Extracting Quantum Entanglement
(General Entanglement Purification Protocols)

Andris Ambainis*
Institute for Advanced Study
ambainis@ias.edu

Adam Smith†
MIT
adsmith@mit.edu

Ke Yang‡
Carnegie Mellon University
yangke@cs.cmu.edu

Abstract

We study the problem of extracting EPR pairs from a general source of entanglement. Suppose Alice and Bob share a bipartite state ρ which is "reasonably close" to perfect EPR pairs. The only information Alice and Bob possess is a lower bound on the fidelity of ρ and a maximally entangled state. They wish to "purify" ρ using local operations and classical communication and output a state that is arbitrarily close to EPR pairs. We prove that on average, Alice and Bob cannot increase the fidelity of the input state significantly. On the other hand, there exist protocols that may fail with a small probability, and otherwise will output states arbitrarily close to EPR pairs with very high probability. These protocols come from the "purity-testing protocols" of Barnum et al [2].

1 Introduction

Random bits are an important computational resource in randomized computation. There has been a lot of work on extracting good random bits from imperfect sources of randomness. Von Neumann [16] showed that a linear number of perfect random bits can be extracted from independent tosses of a biased coin. More recent research has constructed extractors [19, 23] which can extract almost perfect random bits from any source with a certain min-entropy, without any other assumptions. The best constructions extract a number of random bits close to the min-entropy of the random source, given a polylogarithmic number of perfect random bits as auxiliary input [22].

Quantum entanglement is an important resource in quantum computation, similar to random bits in probabilistic computation. It comes in the form of Einstein-Podolsky-Rosen [8] (EPR) pairs. An EPR pair is the state of two quantum bits $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ shared by two parties, with one party (Alice) holding one quantum bit and the other party (Bob) holding the second bit. This is the quantum counterpart of a random bit shared by two parties.

Besides being conceptually interesting in quantum mechanics, EPR pairs are also very useful in quantum information theory. Using an EPR pair, Alice and Bob can perform quantum teleportation. By performing only local operations and classical communication (LOCC), Alice can "transport" a qubit to Bob, who could be miles away from Alice [3]. So EPR pairs, along with a classical communication channel, effectively constitute a quantum channel. Conversely, "superdense coding" is possible with EPR pairs: if Alice and Bob share an EPR pair, then Alice can transport 2 classical bits to Bob by just sending one qubit [7].

For the teleportation and dense coding to work perfectly, perfect EPR pairs are needed. Individual qubits are prone to errors, which make for imperfect pairs. This creates the need for generating perfect (or almost perfect) EPR pairs from imperfect ones. This problem of extracting EPR pairs is sometimes known as "entanglement distillation" or "entanglement purification". It has been the focus of much research; we list the most relevant works here.

Bennett et. al. [4] gave a protocol for the case that Alice and Bob share identical copies of the pure state $|\phi\rangle = (\cos\theta|01\rangle + \sin\theta|10\rangle)$. This was extended to the case when Alice and Bob share identical copies of a mixed state [5, 6, 11]. Vidal [24], and subsequently, Jonathan and Plenio [12], Hardy [10], and Vidal, Jonathan, and Nielsen [25] considered extracting entanglement from a single copy of an arbitrary pure state, assuming that we know a complete description of the state. All these works use relatively simple models for imperfect EPR pairs, and have counterparts in the classical problem of randomness extraction. The model where Alice and Bob share identical copies of the same state corresponds to generating perfect random bits from the sequence of i.i.d. biased coin flips. Extracting entanglement from a single copy of a known state corresponds to con-

*Supported by NSF Grant CCR-9987845 and the State of New Jersey.

†Supported by the US DoD MURI program administered by the Army Research Office under grant DAAD19-00-1-0177.

‡Partially supported by the CMU SCS Alumni Fellowship and the NSF Aladdin center, Grant CCR-0122581.

structing uniform/almost uniform random bits from a biased distribution when we know a complete description of the distribution. Both of those are very easy tasks classically. Dealing with quantum states makes them much harder, but their applicability remains limited.

More general error models were considered in quantum cryptography. In the context of quantum key distribution, Lo and Chau [14], and later Shor and Preskill [21], considered a problem similar to entanglement extraction. They proved that there exist "testing protocols" involving only LOCC, such that perfect EPR pairs will "pass" the protocol with certainty, while any state far from EPR pairs "fail" with very high probability. This allows one to conclude that a state that passes a verification protocol will have very high fidelity with the perfect EPR pairs. This *testing* problem was formulated more explicitly by Barnum et al. [2], in the setting of authenticating quantum messages. They also give a more efficient construction. In this paper, we consider the problem of *extraction* with respect to general errors. Nonetheles, the approach and results of [14, 21, 2] will prove very useful.

