq \epsilon Pr(S)$. So, we derive the desired result. □

Concluding the proof of Theorem 4.3.3, we can use the two claims to find some $(x,y,a,b,u) \in S$ such that both propositions hold. Then, suppose that for k_1 number of rounds it holds that

$$Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \leq 1 - 4c_1\zeta$$

and for k_2 number of rounds it holds that

$$Pr(WIN_i \mid (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \leq \omega_q(G) - 2c_2\delta$$

Then, we have that

$$1 - 4\zeta \leq \frac{1}{w'} \sum_{i\in[n],u_i=1,x_i=x_0} Pr(A_i = a_i \mid X_i = x_0, (A,B,X,Y)_{<i} = (a,b,x,y)_{<i})$$

$$\leq \frac{1}{w'}[(w'-k_1)\cdot 1 + k_1(1-4c_1\zeta)] = 1 - \frac{4k_1c_1\zeta}{w'} \iff k_1 \leq \frac{w'}{c_1}$$

and since from the game G we have $Pr(WIN_i \mid (A,B,X,Y)_{<i} = (a,b,x,y)_{<i}) \leq \omega_q(G)$, we have

$$\omega_q(G) - 2\delta \leq \frac{1}{w} \sum_{i\in[n],u_i=1} Pr(WIN_i \mid (A,B,X,Y)_{<i} = (a,b,x,y)_{<i})$$

$$\leq \frac{1}{w}[(w-k_2)\omega_q(G) + k_2(\omega_q(G) - 2c_2\delta)] = \omega_q(G) - \frac{2k_2c_2\delta}{w}$$

However, we also have from the definition of S:

$$w \leq \frac{4\gamma n}{3} = \frac{2}{p_0}\frac{2p_0\gamma n}{3} \leq \frac{2}{p_0}w'$$

thus we get

$$\omega_q(G) - 2\delta \leq \omega_q(G) - \frac{2k_2c_2\delta}{w} \leq \omega_q(G) - \frac{2k_2c_2p_0\delta}{2w'} \iff k_2 \leq \frac{2w'}{c_2p_0}$$

Overall, we have that

$$k_1 + k_2 \leq w'\left(\frac{1}{c_1} + \frac{2}{c_2p_0}\right)$$

If we put $c_1 = 2$ and $c_2 = 4/p_0$, then $k_1 + k_2 < w'$, so we can find some round i, such that both of the relations hold:

$$Pr(A_i = a_i \mid X_i = x_0, (A, B, X, Y)_{<i} = (a, b, x, y)_{<i}) \geq 1 - 8\zeta$$

$$Pr(WIN_i \mid (A, B, X, Y)_{<i} = (a, b, x, y)_{<i}) \geq \omega_q(G) - 8\delta/p_0$$

We can now find a strategy for a single round of the game. Before the game starts, the players execute the protocol with inputs (x, y) until round i. If they do not get outputs (a, b), they abort and start again. When they have achieved these outputs, they stop communicating and play a round of the game G. The properties of their strategy now are those from the relations above. □

4.4 Upper Bounds on Randomness

We have proved that we can have exponential randomness expansion using the Vazirani-Vidick protocol, which uses iteratively the CHSH game. Nonetheless, if the protocol is non-adaptive, which means that we cannot use its output to produce some input in a later round, there are some limitations on the randomness we can generate. Next we will present two upper bounds on these randomness expansion protocols, which were suggested by Coudron, Vidick and Yuen [15]. We will prove these bounds by showing that the two players have cheating strategies, so that they can make the referee accept, but generating less randomness.

The main idea behind the cheating strategies is that after a number of rounds, there must be some correlations among the inputs, which are independent from the random seed. The players can use these correlations to generate outputs without using quantum properties, so the output entropy will not increase.

In order to observe better the correlations among the inputs, we can define the input matrix.

Definition 4.4.1. If P is a non-adaptive protocol for randomness expansion with n rounds and seed length m, then we define the input matrix M_p as a $n \times 2^m$ matrix with elements $M_p(i, \sigma) = (X(\sigma)_i, Y(\sigma)_i)$, where $X(\sigma)_i$ and $Y(\sigma)_i$ are the input sequences for D_1 and D_2 respectively resulting from the seed $\sigma \in \{0, 1\}^m$.

So, the input matrix is a matrix, in which every column is the sequence of inputs corresponding to a specific value of the seed. It is clear that every entry of the matrix is well defined, as for a specific choice of seed we have a defined sequence of inputs. Using this matrix, it is easier to find correlations among the inputs.

4.4.1 A doubly exponential bound for perfect games

This doubly exponential upper bound is based on games with $\omega_q(G) = 1$, so the referee's test is checking that the players succeed in every round.

Theorem 4.4.2. *Let G be a game such that $\omega_q(G) = 1$ and P be a randomness expansion protocol with input alphabets $\mathcal{X}, \mathcal{Y}$ and output alphabets $\mathcal{A}, \mathcal{B}$. Suppose that the test of the referee is that the devices win every round of the protocol and completeness and soundness hold with quantum devices. Then the expansion achieved by P must satisfy*

$$g(m) \leq |\mathcal{X} \times \mathcal{Y}|^{2^m} \log |\mathcal{A} \times \mathcal{B}| - \log(1 - 2\epsilon)$$

where ϵ is the smoothness of the protocol P.

Proof. We define $M_i \in (\mathcal{X} \times \mathcal{Y})^{2^m}$ to be the i-th row of the input matrix and also the set $F(M) \subset [n]$ as the set of round indices i such that $M_i \neq M_j$ for all $j < i$. It is easy to see that $|F(M)| \leq |\mathcal{X} \times \mathcal{Y}|^{2^m}$ and since the input matrix of the protocol is known to the devices, they are able to find this set.

So, the cheating strategy is as follows: in every round i, the devices can check if $i \in F(M)$. If it is, then they play the game honestly with the perfect strategy. If not, then they can find another round j, such that $M_j = M_i$ (from the definition of $F(M)$). In this round the inputs are the same as in round i regardless of the seed, because the whole row is the same. Thus, for round i they can use the same output as in round j, which is sure to pass the referee's test. As a result, the only rounds that entropy is generated are the rounds in $F(M)$.

We conclude that since the support of this probability space is

$$|\mathcal{A} \times \mathcal{B}|^{|\mathcal{X} \times \mathcal{Y}|^{2^m}}$$

the max-entropy is

$$H_0(A, B|X, Y, WIN) = |\mathcal{X} \times \mathcal{Y}|^{2^m} \cdot \log |\mathcal{A} \times \mathcal{B}|$$

As we know, the expansion of the protocol is the smooth min-entropy of the outputs. Therefore, in order to have the final result, we can bound the smooth min-entropy with the following lemma:

Lemma 4.4.3. *If X is a discrete random variable and $\epsilon \in [0, 1)$, then*

$$H^{\epsilon}_{\infty}(X) \leq H_0(X) - \log(1 - 2\epsilon)$$

Proof. If $\mu = H^{\epsilon}_{\infty}(X)$, then there must be by definition another discrete random variable Y such that $\|Y - X\|_1 \leq \epsilon$ and $H_\infty(Y) = \mu$. Hence, we have for every $u \in supp(X)$, $Pr(Y = u) \leq 2^{-\mu}$ and also

$$\|Y - X\|_1 = \frac{1}{2} \sum_{u \in supp(Y) \cup supp(X)} |Pr_Y(u) - Pr_X(u)| \geq \frac{1}{2} \sum_{u \in supp(X)} |Pr_Y(u) - Pr_X(u)|$$

$$\geq \frac{1}{2} \left| \sum_{u \in supp(X)} Pr_Y(u) - \sum_{u \in supp(X)} Pr_X(u) \right| = \frac{1}{2} \left(1 - \sum_{u \in supp(X)} Pr_Y(u) \right) \geq \frac{1}{2}(1 - |supp(X)| 2^{-\mu})$$

Therefore, $|supp(X)|2^{-\mu} \geq 1 - 2\epsilon$ and taking logarithm in this relation and using $H_0(X) = \log |supp(X)|$, we get the desired result. □

Using lemma 4.4.3 and the value of max-entropy, we finally get

$$H^{\epsilon}_{\infty}(A, B|X, Y, WIN) \leq |\mathcal{X} \times \mathcal{Y}|^{2^m} \log |\mathcal{A} \times \mathcal{B}| - \log(1 - 2\epsilon)$$

□

4.4.2 A doubly exponential bound for robust protocols

We can generalise the cheating strategy above, so that we can prove a doubly exponential upper bound for any non-adaptive and robust randomness expansion protocol. In this case, we do not need the game G to have a perfect strategy and the referee's test to be checking if the players win in all rounds. This fact rules out the strategy that the players used in the previous case, because the referee could change their test, so that they can detect patterns or repetitions in the players' answers. However, there is a more elaborate cheating strategy that can pass any type of tests. The general idea is instead of finding a specific output to each question pair, to apply their strategy for multiple times and generate an approximate distribution of the real distribution of outputs to this question. After having this distribution, they can use shared randomness to sample from it in order to generate their answers.

