Noisy Embezzlement of Entanglement and Applications to Entanglement Dilution

by

Dariusz Lasecki

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization (Quantum Information)

Waterloo, Ontario, Canada, 2019

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis we present the concept of embezzlement of entanglement [17], its properties, efficiency, possible generalizations [14]. and propose the linear programming characterization of this phenomenon. Then, we focus on the noisy setting of embezzlement of entanglement. We provide the detailed proof of the quantum correlated sampling lemma [12] which can be considered a protocol for noisy embezzlement of entanglement. Next, we propose a classical synchronization scheme for two spatially separated parties which do not communicate and use shared randomness to synchronize their descriptions of a quantum state. The result, together with the canonical embezzlement of entanglement [17], improves the quantum correlated sampling lemma [12] for small quantum states in terms of the probability of success and distance between desired and final states. Then, we discuss the role of entanglement spread [9, 11] in dilution of entanglement [15]. We propose an explicit protocol for the task of dilution of entanglement without communication. The protocol uses EPR pairs and an embezzling state of size $O(\sqrt{n}/\epsilon)$ qubits for the task of diluting n partially entangled states up to infidelity ϵ. We modify the protocol to work in a noisy setting where the classical synchronization scheme finds its application.

Acknowledgements

I would like to express my gratitude to my supervisor, Debbie Leung, for inspiring the topic of my research, indispensable discussions regarding results included in this thesis and useful feedback on the draft of this document.

I would like to thank Jon Yard, Anurag Anshu, Shima Bab Hadiashar and Ala Shayeghi for very helpful discussions and also Ashwin Nayak and Jon Yard for agreeing to be the readers of this thesis and for providing valuable comments.

At the end, I would like to appreciate the Institute for Quantum Computing for providing a unique, vibrant and pleasant environment for my research and also my family and friends for love, understanding and continuous support.

Dedication

I dedicate this thesis to my family.

Table of Contents

Chapter 1

Introduction

In this chapter we will introduce basic concepts, definitions and notation used throughout this thesis. We will mostly focus on aspects related to quantum entanglement and its applications. It will give us motivation for why is it so interesting to 'embezzle' entanglement.

1.1 Preliminaries

1.1.1 Notation

In this subsection we introduce and explain mathematical notation used throughout this thesis.

- The d-dimensional Hilbert space is denoted by $\mathcal{H}^d$. Sometimes, the dimension is omitted.
- The set of bounded linear operators from a Hilbert space $\mathcal{H}_1$ to a Hilbert space $\mathcal{H}_1$ is denoted by $L(\mathcal{H}_1, \mathcal{H}_2)$.
- A vector in a Hilbert space is denoted, in the Dirac notation, by $|\cdot\rangle$. A corresponding vector in a dual Hilbert space is denoted by $\langle\cdot| = |\cdot\rangle^\dagger$, where by $\dagger$ we mean the Hermitian conjugate operation.
- The set of density operators on $\mathcal{H}$ is denoted by $D(\mathcal{H})$.
- The trace of an operator is denoted by $\mathrm{Tr}(\cdot)$.

- For the sake of compactness, tensor product symbols $\otimes$ and identity operators might be omitted if their presence can be concluded from the context.

1.1.2 Qubits and Qudits

One of the central concepts in the quantum information theory is a quantum generalization of a classical bit, called a qubit. It originates from the observation that physical systems which follow the rules of quantum mechanics may have degrees of freedom with states distinguishable between each other through some interaction with the system. We can assign logical values to these states and use them for information processing tasks, and this abstraction has far-reaching consequences. A classical bit always has a deterministic state which does not change upon a readout. The situation is more complex and subtle in case of a qubit, due to the laws of quantum mechanics. The qubit state is defined as a linear combination of basis states, called a superposition. Reading a qubit state gives a fundamentally probabilistic result which corresponds to one of the states from the superposition. What is more, measuring a qubit also affects its state which then collapses from a linear combination of states to a state that was just observed. In this sense, qubits can be considered very fragile. However, it still possible to manipulate their state in such a way that the final measurement will yield insightful information regarding our information processing task. Mathematically, using the Dirac notation introduced in the previous subsection, we define a qubit as follows.

Definition 1 (Qubit)**.** *A qubit is an abstraction of a two-dimensional system described by a quantum state* $|\psi\rangle \in \mathcal{H}^2$ *as follows*

$$|\psi\rangle = \alpha_1 |0\rangle + \alpha_2 |1\rangle ,$$

where $\alpha_1, \alpha_2 \in \mathbb{C}$*,* $|\alpha_1|^2 + |\alpha_2|^2 = 1$ *and* $\{|0\rangle , |1\rangle\}$ *is an orthonormal basis of a two-dimensional Hilbert space.*

A qudit is a straightforward generalization of a qubit to a d-dimensional system.

Definition 2 (Qudit)**.** *A qudit is an abstraction of a d-dimensional system described by a quantum state* $|\psi\rangle \in \mathcal{H}^d$ *as follows*

$$|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle ,$$

where $\alpha_0, \alpha_1, \ldots, \alpha_{d-1} \in \mathbb{C}$*,* $\sum_{i=0}^{d-1} |\alpha_i|^2 = 1$ *and* $\{|0\rangle , |1\rangle , \ldots , |d-1\rangle\}$ *is an orthonormal basis of a d-dimensional Hilbert space.*

1.1.3 Quantum Measurements

In this section we define quantum measurements for pure states. This definition is sometimes called the third postulate of quantum mechanics and is at the heart of the probabilistic nature of quantum mechanics.

Definition 3. *A quantum measurement is defined by measurement operators* $\{M_i\}$ *which act on quantum states and satisfy the completeness relation, i.e.,* $\sum_i M_i^\dagger M_i = I$. *When a pure state* $|\psi\rangle$ *is measured, the measurement outcome* i *occurs with probability* $\Pr(i) = \langle\psi| M_i^\dagger M_i |\psi\rangle$ *and the post-measurement state is* $\frac{1}{\sqrt{\Pr(i)}} M_i |\psi\rangle$.

Therefore, the third postulate of quantum mechanics tells us that a measurement of a pure quantum state yield a fundamentally probabilistic result. The above definition can be generalized to density operators, introduced in the next subsection, however, it is not needed in this thesis.

1.1.4 Density Operator Formalism

So far, we have defined quantum states as vectors in the Hilbert space. Quantum states described in this way are called *pure quantum states*. This description, although probabilistic with respect to the outcome of a measurement, assumes that we have a perfect knowledge of coefficients that characterize the state. In practice, however, this knowledge might be incomplete, e.g. equipment used for a state preparation was not perfect, a state was transmitted to us through a noisy channel or we are not sure on how was the state prepared. Therefore, it is of great interest to have a way of including these sometimes inevitable uncertainties in the description of quantum states. The formalism defined for this purpose is called the density operators formalism.

We start by defining the ensemble of quantum states. When we are not sure about the exact description of a quantum state that we possess, we might try use some partial information that we do possess to assign probabilities to descriptions of quantum states that might match our actual quantum state. It can be modelled by introducing a random variable X with a probability distribution p_X over the elements of some alphabet $\mathcal{X}$ which is used to label possible pure quantum states. This probability distribution can be thought of as representing our belief of which pure quantum we actually have.

Definition 4 (Ensemble of Quantum States)**.** *An ensemble* $\mathcal{E}$ *of quantum states is defined as*

$$\mathcal{E} = \{p_X(x), |\psi_x\rangle\}_{x\in\mathcal{X}}$$

Using an ensemble, a density operator is defined as follows.

Definition 5 (Density Operator)**.** *A density operator ρ is defined as*

$$\rho = \sum_{x \in \mathcal{X}} p_X(x) \, |\psi_x\rangle\langle\psi_x|$$

A density operator is considered to be a noisy quantum state and is therefore manipulable, i.e., it can be transformed to another quantum state or measured. We notice that a pure state $|\psi\rangle$ also has a corresponding density operator which is $|\psi\rangle\langle\psi|$.

1.1.5 Distances Between Quantum States

In this subsection we will define what does it mean for two quantum states to be close to each other. We will introduce selected distance measures and their properties, especially their relationship to the fidelity between pure quantum states. These facts will be frequently used throughout this thesis.

We start by defining the Schatten-p norm.

Definition 6 (Schatten-p norm)**.** *Let $A \in L(\mathcal{H})$. The Schatten-p norm of A is defined as*

$$||A||_p = \left[\mathrm{Tr} \left(\left(A^\dagger A \right)^{p/2} \right) \right]^{1/p}.$$

Based on the Schatten-p norm defined above, we will introduce Schatten-1 and Schatten-2 distances between density operators. These are norms commonly used in quantum information theory to quantify distances between density operators representing quantum states.

Definition 7 (Schatten-1 distance)**.** *Let $\rho, \sigma \in D(\mathcal{H})$ be density operators. The Schatten-1 distance is defined as*

$$||\rho - \sigma||_1 = \mathrm{Tr} \, |\rho - \sigma|.$$

The half of the Schatten-1 distance is also known as trace distance.

Definition 8 (Schatten-2 distance)**.** *Let $\rho, \sigma \in D(\mathcal{H})$ be density operators. The Schatten-2 distance is defined as*

$$||\rho - \sigma||_2 = \sqrt{\mathrm{Tr}[(\rho - \sigma)^2]}.$$

The Schatten-2 distance is also known as Hilbert-Schmidt distance or Frobenius distance.

We will now introduce a useful fact which holds for any Schatten norm of a bounded linear operator acting on a Hilbert space. It relates the norm of such an operator with the norm of the vector of its singular values. Singular values are defined later in this Chapter in the Schmidt Decomposition Theorem.

Lemma 1. *Suppose $A \in L(\mathcal{H}_1, \mathcal{H}_2)$. Then*

$$||A||_p = ||s(A)||_p,$$

where $||s(A)||_p$ is the vector p-norm of the vector of singular values of A.

Proof. We consider the definition of the Schatten-p norm and apply a Singular Values Decomposition to A which yields $A = UDV$, where U, V are unitary operations and D is a diagonal matrix with strictly positive entries. Then,

$$||A||_p = \left[\mathrm{Tr}\left(\left(V^\dagger D U^\dagger U D V\right)^{p/2}\right)\right]^{1/p} = \left[\mathrm{Tr}\left(\left(D^2\right)^{p/2}\right)\right]^{1/p} =$$

$$= \left[\mathrm{Tr}\, D^p\right]^{1/p} = \left(\sum_i s_i^p(A)\right)^{1/p} = ||s(A)||_p.$$

□

When dealing with pure quantum states, we may quantify distances between them by an Euclidean distance between the vectors of coefficients characterizing them.

Definition 9 (Euclidean distance)**.** *Let $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ be arbitrary qudits. Then the Euclidean distance between $|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle$ and $|\phi\rangle = \sum_{i=0}^{d^-1} \beta_i |i\rangle$ is defined as*

$$|| \, |\psi\rangle - |\phi\rangle \, ||_2 = \sqrt{\sum_{i=0}^{d-1} (\alpha_i - \beta_i)^2}.$$

As discussed, Schatten norms are widely used as distance measures between density operators. However, in practice it is often easier to use the concept of fidelity between quantum states which can be intuitively understood as an overlap between quantum states. The fidelity itself is not a distance measure in the mathematical sense, however, there is a clear relationship between the fidelity and Schatten distances for density operators. Therefore, it is usually easier to work with fidelity and then, if necessary, switch to the Schatten distance (e.g. to make use of the triangle inequality). The fidelity between quantum states is defined as follows.

Definition 10 (Fidelity)**.** *A fidelity F between quantum states $\rho, \sigma \in D(\mathcal{H})$ is defined as*

$$F(\rho, \sigma) = \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}.$$

In case of pure states $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, this definition reduces to $F(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle$. Sometimes, another definition $F'(\rho, \sigma)$ is used instead of $F(\rho, \sigma)$, related as $F'(\rho, \sigma) = F^2(\rho, \sigma)$.

The relationship between the Schatten distance and fidelity of two pure states is given in the following Lemma.

Lemma 2. *Let $|\psi\rangle, |\phi\rangle$ be arbitrary qudits. Then the following holds*

$$|| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \, ||_p = 2^{1/p}\sqrt{1 - F^2(|\phi\rangle, |\psi\rangle)}.$$

Proof. This proof follows [19]. Let $N = |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$. We notice that N is Hermitian, thus also normal. Therefore, the singular values of N are absolute values of those eigenvalues of N which are non-zero. We calculate

$$\mathrm{Tr}\big(N^2\big) = \mathrm{Tr}\left[(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)\right] =$$

$$= \mathrm{Tr}\left[|\psi\rangle\langle\psi| - \langle\psi|\phi\rangle\,|\psi\rangle\langle\phi| - \langle\phi|\psi\rangle\,|\phi\rangle\langle\psi| + |\phi\rangle\langle\phi|\right] =$$
$$= 2 - |\,\langle\psi|\phi\rangle\,|^2 - |\,\langle\phi|\psi\rangle\,|^2 = 2 - 2F^2(|\phi\rangle, |\psi\rangle).$$

Since $\mathrm{Tr}(N) = 0$ and $\mathrm{rank}(N) \leq 2$, it follows that N has two non-zero eigenvalues which are $\pm\lambda$. In this case, we have

$$\mathrm{Tr}\big(N^2\big) = 2\lambda^2.$$

Therefore,

$$2\lambda^2 = 2 - 2F^2(|\phi\rangle, |\psi\rangle),$$
$$\lambda = \sqrt{1 - F^2(|\phi\rangle, |\psi\rangle)},$$
$$|| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \, ||_p = \left(2\left(1 - F^2(|\phi\rangle, |\psi\rangle)\right)^{p/2}\right)^{1/p} = 2^{1/p}\sqrt{1 - F^2(|\phi\rangle, |\psi\rangle)}.$$

□

In the following lemma we state and prove a useful fact that pure quantum states which are close to each other in Euclidean distance, are also close to each other in trace distance.

Lemma 3. *Let* $|\psi\rangle, |\phi\rangle$ *be arbitrary qudits. Suppose that* $|| \, |\psi\rangle - |\phi\rangle \, ||_2 \leq \epsilon$. *Then the following holds*

$$|| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \, ||_2 \leq O(\epsilon),$$

and, consequently,

$$|| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \, ||_1 \leq O(\epsilon).$$

Proof. Suppose $|\psi\rangle = \sum_{j=1}^{n} \left(a_j + ia'_j\right) |jj\rangle$ and $|\phi\rangle = \sum_{j=1}^{n} \left(b_j + ib'_j\right) |jj\rangle$. Then

$$\langle\psi|\phi\rangle = \left(\sum_{k=1}^{n} \left(a_k - ia'_k\right) \langle kk|\right) \left(\sum_{j=1}^{n} \left(b_j + ib'_j\right) |jj\rangle\right) = \sum_{j=1}^{n} \left(a_j b_j + ia_j b'_j - ia'_j b_j + a'_j b'_j\right),$$

$$\mathrm{Re}\, \langle\psi|\phi\rangle = \sum_{j=1}^{n} \left(a_j b_j + a'_j b'_j\right).$$

Since we assumed that both quantum states are normalized, we have

$$\langle\psi|\psi\rangle = \sum_{j=1}^{n} \left(a_j^2 + (a'_j)^2\right) = 1,$$

$$\langle\phi|\phi\rangle = \sum_{j=1}^{n} \left(b_j^2 + (b'_j)^2\right) = 1.$$

Therefore,

$$|| \, |\psi\rangle - |\phi\rangle \, ||_2^2 = \sum_{j=1}^{n} \left(\left(a_j - b_j\right)^2 + \left(a'_j - b'_j\right)^2\right) =$$

$$= \sum_{j=1}^{n} \left(a_j^2 - 2a_j b_j + b_j^2 + (a'_j)^2 - 2a'_j b'_j + (b'_j)^2\right) = 2 - 2\,\mathrm{Re}\, \langle\psi|\phi\rangle.$$

By the Lemma 2 we have

$$|| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \, ||_2^2 = 2 - 2|\, \langle\psi|\phi\rangle \, |^2.$$

Now, suppose that $|| \, |\psi\rangle - |\phi\rangle \, ||_2^2 \leq \epsilon$. It implies that

$$\mathrm{Re}\, \langle\psi|\phi\rangle \geq 1 - \frac{\epsilon}{2}.$$

Thus,

$$|\langle\psi|\phi\rangle|^2 = (\mathrm{Re}\,\langle\psi|\phi\rangle)^2 + (\mathrm{Im}\,\langle\psi|\phi\rangle)^2 \geq \left(1-\frac{\epsilon}{2}\right)^2 + (\mathrm{Im}\,\langle\psi|\phi\rangle)^2 \geq \left(1-\frac{\epsilon}{2}\right)^2 = 1 - O(\epsilon).$$

Therefore,

$$||\,|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\,||_2^2 \leq 2 - 2 + O(\epsilon) = O(\epsilon).$$

We proved that

$$||\,|\psi\rangle - |\phi\rangle\,||_2 \leq \epsilon \implies ||\,|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\,||_2 \leq O(\epsilon).$$

Since $\sqrt{2}||\,|\psi\rangle - |\phi\rangle\,||_2 = ||\,|\psi\rangle - |\phi\rangle\,||_1$, we also have

$$||\,|\psi\rangle - |\phi\rangle\,||_2 \leq \epsilon \implies ||\,|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\,||_1 \leq O(\epsilon).$$

□

1.1.6 Pauli Gates

In this subsection we define basic operations which can be performed on qubits and which will appear in this thesis. They are commonly referred to as Pauli gates.

Definition 11 (Pauli-X gate)**.** *A Pauli-X gate, also known as a bit-flip gate, is defined by having the following action on a computational basis:* $X|0\rangle = |1\rangle$, $X|1\rangle = |0\rangle$. *Its matrix representation in the computational basis is*

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Definition 12 (Pauli-Y gate)**.** *A Pauli-Y gate is defined by having the following action on a computational basis:* $Y|0\rangle = i|1\rangle$, $Y|1\rangle = -i|0\rangle$. *Its matrix representation in the computational basis is*

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

Definition 13 (Pauli-Z gate)**.** *A Pauli-Z gate, also known as a phase-flip gate, is defined by having the following action on a computational basis:* $Z|0\rangle = |0\rangle$, $Z|1\rangle = -|1\rangle$. *Its matrix representation in the computational basis is*

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

1.2 Quantum Entanglement

In this section we define entangled quantum states. Then, we introduce and prove Schmidt Decomposition Theorem which is an important tool for expressing them. Finally, we explain why quantum entanglement can be considered a resource in quantum information theory.

1.2.1 Separable and Entangled Quantum States

Definition 14 (Separable state)**.** *A pure state $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ is called separable if there exist pure quantum states $|\psi_A\rangle \in \mathcal{H}_\mathcal{A}$ and $|\psi_B\rangle \in \mathcal{H}_\mathcal{B}$ such that*

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle .$$

Definition 15 (Entangled state)**.** *A pure state $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ is called entangled if it is not separable.*

Definition 16 (Bell states)**.** *The following four types of maximally entangled qubits are called Bell states*

$$\begin{aligned}
\left|\Phi_2^+\right\rangle_{AB} &= \frac{1}{\sqrt{2}} |00\rangle_{AB} + \frac{1}{\sqrt{2}} |11\rangle_{AB} , \\
\left|\Phi_2^-\right\rangle_{AB} &= \frac{1}{\sqrt{2}} |00\rangle_{AB} - \frac{1}{\sqrt{2}} |11\rangle_{AB} , \\
\left|\Psi_2^+\right\rangle_{AB} &= \frac{1}{\sqrt{2}} |01\rangle_{AB} + \frac{1}{\sqrt{2}} |10\rangle_{AB} , \\
\left|\Psi_2^-\right\rangle_{AB} &= \frac{1}{\sqrt{2}} |01\rangle_{AB} - \frac{1}{\sqrt{2}} |10\rangle_{AB} .
\end{aligned}$$

We note that $\left|\Phi_2^+\right\rangle_{AB}$ is often referred to as an Einstein-Podolsky-Rosen (EPR) pair and $\left|\Phi_2^-\right\rangle_{AB}$ is often referred to as a singlet state.

1.2.2 Schmidt Decomposition

An extremely useful tool for analyzing pure bipartite quantum states is the Schmidt Decomposition. It allows us to characterize a quantum state by real positive coefficients and corresponding bases. Moreover, it limits the size of the representation to the size of the smaller one of two Hilbert spaces involved.