This current understanding of quantum entanglement extraction is quite different from that of classical randomness extraction, where an extractor works with a very general model of randomness: so long as the random source has an certain min entropy, the extractor is guaranteed to extract almost perfect random bits, regardless of the input distribution. Moreover, extractors can be made very efficient: the number of perfect random bits they output is almost the min entropy of their input, and only a logarithmic number of perfect random bits are invested as auxiliary input[1]. Can we have an "entanglement extraction protocol" in the quantum world that matches the extractor in the classical world? In other words, can we have a protocol that works regardless of the input state, and can our protocol be made efficient, namely, output as many near-perfect EPR pairs as possible and invest as few perfect EPR pairs as possible? These are the questions we set out to consider in this paper.

1.1 Models of Imperfect EPR Pairs

We use the following model for the imperfect EPR pairs in our paper: Alice and Bob start by sharing a state of n perfect EPR pairs, and then some "distortion operator" $\mathcal{D}$ is applied to the state. This operator $\mathcal{D}$ isn't necessarily a unitary operator and thus the state Alice and Bob end up with could be a mixed state. The only assumption that we have is that the distortion is not "very large". More precisely, we assume that Alice and Bob share a state ρ with fidelity[2] at least $(1-\epsilon)$. We call this model of imperfect EPR pairs the "General Error" model. We call the protocols for this model General Entanglement Purification Protocols (GEPPs)

As the readers might have noticed, our model of imperfect EPR pairs doesn't look similar to the model an extractor uses. An extractor works with *any* random source with enough entropy, and thus it is tempting to require an entanglement extraction protocol to work with *any* bipartite state with enough entanglement. Ideally, we would like a protocol that takes any bipartite system that has a certain amount of entanglement and outputs some near-perfect EPR pairs. However, our model doesn't allow this — we only require our protocol to work with states that are "close" to perfect EPR pairs. This "closeness" condition seems to be a serious constraint. Nevertheless, as we will show later in this paper, this constraint is necessary: there simply don't exist protocols that will product near-perfect EPR pairs on any input of certain entanglement via LOCC. This situation is very different from that of randomness extraction.

1.2 Our Contribution

We consider 3 types of GEPPs. Roughly speaking, a GEPP is *absolutely successful* (AS), if it never fails, and always outputs a state of very high fidelity. A GEPP is *conditionally successful* (CS), if the probability it fails is small, and when it doesn't fail, it outputs a state of very high *expected* fidelity. A *deterministic conditionally successful* (DCS) GEPP, on the other hand, outputs a state of high fidelity with probability 1, conditioned on not failing.

1. There do not exist absolutely successful GEPPs with "interesting" parameters. More precisely: Suppose Alice and Bob share a state of fidelity $1-\epsilon$ to Ψ_N, a maximally entangled state of dimension N. Suppose they also have an auxiliary input Ψ_K. They then perform LOCC to create a state σ in a subspace of dimension $M \times M$. Then the maximal fidelity of the state σ which Alice and Bob can guarantee is $1 - \frac{N}{N-1}(1 - \frac{K}{M})\epsilon$. If K is significantly smaller than M improvement of fidelity is very small. In other words, Alice and Bob cannot arbitrarily increase the *average* fidelity of the input state.

2. Good CS and good DCS protocols exist. In fact, they are closely linked to the purity-testing schemes of Barnum et al. [2]. Informally, a purity testing protocol is an LOCC protocol where the input is joint state shared by Alice and Bob which they think might be the EPR state $|\Phi^{(n)}\rangle = |\Phi^{+}\rangle^{\otimes n}$. Alice and Bob want to test if their

[1] Furthermore, it is proven [20] that it is optimal: the number of random bits output by an extractor cannot exceed the min entropy of the input distribution, and the number of perfect random bits an extractor invests must be at least logarithmic in the input length.

[2] Most of the time in this paper, we are interested in the fidelity of a state ρ and a (pre-defined) maximally entangled state (e.g., an EPR pair). In this case, we simply use the "fidelity of state ρ" to denote the fidelity of ρ and the pre-defined maximally entangled state of the appropriate dimension.

shared state is indeed $|\,\Phi^{(n)}\rangle$, while sacrificing the least number of EPR pairs. We observe that any purity-testing scheme immediately yields a GEPP. In particular, this will yield a CS protocol that, on any state of n imperfect EPR pairs with fidelity $1-\epsilon$, will fail with probability at most ϵ. When the protocol doesn't fail, it outputs $(n-s)$ near-perfect EPR pairs of expected fidelity at least $1-2^{-s}/(1-\epsilon)$. So as we increase s, we can get EPR pairs that are arbitrarily close to perfect. This protocol doesn't need any perfect EPR pairs as auxiliary input. Using the construction of [2], we also obtain a DCS protocol that requires $(2s+1)$ additional perfect EPR pairs and outputs a state of fidelity $1-2^{-s+\log n}/(1-\epsilon)$ with probability 1, conditioned on not failing.