Theorem 4.4.4. *If P is a non-adaptive η-robust protocol for randomness expansion with seed length m, input alphabets $\mathcal{X}, \mathcal{Y}$ and output alphabets $\mathcal{A}, \mathcal{B}$, such that completeness and soundness hold with quantum devices, then the expansion achieved by P must satisfy:*

$$g(m) \leq K \cdot |\mathcal{X} \times \mathcal{Y}|^{2^m} \cdot \log |\mathcal{A} \times \mathcal{B}| - \log(1 - 2\epsilon)$$

where $K = \Theta\left(\frac{|\mathcal{A}\times\mathcal{B}|^2}{\eta^2} \log \frac{|\mathcal{A}\times\mathcal{B}|\cdot|\mathcal{X}\times\mathcal{Y}|^{2^m}|}{\eta}\right)$ and ϵ is the smoothness of the protocol P.

As a result of the theorem, any robust randomness expansion protocol can have expansion at most $g(m) = 2^{O(2^m)}$, so unbounded randomness expansion is impossible for non-adaptive protocols, if we allow the devices used to have memory. However, we will show in the next chapter that unbounded randomness is possible with adaptive protocols, if we increase the number of devices used to 3.

Proof. Since completeness and soundness of G hold with quantum devices on the game G, then there is an ideal quantum strategy for the players in a single round of the game that can be used in the protocol, let it be S_G. Based on this strategy, we will find a new cheating strategy S'.

We consider the input matrix M, as defined above. For any round i, the players check if $i \in F(M)$, which means that there is no previous round with the same row in the input matrix. If it is, then they perform the sampling step: they play the game G repeatedly using the ideal strategy S_G for K times, where K is the value defined above. As a result they produce sequences of outputs $a^{(i)} = (a^i_k)_{k=1,..,K}$ (for Alice) and $b^{(i)} = (b^i_k)_{k=1,..,K}$ (for Bob), which they store locally. Then they continue with the replay step that we define below. If $i \notin F(M)$, they perform only the replay step.

In the replay step, if it is round i, the players find the round $j \in F(M)$ such that $M_i = M_j$. Then they use shared randomness to choose some $k \in \{1, ..., K\}$ and they output $a^{(j)}_k$ and $b^{(j)}_k$

as their answers to the referee. As we see the distribution of their answers has density function

$$q_i(a,b) = \frac{1}{K}\sum_{k=1}^{K} \mathbf{1}\left[(a_k^{(i)}, b_k^{(i)}) = (a,b)\right]$$

Since the protocol P is η-robust, if we want strategy S' to be successful with referee's test, we need to prove that with high probability this distribution has statistical distance at most η with the distribution generated by the strategy S_G for the same question pair (x_i, y_i).

Using Hoeffding's inequality (Lemma 4.2.2), we get:

$$Pr\left(|q_i(a,b) - S_G(a,b|x_i,y_i)| > \frac{\eta}{2|\mathcal{A}\times\mathcal{B}|}\right)$$

$$\leq Pr\left(\sum_{k=1}^{K}\left|\frac{1}{K}\mathbf{1}[(a_k^{(i)}, b_k^{(i)}) = (a,b)] - S_G(a,b|x_i,y_i)\right| > \frac{\eta}{2|\mathcal{A}\times\mathcal{B}|}\right)$$

$$\leq 2exp\left(-\frac{\eta^2}{2K\cdot 1/K^2\cdot|\mathcal{A}\times\mathcal{B}|^2}\right)$$

Using the union bound for any pair (a,b), we have

$$Pr(\|q_i - S_G(\cdot,\cdot|x_i,y_i)\|_1 > \eta) \leq |\mathcal{A}\times\mathcal{B}|\cdot exp\left(-\Theta\left(\frac{\eta^2 K}{|\mathcal{A}\times\mathcal{B}|^2}\right)\right)$$

Thus, again using the union bound, the probability that there is a round $i \in F(M)$ such that $|q_i - S_G(\cdot,\cdot|x_i,y_i)| > \eta$ is at most $|F(M)|\cdot|\mathcal{A}\times\mathcal{B}|\cdot exp\left(-\Theta\left(\frac{\eta^2 K}{|\mathcal{A}\times\mathcal{B}|^2}\right)\right)$. Using the value for K, we have that this probability is less than $\eta/2$.

We have ensured that the cheating strategy S' will be accepted by the referee. It only suffices to prove the upper bound on the randomness expansion. Firstly, we can derandomise the replay step. We assume that the shared randomness the two players use is r. Using strategy S' the success probability is $Pr(WIN) \geq c$ for some c that is accepted by the referee. Using the probabilistic method, we see that there is a fixed random string r^* that can achieve $Pr(WIN) \geq c$. So, the players can precompute r^* and use this instead of shared randomness. Consequently, there is no randomness generated at the replay step. So, the only randomness produced is at the sampling step, which is a K-times repetition of the strategy used in the previous case, so we get the desired result. □

Chapter 5

Infinite Randomness Expansion

5.1 Introduction

In the last chapter, we saw that the randomness generated by any non-adaptive robust protocol is bounded (double exponentially). In this chapter, we see what we can do with adaptive protocols. We do not know yet if we can compose protocols that use the same devices, because they have much information from their previous output that can be used on their new input. So, we look what is possible with more than 2 devices.

The first major result came from Coudron and Yuen in 2014 [17]. Their protocol used 8 non-signalling devices and achieved infinite (unbounded) randomness expansion. That is with an initial seed, we can use randomness expansion subprotocols iteratively for as many times as we desire. In their paper, they mention 4 problems for infinite randomness expansion:

The Input Security Problem The VV protocol (from [16]) requires the seed to be uniform to the eavesdropper, but this is not the case in an adaptive protocol.

The Extractor Seed Problem We need to extract randomness from the string and the seed that the extractor uses may not be secure.

The Conditioning Security Problem The output guarantees only hold conditioned on the protocol succeeding, which can probably skew the distribution.

The Compounding Error Problem Errors will accumulate with each iteration of the protocol.

The 4th problem is solved in the analysis of the protocol. The other 3 problems are solved with the help of 4 devices that operate to secure the output of every randomness expansion subprotocol.

In detail, their idea is to use 2 clusters of 4 devices, let them be C_1 and C_2. Each cluster has two devices that execute an exponential randomness expansion protocol, in their case the VV protocol. The random output of this subprotocol is used with the other 2 devices of the cluster, which execute a subprotocol called RUV (from the paper of Reichardt, Unger and Vazirani [36]). This subprotocol uses a property of nonlocal games called rigidity and reduces the entropy, but at the same time makes it secure from the devices of the other cluster. As a result, the cluster achieves to expand randomness which is completely secure from the devices of the other cluster. Thus, we can use the other cluster to further expand randomness. We can do this procedure for unlimited number of times and it is proved that the errors accumulate to a small upper bound.

Later, in a paper from Chung, Shi and Wu [19], it is proved that any randomness expansion protocol can guarantee the same security and performance, even if its uniform-to-all input is replaced by a uniform-to-device input, thus solving the 3 first problems. In combination with the exponential randomness expansion protocol of Miller and Shi [18], this theorem, the Equivalence Lemma, which will be explained in Section 5.2, makes possible to have infinite randomness expansion with 4 devices; we just have 2 clusters of 2 devices that execute the Miller and Shi protocol and cross-feed each other.