Theorem 1.2.1 (Schmidt decomposition)**.** *Suppose* $|\psi\rangle_{AB} \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ *is a pure bipartite quantum state and* $\mathcal{H}_\mathcal{A}, \mathcal{H}_\mathcal{B}$ *are Hilbert spaces of finite, possibly different dimensions. Then, it is possible to find a vector of real and strictly positive coefficients* $\{\lambda_i\}$ *and orthonormal bases* $\{|i\rangle_A\}$ *and* $\{|i\rangle_B\}$ *in* $\mathcal{H}_\mathcal{A}$ *and* $\mathcal{H}_\mathcal{B}$ *respectively, such that*

$$|\psi\rangle_{AB} = \sum_{i=0}^{d-1} \lambda_i |i\rangle_A |i\rangle_B .$$

We say that d *is the Schmidt rank of* $|\psi\rangle_{AB}$ *and* $\{\lambda_i\}$ *are Schmidt coefficients. Moreover,* $d \leq \min\{\dim(\mathcal{H}_\mathcal{A}), \dim(\mathcal{H}_\mathcal{B})\}$.

Proof. This proof follows [21]. Consider a pure bipartite quantum state $|\psi\rangle_{AB} \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. Generally, it can be written in terms of some orthonormal bases $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$ of $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$ respectively as

$$|\psi\rangle_{AB} = \sum_{i=0}^{d_A-1} \sum_{j=0}^{d_B-1} a_{ij} |i\rangle_A |j\rangle_B ,$$

where $d_A = \dim(\mathcal{H}_\mathcal{A})$ and $d_B = \dim(\mathcal{H}_\mathcal{B})$. We define a matrix of coefficients C as follows

$$[C]_{ij} = a_{ij}.$$

We use the Singular Value Decomposition to express the matrix C as a product of three matrices

$$C = UDV,$$

where D is a diagonal matrix with strictly positive entries (unique up to reordering) and U, V are unitary. We define $[D]_{ii} = \lambda_i$, $[U]_{ik} = u_{ik}$ and $[V]_{kj} = v_{kj}$. Then, the singular value decomposition is equivalent to

$$a_{ij} = \sum_{k=0}^{d-1} u_{ik} \lambda_k v_{kj}.$$

We substitute above into the general description of $|\psi\rangle_{AB}$ and obtain

$$|\psi\rangle_{AB} = \sum_{i=0}^{d_A-1} \sum_{j=0}^{d_B-1} \left(\sum_{k=0}^{d-1} u_{ik} \lambda_k v_{kj} \right) |i\rangle_A |j\rangle_B =$$

$$= \sum_{k=0}^{d-1} \lambda_k \left(\sum_{i=0}^{d_A-1} u_{ik} |i\rangle_A \right) \left(\sum_{j=0}^{d_B-1} v_{kj} |j\rangle_B \right) = \sum_{k=0}^{d-1} \lambda_k |k\rangle_A |k\rangle_B ,$$

where we defined $|k\rangle_A = \sum_{i=0}^{d_A-1} u_{ik} |i\rangle_A$ and $|k\rangle_B = \sum_{j=0}^{d_B-1} u_{kj} |j\rangle_B$. We shall also verify that our new bases $\{|k\rangle_A\}$ and $\{|k\rangle_B\}$ are both orthonormal.

$$\langle k|l\rangle_A = \left(\sum_{m=0}^{d_A-1} u_{mk}^\dagger \langle m|_A\right)\left(\sum_{i=0}^{d_A-1} u_{il} |i\rangle_A\right) = \sum_{m=0}^{d_A-1}\sum_{i=0}^{d_A-1} u_{mk}^\dagger u_{il} \langle m|i\rangle_A = \sum_{i=0}^{d_A-1} u_{ik}^\dagger u_{il} = \delta_{kl},$$

where we used the fact that columns of the unitary matrix U form an orthonormal basis. Analogously, $\{|k\rangle_B\}$ is an orthonormal basis. □

Now, we state the fact which can be seen as a conservation law for Schmidt coefficients under local unitary transformations.

Lemma 4. *Schmidt coefficients of a pure bipartite quantum state $|\psi\rangle_{AB} \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ are invariant under local unitary transformations.*

Proof. Consider a Schmidt Decomposition of a quantum state $|\psi\rangle_{AB} = \sum_{i=0}^{d-1} \lambda_i |i\rangle_A |i\rangle_B$. We recall that $\lambda_i > 0$ and bases $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ in $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$ are orthonormal. By applying any unitary local transformation $U_A \otimes U_B$ we obtain

$$U_A \otimes U_B |\psi\rangle_{AB} = \sum_{i=0}^{d-1} \lambda_i U_A |i\rangle_A U_B |i\rangle_B = \sum_{i=0}^{d-1} \lambda_i |i'\rangle_A |i'\rangle_B ,$$

where we defined new bases $|i'\rangle_A = U_A |i\rangle_A$ and $|i'\rangle_B = U_B |i\rangle_B$. It is easy to see that the new bases are also orthonormal

$$\langle j'|i'\rangle_A = \langle j| U_A^\dagger U_A |i\rangle_A = \langle j|i\rangle_A = \delta_{ij},$$

and analogously for $\{|i'\rangle_B\}$. Therefore, we see that by applying local unitary transformations we obtained another valid Schmidt decomposition, we the same, possibly reordered, Schmidt coefficients. □

1.2.3 Quantifying Entanglement

Having introduced entangled states, we would like to have a measure which quantifies the amount of entanglement in them. For instance, we would like to see separable states as states with no entanglement and Bell states, maximally entangled states of two qubits, as states with more entanglement than any other state of two qubits. One such a measure, which we will use in this thesis, is the von Neumann entanglement entropy which we define a follows.

Definition 17 (von Neumann Entanglement Entropy)**.** *Let ρ_{AB} be a density operator of a bipartite quantum state. The von Neumann entanglement entropy of ρ_{AB} is defined as*

$$S(\rho_A) = -\operatorname{Tr}(\rho_A \log \rho_A),$$

where $\rho_A = \operatorname{Tr}_B \rho_{AB}$ is obtained by a partial trace operation. It holds that $S(\rho_A) = S(\rho_B)$.

1.2.4 Entanglement as an Interconvertible Resource

Quantum entanglement has a wide spectrum of applications and is in fact one of the essential elements of quantum information theory which offers advantages over classical theories. It is therefore well-motivated to investigate whether quantum entanglement can be considered as a tangible resource. One of the properties that we might be interested in this context is inter-convertibility between different forms of entanglement. This problem was studied by Bennett, Bernstein, Popescu and Schumacher in [2]. They showed the feasibility of the so called concentration and dilution of quantum entanglement. The concentration of entanglement is a task in which Alice and Bob share some number of partially entangled pure states and would like to transform them into some number of maximally entangled pure states. Formally, it can be defined as follows.

Definition 18. *The concentration of entanglement is a quantum process in which the following transformation occurs*

$$|\psi\rangle_{AB}^{\otimes R_c n} \xrightarrow{LO} \left|\Phi_2^+\right\rangle_{AB}^{\otimes n},$$

where R_c is the rate of concentration.

The dilution of entanglement is a task in which Alice and Bob share some number of maximally entangled pure states and would like to transform them into some number of partially entangled pure states using Local Operations and Classical Communication (LOCC). The dilution of entanglement will be studied a lot in this thesis. Formally, it can be defined as follows.

Definition 19. *Dilution of entanglement is a quantum process in which the following transformation occurs*

$$\left|\Phi_2^+\right\rangle_{AB}^{\otimes R_d n} \xrightarrow{LOCC} |\psi\rangle_{AB}^{\otimes n},$$

where R_d is the rate of dilution.

The authors of [2] showed that we can concentrate and dilute entanglement back and forth in the asymptotic limit of n. Therefore, quantum entanglement is an interconvertible resource.

1.3 Applications of Quantum Entanglement

As we briefly discussed, quantum entanglement is a fungible resource in quantum information theory. It is a resource which is also extremely important because it allows to perform certain tasks which are impossible within the classical framework. In this section, we will describe three famous protocols for which quantum entanglement is essential: quantum teleportation, quantum superdense coding and quantum key distribution.

1.3.1 Quantum Teleportation

The quantum teleportation protocol [3] allows for a perfect transmission of a qubit by using local unitary operations and classical communication. A qubit to be transmitted may not be known to a sender and/or a receiver.

Quantum teleportation protocol [3]

Resources:
Alice: an unknown quantum state $|\psi\rangle_{A'} = \alpha |0\rangle_{A'} + \beta |1\rangle_{A'}$,
Shared by Alice and Bob: an EPR pair $\left|\Phi_2^+\right\rangle_{AB}$ and a noiseless classical communication channel.
Goal: Bob possesses a state $|\psi\rangle_{B'}$.
Protocol:

1. Alice measures her part of the state $|\psi\rangle_{A'} \left|\Phi_2^+\right\rangle_{AB}$ in the Bell basis, i.e., $\left\{\left|\Phi_2^+\right\rangle\!\left\langle\Phi_2^+\right|_{AA'}, \left|\Phi_2^-\right\rangle\!\left\langle\Phi_2^-\right|_{AA'}, \left|\Psi_2^+\right\rangle\!\left\langle\Psi_2^+\right|_{AA'}, \left|\Psi_2^-\right\rangle\!\left\langle\Psi_2^-\right|_{AA'}\right\}$.
2. Depending on the result of the measurement, Alice sends two classical bits to Bob; 00 for $\left|\Phi_2^+\right\rangle_{AA'}$, 01 for $\left|\Phi_2^-\right\rangle_{AA'}$, 10 for $\left|\Psi_2^+\right\rangle_{AA'}$ and 11 for $\left|\Psi_2^-\right\rangle_{AA'}$.
3. Depending on the classical bits received, Bob applies a unitary to his part of the maximally entangled state; I if 00, Z if 01, X if 10 and XZ if 11.

Theorem 1.3.1. *In the quantum teleportation protocol, an unknown qubit state $|\psi\rangle$ is teleported from Alice to Bob by using local unitary operations and sending two bits of classical communication from Alice to Bob.*

Proof. Suppose Alice possesses a quantum state $|\psi\rangle_{A'} = \alpha |0\rangle_{A'} + \beta |1\rangle_{A'}$ and shares a maximally entangled state $\left|\phi_2^+\right\rangle_{AB} = \frac{1}{\sqrt{2}} |00\rangle_{AB} + \frac{1}{\sqrt{2}} |11\rangle_{AB}$ with Bob. Then, their joint state is

$$|\psi\rangle_{A'} \left|\phi_2^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}(\alpha |000\rangle_{A'AB} + \beta |100\rangle_{A'AB} + \alpha |011\rangle_{A'AB} + \beta |111\rangle_{A'AB}).$$

We notice that the following identities hold

$$|00\rangle_{AB} = \frac{1}{\sqrt{2}}\left(\left|\Phi_2^+\right\rangle_{AB} + \left|\Phi_2^-\right\rangle_{AB}\right)$$
$$|01\rangle_{AB} = \frac{1}{\sqrt{2}}\left(\left|\Psi_2^+\right\rangle_{AB} + \left|\Psi_2^-\right\rangle_{AB}\right)$$
$$|10\rangle_{AB} = \frac{1}{\sqrt{2}}\left(\left|\Psi_2^+\right\rangle_{AB} - \left|\Psi_2^-\right\rangle_{AB}\right)$$
$$|11\rangle_{AB} = \frac{1}{\sqrt{2}}\left(\left|\Phi_2^+\right\rangle_{AB} - \left|\Phi_2^-\right\rangle_{AB}\right).$$

Therefore, our state can be rewritten as

$$\begin{aligned}|\psi\rangle_{A'}\left|\phi_2^+\right\rangle_{AB} = \frac{1}{2}[&\left|\Phi_2^+\right\rangle_{AB}(\alpha|0\rangle_B + \beta|1\rangle_B) + \left|\Phi_2^-\right\rangle_{AB}(\alpha|0\rangle_B - \beta|1\rangle_B) + \\ &+ \left|\Psi_2^+\right\rangle_{AB}(\alpha|1\rangle_B + \beta|0\rangle_B) + \left|\Psi_2^-\right\rangle_{AB}(\alpha|1\rangle_B - \beta|0\rangle_B)] = \\ = \frac{1}{2}\Big(&\left|\Phi_2^+\right\rangle_{AB}|\psi\rangle_B + \left|\Phi_2^-\right\rangle_{AB} Z|\psi\rangle_B + \left|\Psi_2^+\right\rangle_{AB} X|\psi\rangle_B + \left|\Psi_2^-\right\rangle_{AB} XZ|\psi\rangle_B\Big).\end{aligned}$$

Therefore, it is easy to see, that after Alice measuring in the Bell basis, the residual state in the Bob's register B is one from the set $\{|\psi\rangle_B, Z|\psi\rangle_B, X|\psi\rangle_B, XZ|\psi\rangle_B\}$. Then, it is enough for Alice to send two classical bits to Bob, which depend on her outcome of the measurement, such that he knows which unitary transformation to apply to obtain the state $|\psi\rangle_B$. □

1.3.2 Quantum Superdense Coding

The quantum superdense coding protocol [4] uses a noiseless qubit channel and local unitary operations to communicate classical information. It can be considered a protocol complementary to the quantum teleportation protocol.

Quantum superdense coding protocol [4]

Resources:
Alice: two classical bits,
Shared by Alice and Bob: an EPR pair $\left|\phi_2^+\right\rangle_{AB}$ and a noiseless quantum communication channel.
Goal: Bob learns Alice's two classical bits.
Protocol:

1. Depending on two classical bits to be communicated, Alice applies a unitary to her part of the

state $\left|\Phi_2^+\right\rangle_{AB}$; I if 00, Z if 01, X if 10 and XZ if 11.

2. Alice transmits her qubit to Bob.

3. Bob measures both of his qubits in the Bell basis, i.e., $\left\{\left|\Phi_2^+\right\rangle\!\left\langle\Phi_2^+\right|_{AB}, \left|\Phi_2^-\right\rangle\!\left\langle\Phi_2^-\right|_{AB}, \left|\Psi_2^+\right\rangle\!\left\langle\Psi_2^+\right|_{AB}, \left|\Psi_2^-\right\rangle\!\left\langle\Psi_2^-\right|_{AB}\right\}$.

4. Bob interprets the result of the measurement as bits that Alice wanted to communicate. 00 for $\left|\Phi_2^+\right\rangle_{AB}$, 01 for $\left|\Phi_2^-\right\rangle_{AB}$, 10 for $\left|\Psi_2^+\right\rangle_{AB}$ and 11 for $\left|\Psi_2^-\right\rangle_{AB}$.

Theorem 1.3.2. *In the quantum superdense coding protocol, a single use of a noiseless qubit channel and applications of local unitary operations is enough to communicate two classical bits from Alice to Bob.*

Proof. We notice that the following identities hold

$$I\left|\Phi_2^+\right\rangle_{AB} = \left|\Phi_2^+\right\rangle_{AB}$$

$$X\left|\Phi_2^+\right\rangle_{AB} = \left|\Phi_2^-\right\rangle_{AB}$$

$$Z\left|\Phi_2^+\right\rangle_{AB} = \left|\Psi_2^+\right\rangle_{AB}$$

$$XZ\left|\Phi_2^+\right\rangle_{AB} = \left|\Psi_2^-\right\rangle_{AB}.$$

Therefore, once Alice sends her qubit to Bob through a noiseless qubit channel, he possesses one of the states on the right hand side above, i.e., one of the Bell states. Since Bell states are mutually orthogonal, they can be perfectly distinguished between each other. Indeed, a measurement in the Bell basis allows Bob to learn which state he possesses. This way, he learns which two classical bits Alice intended to communicate. □

1.3.3 Quantum Key Distribution (E91 protocol)

The E91 protocol [7], proposed by Ekert, is one of the first important applications of quantum entanglement. It allows Alice and Bob to establish a secret key or detect the presence of an adversary. Such a provably secret key can be then used for provably secure communication using the so called one-time pad method. The E91 protocol is the modification of the well-known BB84 protocol [1] which accomplishes the same task. The main difference between them is that the BB84 protocol requires quantum communication between Alice and Bob and for the E91 protocol it is enough for Alice and Bob to have shared entanglement.

E91 protocol [7]

Input:
Alice and Bob: supplied n quantum states claimed to be maximally entangled states $\left|\Psi_2^-\right\rangle_{AB}$ (singlets)
Resources:
Shared by Alice and Bob: a public classical communication channel.
Goal: Alice and Bob establish a shared secret key
Protocol:

1. Suppose Alice and Bob can measure a qubit along one of the vectors $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ respectively, which correspond to measurements in a computational basis $\{|0\rangle, |1\rangle\}$ rotated by angles $\left\{0, \frac{\pi}{4}, \frac{\pi}{2}\right\}$ and $\left\{\frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}$ respectively. Alice and Bob measure their share of maximally entangled states along one of the vectors uniformly at random and independently of each other.

2. Alice and Bob publicly announce their measurement vectors. They keep measurement results of qubits for which they used the same basis, i.e., along $(\mathbf{a}_2, \mathbf{b}_1)$ and $(\mathbf{a}_3, \mathbf{b}_2)$, as a potential secret key.

3. The rest of the measurement results are announced publicly and analyzed to detect a potential adversary. Alice and Bob calculate the empirical value of the correlation coefficient

$$S = E(\mathbf{a}_1, \mathbf{b}_1) - E(\mathbf{a}_1, \mathbf{b}_3) + E(\mathbf{a}_3, \mathbf{b}_1) + E(\mathbf{a}_3, \mathbf{b}_3),$$

   where $E(\mathbf{a}_i, \mathbf{b}_j) = P_{++}(\mathbf{a}_i, \mathbf{b}_j) + P_{--}(\mathbf{a}_i, \mathbf{b}_j) - P_{+-}(\mathbf{a}_i, \mathbf{b}_j) - P_{-+}(\mathbf{a}_i, \mathbf{b}_j)$ and $P_{\pm\pm}(\mathbf{a}_i, \mathbf{b}_j)$ means the probability of obtaining results ± 1 and ± 1 when measuring along $\mathbf{a}_i$ and $\mathbf{b}_j$.

4. If Alice and Bob obtain $S \approx 2\sqrt{2}$, they use perfectly anti-correlated (and secret) results of measurements along $(\mathbf{a}_2, \mathbf{b}_1)$ and $(\mathbf{a}_3, \mathbf{b}_2)$ as their secret key. Otherwise, they assume that an eavesdropper tempered with singlets and they abort the protocol.

