# Adaptive Protocols for the Quantum Depolarizing Channel

by

Alan W. Leung

B.A., The University of Cambridge, 2002
M.A., The University of Cambridge, 2006

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author ..................................................
Department of Mathematics
May 1, 2007

Certified by ..................................................
Peter W. Shor
Morss Professor of Applied Mathematics
Thesis Supervisor

Accepted by ..........................
Alar Toomre
Chairman, Applied Mathematics Committee

Accepted by ......................
Pavel I. Etingof
Chairman, Department Committee on Graduate Students

# Adaptive Protocols for the Quantum Depolarizing Channel

by

Alan W. Leung

Submitted to the Department of Mathematics
on May 1, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

## Abstract

In the first part, we present a family of entanglement purification protocols that generalize four previous methods, namely the recurrence method, the modified recurrence method, and the two methods proposed by Maneva-Smolin and Leung-Shor. We will show that this family of protocols have improved yields over a wide range of initial fidelities F, and hence imply new lower bounds on the quantum capacity assisted by two-way classical communication of the quantum depolarizing channel. In particular, we show that the yields of these protocols are higher than the yield of universal hashing for F less than 0.99999 and as F goes to 1.

In the second part, we define, for any quantum discrete memoryless channel, quantum entanglement capacity with classical feedback, a quantity that lies between two other well-studied quantities. These two quantities - namely the quantum capacity assisted by two-way classical communication and the quantum capacity with classical feedback - are widely conjectured to be different. We then present adaptive protocols for this newly-defined quantity on the quantum depolarizing channel. These protocols in turn imply new lower bounds on the quantum capacity with classical feedback.

Thesis Supervisor: Peter W. Shor
Title: Morss Professor of Applied Mathematics

# Acknowledgments

I am deeply indebted to my thesis advisor Peter Shor, without whom this thesis would not have been possible. I would like to thank him for being my advisor, his teaching, ingenious insights, meetings and discussions. All these are privileges I am very grateful for.

I would like to thank Prof. Seth Lloyd and Prof. Daniel J. Kleitman for serving on the thesis committee.

Over the years, many faculties contributed to my research, education and general well-being in the department, and I want to thank these professors. Among them are Prof. Rogers for working on 18.022 and RSI programs, Prof. Pak for passing my qualifying exam and research advising, Prof. Freedman for working on 18.03 and Prof. Hesselholt for working on 18.022. They are excellent researchers and also extraordinary folks to be acquainted with.

I would like to thank Joanne, Stevie and Debbie from the undergraduate math office for their helpfulness and friendliness. Thanks also go to Linda and Michelle from the graduate student office for their reminders and assistance in various administrative tasks.

I am grateful that during the past 5 years, I get to know many other talented graduate students and fabulous friends. Especially, Fu Liu, Thomas Lam, Huadong Pang and Joungkeun Lim. I thank them for encouragements and the joys and tears that we shared.

I would like to thank my parents and my sister Cary. Their undivided and unconditional supports have made the overseas study in the UK and the pursuit of an advanced degree in the US possible.

Finally, I want to dedicate the thesis to Vivian whose love has become my most important discovery, and the mystery and depth of which are $\infty^{\infty}$.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Preliminaries

## 1.1 Introduction

Quantum information theory studies the information processing power one can achieve by harnessing quantum mechanical principles[8, 35, 38]. Many important results such as quantum teleportation, superdense coding, factoring and search algorithms make use of quantum entanglements as fundamental resources[3, 11, 22, 41]. There is no complete theory to categorize and quantify the amount of entanglements in $N$ spin-$\frac{1}{2}$ particles in general. Among the prominent measures are entanglement cost, entanglement of formation, relative entropy of entanglement and distillable entanglement[14, 20, 21, 45].

In studying the entanglement-assisted capacities of a quantum discrete memoryless channel (QDMC) and the trade-offs between resources in attaining them, quantum entanglement is expressed in terms of an ebit - a pair of maximally entangled spin-$\frac{1}{2}$ particles - shared by the sender Alice and receiver Bob[9, 10, 18, 23, 40]. For example,

$$\begin{aligned}
00:&\quad |\Phi^+\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)\\
01:&\quad |\Psi^+\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\\
10:&\quad |\Phi^-\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle)\\
11:&\quad |\Psi^-\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \qquad (1.1)
\end{aligned}$$

are the so-called Bell basis and each of these states is considered equivalent to an ebit. However, these maximally-entangled pure states are only special cases of a general two-particle mixed state. In fact, any pure states, entangled or not, become mixed states once they are exposed to noise. Therefore, it is important to study procedures by which the sender and receiver can extract pure-state entanglement out of some shared mixed entangled states. These procedures are called entanglement purification protocols (EPP).

EPP are divided into 1-EPP and 2-EPP according to whether the sender and receiver are allowed to communicate uni- or bi-directionally. The scenario in which we study EPP is described as follows:

At the beginning of these entanglement purification protocols, Alice and Bob share a large number of the generalized Werner states[47]

$$\rho_F = F\,|\Phi^+\rangle\langle\Phi^+| + \frac{1-F}{3}\big(|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|\big), \qquad (1.2)$$

say $\rho_F^{\otimes N}$, and they are allowed to communicate classically, apply unitary transformations and perform projective measurements. We place no restriction on the size of their ancilla systems so that we lose no generality in restricting their local operations to unitaries and projective measurements. In the end, the quantum states $\Upsilon$ shared by Alice and Bob are to be a close approximation of the maximally entangled states $(|\Phi^+\rangle\langle\Phi^+|)^{\otimes M}$, or more precisely we require the fidelity between $\Upsilon$ and $(|\Phi^+\rangle\langle\Phi^+|)^{\otimes M}$ approaches one as $N$ goes to infinity. We define the yield of such protocols to be $M/N$.

There are two main reasons why this is considered a general scenario. The first reason is, by a preprocessing operation known as "twirl", any two-particle mixed state can be converted to a classical mixture of the four Bell basis states, and this alone is sufficient as an input state to all the protocols discussed in this thesis[7]. However, it is more convenient to equalize three of the four Bell states and prepare the input as the Werner state $\rho_F$, even though the equalization only adds unnecessary entropy to the mixture.

The second reason is the equivalence between an entanglement purification protocol on the Werner state $\rho_F$ and a protocol to faithfully transmit quantum states through the $\frac{(4F-1)}{3}$-depolarizing channel established in [7]. A $p$-depolarizing channel is a simple qubit channel such that a qubit passes through the channel undisturbed with probability $p$ and outputs as a completely random qubit with probability $1-p$. Specifically, the yield of a 1-EPP on the Werner state $\rho_F$ is equal to the unassisted quantum capacity of a $\frac{(4F-1)}{3}$-depolarizing channel ($Q$); and the yield of a 2-EPP on the Werner state $\rho_F$ is equal to the quantum capacity assisted by two-way classical communication of a $\frac{(4F-1)}{3}$-depolarizing channel ($Q_2$).

## 1.2 Elementary notions in quantum information theory

In this section, we review some elementary notions in quantum information theory. Not only do they provide background materials, but also introduce notations to facilitate discussion in this thesis. Most of these notations and discussions follow [30, 35].

### 1.2.1 Quantum states

As a postulate of quantum mechanics, one can associate any isolated physical system with a Hilbert space known as the state space of the system. The system is then completely characterized by a state vector. In quantum computation and quantum information theory, the conventional unit is a qubit - analogous to a classical bit in classical information theory [16] - and its state is represented by a unit vector in a two-dimensional Hilbert space, $\mathcal{H}_2$. Unlike its classical counterpart that has a state of either 0 or 1, however, a qubit can be a linear combination of the basis states $|0\rangle$ and $|1\rangle$. For example, $\alpha_0 |0\rangle + \alpha_1 |1\rangle \in \mathcal{H}_2$ where $\alpha_0$ and $\alpha_1$ are complex numbers. The basis states $|0\rangle$ and $|1\rangle$ are also known as the computational basis. In fact, we have already seen examples of a state vector in higher dimension. The state vector $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, one of the four Bell states in (1.1), is in a four-dimensional Hilbert space, $\mathcal{H}_4 \equiv \mathcal{H}_2 \otimes \mathcal{H}_2$ and thus describes the state of two qubits.

The density operator language is very useful in describing quantum states that are not completely known[30]. For example, we use the notation $\frac{1}{3} |0\rangle \langle 0| + \frac{2}{3} |1\rangle \langle 1|$ to describe a qubit that has a probability $\frac{1}{3}$ to be in the state $|0\rangle$ and a probability of $\frac{2}{3}$ to be in the state $|1\rangle$. In general, if a quantum system has a probability of $p_i$ to be in the state $|\psi_i\rangle$ for $i = 1, 2, \ldots, n$, then we say the system is an ensemble of pure states $\{p_i, |\psi_i\rangle\}_{i=1}^{n}$ and is described by the density operator

$$\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle \langle \psi_i| \in \mathcal{B}(\mathcal{H}_d) \tag{1.3}$$

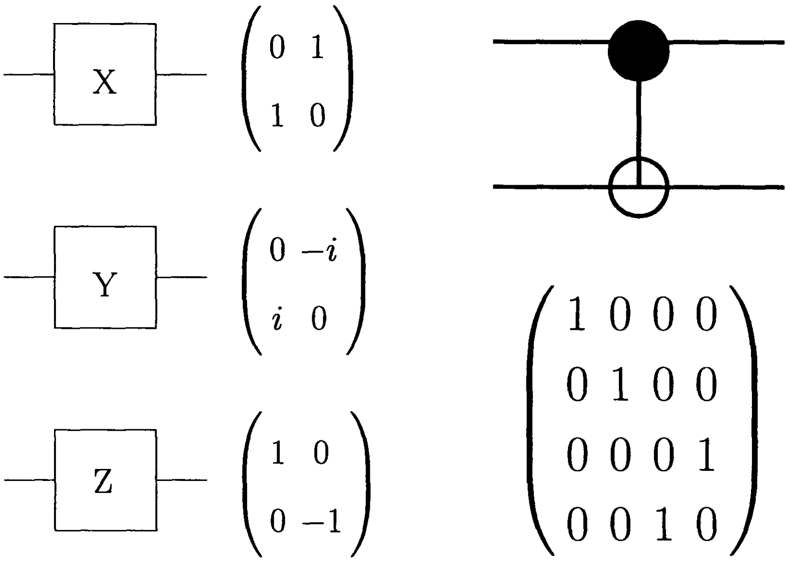
where $|\psi_i\rangle \in \mathcal{H}_d$ and $\mathcal{B}(\mathcal{H}_d)$ is the bounded algebra on a d-dimensional Hilbert space. Mathematically, any operator $\rho$ is the density operator associated to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if

1. $\rho$ has trace equal to one; and

2. $\rho$ is a positive operator.

### 1.2.2 Quantum gates

In this section, we introduce circuit notations for common quantum gates such as the single-qubit Pauli matrices and the two-qubit Controlled-NOT gate. We also illustrate the BXOR operation - bilateral application of Controlled-NOT on a bipartite quantum state - which will be used extensively. Note that all matrices in this section and throughout this thesis are expressed in the computational basis unless stated otherwise. In figure 1-1, we show examples of quantum gates that act on one or two qubits.


X
$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Y
$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$
Z
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

(a) Pauli matrices: $\sigma_x, \sigma_y, \sigma_z$.

(b) Controlled-NOT gate.

Figure 1-1: Single-qubit and two-qubit quantum gates.

When we study entanglement purification of bipartite quantum states, we will often use the BXOR operation. Suppose two persons whom we call Alice and Bob share the two bipartite states $|\Phi^+\rangle$ and $|\Phi^-\rangle$, we say Alice and Bob apply BXOR on $|\Phi^+\rangle$ and $|\Phi^-\rangle$ and that $|\Phi^+\rangle$ is the 'source' and $|\Phi^-\rangle$ is the 'target' when the scenario in figure 1-2 occurs. In table 1.1, we list all possibilities of applying BXOR to the four Bell states in (1.1) as these will be useful in the discussion of entanglement purification protocols.

### 1.2.3 Quantum measurements

A projective measurement, a special case of what is known as POVW measurements, is represented by a Hermitian operator $M$ on the state space of the system on which we

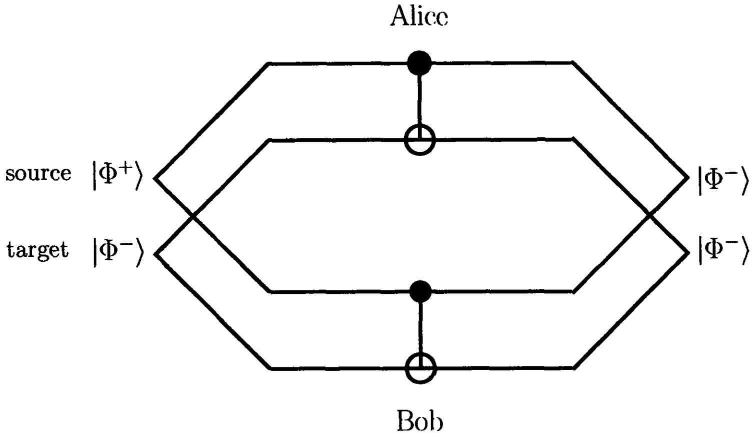


Figure 1-2: BXOR operation.

| input | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | |
|---|---|---|---|---|---|
| $\|\Phi^+\rangle$ | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | (source) |
| | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | (target) |
| $\|\Psi^+\rangle$ | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | (source) |
| | $\|\Psi^+\rangle$ | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | $\|\Phi^+\rangle$ | (target) |
| $\|\Phi^-\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | (source) |
| | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | (target) |
| $\|\Psi^-\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | $\|\Phi^+\rangle$ | $\|\Psi^+\rangle$ | (source) |
| | $\|\Psi^-\rangle$ | $\|\Phi^-\rangle$ | $\|\Psi^-\rangle$ | $\|\Phi^-\rangle$ | (target) |

Table 1.1: Outputs of the BXOR operations for Bell states 'source' (the top row) and 'target' (the leftmost column) inputs.

would like to take a measurement. This Hermitian operator, also called an observable in this case, has a spectral decomposition,

$$M = \sum_m m P_m,$$

where $P_m$ is the projector onto the eigenspace of $M$ with eigenvalue $m$. The different values of $m$ are the possible measurement outcomes. If the quantum state is represented by a state vector $|\psi\rangle$, then the probability of getting the measurement result $m$ is given by

$$\text{prob}(m) = \langle\psi| P_m |\psi\rangle .$$

Given that the measurement outcome is $m$, the quantum system immediately after the measurement is described by

$$\frac{P_m |\psi\rangle}{\sqrt{\text{prob}(m)}}.$$

For example, when the quantum state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ is measured by the observable $\sigma_z = (+1)|0\rangle\langle 0| + (-1)|1\rangle\langle 1|$, there is a probability 0.5 that the outcome is +1 and the quantum state is $|0\rangle$; and a probability 0.5 that the outcome is $-1$ and the quantum state is $|1\rangle$. When the observable is $\sigma_x$ (or respectively $\sigma_z$), we say we measure a quantum state along the $x$-axis (or respectively $z$-axis). In section 1.2.1, we learned that a mixed quantum state can be conveniently represented by a density operator $\rho$. Then the probability of getting measurement result $m$ is given by

$$\text{prob}(m) = tr(P_m^\dagger P_m \rho)$$

and the quantum system immediately after the measurement is described by

$$\frac{P_m \rho P_m^\dagger}{tr(P_m^\dagger P_m \rho)}.$$

### 1.2.4 Quantum discrete memoryless channel (QDMC) and its various capacities

Discrete memoryless channel (DMC) can be described by a probability transition matrix and its capacity is uniquely defined[16, 38]. Quantum discrete memoryless channel (QDMC) can be described in many different ways and has various well-defined capacities depending on the availability of auxiliary resources such as classical communication or shared entanglements.

Mathematically, QDMC can be defined as a trace-preserving, completely positive linear map from the bounded algebra of an input Hilbert space to the bounded algebra of an output Hilbert space,

$$\mathcal{N} : \mathcal{B}(\mathcal{H}_{d_1}) \longrightarrow \mathcal{B}(\mathcal{H}_{d_2})$$

and any such map $\mathcal{N}$ can be given an operator-sum representation [30, 35] which we state as follows,

$$\mathcal{N}(\rho) = \sum_j E_j \rho E_j^\dagger \tag{1.4}$$

where $\{E_j\}$ is a set of linear operators which map the input Hilbert space $\mathcal{H}_{d_1}$ to the output Hilbert space $\mathcal{H}_{d_2}$ and $\sum_j E_j^\dagger E_j = I$. Hence if we represent a general input state as a density operator $\big($c.f. equation (1.3)$\big)$, the output state is

$$\rho = \sum_{i=1}^{n} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| \quad \mapsto \quad \rho' = \sum_{i,j} p_i E_j \left|\psi_i\right\rangle \left\langle\psi_i\right| E_j^\dagger. \tag{1.5}$$

The QDMC we study in this thesis is the quantum depolarizing channel. A $p$-depolarizing channel $\mathcal{E}_p : \mathcal{B}(\mathcal{H}_2) \longrightarrow \mathcal{B}(\mathcal{H}_2)$ has the following set of linear operator elements:

$$\{E_0 = \sqrt{\frac{1+3p}{4}} I, E_1 = \sqrt{\frac{1-p}{4}} \sigma_x, E_2 = \sqrt{\frac{1-p}{4}} \sigma_y, E_3 = \sqrt{\frac{1-p}{4}} \sigma_z\}.$$

Simple algebra shows that

$$\mathcal{E}_p(\rho) = \sum_{j=0}^{3} E_j \rho E_j^\dagger = p \times \rho + (1-p) \times \frac{I}{2},$$

i.e. with probability $p$ the quantum state passes the channel unaffected and with probability $1-p$ the quantum state is replaced by a completely random state $\frac{I}{2}$.

While the capacity of a DMC is given by a single numerical value representing the amount of information that can be transmitted asymptotically without error per channel use and that this value is unaffected by the use of classical feedback, for quantum discrete memoryless channels, the analogous statements are not true. Capacities are affected by side classical communication and shared entanglement[5, 9]; and QDMC can be used to transmit either classical or quantum information.

We can then define, for every quantum discrete memoryless channel, various capacities: $C$, unassisted classical capacity; $C_B$, classical capacity assisted by classical feedback; $C_2$, classical capacity assisted by independent classical information; $C_E$, entanglement-assisted classical capacity; $Q$, unassisted quantum capacity; $Q_B$, quantum capacity assisted by classical feedback; $Q_2$, quantum capacity assisted by independent classical information; and finally $Q_E$, entanglement-assisted quantum capacity. So far, some progress has been made to compute the capacities for specific channels[6, 9, 27] and to study their relations[5]. However, search for a general formula only succeeded in a few cases[10, 17, 24, 37, 40], and progress in this direction has been hindered by the additivity conjecture[39]. In particular, the following capacities (of the quantum depolarizing channel) will be studied in this thesis,

- $C$: the rate at which the sender can transmit classical information to the receiver asymptotically without error;

- $Q$: the rate at which the sender can transmit quantum information to the receiver asymptotically without error;

- $Q_B$: the rate at which the sender can transmit quantum information to the receiver asymptotically without error when a classical communication channel from Bob to Alice is available; and

- $Q_2$: the rate at which the sender can transmit quantum information to the receiver asymptotically without error when a bidirectional classical communication channel between Alice and Bob is available.

Whilst the first two capacities can easily be described mathematically, the last two capacities cannot because the protocols to achieve the capacities may be interactive, i.e. Alice and Bob can communicate classically after each channel use. In this thesis, we will improve the lower bounds of the last two capacities for a p-depolarizing channel.

## 1.3 Previous works

In this section, we review the best known entanglement purification protocols, namely the universal hashing, the recurrence method and the Maneva-Smolin method. Universal hashing is a 1-EPP and the last two methods are 2-EPP. As aforementioned, it is known that the yield of any 1-EPP on the Werner state $\rho_F$ is the same as the unassisted quantum capacity of a $\frac{(4F-1)}{3}$-depolarizing channel $\big(Q(\mathcal{E}_p)\big)$; and the yield of any 2-EPP on the Werner state $\rho_F$ is the same as the quantum capacity assisted by two-way classical communication of a $\frac{(4F-1)}{3}$-depolarizing channel $\big(Q_2(\mathcal{E}_p)\big)$.

### 1.3.1 Universal hashing

Universal hashing, introduced in [7], requires only one-way classical communication and hence is a 1-EPP. The hashing method works by having Alice and Bob each perform some local unitary operations on the corresponding members of the shared bipartite quantum states. They then locally measure some of the pairs to gain classical information about the identities of the the remaining unmeasured pairs. It was shown that each measurement can be made to reveal almost 1 bit of information about the unmeasured Bell states pairs. Since the information associated with a quantum state $\rho_F$ is given by its von Neumann entropy $S(\rho_F)$, we know from typical subspace argument that, with probability approaching 1 and by measuring $NS(\rho_F)$ pairs, Alice and Bob can figure out the identities of all pairs including the unmeasured ones. Once the identities of the Bell states are known, Alice and Bob can convert them into the standard states $\Phi^+$ easily. Therefore this protocol distills a yield of $\Big(N - NS(\rho_F)\Big)/N = 1 - S(\rho_F)$.

### 1.3.2 The recurrence method and the modified recurrence method

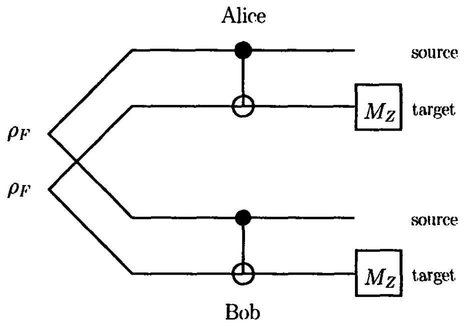


Figure 1-3: The recurrence method.