The goal here is to reduce the number of devices to 3. The general idea is to use the Equivalence Lemma and a subprotocol that expands randomness, but keeping it secure from the one of the two devices. So, we can have 3 clusters that we alternate: $\{D_1, D_2\}$, $\{D_2, D_3\}$ and $\{D_3, D_1\}$ and the first device of the pair is the device used for the output randomness, while the output must remain uniformly random from the second device and some eavesdropper, which also contains the third unused device. Hence, since the new input is uniform to the devices of the next cluster, using the Equivalence Lemma, we get the desired security.

It is clear that we need a new randomness expansion protocol that makes its input secure for the one of the two devices, as well as the eavesdropper. We call this protocol (or subprotocol of the final infinite expansion protocol), Blind Randomness Expansion Protocol (BRE). In order to prove that such a protocol can exist, we use the Generalised Entropy Accumulation Theorem [20]. This is a new updated version of the Entropy Accumulation Theorem, which we mentioned in the previous chapter, and can be used for various applications.

Consequently, we present the two main theorems, the Equivalence Lemma and the Generalised Entropy Accumulation Theorem.

5.2 Equivalence Lemma

We will use the notation of the original paper [19] to define the physical randomness extractors. We have a physical system $\mathcal{S} = (X, D, E)$, where X is the source, a classical system that corresponds to the seed we have seen in randomness expansion protocols, D is the quantum system, which consists of t quantum devices $(D_1, D_2, ..., D_t)$ with some specific operations, which do not allow communication with each other, and E is a quantum adversary.

Definition 5.2.1. (Uniform source) The system X is an (n, k) source, if the seed is a string of n bits with min-entropy k (the probability of any string is less than 2^{-k}). However, randomness is relative, so this min-entropy value is conditioned on the side information of the other components. If the min-entropy is taken with the scope of all subsystems of $\mathcal{S}$ (except for X), then we say that the source is uniform-to-all random. If it is taken on the devices $D = (D_1, D_2, ..., D_t)$, then the source is uniform-to-devices.

The formal definition of a physical randomness extractor is below.

Definition 5.2.2. (Physical Randomness Extractor) A physical randomness extractor for a physical system $\mathcal{S}(X, D, E)$ is a classical deterministic algorithm, which uses the source X as input to set its parameters and classically interacts with the devices of D, in order them to execute their operations. The output of the algorithm is a bit $A \in \{0,1\}$, where 1 is for accepting and 0 for rejecting and a string $Z \in \{0,1\}^*$, which corresponds to the desired random string.

Theorem 5.2.3. (Equivalence Lemma) *[19]*

Any seeded PRE(Physical Randomness Extractor) for uniform-to-all seeds is also a seeded PRE for uniform-to-devices seeds with the same performance parameters using the same implementation.

This result is very strong, since it let us compose different protocols for randomness expansion only by ensuring that the input to a protocol is uniform to its devices.

5.3 Generalised Entropy Accumulation Theorem

At first, we need to describe the setup. We have two quantum registers, R (user's register) and E (eavesdropper's register). Initially they are at states R_0 and E_0 and the density matrix of the initial joint state is $\rho_{R_0E_0}$. In order to simulate the evolution of the two systems, we consider a sequence of channels $\{\mathcal{M}_i\}_{i=1}^n$ with $\mathcal{M}_i \in CPTP(R_{i-1}E_{i-1}, C_iA_iR_iE_i)$ for each $i = 1, .., n$ which represent the iterative process that will finally accumulate entropy. C_i's are classical systems, which will be used to denote if a round is successful or not, and A_i's are the source of entropy.

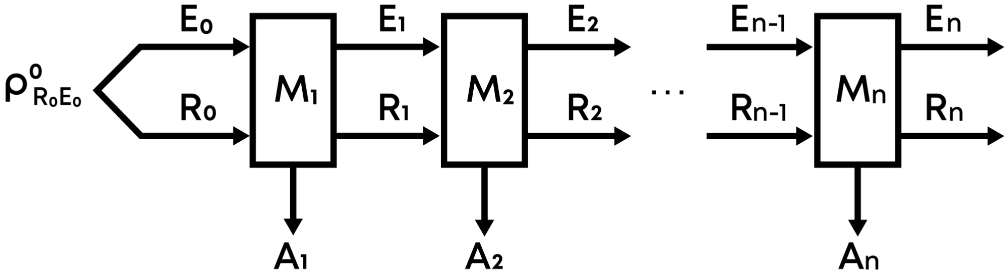


Figure 5.1: The setup for the Generalised Entropy Accumulation Theorem: we have two system E and R that evolve through the operation of the quantum channels $\mathcal{M}_i$, which also produce a classical outcome A_i.

We require two conditions for the channels $\mathcal{M}_i$. We will call the first condition, the non-signalling condition, which states that $\forall \mathcal{M}_i$, there exists a channel $\mathcal{R}_i \in CPTP(E_{i-1}, E_i)$ such that $Tr_{A_iR_iC_i} \circ \mathcal{M}_i = \mathcal{R}_i \circ Tr_{R_{i-1}}$. This means that the evolution of the system E_{i-1} is not affected by the state R_{i-1}, so we have no form of communication between R and E.

For the second condition, if we define $\mathcal{M}'_i = Tr_{C_i} \circ \mathcal{M}_i$, there must exist a channel $\mathcal{T} \in CPTP(A^nE_n, C^nA^nE_n)$ such that

$$\mathcal{M}_n \circ ... \circ \mathcal{M}_1 = \mathcal{T} \circ \mathcal{M}'_n \circ ... \circ \mathcal{M}'_1 \tag{5.1}$$

and $\mathcal{T}$ is of the form

$$\mathcal{T}(\Omega_{A^nE_n}) = \sum_{y\in\mathcal{Y},z\in\mathcal{Z}} (\Pi^{(y)}_{A^n} \otimes \Pi^{(z)}_{E_n})\Omega_{A^nE_n}(\Pi^{(y)}_{A^n} \otimes \Pi^{(z)}_{E_n}) \otimes |r(y,z)\rangle \langle r(y,z)|_{C^n} \tag{5.2}$$

where $\{\Pi^{(y)}_{A^n}\}_y$ and $\{\Pi^{(z)}_{E_n}\}_z$ are families of mutually orthogonal projectors on A^n and E_n, and $r : \mathcal{Y} \times \mathcal{Z} \to \mathcal{C}$ is a deterministic function. Intuitively, this condition says that the classical statistics generated in every round (the classical system C_i) can be reconstructed at the end of the protocol with the information available on the final states A^n and E_n, which is the case in most applications.

We define $\mathbb{P}$ to be the set of probability distributions on the alphabet $\mathcal{C}$ of the classical systems C_i's. From now on, let $\tilde{E}_{i-1}$ be a system isomorphic to $R_{i-1}E_{i-1}$. For any $q \in \mathbb{P}$ we define the set of quantum states

$$\Sigma_i(q) = \{v_{C_iA_iR_iE_i\tilde{E}_{i-1}} = \mathcal{M}_i(\omega_{R_{i-1}E_{i-1}\tilde{E}_{i-1}})|\ \omega \in S(R_{i-1}E_{i-1}\tilde{E}_{i-1})\ \ and\ \ v_{C_i} = q\}, \tag{5.3}$$

where v_{C_i} denotes the probability distribution over $\mathcal{C}$ that emerge from the reduced state of the outcome as $Pr[c] = \langle c| v_{C_i} |c\rangle$. As a result, a state σ belongs to the set $\Sigma_i(q)$, if and only if the state $\mathcal{M}_i(\sigma)$ has output distribution q on the system C_i.

Considering the above definition 5.3, we define min-tradeoff function.