Theorem 1.3.3. *E91 protocol establishes a secret key of length about $\frac{n}{3}$ or provides statistical evidence for the presence of an adversary.*

Proof. This proof follows [7]. We consider the correlation coefficient S from the protocol, $S = E(\mathbf{a}_1, \mathbf{b}_1) - E(\mathbf{a}_1, \mathbf{b}_3) + E(\mathbf{a}_3, \mathbf{b}_1) + E(\mathbf{a}_3, \mathbf{b}_3)$. We now show, that if singlets are not tempered with, the coefficient S has a value $S = -2\sqrt{2}$. Suppose $\mathbf{a}_i$ and $\mathbf{b}_j$ are measurements in a computational basis rotated by angles α and β respectively. We introduce rotation matrices given by

$$S_\alpha = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix},$$

$$S_\beta = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix}.$$

The spin-$\frac{1}{2}$ measurement of two particles along directions α and β is given by $M_{\alpha,\pm}\otimes M_{\beta,\pm} = \frac{1}{4}(I \pm S_\alpha)\otimes(I \pm S_\beta)$. Since it is a symmetric projection, the probabilities from the protocol can be calculated as follows

$$P_{++}(\mathbf{a}_i, \mathbf{b}_j) = \left\langle \Psi_2^- \right| M_{\alpha,+} \otimes M_{\beta,+} \left| \Psi_2^- \right\rangle,$$
$$P_{+-}(\mathbf{a}_i, \mathbf{b}_j) = \left\langle \Psi_2^- \right| M_{\alpha,+} \otimes M_{\beta,-} \left| \Psi_2^- \right\rangle,$$
$$P_{-+}(\mathbf{a}_i, \mathbf{b}_j) = \left\langle \Psi_2^- \right| M_{\alpha,-} \otimes M_{\beta,+} \left| \Psi_2^- \right\rangle,$$
$$P_{--}(\mathbf{a}_i, \mathbf{b}_j) = \left\langle \Psi_2^- \right| M_{\alpha,-} \otimes M_{\beta,-} \left| \Psi_2^- \right\rangle.$$

Using the results above, we obtain

$$E(\mathbf{a}_i, \mathbf{b}_j) = P_{++}(\mathbf{a}_i, \mathbf{b}_j) - P_{+-}(\mathbf{a}_i, \mathbf{b}_j) - P_{-+}(\mathbf{a}_i, \mathbf{b}_j) + P_{--}(\mathbf{a}_i, \mathbf{b}_j) = -\cos(\alpha - \beta) = -\mathbf{a}_i \cdot \mathbf{b}_j.$$

Therefore, using the actual angles of measurements and remembering that we deal with unit vectors, we obtain

$$S = -(\mathbf{a}_1 \cdot \mathbf{b}_1 - \mathbf{a}_1 \cdot \mathbf{b}_3 + \mathbf{a}_3 \cdot \mathbf{b}_1 + \mathbf{a}_3 \cdot \mathbf{b}_3) = -\cos\left(0 - \frac{\pi}{4}\right) + \cos\left(0 - \frac{3\pi}{4}\right) -$$

$$-\cos\left(\frac{\pi}{2} - \frac{\pi}{4}\right) - \cos\left(\frac{\pi}{2} - \frac{3\pi}{4}\right) = -3\cos\left(\frac{\pi}{4}\right) + \cos\left(\frac{3\pi}{4}\right) = -2\sqrt{2}.$$

Let us now consider the case in which an eavesdropper interferes with singlets to later obtain some information about the secret key. All the eavesdropper can do is to try measuring qubits that form a singlet along a certain direction which may vary from pair to pair, depending on whatever malicious strategy the eavesdropper may have. In this scenario, the correlation coefficient is of the form

$$S = \int \rho(\mathbf{n}_a, \mathbf{n}_b) d\mathbf{n}_a d\mathbf{n}_b [(\mathbf{a}_1 \cdot \mathbf{n}_a)(\mathbf{b}_1 \cdot \mathbf{n}_b) - (\mathbf{a}_1 \cdot \mathbf{n}_a)(\mathbf{b}_3 \cdot \mathbf{n}_b) +$$

$$+(\mathbf{a}_3 \cdot \mathbf{n}_a)(\mathbf{b}_1 \cdot \mathbf{n}_b) + (\mathbf{a}_3 \cdot \mathbf{n}_a)(\mathbf{b}_3 \cdot \mathbf{n}_b)],$$

where $\mathbf{n}_a, \mathbf{n}_b$ are directions along which measurements were performed by an eavesdropper on Alice's and Bob's particles respectively. In our protocol, it can be further simplified by substituting the actual values for measurement directions. Suppose that measurements along $\mathbf{n}_a$ and $\mathbf{n}_b$ are parametrized by angles α and β respectively. Remembering that we deal with unit vectors, we have

$$S = \int \rho(\mathbf{n}_a, \mathbf{n}_b) d\mathbf{n}_a d\mathbf{n}_b \left[\cos(\alpha - 0)\cos\left(\beta - \frac{\pi}{4}\right) - \cos(\alpha - 0)\cos\left(\beta - \frac{3\pi}{4}\right) +\right.$$

$$+\cos\Big(\alpha-\frac{\pi}{2}\Big)\cos\Big(\beta-\frac{\pi}{4}\Big)+\cos\Big(\alpha-\frac{\pi}{2}\Big)\cos\Big(\beta-\frac{3\pi}{4}\Big)\Big]=$$

$$=\int\rho(\mathbf{n}_a,\mathbf{n}_b)d\mathbf{n}_a d\mathbf{n}_b\left[\sqrt{2}\cos(\alpha-\beta)\right]=\int\rho(\mathbf{n}_a,\mathbf{n}_b)d\mathbf{n}_a d\mathbf{n}_b\left[\sqrt{2}\mathbf{n}_a\cdot\mathbf{n}_b\right]=$$

$$=\sqrt{2}\int\rho(\mathbf{n}_a,\mathbf{n}_b)d\mathbf{n}_a d\mathbf{n}_b\left[\mathbf{n}_a\cdot\mathbf{n}_b\right].$$

By examining the integral, we see that it is lower and upper bounded by -1 and 1 respectively. Thus, the coefficient S can take values from the range

$$-\sqrt{2}\leq S\leq\sqrt{2}.$$

Therefore, we showed that based on the value of S that Alice and Bob calculate empirically, they can obtain the statistical evidence of the presence of an eavesdropper. If they are convinced that they are not present, they can use their secret results from measurements along the same direction to establish a secret binary key. We notice that if Alice and Bob were sent n singlet pairs and performed n measurements on them, on average one third of their measurements should happen in the same basis. Therefore, the number of bits that can constitute for a secret key is roughly $\frac{n}{3}$. □

Chapter 2

Embezzlement of Entanglement

As explained in the previous chapter, Schmidt coefficients are invariant under local unitary transformations. It turns out, however, that this conservation law can be broken approximately (but with arbitrary accuracy). It is possible if Alice and Bob posses an additional resource called an embezzling quantum state. This result is referred to as the embezzlement of entanglement and was proposed by van Dam and Hayden [17]. In this chapter, we will provide the statement and the proof of this result. Next, following results of Leung and Wang [14], we will consider generalized characterizations of embezzling states and discuss their efficiency. First, we will define universal embezzling families, i.e., those allowing for embezzlement using the same embezzling state independent of the target state and then non-universal embezzling families and discuss their efficiency. Then, we describe a non-universal embezzlement scheme due to Leung, Toner and Watrous [13]. Finally, we provide a linear programming characterization of embezzlement of entanglement.

2.1 Canonical Universal Embezzling Family

In this section we present the seminal result about embezzlement of entanglement from [17]. It shows that Alice and Bob can prepare any pure bipartite entangled state up to arbitrary accuracy using only local unitary operations, as long as they share a universal (i.e. not target state dependent) state of a specific form and size large enough.

Theorem 2.1.1. *For every entangled bipartite quantum state with a Schmidt decomposition* $|\phi\rangle_{AB} = \sum_{i=1}^{d} \alpha_i |i\rangle_A |i\rangle_B$*, there exist* n *and local unitary operations* U_A, U_B *such that*

$$F(U_A^\dagger \otimes U_B^\dagger |\mu(n)\rangle_{AB} |00\rangle_{AB}, |\mu(n)\rangle_{AB} |\phi\rangle_{AB}) \geq 1 - \epsilon,$$

where $|\mu(n)\rangle_{AB} = \frac{1}{\sqrt{C(n)}} \sum_{i=1}^{n} \frac{1}{\sqrt{j}} |jj\rangle_{AB}$, $C(n)$ *is a normalization constant,* $n > d^{1/\epsilon}$ *and* $\epsilon = \frac{\log d}{\log n}$.

Proof. The proof follows [17]. We define the state $|\omega(n)\rangle := \sum_{j=1}^{dn} \omega_j |jj\rangle_{AB}$ which has the same Schmidt basis and coefficients as our target state $|\mu(n)\rangle_{AB} |\phi\rangle_{AB}$ but sorted in decreasing order. We start the proof by showing that first n coefficients of $|\omega(n)\rangle$ are smaller than those of $|\mu(n)\rangle_{AB} |00\rangle_{AB}$. These coefficients of $|\omega(n)\rangle$ are of the form $\frac{\alpha_i}{\sqrt{jC(n)}}$. Suppose we fix numbers i and t and define N_i^t to be the number of coefficients $\frac{\alpha_i}{\sqrt{jC(n)}}$ that are strictly greater than $\frac{1}{\sqrt{tC(n)}}$. This inequality implies the restriction that $1 \leq j \leq \alpha_i^2 t$. Therefore, we have $N_i^t < \alpha_i^2 t$. By summing this inequality over i, we obtain $\sum_{i=1}^{m} N_i^t < \sum_{i=1}^{m} \alpha_i^2 t = t$, where we used the fact that α_i's describe a normalized quantum state. Consider the case when $t = 1$. Then, $\sum_{i=1}^{m} N_i^1 < 1$, which means that no coefficient of the form $\frac{\alpha_i}{\sqrt{jC(n)}}$ is greater than $\frac{1}{\sqrt{C(n)}}$. We recall the order $\omega_1, \omega_2, \ldots, \omega_{dn}$ that we assumed on coefficients ω_j's. Combining it with the previous statement, we know that $\omega_1 \leq \frac{1}{\sqrt{C(n)}}$. Considering next values of t, it is easy to see that $\omega_j \leq \frac{1}{\sqrt{jC(n)}}$ for all $1 \leq j \leq n$. This inequality lets us bound the fidelity between $|\omega(n)\rangle$ and $|\mu(n)\rangle_{AB} |00\rangle_{AB}$ in the following way

$$F(|\omega(n)\rangle, |\mu(n)\rangle_{AB} |00\rangle_{AB}) = \sum_{j=1}^{n} \frac{\omega_j}{\sqrt{jC(n)}} \geq \sum_{j=1}^{n} \omega_j^2.$$

Now, we would like to show that the fidelity above is close to 1 for n large enough. To do so, we consider the maximally entangled state of Schmidt rank d, $|\Phi_d\rangle := \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle_{AB}$ and define $|\psi(n)\rangle := |\mu(n)\rangle |\Phi_d\rangle$. It is easy to see that $|\omega(n)\rangle$ majorizes $|\psi(n)\rangle$, i.e., $|\omega(n)\rangle \succ |\psi(n)\rangle$ which implies $\sum_{j=1}^{n} \omega_j^2 \geq \sum_{j=1}^{n} \beta_j^2$, where β_j's are coefficients of $|\psi(n)\rangle$ in decreasing order. The sum of n biggest coefficients of $|\psi(n)\rangle$ can be bounded as follows

$$\sum_{j=1}^{n} \beta_j^2 \geq \sum_{j=1}^{\lfloor n/d \rfloor} \sum_{i=1}^{d} \frac{1}{jC(n)d} \geq 1 - \frac{\log d}{\log n}.$$

Therefore, we obtained that

$$F(|\omega(n)\rangle, |\mu(n)\rangle_{AB} |00\rangle_{AB}) \geq 1 - \frac{\log d}{\log n}.$$

To achieve a fidelity of at least $1 - \epsilon$, we set $\epsilon = \frac{\log d}{\log n}$ which yields a constraint $n > d^{1/\epsilon}$ on the Schmidt rank of the embezzling state. We recall that $|\omega(n)\rangle$ is the state $|\mu(n)\rangle |\phi\rangle$ with

sorted coefficients, therefore $U_A \otimes U_B |\mu(n)\rangle |\phi\rangle = |\omega(n)\rangle$ for some local unitary operations U_A, U_B. It follows that

$$F(|\mu(n)\rangle |\phi\rangle , U_A^\dagger \otimes U_B^\dagger |\mu(n)\rangle_{AB} |00\rangle_{AB}) \geq 1 - \epsilon.$$

□

2.2 Properties of Embezzling Families

In this section we present and prove certain interesting properties of embezzling families. First, we will show that it is possible to embezzle any bipartite state $|\phi\rangle$ in superposition which is the result due to [14].

Lemma 5. *Let $|\mu(n)\rangle$ be a state capable of embezzling any state of Schmidt rank m with infidelity at most ϵ. Suppose we are given a set $\{|\phi_j\rangle\}_{j=1}^k$ of normalized states of the form*

$$|\phi_j\rangle = \sum_{p=1}^{m} \phi_{j,p} |m(j-1)+p\rangle |m(j-1)+p\rangle ,$$

where $\phi_{j,p} \in \mathbb{C}$ for $p = 1, \ldots, m$ and $j = 1, \ldots, k$. Then, there exist local unitary operations U_A, U_B such that

$$F(U_A \otimes U_B |\mu(n)\rangle_{AB} \sum_{j=1}^{k} \alpha_j |jj\rangle_{AB} , |\mu(n)\rangle_{AB} \sum_{j=1}^{k} \alpha_j |\phi_j\rangle_{AB}) \geq 1 - \epsilon,$$

where $\alpha_j \in \mathbb{C}$ for $j = 1, \ldots, k$ and $\sum_{j=1}^k |\alpha_j|^2 = 1$.

Proof. We will follow the proof given in [14]. By the embezzlement result, we know that

$$\forall j \quad \exists U_j : F(U_j \otimes U_j |\mu(n)\rangle |00\rangle , |\mu(n)\rangle \sum_{p=1}^{m} \phi_{j,p} |pp\rangle) \geq 1 - \epsilon.$$

Then, for every j we can construct a modified unitary U_j' which acts as

$$F(U_j' \otimes U_j' |\mu(n)\rangle |jj\rangle , |\mu(n)\rangle |\phi_j\rangle) \geq 1 - \epsilon.$$

We demand that U'_j has the property that $U'_j |\xi\rangle |j'\rangle = 0$ for all $|\xi\rangle$ whenever $j \neq j'$. Then we construct an isometry $U = \sum_j U'_j$ which satisfies

$$\left[\sum_{j'=1}^{k} \alpha_{j'}^{\dagger} \langle\mu(n)| \langle\phi'_j|\right] \left[U \otimes U \sum_{j=1}^{k} \alpha_j |\mu(n)\rangle |jj\rangle\right] =$$

$$= \left[\sum_{j'=1}^{k} \alpha_{j'}^{\dagger} \langle\mu(n)| \langle\phi'_j|\right] \left[\sum_{j=1}^{k} \alpha_j U'_j \otimes U'_j |\mu(n)\rangle |jj\rangle\right] \geq 1 - \epsilon.$$

□

Lemma 6. *Let $|\mu\rangle$ be an embezzling state which lets us embezzle any quantum state of Schmidt rank 2 with fidelity at least F. Then, it is possible to use it to embezzle any quantum state $|\phi\rangle$ of Schmidt rank m with fidelity at least F_m, where $1 - F_m^2 \leq \lceil \log_2 m \rceil^2 (1 - F^2)$.*

Proof. The proof can be found in [14]. □

2.3 Regular Universal Embezzling Families

The canonical embezzling family $\{|\mu(n)\rangle\}_{n=1}^{\infty}$ described in the earlier section is not the only one which has embezzling properties. Actually, authors in [14] introduced a generalized version whose coefficients are characterized by a decreasing function.

Definition 20. *An embezzling family $\{|\mu_R(n)\rangle\}_{n=1}^{\infty}$ is called regular if it consists of the states of the form*

$$|\mu_R(f,n)\rangle = \frac{1}{\sqrt{C(f,n)}} \sum_{i=1}^{n} f(i) |ii\rangle_{AB},$$

where $f : \mathbb{N} \to \mathbb{R}^+$ is a decreasing function of i and $C(f,n)$ is a normalization constant.

The following lemma, due to [14], provides a necessary condition for being a regular universal embezzling family.

Lemma 7. *If $\{|\mu(n)\rangle\}_{n=1}^{\infty}$ is a regular embezzling family, then $\lim_{n\to\infty} C(f,n) = \infty$.*

Proof. This result follows immediately from Lemma 9 in the next section. □

The following lemma, due to [14], provides a sufficient condition for being a regular universal embezzling family.

Lemma 8. *Suppose* $\lim_{n\to\infty} C(f,n) = \infty$. *If* $\forall |\varphi\rangle \lim_{n\to\infty} \rho(|\varphi\rangle, f, i) = 1$, *then* $\{|\mu(n)\rangle\}_{n=1}^{\infty}$ *is a regular universal embezzling family, where* $\rho(|\varphi\rangle, f, i) = \frac{\omega(i,n)}{\mu(i,n)}$, $\mu(i,n) = \frac{f(i)}{\sqrt{C(f,n)}}$ *and* $|\omega\rangle$ *is the state obtained after embezzlement but with coefficients sorted in a decreasing order, denoted by* $\omega(i,n)$.

Proof. We will follow the proof given in [14]. We assume that $\forall |\varphi\rangle \lim_{n\to\infty} \rho(|\varphi\rangle, f, i) = 1$. Then, for any $\epsilon > 0$, $\exists n_\epsilon$ such that $(1-\epsilon)\mu(i,n) < \omega(i,n) < (1+\epsilon)\mu(i,n)$ for all $i > n_\epsilon$. Then, we have

$$F(|\mu(f,n)\rangle, |\omega\rangle) = \sum_{i=1}^{2n} \mu(i,n)\omega(i,n) = \sum_{i=1}^{n_\epsilon} \mu(i,n)\omega(i,n) + \sum_{i=n_\epsilon+1}^{2n} \mu(i,n)\omega(i,n)$$

$$> \sum_{i=n_\epsilon+1}^{2n} \mu(i,n)\omega(i,n) > (1-\epsilon) \sum_{i=n_\epsilon+1}^{2n} \mu(i,n)^2 > (1-\epsilon) - \sum_{i=1}^{n_\epsilon} \mu(i,n)^2 > (1-\epsilon) - \frac{C(f,n_\epsilon)}{C(f,n)}.$$

Since we assumed that $\lim_{n\to\infty} C(f,n) = \infty$ and n_ϵ is not a function of n, we have that $\lim_{n\to\infty} F(|\mu(f,n)\rangle, |\omega\rangle) = 1$. Therefore, $\{|\mu(n)\rangle\}_{n=1}^{\infty}$ is a regular universal embezzling family. □

2.4 General Universal Embezzling Families

Extending the generalization from the previous section, we may assume that a function that characterizes coefficients is dependent not only on i but also on n. Embezzling families of this form are called general universal embezzling families and were also introduced in [14].

Definition 21. *An embezzling family* $\{|\mu_G(n)\rangle\}_{n=1}^{\infty}$ *is called general if it consists of the normalized quantum states of the form*

$$|\mu_G(n)\rangle = \sum_{i=1}^{n} \mu(i,n) |ii\rangle_{AB},$$

where $\mu : \mathbb{N} \times \mathbb{N} \to \mathbb{R}^+$ *is a decreasing function of* i.

The following lemma, due to [14], provides a necessary condition for being a general universal embezzling family.

Lemma 9. *If* $\{|\mu(n)\rangle\}_{n=1}^{\infty}$ *is a universal embezzling family, then* $\lim_{n\to\infty}\mu(1,n)=0$.

Proof. We will follow the proof given in [14]. Suppose $|\mu(n)\rangle$ is an embezzling state which can embezzle a state of Schmidt rank 2 with fidelity at least F. We will investigate a lower bound on fidelity which arises from considering a quantum state $|\phi\rangle = \frac{1}{\sqrt{2}}(|11\rangle + |22\rangle)$. Then, we have

$$1 - F(|\mu(n)\rangle, |\omega\rangle) = 1 - \sum_{i=1}^{2n}\mu(i,n)\omega(i,n) = \frac{1}{2}\sum_{i=1}^{2n}\mu(i,n)^2 + \frac{1}{2}\sum_{i=1}^{2n}\omega(i,n)^2$$

$$-\frac{1}{2}\sum_{i=1}^{2n}2\mu(i,n)\omega(i,n) = \frac{1}{2}\sum_{i=1}^{2n}(\mu(i,n)-\omega(i,n))^2 \geq \frac{1}{2}(\mu(1,n)-\omega(1,n))^2 =$$

$$= \frac{1}{2}\left(\mu(1,n) - \frac{\mu(1,n)}{\sqrt{2}}\right)^2 = \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)^2\mu(1,n)^2,$$

where $|\omega\rangle$ is the state obtained after embezzlement but with coefficients sorted in a decreasing order, denoted by $\omega(i,n)$. In the equation above we used normalization of coefficients, lower-bounded the sum by the biggest term and used the fact that $\omega_1 = \frac{\mu(1,n)}{\sqrt{2}}$. Since $\mu(1,n) > 0$ and in embezzlement we expect that $\lim_{n\to\infty} F(|\mu(n)\rangle, |\omega\rangle) = 1$ (e.g. consider the limit of the canonical embezzlement result), together with the equation above it implies that

$$\lim_{n\to\infty}\mu(1,n) = 0.$$

The result is proved for a state with Schmidt rank 2. However, by Lemma 6 we can extend it to a quantum state of any finite Schmidt rank m because the resulting fidelity will only differ by a constant dependent on m. □

2.5 Non-universal Embezzling Families

In this section we present a method for embezzlement with an embezzling state dependent on a target state, which was proposed by Leung, Toner and Watrous in [13]. In this approach, the embezzling state is of the form

$$|\mu\rangle_{AB}^{\psi} = \frac{1}{\sqrt{C}}\sum_{r=1}^{n}|\phi\rangle_{A_1B_1}\cdots|\phi\rangle_{A_rB_r}|\psi\rangle_{A_{r+1}B_{r+1}}\cdots|\psi\rangle_{A_{n+1}B_{n+1}} = \frac{1}{\sqrt{C}}\sum_{r=1}^{n}|\phi\rangle_{AB}^{\otimes r}|\psi\rangle_{AB}^{\otimes(n-r)},$$

where $|\phi\rangle$ can be any bipartite quantum state, not necessarily entangled. To cover the case when the overlap $\langle\phi|\psi\rangle$ is complex, i.e., of the form $ae^{i\theta}$, we assume that $|\psi\rangle$ is the state that we intend to embezzle shifted by a proper global phase $e^{-i\theta}$. In the following theorem, due to [13], we state and prove the Non-universal embezzlement result.