The recurrence method[4, 7] is illustrated in figure 1-3. Alice and Bob put the quantum states $\rho_F^{\otimes N}$ into groups of two and apply XOR operations to the corresponding members of the quantum states $\rho_F^{\otimes 2}$, one as the source and one as the target. They then take projective measurements on the target states along the z-axis, and compare their measurement results with the side classical communication channel. If they get identical results, the source pair "passed"; otherwise the source pair "failed". Alice and Bob then collect all the "passed" pairs, and apply a unilateral $\pi$ rotation $\sigma_x$ followed by a bilateral $\pi/2$ rotation $B_x$[1]. This process is iterated until it becomes more beneficial to pass on to the universal hashing. If we denote the quantum states by $\rho = p_{00} |\Phi^+\rangle \langle\Phi^+| + p_{01} |\Psi^+\rangle \langle\Psi^+| + p_{10} |\Phi^-\rangle \langle\Phi^-| + p_{11} |\Psi^-\rangle \langle\Psi^-|$, then this protocol has the following recurrence relation:

$$\begin{aligned} p'_{00} &= (p_{00}^2 + p_{10}^2)/p_{pass}; \; p'_{01} = (p_{01}^2 + p_{11}^2)/p_{pass}; \\ p'_{10} &= 2p_{01}p_{11}/p_{pass}; \quad p'_{11} = 2p_{00}p_{10}/p_{pass}; \end{aligned} \tag{1.6}$$

and

[1] As mentioned in [7], the application of a $\sigma_x$ and $B_x$ rather than a twirl was proposed by C. Macchiavello. This is known as the modified recurrence method

$$p_{pass} = p_{00}^2 + p_{01}^2 + p_{10}^2 + p_{11}^2 + 2p_{00}p_{10} + 2p_{01}p_{11}. \tag{1.7}$$

### 1.3.3 The Maneva-Smolin method

The Maneva-Smolin method[34] is illustrated in figure 1-4. Alice and Bob first choose a block size $m$ and put the quantum states into groups of $m$. They then apply bipartite XOR gates between each of the first $m-1$ pairs and the $m$th pairs. After that, they take measurements on these $m$th pairs along the z-axis, and compare their results with side classical communication channel. If they get identical results, they perform universal hashing on the corresponding $m-1$ remaining pairs; if they get different results, they throw away all m pairs. The yield for this method is:

$$p_{pass}\frac{m-1}{m}\left(1-\frac{H(\text{passed source states})}{m-1}\right)$$

where $H(P)$ is the Shannon entropy of the probability distribution $P$.

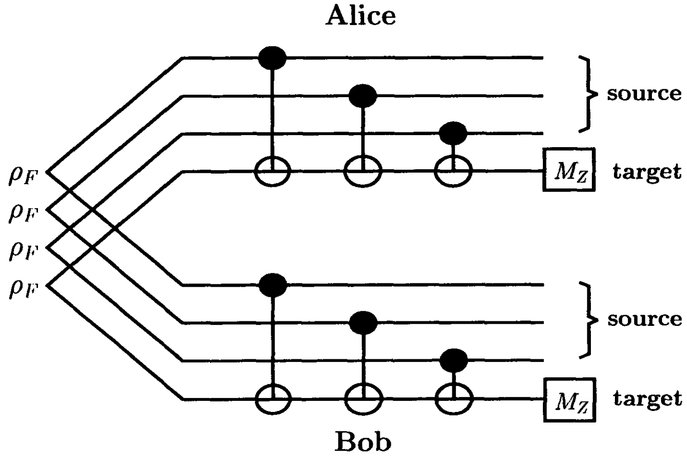


Figure 1-4: The Maneva-Smolin method when $m = 4$.

# Chapter 2

# Adaptive entanglement purification protocols

In this chapter, we study 2-EPP, entanglement purification protocols when the two parties, Alice and Bob, are allowed to communicate classically. In section 2.1, we introduce a new 2-EPP [32]. We compute its yield for the Werner state $\rho_F$ and compare with the 2-EPP introduced in section 1.3. We also give a closed-form expression for general Bell-diagonal input states. In section 2.2, we present a family of 2-EPP that generalizes the previous methods in section 1.3 and the method in section 2.1. We show this family of protocols have improved yields over a wide range of initial fidelities $F$. In particular, the yield of this family of protocols on the Werner state $\rho_F$ is higher than that of universal hashing for $F < 0.993$ and as $F \longrightarrow \infty$. In section 2.3, we established the '$F \longrightarrow \infty$' part of the previous statement analytically. In section 2.4, we modify the family of protocols to achieve higher yields. Finally, we discuss the results of 2-EPP, some recent progresses[25, 46] and directions for further research.

Our protocols work for any Bell-diagonal states, and we adopt the 2-bit representation of the Bell states in (1.1). As a result, the Werner state $\rho_F$ is simply a probability distribution over the four Bell states, $00, 01, 10$ and $11$. Similarly, when Alice and Bob share $N$ bipartite states that are Bell diagonal, a probability distribution over a binary string of length $2N$ provides a complete description.

## 2.1 The Leung-Shor method

In section 2.1.1, we present a new entanglement purification protocol and compare its yield with the yields obtained by the recurrence method [7] and the Maneva-Smolin method [34]. In section 2.1.2, we give a closed-form expression for the yield of this new protocol.

### 2.1.1 New protocol with improved yield

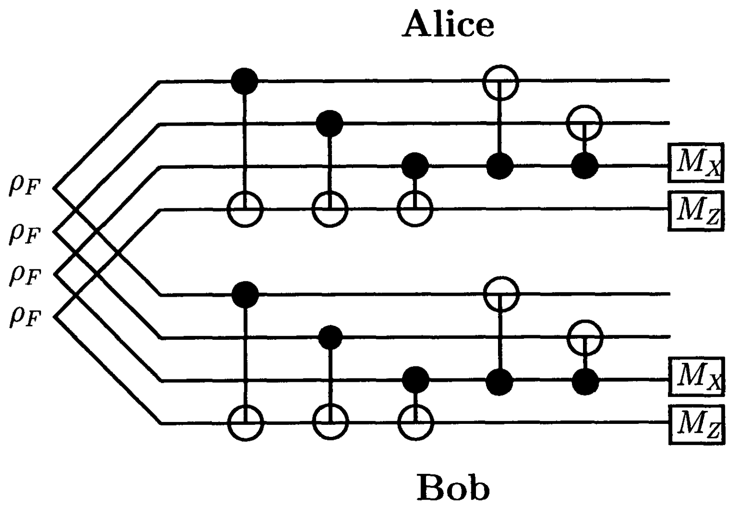


Figure 2-1: The Leung-Shor method. The yield of this protocol is shown in figure 2-2.

The new protocol is illustrated in figure 2-1. Alice and Bob share the quantum states $\rho_F^{\otimes N}$ and put them into groups of four. They then apply the quantum circuit shown in figure 2-1 and take measurements on the third and fourth pairs along the x- and z-axis respectively. Using the side classical communication channel, they can compare their results with each other. If they get identical results on both measurements, they keep the first and second pairs and apply universal hashing[7]. If either of the two results disagrees, they throw away all four pairs.

The four pairs can be described by an 8-bit binary string, and since these are mixed states they are in fact probability distribution over all 256(= $2^8$) possible 8-bit binary strings. The quantum circuit consists only of XOR gates and therefore maps the

8-bit binary strings, along with their underlying probability distribution, bijectively to themselves. Let us call these probability distributions $P(a_1a_2b_1b_2c_1c_2d_1d_2)$ and $P'(a_1a_2b_1b_2c_1c_2d_1d_2)$.

The quantum measurements on the third and fourth pairs are simply checking the 5th bit (measurement on the third pair along x-axis) and the 8th bit (measurement on the fourth pair along z-axis), where a "0" means Alice and Bob getting identical results and a "1" means their getting opposite results. For example, if the 8-bit binary string is "$a_1a_2b_1b_2c_1c_2d_1d_2 = 00100111$", which corresponds to the quantum states $\Phi^+\Phi^-\Psi^+\Psi^-$, then Alice and Bob will get identical results on the third pair but opposite results on the fourth. The "pass" probability is $p_{pass} = \sum_{a_1,a_2,b_1,b_2,c_2,d_1\in\{0,1\}} P'(a_1a_2b_1b_20c_2d_10)$ and the post-measurement probability distribution is $Q(a_1a_2b_1b_2) = \sum_{c_2,d_1\in\{0,1\}} P'(a_1a_2b_1b_20c_2d_10)/p_{pass}$. The yield of this method[34] is:

$$\frac{p_{pass}}{2}\left(1-\frac{H(Q(a_1a_2b_1b_2))}{2}\right) \tag{2.1}$$

where $H(Q(a_1a_2b_1b_2))$ is the Shannon entropy function. Figure 2-2 compares the yield of our new method with the recurrence method and the Maneva-Smoline method.

### 2.1.2 Closed-form expression

The quantum circuit that Alice and Bob apply to the quantum states $\rho_F^{\otimes 4}$ consists only of XOR gates and therefore maps the 8-bit binary strings bijectively to themselves. Let us call this bijection $f$:

$$\begin{aligned} f : \{0,1\}^8 &\longrightarrow \{0,1\}^8 \\ (a_1,a_2,b_1,b_2,c_1,c_2,d_1,d_2) &\longmapsto (a_1\oplus d_1, a_2\oplus c_2, b_1\oplus d_1, b_2\oplus c_2, \\ &\qquad a_1\oplus b_1\oplus c_1\oplus d_1, c_2, d_1, a_2\oplus b_2\oplus c_2\oplus d_2) \end{aligned}$$

Table 2.1: The quantum states that lead to identical results for Alice and Bob. $G = (1-F)/3; \Phi^+ = 00; \Psi^+ = 01; \Phi^- = 10; \Psi^- = 11.$

| $P(a_1a_2b_1b_2c_1c_2d_1d_2)$ | $a_1a_2b_1b_2c_1c_2d_1d_2$ | $f(a_1a_2b_1b_2c_1c_2d_1d_2)$ | $tr_{c,d}(f(a_1a_2b_1b_2c_1c_2d_1d_2))$ |
|---|---|---|---|
| $F^4$ | 00000000 | 00000000 | 0000 |
| $G^4$ | 01010101 | 00000100 | 0000 |
| $G^4$ | 10101010 | 00000010 | 0000 |
| $G^4$ | 11111111 | 00000110 | 0000 |
| $F^2G^2$ | 00010001 | 00010000 | 0001 |
| $F^2G^2$ | 01000100 | 00010100 | 0001 |
| $G^4$ | 10111011 | 00010010 | 0001 |
| $G^4$ | 11101110 | 00010110 | 0001 |
| $F^2G^2$ | 00101000 | 00100000 | 0010 |
| $G^4$ | 01111101 | 00100100 | 0010 |
| $F^2G^2$ | 10000010 | 00100010 | 0010 |
| $G^4$ | 11010111 | 00100110 | 0010 |
| $FG^3$ | 00111001 | 00110000 | 0011 |
| $FG^3$ | 01101100 | 00110100 | 0011 |
| $FG^3$ | 10010011 | 00110010 | 0011 |
| $FG^3$ | 11000110 | 00110110 | 0011 |
| $F^2G^2$ | 00010100 | 01000100 | 0100 |
| $F^2G^2$ | 01000001 | 01000000 | 0100 |
| $G^4$ | 10111110 | 01000110 | 0100 |
| $G^4$ | 11101011 | 01000010 | 0100 |
| $F^2G^2$ | 00000101 | 01010100 | 0101 |
| $F^2G^2$ | 01010000 | 01010000 | 0101 |
| $G^4$ | 10101111 | 01010110 | 0101 |
| $G^4$ | 11111010 | 01010010 | 0101 |
| $F^2G^2$ | 00111100 | 01100100 | 0110 |
| $G^4$ | 01101001 | 01100000 | 0110 |
| $G^4$ | 10010110 | 01100110 | 0110 |
| $F^2G^2$ | 11000011 | 01100010 | 0110 |
| $FG^3$ | 00101101 | 01110100 | 0111 |
| $FG^3$ | 01111000 | 01110000 | 0111 |
| $FG^3$ | 10000111 | 01110110 | 0111 |
| $FG^3$ | 11010010 | 01110010 | 0111 |
| $F^2G^2$ | 00100010 | 10000010 | 1000 |
| $G^4$ | 01110111 | 10000110 | 1000 |
| $F^2G^2$ | 10001000 | 10000000 | 1000 |
| $G^4$ | 11011101 | 10000100 | 1000 |
| $F^2G^2$ | 00110011 | 10010010 | 1001 |
| $G^4$ | 01100110 | 10010110 | 1001 |
| $G^4$ | 10011001 | 10010000 | 1001 |
| $F^2G^2$ | 11001100 | 10010100 | 1001 |
| $F^2G^2$ | 00001010 | 10100010 | 1010 |
| $G^4$ | 01011111 | 10100110 | 1010 |
| $F^2G^2$ | 10100000 | 10100000 | 1010 |
| $G^4$ | 11110101 | 10100100 | 1010 |
| $FG^3$ | 00011011 | 10110010 | 1011 |
| $FG^3$ | 01001110 | 10110110 | 1011 |
| $FG^3$ | 10110001 | 10110000 | 1011 |
| $FG^3$ | 11100100 | 10110100 | 1011 |
| $FG^3$ | 00110110 | 11000110 | 1100 |
| $FG^3$ | 01100011 | 11000010 | 1100 |
| $FG^3$ | 10011100 | 11000100 | 1100 |
| $FG^3$ | 11001001 | 11000000 | 1100 |
| $FG^3$ | 00100111 | 11010110 | 1101 |
| $FG^3$ | 01110010 | 11010010 | 1101 |
| $FG^3$ | 10001101 | 11010100 | 1101 |
| $FG^3$ | 11011000 | 11010000 | 1101 |
| $FG^3$ | 00011110 | 11100110 | 1110 |
| $FG^3$ | 01001011 | 11100010 | 1110 |
| $FG^3$ | 10110100 | 11100100 | 1110 |
| $FG^3$ | 11100001 | 11100000 | 1110 |
| $F^2G^2$ | 00001111 | 11110110 | 1111 |
| $G^4$ | 01011010 | 11110010 | 1111 |
| $G^4$ | 10100101 | 11110100 | 1111 |
| $F^2G^2$ | 11110000 | 11110000 | 1111 |

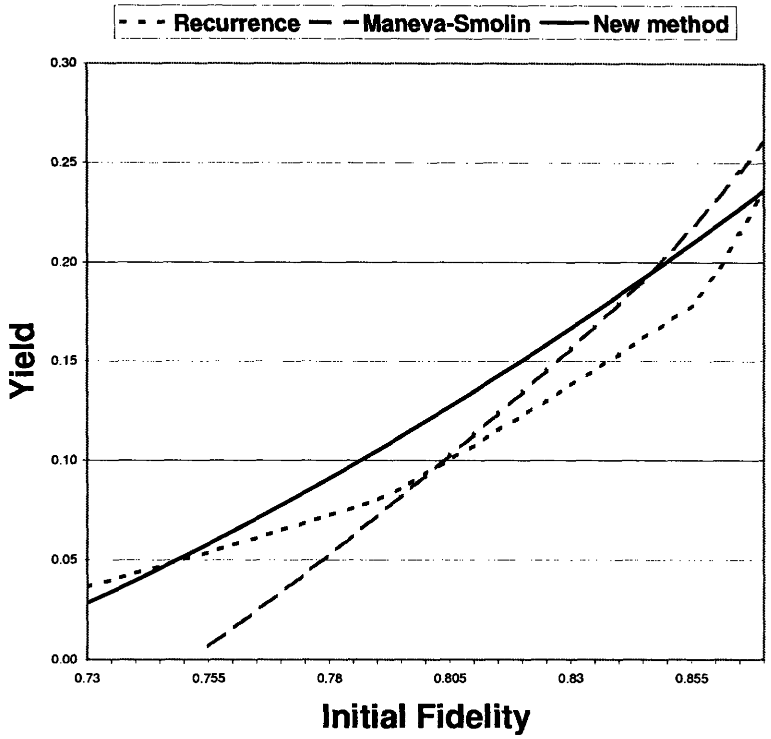


Figure 2-2: The dotted line is the yield for modified recurrence method [7]; the dash line is for the Maneva-Smolin method [34]. The yield of our new method is represented by the solid line, and there is an improvement over the previous methods when the initial fidelity is between 7.5 and 8.45.

In table 2.1, we list the quantum states that lead to identical measurement results for both measurements and the associated probabilities in the ensemble $\rho_F^{\otimes 4}$. From that, we can write down expressions for $p_{pass}$ and $H(Q(a_1a_2b_1b_2)$ in equation(2.1) as follows:

$$\begin{aligned} H(Q(a_1a_2b_1b_2)) &= -\Big(\frac{F^4+3G^4}{p_{pass}}\Big)\log_2\Big(\frac{F^4+3G^4}{p_{pass}}\Big) - 9\Big(\frac{2F^2G^2+2G^4}{p_{pass}}\Big)\log_2\Big(\frac{2F^2G^2+2G^4}{p_{pass}}\Big) \\ &\quad -6\Big(\frac{4FG^3}{p_{pass}}\Big)\log_2\Big(\frac{4FG^3}{p_{pass}}\Big) \qquad (2.2) \\ p_{pass} &= F^4+18F^2G^2+24FG^3+21G^4 \qquad (2.3) \end{aligned}$$

where $G=(1-F)/3$. So far, we have applied the new protocol to the quantum states

$$\rho_F = F\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \frac{1-F}{3}\Big(\left|\Phi^-\right\rangle\left\langle\Phi^-\right| + \left|\Psi^+\right\rangle\left\langle\Psi^+\right| + \left|\Psi^-\right\rangle\left\langle\Psi^-\right|\Big);$$

however, our method works for any Bell-diagonal states

$$\rho = p_{00}\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + p_{01}\left|\Psi^+\right\rangle\left\langle\Psi^+\right| + p_{10}\left|\Phi^-\right\rangle\left\langle\Phi^-\right| + p_{11}\left|\Psi^-\right\rangle\left\langle\Psi^-\right|.$$

Equation (2.2) and (2.3) then become

$$\begin{aligned}
H(Q(a_1a_2b_1b_2)) = & -\left(\frac{p_{00}^4+p_{01}^4+p_{10}^4+p_{11}^4}{p_{pass}}\right)\log_2\left(\frac{p_{00}^4+p_{01}^4+p_{10}^4+p_{11}^4}{p_{pass}}\right) \\
& -6\times\left(\frac{4p_{00}p_{01}p_{10}p_{11}}{p_{pass}}\right)\log_2\left(\frac{4p_{00}p_{01}p_{10}p_{11}}{p_{pass}}\right) \\
& -3\times\left(\frac{2p_{00}^2p_{01}^2+2p_{10}^2p_{11}^2}{p_{pass}}\right)\log_2\left(\frac{2p_{00}^2p_{01}^2+2p_{10}^2p_{11}^2}{p_{pass}}\right) \\
& -3\times\left(\frac{2p_{00}^2p_{10}^2+2p_{01}^2p_{11}^2}{p_{pass}}\right)\log_2\left(\frac{2p_{00}^2p_{10}^2+2p_{01}^2p_{11}^2}{p_{pass}}\right) \\
& -3\times\left(\frac{2p_{00}^2p_{11}^2+2p_{01}^2p_{10}^2}{p_{pass}}\right)\log_2\left(\frac{2p_{00}^2p_{11}^2+2p_{01}^2p_{10}^2}{p_{pass}}\right) \\
p_{pass} = & \left(p_{00}^4+p_{01}^4+p_{10}^4+p_{11}^4\right)+6\times 4p_{00}p_{01}p_{10}p_{11}+3\times\sum_{\substack{i,j\in\{0,1\}^2\\ i\neq j}} 2p_i^2p_j^2.
\end{aligned}$$

With these equations, we can combine the recurrence method and our new method: we start with the recurrence method and pass on to our new method rather than universal hashing. Indeed, there are improvements, but they occur over segments of narrow regions and the improvements are insignificant. Therefore we believe these improvements have only to do with the number of recurrence steps performed before passing on to universal hashing, and we will spare the readers with the details.

## 2.2 Adaptive Entanglement Purification Protocols (AEPP)

In this section, we will present a family of entanglement purification protocols that generalize four previous methods, namely the recurrence method, the modified recurrence method, and the two methods proposed by Maneva-Smolin and Leung-Shor. We will show that this family of protocols have improved yields over a wide range of initial fidelities F, and hence imply new lower bounds on the quantum capacity assisted by two-way classical communication of the quantum depolarizing channel. In particular, the yields of these protocols are higher than the yield of universal hashing for F less than 0.993 and as F goes to 1.

The sender Alice and receiver Bob will often apply the BXOR operation on two of their N bipartite quantum states. These N states are mixtures of Bell diagonal states and can be represented by a probability distribution over a string of '0' and '1' of length 2N. Using the two classical bit notations, we write

$$\begin{array}{rcl} BXOR(i,j) : \{0,1\}^{2N} & \rightarrow & \{0,1\}^{2N} \\ (a_i, b_i) & \mapsto & (a_i \oplus a_j, b_i) \\ (a_j, b_j) & \mapsto & (a_j, b_i \oplus b_j) \\ (a_k, b_k) & \mapsto & (a_k, b_k) \text{ if } k \neq i, j \end{array}$$

when Alice and Bob share apply BXOR to the ith pair as 'source' and the jth pair as 'target'.