Definition 5.3.1. A function $f : \mathbb{P} \to \mathbb{R}$ is called a min-tradeoff function for the sequence of channels $\{\mathcal{M}_i\}$ if

$$f(q) \leq \min_{v\in\Sigma_i(q)} H(A_i|E_i\tilde{E}_{i-1})_v,\ \forall i = 1, ..., n. \tag{5.4}$$

This means that the function f is upper bounded by any possible entropy of a state with output distribution of C_i equal to q. Note that if $\Sigma_i(q) = \emptyset$, then $f(q)$ can be chosen arbitrarily.

We define two important distributions, δ_x with $x \in \mathcal{C}$ which is distribution with all the weight on element x, and $freq(C^n)$, where $C^n \in \mathcal{C}^n$ which is the distribution on $\mathcal{C}$ defined by the frequency of its letters: $freq(C^n)(c) = \frac{|\{i \in 1,...,n: C_i = c\}|}{n}$.

We also define some quantities related to a min-tradeoff function that will be used in the final result:

$$Max(f) := \max_{q\in\mathbb{P}} f(q),$$
$$Min(f) := \min_{q\in\mathbb{P}} f(q),$$
$$Min_\Sigma(f) := \min_{q:\Sigma(q)\neq\emptyset} f(q)$$
$$Var(f) := \max_{q:\Sigma(q)\neq\emptyset} \sum_{x\in\mathcal{C}} q(x)f(\delta_x)^2 - \left(\sum_{x\in\mathcal{C}} q(x)f(\delta_x)\right)^2$$

where $\Sigma(q) = \cup_i \Sigma_i(q)$.

Finally in this context, an event Ω is defined as a subset of $\mathcal{C}^n$, and for a final state $\rho_{C^nA^nE_nR_n}$ we have $Pr_\rho[\Omega] = \sum_{c^n\in\Omega} Tr[\Pi_{c^n}\rho_{C^nA^nE_nR_n}\Pi_{c^n}]$ for the probability of the event Ω and

$$\rho_{C^nA^nE_nR_n|\Omega} = \frac{1}{Pr_\rho[\Omega]} \sum_{c^n\in\Omega} |c^n\rangle \langle c^n|_{C^n} \otimes \Pi_{c^n}\rho_{C^nA^nE_nR_n}\Pi_{c^n} \tag{5.5}$$

for the state conditioned on Ω, where Π_{c^n} is a projector of the system C^n on the event $C^n = c^n$.

We now state the main theorem.

Theorem 5.3.2. *(from [20]) Consider a sequence of channels $\mathcal{M}_i \in CPTP(R_{i-1}E_{i-1}, C_iA_iR_iE_i)$ for $i \in \{1, ..., n\}$, which satisfy the non-signalling condition and (5.1). Let $\epsilon \in (0,1)$, $\alpha \in (1, 3/2)$, $\Omega \subset \mathcal{C}^n$, $\rho_{R_0E_0} \in S(R_0E_0)$, and f be an affine min-tradeoff function with $h = \min_{c^n \in \Omega} f(freq(c^n))$. Then,*

$$
\begin{aligned}
&H^{\epsilon}_{min}(A^n|E_n)_{\mathcal{M}_n \circ ... \circ \mathcal{M}_1(\rho_{R_0E_0})_{|\Omega}} \\
&\geq nh - n\frac{\alpha-1}{2-\alpha}\frac{ln(2)}{2}V^2 - \frac{g(\epsilon) + \alpha \log(1/Pr_{\rho^n}[\Omega])}{\alpha - 1} - n\left(\frac{\alpha-1}{2-\alpha}\right)K'(\alpha)
\end{aligned}
$$

where $Pr[\Omega]$ is the probability of observing event Ω, and

$$g(\epsilon) = -\log\Big(1 - \sqrt{1-\epsilon^2}\Big),$$

$$V = \log\big(2d_A^2A + 1\big) + \sqrt{2 + Var(f)},$$

$$K'(\alpha) = \frac{(2-\alpha)^3}{6(3-2\alpha)^3 ln(2)} 2^{\frac{\alpha-1}{2-\alpha}(2\log d_A + Max(f) - Min_{\Sigma}(f))} ln^3\left(2^{\frac{\alpha-1}{2-\alpha}(2\log d_A + Max(f) - Min_{\Sigma}(f))} + e^2\right),$$

with $d_A = \max_i dim(A_i)$.

Corollary 5.3.3. *For the same settings given above we have*

$$H^{\epsilon}_{min}(A^n|E_n)_{\mathcal{M}_n \circ ... \circ \mathcal{M}_1(\rho_{R_0E_0})_{|\Omega}} \geq nh - c_1\sqrt{n} - c_0$$

for c_1 and c_0 defined as

$$c_1 = \sqrt{\frac{2ln(2)V^2}{\eta}\left(g(\epsilon) + (2-\eta)\log(1/Pr_{\rho^n}[\Omega])\right)}$$

$$c_0 = \frac{(2-\eta)\log(1/Pr_{\rho^n}[\Omega]) + \eta^2 g(\epsilon)}{3ln^2(2)V^2(2\eta-1)^3} 2^{\frac{1-\eta}{\eta}(2\log d_A + Max(f) - Min_{\Sigma}(f))} ln^3\left(2^{2\log d_A + Max(f) - Min_{\Sigma}(f)} + e^2\right)$$

with $\eta = \dfrac{2ln(2)}{1+2ln(2)}$, $g(\epsilon) = \log\Big(1 - \sqrt{1-\epsilon^2}\Big)$, $V = \log\big(2d_A^2 + 1\big) + \sqrt{2 + Var(f)}$.

We will use Corollary 5.3.3 to prove Blind Randomness Expansion.

5.4 Blind/Local Randomness Expansion Protocol (BRE)

For the following protocol, we need to consider Bob and Eve as one player, let it be Eve, because we require the generated randomness to be hidden from both Bob and Eve. The referee sends the corresponding part of the question to Alice, whose output will be the generated randomness of the protocol, and the other one to Eve. There will be two kinds of rounds: the generation rounds with a specific pair of inputs, where the referee collects the outputs for the final randomness, and the test rounds, where the referee checks if the players play honestly (checking the condition of the nonlocal game). The statement of the protocol is similar with the one in Protocol 4.3.1, but it is shown below for convenience.

Blind Randomness Expansion Protocol (BRE)

Protocol arguments

- G : two-player non-local game, specified by a question set $\mathcal{X} \times \mathcal{Y}$, a probability distribution q on $\mathcal{X} \times \mathcal{Y}$, an answer set $\mathcal{A} \times \mathcal{B}$, and a winning condition $\omega : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \rightarrow \{0, 1\}$
- $x^* \in \mathcal{X}$, $y^* \in \mathcal{Y}$: fixed inputs used for generation rounds
- D : untrusted device capable of playing one side of G repeatedly
- $n \in \mathbb{N}$: number of rounds
- $\gamma \in (0, 1]$: expected fraction of test rounds
- ω_{exp} : expected winning probability in G
- δ : error tolerance

Protocol steps

For rounds $i = 1, ..., n$, the referee performs the following steps:

1. They choose $T_i \in \{0, 1\}$ with $Pr[T_i = 1] = \gamma$. If $T_i = 1$, the referee chooses $X_i, Y_i \in \mathcal{X} \times \mathcal{Y}$ according to the question distribution q. If $T_i = 0$, the referee chooses $X_i = x^*$, $Y_i = y^*$.
2. They send input X_i to Alice and Y_i to Eve. They receive answers A_i and B_i, respectively.
3. If $T_i = 0$, the referee sets $C_i = \perp$. If $T_i = 1$, they set $C_i = \omega(X_i, Y_i, A_i, B_i)$.

At the end of the protocol, the referee aborts if $|\{i \;\; s.t. \;\; C_i = 0\}| > (1 - \omega_{exp} + \delta) \cdot \gamma n$.

Theorem 5.4.1. *(from [20]*

The referee executes Protocol BRE. Alice and the adversary Eve cannot communicate with each other. We define R_i and E'_i to be the states of their quantum systems respectively and $E_i := T^i X^i Y^i B^i E'^i$ to be the side-information available to Eve after the i-th round of the protocol. Also, $N_i \in CPTP(R_{i-1}E_{i-1}, C_i A_i R_i E_i)$ is a quantum channel that corresponds to the i-th round and N_i^{test} is the same as N_i, but only when $T_i = 1$. Let $\rho_{A^n C^n R_n E_n}$ be the final state of the systems and Ω the event that the referee does not abort (like it was defined in 5.5).