Theorem 2.5.1. *For every entangled state* $|\psi\rangle_{AB} = \sum_{i=1}^{d} \alpha_i |i\rangle_A |i\rangle_B$, *there exist local unitary operations* $U_{AA'}, U_{BB'}$ *such that*

$$F\left((U_{AA'} \otimes U_{BB'}) |\mu\rangle_{AB}^{\psi} |\phi\rangle_{A'B'}, |\mu\rangle_{AB}^{\psi} |\psi\rangle_{A'B'}\right) \geq 1 - \epsilon,$$

where $n > \frac{1}{\epsilon}$.

Proof. The proof follows [13]. Alice and Bob, having the state $|\mu\rangle_{AB}^{\psi} |\phi\rangle_{A'B'}$, shift every state to the neighbouring register (to the right). The state from the last register is shifted to the first register. This transformation is clearly reversible and executed locally by each of two parties, therefore can be implemented by local unitary operations $U_{AA'}$ and $U_{BB'}$ such that

$$(U_{AA'} \otimes U_{BB'}) |\mu\rangle_{AB} |\phi\rangle_{A'B'} = \left(\frac{1}{\sqrt{C}} \sum_{r=1}^{n} |\phi\rangle_{AB}^{\otimes(r+1)} |\psi\rangle_{AB}^{\otimes(n-r)}\right) |\psi\rangle_{A'B'}.$$

We can verify that

$$C = n\frac{1 + \langle\phi|\psi\rangle}{1 - \langle\phi|\psi\rangle} - 2\langle\phi|\psi\rangle \frac{1 - \langle\phi|\psi\rangle^n}{(1 - \langle\phi|\psi\rangle)^2}.$$

Then, we have

$$F\left((U_{AA'} \otimes U_{BB'}) |\mu\rangle_{AB} |\phi\rangle_{A'B'}, |\mu\rangle_{AB} |\psi\rangle_{A'B'}\right) = 1 - \frac{1 - \langle\phi|\psi\rangle}{C} \geq 1 - \frac{1 - \langle\phi|\psi\rangle}{n} \geq 1 - \frac{1}{n},$$

where we used that $n \leq C$. Therefore, it is enough to take $n > \frac{1}{\epsilon}$. □

2.6 Efficiency of Embezzling Families

In this Chapter we first introduced the canonical embezzling family which originates from the seminal paper regarding embezzlement of entanglement and is universal. Then, we defined two generalizations of this family. Finally, we showed how to perform embezzlement with a state-dependent catalyst. We might wonder which family is the most efficient, i.e.,

offers the best trade-off of fidelity of embezzlement to the required size of an embezzling state. In the canonical embezzlement paper [17], authors proved that we should not expect any other family to be much more efficient. They even significantly relaxed the setting, by allowing classical communication and a state-dependent catalyst. They showed that their family is still asymptotically optimal, up to a constant factor, when compared to any embezzling family in the relaxed setting. In this Section, we reproduce their argument.

First, we state the Fannes Inequality for density operators which will be crucial for proving the result about efficiency.

Theorem 2.6.1 (Fannes Inequality)**.** *Let ρ, σ be density matrices of dimension d. Then*

$$|S(\rho) - S(\sigma)| \leq \frac{1}{2}||\rho - \sigma||_1 \log(d-1) + H(||\rho - \sigma||_1),$$

where H is the binary entropy function and $S(\rho) = H((\lambda_i))$ is the von Neumann entropy defined on a matrix ρ whose eigenvalues form the vector (λ_i).

Proof. See [8]. □

Suppose that we want to embezzle a state $|\varphi\rangle_{AB}$ with a Schmidt rank m by using a state-dependent catalyst $|\xi_\varphi\rangle$ with a Schmidt rank n. Assume that we use the optimal protocol which uses LOCC (Local Operations and Classical Communication) and that for our case it produces a quantum state σ_{AB}. By the result in [18], σ_{AB} can be assumed to be a density operator representing a pure state which has the same Schmidt basis as $|\xi\rangle_{AB} \otimes |\varphi\rangle_{AB}$. Since an LOCC protocol cannot increase the amount of entanglement, we know that $S(\sigma_A) \leq S(\xi_A)$. Assuming that the embezzlement protocol guarantees that we obtain a state which is δ close to a desired state, i.e., $\operatorname{Tr}|\sigma_A - \xi_A \otimes \varphi_A| = \delta$, we can apply the Fannes inequality, as long as $\delta < \frac{1}{e}$, as follows

$$S(\varphi_A) \leq |S(\xi_A \otimes \varphi_A) - S(\sigma_A)| < \delta(\log m + \log n) - \delta \log \delta,$$

which implies

$$\frac{S(\varphi_A) + \delta \log \delta}{\log m + \log n} < \delta.$$

For the canonical embezzlement protocol, assuming that we perform the task with the Schatten-1 norm between states given by $\operatorname{Tr}|\omega(n)_A - \mu(n)_A| = \delta$ (for definitions of these states see the fragment of this thesis regarding canonical embezzlement), we obtain

$$\delta = \operatorname{Tr}|\omega(n)_A - \mu(n)_A| = \sum_{j=1}^{n}(\mu_j^2 - \omega_j^2) + \sum_{j=n+1}^{nm}\omega_j^2 \leq \frac{\log m}{\log n} - 1 + \sum_{j=1}^{n}\mu_j^2 + \sum_{j=n+1}^{nm}\omega_j^2 \leq$$

$$\leq \frac{\log m}{\log n} + \sum_{j=n+1}^{nm} \omega_j^2 \leq \frac{\log m}{\log n} + \sum_{j=1}^{n} \omega_j^2 \leq \frac{2\log m}{\log n},$$

where we used the facts that coefficients μ_j are normalized, coefficients ω_j are in a non-increasing order, $\mu_j = 0$ for $j > n$ and the bound from the proof of the canonical embezzlement. For the Schatten-1 distance, we used the fact that both states have the same Schmidt basis and that for pure states singular values are squares of eigenvalues of corresponding density operators. Therefore, for a fixed state that we want to embezzle, we have

$$\delta \geq \Omega\left(\frac{1}{\log n}\right),$$

for the best LOCC protocol with state-dependent catalyst, and

$$\delta \leq O\left(\frac{1}{\log n}\right),$$

for the canonical embezzlement. Thus, the canonical embezzlement is optimal up to a constant factor in δ achievable for a given n.

A detailed study of efficiency of embezzling families was conducted in [14]. They explicitly constructed another embezzling family which seems to be outperforming the canonical embezzling family for small sizes of a catalyst, based on numerical evidence.

2.7 Linear Programming Characterization of Embezzlement of Entanglement

In this section we present the characterization of embezzlement of entanglement as a linear program which is the first creative contribution of this thesis. The solution of this program gives an explicit unitary that Alice and Bob should use to achieve the maximal fidelity of the embezzlement protocol for a given embezzling family, its size and a particular target state that they want to embezzle.
First, we state several definitions and a theorem which will be crucial in proving the correctness of our linear program.

Definition 22 (Doubly-stochastic matrix)**.** *An $n \times n$ matrix M is doubly-stochastic if*

$$\forall i = 1, 2, \ldots, n \quad \sum_{j=1}^{n} M_{ij} = 1,$$

$$\forall j = 1, 2, \ldots, n \quad \sum_{i=1}^{n} M_{ij} = 1,$$
$$\forall i, j = 1, 2, \ldots, n \quad M_{ij} \geq 0.$$

Definition 23 (Permutation matrix)**.** *An $n \times n$ matrix is called a permutation matrix if it can be obtained by permuting rows (or columns) of an $n \times n$ identity matrix.*

Theorem 2.7.1 (The Birkhoff-von Neumann Theorem)**.** *The set of $n \times n$ doubly stochastic matrices defines a convex polytope with vertices being $n \times n$ permutation matrices.*

Proof. The theorem is stated and proved as Theorem 4.28 in [20]. □

Now, we are ready to state and prove our theorem, which gives the linear programming characterization of embezzlement of entanglement.

Theorem 2.7.2. *Let $|\mu(n)\rangle$ be any embezzling state of Schmidt rank n, characterized by the vector $\boldsymbol{c}$ of its Schmidt coefficients, padded with zeros to a proper dimension, let $|\phi\rangle$ be a bipartite quantum state of Schmidt rank m that we want to embezzle and let $|\mu(n)\rangle \otimes |\phi\rangle$ be the bipartite quantum state that we would like to have at the end of the protocol, characterized by the vector of Schmidt coefficients $\boldsymbol{t}$. Then, the following linear program computes the maximal fidelity of embezzling $|\phi\rangle$ with $|\mu(n)\rangle$ and when solved by the Simplex method provides an explicit unitary to do so*

$$\max \quad \boldsymbol{t}^T X \boldsymbol{c}$$
$$\forall i = 1, 2, \ldots, mn \quad \sum_{j=1}^{mn} X_{ij} = 1,$$
$$\forall j = 1, 2, \ldots, mn \quad \sum_{i=1}^{mn} X_{ij} = 1,$$
$$\forall i, j = 1, 2, \ldots, mn \quad X_{ij} \geq 0,$$

where X is an $(mn) \times (mn)$ matrix of decision variables.

Proof. The feasible region of the linear program defines the set of all doubly-stochastic matrices which clearly is a non-empty set for all $m, n > 0$. Due to the Birkhoff-von Neumann Theorem, it is a convex polytope with extreme points being permutation matrices. Since in feasible convex optimization problems there always exists an extreme point of a

feasible region which gives the optimal solution and the Simplex method only visits extreme points of a feasible region, the optimal solution of the linear program above will be given in terms of a permutation matrix. It is a known fact that permutation matrices are unitary. The objective function is a fidelity between a reshuffled embezzling state and a target state. Since it is a maximization problem, the unitary X is chosen such that this fidelity is maximized. □

2.8 Perfect Embezzlement

In this section we will briefly present results about perfect embezzlement of entanglement by Cleve, Liu and Paulsen [6]. We will focus on stating the results and describing frameworks within which they are valid, namely the tensor product framework and the commuting operator framework. These developments are based on the theory of C^*-algebras which are beyond the scope of this thesis.

Definition 24. *Perfect embezzlement is an embezzlement in which the fidelity between the desired and the final state of the protocol is exactly* 1*.*

2.8.1 Tensor Product Framework

The tensor product framework is the one widely used in quantum information theory. It assumes that a Hilbert space $\mathcal{H}$ associated with the system of interest can be partitioned into smaller Hilbert spaces. In particular, in a bipartite case, we may associate Alice with a Hilbert space $\mathcal{H}_A$ and Bob with a Hilbert space $\mathcal{H}_B$ such that $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Moreover, in case of embezzlement of entanglement, we may say that an embezzling state belongs to a Hilbert space $\mathcal{R} = \mathcal{R}_A \otimes \mathcal{R}_B$ and a target state will be created in a Hilbert space $\mathcal{T} = \mathcal{T}_A \otimes \mathcal{T}_B$ such that $\mathcal{H} = \mathcal{T}_A \otimes \mathcal{R}_A \otimes \mathcal{R}_B \otimes \mathcal{T}_B$. This framework captures the notion of locality, i.e., the fact that both parties might be spatially separated. Then Alice can only apply transformations to the Hilbert space accessible to her, i.e., $\mathcal{H}_A = \mathcal{T}_A \otimes \mathcal{R}_A$ and similarly Bob has only access to $\mathcal{H}_B = \mathcal{T}_B \otimes \mathcal{R}_B$. We might also generalize this setting by allowing Alice and Bob to have auxiliary systems (ancillas) residing in Hilbert spaces $\mathcal{A}_A$ and $\mathcal{A}_B$ such that $\mathcal{H} = \mathcal{A}_A \otimes \mathcal{T}_A \otimes \mathcal{R}_A \otimes \mathcal{R}_B \otimes \mathcal{T}_B \otimes \mathcal{A}_B$. It was proved in [6] that perfect embezzlement is not possible in this kind of framework, even in the generalized setting assuming infinite-dimensional Hilbert spaces $\mathcal{A}_A$ and $\mathcal{A}_B$.

Theorem 2.8.1. *Perfect embezzlement is impossible in the tensor product framework, even if Alice and Bob are allowed to use ancillas.*

Proof. This proof follows [6]. Consider the embezzlement scenario with auxiliary systems. In this case, Alice and Bob start with the state $|A\rangle_A |A\rangle_B |0\rangle_A |0\rangle_B |\mu\rangle_{AB} \in \mathcal{A}_A \otimes \mathcal{A}_B \otimes \mathcal{T}_A \otimes \mathcal{T}_B \otimes \mathcal{R}_A \otimes \mathcal{R}_B$ and hope to obtain the joint state $|A'\rangle_A |A'\rangle_B |\psi\rangle_{AB} |\mu\rangle_{AB} \in \mathcal{A}_A \otimes \mathcal{A}_B \otimes \mathcal{T}_A \otimes \mathcal{T}_B \otimes \mathcal{R}_A \otimes \mathcal{R}_B$. We can use the Schmidt Decomposition to characterize both states with a vector of positive real coefficients and certain orthonormal bases. Since Schmidt coefficients do not change under local unitary transformations, which is the case in embezzlement, both sorted vectors of coefficients must be the same to achieve perfect embezzlement. Suppose the biggest Schmidt coefficient of the initial state is α_0. Then, since $|\psi\rangle$ is entangled, the biggest Schmidt coefficient of the final state is $\alpha < \alpha_0$. Therefore, perfect embezzlement is impossible in the tensor product framework. □

2.8.2 Commuting Operator Framework

In the commuting operator framework we assume that there is one joint Hilbert space for the system of interest which we do not partition any more. Therefore, the notion of locality is not present and we assume that both Alice and Bob can operate on the whole Hilbert space $\mathcal{H}$. Nevertheless, we still want to capture the notion of independence between Alice's and Bob's actions to emphasize that they are distinct parties. To do so, we assume that their actions commute, i.e., Alice and Bob can perform them at any time with respect to the other party and the global result will be the same. We notice that it is also the case in the tensor product framework. It was proved in [6] that perfect embezzlement is possible in this framework.

Theorem 2.8.2. *Perfect embezzlement is possible in the commuting operators framework but it requires infinite-dimensional Hilbert spaces.*

Proof. See Sections 3 and 4 of [6]. □

Chapter 3

Noisy Embezzlement of Entanglement

In the previous chapter we discussed the embezzlement of entanglement in which we assumed that Alice and Bob both have the same classical description of the state that they intend to embezzle. We may generalize this assumption and say that Alice and Bob have descriptions which slightly differ from each other and they are both aware of a bound for this discrepancy. It can happen in practice if Alice and Bob obtain their descriptions through computing devices which rely on slightly different implementations of the floating point arithmetic to represent real numbers or different software platforms. Another possibility for this situation is when descriptions are transmitted to Alice and Bob through imperfect classical channels. This kind of a noisy scenario is outside of the scope of the result by van Dam and Hayden [17] and the canonical embezzlement procedure can actually fail as discussed by Dinur, Steurer and Vidick in Section 5 of [12]. In this chapter, we will discuss two approaches to perform the embezzlement with a discrepancy in the input. The first one is the quantum correlated sampling lemma [12] which is a quantum protocol and the second one is the classical synchronization scheme which can be followed by a standard embezzlement procedure.

3.1 Quantum Correlated Sampling Protocol

The quantum correlated sampling protocol is a quantum procedure, which can be seen as a robust version of embezzlement of entanglement. It was proposed by Dinur, Steurer

and Vidick in Section 5 of [12]. It uses the canonical embezzlement as a subroutine and, additionally, requires quantum entanglement and shared randomness.

Quantum correlated sampling protocol [12]

Input:
Alice: a biased classical description $|\psi_1\rangle$ of a bipartite quantum state $|\psi\rangle$,
Bob: a biased classical description $|\psi_2\rangle$ of a bipartite quantum state $|\psi\rangle$,
Alice and Bob: $\delta > 0, \eta = O(\delta^{1/5})$.
Shared resources: maximally entangled states $|\phi_d\rangle$, an embezzling state $|\mu\rangle$ of size large enough, shared randomness (uniformly distributed).
Promise: $|| \, |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| \, ||_1^2 \leq \delta$, $|| \, |\psi\rangle\langle\psi| - |\psi_1\rangle\langle\psi_1| \, ||_1^2 \leq \delta$.
Goal: Alice and Bob share an approximation $|\xi\rangle$ of $|\psi\rangle$ such that $|| \, |\psi\rangle\langle\psi| - |\xi\rangle\langle\xi| \, ||_1 \leq O(\delta^{1/10})$.
Protocol:

1. Alice and Bob use shared randomness to obtain a discretization $\tau_0, \ldots, \tau_K$ of the interval $[0, 1]$, where $K = \frac{\log(d/\delta)}{\log(1+\eta)}$. They set $\tau_0 = 1$, $\tau_{K+1} = 0$ and for $j = 1, \ldots, K$, they sample τ_j uniformly at random from the interval $[(1+\eta)^{-j}, (1+\eta)^{-j+1}]$.

2. Based on the discretization, they compute a classical description of a quantum state

$$|\xi_0\rangle \propto \sum_{j=0}^{K} \tau_j \, |jj\rangle_{AB} \, |\Phi_d\rangle_{AB} \, ,$$

   where $|\Phi_d\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle_{AB}$.

3. Alice and Bob use $|\mu\rangle$ for canonical embezzlement to create a quantum state $|\xi_0\rangle^{\otimes N}$, where $N = (2\delta d \sum_j \tau_j^2)^{-2}$. Part of the state can be created using pre-shared maximally entangled states.

4. Alice and Bob compute Schmidt decompositions of their states $|\psi_1\rangle = \sum_i \lambda_i \, |u_i\rangle \, |u_i'\rangle$ and $|\psi_2\rangle = \sum_i \mu_i \, |v_i\rangle \, |v_i'\rangle$. They build sets $S_j := \{i : \lambda_i \in [\tau_{j+1}, \tau_j)\}$ and $T_j := \{i : \lambda_i \in [\tau_{j+1}, \tau_j)\}$ respectively. For every j, they build projectors P_j on the span of $|u_i\rangle : i \in S_j$ and Q_j on the span of $|v_i\rangle : i \in T_j$ respectively.

5. Alice uses the two-outcome measurement $\{P_A, I - P_A\}$ where $P_A := \sum_j |j\rangle \langle j| \otimes P_j$ to measure her part of the first copy of $|\xi_0\rangle$. Bob does the same with $P_B := \sum_j |j\rangle \langle j| \otimes Q_j$. If either of them obtains the first outcome they proceed to Step 6 of the protocol. Otherwise, they move onto the next copy of $|\xi_0\rangle$. The protocol is aborted if either player has measured all N copies of the state without obtaining the first outcome.

6. Alice (resp. Bob) erases $|j\rangle$ in the first register of $|\xi_0\rangle$ (the copy for which they obtained a successful outcome) controlled on the second register. All qubits are discarded except the remaining register of $|\xi_0\rangle$. Bob transforms $|v_i\rangle$ to $|v_i'\rangle$ by applying a unitary.

Theorem 3.1.1. *Let d be an integer and $\delta > 0$. Given classical descriptions of quantum states $|\psi_1\rangle_{AB}, |\psi_2\rangle_{AB} \in \mathcal{H}^d \otimes \mathcal{H}^d$ such that $|| \, |\psi_1\rangle\langle\psi_1|_{AB} - |\psi_2\rangle\langle\psi_2|_{AB} \, ||_1^2 \leq \delta$, there exists an integer n and local unitary operations U_A, U_B such that*

$$||U_A \otimes U_B \, |\mu(n)\rangle\langle\mu(n)|_{AB} \otimes |00\rangle\langle 00|_{AB} \, U_A^\dagger \otimes U_B^\dagger - |\mu(n)\rangle\langle\mu(n)|_{AB} \otimes |\psi\rangle\langle\psi|_{AB} \, ||_1 = O(\delta^{1/10})$$

holds with probability at least $1 - O(\delta^{1/5})$, where $|\mu(n)\rangle$ is an embezzling state and U_A, U_B are defined by the quantum correlated sampling protocol.

We will now prove several auxiliary lemmas which together prove the Theorem stated above. All of the lemmas and proofs come from [12], however, most proofs presented here are more elaborate. We start by proving that input states and their discretized versions are close to each other in terms of the Schatten-1 distance.