### 2.2.1 Description of AEPP

1. **AEPP(a,2):** Alice and Bob put the bipartite quantum states $\rho_F^{\otimes N}$ into groups of two, apply BXOR(1,2)

$$(a_1, b_1, a_2, b_2) \mapsto (a_1 \oplus a_2, b_1, a_2, b_1 \oplus b_2)$$

and take projective measurements on the second pair along the z-axis. Using two-way classical communication channel, they can compare their measurement results. If the measurement results agree($b_1 \oplus b_2 = 0$), then it is likely that there has been no amplitude error and Alice and Bob will perform universal hashing on the first pair; if the results disagree($b_1 \oplus b_2 = 1$), they throw away the first pair because it is likely that an amplitude error has occurred. We give a graphical representation of this protocol in figure 2-3.

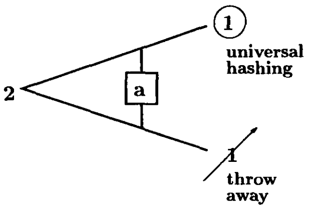


Figure 2-3: AEPP(a,2).

**2. AEPP(a,4):** Alice and Bob put the bipartite quantum states $\rho_F^{\otimes N}$ into groups of four, apply BXOR(1,4), BXOR(2,4), BXOR(3,4)

$$(a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4) \mapsto$$
$$(a_1 \oplus a_4, b_1, a_2 \oplus a_4, b_2, a_3 \oplus a_4, b_3, a_4, b_1 \oplus b_2 \oplus b_3 \oplus b_4)$$

and take projective measurements on the fourth pair along the z-axis. Using two-way classical communication channel, they can compare their measurement results. If the measurement results agree($b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 0$), then it is likely that there has been no amplitude error and Alice and Bob will perform universal hashing on the first three pairs together.

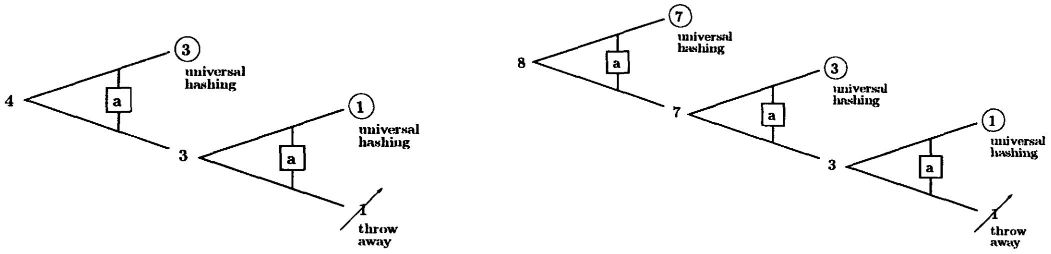
On the other hand, if the results disagree($b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 1$), it is likely that there is one amplitude error and Alice and Bob want to locate this amplitude error.

They do so by applying BXOR(2,1)

$$\begin{aligned}(a_1 \oplus a_4, b_1, a_2 \oplus a_4, b_2, a_3 \oplus a_4, b_3, a_4, 1) &\mapsto \\ (a_1 \oplus a_4, b_1 \oplus b_2, a_1 \oplus a_2, b_2, a_3 \oplus a_4, b_3, a_4, 1)\end{aligned} \tag{2.4}$$

and taking projective measurements on the first pair along the z-axis. Note that the second pair$(a_1 \oplus a_2, b_2)$ and the third pair$(a_3 \oplus a_4, b_3)$ are no longer entangled. Alice and Bob then use classical communication channel to compare their results. If the results agree$(b_1 \oplus b_2 = 0)$, then the amplitude error detected by the first measurements is more likely to be on either the third or the fourth pair than on the first two. Therefore Alice and Bob perform universal hashing on the second pair and throw away the third pair. If the results disagree$(b_1 \oplus b_2 = 1)$, then the amplitude error is more likely to be on the first two pairs. In this case, Alice and Bob perform universal hashing on the third pair and throw away the second pair.

Note that the amplitude error could have been on the fourth pair but this protocol works well even if that is the case; and also that with this procedure we always end up with one pair on which Alice and Bob can perform universal hashing when the first measurement results disagree. We represent this protocol graphically in figure 2-4(a).



(a) AEPP(a,4)

(b) AEPP(a,8)

Figure 2-4: AEPP(a,4) and AEPP(a,8).

**3. AEPP(a,8):** Alice and Bob put the bipartite quantum states $\rho_F^{\otimes N}$ into groups of eight, apply BXOR(1,8), BXOR(2,8), BXOR(3,8), ..., BXOR(7,8)

$$(a_1, b_1, a_2, b_2, \ldots, a_7, b_7, a_8, b_8) \mapsto$$
$$(a_1 \oplus a_8, b_1, a_2 \oplus a_8, b_2, \ldots, a_7 \oplus a_8, b_7, a_8, b_1 \oplus \ldots \oplus b_8)$$

and take projective measurements on the eighth pair along the z-axis. Using classical communication channel, Alice and Bob compare their measurement results. If the results agree($b_1 \oplus \ldots \oplus b_8 = 0$), then an amplitude error is not likely and they perform universal hashing on the first seven pairs together.

On the other hand, if the measurement results disagree($b_1 \oplus \ldots \oplus b_8 = 1$), then Alice and Bob want to catch this amplitude error and they do that by applying BXOR(2,1), BXOR(3,1), BXOR(4,1)

$$(a_1 \oplus a_8, b_1, a_2 \oplus a_8, b_2, \ldots, a_7 \oplus a_8, b_7, a_8, 1) \mapsto$$
$$(a_1 \oplus a_8, b_1 \oplus b_2 \oplus b_3 \oplus b_4, a_1 \oplus a_2, b_2, a_1 \oplus a_3, b_3,$$
$$a_1 \oplus a_4, b_4, a_5 \oplus a_8, b_5, a_6 \oplus a_8, b_6, a_7 \oplus a_8, b_7, a_8, 1)$$

and taking projective measurements on the first pair along the z-axis. Note that the second, third and fourth pairs are not entangled with the fifth, sixth and seventh pairs. After Alice and Bob compare their results with classical communication channel and if the results disagree ($b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 1$), they perform universal hashing on the fifth, sixth and seventh pairs because $b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 1$ and $b_1 \oplus \ldots \oplus b_8 = 1$ together imply $b_5 \oplus b_6 \oplus b_7 \oplus b_8 = 0$. The first four pairs are now represented by $(a_1 \oplus a_8, 1, a_1 \oplus a_2, b_2, a_1 \oplus a_3, b_3, a_1 \oplus a_4, b_4)$, and it can be easily seen that we are in the same situation as the left hand side of equation (2.4): Alice and Bob know that $b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 1$ and the pair on which they measured to find out this information has its phase error added to the other three pairs. Therefore Alice and Bob can apply the same procedure as equation (2.4) and end up with one pair that they will perform universal hashing on. Now if the results actually agree ($b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 0$), the same procedure still applies but we need to switch the roles played by the first four pairs and by the last four pair. We represent this protocol graphically in figure 2-4(b).

**4. AEPP(a,N=$2^n$) and AEPP(p,N=$2^n$):** Clearly, the above procedures generalize to AEPP(a,N=$2^n$) and can be proved inductively. The procedures - AEPP(a,N=$2^n$) - we discussed so far focus on amplitude error. If we instead try to detect phase error by switching the source pairs and target pairs in all the BXOR operations and measuring along the x-axis rather than the z-axis, AEPP(p,N=$2^n$) can be defined analogously. We represent the protocols AEPP(p,N=$2^n$) graphically in figure 2-5, and we present the yields of AEPP(a,$N = 2^n$) for $n = 2, 3, 4, 5, 6$ in figure 2-6.

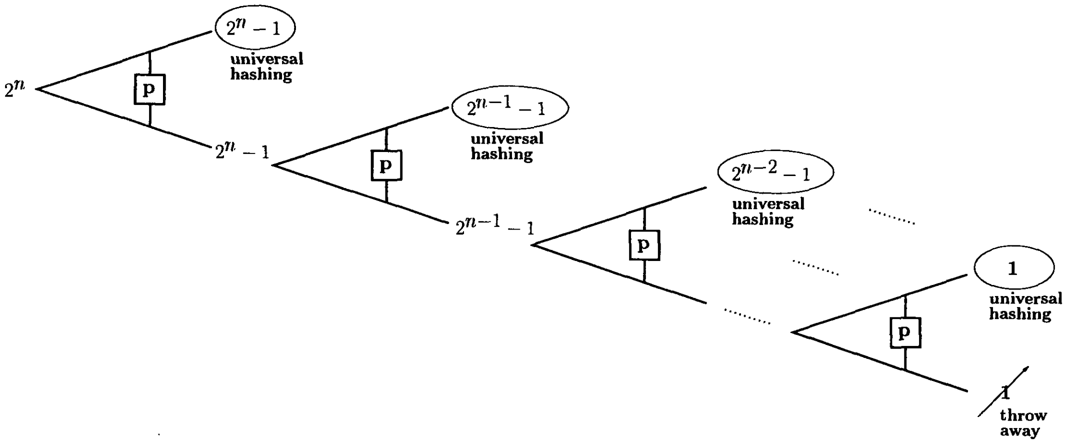


Figure 2-5: AEPP(p,N=$2^n$).

### 2.2.2 Generalization of previous methods

We show that four previous protocols - the recurrence method, the modified recurrence method and the two methods proposed by Maneva-Smolin and Leung-Shor - all belong to the family AEPP(a/p,$N = 2^n$).

**1. The recurrence method:** The recurrence method[7] is the repeat applications of AEPP(a,2). When Alice and Bob have identical measurement results, rather than applying universal hashing right away, they repeatedly apply AEPP(a,2) until it is more beneficial to switch to hashing.

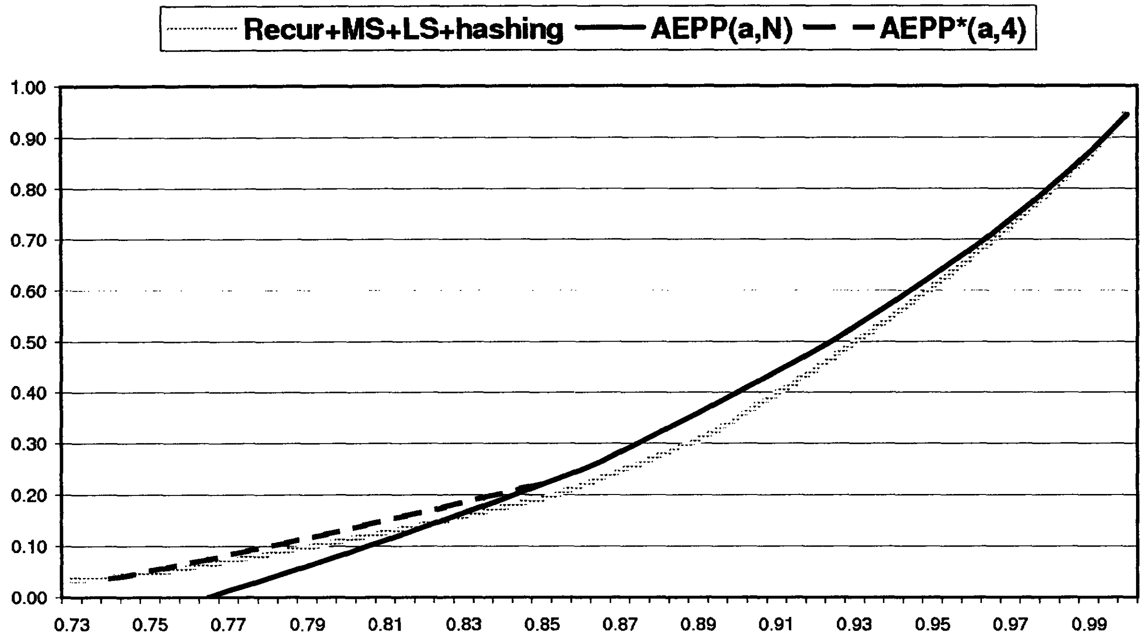


Figure 2-6: Comparison of AEPP and previous methods: The lightly colored line is the yield of the four methods discussed in sec.2.2.1; the solid line represents the yields of AEPP(a,$N = 2^n$) where $n = 2, 3, 4, 5, 6$; the dashed line represents the optimized AEPP(a,4), which is denoted by AEPP*(a,4) (see section 2.2.3 for details).

**2. The modified recurrence method:** The modified recurrence method[7] is the repeat, alternate applications of AEPP(a,2) and AEPP(p,2). After Alice and Bob apply AEPP(a,2) and obtain identical measurement results, rather than applying universal hashing right away, they repeatedly and alternately apply AEPP(p,2), AEPP(a,2) and so forth until it becomes more beneficial to switch to universal hashing.

**3. The Maneva-Smolin method:** The Maneva-Smolin method[34] is to apply the first step of AEPP(a,N). Perform universal hashing on the N-1 pairs if the measurement results agree but throw away all the N-1 pairs if they do not. This is illustrated in figure 2-7.

**4. The Leung-Shor method:** The Leung-Shor method(section 2.1.1 and [32]) is a combination of the first step AEPP(a,4) and AEPP(p,4); however, this method fails to utilize all entanglements by throwing away the 3 pairs if the first measurement results disagree. This is illustrated in figure 2-8.

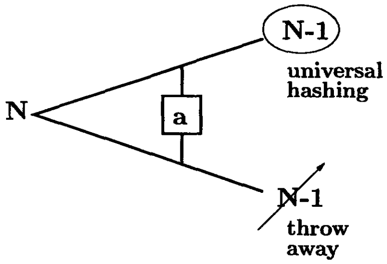


Figure 2-7: The Maneva-Smolin method[34].

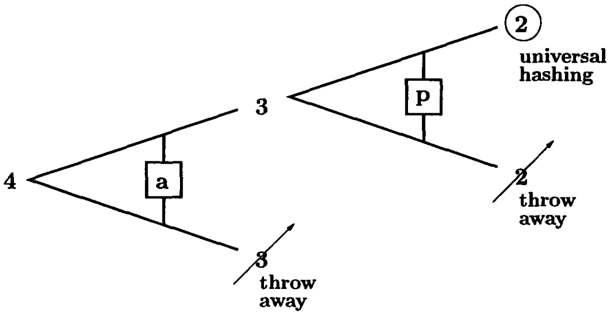


Figure 2-8: The Leung-Shor method: section 2.1.1 and [32].

### 2.2.3 Optimization

After we apply AEPP(a,N=$2^n$), we might end up with either $2^n-1$ pairs or n-1 groups of pairs ($2^{n-1}-1$, $2^{n-2}-1$, $\dots 2^k-1$, $\dots 3$ and 1) pairs depending on the results of the first measurements. It is possible to treat these n-1 groups differently because they are not entangled to each other. We can either perform universal hashing(as in the Maneva-Smolin method) or apply AEPP(p,$2^k-1$)(as in the Leung-Shor method). If we do apply AEPP(p,$2^k-1$), we will end up with two groups of quantum states of different sizes because we started with $N=2^k-1$ rather than $N=2^k$. In figure 2-9, we show two such possibilities as shown for $N=4$ , and higher yields are achieved for $F>0.74$ as shown in figure 2-6.

### 2.2.4 Higher yield than universal hashing

As we can see from figure 2-6, the yields of AEPP(a,$N=2^n$) exceed the yield of universal hashing for $F<0.993$. In section 2.3, we prove the following theorem:

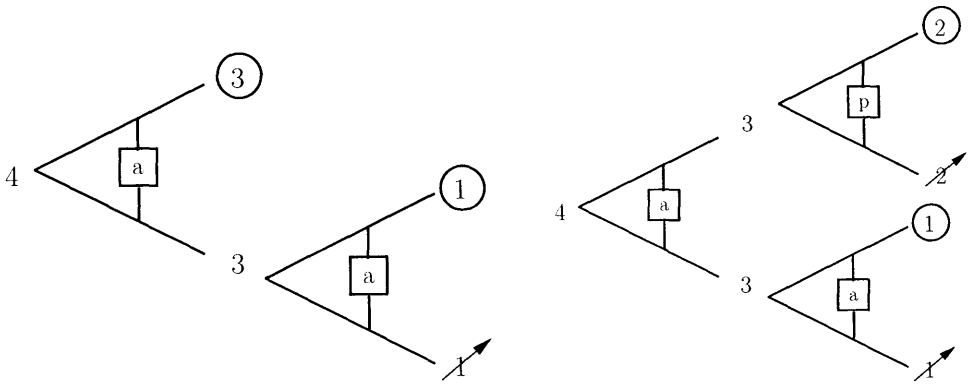


Figure 2-9: AEPP*(4,a).

**Theorem 1.** *Let $N = 2^n$ where $n$ is a positive integer. Denote by $Y_{AEPP}$ the yields of AEPP(a,N) on the Werner state $\rho_F$. Then*

$$\begin{aligned} Y_{AEPP} &= p \times \frac{N-1}{N} \times \left(1 - \frac{S_{N-1}}{N-1}\right) + (1-p) \times \left[\frac{N/2-1}{N} \times \left(1 - \frac{S_{\frac{N}{2}-1}}{N/2-1}\right)\right. \\ &\left. + \frac{N/4-1}{N} \times \left(1 - \frac{S_{\frac{N}{4}-1}}{N/4-1}\right) + \ldots + \frac{2-1}{N} \times \left(1 - \frac{S_1}{2-1}\right)\right] \\ &= 1 - \frac{p}{N}(1 + S_{N-1}) - \frac{1-p}{N}(n + 1 + S_{\frac{N}{2}-1} + S_{\frac{N}{4}-1} + \ldots + S_3 + S_1) \end{aligned}$$

*where* $p = prob(b_1 \oplus b_2 \oplus \ldots b_N = 0)$ *and* $S_{K-1} = H(a_1 \oplus a_K, b_1, a_2 \oplus a_K, b_2, \ldots, a_{K-1} \oplus a_K, b_{K-1} | b_1 \oplus \ldots \oplus b_K = 0)$ *for* $K = 2, 4, 8, \ldots, 2^n$*. Furthermore, let* $F = \frac{2^n-1}{2^n}$ *and* $G = \frac{1-F}{3}$*. Then*

$$\begin{aligned} \lim_{n \to \infty} Y_{AEPP} &\geq 1 - H(F,G,G,G) + \frac{H(p^*) - p^*}{N} \\ &> 1 - H(F,G,G,G) \\ &= \textit{Yield of universal hashing on } \rho_F \end{aligned}$$

*where*

$$p^* = \lim_{n \to \infty} p = \frac{1}{2}(1 + e^{-\frac{4}{3}}).$$

## 2.3 Proof of theorem 1

**Lemma 1.**

$$p^* = \lim_{n \to \infty} \mathit{prob}(b_1 \oplus b_2 \oplus \ldots b_N = 0) = \frac{1}{2}(1 + e^{-\frac{4}{3}}).$$

*Proof.* $b_i$'s are the amplitude error bits and, for any i, prob$(b_i = 1) = 2G = \frac{2}{3N}$. When we have $N$ qubit pairs,

$$\begin{aligned}
\text{prob(no error)} &= \left(1 - \frac{2}{3N}\right)^N \approx e^{-2/3} && (2.5)\\
\text{prob(1 error)} &= N\left(1 - \frac{2}{3N}\right)^{N-1}\left(\frac{2}{3N}\right)\\
&= \frac{2}{3}\left(1 - \frac{2}{3N}\right)^{N-1} \approx \frac{2}{3}e^{-2/3} && (2.6)\\
\text{prob(2 errors)} &= \frac{N(N-1)}{2!}\left(1 - \frac{2}{3N}\right)^{N-1}\left(\frac{2}{3N}\right)^2\\
&= \frac{N-1}{N}\left(1 - \frac{2}{3N}\right)^{-2}\frac{(2/3)^2}{2!}\left(1 - \frac{2}{3N}\right)^N \approx \frac{(2/3)^2}{2!}e^{-2/3} && (2.7)\\
\text{prob(k errors)} &= \frac{N(N-1)\ldots(N-k+1)}{N^k}\left(1 - \frac{2}{3N}\right)^{-k}\frac{(2/3)^k}{k!}\left(1 - \frac{2}{3N}\right)^N\\
&\approx \frac{(2/3)^k}{k!}e^{-2/3}\left(1 + O\left(\frac{k^2}{N}\right)\right) && (2.8)
\end{aligned}$$

In equation (2.5), (2.6) and (2.7), we did not include the error term $O\left(\frac{k^2}{N}\right)$ because $\sum_{k=0}^{N} \frac{(2/3)^k}{k!}\frac{k^2}{N}$ is negligible. In the following calculation, we will drop the error terms for brevity.

$$\begin{aligned}
p &= \text{prob}(b_1 \oplus b_2 \oplus \ldots b_N = 0) \\
&= \sum_{k \text{ is even}} \frac{(2/3)^k}{k!} e^{-2/3} \\
&= \frac{1}{2} \sum_{k=0}^{N} \left( \frac{(2/3)^k}{k!} e^{-2/3} + \frac{(-2/3)^k}{k!} e^{-2/3} \right) \\
&= \frac{e^{-2/3}}{2} \sum_{k=0}^{N} \left( \frac{(2/3)^k}{k!} + \frac{(-2/3)^k}{k!} \right) \\
\lim_{n \to \infty} p &= \frac{1 + e^{-4/3}}{2}
\end{aligned}$$

□

**Lemma 2.** *For $K = 2, 4, 8, \ldots, 2^n$, let*

$$\begin{aligned}
S_{K-1} &= H(a_1 \oplus a_K, b_1, a_2 \oplus a_K, b_2, \ldots, a_{K-1} \oplus a_K, b_{K-1} | b_1 \oplus \ldots \oplus b_K = 0) \text{ and} \\
T_{K-1} &= H(a_1 \oplus a_K, b_1, a_2 \oplus a_K, b_2, \ldots, a_{K-1} \oplus a_K, b_{K-1} | b_1 \oplus \ldots \oplus b_K = 1).
\end{aligned}$$