Let $g : \mathbb{P}(\{0, 1\}) \rightarrow \mathbb{R}$ be an affine function satisfying the conditions

$$g(p) \leq \inf_{\omega \in S(R_{i-1}E_{i-1}\tilde{E}_{i-1}) : N_i^{test}(\omega)_{C_i} = p} H(A_i | E_i \tilde{E}_{i-1})_{N_i(\omega)},$$

$$Max(g) = g(\delta_1),$$

where $\tilde{E}_{i-1} \equiv R_{i_1} E_{i-1}$ is a purifying system and δ_x is the distribution with all the weight on x. Then, for any $\epsilon_\alpha, \epsilon_s \in (0, 1)$ (ϵ_α for soundness and ϵ_s for smoothness), either $Pr[\Omega] \leq \epsilon_\alpha$ or

$$H_{min}^{\epsilon_s}(A^n | E_n) \geq nh - c_1\sqrt{n} - c_0 \tag{5.6}$$

for $c_1, c_0 \geq 0$ independent of n and

$$h = \min_{p' \in \mathbb{P}(\{0,1\}) : p'(0) \leq 1 - \omega_{exp} + \delta} g(p'),$$

where ω_{exp} is the expected winning probability and δ the error tolerance from Protocol BRE. If we treat $\epsilon_s, \epsilon_\alpha, dim(A_i), \delta, Max(g)$, and $Min(g)$ as constants, then $c_1 = O(1/\sqrt{\gamma})$ and $c_0 = O(1)$. Furthermore, if there exists a quantum strategy that wins the game G with probability ω_{exp}, there is an honest behaviour of Alice and Eve for which $Pr[\Omega] \geq 1 - exp(-\frac{\delta^2}{1-\omega_{exp}+\delta}\gamma n)$.

Proof. We will show that we have the necessary conditions to use Corollary 5.3.3. Firstly, we need the sequence of channels $\mathcal{N}_i$ to satisfy the two required conditions. In order to check the non-signalling condition, we need to define a channel $\mathcal{R}_i \in CPTP(E_{i-1}, E_i)$: in the first step $\mathcal{R}_i$ samples T_i, X_i and Y_i exactly as in Step 1 of the protocol, so as to have the same results. After that it executes the part of Eve, which only requires Y_i and E_{i-1}, since R_{i-1} cannot affect the system because of the no communication requirement. Thus, $Tr_{A_i R_i C_i} \circ \mathcal{N}_i = \mathcal{R}_i \circ Tr_{R_{i-1}}$. Furthermore, C_i is a deterministic function of the variables X_i, Y_i, A_i and B_i, so it is easy to see that the second condition holds, as well.

Afterwards, we need to construct a min-tradeoff function. We define $\mathcal{N}_i = \gamma \mathcal{N}_i^{test} + (1-\gamma)\mathcal{N}_i^{data}$, with $\mathcal{N}_i^{test}$ always picking $T_i = 1$ and $\mathcal{N}_i^{data}$ always picking $T_i = 0$. Then, using Lemma A.0.1 and the condition $Max(g) = g(\delta_1)$ we get the function

$$f(\delta_0) = g(\delta_1) + \frac{1}{\gamma}(g(\delta_0) - g(\delta_1)),$$
$$f(\delta_1) = f(\delta_\perp) = g(\delta_1)$$

which is an affine min-tradeoff function for $\{\mathcal{N}_i\}$.

Then, since the event Ω is a subset of the random variable C^n and $c^n \in \Omega$ if and only if $freq(c^n)(0) \leq (1 - \omega_{exp} + \delta)\gamma$ because of the abort condition, we get (denoting $p = freq(c^n)$ for short):

$$f(freq(c^n)) = p(0)f(\delta_0) + (1-p(0))f(\delta_1) = \frac{p(0)}{\gamma}g(\delta_0) + \left(1 - \frac{p(0)}{\gamma}\right)g(\delta_1) \geq h,$$

The last inequality holds because g is affine (for any $\lambda \in [0,1]$ and x_1, x_2, $\lambda g(x_1) + (1-\lambda)g(x_2) = g(\lambda x_1 + (1-\lambda)x_2)$) and the distribution $p'(0) = p(0)/\gamma$, $p'(1) = 1 - p(0)/\gamma$ satisfies $p'(0) \leq 1 - \omega_{exp} + \delta$. Thus, from Corollary 5.3.3 gives us Equation 5.6. We can obtain c_1, c_0 from the expressions in Corollary 5.3.3.

Finally, we need to show that the honest strategy for both players will only fail with small probability. We define the random variable F_i by $F_i = 1$ if $C_i = 0$, and $F_i = 0$ otherwise. If both players play the game honestly, they win with probability ω_{exp}, hence $\mathbb{E}[F_i] = (1 - \omega_{exp})\gamma$. So using the Chernoff bound and the abort condition, we have

$$Pr[abort] = Pr\left[\sum_{i=1}^{n} F_i > (1 - \omega_{exp} + \delta) \cdot \gamma n\right] =$$
$$Pr\left[\sum_{i=1}^{n} F_i > \left(1 + \frac{\delta}{1 - \omega_{exp}}\right) \cdot \mathbb{E}\left[\sum_{i=1}^{n} F_i\right]\right] \leq e^{-\frac{\delta^2}{1-\omega_{exp}+\delta}\gamma n}$$

□

The only missing part of the proof is finding the function g from the statement of the Theorem 5.4.1. In [20], they show that there is such a function for the CHSH game such that we can generate $\Omega(n)$ bits with a $polylog(n)$ seed. However, for our purpose here, we can use whichever function that expands an m-bits seed to $\Omega(m^2)$ uniformly random bits.

5.5 Infinite Randomness Protocol

Next, we will analyze the protocol for Infinite Randomness Expansion with 3 devices, which we call IE3.

Infinite Randomness Expansion Protocol with 3 devices (IE3)

Protocol arguments

- The BRE Protocol for Blind Randomness Expansion, which is executed by 2 quantum devices, the first one which is the main device (denoted as Alice in the BRE protocol) and the second one which is the "blind" device and has no side information about the output randomness (denoted as Eve in the BRE protocol)
- 3 quantum devices D_1, D_2, D_3
- 3 ordered clusters of the devices $C_1 = (D_1, D_2)$, $C_2 = (D_2, D_3)$, $C_3 = (D_3, D_1)$ (with ordered we mean that the first device has a different function from the second one)
- A (close to) uniformly random string r_0 of m bits

Protocol steps

For rounds $i = 1, 2, ...$ until the desired amount of randomness is generated, the referee performs the following steps:

1. They set $j = i \ (mod\ 3)$. Then they use the generated random string r_{i-1} from the previous round to produce new inputs for the BRE subprotocol. If $i = 1$, they use the initial seed instead.
2. The referee executes the BRE subprotocol using the input produced for the cluster C_j. The first device of the cluster is the main device (the one that they take the output randomness from) and the second device is the "blind" device (the one that they want to hide the output from).
3. They use a randomness extractor to convert the output of the first device to a uniformly random string r_i (it is uniformly random for the other 2 devices and any quantum adversary).

We use the function $g(m)$ for the length of the output randomness of the BRE subprotocol on a seed with length m and $g^{(k)}(m)$ to denote the k-fold composition of $g(m)$ (i.e. $g^{(1)}(m) = g(m), g^{(2)}(m) = g(g(m))$, etc.).

We proceed with proving a theorem for the completeness of the protocol and a theorem for its soundness.