Lemma 10. *Let $|\Psi\rangle := C \sum_j \tau_j \sum_{i \in S_j} |u_i\rangle \, |u_i'\rangle$ and $|\Phi\rangle := C' \sum_j \tau_j \sum_{i \in T_j} |v_i\rangle \, |v_i'\rangle$ be discretized classical descriptions of Alice's and Bob's input. It holds that*

$$(1+\eta)^{-2} \leq C, C' \leq 1,$$

and

$$\max\left\{ || \, |\psi\rangle\langle\psi| - |\Psi\rangle\langle\Psi| \, ||_1^2, || \, |\phi\rangle\langle\phi| - |\Phi\rangle\langle\Phi| \, ||_1^2 \right\} = O(\eta).$$

Proof. The normalization constant C can be evaluated as $C^{-2} = \sum_k \tau_k^2 s_k$. Using the discretization we can bound it in terms of λ_i (coefficients which give rise to a specific sum $\sum_k \tau_k^2 s_k$) and then in terms of η. We notice that $\tau_{k+1} \leq \lambda_i$ which implies $\tau_k (1+\eta)^{-2} \leq \tau_{k+1} \leq \lambda_i$. Then, we have $\tau_k \leq (1+\eta)^2 \lambda_i$ and thus

$$\frac{1}{C^2} \leq \sum_i \lambda_i^2 (1+\eta)^4 = (1+\eta)^4 \sum_i \lambda_i^2 = (1+\eta)^4,$$

$$C \geq (1+\eta)^{-2}.$$

Analogously,

$$C' \geq (1+\eta)^{-2}.$$

On the other hand, we have that $\lambda_i \leq \tau_k$ and, due to the normalization condition, we have

$$1 = \sum_i \lambda_i^2 \leq \sum_k \tau_k^2 s_k = C^{-2}.$$

Therefore, $C \leq 1$, and similarly $C' \leq 1$. Next, we can bound the distance between original and discrete versions of input states.

$$|| \, |\psi\rangle - |\Psi\rangle \, ||_2^2 = || \sum_i \lambda_i \, |u_i\rangle \, |u_i'\rangle - C \sum_k \tau_k \sum_{i \in S_k} |u_i\rangle \, |u_i'\rangle \, ||_2^2 =$$

$$= || \sum_k \sum_{i \in S_k} \lambda_i \, |u_i\rangle \, |u_i'\rangle - C \sum_k \tau_k \sum_{i \in S_k} |u_i\rangle \, |u_i'\rangle \, ||_2^2 = || \sum_k \sum_{i \in S_k} (\lambda_i - C\tau_k) \, |u_i\rangle \, |u_i'\rangle \, ||_2^2 =$$

$$= \sum_k \sum_{i \in S_k} (\lambda_i - C\tau_k)^2 = \sum_k \sum_{i \in S_k} (\lambda_i^2 - 2C\lambda_i\tau_k + C^2\tau_k^2) =$$

$$= \sum_k \sum_{i \in S_k} \lambda_i^2 - \sum_k \sum_{i \in S_k} 2C\lambda_i\tau_k + \sum_k \sum_{i \in S_k} C^2\tau_k^2 = 2 - \sum_k \sum_{i \in S_k} 2C\lambda_i\tau_k \leq$$

$$\leq 2 - 2C \sum_k \sum_{i \in S_k} \tau_{k+1}\tau_k \leq 2 - 2C \sum_k \sum_{i \in S_k} (1+\eta)^{-k-1}\tau_k = 2 - 2C(1+\eta)^{-2} \sum_k \sum_{i \in S_k} (1+\eta)^{-k+1}\tau_k \leq$$

$$\leq 2 - 2C(1+\eta)^{-2} \sum_k \sum_{i \in S_k} \tau_k^2 = 2 - 2(1+\eta)^{-2} \frac{1}{C} C^2 \sum_k \sum_{i \in S_k} \tau_k^2 = 2 - 2(1+\eta)^{-2} \frac{1}{C} \leq$$

$$\leq 2 - 2(1-\eta)^{-2} = 2\left(1 - \frac{1}{(1+\eta)^2}\right) = O(\eta).$$

Therefore, by Lemma 3, we proved that

$$|| \, |\psi\rangle\langle\psi| - |\Psi\rangle\langle\Psi| \, ||_1^2 = O(\eta),$$

and similarly

$$|| \, |\phi\rangle\langle\phi| - |\Phi\rangle\langle\Phi| \, ||_1^2 = O(\eta).$$

□

We state the Markov's inequality which will be useful in proving the next lemma.

Definition 25 (Markov's inequality)**.** *Let X be a non-negative random variable and $a > 0$ be a real number. Then, the following holds*

$$\Pr(X \geq a) \leq \frac{E(X)}{a}.$$

Next, we will need a lemma which probabilistically bounds the expression which is proportional to the probability that Alice and Bob simultaneously obtain the desired outcome of the measurement in the Step 5 of the protocol.

Lemma 11. *The following statement holds with probability at least* $1 - O(\delta^{1/3}\eta^{-2/3})$ *over the choice of* τ_j*:*

$$\sum_j \tau_j^2 \operatorname{Tr}(P_j Q_j) = 1 - O(\eta) - O(\delta^{1/3}\eta^{-2/3}).$$

Proof. By Lemma 10 we know that $|| \, |\psi\rangle\langle\psi| - |\Psi\rangle\langle\Psi| \, ||_1 = O(\sqrt{\eta})$ and $|| \, |\phi\rangle\langle\phi| - |\Phi\rangle\langle\Phi| \, ||_1 = O(\sqrt{\eta})$. By the promise we have $|| \, |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \, ||_1 = O(\sqrt{\delta})$. Therefore, by using the triangle inequality we obtain

$$O(\sqrt{\eta}) + O(\sqrt{\delta}) \geq || \, |\psi\rangle\langle\psi| - |\Psi\rangle\langle\Psi| \, ||_1 + || \, |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \, ||_1 + || \, |\phi\rangle\langle\phi| - |\Phi\rangle\langle\Phi| \, ||_1 \geq$$

$$\geq || \, |\Phi\rangle\langle\Phi| - |\Psi\rangle\langle\Psi| \, ||_1.$$

Since $\eta = \delta^{1/5}$, we conclude that

$$|| \, |\Phi\rangle\langle\Phi| - |\Psi\rangle\langle\Psi| \, ||_1 \leq O(\sqrt{\eta}).$$

We use the relationship between fidelity and Schatten-1 distance for quantum states to get

$$F(|\Phi\rangle, |\Psi\rangle) = \sqrt{1 - \frac{1}{4}|| \, |\Phi\rangle\langle\Phi| - |\Psi\rangle\langle\Psi| \, ||_1^2} \geq \sqrt{1 - O(\eta)} = 1 - O(\eta).$$

We also calculate that

$$F(|\Phi\rangle, |\Psi\rangle) = CC'(\sum_k \tau_k \sum_{j \in T_k} \langle v_j| \langle v_j'|)(\sum_{k'} \tau_{k'} \sum_{i \in S_{k'}} |u_i\rangle |u_i'\rangle) = CC' \sum_{kk'} \tau_k \tau_{k'} \sum_{ij} |\langle v_j|u_i\rangle|^2 =$$

$$= CC' \sum_{kk'} \tau_k \tau_{k'} \operatorname{Tr}(P_k Q_{k'}) = CC' \left(\sum_{k \neq k'} \tau_k \tau_{k'} \operatorname{Tr}(P_k Q_{k'}) + \sum_k \tau_k^2 \operatorname{Tr}(P_k Q_{k'}) \right).$$

Thus, we obtained

$$\sum_{k \neq k'} \tau_k \tau_{k'} \operatorname{Tr}(P_k Q_{k'}) + \sum_k \tau_k^2 \operatorname{Tr}(P_k Q_{k'}) \geq$$

$$\geq CC' \left(\sum_{k \neq k'} \tau_k \tau_{k'} \operatorname{Tr}(P_k Q_{k'}) + \sum_k \tau_k^2 \operatorname{Tr}(P_k Q_{k'}) \right) \geq 1 - O(\eta),$$

where we also used upper bounds on normalization constants C, C' from Lemma 10.

To prove the current Lemma, we are interested in the term $\sum_k \tau_k^2 \operatorname{Tr}(P_k Q_{k'})$ from the equation above. Therefore, we will now bound the term $\sum_{k \neq k'} \tau_k \tau_{k'} \operatorname{Tr}(P_k Q_{k'})$. To do

so, we split it into two sums based on the inequality dependent on λ_i and μ_i which is $|\lambda_i - \mu_j|^2 \geq \theta\mu_j\lambda_i$ for some threshold parameter θ.

$$\sum_{k\neq k'} \tau_k\tau_{k'} \operatorname{Tr}(P_kQ_{k'}) = \sum_{k\neq k'} \tau_k\tau_{k'} \sum_{i\in S_k, j\in T_{k'}} |\langle v_j|u_i\rangle|^2 \leq (1+\eta)^2 \sum_{k\neq k'} \sum_{i\in S_k, j\in T_{k'}} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 =$$

$$= (1+\eta)^2 \left(\sum_{\substack{i\in S_k, j\in T_{k'}, k\neq k' \\ |\lambda_i-\mu_j|^2\geq\theta\mu_j\lambda_i}} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 + \sum_{\substack{i\in S_k, j\in T_{k'}, k\neq k' \\ |\lambda_i-\mu_j|^2<\theta\mu_j\lambda_i}} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 \right).$$

where we used the fact proved in Lemma 3.1 that $\tau_k \leq (1+\eta)\lambda_i$ and $\tau'_k \leq (1+\eta)\mu_i$. Each sum above will be now bounded separately. We start with the first one and bound it as follows

$$\sum_{\substack{i\in S_k, j\in T_{k'}, k\neq k' \\ |\lambda_i-\mu_j|^2\geq\theta\mu_j\lambda_i}} \lambda_i\mu_i |\langle v_j|u_i\rangle|^2 \leq \sum_{\substack{i,j \\ |\lambda_i-\mu_j|^2\geq\theta\mu_j\lambda_i}} \lambda_i\mu_i |\langle v_j|u_i\rangle|^2 \leq \sum_{\substack{i,j \\ |\lambda_i-\mu_j|^2\geq\theta\mu_j\lambda_i}} \frac{|\lambda_i-\mu_j|^2}{\theta} |\langle v_j|u_i\rangle|^2 \leq$$

$$\theta^{-1} \sum_{i,j} |\lambda_i - \mu_j|^2 |\langle v_j|u_i\rangle|^2 \leq \theta^{-1} \sum_{i,j} |\lambda_i - \mu_j|^2 \leq \theta^{-1} ||\psi - \phi||_2^2 \leq \theta^{-1}\delta.$$

Now, we proceed with bounding the second sum. We notice that if θ is at most η multiplied by a small constant, then $k' = k-1$ or $k' = k+1$. First, suppose $k' = k-1$, then

$$|\lambda_i - \mu_j|^2 \leq \theta\mu_j\lambda_i \leq \theta\tau_{k-1}\tau_k \leq \theta(1+\eta)^{-k+2}\tau_k = (1+\eta)^2\theta(1+\eta)^{-k} \leq (1+\eta)^2\theta\tau_k^2.$$

Suppose $k' = k+1$, then

$$|\lambda_i - \mu_j|^2 \leq \theta\mu_j\lambda_i \leq \theta\tau_{k'}\tau_k = \theta\tau_{k+1}\tau_k \leq \theta(1+\eta)^{-k}\tau_k \leq \theta\tau_k^2 \leq (1+\eta)^2\theta\tau_k^2.$$

Therefore, we can bound both cases by $(1+\eta)^2\theta\tau_k^2$. Next, we calculate that τ_k is uniformly sampled from the interval of length

$$L = (1+\eta)^{-k+1} - (1+\eta)^{-k} = (1+\eta)^{-k}(1+\eta-1) = \eta(1+\eta)^{-k} =$$

$$= \eta(1+\eta)^{-k+1}(1+\eta)^{-1} \leq \tau_k\eta(1+\eta)^{-1}.$$

Since the bound $|\lambda_i - \mu_j|^2 \leq (1+\eta)^2\theta\tau_k^2$ is implied directly by the constraint from the summation that $|\lambda_i - \mu_j|^2 < \theta\mu_j\lambda_i$, the probability of τ_k satisfying it, together with satisfying $\lambda_i \leq \tau_k \leq \mu_j$ is

$$\frac{|\lambda_i - \mu_j|}{L} \leq \frac{\sqrt{\theta}\tau_k(1+\eta)^2}{\tau_k\eta} = O\left(\sqrt{\theta}\eta^{-1}\right).$$

It implies that on expectation over the choice of τ_k, we have the bound

$$\mathbb{E}_{\tau_k} \sum_{\substack{i\in S_k, j\in T_{k'}, k\neq k' \\ |\lambda_i-\mu_j|^2<\theta\mu_j\lambda_i}} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 \leq O\left(\sqrt{\theta}\eta^{-1}\right) \sum_{i,j} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 = O\left(\sqrt{\theta}\eta^{-1}\right).$$

By applying Markov's inequality, we can also see that it is very unlikely that the actual sum (without taking the expectation) is significantly larger (we introduce a big constant $\frac{1}{\eta}$) than its expected value.

$$\Pr\left(\sum_{\substack{i\in S_k, j\in T_{k'}, k\neq k' \\ |\lambda_i-\mu_j|^2<\theta\mu_j\lambda_i}} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 \geq \frac{O\left(\sqrt{\theta}\eta^{-1}\right)}{\eta} \right) \leq \eta.$$

Therefore, it implies that

$$\Pr\left(\sum_{\substack{i\in S_k, j\in T_{k'}, k\neq k' \\ |\lambda_i-\mu_j|^2<\theta\mu_j\lambda_i}} \lambda_i\mu_j |\langle v_j|u_i\rangle|^2 \leq \frac{O\left(\sqrt{\theta}\eta^{-1}\right)}{\eta} \right) \geq 1-\eta.$$

Overall, we ended up with the following

$$1 - O(\eta) = F(|\,|\Phi\rangle\,, |\Psi\rangle) \leq (1+\eta)^4[\delta\theta^{-1} + O(\sqrt{\theta}\eta^{-1})] + \sum_k \tau_k^2 \operatorname{Tr}(P_k Q_k),$$

$$\sum_k \tau_k^2 \operatorname{Tr}(P_k Q_k) \geq 1 - O(\eta) - (1+\eta)^4[\delta\theta^{-1} + O(\sqrt{\theta}\eta^{-1})],$$

with probability at least $1 - O\left(\sqrt{\theta}\eta^{-1}\right)$. We choose $\theta = (\delta\eta)^{2/3}$ and obtain

$$\sum_k \tau_k^2 \operatorname{Tr}(P_k Q_k) \geq 1 - O(\eta) - (1+\eta)^4[\delta(\delta\eta)^{-2/3} + O((\delta\eta)^{1/3}\eta^{-1})] = 1 - O(\eta) - O(\delta^{1/3}\eta^{-2/3}),$$

with probability at least $1 - O(\eta)$.

We recall that at the beginning of this proof we assumed that $\theta \ll \eta$. With the particular choice $\theta = (\delta\eta)^{2/3}$ that we made, it is equivalent to $\delta \ll \eta^{1/2}$. Later in the proof we will actually choose $\eta = \delta^{1/5}$, which is $\delta = \eta^5$ and therefore it will satisfy the condition. □

Finally, we will show that after performing the protocol, Alice and Bob share a state which is close to one of the input states in trace distance with high probability.

Lemma 12. *Given $|\psi\rangle$ and $|\phi\rangle$ such that $|| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \, ||_1^2 \leq \delta$ and $\eta = \delta^{1/5}$, the quantum correlated sampling protocol terminates with Alice and Bob sharing a state $|\xi\rangle$ such that $|| \, |\xi\rangle\langle\xi| - |\psi\rangle\langle\psi| \, ||_1 = O(\delta^{1/10})$ with probability at least $1 - O(\delta^{1/5})$.*

Proof. Suppose that in the Step 5 of the quantum correlated sampling protocol Alice and Bob both obtain outcomes which let them synchronously proceed to the next step of the protocol. Then, they share the following state after that measurement

$$P_A \otimes P_B \, |\xi_0\rangle = C'' \left(\sum_k |k\rangle \langle k|_A \otimes P_k \right) \otimes \left(\sum_j |j\rangle \langle j|_B \otimes Q_j \right) \left(\sum_{mi} \tau_m \, |m,m\rangle_{AB} \, |i,i\rangle_{AB} \right) =$$

$$= C'' \sum_k \tau_k \, |k,k\rangle_{AB} \sum_i P_k \, |i\rangle_A \, Q_k \, |i\rangle_B =$$

$$= C'' \sum_k \tau_k \, |k,k\rangle_{AB} \sum_i \left(\sum_{j \in S_k} |u_j\rangle \langle u_j|i\rangle_A \right) \left(\sum_{m \in T_k} |v_m\rangle \langle v_m|i\rangle_B \right) =$$

$$= C'' \sum_k \tau_k \, |k,k\rangle_{AB} \sum_{j \in S_k, m \in T_k} \left(\sum_i \langle u_j|i\rangle \langle v_m|i\rangle \right) |u_j\rangle \, |v_m\rangle =$$

$$= C'' \sum_k \tau_k \, |k,k\rangle_{AB} \sum_{j \in S_k, m \in T_k} \left(\sum_i \langle u_j|i\rangle \langle i|v_m\rangle \right) |u_j\rangle \, |v_m\rangle =$$

$$= C'' \sum_k \tau_k \, |k,k\rangle_{AB} \sum_{j \in S_k, m \in T_k} \langle u_j|v_m\rangle \, |u_j\rangle \, |v_m\rangle = C'' \sum_k \tau_k \, |k,k\rangle_{AB} \sum_{i \in S_k, j \in T_k} \langle u_i|v_j\rangle \, |u_i\rangle \, |v_j\rangle \, .$$

According to the protocol, Alice and Bob discard the $|k,k\rangle_{AB}$ register, Bob applies a unitary to his register and they end up with the state

$$|\xi\rangle = C'' \sum_k \tau_k \sum_{i \in S_k, j \in T_k} \langle u_i|v_j\rangle \, |u_i\rangle \, |v'_j\rangle \, .$$

The normalization constant C'' evaluates as

$$(C'')^{-2} = \sum_k \tau_k^2 \sum_{i \in S_k, j \in T_k} |\, \langle u_i|v_j\rangle \,|^2 = \sum_k \tau_k^2 \, \mathrm{Tr}(P_k Q_k) = 1 - O(\eta) - O(\delta^{1/3} \eta^{-2/3}),$$

where the last step follows from the Lemma 11. We equate the second and the third terms above to obtain $\eta = \delta^{1/5}$. Therefore, we proved

$$(C'')^{-2} = \sum_k \tau_k^2 \operatorname{Tr}(P_k Q_k) = 1 - O(\delta^{1/5}).$$

with probability at least $1 - O(\delta^{1/5})$.
Now, we calculate the fidelity between the obtained state $|\xi\rangle$ and the discretized description $|\Phi\rangle$ of one of the input states

$$F(|\xi\rangle, |\Phi\rangle) = C''C \left(\sum_k \tau_k \sum_{i \in S_k, j \in T_k} \langle u_i | v_j \rangle \langle u_i | \langle v'_j |) (\sum_p \tau_p \sum_{m \in S_p} |v_m\rangle |v'_m\rangle \right) =$$

$$= C''C \sum_k \tau_k^2 \sum_{i \in S_k, j \in T_k} |\langle u_i | v_j \rangle|^2 = \frac{C''C}{(C'')^2} = \frac{C}{C''} \geq \frac{1 - O(\delta^{1/5})}{(1+\eta)^2} = \frac{1 - O(\delta^{1/5})}{(1+\delta^{1/5})^2} = 1 - O(\delta^{1/5}).$$

We use the inequality between fidelity and Schatten-1 distance to obtain the following

$$2\sqrt{O(\delta^{1/5}) - O(\delta^{2/5})} \geq || \, |\xi\rangle\langle\xi| - |\Phi\rangle\langle\Phi| \, ||_1,$$

and therefore

$$|| \, |\xi\rangle\langle\xi| - |\Phi\rangle\langle\Phi| \, ||_1^2 \leq O(\delta^{1/5}).$$

By the triangle inequality we have

$$|| \, |\psi\rangle\langle\psi| - |\xi\rangle\langle\xi| \, ||_1 \leq || \, |\psi\rangle\langle\psi| - |\Phi\rangle\langle\Phi| \, ||_1 + || \, |\Phi\rangle\langle\Phi| - |\xi\rangle\langle\xi| \, ||_1 \leq O(\sqrt{\eta}) + O(\delta^{1/10}) = O(\delta^{1/10}).$$

We notice that there is another source of error in the protocol, namely the embezzlement error which is created in the Step 2 on the protocol. However, this error can be easily suppressed by choosing the size of an embezzling state large enough and we will assume that it is the case in the protocol. The authors of the quantum correlated sampling protocol suggest that an embezzling state of size $O((d/\delta)^2)$ qubits should be sufficient, even if we decide to embezzle EPR pairs that are part of $|\xi_0\rangle$.