*Then*

$$N \times H(F, G, G, G) \geq H(p, 1-p) + pS_{N-1} + (1-p)T_{N-1} \text{ and} \tag{2.9}$$

$$T_{K-1} \geq S_{K/2-1} + T_{\frac{K}{2}-1} + 1 \tag{2.10}$$

*Proof.* To prove (2.9), note that $N \times H(F, G, G, G) = S(\rho_F^{\otimes N})$. Let $p = \text{prob}(b_1 \oplus b_2 \oplus \ldots b_N = 0)$. Then we can write $\rho_F^{\otimes N} = p \times \rho_0 + (1-p) \times \rho_1$, where $\rho_0$ is the state whose support lies on Bell states that are characterized by $b_1 \oplus b_2 \oplus \ldots b_N = 0$ and $\rho_1$ is the state whose support lies on Bell states that are characterized by $b_1 \oplus b_2 \oplus \ldots b_N = 1$. It is clear then $\rho_0$ and $\rho_1$ have orthogonal supports as any two distinct Bell states are. Then we have

$$
\begin{aligned}
&N \times H(F, G, G, G) = S(\rho_F^{\otimes N}) \\
&= S(p \times \rho_0 + (1-p) \times \rho_1) \\
&= pS(\rho_0) + (1-p)S(\rho_1) + H(p) \\
&= pH(a_1, b_1, a_2, \ldots, b_{N-1}, a_N, b_N | b_1 \oplus \ldots \oplus b_N = 0) + \\
&\quad (1-p)H(a_1, b_1, a_2, \ldots, b_{N-1}, a_N, b_N | b_1 \oplus \ldots \oplus b_N = 1) + H(p) \\
&= pH(a_1 \oplus a_N, b_1, \ldots, a_{N-1} \oplus a_N, b_{N-1}, a_N, b_1 \oplus \ldots \oplus b_N | b_1 \oplus \ldots \oplus b_N = 0) + \\
&\quad (1-p)H(a_1 \oplus a_N, b_1, \ldots, a_{N-1} \oplus a_N, b_{N-1}, a_N, b_1 \oplus \ldots \oplus b_N | \\
&\quad b_1 \oplus \ldots \oplus b_N = 1) + H(p) \qquad (2.11)
\end{aligned}
$$

where the last equality holds because the operations $BXOR(1, N)$, $BXOR(2, N) \ldots$ $BXOR(N-1, N)$ are all unitary and hence preserve entropy. Since the function $-p \log_2 p$ is subadditive for $p \leq 1$, we have

$$
\begin{aligned}
S_{N-1} &\equiv H(a_1 \oplus a_N, b_1, \ldots, a_{N-1} \oplus a_N, b_{N-1} | b_1 \oplus \ldots \oplus b_N = 0) \\
&\leq H(a_1 \oplus a_N, b_1, \ldots, a_{N-1} \oplus a_N, b_{N-1}, a_N, b_1 \oplus \ldots \oplus b_N | b_1 \oplus \ldots \oplus b_N = 0) \text{ and} \\
T_{N-1} &\equiv H(a_1 \oplus a_N, b_1, \ldots, a_{N-1} \oplus a_N, b_{N-1} | b_1 \oplus \ldots \oplus b_N = 1) \\
&\leq H(a_1 \oplus a_N, b_1, \ldots, a_{N-1} \oplus a_N, b_{N-1}, a_N, b_1 \oplus \ldots \oplus b_N | b_1 \oplus \ldots \oplus b_N = 1).
\end{aligned}
$$

Substituting these into equation (2.11) yields $N \times H(F, G, G, G) \geq H(p) + pS_{N-1} + (1-p)T_{N-1}$.

To prove (2.10), note that for K qubit pairs shared between Alice and Bob where $K = 2, 4, 8, \ldots, N$, when they apply the unitary operations $BXOR(1, K)$, $BXOR(2, K) \ldots BXOR(K-1, K)$ and get different results for measuring the Kth pair along the z-axis, then the entropy of the resulting quantum state $\Gamma$ can be described by $S(\Gamma) = H(a_1 \oplus a_K, b_1, \ldots, a_{K-1} \oplus a_K, b_{K-1} | b_1 \oplus \ldots \oplus b_K = 1))$. By symmetry, prob $(b_1 \oplus \ldots \oplus b_{K/2} = 1) =$ prob $(b_{K/2+1} \oplus \ldots \oplus b_K = 1) = \frac{1}{2}$. Therefore, $\Gamma = \frac{1}{2}\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)} + \frac{1}{2}\Gamma_{(b_{K/2+1} \oplus \ldots \oplus b_K=1)}$ where $\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)}$ and $\Gamma_{(b_{K/2+1} \oplus \ldots \oplus b_K=1)}$ have orthogonal supports:

$$\begin{aligned} T_{K-1} &\equiv H(a_1 \oplus a_K, b_1, \ldots, a_{K-1} \oplus a_K, b_{K-1} | b_1 \oplus \ldots \oplus b_K = 1) \\ &= S(\Gamma) \\ &= S\Big(\frac{1}{2}\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)} + \frac{1}{2}\Gamma_{(b_{K/2+1} \oplus \ldots \oplus b_K=1)}\Big) \\ &= \frac{1}{2}S\Big(\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)}\Big) + \frac{1}{2}S\Big(\Gamma_{(b_{K/2+1} \oplus \ldots \oplus b_K=1)}\Big) + H\Big(\frac{1}{2}\Big). \qquad (2.12) \end{aligned}$$

Note that

$$\begin{aligned} & S\Big(\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)}\Big) \\ &= H\Big(a_1 \oplus a_K, b_1, \ldots, a_{K-1} \oplus a_K, b_{K-1} \Big| (b_1 \oplus \ldots \oplus b_{K/2} = 1) \wedge (b_{K/2+1} \oplus \ldots \oplus b_K = 0)\Big) \\ &= H\Big(a_1 \oplus a_K, b_1 \oplus \ldots \oplus b_{K/2}, a_2 \oplus a_1, b_2, \ldots, a_{K/2} \oplus a_1, b_{K/2}, a_{K/2+1} \oplus a_K, \\ &\quad b_{K/2+1}, \ldots, a_{K-1} \oplus a_K, b_{K-1} \Big| (b_1 \oplus \ldots \oplus b_{K/2} = 1) \wedge (b_{K/2+1} \oplus \ldots \oplus b_K = 0)\Big) \\ &= H\Big(a_1 \oplus a_K, b_1 \oplus \ldots \oplus b_{K/2}, a_2 \oplus a_1, b_2, \ldots, a_{K/2} \oplus a_1, b_{K/2} \Big| (b_1 \oplus \ldots \oplus b_{K/2} = 1)\Big) \\ &\quad H\Big(a_{K/2+1} \oplus a_K, b_{K/2+1}, \ldots, a_{K-1} \oplus a_K, b_{K-1} \Big| (b_{K/2+1} \oplus \ldots \oplus b_K = 0)\Big) \\ &= H\Big(a_1 \oplus a_K, b_1 \oplus \ldots \oplus b_{K/2}, a_2 \oplus a_1, b_2, \ldots, a_{K/2} \oplus a_1, b_{K/2} \Big| (b_1 \oplus \ldots \oplus b_{K/2} = 1)\Big) \\ &\quad + S_{K/2-1} \qquad (2.13) \end{aligned}$$

where the second equality was obtained by applying the operations $BXOR(2,1)$, $BXOR(3,1)$,..., $BXOR(K/2-1,1)$, $BXOR(K/2,1)$. And since $-p\log_2 p$ is subadditive for $p \leq 1$, the first term in (2.13)

$$H\Big(a_1 \oplus a_K, b_1 \oplus \ldots \oplus b_{K/2}, a_2 \oplus a_1, b_2, \ldots, a_{K/2} \oplus a_1, b_{K/2}\Big|\big(b_1 \oplus \ldots \oplus b_{K/2} = 1\big)\Big)$$
$$\geq H\Big(a_2 \oplus a_1, b_2, \ldots, a_{K/2} \oplus a_1, b_{K/2}\Big|\big(b_1 \oplus \ldots \oplus b_{K/2} = 1\big)\Big)$$
$$= T_{K/2-1}. \tag{2.14}$$

Therefore, $S\Big(\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)}\Big) \geq T_{K/2+1} + S_{K/2-1}$. Using similar argumnts, one can show $S\Big(\Gamma_{(b_{K/2+1} \oplus \ldots \oplus b_K=1)}\Big) \geq S_{K/2-1} + T_{K/2+1}$. Putting these back to equation (2.12),

$$\begin{aligned} & T_{K-1} \\ = & \frac{1}{2}S\Big(\Gamma_{(b_1 \oplus \ldots \oplus b_{K/2}=1)}\Big) + \frac{1}{2}S\Big(\Gamma_{(b_{K/2+1} \oplus \ldots \oplus b_K=1)}\Big) + H\Big(\frac{1}{2}\Big) \\ \geq & \frac{1}{2}\big(T_{K/2+1} + S_{K/2-1}\big) + \frac{1}{2}\big(S_{K/2-1} + T_{K/2+1}\big) + 1 \\ = & S_{K/2-1} + T_{K/2-1} + 1. \end{aligned}$$

□

We are now ready to prove theorem 1. We first apply the above lemmas:

$$\begin{aligned} & N \times H(F, G, G, G) \\ \geq & H(p) + pS_{N-1} + (1-p)T_{N-1} \\ \geq & H(p) + pS_{N-1} + (1-p)\Big(1 + S_{N/2-1} + T_{N/2-1}\Big) \\ \geq & H(p) + pS_{N-1} + (1-p)\Big(1 + S_{N/2-1} + 1 + S_{N/4-1} + T_{N/4-1}\Big) \\ \geq & H(p) + pS_{N-1} + (1-p)\Big((n-1) + S_{N/2-1} + S_{N/4-1} + \ldots + S_3 + S_1 + T_1\Big) \end{aligned}$$

$$\Rightarrow pS_{N-1}+(1-p)(S_{N/2-1}+\ldots+S_3+S_1 \leq N\times H(F,G,G,G)-H(p)-(1-p)\Big(n-1+T_1\Big)$$

Simple calculation shows $T_1 = 2$. Therefore,

$$\begin{aligned}
Y_{AEPP} &= 1-\frac{p}{N}(1+S_{N-1})-\frac{1-p}{N}(n+1+S_{N/2-1}+S_{N/4-1}+\ldots+S_3+S_1) \\
&= 1-\frac{p}{N}-\frac{1-p}{N}(n+1)-\frac{p}{N}S_{N-1}-\frac{1-p}{N}(S_{N/2-1}+\ldots+S_3+S_1) \\
&\geq 1-\frac{p}{N}-\frac{1-p}{N}(n+1)-\frac{1}{N}\Big(N\times H(F,G,G,G)-H(p)-(1-p)(n-1+T_1)\Big) \\
&\geq 1-H(F,G,G,G)+\frac{H(p)-p}{N} \\
&> 1-H(F,G,G,G).
\end{aligned}$$

This completes the proof of theorem 1.

## 2.4 Modified AEPP

In this section, we modify AEPP from the previous section to achieve higher yields. Recall that, as the first steps of AEPP(a,N=$2^n$), Alice and Bob apply $BXOR(1,N)$, $BXOR(2,N)$, ..., $BXOR(N-1,N)$ to obtain the quantum state $(a_1 \oplus a_N, b_1, a_2 \oplus a_N, b_2, \ldots, a_{N-1} \oplus a_N, b_{N-1}, a_N, b_1 \oplus b_2 \oplus \ldots \oplus b_{N-1} \oplus b_N)$ and take measurements on the Nth qubit pair, $(a_N, b_1 \oplus b_2 \oplus \ldots \oplus b_{N-1} \oplus b_N)$. However, if the entropy of the Nth qubit pair is small, or more precisely if $S(a_N, b_1 \oplus b_2 \oplus \ldots \oplus b_{N-1} \oplus b_N) \leq 1$, they can perform universal hashing instead of taking measurements along the z-axis. Specifically, Alice and Bob can apply AEPP(a,N=$2^n$) to M blocks of N qubit pairs and apply universal hashing to the M Nth qubit pairs as shown in figure 2-10. This modification has two immediate advantages:

1. As a result of hashing, there are an extra amount of EPR pairs equal to $\Big(1 -$

$S(a_N, b_1 \oplus b_2 \oplus \ldots \oplus b_{N-1} \oplus b_N)\Big)/N.$

2. Taking measurements in the original AEPP protocols destroys the information in $a_N$. However, universal hashing not only reveals the identity of $b_1 \oplus b_2 \oplus \ldots \oplus b_{N-1} \oplus b_N$ but the value of $a_N$ as well. As a result, if $a_N = 0$ and $b_1 \oplus \ldots \oplus b_N = 0$, the $N-1$ remaining qubit pairs are represented by $(a_1, b_1, a_2, b_2, \ldots, a_{N-1}, b_{N-1})$; and if $a_N = 1$ and $b_1 \oplus \ldots \oplus b_N = 0$, the qubits are represented by $(a_1 \oplus 1, b_1, a_2 \oplus 1, b_2, \ldots, a_{N-1} \oplus 1, b_{N-1})$. Alice and Bob can collect a large number of these two distinct groups and apply universal hashing separately. By the concavity of entropy function, $S(p_1\rho_1 + p_2\rho_2) \geq p_1 S(\rho_1) + p_2 S(\rho_2)$, the entropy is smaller and hence a higher yield will be obtained by hashing.

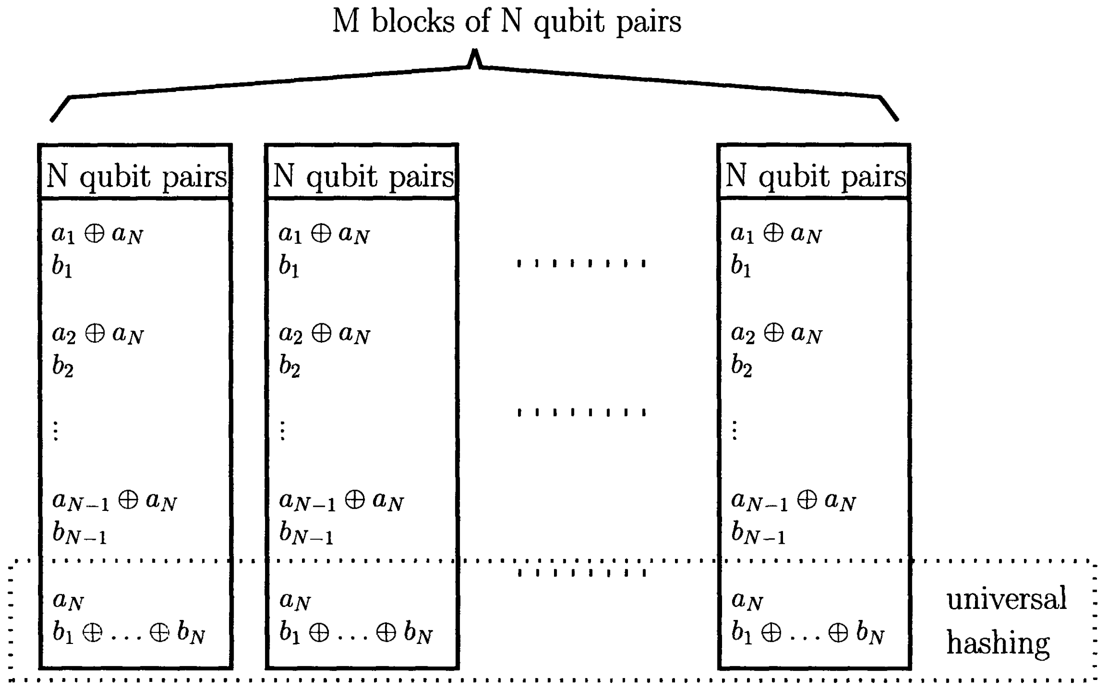


Figure 2-10: New-AEPP(a, N=$2^n$). Alice and Bob replace measurements along $Z$-axis by universal hashing wherever the entropy of the qubit pair is less than 1.

In AEPP(a, N=$2^n$), there are $n-2$ more measurements if the first measurement reveals $b_1 \oplus \ldots \oplus b_N = 1$. Obviously we should also replace these measurements by hashing whenever possible, and take measurements only if the entropy of the qubit pairs is greater than 1. We now explain how to compute the yield for New-AEPP(a,4) before we give a recursing procedure to compute the yields of New-AEPP(a,N=$2^n$).

### 2.4.1 New-AEPP(a,4)

After Alice and Bob apply $BXOR(1,4)$,$BXOR(2,4)$,$BXOR(3,4)$, the quantum states become

$$\begin{array}{l} a_1 \oplus a_4 \\ b_1 \\ a_2 \oplus a_4 \\ b_2 \\ a_3 \oplus a_4 \\ b_3 \\ a_4 \\ b_1 \oplus b_2 \oplus b_3 \oplus b_4 \end{array}$$

Denote by m the value of $(a_4, b_1 \oplus b_2 \oplus b_3 \oplus b_4)$. Then the yield of New-AEPP(a,4) is given by

$$\begin{aligned} & \text{prob}(m=00) \times \ \text{HASH-00}(4) + \ \text{prob}(m=01) \times \ \text{HASH-01}(4) \\ + & \ \text{prob}(m=10) \times \ \text{HASH-10}(4) + \ \text{prob}(m=11) \times \ \text{HASH-11}(4) \\ + & \Big(1 - H\big( \ \text{prob}(m=00), \ \text{prob}(m=01), \ \text{prob}(m=10), \ \text{prob}(m=11)\big)\Big)/4 \end{aligned}$$

**1. HASH-00(4):** Conditioned on $m \equiv (a_4, b_1 \oplus b_2 \oplus b_3 \oplus b_4) = 00$, the 3 remaining qubit pairs are $(a_1, b_1, a_2, b_2, a_3, b_3)$ and by universal hashing the yield is

$$\text{HASH-00}(4) = \left(1 - H\left(a_1, b_1, a_2, b_2, a_3, b_3 \middle| (a_4 = 0) \wedge (b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 0)\right)/3\right) \times \frac{3}{4}.$$

**2. HASH-10(4):** Conditioned on $m \equiv (a_4, b_1 \oplus b_2 \oplus b_3 \oplus b_4) = 10$, the 3 remaining qubit pairs are $(a_1 \oplus 1, b_1, a_2 \oplus 1, b_2,\ a_3 \oplus 1, b_3)$ and by universal hashing the yield is

$$\text{HASH-10}(4) = \left(1 - H\left(a_1 \oplus 1, b_1, a_2 \oplus 1, b_2, a_3 \oplus 1, b_3 \middle| (a_4 = 1) \wedge (b_1 \oplus b_2 \oplus b_3 \oplus b_4 = 0)\right)/3\right) \times \frac{3}{4}.$$

**3. HASH-01(4):** Conditioned on $m \equiv (a_4, b_1 \oplus b_2 \oplus b_3 \oplus b_4) = 01$, the 3 remaining qubit pairs are

$$\begin{aligned} a_1 &\longmapsto a_1 \\ b_1 &\longmapsto b_1 \oplus b_2 \\ a_2 &\longmapsto a_2 \oplus a_1 \\ b_2 &\longmapsto b_2 \\ a_3 &\longmapsto a_3 \\ b_3 &\longmapsto b_3 \end{aligned}$$

where the mapping is achieved by applying $BXOR(2,1)$. Denote by q the value of $(a_1, b_1 \oplus b_2)$. Depending on the entropy of this qubit pair, Alice and Bob can choose to apply universal hashing or take a measurement. First, assume the entropy of this pair is less than 1, then universal hashing is applied. If $q = 00$, the second pair becomes $(a_2, b_2)$ conditioned on $a_1 = 0$ and $b_1 \oplus b_2 = 0$; if $q = 10$, the second pair becomes $(a_2 \oplus 1, b_2)$ conditioned on $a_1 = 1$ and $b_1 \oplus b_2 = 0$; if $q = 01$, the third pair becomes $(a_3, b_3)$ conditioned on $a_4 = 0$ and $b_3 \oplus b_4 = 0$; finally, if $q = 11$, the third pair becomes $(a_3, b_3)$ conditioned on $a_4 = 0$ and $b_3 \oplus b_4 = 0$. Therefore, if

$H\Big(\ \mathrm{prob}(q=00),\ \mathrm{prob}(q=01),\ \mathrm{prob}(q=10),\ \mathrm{prob}(q=11)\Big) \leq 1$, then

$$\begin{aligned}\text{HASH-01}(4) = &\Big(\ \mathrm{prob}(q=00) +\ \mathrm{prob}(q=01) +\ \mathrm{prob}(q=11)\Big)\\ &\times \Big(1 - H\Big(a_2, b_2\Big|(a_1=0)\wedge(b_1\oplus b_2=0)\Big)\Big)/4\\ &+\ \mathrm{prob}(q=10)\times\Big(1 - H\Big(a_2\oplus 1, b_2\Big|(a_1=1)\wedge(b_1\oplus b_2=0)\Big)\Big)/4\\ &+\Big(1 - H\Big(\ \mathrm{prob}(q=00),\ \mathrm{prob}(q=01),\ \mathrm{prob}(q=10),\ \mathrm{prob}(q=11)\Big)\Big)/4.\end{aligned}$$