Theorem 5.5.1. (Completeness of the IE3 protocol) *Suppose we execute the protocol IE3(C, S) with $C = \{D_1, D_2, D_3\}$ and S being a uniformly random m-bit seed that is secure against D_1, D_2, D_3. Then, there exists a quantum strategy for the devices, where the devices play*

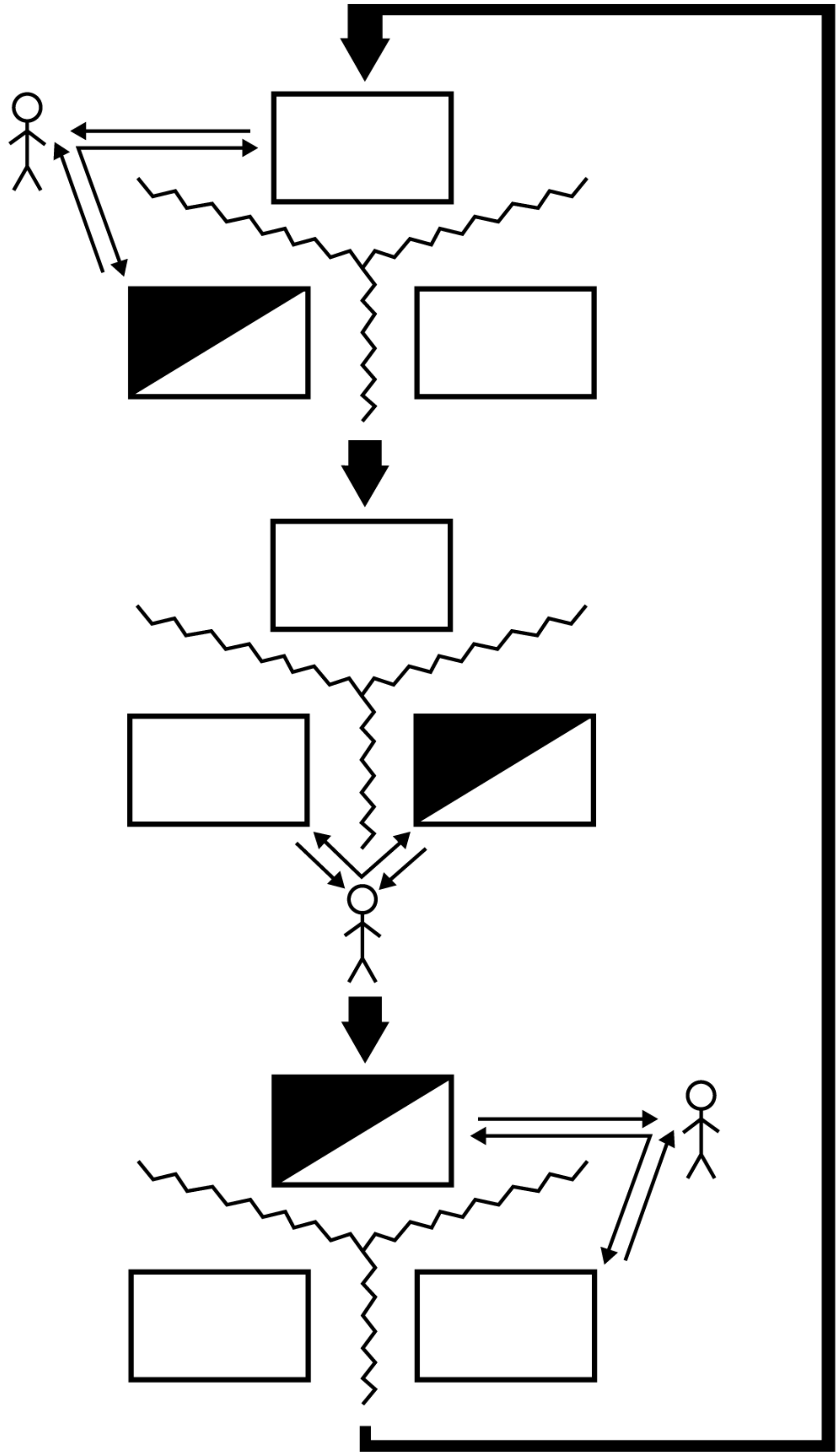


Figure 5.2: The protocol for infinite randomness with 3 devices: at each round the referee executes the BRE subprotocol with one pair (cluster) of devices, from one of which the output randomness is hidden (half-black box), so the referee can use it for the next round.

honestly and do not communicate, such that the probability that the referee aborts in any round i of the protocol is at most $exp(-\Omega(m))$.

Proof. In every round the pair that executes the protocol uses the ideal CHSH strategy to pass the BRE subprotocol. For a fixed round i, we have from Theorem 5.4.1 that the probability that the devices playing fail in the BRE subprotocol is at most $exp(-\Omega(m_i))$, where $m_i = g^{(i)}(m)$. Thus, by the union bound, the probability of the referee aborting any round i is at most

$$\sum_{i=0}^{\infty} exp(-\Omega(m_i)) = \sum_{i=0}^{\infty} exp(-\Omega(g^{(i)}(m))) \leq exp(-\Omega(m))$$

where we used that $g(m) = \Omega(m^2)$ for the last inequality. So, the completeness holds for any large enough expansion $g(m)$. □

Before continuing with the soundness, we need some extra definitions for the convenience of the proof.

Definition 5.5.2. (cq-state) A classical-quantum state (or cq-state) is the composition (using the tensor product) of two systems with one being classical (the density matrix has nonzero elements only in the main diagonal) and one being quantum. If X is the classical system (or equivalently a random variable) and E is the quantum one, then the density matrix of the cq-state is

$$\rho_{XE} = \sum_x p_x \left|x\right\rangle \left\langle x\right| \otimes \rho_B^x$$

where p_x is the probability of the event $X = x$ and $\{\left|x\right\rangle\}$ is an orthonormal basis.

We denote with U_m the density matrix of the uniformly random variable over all strings with length m. It holds that $U_m = 2^{-m}\mathbb{I}_{2^m}$.

Definition 5.5.3. (trace norm) For a matrix A, we define its trace norm as $\|A\|_{tr} = \frac{1}{2}\sqrt{A^*A}$

We can use the trace norm of the difference of two density matrices as a measure of their proximity.

We can use these definitions to provide a different way to describe smoothness. The cq-state $\sigma = U_m \otimes \rho_E$ is a state with a uniform random variable of m bits and a system for a quantum adversary E, which cannot extract any side information about the outcome of X. So, the state σ has min-entropy $H_\infty(X|E) = m$. Our goal is to generate a state that is ϵ-close to this state (so we can have $H_\infty^\epsilon(X|E) = m$), thus we define the notion of ϵ-security, which is similar to ϵ-smoothness:

Definition 5.5.4. (Secure cq-state) If E is a quantum adversary system and ρ_{XE} a cq-state with X being the classical output of a randomness expansion protocol, then X is ϵ-secure against E if and only if

$$\left\|\rho_{XE} - U_{|X|} \otimes \rho_E\right\|_{tr} \leq \epsilon.$$

We will use this notion of security to prove the soundness of the IE3 protocol. This is the main technical contribution of the thesis and is inspired by the corresponding proof from Coudron and Yuen [17].

Theorem 5.5.5. (Soundness of the IE3 protocol) *Let D_1, D_2, D_3 be 3 non-signaling quantum devices. Alice executes the IE3(C, S) protocol with $C = \{D_1, D_2, D_3\}$ and S an m-bit seed. We define WIN_i to be the event that the referee did not abort the IE3 protocol in the i-th round, and $WIN_{\leq i} = WIN_1 \wedge ... \wedge WIN_i$. Let E be the eavesdropper system that may be entangled with D_1, D_2, D_3, but cannot communicate with them and ρ^0_{SC} be the initial state of the seed and the devices. If $\rho^0_{SC} = U_m \otimes \rho^0_C$ (uniformly random for the devices), and $Pr(WIN_{\leq n}) \geq \epsilon_a$ for all $n \in \mathbb{N}$ (ϵ_a is the soundness of the BRE subprotocol), then*

$$\left\| \rho^n_{X_n E} - U_{g^{(n)}(m)} \otimes \rho^n_E \right\|_{tr} \leq \frac{2\epsilon_1}{\epsilon_a},$$

where

- *ϵ_1 is the smoothness for the first execution of the BRE subprotocol and*
- *$\rho^n_{X_n E}$ denotes the joint state of the register X_n that holds the output randomness and the system of the eavesdropper after n rounds of the IE3(C, S) Protocol, conditioned on the event $WIN_{\leq n}$.*

For the proof of the Theorem 5.5.5, we take for granted the correctness of the BRE subprotocol (Theorem 5.4.1), so we have that for any round i the output of the cluster C_i (where C_i denotes the cluster $C_{i\ (mod\ 3)}$) is approximately secure against the next cluster C_{i+1}. Thus, using the Equivalence Lemma, since BRE is a Physical Randomness Extractor protocol, it is possible to compose the executions of BRE by consecutive clusters and to produce nearly uniform output at each round. Moreover, the approximation errors accumulate linearly with each iteration.