We will now prove the claimed probability of success of the protocol. It is the probability that in Step 5 of the quantum correlated sampling protocol Alice and Bob synchronously proceed to Step 6. This event, denoted *sync*, happens if Alice and Bob obtain outcomes P_A and P_B when measuring the same copy of $|\xi_0\rangle$ and Pr(sync) is given by

$$\Pr(\text{sync}) = \sum_{i=0}^{N-1} \Pr(\text{outcome } (I - P_A) \text{ and } (I - P_B))^i \Pr(\text{outcome } P_A \text{ and } P_B) =$$

$$= \Pr(\text{outcome } P_A \text{ and } P_B) \sum_{i=0}^{N-1} \Pr(\text{outcome } (I-P_A) \text{ and } (I-P_B))^i =$$

$$= \Pr(\text{outcome } P_A \text{ and } P_B) \frac{1 - \Pr(\text{outcome } (I-P_A) \text{ and } (I-P_B))^N}{1 - \Pr(\text{outcome } (I-P_A) \text{ and } (I-P_B))}.$$

We can choose N, the number of states $|\xi_0\rangle$ embezzled by Alice and Bob, large enough such that with probability at least $1-\delta^2$ both of them obtain a successful outcome before the number N of copies of $|\zeta_0\rangle$ runs out. Therefore,

$$\Pr(\text{outcome } (I-P_A) \text{ and } (I-P_B))^N \leq \delta^2.$$

We also have that

$$\Pr(\text{outcome } (I-P_A) \text{ and } (I-P_B)) = C_d \langle\zeta_0| (I-P_A) \otimes (I-P_B) |\zeta_0\rangle =$$

$$= 1 - C_d \langle\zeta_0| P_A \otimes I) |\zeta_0\rangle - C_d \langle\zeta_0| I \otimes P_B) |\zeta_0\rangle + C_d \langle\zeta_0| P_A \otimes P_B |\zeta_0\rangle =$$

$$= 1 - 2C_d \langle\zeta_0| P_A \otimes I) |\zeta_0\rangle + C_d \langle\zeta_0| P_A \otimes P_B |\zeta_0\rangle =$$

$$= 1 - 2\Pr(\text{outcome } P_A) + \Pr(\text{outcome } P_A \text{ and } P_B),$$

where $C_d = (d\sum_k \tau_k^2)^{-1}$. Therefore,

$$\Pr(sync) \geq \Pr(\text{outcome } P_A \text{ and } P_B) \frac{1-\delta^2}{2\Pr(\text{outcome } P_A) - \Pr(\text{outcome } P_A \text{ and } P_B)} \geq$$

$$\geq \frac{C_d(1-O(\delta^{1/5}))(1-\delta^2)}{2C_d\sum_k \tau_k^2 s_k - C_d(1-O(\delta^{1/5}))} = \frac{(1-O(\delta^{1/5}))(1-\delta^2)}{2\frac{1}{C^2} - (1-O(\delta^{1/5}))} \geq \frac{(1-O(\delta^{1/5}))(1-\delta^2)}{2-(1-O(\delta^{1/5}))} =$$

$$= \frac{(1-O(\delta^{1/5}))(1-\delta^2)}{1+O(\delta^{1/5})} = \frac{1-O(\delta^{1/5})}{1+O(\delta^{1/5})} = 1 - O(\delta^{1/5}).$$

Therefore, we proved that with probability at least $1-O(\delta^{1/5})$ Alice and Bob, at the end of the protocol, have the state $|\xi\rangle$ such that $|| \, |\psi\rangle\langle\psi| - |\xi\rangle\langle\xi| \, ||_1 \leq O(\delta^{1/10})$. □

3.2 Classical Synchronization Scheme and Canonical Embezzlement

In this section we will introduce the classical synchronization scheme for descriptions of quantum states. Most of the contribution to this protocol come from Debbie Leung and

Richard Cleve. The protocol relies on randomness shared between Alice and Bob that lets them compute common discretizations of the real line. They express every coefficient of their descriptions in terms of a point from a discretization. It allows them to share the same classical description of a quantum state with high probability at the end of the protocol, without using any form of communication. Moreover, the description is close to their initial inputs. Based on the synchronized description, they can then perform the canonical embezzlement procedure to establish a shared bipartite entangled state between them. Therefore, the classical synchronization scheme together with the canonical embezzlement achieves the same result as the quantum correlated sampling lemma, with similar assumptions.

Classical synchronization scheme

Input:
Alice: a classical description of a quantum state $|\psi_1\rangle = \sum_{j=0}^{d^2-1}(a_{2j+1} + ia_{2j+2})\,|j\rangle_{AB}$, $\delta > 0$
Bob: a classical description of a quantum state $|\psi_2\rangle = \sum_{j=0}^{d^2-1}(b_{2j+1} + ia_{bj+2})\,|j\rangle_{AB}$, $\delta > 0$
Shared resources: shared randomness (uniformly distributed)
Promise: $||\,|\psi_1\rangle - |\psi_2\rangle\,||_2 \leq \delta$.
Goal: With high probability, Alice and Bob share the classical description of a quantum state $|\psi_1'\rangle$ such that $||\,|\psi_1\rangle - |\psi_1'\rangle\,||_2 \leq O(d\sqrt{\delta})$.
Protocol:

1. Using shared randomness, Alice and Bob prepare $\mathbf{s} \in [0,\mu]^{2d^2}$ and $\forall i = 1 \ldots 2d^2 \quad \mathbf{v}^i : \mathbf{v}^i_j = s_i + k_j\mu : k_j \in \left[\lfloor \frac{\min(a_i,b_i)}{\mu} \rfloor - 1, \lceil \frac{\max(a_i,b_i)}{\mu} \rceil + 1\right] \cap \mathbb{Z}$, where $\mu = O(\sqrt{\delta})$.

2. Alice (Bob) $\forall i = 1 \ldots 2d^2$ chooses $\alpha_i \in \mathbf{v}^i : \alpha_i = \arg\min_{\alpha \in \mathbf{v}^i} |a_i - \alpha|$ ($\beta_i \in \mathbf{v}^i : \beta_i = \arg\min_{\beta \in \mathbf{v}^i} |b_i - \beta|$).

3. Alice and Bob obtain vectors $|v_1\rangle = \sum_{j=0}^{d^2-1}(\alpha_{2j+1} + i\alpha_{2j+2})\,|j\rangle_{AB}$ and $|v_2\rangle = \sum_{j=0}^{d^2-1}(\beta_{2j+1} + i\beta_{2j+2})\,|j\rangle_{AB}$ respectively. Alice updates her classical description to a unit vector $|\psi_1'\rangle \propto \sum_{j=0}^{d^2-1}(\alpha_{2j+1} + i\alpha_{2j+2})\,|j\rangle_{AB}$ and Bob updates his classical description to a unit vector $|\psi_2'\rangle \propto \sum_{j=0}^{d^2-1}(\beta_{2j+1} + i\beta_{2j+2})\,|j\rangle_{AB}$.

Theorem 3.2.1. *In the classical synchronization scheme with unlimited shared randomness,* $\Pr(|\psi_1'\rangle = |\psi_2'\rangle) \geq 1 - O(d\sqrt{\delta})$ *and* $||\,|\psi_1'\rangle\langle\psi_1'| - |\psi_1\rangle\langle\psi_1|\,||_1 \leq O(d\sqrt{\delta})$.

Proof. Let us consider any pair of coefficients a_i and b_i from the protocol. They are updated by Alice and Bob to α_i and β_i correspondingly. Without loss of generality, we assume that

$a_i \leq b_i$. We calculate the probability that $s_i + k_x\mu = \alpha_i \neq \beta_i = s_i + k_{x+1}\mu$ (a_i, b_i are mapped to neighbouring discretization points because $\mu = O(\sqrt{\delta}) > \delta$). Since the only component which is not fixed is s_i (it arises randomly), we wonder for which values of s_i coefficients a_i and b_i are closest to different values on a discretized real line. It happens for $s_i + k_x\mu + \frac{\mu}{2} \in [a_i, b_i]$. Therefore, recalling that $s_i \in [0, \mu]$, $\Pr(a_i \neq b_i) = \frac{|a_i - b_i|}{\mu}$. By the union bound,

$$\Pr(\text{at least 1 pair of coefficients differs}) \leq \sum_i \frac{|a_i - b_i|}{\mu} \leq \frac{\sqrt{2d^2}}{\mu}\sqrt{\sum_i (a_i - b_i)^2} \leq \frac{\sqrt{2d^2}\delta}{\mu}.$$

Therefore, $\Pr\left(|\psi_1'\rangle = |\psi_2'\rangle\right) \geq 1 - \frac{\sqrt{2}d\delta}{\mu}$. Next, we want to bound $|| \, |\psi_1'\rangle - |\psi_1\rangle \, ||_2$. We notice that

$$|| \, |\psi_1\rangle - |v_1\rangle \, ||_2 = \sqrt{\sum_i |\alpha_i - a_i|^2} \leq \sqrt{\sum_i \left(\frac{\mu}{2}\right)^2} = \sqrt{2d^2\frac{\mu^2}{4}} = \frac{\sqrt{2}d\mu}{2}.$$

By the triangle inequality, we have $|| \, |\psi_1\rangle \, || + || \, |\psi_1\rangle - |v_1\rangle \, ||_2 \geq || \, |v_1\rangle \, ||$, and thus $||v_1|| \leq 1 + \frac{\sqrt{2}d\mu}{2}$. It implies that

$$|| \, |v_1\rangle - |\psi_1'\rangle \, ||_2 = \left|\left||v_1\rangle - \frac{|v_1\rangle}{|| \, |v_1\rangle \, ||_2}\right|\right|_2 = \left(1 - \frac{1}{|| \, |v_1\rangle \, ||_2}\right) || \, |v_1\rangle \, ||_2 = || \, |v_1\rangle \, ||_2 - 1 \leq \frac{\sqrt{2}d\mu}{2}.$$

By the triangle inequality, we have $|| \, |\psi_1'\rangle - |\psi_1\rangle \, ||_2 \leq || \, |\psi_1\rangle - |v_1\rangle \, ||_2 + || \, |v_1\rangle - |\psi_1'\rangle \, ||_2 \leq \frac{\sqrt{2}d\mu}{2} + \frac{\sqrt{2}d\mu}{2} = \sqrt{2}d\mu$. We equate probability and distance errors $\frac{\sqrt{2}d\delta}{\mu} = \sqrt{2}d\mu$ to obtain $\mu = \sqrt{\delta}$. It implies that $|| \, |\psi_1'\rangle - |\psi_1\rangle \, ||_2 \leq O(d\sqrt{\delta})$ and $\Pr\left(|\psi_1'\rangle = |\psi_2'\rangle\right) \geq 1 - O(d\sqrt{\delta})$. By Lemma 3, we obtain $|| \, |\psi_1'\rangle\langle\psi_1'| - |\psi_1\rangle\langle\psi_1| \, ||_1 \leq O(d\sqrt{\delta})$ □

Theorem 3.2.2. *In the classical synchronization scheme with* $d^2 \log_2\left(\frac{4d^2}{\delta}\right)$ *bits of shared randomness,* $\Pr\left(|\psi_1'\rangle = |\psi_2'\rangle\right) \geq 1 - O(d\sqrt{\delta})$ *and* $|| \, |\psi_1'\rangle\langle\psi_1'| - |\psi_1\rangle\langle\psi_1| \, ||_1 \leq O(d\sqrt{\delta})$.

Proof. Let $s_i = x\Delta s$, $x = 0, 1, \ldots, \frac{\mu}{\Delta s}$ be the discretization of s_i from the classical synchronization scheme. The discretization does not affect the bound for the distance from Theorem 1 but affects the bound for the probability of error which we will now update. If $\Delta s > \delta$, then coefficients a_i and b_i are closest to different values on a discretized real line for at most one value of x, i.e., $\Pr(a_i \neq b_i) \leq \frac{\Delta s}{\mu}$. Else, if $\Delta s \leq \delta$, we can use the bound from Theorem 1. Let $\mathbb{S} = \{i : |a_i - b_i| < \Delta s\}$, then, by the union bound,

$$\Pr\left(|\psi_1'\rangle \neq |\psi_2'\rangle\right) \leq \sum_{i \in \mathbb{S}} \frac{\Delta s}{\mu} + \sum_{i \notin \mathbb{S}} \frac{|a_i - b_i|}{\mu} \leq 2d^2\frac{\Delta s}{\mu} + \sqrt{2d^2}\frac{\delta}{\mu}.$$

We choose $2d^2\frac{\Delta s}{\mu} = \sqrt{2d^2}\frac{\delta}{\mu}$, which implies $\Delta s = \frac{\delta}{\sqrt{2d^2}}$. Then, $\Pr\left(|\psi_1'\rangle \neq |\psi_2'\rangle\right) \leq 2\sqrt{2d^2}\frac{\delta}{\mu}$. We recall $|| \, |\psi_1'\rangle - |\psi_1\rangle \, ||_2 \leq \sqrt{2}d\mu$ and equate probability and distance errors $2\sqrt{2d^2}\frac{\delta}{\mu} = \sqrt{2d^2}\mu$ to obtain $\mu = \sqrt{2\delta}$. It follows that $|| \, |\psi_1'\rangle - |\psi_1\rangle \, ||_2 \leq 2d\sqrt{\delta}$ and $\Pr\left(|\psi_1'\rangle = |\psi_2'\rangle\right) \geq 1 - O(d\sqrt{\delta})$. By Lemma 3, it implies that $|| \, |\psi_1'\rangle\langle\psi_1'| - |\psi_1\rangle\langle\psi_1| \, ||_1 \leq O(d\sqrt{\delta})$. For $2d^2$ of s_i coefficients, we need the following number of bits of shared randomness

$$2d^2 \log_2(\frac{\mu}{\Delta s}) = 2d^2 \log_2 \frac{\sqrt{2\delta}}{\frac{\delta}{\sqrt{2d^2}}} = 2d^2 \log_2 \frac{2\sqrt{d^2}}{\sqrt{\delta}} = d^2 \log_2 \left(\frac{4d^2}{\delta}\right).$$

□

Having proved the classical synchronization scheme, we can now formally show how to use it together with the canonical embezzlement to perform a noisy embezzlement.

Classical synchronization with canonical embezzlement protocol

Input:
Alice: a classical description of a quantum state $|\psi_1\rangle = \sum_{j=0}^{d^2-1} (a_{2j+1} + ia_{2j+2}) \, |j\rangle_{AB}$, $\delta, \epsilon > 0$
Bob: a classical description of a quantum state $|\psi_2\rangle = \sum_{j=0}^{d^2-1} (b_{2j+1} + ib_{2j+2}) \, |j\rangle_{AB}$, $\delta, \epsilon > 0$
Shared resources: shared randomness (uniformly distributed), an embezzling state $|\mu(n)\rangle$
Promise: $|| \, |\psi_1\rangle - |\psi_2\rangle \, ||_2 \leq \delta$, a quantum state described in a fixed basis $\{|j\rangle_{AB}\}$.
Goal: With high probability, Alice and Bob share a quantum state $|\psi_1'\rangle$ such that $|| \, |\psi_1\rangle\langle\psi_1|_{AB} - |\psi'\rangle\langle\psi'|_{AB} \, ||_1 \leq O(\sqrt{\epsilon}) + O(d\sqrt{\delta})$.
Protocol:

1. Alice and Bob use the classical synchronization scheme to obtain descriptions of quantum states $|\psi_{CSS1}\rangle$ and $|\psi_{CSS2}\rangle$ respectively, which are the same with high probability.
2. Alice and Bob calculate the Schmidt decomposition of their synchronized state.
3. Alice and Bob use the canonical embezzlement based on their classical descriptions from Step 2 to obtain a quantum state $|\psi'\rangle_{AB}$ shared between them.

Theorem 3.2.3. *In the classical synchronization with canonical embezzlement protocol,* $|| \, |\psi_1\rangle\langle\psi_1|_{AB} - |\psi'\rangle\langle\psi'|_{AB} \, ||_1 \leq O(\sqrt{\epsilon}) + O(d\sqrt{\delta})$ *with probability at least* $1 - O(d\sqrt{\delta})$ *and* $n > d^{2/\epsilon}$.

Proof. The probability of success follows directly from Theorem 3.2.2. The same theorem tells us that $|| \, |\psi\rangle\langle\psi|_{AB} - |\psi_{CSS1}\rangle\langle\psi_{CSS1}|_{AB} \, ||_1 \leq O(d\sqrt{\delta})$. From Theorem 2.1.1 about

canonical embezzlement, we know that $F(|\psi_{CSS1}\rangle_{AB}, |\psi'\rangle_{AB}) \geq 1-\epsilon$. We can use the relationship between fidelity and the Schatten-1 distance for pure states to obtain

$$2\sqrt{2\epsilon-\epsilon^2} \geq || \, |\psi_{CSS1}\rangle\langle\psi_{CSS1}|_{AB} - |\psi'\rangle\langle\psi'|_{AB} \, ||_1.$$

By the triangle inequality we obtain

$$|| \, |\psi_1\rangle\langle\psi_1|_{AB} - |\psi'\rangle\langle\psi'|_{AB} \, ||_1$$

$$\leq || \, |\psi_1\rangle\langle\psi_1|_{AB} - |\psi_{CSS1}\rangle\langle\psi_{CSS1}|_{AB} \, ||_1 + || \, |\psi_{CSS1}\rangle\langle\psi_{CSS1}|_{AB} - |\psi'\rangle\langle\psi'|_{AB} \, ||_1 \leq O(\sqrt{\epsilon}) + O(d\sqrt{\delta}).$$

The minimum size of the embezzling state to achieve the fidelity of at least $1-\epsilon$ follows directly from Theorem 2.1.1 and is $n > d^{2/\epsilon}$ □

3.3 Noisy Embezzlement Protocols - Comparison

We can immediately notice that the classical synchronization with canonical embezzlement protocol may offer a significant improvement when compared to the quantum correlated sampling protocol. To prepare grounds for a quantitative comparison, we need to notice that the former protocol operates on a promise that $|| \, |\psi_1\rangle - |\psi_2\rangle \, ||_2 \leq \delta$. which implies $|| \, |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| \, ||_1 \leq O(\delta)$ whereas the latter assumes that $|| \, |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| \, ||_1^2 \leq \delta$. Therefore, to unify the assumptions, we will assume that the quantum correlated sampling lemma has a promise $|| \, |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| \, ||_1 \leq \sqrt{\delta}$. In this case, this protocol guarantees that it produces the state $|\psi_1'\rangle$ such that $\Pr(|\psi_1'\rangle = |\psi_2'\rangle) \geq 1 - O(\delta^{1/10})$ and $|| \, |\psi_1'\rangle\langle\psi_1'| - |\psi_1\rangle\langle\psi_1| \, ||_1 \leq O(\delta^{1/20})$, where $|\psi_1\rangle$ is one of the inputs. The classical synchronization with canonical embezzlement protocol proposed in this thesis, guarantees $\Pr(|\psi_1'\rangle = |\psi_2'\rangle) \geq 1 - O(d\sqrt{\delta})$ and $|| \, |\psi_1'\rangle\langle\psi_1'| - |\psi_1\rangle\langle\psi_1| \, ||_1 \leq O(d\sqrt{\delta})$. Therefore, in most cases we can achieve the same goal with much higher probability of success, while obtaining a state which is much closer to the desired one. The only caveat of the classical synchronization with canonical embezzlement protocol is the dependence on the size of the state which is not present in the quantum correlated sampling protocol. It means that in the regime of high-dimensional states, it may still be better to use the latter. We can come up with a straightforward criterion which tells us which protocol is more suitable for a given state dimension and discrepancy. The classical synchronization with canonical embezzlement protocol is better, in terms of distance, for states which size satisfy

$$O(d\sqrt{\delta}) < O(\delta^{1/20}) \implies d < O(\delta^{-9/20}).$$

When it comes to other advantages of the method that we propose in this thesis, it is worth noting that our solution is more 'classical' and quantum operations are actually only limited to the canonical embezzlement. It is a benefit because quantum operations and resources are usually considered to be more expensive and more difficult to execute in practice than classical ones. Major differences include the facts that the quantum correlated sampling protocol requires additional quantum entanglement in the state in Step 2 and assumes that Alice and Bob have apparatuses to measure their states. What is more, since it asks to embezzle N copies of the state $|\xi_0\rangle$, the size of the embezzling state is much bigger when compared to our approach where only one state of the size of the target state is embezzled which is $O(d^2/\delta)$ qubits versus $O(\log d/\delta)$ qubits.