If the entropy is greater than one and they have to take measurements on the first qubit pair, then the probability of getting identical results is simply $\mathrm{prob}(q=00)+\ \mathrm{prob}(q=10)$ and that of getting different results is $\mathrm{prob}(q=01)+\ \mathrm{prob}(q=11)$. If $b_1\oplus b_2=0$, Alice and Bob can apply hashing on the second pair, $(a_1\oplus a_2, b_2)$ and the yield is $(1-S_1)/4$ using the notation from the previous section; if $b_1\oplus b_2=1$, then the yield of applying hashing on the third pair is $\Big(1-H(a_3,b_3|(a_4=0)\wedge(b_3\oplus b_4=0))\Big)/4$. Therefore, if $H\Big(\ \mathrm{prob}(q=00),\ \mathrm{prob}(q=01),\ \mathrm{prob}(q=10),\ \mathrm{prob}(q=11)\Big) > 1$, then

$$\begin{aligned}\text{HASH-11}(4) = &\Big(\ \mathrm{prob}(q=00)+\ \mathrm{prob}(q=10)\Big)\times(1-S_1)/4+\Big(\ \mathrm{prob}(q=01)\\ &+\ \mathrm{prob}(q=11)\Big)\times\Big(1-H\Big(a_3,b_3\Big|(a_4=0)\wedge(b_3\oplus b_4=0)\Big)\Big)/4.\end{aligned}$$

**4. HASH-11(4):** Conditioned on $m\equiv(a_4, b_1\oplus b_2\oplus b_3\oplus b_4)=11$, the 3 remaining qubit pairs are

$$\begin{aligned} a_1 \oplus 1 &\longmapsto a_1 \oplus 1 \\ b_1 &\longmapsto b_1 \oplus b_2 \\ a_2 \oplus 1 &\longmapsto a_2 \oplus a_1 \\ b_2 &\longmapsto b_2 \\ a_3 \oplus 1 &\longmapsto a_3 \oplus 1 \\ b_3 &\longmapsto b_3 \end{aligned}$$

where the mapping is achieved by applying $BXOR(2, 1)$. Denote by r the value of $(a_1 \oplus 1, b_1 \oplus b_2)$. Depending on the entropy of this qubit pair, Alice and Bob can choose to apply hashing or to take a measurement. First, assume the entropy of this pair is less than 1, then universal hashing is applied. If $r = 00$, the second pair becomes $(a_2 \oplus 1, b_2)$ conditioned on $a_1 = 1$ and $b_1 \oplus b_2 = 0$; if $r = 01$, the third pair becomes $(a_3 \oplus 1, b_3)$ conditioned on $a_4 = 1$ and $b_3 \oplus b_4 = 0$; if $r = 11$, the third pair becomes $(a_3 \oplus 1, b_3)$ conditioned on $a_4 = 1$ and $b_3 \oplus b_4 = 0$; finally, if $r = 10$, the second pair becomes $(a_2, b_2)$ conditioned on $a_1 = 0$ and $b_1 \oplus b_2 = 0$. Therefore, if $H\big(\text{ prob}(r = 00),\ \text{prob}(r = 01),\ \text{prob}(r = 10),\ \text{prob}(r = 11)\big)\Big) \leq 1$, then

$$\begin{aligned} \text{HASH-11}(4) = &\Big(\text{ prob}(r = 00) + \text{ prob}(r = 01) + \text{ prob}(r = 11)\Big) \\ &\times \Big(1 - H\Big(a_2 \oplus 1, b_2 \Big| (a_1 = 1) \wedge (b_1 \oplus b_2 = 0)\Big)\Big)/4 \\ &+ \text{ prob}(r = 10) \times \Big(1 - H\Big(a_2, b_2 \Big| (a_1 = 0) \wedge (b_1 \oplus b_2 = 0)\Big)\Big)/4 \\ &+ \Big(1 - H\Big(\text{ prob}(r = 00),\ \text{prob}(r = 01),\ \text{prob}(r = 10),\ \text{prob}(r = 11)\Big)\Big)/4. \end{aligned}$$

If Alice and Bob have to take measurements on the first qubit pair, then the probability of getting identical results is simply prob$(r = 00) +$ prob$(r = 10)$ and that of getting different results is prob$(r = 01) +$ prob$(r = 11)$. If $b_1 \oplus b_2 = 0$, Alice and Bob can apply hashing on the second pair, $(a_1 \oplus a_2, b_2)$ and the yield is $(1 - S_1)/4$; if $b_1 \oplus b_2 = 1$, then the yield of hashing on the third pair is $\Big(1 - H(a_3 \oplus 1, b_3|(a_4 = 1) \wedge (b_3 \oplus b_4 = 0))\Big)/4$. Therefore, if $H\big($ prob$(r = 00)$, prob$(r = 01)$, prob$(r = 10)$, prob$(r = 11)\big)\Big) > 1$, then

$$\begin{aligned}\text{HASH-11}(4) = {} & \Big(\text{ prob}(r = 00) + \text{ prob}(r = 10)\Big) \times (1 - S_1)/4 + \Big(\text{ prob}(r = 01) \\ & + \text{ prob}(r = 11)\Big) \times \Big(1 - H\Big(a_3 \oplus 1, b_3\Big|(a_4 = 1) \wedge (b_3 \oplus b_4 = 0)\Big)\Big)/4.\end{aligned}$$

### 2.4.2 New-AEPP(a,N=2$^{\text{n}}$)

After Alice and Bob apply $BXOR(1, N)$,$BXOR(2, N)$,..., $BXOR(N - 1, N)$, the quantum states become

$$\begin{array}{l} a_1 \oplus a_N \\ b_1 \\ a_2 \oplus a_N \\ b_2 \\ \vdots \\ a_{N-1} \oplus a_N \\ b_{N-1} \\ a_N \\ b_1 \oplus \ldots \oplus b_N \end{array}$$

Denote by m the value of $(a_N, b_1 \oplus \ldots \oplus b_N)$. Then the yield of New-AEPP(a,N=$2^n$) is given by

$$\begin{aligned} & \text{prob}(m = 00) \times \text{ HASH-00}(N) + \text{ prob}(m = 01) \times \text{ HASH-01}(N) \\ + & \text{ prob}(m = 10) \times \text{ HASH-10}(N) + \text{ prob}(m = 11) \times \text{ HASH-11}(N) \\ + & \Big(1 - H\big( \text{ prob}(m = 00), \text{ prob}(m = 01), \text{ prob}(m = 10), \text{ prob}(m = 11)\big)\Big)/N. \end{aligned}$$

where HASH-00(N), HASH-01(N), HASH-10(N) and HASH-11(N) can be evaluated using the following recurrence formulas for $K = 4, 8, 16, 32, \ldots, N$:

1. **HASH-00(K):** Conditioned on $m \equiv (a_K, b_1 \oplus \ldots \oplus b_K) = 00$, the $K-1$ remaining qubit pairs are $(a_1, b_1, \ldots, a_{K-1}, b_{K-1})$ and by universal hashing the yield is

$$\begin{aligned} & \text{HASH-00}(K) \\ = & \left(1 - \frac{H\Big(a_1, b_1, \ldots, a_{K-1}, b_{K-1} \Big| (a_K = 0) \wedge (b_1 \oplus \ldots \oplus b_K = 0)\Big)}{K-1}\right) \times \frac{K-1}{K}. \end{aligned}$$

2. **HASH-10(K):** Conditioned on $m \equiv (a_K, b_1 \oplus \ldots \oplus b_K) = 10$, the $K-1$ remaining qubit pairs are $(a_1 \oplus 1, b_1, \ldots, a_{K-1} \oplus 1, b_{K-1})$ and by universal hashing the yield is

$$\begin{aligned} & \text{HASH-10}(K) \\ = & \left(1 - \frac{H\Big(a_1 \oplus 1, b_1, \ldots, a_{K-1} \oplus 1, b_{K-1} \Big| (a_K = 1) \wedge (b_1 \oplus \ldots \oplus b_K = 0)\Big)}{K-1}\right) \times \frac{K-1}{K}. \end{aligned}$$

3. **HASH-01(K):** Conditioned on $m \equiv (a_K, b_1 \oplus \ldots \oplus b_K) = 01$, the K-1 remaining qubit pairs are

$$
\begin{aligned}
a_1 &\longmapsto a_1\\
b_1 &\longmapsto b_1 \oplus \ldots \oplus b_{K/2}\\
a_2 &\longmapsto a_2 \oplus a_1\\
b_2 &\longmapsto b_2\\
&\vdots\\
a_{K/2} &\longmapsto a_{K/2} \oplus a_1\\
b_{K/2} &\longmapsto b_{K/2}\\
a_{K/2+1} &\longmapsto a_{K/2+1}\\
b_{K/2+1} &\longmapsto b_{K/2+1}\\
&\vdots\\
a_{K-1} &\longmapsto a_{K-1}\\
b_{K-1} &\longmapsto b_{K-1}
\end{aligned}
$$

where the mapping is achieved by applying $BXOR(2,1)$, $BXOR(3,1), \ldots, BXOR(\frac{K}{2},1)$. Denote by q the value of $(a_1, b_1 \oplus \ldots \oplus b_{K/2})$. Depending on the entropy of this qubit pair, Alice and Bob can choose to apply universal hashing or take a measurement like the discussion of New-AEPP(a,4). It can be verified that if $H\Big(\text{ prob}(q = 00),\ \text{prob}(q = 01),\ \text{prob}(q = 10),\ \text{prob}(q = 11)\Big) \leq 1$ and they apply hashing, then

$$
\begin{aligned}
\text{HASH-01}(K) = \Bigg(&\Big(\text{ prob}(q = 00) + \text{ prob}(q = 01) + \text{ prob}(q = 11)\Big) \times \text{HASH-00(K/2)}\\
&+ \Big(\text{ prob}(q = 00) + \text{ prob}(q = 01) + \text{ prob}(q = 10)\Big) \times \text{HASH-01(K/2)}\\
&+ \text{ prob}(q = 10) \times \text{HASH-10(K/2)} + \text{ prob}(q = 11) \times \text{HASH-11(K/2)}\Bigg)/2\\
&+ \Big(1 - H\big(\text{ prob}(q = 00),\ \text{prob}(q = 01),\ \text{prob}(q = 10),\ \text{prob}(q = 11)\big)\Big)/K.
\end{aligned}
$$

If the entropy is greater than one and they have to take measurements on the first qubit pair, then the probability of getting identical results is simply $\text{prob}(q = 00) + \text{prob}(q = 10)$ and that of getting different results is $\text{prob}(q = 01) + \text{prob}(q = 11)$. Therefore, if $H\big(\text{prob}(q = 00), \text{prob}(q = 01), \text{prob}(q = 10), \text{prob}(q = 11))\big) > 1$, then

$$\begin{aligned}
&\text{HASH-01}(K)\\
&= \left(\text{prob}(q = 00) + \text{prob}(q = 10)\right) \times \left(\text{HASH-01(K/2)/2} + \left(1 - \frac{S_{\frac{K}{2}-1}}{\frac{K}{2}-1}\right) \times \frac{\frac{K}{2}-1}{K}\right)\\
&+ \left(\text{prob}(q = 01) + \text{prob}(q = 11)\right) \times \left(\text{HASH-00(K/2)/2} + \left(\left(1 - \frac{S_{\frac{K}{4}-1}}{\frac{K}{4}-1}\right) \times \frac{\frac{K}{4}-1}{K}\right.\right.\\
&+ \left(1 - \frac{S_{\frac{K}{8}-1}}{\frac{K}{8}-1}\right) \times \frac{\frac{K}{8}-1}{K} + \ldots + \left.\left.\left(1 - \frac{S_1}{2-1}\right) \times \frac{2-1}{K}\right)\right).
\end{aligned}$$

4. **HASH-11(K):** Conditioned on $m \equiv (a_K, b_1 \oplus \ldots \oplus b_K) = 11$, the K-1 remaining qubit pairs are

$$\begin{aligned}
a_1 \oplus 1 &\longmapsto a_1\\
b_1 &\longmapsto b_1 \oplus \ldots \oplus b_{K/2}\\
a_2 \oplus 1 &\longmapsto a_2 \oplus a_1\\
b_2 &\longmapsto b_2\\
&\vdots\\
a_{K/2} \oplus 1 &\longmapsto a_{K/2} \oplus a_1\\
b_{K/2} &\longmapsto b_{K/2}
\end{aligned}$$

$$
\begin{aligned}
a_{K/2+1} \oplus 1 &\longmapsto a_{K/2+1} \oplus 1 \\
b_{K/2+1} &\longmapsto b_{K/2+1} \\
&\vdots \\
a_{K-1} \oplus 1 &\longmapsto a_{K-1} \oplus 1 \\
b_{K-1} &\longmapsto b_{K-1}
\end{aligned}
$$

where the mapping is achieved by applying $BXOR(2,1), BXOR(3,1), \ldots, BXOR(\frac{K}{2},1)$. Denote by r the value of $(a_1, b_1 \oplus \ldots \oplus b_{K/2})$. It can be verified that if $H\big(\text{ prob}(r = 00),\ \text{prob}(r = 01),\ \text{prob}(r = 10),\ \text{prob}(r = 11)\big)\Big) \leq 1$ and they apply hashing, then

$$
\begin{aligned}
\text{HASH-11}(K) = \Bigg(&\Big(\text{ prob}(r = 00) + \text{ prob}(r = 01) + \text{ prob}(r = 11)\Big) \times \text{HASH-10(K/2)} \\
&+ \Big(\text{ prob}(r = 00) + \text{ prob}(r = 01) + \text{ prob}(r = 10)\Big) \times \text{HASH-11(K/2)} \\
&+ \text{ prob}(r = 10) \times \text{HASH-00(K/2)} + \text{ prob}(r = 11) \times \text{HASH-01(K/2)}\Bigg)/2 \\
&+ \Big(1 - H\big(\text{ prob}(r = 00),\ \text{prob}(r = 01),\ \text{prob}(r = 10),\ \text{prob}(r = 11)\big)\Big)/K.
\end{aligned}
$$

If the entropy is greater than one and they have to take measurements on the first qubit pair, then the probability of getting identical results is simply $\text{prob}(r = 00) + \text{prob}(r = 10)$ and that of getting different results is $\text{prob}(r = 01) + \text{prob}(r = 11)$. Therefore, if $H\big(\text{ prob}(r = 00),\ \text{prob}(r = 01),\ \text{prob}(r = 10),\ \text{prob}(r = 11)\big)\Big) > 1$, then

$$
\begin{aligned}
&\text{HASH-11}(K)\\
&= \left(\text{prob}(r=00)+\text{ prob}(r=10)\right)\times\left(\text{HASH-11(K/2)/2}+\left(1-\frac{S_{\frac{K}{2}-1}}{\frac{K}{2}-1}\right)\times\frac{\frac{K}{2}-1}{K}\right)\\
&+\left(\text{prob}(r=01)+\text{ prob}(r=11)\right)\times\left(\text{HASH-10(K/2)/2}+\left(\left(1-\frac{S_{\frac{K}{4}-1}}{\frac{K}{4}-1}\right)\times\frac{\frac{K}{4}-1}{K}\right.\right.\\
&\left.\left.+\left(1-\frac{S_{\frac{K}{8}-1}}{\frac{K}{8}-1}\right)\times\frac{\frac{K}{8}-1}{K}+\ldots+\left(1-\frac{S_1}{2-1}\right)\times\frac{2-1}{K}\right)\right).
\end{aligned}
$$

### 2.4.3 Yield of New-AEPP(a,N=$2^{\text{n}}$) on the Werner state

In figure 2-11, we compute the yield of New-AEPP(a,N=$2^{\text{n}}$) on the Werner state for $n = 2, 3, 4, 5, 6$. Note that the protocols as described in the previous section are only well-defined when $S(a_N, b_1 \oplus \ldots \oplus b_N) \leq 1$. Also, with this modification, even for $N = 4$, the yield of New-AEPP(a,N=$2^{\text{n}}$) is higher than that of universal hashing for any $F \leq 0.99999$.

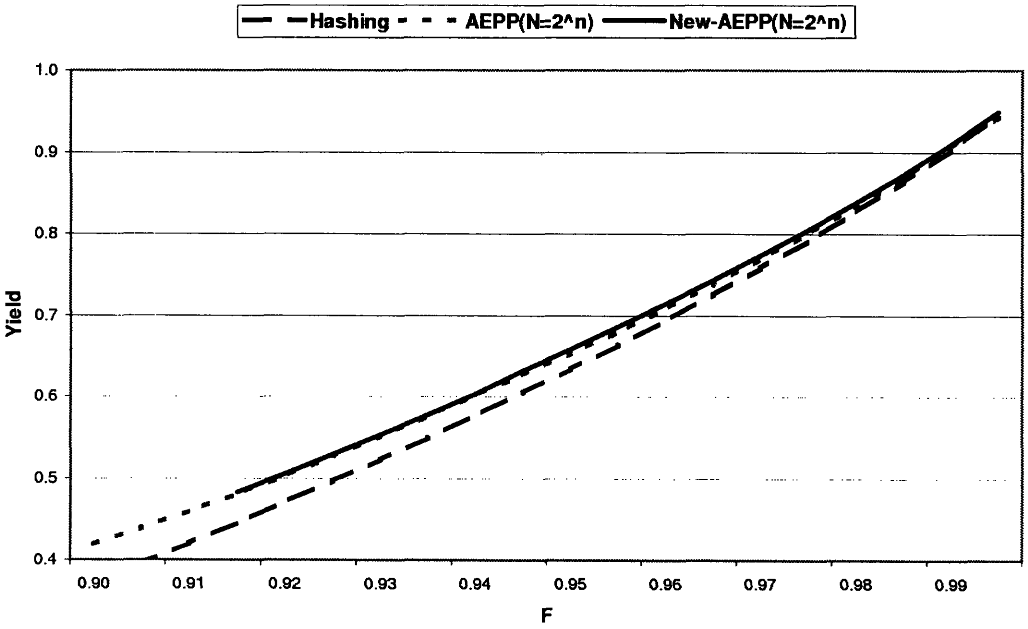


Figure 2-11: Yield of New-AEPP(a, N=$2^n$) on the Werner state $\rho_F$ for $n = 2, 3, 4, 5, 6$.

## 2.5 Discussion on 2-EPP

We presented a family of entanglement purification protocols AEPP(a,N) with improved yields over previous two-way entanglement purification protocols. Moreover, the yields of these protocols are higher than the yield of universal hashing for $F < 0.993$ (computed numerically) and as F goes to 1 (shown analytically in section 2.3). We modified AEPP(a,N) by replacing measurement with hashing and obtain yields higher than the yield of universal hashing for $F < 0.99999$. There are other recent progresses in this direction[25, 46]. It is worth studying whether one can combine AEPP with these works to achieve even higher yields.

# Chapter 3

# Quantum entanglement capacity with classical feedback

It is an open question whether the quantum capacity with classical feedback $Q_B$ and the quantum capacity with two-way classical communication $Q_2$ are equal to one another[5, 8]. In section 3.1, we define a new quantity called quantum entanglement capacity with classical feedback $E_B$ and this quantity is shown to lie between $Q_B$ and $Q_2$. We will then give an alternate operational meaning of $E_B$. In section 3.2, we describe how one can turn a QECC into an $E_B$ protocol. We demonstrate the idea with Cat code and Shor code, and we modify some of the 2-EPP in the last chapter to $E_B$ protocols. In section 3.3, we compute new lower bounds on $Q_B$ implied by these $E_B$ protocols. We then discuss some characteristics of Cat code and discuss further research directions.

## 3.1 A quantity that lies between $Q_B$ and $Q_2$

In this section, we define, for any quantum discrete memoryless channel, a quantity called quantum entanglement capacity with classical feedback $E_B$. We will show that this quantity is less than the quantum capacity with two-way classical communication $Q_2$ and is greater than the quantum capacity with classical feedback $Q_B$.