Proof. Define $j := i\ (mod\ 3)$. From Bayes' Rule, we can split the overall probability of success, $p = Pr(WIN_{\leq k})$, into conditional probabilities $p_i = Pr(WIN_i | WIN_{\leq i-1})$, so as to have $p = \prod p_i \geq \epsilon_a$. We also define recursively a function $\delta(i)$ to operate as an error bound, such that $\delta(i) := \epsilon_s(g^{(i-1)}(m)) + \delta(i-1)/p_i = \epsilon_i + \delta(i-1)/p_i$ and $\delta(1) := \epsilon_s(m)$. With $\epsilon_s(m)$ we denote the smoothness of the BRE subprotocol on input with length m or in other words how close the final state is to a state which is uniformly random.

For the desired result, there must exist a state $\mu^i_{XD_iC_{i+1}E}$ for every $i = 1, ..., k-1$, such that $\mu^i_{XC_{i+1}E} = U_{g^{(i)}(m)} \otimes \mu^i_{C_{i+1}E}$ and

$$\left\| \rho^i_{X_iCE} - \mu^i_{X_iCE} \right\|_{tr} \leq \delta(i),$$

where the state $\rho^i_{X_iCE}$ is taken conditioned on $WIN_{\leq i}$.

We will prove this by induction:

For $k = 1$, we use Theorem 5.4.1 with $D = D_1$ and the adversary Eve to be C_2 and E together. Thus, we obtain that there exists a state $\mu^1_{X_1CE}$ such that $\mu^1_{X_1C_2E} = U_{g(m)} \otimes \mu^1_{C_2E}$, and

$$\left\| \rho^1_{X_1CE} - \mu^1_{X_1CE} \right\|_{tr} \leq \epsilon_s(m) = \delta(1).$$

Using this as the base case, we suppose that after running $k-1$ rounds of the protocol our hypothesis holds. Due to the Equivalence Lemma (5.2.3), the input to cluster C_k for running the k-th round of the protocol (the BRE subprotocol) needs to be uniformly random for the cluster

C_k, which is true by the hypothesis. Hence, we can use again Theorem 5.4.1 along with Lemma A.0.2 to conclude that there exists a state $\mu^k_{X_kCE}$ such that $\mu^k_{XC_{k+1}E} = U_{g^{(k)}(m)} \otimes \mu^k_{C_{k+1}E}$ and

$$\left\| \rho^k_{X_kCE} - \mu^k_{X_kCE} \right\|_{tr} \leq \epsilon_s(g^{(k-1)}(m)) + \delta(k-1)/p_k := \delta(k).$$

So, the induction argument is complete. We bound $\delta(k)$ by:

$$\delta(k) = \epsilon_k + \frac{1}{p_k}\left(\epsilon_{k-1} + \frac{1}{p_{k-1}}(\epsilon_{k-2} + ...)\right) \leq \frac{1}{\epsilon_a}(\epsilon_k + \epsilon_{k-1} + ... + \epsilon_1) \leq \frac{2\epsilon_1}{\epsilon_a}$$

where we used that $\prod p_i \geq \epsilon_a$ and for each ϵ_i (the smoothness of BRE subprotocol in round i) it holds that $\epsilon_i \leq 2\epsilon_{i-1}$ (we can modify the parameters of Theorem 5.4.1 to have this relation).

Finally, for every k, we have that

$$\left\| \rho^k_{X_kE} - U_{g^{(k)}(m)} \otimes \rho^k_E \right\|_{tr} \leq \left\| \rho^k_{X_kE} - \mu^k_{X_kE} \right\|_{tr} + \left\| \mu^k_{X_kE} - U_{g^{(k)}(m)} \otimes \rho^k_E \right\|_{tr}$$

$$\leq \delta(k) + \left\| U_{g^{(k)}(m)} \otimes \mu^k_E - U_{g^{(k)}(m)} \otimes \rho^k_E \right\|_{tr} = \delta(k) + \left\| \mu^k_E - \rho^k_E \right\|_{tr} \leq 2\delta(k).$$

where for the last inequality we used that tracing over some subsystems on a density matrix only reduces its trace norm, and the proof is complete. □

Appendix A

Lemmas

The first lemma is taken from [37]:

Lemma A.0.1. *Suppose $\mathcal{M}_i \in CPTP(R_{i-1}E_{i-1}, C_iA_iR_iE_i)$ are channels with the same conditions as in Theorem 5.3.2 that are of the form:*

$$\mathcal{M}_i = \gamma \mathcal{M}^{test}_{i,R_{i-1}E_{i-1}\to C_iA_iR_iE_i} + (1-\gamma)\mathcal{M}^{test}_{i,R_{i-1}E_{i-1}\to A_iR_iE_i} \otimes |\perp\rangle\langle\perp|_{C_i}$$

where we extended alphabet $\mathcal{C}$ to $\mathcal{C}' = \mathcal{C} \cup \{\perp\}$. Also, let the affine function $g : \mathbb{P}(\mathcal{C}') \to \mathbb{R}$ satisfy for any $q \in \mathbb{P}(\mathcal{C}')$ and any index i:

$$g(q) \leq \min_{\omega \in S(R_{i-1}E_{i-1}\tilde{E}_{i-1}} \{H(A_i|E_i\tilde{E}_{i-1})_{\mathcal{M}_i(\omega)} : (\mathcal{M}_i(\omega))_{C_i} = q\}$$

where $\tilde{E}_{i-1} \equiv R_{i-1}E_{i-1}$ is a purifying system. Then, we can define a new affine function $f : \mathbb{P}(\mathcal{C}) \to \mathbb{R}$ by

$$f(\delta_x) = Max(g) + \frac{1}{\gamma}(g(\delta_x) - Max(g)) \quad \forall x \in \mathcal{C}'$$
$$f(\delta_\perp) = Max(g)$$

Then, the function f is a min-tradeoff function for $\{\mathcal{M}_i\}$ with properties:

$$Max(f) = Max(g)$$
$$Min(f) = \left(1 - \frac{1}{\gamma}\right) Max(g) + \frac{1}{\gamma} Min(g)$$
$$Min_\Sigma(f) \geq Min(g)$$
$$Var(f) \leq \frac{1}{\gamma}(Max(g) - Min(g))^2.$$

The second lemma is taken from [17]:

Lemma A.0.2. *Suppose we have a randomness expansion protocol P with F being a binary register with value 1 if the referee does not abort, S the register of the seed, X the register of the outcome, D the device used in the protocol and E an arbitrary quantum system that may be entangled with D. We also have the density matrix $\sigma_{FSX} := |0\rangle\langle 0|_F \otimes U_{|S|} \otimes |0\rangle\langle 0|_X$.*

We also define the quantum operation $\mathcal{F}$ and $\mathcal{E}$. $\mathcal{E}$ is some unitary map applied to the state ρ_{FSXD}, which represents the actions of the protocol P, and $\mathcal{F}$ takes the state ρ_{FSXD} and outputs the state $\rho_{XD|F=1}$, which results from the post-measurement state of ρ_{FSXD} conditioned on measuring $|1\rangle$ in F and tracing out F and S.