The quantum correlated sampling protocol proposed in [12] was immediately used to prove the Parallel Repetition Theorem for entangled projection games in the non-expanding case in the same work. From the discussion above it follows that our classical synchronization with canonical embezzlement protocol could be, to some extent, applied to prove it as well, possibly improving the result in certain regimes which is discussed in the next section.

3.4 Application to the Quantum Parallel Repetition Theorem for projection nonlocal games

In this section we briefly describe the application of the noisy embezzlement to the Quantum Parallel Repetition Theorem for projection games proved in [12]. We will begin with necessary definitions based on [5].

Definition 26 (Nonlocal game)**.** *A nonlocal game between two players is defined as* $G = (S, T, A, B, \pi, V)$*, where* S *and* T *denote finite sets of possible questions to the first and the second player correspondingly,* π *is a probability distribution defined on* $S \times T$*,* A *and* B *denote finite sets of possible answers given by the first and the second player correspondingly and* $V : A \times B \times S \times T \to \mathbb{R}$ *is the acceptance criterion (also called a payoff function). It is assumed that players cooperate with each other to maximize the payoff function but cannot communicate.*

Definition 27 (Projection game)**.** *A projection game is a nonlocal game such that for any pair of questions, any answer given by a player determines at most one valid answer of another player.*

As in [12], we also define the following.

Definition 28 (Square of a game)**.** *Let G be a nonlocal game. In the square of a game G, denoted $G^\dagger G$, the referee samples a question u for the first player and then independently samples questions v, v' for the second player using the distribution π but conditioned on u. Each player in $G^\dagger G$ is given a question v and v' respectively. Players should produce answers b, b' respectively, such that there exists a valid answer a such that (a, b) satisfies (u, v) and (a, b') satisfies (u, v').*

Definition 29 (Expanding game)**.** *Suppose G is a projection game. Let H be the weighted adjacency matrix for $G^\dagger G$ with (v, v')-th entry given by $\pi(v, v') = \sum_u \pi(u)\pi(v'|u)\pi(v|u)$. Let D be the diagonal matrix, with $\pi_R(v)$ on its diagonal, where $\pi_R(v)$ is the marginal distribution on the set questions of for the second player. We define the Laplacian L associated with $G^\dagger G$ as $L = I - D^{-1/2} H D^{-1/2}$. The family of games (G_n), where G_n is of size n, is called expanding if the second smallest eigenvalue of $L(G_n)$ is a positive constant with no dependence on n.*

Definition 30 (Quantum strategy)**.** *A quantum strategy for a nonlocal game G is defined as $p_G = \left(|\psi\rangle, \left\{\left(A_s^a\right)_{a\in A}\right\}_{s\in S}, \left\{\left(B_t^b\right)_{b\in B}\right\}_{t\in T}\right)$, where $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a bipartite entangled quantum state shared between players and $\left\{\left(A_s^a\right)_{a\in A}\right\}_{s\in S}, \left\{\left(B_t^b\right)_{b\in B}\right\}_{t\in T}$ are sets of POVM measurements that, depending on questions asked, can be used to measure $|\psi\rangle$ by player A and player B correspondingly, yielding answers a and b. A quantum strategy can be assigned a value, given by*

$$\mathrm{VAL}(p_G) = \sum_{(s,t)\in S\times T} \pi(s,t) \sum_{(a,b)\in A\times B} \langle\psi| A_s^a \otimes B_t^b |\psi\rangle V(a,b,s,t).$$

Definition 31 (Entangled value of a game)**.** *An entangled value of a game, $\mathrm{VAL}^*(G)$, is*

$$\mathrm{VAL}^*(G) = \sup_{p_G\in P} \mathrm{VAL}(p_G),$$

where P is the set of all possible strategies.

Theorem 3.4.1 (Parallel Repetition Theorem [12])**.** *There exists constants $c, C > 0$ such that for any projection game G,*

$$\mathrm{VAL}^*(G^{\otimes k}) \leq \left(1 - C\left(1 - \mathrm{VAL}^*(G)\right)^c\right)^{k/2},$$

where $G^{\otimes k}$ means that k games G are played in parallel by the players, $c = 1$ in case of expanding games and $c \leq 10$ in the non-expanding case.

We note that usually when considering optimal quantum strategies for nonlocal games, a fixed quantum state is considered to be the part of it. In [12], it is assumed that players share an embezzling state as a resource, therefore, they are capable of generating any bipartite entangled quantum state on demand. In such a situation, it is reasonable to drop the requirement that a quantum state for which we evaluate the value of a game is universal for all pairs of questions. The authors of [12] proved that in an expanding case, players can use roughly the same entangled quantum state as part of their strategy for a near-optimal performance, independent of the questions being ask by a referee. It is not the case however, in case of non-expanding games. Then, players need to be capable of preparing different entangled quantum states that depend on questions given. Since in quantum games no form of communication is allowed, players cannot prepare such a tailored bipartite entangled state unless they share an embezzling state as a resource. What is more, descriptions of the state that they need to prepare for a particular set of questions differ slightly from each other. It is because a player does not know a question asked to another player, while the state to be embezzled in general depends on both questions. Luckily, the authors proved in Lemma 15 in [12] that in their setting, which assumes that players are only asked neighbouring questions (neighbouring in the constraints graph induced by $G^\dagger G$), each player can assume that a state does not depend on a question of another player, without giving up much on an overall performance - the descriptions of quantum states that they obtain are close to each other. In this case, the canonical embezzlement of entanglement result does not apply but the authors use the quantum correlated sampling protocol which they developed for this purpose.

We note that for the non-expanding case of the Parallel Repetition Theorem, we can also use our classical synchronization scheme with canonical embezzlement instead of the quantum correlated sampling protocol. Our protocol improves the dependence on δ, however introducing a dependence on the size of the state to be embezzled. It results in $c = \frac{1}{1+\log_{\sqrt{\delta}} d}$ in Theorem 3.4.1. shared randomness can be generated by the players by additionally embezzling a sufficient number of EPR pairs and then measuring them.

Chapter 4

Embezzlement in Dilution of Entanglement

In this chapter we focus on dilution of entanglement assisted by an embezzling state. First, we describe developments which proved that this task without such assistance requires classical communication. We introduce the concept of strong typicality and prove its properties which we then use to provide a new explicit protocol for dilution of entanglement assisted by an embezzling state without communication. Following existing literature, we discuss why is it possible, using the concept of the entanglement spread. Then, we apply our classical synchronization scheme to generalize our protocol of dilution of entanglement assisted by an embezzling state to a noisy scenario.

4.1 Communication Cost of Dilution of Entanglement

Concentration and dilution of entanglement, discussed in the Introduction, can be considered inverses of each other. However, we may notice a curious asymmetry, namely that for the concentration of entanglement it is enough to use local operations only, whereas in case of the dilution of entanglement, the classical communication is required as well. To be precise, the protocol from [2] requires $O(n)$ classical bits. Motivated to reduce the amount of classical communication, Lo and Popescu showed in [15] that there is a protocol that requires only $O(\sqrt{n})$ bits of classical communication for the task of dilution of entanglement. To do so, they used the concept of strong typicality and the teleportation of entanglement. Next, it was proved by Harrow and Lo in [10] and by Hayden and Winter in [11] that the

task actually requires $\Omega(\sqrt{n})$ bits of classical communication and therefore no protocol can use less than that. The proof in [11] relies on the concept of entanglement spread which relationship to the necessity of classical communication in entanglement transformations is then emphasized and discussed by Harrow in [9]. We explain this relationship in the next section.

4.2 Entanglement Spread and Classical Communication Cost

We define the entanglement spread introduced in [11]. Intuitively, it is intended to capture variations in eigenvalues of a quantum state.

Definition 32 (Entanglement spread)**.** *The entanglement spread of a bipartite quantum state $\rho_{AB} \in D(\mathcal{H})$ is defined as*

$$\Delta(\rho_{AB}) = \log \operatorname{rank} \rho_A + \log ||\rho_A||_\infty,$$

where $\rho_A = \operatorname{Tr}_B(\rho_{AB})$, $\operatorname{rank} \operatorname{Tr}_B(\rho_{AB})$ is the number of non-zero eigenvalues of ρ_{AB} and $||\rho_{AB}||_\infty$ is the largest eigenvalue of ρ_{AB}.

It is useful to notice the following properties of the entanglement spread

$$\Delta(\rho_{AB} \otimes \sigma_{AB}) = \Delta(\rho_{AB}) + \Delta(\sigma_{AB}),$$

$$\forall \rho_{AB} \quad \Delta(\rho_{AB}) \geq 0,$$

$$\Delta(\rho_{AB}) = 0 \iff \rho_{AB} \text{ is a maximally entangled or a product state.}$$

The same authors introduce the concept of an ϵ-perturbed entanglement spread. It is because the entanglement spread is a quantity which is not robust against small perturbations of a quantum state.

Definition 33 (ϵ-perturbed entanglement spread)**.** *Let $\epsilon \geq 0$. Then the ϵ-perturbed entanglement spread of a bipartite quantum state ρ_{AB} is defined as*

$$\Delta_\epsilon(\rho_{AB}) = \min_P \left\{ \log \operatorname{Tr} P + \log ||P \rho_A P||_\infty : \operatorname{Tr} P \rho_A \geq 1 - \epsilon \right\},$$

where $\rho_A = \operatorname{Tr}_B(\rho_{AB})$ and the minimization is over all projectors P.

The Theorem which we will now state provided a lower bound on classical communication necessary for entanglement transformations in terms of the entanglement spread and the ϵ-perturbed entanglement spread. It is used in [11] to prove the lower bound on classical communication necessary for the process of dilution of entanglement, mentioned in the previous section. The proof relies on the fact that the entanglement spread of EPR pairs, the initial state of the protocol, is zero. Then, a lower bound on the ϵ-perturbed entanglement spread of the final state is derived using typicality arguments.

Theorem 4.2.1. *Suppose we transform a bipartite quantum state $|\phi\rangle_{AB}$ into a bipartite quantum state $|\psi\rangle_{AB}$ with fidelity $1-\epsilon$, using local operations and C bits of classical communication (in either direction). Then*

$$C \geq \Delta_\delta(\rho^\psi_{AB}) - \Delta_0(\rho^\phi_{AB}) + 2\log(1-\delta),$$

where $\delta = (4\epsilon)^{1/8}$.

Proof. See [11]. □

The bound in the Theorem above suggests a strong relationship between the classical communication cost of entanglement transformations and the entanglement spread of pure quantum states involved. Indeed, we can recall the invariance of Schmidt coefficients (i.e. also of eigenvalues in case of pure states) under local operations presented in the Introduction. Since local operations can only reshuffle Schmidt coefficients, it poses restrictions on possible transformations of bipartite pure quantum states, including the invariance of the entanglement spread under their action (the number of non-zero eigenvalues and the greatest eigenvalue do not change). Then, classical communication is the only component of the assumed framework which can potentially change the entanglement spread. In [9], Harrow argues that it is indeed the fundamental reason for the asymmetry, in terms of classical communication cost, between the concentration and dilution of entanglement. Because of that, given some other source of entanglement spread in the protocol, the classical communication cost can be reduced to zero. A natural candidate for such a source that he suggests is an embezzling state which is known to be the state of high entanglement spread. An embezzling state offers coefficients of varying (to be precise, strictly decreasing) values and therefore gives us a lot of freedom to 'move' some of them with local operations to blank quantum registers, 'creating' quantum states of basically arbitrary entanglement spread (depending on the size of an available embezzling state). Moreover, as proved in the canonical embezzlement result, it can be done with arbitrarily high fidelity (again depending on the size of an available embezzling state). It should be stressed that it does not necessarily mean creating the whole target state through embezzlement from

scratch. Depending on the particular transformation considered and the initial quantum state available (apart from an embezzling state), it might be enough to just introduce the missing or destroy the excessive amount of entanglement spread, which would use a smaller embezzling state than that required in the canonical embezzlement of the whole target state. Indeed, Harrow claims in [9] that the dilution of entanglement assisted by an embezzling state of size $O(\sqrt{n}/\epsilon)$ qubits is enough to dilute n partially entangled states from EPR pairs with infidelity ϵ and without classical communication. We note that the canonical embezzlement of n such partially entangled states would require $O(n/\epsilon)$ qubits. In the following sections of this thesis we propose an explicit protocol for the dilution of entanglement assisted by an embezzling state of size $O(\sqrt{n}/\epsilon)$ qubits. As a starting point to do so, we explain the concept of strong typicality in the next section.

4.3 Strong Typicality

Following [16], we will now introduce the concept of strong typicality, also known as letter typicality. It will be crucial for understanding protocols presented in the next subsections. We consider a memoryless source $\mathcal{X}$ which generates letters from the set $\{a_1, a_2, \ldots, a_M\}$. The probability distribution associated with the source is denoted by p, i.e., $\Pr(X_k = a_i) := p_i$ and $\sum_{i=1}^{M} p_i = 1$. We say x^n is a sequence which is an instance of n independent identically distributed random variables $\{X_k\}_{k=1}^{n}$ such that $X_k \in \mathcal{X} \quad \forall k = 1, 2, \ldots, n$.

Definition 34 (K-letter typical sequence)**.** *Suppose we are given $k > 0$. A sequence x^n is called k-letter typical if and only if*

$$\left| \frac{f_i(x^n) - np_i}{\sqrt{np_i(1-p_i)}} \right| < k \quad \forall n, i = 1, 2, \ldots, M,$$

where $f_i(x^n)$ is the number of positions in a sequence x^n which are equal to a_i. The set of all k-letter typical sequences is defined as follows

$$\mathcal{T}^k := \{x^n \in \mathcal{X}^n : x^n \text{ is } k\text{-letter typical}\}.$$

Before we state the theorem regarding k-letter typicality, we state the Chebyshev's inequality which we will use to prove the theorem.

Lemma 13 (Chebyshev's inequality)**.** *Let X be a random variable with expected value μ and non-zero variance σ^2 which are both finite. For any real number $k > 0$ we have*

$$\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Now, we formally state and prove the theorem regarding k-letter typicality.

Theorem 4.3.1. *For any $\epsilon > 0$, $k > \sqrt{\frac{M}{\epsilon}}$ and $n \geq 1$, we have*

1. $$\Pr\big(X^n \in \mathcal{T}^k\big) > 1 - \epsilon,$$

2. $$\exists \quad A(k,p) > 0 : 2^{-nH - A\sqrt{n}} < \Pr(X^n = x^n) < 2^{-nH + A\sqrt{n}} \quad \forall x^n \in \mathcal{T}^k,$$

   where $H := -\sum_{i=1}^{M} p_i \log_2 p_i$ is an entropy of the source.

Proof. 1.

$$\Pr\big(X^n \notin \mathcal{T}^k\big) = \Pr\left(\exists i : \left|\frac{f_i(X^n) - np_i}{\sqrt{np_i(1-p_i)}}\right| \geq k\right) \leq$$

$$\sum_{i=1}^{M} \Pr\left(\left|\frac{f_i(X^n) - np_i}{\sqrt{np_i(1-p_i)}}\right| \geq k\right) \leq \sum_{i=1}^{M} \frac{1}{k^2} = \frac{M}{k^2} < \epsilon,$$

where in the second to last inequality we used Chebyshev's inequality. In our case, the random variable in the Chebyshev's inequality is $X_i = f_i(X^n)$. We notice that our random variable has a multinomial distribution which is known to have the mean $\mu_i = np_i$ and variance $\sigma_i = np_i(1-p_i)$. Then, it follows that

$$\Pr\big(X^n \in \mathcal{T}^k\big) > 1 - \epsilon,$$

when we choose $k > \sqrt{\frac{M}{\epsilon}}$.

2. Let us consider a sequence x^n which is k-letter typical. From the definition of k-typicality, we know that

$$np_i - k\sqrt{np_i(1-p_i)} < f_i(x^n) < np_i + k\sqrt{np_i(1-p_i)},$$

which gives us a lower and an upper bound on $f_i(x^n)$. Since we assumed that X^n is a sequence of independent identically distributed random variables, we know that

$$\Pr(X^n = x^n) = \Pr(X_1 = x_1)\Pr(X_2 = x_2)\ldots\Pr(X_n = x_n) = \prod_{i=1}^{M} p_i^{f_i(x^n)},$$

which, after taking the logarithm of both sides, implies

$$-\log \Pr(X^n = x^n) = -\sum_{i=1}^{M} f_i(x^n) \log p_i.$$

We can now use the lower and upper bounds on $f_i(x^n)$ in the equation above to obtain

$$-\sum_{i=1}^{M} \left(np_i - k\sqrt{np_i(1-p_i)} \right) \log p_i <$$

$$< -\log \Pr(X^n = x^n) < -\sum_{i=1}^{M} \left(np_i + k\sqrt{np_i(1-p_i)} \right) \log p_i,$$

which can be rewritten as

$$nH - A\sqrt{n} < -\log \Pr(X^n = x^n) < nH + A\sqrt{n},$$

where H is an entropy function and $A := -k \sum_{i=1}^{M} \sqrt{np_i(1-p_i)}$. By taking an exponent of both sides we obtain

$$2^{-nH - A\sqrt{n}} < \Pr(X^n = x^n) < 2^{-nH + A\sqrt{n}}.$$

□

4.4 Dilution of Entanglement Without Communication

As discussed in the previous section, the task of diluting n partially entangled states requires classical communication of order $\Theta(\sqrt{n})$. In this section, we will present a new protocol which allows to perform the task of dilution of entanglement without any communication and with arbitrarily high fidelity, using a canonical embezzling state as a resource and given that n is large. It might appear trivial at first because we proved that it is possible to embezzle any bipartite entangled state, in particular n copies of a target partially entangled state. However, in such a case the size of a required embezzling state would be huge, namely $O(n/\epsilon)$ qubits. Our protocol requires an embezzling state of size only $O(\sqrt{n}/\epsilon)$, while achieving comparable fidelity. The protocol with this size of an embezzling state is the realization of a claim made by Harrow in Section 3.1 of [9].

No-communication dilution of entanglement protocol

Input:
Alice and Bob: a classical description of a bipartite quantum state $|\psi\rangle = (a+ia')|00\rangle + (b+ib')|11\rangle$, $\epsilon_1, \epsilon_2 > 0, n$ large.
Shared resources: $(nE - O(\sqrt{n}/\epsilon_1))$ EPR pairs $\left|\Phi_2^+\right\rangle$, a catalyst state $|\mu\rangle$ of size $O(\sqrt{n}/\epsilon_1)$ qubits.
Promise: Descriptions provided in a fixed basis $\{|00\rangle, |11\rangle\}$ known to both parties.
Goal: Alice and Bob share a quantum state $|\psi_n\rangle$ such that $|| \, |\psi_n\rangle\langle\psi_n| - |\psi\rangle\langle\psi|^{\otimes n} \, ||_1 \leq O(\sqrt{\epsilon_1} + \sqrt{\epsilon_2})$.
Protocol:

1. Alice and Bob calculate $E = E(\mathrm{Tr}_A |\psi\rangle)$, the von Neumann entropy of their subsystem of $|\psi\rangle$.

2. Alice and Bob calculate the Schmidt decomposition of $|\psi\rangle$, characterized by coefficients $\{a_S, b_S\}$.

3. For each $k = 0, 1, \ldots, n$, there are $\binom{n}{k}$ coefficients of the form $a_S^k b_S^{n-k}$ in $|\psi\rangle^{\otimes n}$. Alice and Bob build the set $\tau(\epsilon_1)$ of values of k which are strongly typical

$$\tau(\epsilon_1) = \left\{ k : nE - \sqrt{\frac{12}{\epsilon_1}}\sqrt{n} \leq \log \binom{n}{k} \leq nE + \sqrt{\frac{12}{\epsilon_1}}\sqrt{n} \right\}.$$

4. For each $k \in \tau(\epsilon_1)$, Alice and Bob group corresponding coefficients of $|\psi\rangle^{\otimes n}$ into bins of size $2^{nE - 2\sqrt{\frac{12n}{\epsilon_1}}}$ such that

$$n_k 2^{nE - 2\sqrt{\frac{12n}{\epsilon_1}}} \leq \binom{n}{k} \leq (n_k + 1) 2^{nE - 2\sqrt{\frac{12n}{\epsilon_1}}},$$

   where n_k is the number of full bins.