### 3.1.1 Definition of $E_B$

Quantum entanglement capacity with classical feedback of a QDMC can be loosely described as the maximal asymptotic rate at which the sender Alice can share the entangled state $|\Phi^+\rangle \in \mathcal{H}_2^{\otimes 2}$ with the receiver Bob with the assistance of a classical feedback channel. Precisely, let the QDMC be described by

$$\begin{aligned} \mathcal{N} : \mathcal{B}(\mathcal{H}_{d_1}) &\longrightarrow \mathcal{B}(\mathcal{H}_{d_2}) \\ \rho \mapsto \rho' &= \sum_i E_i \rho E_i^\dagger, \end{aligned}$$

where $\sum_j E_j^\dagger E_j = I$ and $\{E_i\}$ is a set of linear operators which map the input Hilbert space $\mathcal{H}_{d_1}$ to the output Hilbert space $\mathcal{H}_{d_2}$. Then in the first round of any $E_B$ protocols, Alice prepares a quantum state $\alpha_1 = |\Upsilon\rangle\langle\Upsilon| \in \mathcal{B}(\mathcal{H}_{d_1}^{\otimes N} \otimes \mathcal{H}_a)$, where $\mathcal{H}_a$ is the Hilbert space representing the ancilla system in her laboratory and she sends the first part of the quantum state to Bob via the quantum channel $\mathcal{N}$:

$$\begin{aligned} \mathcal{N} : \mathcal{B}(\mathcal{H}_{d_1}) &\longrightarrow \mathcal{B}(\mathcal{H}_{d_2}) \\ \rho_1 = tr_{(d_1^{N-1}\times a)}(|\Upsilon\rangle\langle\Upsilon|) &\mapsto \rho_1' = \sum_i E_i \rho_1 E_i^\dagger. \end{aligned}$$

After sending $\rho_1$, Alice's quantum system is described by $\alpha_1' = tr_{d_1}(\alpha_1) \in \mathcal{B}(\mathcal{H}_{d_1}^{\otimes(N-1)} \otimes \mathcal{H}_a)$. On the other hand, Bob is now in possession of the quantum state $\rho_1'$ he just received from Alice as well as the ancilla system in his laboratory, and therefore his quantum system can be described by $\beta_1' = \rho_1' \otimes \beta_1 = \rho_1' \otimes |0\rangle\langle 0|^{\otimes \log_2 b} \in \mathcal{B}(\mathcal{H}_{d_2} \otimes \mathcal{H}_b)$. Next Bob performs local quantum operation on his quantum system:

$$\begin{aligned} \mathbf{B} : \mathcal{B}(\mathcal{H}_{d_2} \otimes \mathcal{H}_b) &\longrightarrow \mathcal{B}(\mathcal{H}_{d_2} \otimes \mathcal{H}_b) \\ \beta_1' &\mapsto \beta_1'' = \sum_i B_i \beta_1' B_i^\dagger \end{aligned}$$

where $\sum_i B_i^\dagger B_i = I$. Bob then uses the feedback channel to send classical information to Alice. Note that if Bob's operation comprised quantum measurements, this classical information could include the measurement results($i$). Upon learning the classical information sent by Bob, Alice's quantum system transforms from $\alpha_1'$ to $\alpha_{1,(i)}'$ and she performs operation on her quantum system:

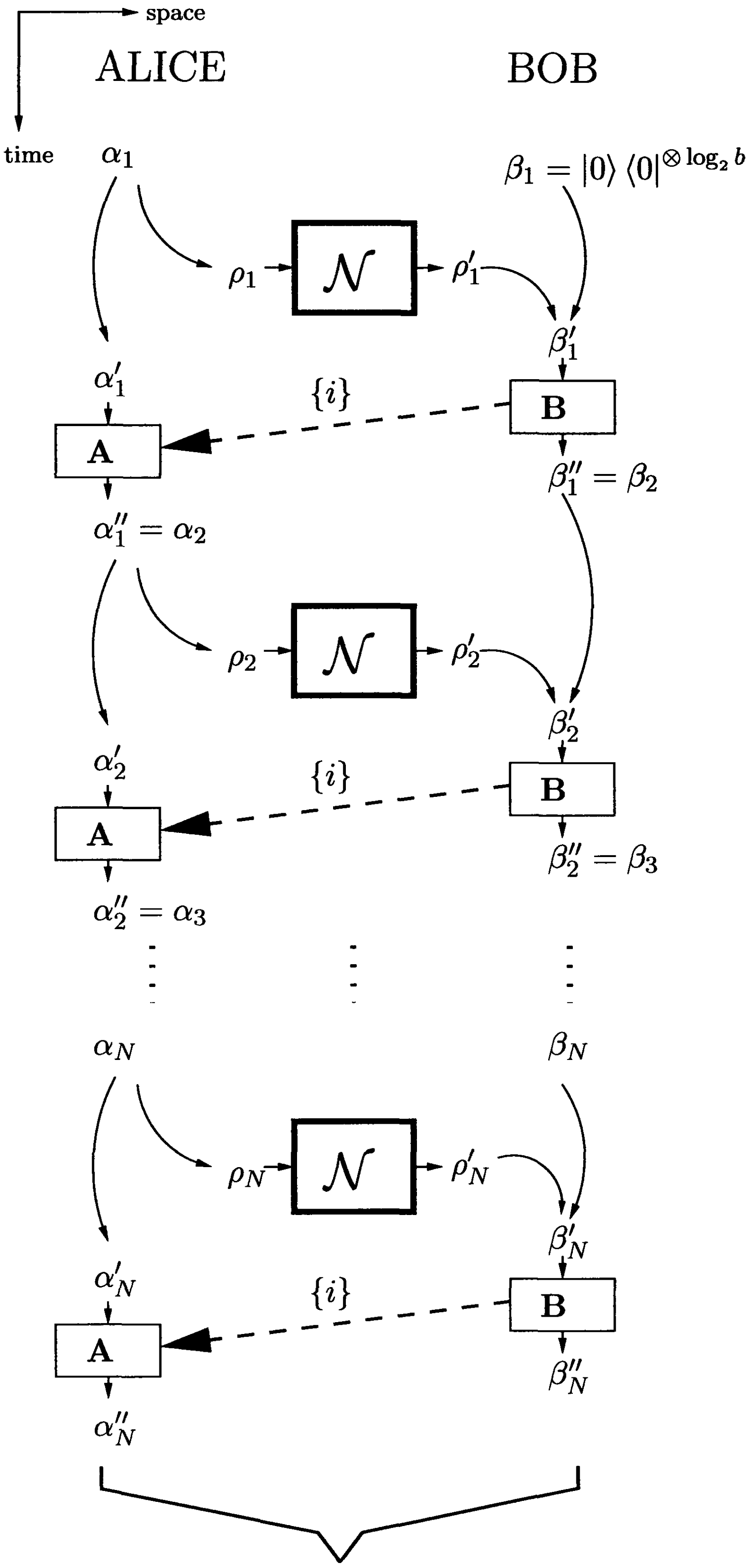
$$\begin{aligned} \mathbf{A_{(i)}} : \mathcal{B}(\mathcal{H}_{d_1}^{\otimes(N-1)} \otimes \mathcal{H}_a) &\longrightarrow \mathcal{B}(\mathcal{H}_{d_1}^{\otimes(N-1)} \otimes \mathcal{H}_a) \\ \alpha_{1,(i)}' &\mapsto \alpha_{1,(i)}'' = \sum_j A_{j,(i)} \alpha_{1,(i)}' A_{j,(i)}^\dagger. \end{aligned}$$

Note that both the quantum system $\alpha_{1,(i)}'$ and Alice's operation $\mathbf{A_{(i)}}$ are dependent on the classical information(i) she received from Bob. This is the end of the first round of any general $E_B$ protocols and can be summarized as:

$$\begin{aligned} \mathbf{LOCC_{A\leftarrow B}}^{(1)} \circ \mathbf{N}^{(1)} : \mathcal{B}(\mathcal{H}_{d_1}^{\otimes N} \otimes \mathcal{H}_a \otimes \mathcal{H}_b) &\longrightarrow \mathcal{B}(\mathcal{H}_{d_1}^{\otimes(N-1)} \otimes \mathcal{H}_a \otimes \mathcal{H}_{d_2} \otimes \mathcal{H}_b) \\ \omega_1 &\mapsto \omega_2. \end{aligned}$$

The second round of the protocols starts with Alice holding $\alpha_2 = tr_{(d_2 \times b)}(\omega_2)$ and Bob holding $\beta_2 = tr_{(d_1^{N-1} \times a)}(\omega_2)$. After $N$ rounds of protocols as seen in figure 3-1, we require the fidelity between the quantum state shared between Alice and Bob, $\omega_{N+1}$, and the quantum state, $(|\Phi^+\rangle \langle \Phi^+|)^{\otimes M}$, to approach 1 as $N$ goes to infinity. Then we define $E_B(\mathcal{N})$ to be the supremum of any attainable $\frac{M}{N \log_2(d_2)}$ - or simply $M/N$ if $d_2 = 2$.

Note that in this section, when we discuss an $E_B$ protocol, for brevity, we often say to compute the $E_B$ associated with the protocol rather than to compute the lower bounds on $E_B(\mathcal{N})$ impled by the protocol.



$$\omega_{N+1} \approx \left(|\Phi^+\rangle\langle\Phi^+|\right)^{\otimes M}$$

Figure 3-1: An $E_B$ protocols for channel $\mathcal{N}$(See the text for details).

### 3.1.2 $E_B \leq Q_2$

To show $E_B \leq Q_2$, we simply convert any $E_B$ protocol to a $Q_2$ protocol with the same rate. Suppose we have a protocol on $\mathcal{N}$ and this $E_B(\mathcal{N})$ protocol achieves $\frac{M}{N\log_2(d_2)}$, then at the end of this protocol Alice and Bob share the quantum state $|\Phi^+\rangle\langle\Phi^+|^{\otimes M}$. Alice now uses the forward classical communication channel to teleport any quantum state $\rho \in \mathcal{H}_2^{\otimes M}$ and therefore this new $Q_2(\mathcal{N})$ protocol achieves $\frac{M}{N\log_2(d_2)}$.

### 3.1.3 $Q_B \leq E_B$

This follows from the fact that $Q_B$ protocols are more restricted than $E_B$ protocols because in defining quantum capacities[2] the sender is required to not only transmit the quantum state $\rho$ but also preserve its entanglement with the environment to which neither the sender nor the receiver has access. In any $E_B$ protocols, the sender is required to transmit half of the maximally entangled states $|\Phi^+\rangle^M$ and is in possession of the other half which she can manipulate in her laboratory. Concisely, one can convert any $Q_B$ protocol to an $E_B$ protocol as follows: Alice prepares $|\Phi^+\rangle\langle\Phi^+|^{\otimes M} \in \mathcal{B}(\mathcal{H}_2^{\otimes M} \otimes \mathcal{H}_2^{\otimes M})$ in her laboratory and performs the $Q_B$ protocol on $\rho = (I/2)^{\otimes M} = tr_{(2^M)}(|\Phi^+\rangle\langle\Phi^+|^{\otimes M})$. At the end of the protocol, Alice and Bob share the bipartite quantum state $|\Phi^+\rangle\langle\Phi^+|^{\otimes M}$ and hence $E_B(\mathcal{N}) \geq Q_B(\mathcal{N}) \geq \frac{M}{N\log_2(d_2)}$.

### 3.1.4 $E_B$ as quantum backward capacity with classical feedback

In section 3.1.1, $E_B$ was defined as the maximal asymptotic rate at which Alice shares the singlet state $|\Phi^+\rangle$ with Bob with the assistance of a classical feedback channel. Alternatively, we can associate $E_B$ with a different operational meaning, namely the asymptotic rate at which Bob can send quantum states to Alice. This is because after any $E_B$ protocols Alice and Bob share the quantum states $(|\Phi^+\rangle\langle\Phi^+|)^M$ and there is a classical channel from Bob to Alice. Therefore, Bob can teleport any quantum states $\rho \in \mathcal{H}_M$ to Alice and this achieves the same yield $\frac{M}{N\log_2(d_2)}$ if we normalize by the dimension of the output Hilbert space or if we assume the input Hilbert space and the output Hilbert space are of the same size. Trivially, if Bob can send quantum states to Alice, Bob can choose to send half of the EPR pair $|\Phi^+\rangle$. Therefore these two notions are equivalent to one another.

## 3.2 Adaptive quantum error-correcting codes (AQECC)

In quantum error-correcting codes[29, 35, 42, 44], quantum states are encoded into the subspace of some larger Hilbert space. Although it has been discovered that quantum states can more generally be encoded into a subsystem rather than a subspace[1, 31], we focus only on subspace encoding. Our aim is to convert any quantum error-correcting codes (QECC) to new adaptive $E_B$ protocols on the quantum depolarizing channel $\mathcal{E}_p$. In section 3.2.1, we briefly review the stabilizer formalism; and in section 3.2.2 we introduce the idea of AQECC. In the rest of the section, we will illustrate with and compute the $E_B(\mathcal{E}_p)$ for two QECC, namely the Cat code and Shor code. We then consider how the recurrence method - a 2-EPP - in chapter 2 can be turned into an $E_B$ protocol. Finally we explain that the Leung-Shor method in chapter 2 is in fact an $E_B$ protocol.

### 3.2.1 Stabilizer formalism for QECC

We briefly review stabilizer formalism and introduce notation. A clear and detailed discussion can be found in [35]. $G_n$ denotes the Pauli group on $n$ qubits, and therefore consists of the n-fold tensor products of Pauli matrices. For example,

$$G_1 = \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}$$

where $X = \sigma_x$, $Y = \sigma_y$ and $Z = \sigma_z$. We use subscripts to denote the qubit that a Pauli matrix acts on. For example, $X_2Y_4$ means $I \otimes X \otimes I \otimes Y \otimes I \otimes \ldots \otimes I \in G_n$. Generators of a subgroup $S \subset G_n$ are independent if for any $i = 1, 2, 3, \ldots, n-k$,

$$< g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_{n-k} > \neq < g_1, \ldots, g_{n-k} > .$$

We say a vector space $V_S \subset \mathcal{H}_2^{\otimes n}$ is stabilized by a subgroup $S \subset G_n$ if for any $|\phi\rangle \in V_S$ and for any $s \in S$,

$$s\,|\phi\rangle = |\phi\rangle\,.$$

The following lemma can be shown easily:

**Lemma 3.** *Let $S =< g_1, \dots, g_{n-k} >$ be generated by $n-k$ independent and commuting elements from $G_n$, and $-I \notin S$. Then $V_S$ is a $2^k$-dimensional vector space.*

Therefore to specify a $2^k$-dimensional subspace for error-correcting codes, we only need to specify $n-k$ independent generators $g_1, \dots, g_{n-k}$. However we still need to specify the logical basis vectors $|x_1, \dots, x_k\rangle_L$ within $V_S$. In this thesis, we only deal with codes where $k = 1$. Therefore, it suffices to specify the logical $\bar{X}$ and logical $\bar{Z}$ such that $\bar{X}\,|0\rangle_L = |1\rangle_L \in \mathcal{H}_n$, $\bar{X}\,|1\rangle_L = |0\rangle_L \in \mathcal{H}_n$, $\bar{Z}\,|0\rangle_L = |0\rangle_L \in \mathcal{H}_n$ and $\bar{Z}\,|1\rangle_L = -\,|1\rangle_L \in \mathcal{H}_n$. Note that in doing so, we indirectly specify $|0\rangle_L$ and $|1\rangle_L$.

### 3.2.2 $E_B$ protocols via AQECC

Recall the aim of any $E_B$ protocols is for Alice to share the bipartite state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ with Bob. We will explain our idea of turning a QECC to an $E_B$ protocol in two steps.

The first step is to simply encode half of the EPR pair $|\Phi^+\rangle$ in an $[n, 1]$ stabilizer code, one that encodes a qubit in an n-dimensional Hilbert space $\mathcal{H}_n$. Alice performs the encoding

$$\begin{aligned} \mathbf{A} : \mathcal{B}(\mathcal{H}_2) &\longrightarrow \mathcal{B}(\mathcal{H}_2^{\otimes n}) \\ tr_2(|\Phi^+\rangle\,\langle\Phi^+|) &\mapsto \alpha_1 \end{aligned}$$

and then sends the $n$ qubits through the $p$-depolarizing channel

$$\begin{aligned}\mathcal{E}_p : \mathcal{B}(\mathcal{H}_2) &\longrightarrow \mathcal{B}(\mathcal{H}_2)\\ \rho &\mapsto \frac{1+3p}{4} \times \rho + \frac{1-p}{4} \times (\sigma_x \rho \sigma_x^\dagger + \sigma_y \rho \sigma_y^\dagger + \sigma_z \rho \sigma_z^\dagger).\end{aligned}$$

Since the error elements of the $p$-depolarizing channel are Pauli matrices, Alice can choose the logical basis states (or alternatively the logical operators $\bar{X}, \bar{Z}$ as we explained in the previous section) in such a way that after the error-correction operation **B**, the encoded qubit has either an $X$ error, a $Y$ error, a $Z$ error or no error. Since $X\left|\Phi^+\right\rangle = \left|\Psi^+\right\rangle$, $Y\left|\Phi^+\right\rangle = \left|\Psi^-\right\rangle$ and $Z\left|\Phi^+\right\rangle = \left|\Phi^-\right\rangle$, the bipartite state between Alice and Bob will be a probabilistic mixture of the four Bell states. Therefore Bob can use the classical feedback channel to perform universal hashing and distill perfect EPR pairs $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$. This first step is illustrated in figure 3-2.

The second step is to modify what has just been described so as to achieve a higher rate. Recall an $[n,1]$ stabilizer code is described by the generators of a subgroup $S =< g_1, g_2, \dots, g_{n-2}, g_{n-1} >$. The error-correcting operation **B** performed by Bob involves measuring the observables $g_1, g_2, \dots, g_{n-1}$ since they are all tenser products of Pauli matrices acting on n qubits. Note that, however, many of the $g_i$'s have identity action on all but a few qubits. For example, in 9-bit Shor code, $g_1 = Z_1 Z_2 (= Z_1 \otimes Z_2 \otimes I_3 \otimes I_4 \otimes I_5 \otimes I_6 \otimes I_7 \otimes I_8 \otimes I_9)$. Also, whenever a measurement result '-1' is obtained, it means some errors have occurred. In the case of Shor code, if Bob takes a measurement on the first two qubits immediately after he receives them from Alice and the measurement result is '-1', it is better for Bob to use the classical feedback channel to inform Alice that some errors have occurred in the first 2 qubits and they should give up this block of transmission and start all over. It is because the quantum state $\omega_{n+1}$ Alice and Bob obtained after n channel uses and decoding will be more mixed if some errors have occurred. It is thus more economical to not continue with this particular block of codes and give up the few qubits that have already been transmitted.

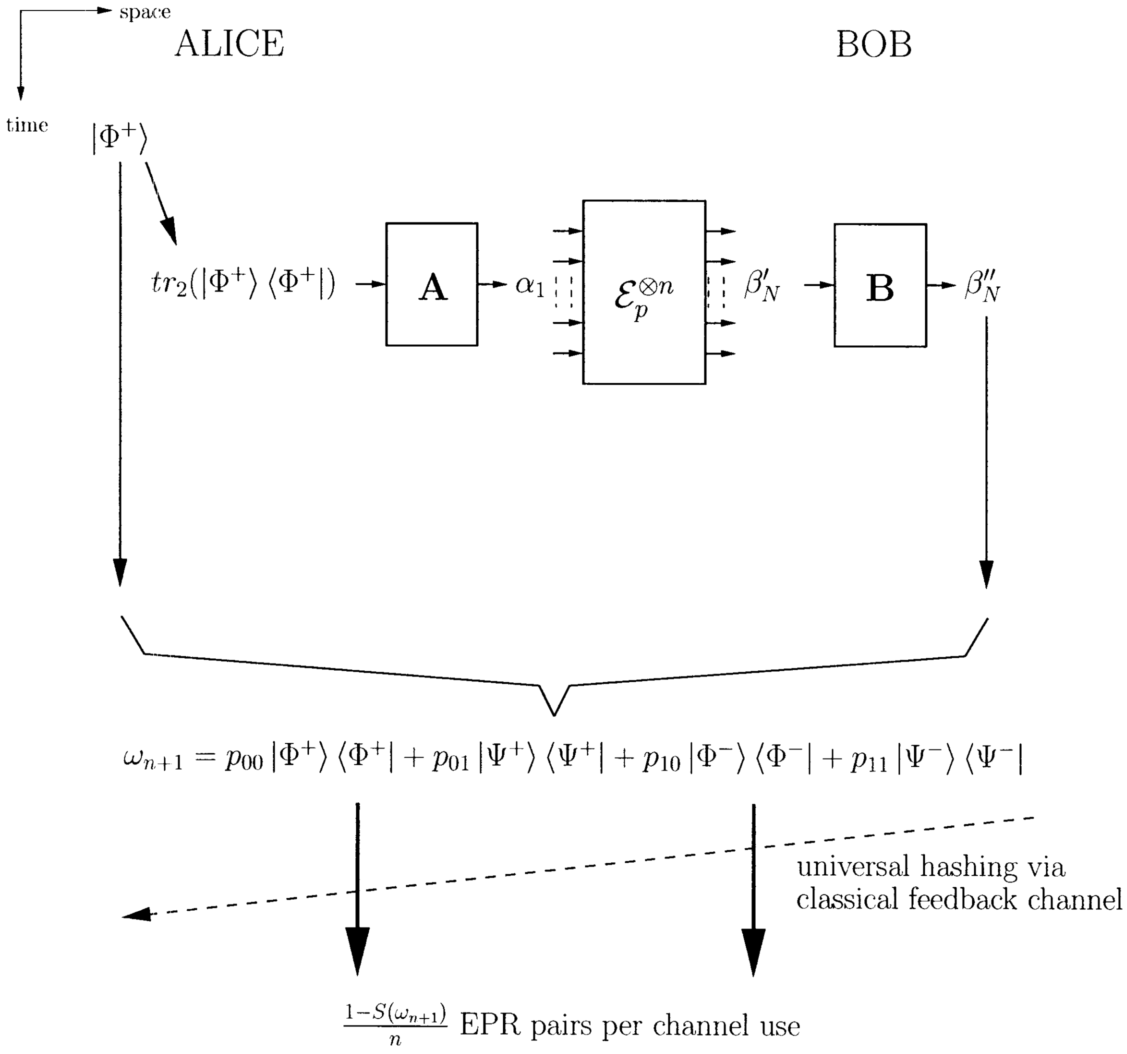

space
time
ALICE
BOB
$|\Phi^+\rangle$
$tr_2(|\Phi^+\rangle\langle\Phi^+|)$
A
$\alpha_1$
$\mathcal{E}_p^{\otimes n}$
$\beta'_N$
B
$\beta''_N$
$\omega_{n+1} = p_{00}\,|\Phi^+\rangle\langle\Phi^+| + p_{01}\,|\Psi^+\rangle\langle\Psi^+| + p_{10}\,|\Phi^-\rangle\langle\Phi^-| + p_{11}\,|\Psi^-\rangle\langle\Psi^-|$
universal hashing via
classical feedback channel
$\frac{1-S(\omega_{n+1})}{n}$ EPR pairs per channel use

Figure 3-2: Encoding half of the EPR pair $|\Phi^+\rangle$ with a QECC .