Furthermore, suppose that for all states σ_{FSXDE} that $\sigma_{FSXD} = \sigma_{FSX} \otimes \sigma_D$ holds, there exists a state τ_{XDE} such that $\tau_{XE} = U_{|X|} \otimes \sigma_E$ and

$$\|\mathcal{F}\mathcal{E} \otimes \mathbb{I}_E(\sigma_{FSXDE}) - \tau_{XDE}\|_{tr} \leq \epsilon$$

Let $\delta, \lambda > 0$ and ρ^i_{FSXDE} the initial state of the system such that $\left\|\rho^i_{FSXDE} - \sigma_{FSXDE}\right\|_{tr} \leq \delta$. If the probability of measuring 1 in the F register of the state $\mathcal{E} \otimes \mathbb{I}_E(\rho^i_{FSXDE})$ is at least λ, then there exists a state μ_{XDE} with $\mu_{XE} = U_{|X|} \otimes \mu_E$ and

$$\left\|\rho^f_{XDE} - \mu_{XDE}\right\|_{tr} \leq \epsilon + \delta/\lambda$$

where $\rho^f_{XDE} = \mathcal{F}\mathcal{E} \otimes \mathbb{I}_E(\rho^i_{FSXDE})$ is the final state.

The intuition behind this lemma is that if the execution of the protocol with a uniformly random seed obtains a state with a random output that is ϵ-close to uniform, then a seed which is δ-close to uniform obtains the same output, but with the additive factor of δ/λ in the trace distance, where λ is the probability of not aborting.

Bibliography

[1] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, 1948. [Online]. Available: https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

[2] P. Benioff, “The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines,” *Journal of Statistical Physics*, vol. 22, no. 5, pp. 563–591, May 1980.

[3] R. P. Feynman, “Simulating Physics with Computers,” *International Journal of Theoretical Physics*, vol. 21, no. 6-7, pp. 467–488, Jun. 1982.

[4] Y. I. Manin, “Computable and Uncomputable,” *Sovetskoe Radio,Moscow*, 1980.

[5] P. W. Shor, “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer,” *SIAM J. Comput.*, vol. 26, no. 5, pp. 1484–1509, 1997. [Online]. Available: https://doi.org/10.1137/S0097539795293172

[6] S. Wiesner, “Conjugate coding,” *SIGACT News*, vol. 15, no. 1, pp. 78–88, 1983. [Online]. Available: https://doi.org/10.1145/1008908.1008920

[7] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” *Theoretical Computer Science*, vol. 560, pp. 7–11, 2014, theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304397514004241

[8] A. K. Ekert, “Quantum cryptography and bell’s theorem,” in *Quantum Measurements in Optics.* Springer, 1992, pp. 413–418.

[9] D. Mayers and A. C. Yao, “Self testing quantum apparatus,” *Quantum Inf. Comput.*, vol. 4, no. 4, pp. 273–286, 2004. [Online]. Available: https://doi.org/10.26421/QIC4.4-3

[10] A. Kolmogorov, “On tables of random numbers,” *Theoretical Computer Science*, vol. 207, no. 2, pp. 387–395, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304397598000759

[11] R. Colbeck, “Quantum and relativistic protocols for secure multi-party computation,” 2009. [Online]. Available: https://arxiv.org/abs/0911.3814

[12] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. M. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. R. Monroe, “Random numbers

certified by bell's theorem," *Nat.*, vol. 464, no. 7291, pp. 1021–1024, 2010. [Online]. Available: https://doi.org/10.1038/nature09008

[13] S. Fehr, R. Gelles, and C. Schaffner, "Security and composability of randomness expansion from bell inequalities," *CoRR*, vol. abs/1111.6052, 2011. [Online]. Available: http://arxiv.org/abs/1111.6052

[14] S. Pironio and S. Massar, "Security of practical private randomness generation," *Physical Review A*, vol. 87, no. 1, p. 012336, 2013.

[15] M. Coudron, T. Vidick, and H. Yuen, "Robust randomness amplifiers: Upper and lower bounds," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 16th International Workshop, APPROX 2013, and 17th International Workshop, RANDOM 2013, Berkeley, CA, USA, August 21-23, 2013. Proceedings*, ser. Lecture Notes in Computer Science, P. Raghavendra, S. Raskhodnikova, K. Jansen, and J. D. P. Rolim, Eds., vol. 8096. Springer, 2013, pp. 468–483. [Online]. Available: https://doi.org/10.1007/978-3-642-40328-6_33

[16] U. V. Vazirani and T. Vidick, "Certifiable quantum dice: or, true random number generation secure against quantum adversaries," in *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, H. J. Karloff and T. Pitassi, Eds. ACM, 2012, pp. 61–76. [Online]. Available: https://doi.org/10.1145/2213977.2213984

[17] M. Coudron and H. Yuen, "Infinite randomness expansion with a constant number of devices," in *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, D. B. Shmoys, Ed. ACM, 2014, pp. 427–436. [Online]. Available: https://doi.org/10.1145/2591796.2591873

[18] C. A. Miller and Y. Shi, "Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices," in *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, D. B. Shmoys, Ed. ACM, 2014, pp. 417–426. [Online]. Available: https://doi.org/10.1145/2591796.2591843

[19] K.-M. Chung, Y. Shi, and X. Wu, "Physical randomness extractors: Generating random numbers with minimal assumptions," 2014. [Online]. Available: https://arxiv.org/abs/1402.4797

[20] T. Metger, O. Fawzi, D. Sutter, and R. Renner, "Generalised entropy accumulation," *CoRR*, vol. abs/2203.04989, 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2203.04989

[21] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition.* Cambridge University Press, 2010.

[22] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Physical review*, vol. 47, no. 10, p. 777, 1935.

[23] J. S. Bell, "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, p. 195, 1964.

[24] J. S. BELL, "On the problem of hidden variables in quantum mechanics," *Rev. Mod. Phys.*, vol. 38, pp. 447–452, Jul 1966. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.38.447

[25] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Physical review letters*, vol. 23, no. 15, p. 880, 1969.

[26] B. S. Cirel'son, "Quantum generalizations of bell's inequality," *Letters in Mathematical Physics*, vol. 4, no. 2, pp. 93–100, 1980.

[27] A. Peres, "Two simple proofs of the kochen-specker theorem," *Journal of Physics A: Mathematical and General*, vol. 24, no. 4, p. L175, 1991.

[28] N. D. Mermin, "Hidden variables and the two theorems of john bell," *Rev. Mod. Phys.*, vol. 65, pp. 803–815, Jul 1993. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.65.803

[29] D. Zuckerman, "General weak random sources," in *Proceedings [1990] 31st Annual Symposium on Foundations of Computer Science.* IEEE, 1990, pp. 534–543.

[30] R. Renner, "Security of quantum key distribution," *International Journal of Quantum Information*, vol. 6, no. 01, pp. 1–127, 2008.

[31] L. Trevisan, "Extractors and pseudorandom generators," *Journal of the ACM*, vol. 48, no. 4, pp. 860–879, 2001.

[32] C. A. Miller and Y. Shi, "Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices," *J. ACM*, vol. 63, no. 4, pp. 33:1–33:63, 2016. [Online]. Available: https://doi.org/10.1145/2885493

[33] ——, "Universal security for randomness expansion from the spot-checking protocol," *SIAM J. Comput.*, vol. 46, no. 4, pp. 1304–1335, 2017. [Online]. Available: https://doi.org/10.1137/15M1044333

[34] R. A. Friedman, R. Renner, and T. Vidick, "Simple and tight device-independent security proofs," *SIAM J. Comput.*, vol. 48, no. 1, pp. 181–225, 2019. [Online]. Available: https://doi.org/10.1137/18M1174726

[35] F. Dupuis, O. Fawzi, and R. Renner, "Entropy accumulation," *CoRR*, vol. abs/1607.01796, 2016. [Online]. Available: http://arxiv.org/abs/1607.01796

[36] B. W. Reichardt, F. Unger, and U. V. Vazirani, "A classical leash for a quantum system: command of quantum systems via rigidity of CHSH games," in *Innovations in Theoretical Computer Science, ITCS '13, Berkeley, CA, USA, January 9-12, 2013*, R. D. Kleinberg, Ed. ACM, 2013, pp. 321–322. [Online]. Available: https://doi.org/10.1145/2422436.2422473

[37] F. Dupuis and O. Fawzi, “Entropy accumulation with improved second-order term,” *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7596–7612, 2019. [Online]. Available: https://doi.org/10.1109/TIT.2019.2929564