5. For each $k \in \tau(\epsilon_1)$, Alice and Bob take first $n_k 2^{nE - 2\sqrt{\frac{12n}{\epsilon_1}}}$ coefficients from initial $\binom{n}{k}$ and put their corresponding vectors into a set V_k.

6. Alice and Bob use the expression

$$|\psi_n\rangle = \left|\phi_2^+\right\rangle^{\otimes (nE - \sqrt{\frac{48n}{\epsilon_1}})} \otimes |\chi\rangle \propto \left(\sum_{i=0}^{2^{nE - \sqrt{\frac{48n}{\epsilon_1}}}} |ii\rangle \right) \otimes \left(\sum_{j=0}^{2^{\sqrt{\frac{48n}{\epsilon_1}}}} \alpha_j |jj\rangle \right),$$

   and solve the system of linear equations $|\psi\rangle^{\otimes n} = |\psi_n\rangle$ to obtain the values of α_j. When they encounter a coefficient of $|\psi\rangle^{\otimes n}$ which corresponds to a vector not from the set $(\bigcup_k V_k)$, they act as if it equaled to 0.

7. Alice and Bob, having the description of $|\chi\rangle$, embezzle it using $|\mu\rangle$ and obtain the state $|\chi_{emb}\rangle$.

8. Alice and Bob share the state $|\eta_{emb}\rangle = \left|\phi_2^+\right\rangle^{\otimes(nE-\sqrt{\frac{48n}{\epsilon_1}})} \otimes |\chi_{emb}\rangle \otimes |\mu\rangle$ which resembles the state $|\eta\rangle = |\psi\rangle^{\otimes n} \otimes |\mu\rangle$.

Theorem 4.4.1. *In the no-communication dilution of entanglement protocol,* $|| \, |\eta\rangle\langle\eta| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1 \leq O(\sqrt{\epsilon_1}+\sqrt{\epsilon_2})$ *in the limit of large* n *and the embezzling state* $|\mu\rangle$ *is of size* $\sqrt{\frac{48}{\epsilon_1\epsilon_2^2}}\sqrt{n}$ *qubits.*

Proof. Steps 1-5 of the protocol directly follow the idea from [15] and exact constants are calculated based on Theorem 4.3.1. To justify exact constants used in Step 2, we refer to the strong typicality which says that typical values of k are those which satisfy $nE - A\sqrt{n} \leq \log \binom{n}{k} \leq nE + A\sqrt{n}$ for $A = -C\sum_{i=1}^{2}\sqrt{p_i(1-p_i)}\log p_i$ and we need to choose C such that $C > \sqrt{\frac{2}{\epsilon}}$. We choose $C = \sqrt{\frac{3}{\epsilon}}$. Variables p_i refer to the probability distribution on symbols in our alphabet which is $\{0,1\}$. The maximum value of $\frac{A}{C}$ is $\frac{2}{e\ln(2)}$ which for clarity we bound by 2. It lets us introduce a universal constant $A_{univ} = \sqrt{\frac{12}{\epsilon}}$. Let us now prove the bound for the distance. Alice and Bob would like to obtain the state $|\eta\rangle$. By the strong typicality, they deduce that they can obtain the state $|\eta'\rangle = \left|\phi_2^+\right\rangle^{\otimes(nE-\sqrt{\frac{48n}{\epsilon_1}})} \otimes |\chi\rangle \otimes |\mu\rangle$ such that $F(|\eta'\rangle, |\eta\rangle) \geq 1-\epsilon_1$. Therefore, they intend to embezzle the state $|\chi\rangle$, however, in practice they obtain the state $|\chi_{emb}\rangle$ such that $F(|\chi\rangle \otimes |\mu\rangle, |\chi_{emb}\rangle \otimes |\mu\rangle) \geq 1-\epsilon_2$ using the embezzling state of size $\sqrt{\frac{48}{\epsilon_1\epsilon_2^2}}\sqrt{n}$ qubits [17]. Thus, $F(|\eta'\rangle, |\eta_{emb}\rangle) \geq 1-\epsilon_2$. We use the relationship between fidelity and Schatten-1 distance for quantum states and get

$$2\sqrt{2\epsilon_1-\epsilon_1^2} + 2\sqrt{2\epsilon_2-\epsilon_2^2} \geq || \, |\eta\rangle\langle\eta| - |\eta'\rangle\langle\eta'| \, ||_1 + || \, |\eta'\rangle\langle\eta'| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1 \geq$$
$$\geq || \, |\eta\rangle\langle\eta| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1.$$

Therefore, $|| \, |\eta\rangle\langle\eta| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1 \leq O(\sqrt{\epsilon_1}+\sqrt{\epsilon_2})$. □

As noted in [15], the typicality technique can be easily generalized for the case of qudits. Therefore, the same applies to the protocol above.

4.5 Noisy Dilution of Entanglement Without Communication

In the previous section we proposed the protocol for dilution of entanglement without communication where we assumed that Alice and Bob have a perfect knowledge of a common bipartite state that they want to obtain. In this section we generalize the protocol to a setting in which there is a small discrepancy between Alice's and Bob's description. For our task, we might use the quantum correlated sampling lemma as a tool for noisy embezzlement, or, we might use the classical synchronization scheme to synchronize Alice's and Bob's description and then make them use the canonical embezzlement. We will choose the latter because, as discussed earlier, it guarantees a much better precision in the regime of low-dimensional states (a qubit state in our case).

Noisy no-communication dilution of entanglement protocol

Input:
Alice: a biased classical description $|\psi_1\rangle$ of a bipartite quantum state $|\psi\rangle$, $|\psi_1\rangle = (a_1 + ia_1')|00\rangle + (b_1 + ib_1')|11\rangle$,
Bob: a biased classical description $|\psi_2\rangle$ of a bipartite quantum state $|\psi\rangle$, $|\psi_2\rangle = (a_2 + ia_2')|00\rangle + (b_2 + ib_2')|11\rangle$,
Alice and Bob: $\epsilon_1, \epsilon_2, \delta > 0, n$ large.
Shared resources: $(nE - O(\sqrt{n}/\epsilon_1))$ EPR pairs $\left|\Phi_2^+\right\rangle$, a catalyst state $|\mu\rangle$ of size $O(\sqrt{n}/\epsilon_1)$ qubits, shared randomness.
Promise: $|| \, |\psi_1\rangle - |\psi_2\rangle \, ||_2 \leq \delta$, $|| \, |\psi\rangle - |\psi_1\rangle \, ||_2 \leq \delta$, descriptions provided in a fixed basis $\{|00\rangle, |11\rangle\}$ known to both parties.
Goal: With high probability, Alice and Bob share a quantum state $|\psi_n\rangle$ such that $|| \, |\psi_n\rangle\langle\psi_n| - |\psi\rangle\langle\psi|^{\otimes n} \, ||_1 \leq O(\sqrt{\epsilon_1} + \sqrt{\epsilon_2} + \sqrt{n\delta})$.
Protocol:

1. Alice and Bob use the classical synchronization scheme to agree on the common classical description of the state $|\psi_{CSS}\rangle$ based on their descriptions $|\psi_1\rangle$ and $|\psi_2\rangle$.

2. Alice and Bob calculate $E = E(\mathrm{Tr}_A \, |\psi_{CSS}\rangle)$, the von Neumann entropy of their subsystem of $|\psi_{CSS}\rangle$.

3. Alice and Bob calculate the Schmidt decomposition of $|\psi_{CSS}\rangle$, characterized by coefficients $\{a_S, b_S\}$.

4. For each $k = 0, 1, \ldots, n$, there are $\binom{n}{k}$ coefficients of the form $a_S^k b_S^{n-k}$ in $|\psi_{CSS}\rangle^{\otimes n}$. Alice and

Bob build the set $\tau(\epsilon_1)$ of values of k which are strongly typical

$$\tau(\epsilon_1) = \left\{ k : nE - \sqrt{\frac{12}{\epsilon_1}}\sqrt{n} \leq \log \binom{n}{k} \leq nE + \sqrt{\frac{12}{\epsilon_1}}\sqrt{n} \right\}.$$

5. For each $k \in \tau(\epsilon_1)$, Alice and Bob group corresponding coefficients of $|\psi_{CSS}\rangle^{\otimes n}$ into bins of size $2^{nE-2\sqrt{\frac{12n}{\epsilon_1}}}$ such that

$$n_k 2^{nE-2\sqrt{\frac{12n}{\epsilon_1}}} \leq \binom{n}{k} \leq (n_k+1)2^{nE-2\sqrt{\frac{12n}{\epsilon_1}}},$$

where n_k is the number of full bins.

6. For each $k \in \tau(\epsilon_1)$, Alice and Bob take first $n_k 2^{nE-2\sqrt{\frac{3n}{\epsilon_1}}}$ coefficients from initial $\binom{n}{k}$ and put their corresponding vectors into a set V_k.

7. Alice and Bob use the expression

$$|\psi^n_{CSS}\rangle = \left|\phi_2^+\right\rangle^{\otimes(nE-\sqrt{\frac{48n}{\epsilon_1}})} \otimes |\chi\rangle \propto \left(\sum_{i=0}^{2^{nE-\sqrt{\frac{48n}{\epsilon_1}}}} |ii\rangle \right) \otimes \left(\sum_{j=0}^{2^{\sqrt{\frac{48n}{\epsilon_1}}}} \alpha_j |jj\rangle \right),$$

and solve the system of linear equations $|\psi_{CSS}\rangle^{\otimes n} = |\psi^n_{CSS}\rangle$ to obtain the values of α_j. When they encounter the coefficient of $|\psi_{CSS}\rangle^{\otimes n}$ which corresponds to a vector not from the set $(\bigcup_k V_k)$, they act as if it equaled to 0.

8. Alice and Bob, having the description of $|\chi\rangle$, embezzle it using $|\mu\rangle$ and obtain the state $|\chi_{emb}\rangle$.

9. Alice and Bob share the state $|\eta_{emb}\rangle = \left|\phi_2^+\right\rangle^{\otimes(nE-\sqrt{\frac{48n}{\epsilon_1}})} \otimes |\chi_{emb}\rangle \otimes |\mu\rangle$ which resembles the state $|\eta\rangle = |\psi\rangle^{\otimes n} \otimes |\mu\rangle$.

Theorem 4.5.1. *In the noisy no-communication dilution of entanglement protocol,* $|| \, |\eta\rangle\langle\eta| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1 \leq O(\sqrt{\epsilon_1} + \sqrt{\epsilon_2} + \sqrt{n\delta})$ *in the limit of large* n *and the embezzling state* $|\mu\rangle$ *is of size* $\sqrt{\frac{48}{\epsilon_1\epsilon_2^2}}\sqrt{n}$ *qubits. The protocol succeeds with probability greater than* $1 - O(\sqrt{\delta})$ *and requires* $O\left(\log_2 \frac{1}{\delta}\right)$ *bits of shared randomness.*

Proof. Steps 2-6 of the protocol directly follow the idea from [15] and exact constants are calculated based on [16]. It follows from the Theorem 3.2.2 that Alice and Bob need $O\left(\log_2 \frac{1}{\delta}\right)$ bits of shared randomness to perform Step 1 of the protocol. Moreover, in this

step they obtain a description $|\psi_{CSS}\rangle$ such that $|| \, |\psi_{CSS}\rangle - |\psi_1\rangle \, ||_2 \leq 4\sqrt{\delta}$ with probability at least $1 - O(\sqrt{\delta})$. By the triangle inequality and using the promise, it follows that $|| \, |\psi_{CSS}\rangle - |\psi\rangle \, ||_2 \leq 4\sqrt{\delta} + \delta$. By using the Lemma 3, we obtain

$$|| \, |\psi_{CSS}\rangle\langle\psi_{CSS}| - |\psi\rangle\langle\psi| \, ||_1 \leq O(\sqrt{\delta}).$$

We use the relationship between the Schatten-1 distance and fidelity of pure states

$$F(|\psi_{CSS}\rangle \, , |\psi\rangle) = \sqrt{1 - \frac{1}{2}|| \, |\psi_{CSS}\rangle\langle\psi_{CSS}| - |\psi\rangle\langle\psi| \, ||_1^2}.$$

We consider the fidelity between n copies of both states which is

$$F(|\psi_{CSS}\rangle^{\otimes n} \, , |\psi\rangle^{\otimes n}) = \left(1 - \frac{1}{2}|| \, |\psi_{CSS}\rangle\langle\psi_{CSS}| - |\psi\rangle\langle\psi| \, ||_1^2\right)^{n/2}.$$

It implies that

$$|| \, |\psi_{CSS}\rangle\langle\psi_{CSS}|^{\otimes n} - |\psi\rangle\langle\psi|^{\otimes n} \, ||_1 = 2\sqrt{1 - \left(1 - \frac{1}{2}|| \, |\psi_{CSS}\rangle\langle\psi_{CSS}| - |\psi\rangle\langle\psi| \, ||_1^2\right)^n} \leq$$

$$\leq 2\sqrt{1 - \left(1 - \frac{1}{2}(4\sqrt{\delta} + \delta)^2\right)^n} = O(\sqrt{n\delta}).$$

The proof of Steps 2-5 is the same as in the Theorem 4.4.1. As a result of these steps, they have a description of $|\psi_{CSS}^n\rangle$ such that $F(|\psi_{CSS}^n\rangle \, , |\psi_{CSS}\rangle^{\otimes n}) \geq 1 - \epsilon_1$ which implies

$$|| \, |\psi_{CSS}^n\rangle\langle\psi_{CSS}^n| - |\psi_{CSS}\rangle\langle\psi_{CSS}|^{\otimes n} \, ||_1 = 2\sqrt{1 - F^2(|\psi_{CSS}^n\rangle \, , |\psi_{CSS}\rangle^{\otimes n})} \leq O(\sqrt{\epsilon_1}).$$

Alice and Bob embezzle the state $|\chi\rangle$ and obtain the state $|\chi_{emb}\rangle$ such that $F(|\chi\rangle \, , |\chi_{emb}\rangle) \geq 1 - \epsilon_2$. Therefore,

$$|| \, |\psi_{CSS}^n\rangle\langle\psi_{CSS}^n| \otimes |\mu\rangle\langle\mu| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1 \leq 2\sqrt{1 - F^2(\psi_{CSS}^n \otimes |\mu\rangle \, , \eta_{emb})} \leq O(\sqrt{\epsilon_2}).$$

We use the triangle inequality twice and obtain

$$O(\sqrt{n\delta} + \sqrt{\epsilon_1} + \sqrt{\epsilon_2}) \geq || \, |\psi_{CSS}\rangle\langle\psi_{CSS}|^{\otimes n} \otimes |\mu\rangle\langle\mu| - |\psi\rangle\langle\psi|^{\otimes n} \otimes |\mu\rangle\langle\mu| \, ||_1 +$$

$$+ || \, |\psi_{CSS}^n\rangle\langle\psi_{CSS}^n| \otimes |\mu\rangle\langle\mu| - |\psi_{CSS}\rangle\langle\psi_{CSS}|^{\otimes n} \otimes |\mu\rangle\langle\mu| \, ||_1 + || \, |\psi_{CSS}^n\rangle\langle\psi_{CSS}^n| \otimes$$

$$\otimes |\mu\rangle\langle\mu| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1 \geq || \, |\eta\rangle\langle\eta| - |\eta_{emb}\rangle\langle\eta_{emb}| \, ||_1.$$

□

As noted in [15], the typicality technique can be easily generalized for the case of qudits. Therefore, the same applies to the protocol above. We note, however, that in case of qudits, the classical synchronization scheme introduces a dependence on the Schmidt rank of a qudit to be diluted. For the regime of high-dimensional states it may significantly affect the performance of the protocol and the quantum correlated sampling protocol could be then used instead and provide better results.

Chapter 5

Summary and Outlook

In this thesis we presented the concept of embezzlement of entanglement [17] which allows two parties to remotely prepare any pure bipartite entangled quantum state without any communication, given a special resource called an embezzling state. We elaborated on its properties [14], generalizations [14], universality [17, 14, 13], efficiency [14] and possibility of achieving a perfect fidelity [6]. Since quantum entanglement is a precious resource [15, 2] in quantum information theory, which we motivated by providing detailed examples of its applications [3, 4, 7], the possibility of its embezzlement is an important development with potentially huge implications for resource theories and quantum protocols.

When it comes to the creative contributions of this thesis, we discovered the linear programming characterization of embezzlement of entanglement that can be used to study its efficiency and to obtain unitary operations that achieve the maximum fidelity of this transformation. We note however, that due to large dimensions of the problem, the practical use of this characterization may be significantly restricted by classical computational resources. Nevertheless, linear programming offers many interesting properties and tools for analyzing problems, e.g. the Weak Duality Theorem or the Strong Duality Theorem and thus the formulation of the problem as a linear program may be a promising starting point for further research. Moreover, we proposed a new protocol for dilution of entanglement [15, 10, 11] in which an embezzling state is treated as a resource and successfully replaces the fundamental necessity for classical communication. This result is attributed to embezzling states being a good source of entanglement spread [9].

Taking into consideration that in real-life scenarios we usually experience noise that alters protocols, we heavily focused on the noisy embezzlement of entanglement. We provided an elaborated proof of the quantum correlated sampling lemma [12] which allows for

a similar result as embezzlement but in case of noisy inputs. We managed, to some extent, to improve this result by introducing the classical synchronization scheme which can be used together with the canonical embezzlement to achieve noisy embezzlement. Our protocol requires less quantum resources and offers a much smaller error which arises from noisy inputs, but introduces a dimensional dependence into it. Therefore, it is more suitable for input states of a small dimension. Using our noisy embezzlement protocol protocol, we also managed to generalize our embezzlement-assisted dilution of entanglement protocol to a noisy setting.

Given developments and applications of embezzlement of entanglement and noisy embezzlement of entanglement explained in this thesis, we believe that there are still many research opportunities regarding this phenomenon. For instance, it is an interesting problem on why does the quantum correlated sampling lemma have no dimensional dependence in its result, whereas in case of our classical synchronization scheme with canonical embezzlement it seemed inevitable. We might suspect that it is related to the former being 'more quantum' than the latter and using entanglement for synchronization between both parties. Another line of research would be to further investigate the benefits of embezzlement-assisted protocols, where embezzlement is used in a non-trivial way (i.e. with a limited size of an embezzling state), as in our protocols for entanglement dilution.

References

[1] Charles Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. volume 560, pages 175–179, 01 1984.

[2] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev.*, A53:2046–2052, 1996.

[3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[4] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[5] Richard Cleve. Lecture notes for QIC 890 Entanglement and Nonlocal Effects. *Institute for Quantum Computing, University of Waterloo*, Winter 2019. Last accessed 03-20-2019.

[6] Richard Cleve, Li Liu, and Vern I. Paulsen. Perfect embezzlement of entanglement. *Journal of Mathematical Physics*, 58(1):012204, 2017.

[7] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[8] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Communications in Mathematical Physics*, 31(4):291–294, dec 1973.

[9] A. W. Harrow. Entanglement spread and clean resource inequalities. 2013.

[10] A. W. Harrow and Hoi-Kwong Lo. A tight lower bound on the classical communication cost of entanglement dilution. *IEEE Transactions on Information Theory*, 50(2):319–327, Feb 2004.

[11] P. Hayden and A. Winter. On the communication cost of entanglement transformations. 2002.

[12] D. Steurer I. Dinur and T. Vidick. A parallel repetition theorem for entangled projection games. *IEEE Conference on Computational Complexity (CCC)*, pages 197–208, 2014.

[13] Debbie Leung, Ben Toner, and John Watrous. Coherent state exchange in multi-prover quantum interactive proof systems. *arXiv e-prints*, page arXiv:0804.4118, Apr 2008.

[14] Debbie Leung and Bingjie Wang. Characteristics of universal embezzling families. *Phys. Rev. A*, 90:042331, Oct 2014.

[15] Hoi-Kwong Lo and Sandu Popescu. Classical communication cost of entanglement manipulation: Is entanglement an interconvertible resource? *Phys. Rev. Lett.*, 83:1459–1462, Aug 1999.

[16] Saif K. Mohammed and Reza Moosavi. Lecture notes in typical sequences. *Communication Systems Division, Department of Electrical Engineering, Linkoping University, Sweden.*, Spring 2012. Last accessed 01-16-2019.

[17] Wim van Dam and Patrick Hayden. Universal entanglement transformations without communication. *Phys. Rev. A*, 67:060302, Jun 2003.

[18] Guifré Vidal, Daniel Jonathan, and M. A. Nielsen. Approximate transformations and robust manipulation of bipartite pure-state entanglement. *Physical Review A*, 62:012304, Jul 2000.

[19] John Watrous. Lecture notes for theory of quantum information. 2011. Last accessed 03-06-2019.

[20] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[21] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.