It is thus important to arrange the order of the measurements $g_1, g_2, \ldots, g_{n-1}$ such that it only involves as few more qubits as possible when one goes down the list. So that when an error is detected early on, Alice and Bob can stop the block and start all over so as to save more channel uses. For example, the generators of Shor code can be arranged as follows:

$$\begin{aligned}
g_1 &= Z_1 \otimes Z_2 \otimes I_3 \otimes I_4 \otimes I_5 \otimes I_6 \otimes I_7 \otimes I_8 \otimes I_9 \\
g_2 &= I_1 \otimes Z_2 \otimes Z_3 \otimes I_4 \otimes I_5 \otimes I_6 \otimes I_7 \otimes I_8 \otimes I_9 \\
g_3 &= I_1 \otimes I_2 \otimes I_3 \otimes Z_4 \otimes Z_5 \otimes I_6 \otimes I_7 \otimes I_8 \otimes I_9 \\
g_4 &= I_1 \otimes I_2 \otimes I_3 \otimes I_4 \otimes Z_5 \otimes Z_6 \otimes I_7 \otimes I_8 \otimes I_9 \\
g_5 &= X_1 \otimes X_2 \otimes X_3 \otimes X_4 \otimes X_5 \otimes X_6 \otimes I_7 \otimes I_8 \otimes I_9 \\
g_6 &= I_1 \otimes I_2 \otimes I_3 \otimes I_4 \otimes I_5 \otimes I_6 \otimes Z_7 \otimes Z_8 \otimes I_9 \\
g_7 &= I_1 \otimes I_2 \otimes I_3 \otimes I_4 \otimes I_5 \otimes I_6 \otimes I_7 \otimes Z_8 \otimes Z_9 \\
g_8 &= I_1 \otimes I_2 \otimes I_3 \otimes X_4 \otimes X_5 \otimes X_6 \otimes X_7 \otimes X_8 \otimes X_9 \qquad (3.1)
\end{aligned}$$

It is conceivable that after a large portion of the qubits in a block have been transmitted, it is better to continue even if an error is detected. It is indeed the case for Shor code when the probability parameter $p$ of the channel $\mathcal{E}_p$ is large. In the next two sections, we will apply this AQECC idea to Cat code and Shor code, and compute the lower bounds on $E_B(\mathcal{E}_p)$ implied by these codes.

### 3.2.3 Cat code and modified Cat code

The n-bit Cat code is an $[n, 1]$ stabilizer code with the following generators

$$\begin{aligned}
g_1 &= Z_1 Z_2\\
g_2 &= Z_2 Z_3\\
g_3 &= Z_3 Z_4\\
\vdots\ &\vdots\ \vdots\\
g_{n-2} &= Z_{n-2} Z_{n-1}\\
g_{n-1} &= Z_{n-1} Z_n
\end{aligned}$$

and we choose the following logical operators

$$\begin{aligned}
\bar{X} &= X_1 X_2 \ldots X_{n-1} X_n\\
\bar{Z} &= Z_1 Z_2 \ldots Z_{n-1} Z_n \text{ if n is odd and}\\
\bar{Z} &= Z_1 Z_2 \ldots Z_{n-1} I_n \text{ if n is even.}
\end{aligned}$$

This in turn determines the logical computational basis

$$|0\rangle_L = |00\ldots00\rangle \in \mathcal{H}_n \text{ and } |1\rangle_L = |11\ldots11\rangle \in \mathcal{H}_n.$$

Therefore, the singlet state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathcal{H}_2^{\otimes 2}$ is encoded as $\frac{1}{\sqrt{2}}(|00\ldots00\rangle + |11\ldots11\rangle) \in \mathcal{H}_2^{\otimes n+1}$ in Alice's laboratory. Alice will send the last n qubits to Bob via the channel $\mathcal{E}_p$. In accordance with the AQECC idea in the previous section, Alice sends the first two qubits first and Bob takes the measurement $g_1$. If the measurement result is '-1', Bob will inform Alice of the result via the classical feedback channel and Alice will discard the n-1 qubits remaining in her laboratory and start all over by encoding another EPR pair and sending the quantum states. If the measurement result is '+1', Bob will inform Alice of the result and

Alice will continue to send the third qubit. Bob will then measure $g_2$. This continues until all n qubits are passed to Bob and Bob gets '+1' in all n-1 measurements $g_1, g_2, \ldots, g_{n-1}$. Alice and Bob will then process a bipartite quantum state $\omega_{n+1} \equiv p_{00} |\Phi^+\rangle \langle\Phi^+| + p_{01} |\Psi^+\rangle \langle\Psi^+| + p_{10} |\Phi^-\rangle \langle\Phi^-| + p_{11} |\Psi^-\rangle \langle\Psi^-|$ that is Bell diagonal. If Alice and Bob repeat the process until they share N copies of $\omega_{n+1}$, i.e. $\omega_{n+1}^{\otimes N}$, they can perform universal hashing on these states and they will have $N\Big(1 - H(p_{00}, p_{01}, p_{10}, p_{11})\Big)$ EPR pairs $|\Phi^+\rangle$. However we are interested in the yield per channel use. Let $p_i$ = prob('+1' for measurement $g_i$). Then the average number of channel uses needed before we successfully pass a block of n-qubit Cat code through the depolarizing channel is given by

$$\begin{aligned} n^* &= \Big(\sum_{i=2}^{n-1} \Big(i \times (\prod_{j=1}^{i-2} p_j) \times (1 - p_{i-1})\Big) + n \times \prod_{i=1}^{n-2} p_i\Big) / \Big(n \times \prod_{i=1}^{n-1} p_i\Big) \\ &= \Big(2 \times (1 - p_1) + 3 \times p_1 \times (1 - p_2) + \ldots + (n-1) \times p_1 \times p_2 \times \ldots p_{n-3} \times (1 - p_{n-2}) \\ &\quad + n \times p_1 \times p_2 \times \ldots \times p_{n-2}\Big) / \Big(n \times p_1 \times \ldots \times p_{n-1}\Big). \end{aligned}$$

From this, the number of EPR pairs per channel use is

$$\begin{aligned} &\frac{1}{N \times n^* \times n} \times N\Big(1 - H(p_{00}, p_{01}, p_{10}, p_{11})\Big) \\ = &\frac{\Big(\prod_{i=1}^{n-1} p_i\Big) \times \Big(1 - H(p_{00}, p_{01}, p_{10}, p_{11})\Big)}{\Big(\sum_{i=2}^{n-1} \Big(i \times (\prod_{j=1}^{i-2} p_j) \times (1 - p_{i-1})\Big) + n \times \prod_{i=1}^{n-2} p_i\Big)}. \end{aligned} \tag{3.2}$$

We now present how to calculate the probabilities $p_1, \ldots, p_{n-1}$ and the quantum state $\omega_{n+1} = p_{00} |\Phi^+\rangle \langle\Phi^+| + p_{01} |\Psi^+\rangle \langle\Psi^+| + p_{10} |\Phi^-\rangle \langle\Phi^-| + p_{11} |\Psi^-\rangle \langle\Psi^-|$. The computation can be given by a simple recurrence relation [19, 43] which can be understood more easily in the language of entanglement purification protocols. Owing to the formal

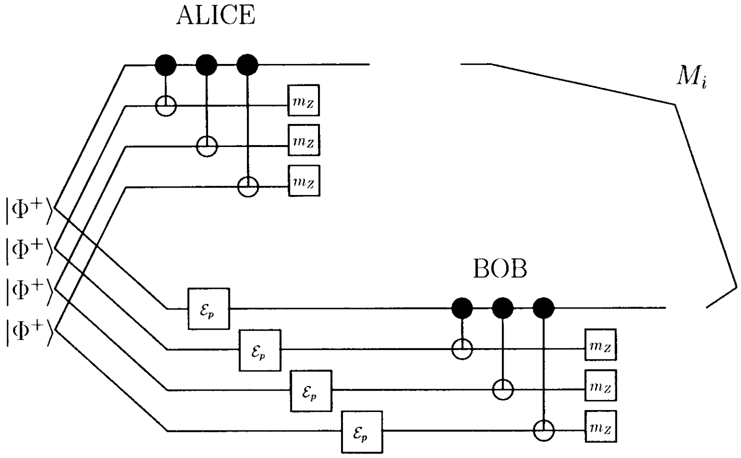


Figure 3-3: 4-bit Cat code in the language of entanglement purification protocols. Note that in our protocols if Bob's measurement results do not agree with Alice's, then not all qubits will be sent through $\mathcal{E}_p$. Alice's measurement results are assumed to be all '+1' so Alice need not send Bob any classical information even though Bob 'compares' his results against Alice's. See the text for details.

equivalence between the measurement of half of a Bell state and the preparation of a qubit, the encoding and decoding of the Cat code can be viewed as a 1-EPP as shown in figure 3-3 for $n = 4$. Note that in order for the purification protocols to work, it appears Alice has to send her measurement results to Bob via a side forward communication channel as in chapter 2. This is in fact not the case because even though the measurement results are non-deterministic, Alice can perform the measurements before she sends the 4 qubits (or generally n qubits). One can pretend Alice takes measurements for as many times as needed until she gets all '+1' before she sends the other halves of the quantum states via $\mathcal{E}_p$. Therefore Alice need not tell Bob the results because Bob already knew the results were all '+1'. (Of course, in reality, Alice can apply unitary operation in her laboratory to transform the states to what she needs even if the measurement results are '-1'.)

Note that applying a CNOT gates on the first and the (i-1)th qubits followed by measuring the (i-1)th qubit along the z-axis as shown in figure 3-3 is the same as measuring $g_i$, and we are interested in keeping track of the quantum state of the first qubit that passed through $\mathcal{E}_p$ after each measurement $g_i$. We are only interested in its quantum state if the measuring result is '+1', since we otherwise discard the states

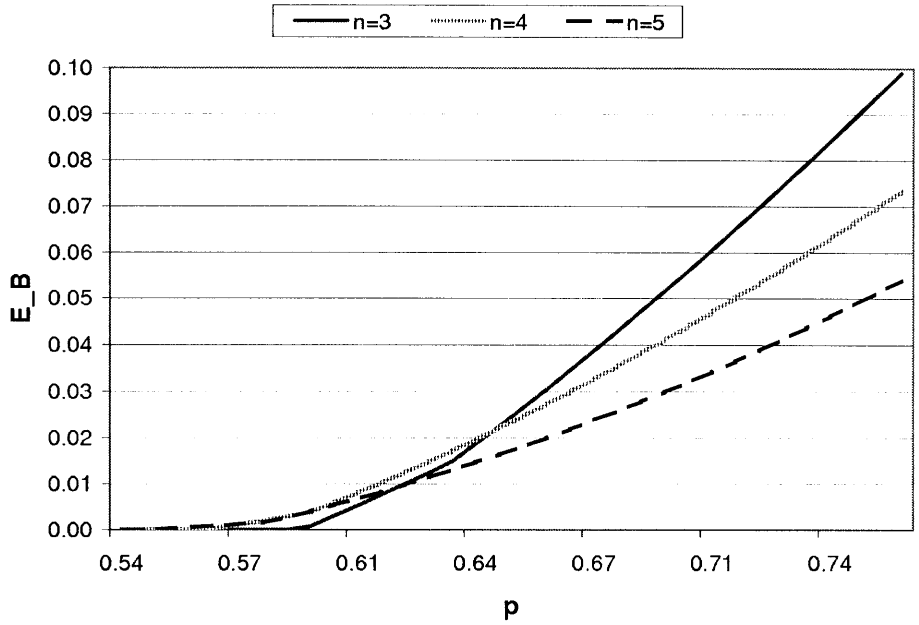


Figure 3-4: Lower bounds on $E_B(\mathcal{E}_p)$ via n-bit Cat code and modified Cat code. See the text for details.

and start all over. Denote this state by $M_i$, and we have the following relations [43] which follow from table 1.1:

$$p_{i+1} = (F+G)\langle\Phi^+|\,M_i\,|\Phi^+\rangle+(2G)\langle\Psi^+|\,M_i\,|\Psi^+\rangle+(F+G)\langle\Phi^-|\,M_i\,|\Phi^-\rangle+(2G)\langle\Psi^-|\,M_i\,|\Psi^-\rangle$$

$$\begin{aligned}\langle\Phi^+|\,M_{i+1}\,|\Phi^+\rangle &= \frac{F\langle\Phi^+|\,M_i\,|\Phi^+\rangle + G\langle\Phi^-|\,M_i\,|\Phi^-\rangle}{p_i}\\ \langle\Psi^+|\,M_{i+1}\,|\Psi^+\rangle &= \frac{G\langle\Psi^+|\,M_i\,|\Psi^+\rangle + G\langle\Psi^-|\,M_i\,|\Psi^-\rangle}{p_i}\\ \langle\Phi^-|\,M_{i+1}\,|\Phi^-\rangle &= \frac{G\langle\Phi^+|\,M_i\,|\Phi^+\rangle + F\langle\Phi^-|\,M_i\,|\Phi^-\rangle}{p_i}\\ \langle\Psi^-|\,M_{i+1}\,|\Psi^-\rangle &= \frac{G\langle\Psi^+|\,M_i\,|\Psi^+\rangle + G\langle\Psi^-|\,M_i\,|\Psi^-\rangle}{p_i}\end{aligned}$$

where $F = \frac{3p+1}{4}$, $G = \frac{1-F}{3}$ and $M_{n-1} = \omega_{n+1}$. From these equations and (3.2), we

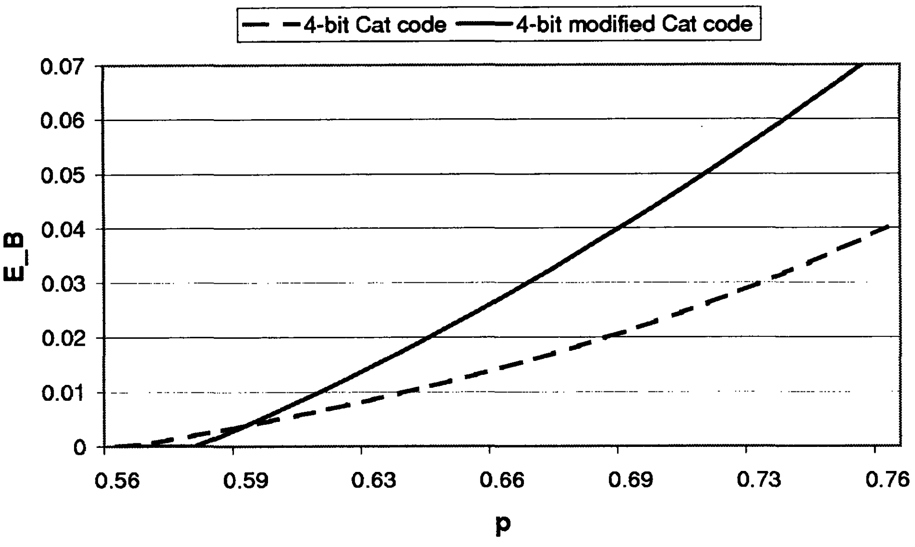


Figure 3-5: 4-bit Cat code vs. 4-bit modified Cat code.

compute the lower bounds on $E_B$ with n-bit Cat code and modified Cat code for $n = 3, 4, 5$ in figure 3-4. Modified Cat code differs from Cat code in the same way that the modified recurrence method differs from the recurrence method. Namely, Bob switches the $|\Phi^-\rangle\langle\Phi^-|$ and $|\Psi^-\rangle\langle\Psi^-|$ components in the probabilistic mixture of Bell states after each measurement. This can be done by first applying a bilateral $\pi/2$ rotation $B_x$ and then a unilateral $\pi$ rotation $\sigma_x$ [7]. Modified Cat code outperforms Cat code when the channel is less noisy(large p), but Cat code performs slightly better when the channel is very noisy and hence achieves a lower threshold value. In figure 3-5, we plot the yield for 4-bit Cat code and modified Cat code separately.

### 3.2.4 Shor code

The generators of Shor code are listed in (3.1). The logical operators and logical computational basis states are as follows:

$$\begin{aligned}\bar{X} &= Z_1Z_2Z_3Z_4Z_5Z_6Z_7Z_8Z_9\\ \bar{Z} &= X_1X_2X_3X_4X_5X_6X_7X_8X_9\end{aligned}$$

$$
\begin{aligned}
|0\rangle_L &= \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} \\
|1\rangle_L &= \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}.
\end{aligned}
$$

As aforementioned, for 9-bit Shor code, the optimal AQECC protocols are slightly different for different levels of noise. We can divide the protocols into 3 regions:

| p | protocol |
|---|---|
| less than 0.75 | start all over if any measurement result is '-1' |
| between 0.75 and 0.78 | start all over if any of the first 7 measurement results is '-1'; otherwise continue with the regular error-correcting operation |
| great than 0.78 | start all over if any of the first 4 measurement results is '-1'; otherwise continue with the regular error-correcting operation |

In the first region (p less than 0.75), one only has to enumerate all $4^9$ error possibilities in the 9 channel uses and adds up all probabilities associated with having an $X$ error, a $Y$ error, a $Z$ error or no error on the encoded qubit. Then the $E_B$ rate achieved for $\mathcal{E}_p$ is given by:

$$
\frac{p_1 \times p_2 \times \ldots \times p_8 \times \big(1 - H(p00, p01, p10, p11)\big)}{n^*}
$$

where

$$
\begin{aligned}
n^* = {} & 2 \times (1 - p_1) + 3 \times p_1(1 - p_2) + 5 \times p_1 p_2 (1 - p_3) + 6 \times p_1 p_2 p_3 (1 - p_4) \\
& +6 \times p_1 p_2 p_3 p_4 (1 - p_5) + 8 \times p_1 p_2 p_3 p_4 p_5 (1 - p_6) + 9 \times p_1 p_2 p_3 p_4 p_5 p_6.
\end{aligned}
$$

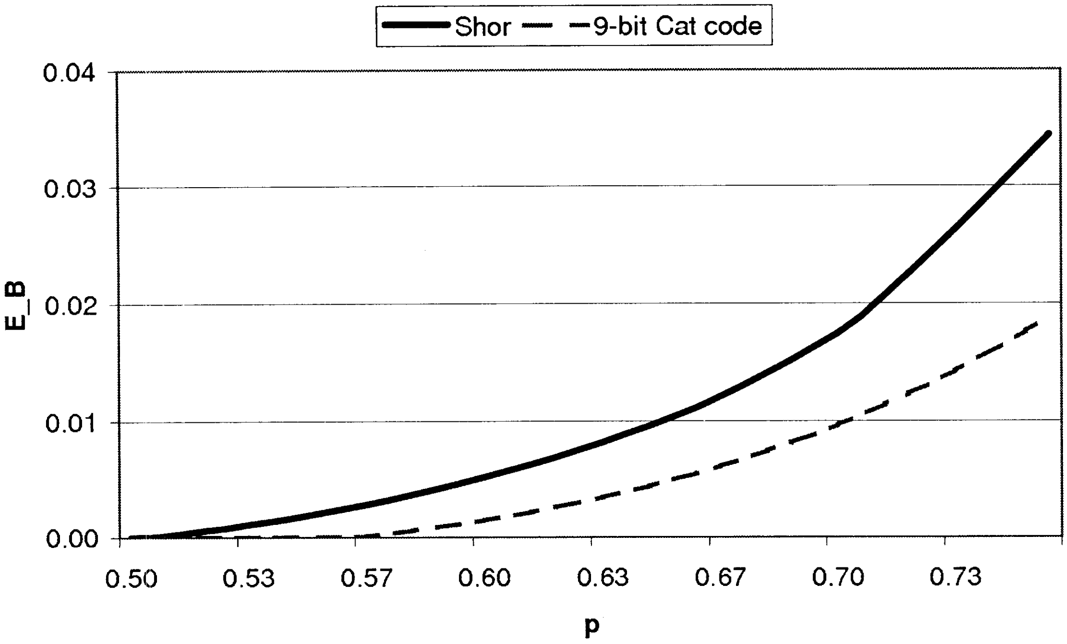


Figure 3-6: Lower bounds on $E_B(\mathcal{E}_p)$ via 9-bit Shor code and 9-bit Cat code.

In the second and the third region, the computation is slightly different. We will illustrate with the third region, and the computation for the second region is similar. Since Alice and Bob will start all over if any of the first 4 measurement results is '-1', there are only $2^{(8-4)} = 16$ possible measurement results given that the whole block of 9 qubits are sent through the channel. Denote the 4-tuple measurement results by $m \in \{0, 1, \ldots, 15\}$. For each measurement result, Bob will carry out error-correcting operation as in the standard 9-bit Shor code and inform Alice which of the 16 measurement results this block of 9 qubits has. Then after a large number of 9-bit blocks are transmitted successfully, Alice and Bob share a large number of each of the 16 types of Bell-diagonal probabilistic mixtures so that they can perform universal hashing on each of these 16 types of mixtures separately. And the $E_B$ rate achieved is given by

$$\sum_m \left( \frac{1}{n^{**}} \text{ prob(measurement result is m)} \left( 1 - H(p00, p01, p10, p11|m) \right) \right)$$

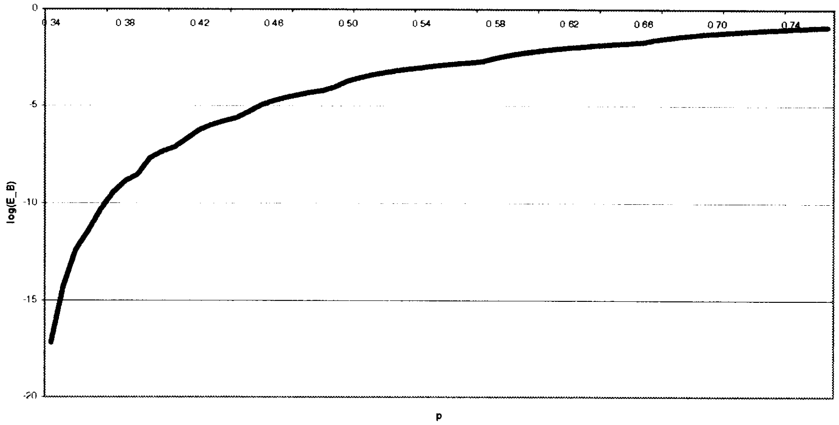


Figure 3-7: Lower bounds on $E_B(\mathcal{E}_p)$ via the modified recurrence method.

where $H(p00, p01, p10, p11|m)$ is the entropy of the probabilistic mixture given a particular measure result $m \in \{0, 1, \ldots, 15\}$ has occurred and $n^{**} = 2 \times (1 - p_1) + 3 \times p_1(1 - p_2) + 5 \times p_1p_2(1 - p_3) + 6 \times p_1p_2p_3(1 - p_4) + 9 \times p_1p_2p_3p_4$. In figure 3-6, we plot the $E_B$ rate achieved; for comparison $E_B$ rate achieved for 9-bit Cat code is also shown.

### 3.2.5 Modified recurrence method

Modified recurrence method[7] as described in chapter 2 is a 2-EPP which requires two-way classical communication. Although Alice can perform the measurement before she sends halves of the EPR pairs $|\Phi^+\rangle$ through $\mathcal{E}_p$ so that Bob need not know her measurement results in the first round, as we discussed in section 3.2.3 and 3.2.4, an iterative process is not possible. In particular, one round of recurrence plus universal hashing via the classical feedback channel achieve positive $E_B$ rate only for $p > 0.638$. If Alice and Bob want to carry out another round of the modified recurrence method, she needs a forward channel to communicate her measurement results to Bob. Since the only forward channel for Alice is $\mathcal{E}_p$, a straightforward extension, therefore, is to use the channel $\mathcal{E}_p$ to send her measurement results to Bob. As a result, from the second round onwards, one classical bit per pair is required for each round of recurrence.

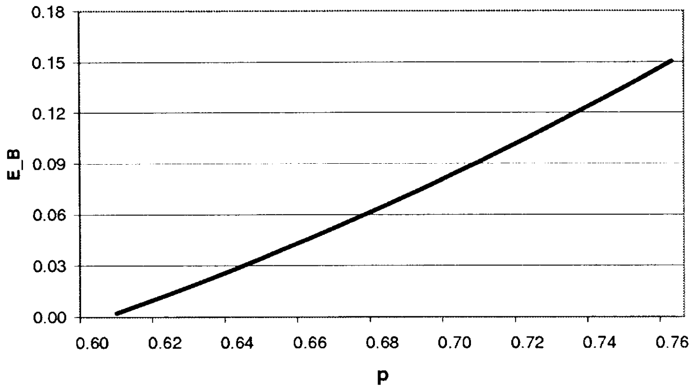


Figure 3-8: Lower bounds on $E_B(\mathcal{E}_p)$ via the Leung-Shor method.

By proving the additivity conjecture for the quantum depolarizing channel $\mathcal{E}_p$, the formula for the classical capacity of $\mathcal{E}_p$ is known[27]:

$$C(\mathcal{E}_p) = 1 + \left(\frac{1+p}{2}\right)\log_2\left(\frac{1+p}{2}\right) + \left(\frac{1-p}{2}\right)\log_2\left(\frac{1-p}{2}\right) = 1 - H\left(\frac{1+p}{2}, \frac{1-p}{2}\right).$$

Then the $E_B$ yield implied by this method for $k$ rounds of recurrence before switching to universal hashing is given by:

$$\left(\frac{p_{pass}^{(1)}}{2}\right) \times \left(\frac{p_{pass}^{(2)}}{2+1/C(\mathcal{E}_p)}\right) \times \cdots \left(\frac{p_{pass}^{(k)}}{2+1/C(\mathcal{E}_p)}\right) \times \left(1 - H\left(p_{00}^{(k)}, p_{01}^{(k)}, p_{10}^{(k)}, p_{11}^{(k)}\right)\right)$$

where $p_{00}^{(k)}$, $p_{01}^{(k)}$, $p_{10}^{(k)}$, $p_{11}^{(k)}$ and $p_{pass}^{(i)}$ for $i = 1, 2, \ldots, k$ are given by the recurrence relations (1.6) and (1.7) in section 1.3.2. In figure 3-7, we plot the $E_B$ rate achieved by this method.

### 3.2.6 Leung-Shor method

The method introduced in section 2.1 is in fact an $E_B$ protocol. Alice only needs to encode the qubits into what they would have been if the measurement results in figure 2-1 were both '+1'. In figure 3-8, we plot the $E_B$ rate achieved. In figure 3-9, we compare the yield of the four methods in this section.

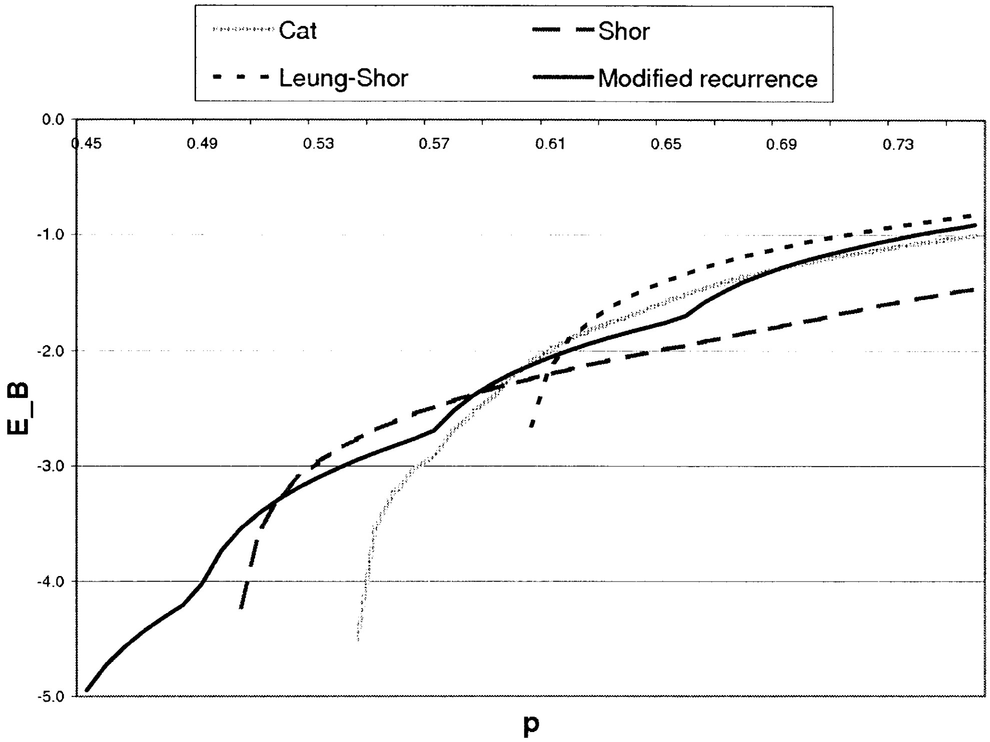


Figure 3-9: Lower bounds on $E_B(\mathcal{E}_p)$.

## 3.3 New lower bounds on $Q_B$

We will establish the following lemma which gives lower bounds on $Q_B$ based on $E_B$ protocols:

**Lemma 4.**

$$Q_B(\mathcal{E}_p) \geq \frac{1}{1 + \frac{C(\mathcal{E}_p)}{E_B(\mathcal{E}_p)}}$$

*where* $C(\mathcal{E}_p) = 1 - H\left(\frac{1+p}{2}, \frac{1-p}{2}\right)$.

*Proof.* In an $E_B$ protocol, Alice and Bob share M EPR pairs $|\Phi^+\rangle$ in N channel uses. Therefore, $E_B(\mathcal{E}_B) = M/N$. To teleport a quantum state $\rho \in \mathcal{H}_2^{\otimes M}$, Alice can use

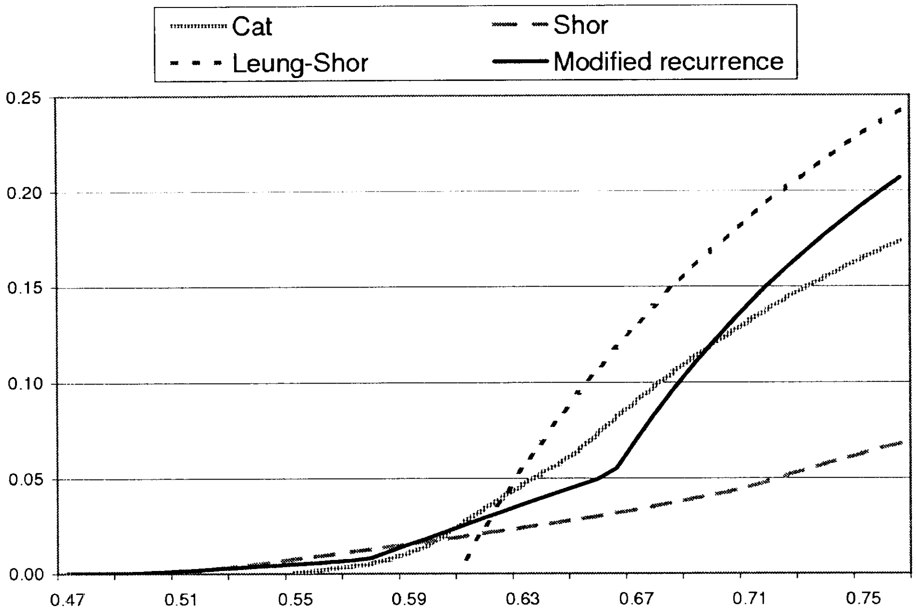


Figure 3-10: Lower bounds on $Q_B(\mathcal{E}_p)$.

the channel $\mathcal{E}_p$ for $\frac{M}{C(\mathcal{E}_p)}$ many times to send M bits of classical information to Bob. Thus,

$$\begin{aligned} Q_B(\mathcal{E}_p) &\geq \frac{M}{N + \frac{M}{C(\mathcal{E}_p)}} = \frac{M/N}{1 + \frac{M/N}{C(\mathcal{E}_p)}} \\ &= \frac{E_B(\mathcal{E}_p)}{1 + \frac{E_B(\mathcal{E}_p)}{C(\mathcal{E}_p)}} = \frac{1}{1 + \frac{C(\mathcal{E}_p)}{E_B(\mathcal{E}_p)}}. \end{aligned}$$

□

From the lemma, any lower bounds on $E_B$ will imply lower bounds on $Q_B$. The lower bounds are presented in figure 3-10.

## 3.4 Threshold of Cat code

It has been shown that in the absence of side classical communication one can achieve non-zero capacity for lower threshold fidelity $F = \frac{3p+1}{4}$ by concatenating 5-bit Cat

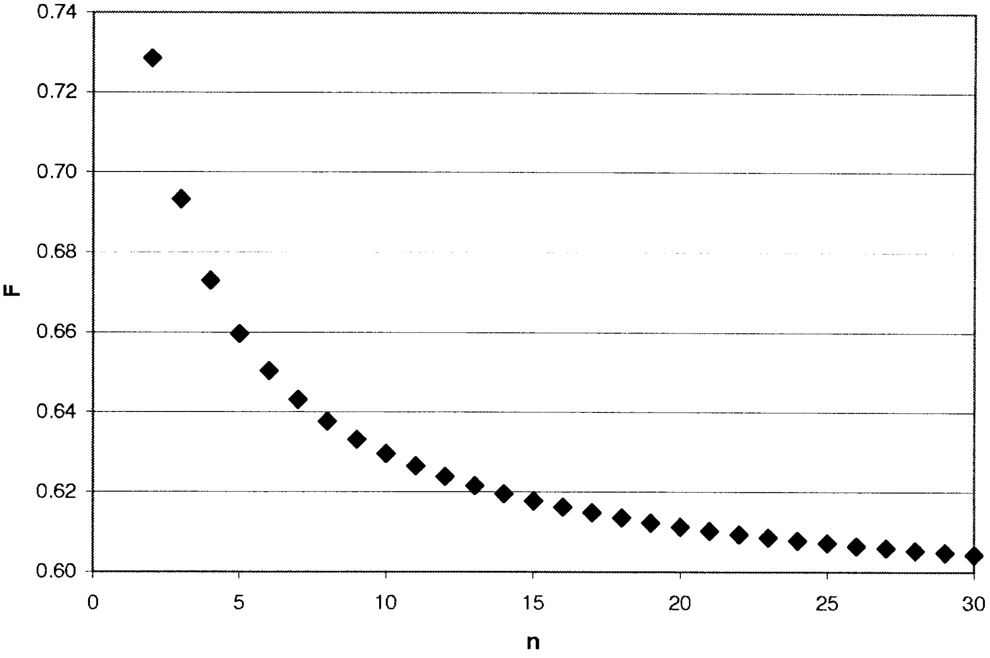


Figure 3-11: Threshold fidelity $F = \frac{3p+1}{4}$ for n-bit Cat code.

code inside a random code (hashing)[19]. Threshold fidelity for concatenating n-bit Cat code into random code was also studied. It was found that threshold fidelities fall into two smooth curves, one for even n and one for odd n, but both curves increase with n, i.e. one does not attain lower threshold by using a longer Cat code. We therefore compute the threshold fidelity for n-bit Cat code in figure 3-11 and we found that these phenomena do not occur in AQECC.

## 3.5 Discussion on $Q_B$, $E_B$ and $Q_2$

In this chapter, we define the quantum entanglement capacity with classical feedback $E_B$ for any quantum discrete memoryless channel. For any channel, this quantity is shown to lie between two other capacities, namely the quantum capacity with classical feedback $Q_B$ and the quantum capacity with two-way classical communication $Q_2$. It is an open question whether these two capacities are equal to one another. While the introduction of this new, intermediate quantity $E_B$ does not simplify the question, it is our hope to shed some light on and provide other means to tackle this open problem. In section 3.1, we provide an alternate operational interpretation of this quantity: it represents the amount of quantum information Bob can send to Alice. It is our hope that, by working with this interpretation, one might be able to prove a non-trivial upper bound on $E_B$ and hence lead to a separation between $Q_B$ and $Q_2$.

We turn many of the well-known QECC into $E_B$ protocols and compute their yields. These in turn lead to new lower bounds on $Q_B$. The QECC that we studied, namely Cat code and Shor code, exhibit different behaviors under this AQECC framework. For example, for Shor code, it is beneficial to not insist on getting no error in all measurements but instead carry out error-correcting procedures after getting no error in the first few measurements. Whereas for Cat code, one has to insist on getting no error in all measurements. It is interesting to study which of these two features is exhibited by other codes.

We also see some connections with 2-EPP. Firstly, even though the Leung-Shor method was introduced in chapter 2 as a 2-EPP, it is in fact an $E_B$ protocol. Secondly, the idea that modified recurrence method applies to Cat code and achieves higher yields.

Finally, one may want to ask whether the threshold fidelity in section 3.4 goes down monotonically and if it does, what value it converges to as n goes to infinity.

# Bibliography

[1] D. Bacon, Operator quantum error-correcting subsystems for self-correcting quantum memories, quant-ph/05-060-23

[2] H. Barnum, E. Knill and M.A. Nielsen, On Quantum Fidelities and Channel Capacities, IEEE Trans. Inform. Theory, vol. 46, p.1317 to 1329, 2000, quant-ph/98-09-010

[3] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. Wootters, Teleporting an unknown quantum state via dual classical and EPR channels, Phys. Rev. Lett., 70, pp. 1895-1899 (1993)

[4] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin and W.K. Wooters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, Phys. Rev. Lett., 76, pp. 722-725 (1996)

[5] C.H. Bennett, I. Devetak, P.W. Shor and J.A. Smolin, Inequalities and Separations among Assisted Capacities of Quantum Channels, quant-ph/04-06-086

[6] C.H. Bennett, D.P. DiVincenzo and J.A. Smolin, Capacities of Quantum Erasure Channels, quant-ph/97-01-015

[7] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin and W.K. Wootters, Mixed State Entanglement and Quantum Error Correction, Phys. Rev. A, 54, pp. 3824-3851 (1996), quant-ph/96-04-024

[8] C.H. Bennett and P.W. Shor, Quantum Information Theory, IEEE Trans. Inform. Theory, vol. 44, p.2724 to 2742, 1998

[9] C.H. Bennett, P.W. Shor, J.A. Smolin and A.V. Thapliyal, Entanglement-Assisted Classical Capacity of Noisy Quantum Channels, Phys. Rev. Lett., vol. 83, p.3081 to 3084, 1999, quant-ph/99-04-023

[10] C.H. Bennett, P.W. Shor, J.A. Smolin and A.V. Thapliyal, Entanglement-Assisted Capacity of a Quantum Channel and the Reverse Shanno Theorem, quant-ph/01-06-052

[11] C.H. Bennett and S.J. Wiesner, Communication via one- and two-particale operators on Einstein-Podolsky-Rsen states, Phys. Rev. Lett., 69, pp. 2881-2884 (1992)

[12] G. Bowen, Quantum Feedback Channels, IEEE Trans. Inform. Theory, vol.50, p.2429 to 2433, 2004, quant-ph/02-09-076

[13] G. Bowen and R. Nagarajan, On Feedback and the Classical Capacity of a Noisy Quantum Channel, IEEE Trans. Inform. Theory, vol.51, p.320 to 324, 2005, quant-ph/03-05-176

[14] D. Bruss, Characterizing Entanglement, quant-ph/01-10-078

[15] C.M. Caves and C.A. Fuchs, Quantum Information: How much information in a state vector, in The Dilemma of Einstein, Podolsky and Rosen - 60 Years Later (An International Symposium in Honour of Nathan Rosen - Haifa, March 1995), edited by A. Mann and M. Revzen, Annals of The Israel Physical Society, vol. 12, p.226 - 257, 1996, quant-ph/96-01-025

[16] T.M. Cover and J.A. Thomas, Elements of Information Theory (John Wiley and Sons, New York, 1991)

[17] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, to appear in IEEE Trans. Inform. Theory, quant-ph/03-04-127

[18] I. Devetak, A.W. Harrow and A. Winter, A family of quantum protocols, quant-ph/03-08-044

[19] D.P. DiVincenzo, P.W. Shor and J.A. Smolin, Quantum channel capapcity of very noisy channels, quant-ph/97-06-061

[20] M.J. Donald, M. Horodecki and B.M. Terhal, The asymptotic entanglement cost of preparing a quantum state, J.Phys.A: Math. Gen., vol. 34, no. 35, pp. 6891-6898 (2001), quant-ph/01-05-017

[21] W. Dür, G. Vidal and J.I. Cirac, Three qubits can be entangled in two inequivalent ways, quant-ph/00-05-115

[22] L. Grover, A fast quantum mechanical algorithm for database search, Proceddings of the 28th Annual ACM Symposium on Theory of Computing, pp. 212-219 (1996)

[23] A. Harrow, P. Hayden and D. Leung, Superdense coding of quantum states, quant-ph/03-07-221

[24] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland and W.K. Wootters, Classical information capacity of a quantum channel, Phys. Rev. A, vol. 54, p.1869 to 1876, 1996

[25] E. Hostens, J. Dehaene and B.D. Moor, Asymptotic adaptive bipartite entanglement distillation protocol, quant-ph/0602205

[26] C. King, Additivity for unital qubit channels, quant-ph/01-03-156

[27] C. King, The capacity of the quantum depolarizing channel, quant-ph/02-04-172

[28] C. King, An application of a matrix inequality in quantum information theory, quant-ph/04-12-046

[29] E. Knill and R. Laflamme, Theory of quantum error-correcting codes, Phys. Rev. A, vol. 55, pp. 900 to 911, 1997

[30] K. Kraus, States, Effects and Operations: Fundamental Notions of Quantum Theory, Lecture Notes in Physics, vol. 190 (Springer-Verlag, Berlin, 1983)

[31] D.W. Kribs, R. Laflamme, D. Poulin and M. Leosky, Operator quantum error correction, quant-ph/05-04-189

[32] A.W. Leung and P.W. Shor, Entanglement Purification with Two-way Classical Communication, quant-ph/07-02-155

[33] A.W. Leung and P.W. Shor, Adaptive Entanglement Purification Protocols with Two-way Classical Communication, quant-ph/07-02-156

[34] E.N. Maneva and J.A. Smolin, Improved two-party and multi-party purification protocols, quant-ph/00-03-099

[35] M.A. Nielsen and I.L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2000)

[36] B.W. Schumacher, Sending entanglement through noisy quantum channels, Phys. Rev. A, vol. 54, p.2614 to 2628, 1996

[37] B. Schumacher and M.D. Westmoreland, Sending classical information via noisy quantum channels, Phys. Rev. A, vol. 54, p.2629 to 2635, 1996

[38] C.E. Shannon, A mathematical theory of communication, The Bell System Tech. J., vol. 27, p. 379 to 423, 623 to 656, 1948

[39] P.W. Shor, Equivalence of Additivity Questions in Quantum Information Theory, quant-ph/03-05-035

[40] P.W. Shor, The Classical Capacity Achievable by a Quantum Channel Assisted by Limited Entanglement, quant-ph/04-02-129

[41] P.W. Shor, Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, SIAM J. Comput., 26, pp. 1484-1509 (1997)

[42] P.W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A, vol. 52, pp.2493, 1995

[43] P.W. Shor and J.A. Smolin, Quantum error-correcting codes need not completely reveal the error syndrome, quant-ph/96-04-006

[44] A.M. Steane, Error correcting codes in quantum theory, Phys. Rev. Lett., vol. 77, pp.793, 1996

[45] G. Vidal, Entanglement monotones, J. Mod. Opt., vol. 47, pp. 355 (2000), quant-ph/98-07-077

[46] K.G.H. Volbrecht and Frank Verstraete, Interpolation of recurrence and hashing entanglement distillation protocols, quant-ph/0404111

[47] R.F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A, 40, pp. 4277-4281 (1989)