2 Notations and Definitions

2.1 General Notations

All logarithms are base-2.

We study quantum systems of finite dimension. We identify a pure state (written in the "ket" notation as $|\,\phi\rangle$) with a (column) vector of unit length. We identify a mixed state with its density matrix. For a quantum system whose states lie in the Hilbert space $\mathcal{H}$ of dimension N, we always assume that it has a canonical computational basis and denote it by $\{|\,0\rangle, |\,1\rangle, ..., |\,N-1\rangle\}$. Furthermore, we often denote $|\,0\rangle \in \mathcal{H}$ by $|\,Z_N\rangle$ to specify the dimension of the Hilbert space.

We are interested in *symmetric, bipartite quantum systems*, namely, systems shared between Alice and Bob, whose states lie in a Hilbert space $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ and $\mathcal{H}^A \equiv \mathcal{H}^B$. Alice can access $\mathcal{H}^A$ and Bob can access $\mathcal{H}^B$. We superscript subspaces and states to distinguish states accessible by Alice and Bob. For example, a general bipartite state $|\,\varphi\rangle$ can written in the following way:

$$|\,\varphi\rangle = \sum_{i,j} \alpha_{ij} |\,i\rangle^A |\,j\rangle^B$$

where $|\,i\rangle^A$ denotes the state of Alice and $|\,j\rangle^B$ denotes the state of Bob. We sometimes subscript a space by its dimension. For example, $\mathcal{H}_N$ means a space of dimension N.

Bell states refer to the following 4 states:

$$\begin{aligned}
\Phi^+ &= \frac{1}{\sqrt{2}}\left(|\,0\rangle^A|\,0\rangle^B + |\,1\rangle^A|\,1\rangle^B\right)\\
\Phi^- &= \frac{1}{\sqrt{2}}\left(|\,0\rangle^A|\,0\rangle^B - |\,1\rangle^A|\,1\rangle^B\right)\\
\Psi^+ &= \frac{1}{\sqrt{2}}\left(|\,0\rangle^A|\,1\rangle^B + |\,1\rangle^A|\,0\rangle^B\right)\\
\Psi^- &= \frac{1}{\sqrt{2}}\left(|\,0\rangle^A|\,1\rangle^B - |\,1\rangle^A|\,1\rangle^B\right)
\end{aligned}$$

These 4 states form a basis of the 2-qubit systems, and all these 4 states are maximally entangled.

A quantum state is *disentangled* if it is of the form $|\,\psi\rangle^A \otimes |\,\psi'\rangle^B$. Any other pure state in $\mathcal{H}^A \otimes \mathcal{H}^B$ is *entangled*. For a pure state $|\,\varphi\rangle$ in a bipartite system, we define its *entanglement* to be the von Neumann entropy of the reduced sub-system of Bob when we trace out Alice:

$$E(|\,\varphi\rangle) = S(\mathrm{Tr}_A(|\varphi\rangle\langle\varphi|)) \tag{1}$$

where $S(\rho) = -\mathrm{Tr}(\rho \log \rho)$ is the von Neumann entropy. We have $S = 0$ if and only the state is disentangled. A mixed state ρ is disentangled if and only if it is equivalent to a state that is a mixture of pure states $|\,\varphi_i\rangle$ with probabilities p_i. Any other mixed state is entangled. However, there is no universally agreed upon definition for the amount of entanglement in a mixed state.

If we denote the dimension of $\mathcal{H}^A$ by N, then the maximum amount of entanglement in this system is $\log N$. We define the state Ψ_N to be

$$\Psi_N = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |\,i\rangle^A |\,i\rangle^B \tag{2}$$

It is a maximally entangled state in $\mathcal{H}^A \otimes \mathcal{H}^B$. Notice it is a state in a space of dimension N^2. In particular, if N is a power of 2: $N = 2^n$, then the state Ψ_N is the state of n EPR pairs. We call this special kind of states *EPR states*.

2.2 Fidelity

For two (mixed) states ρ and σ in the same quantum system, their *fidelity* is defined as[3]

$$F(\rho, \sigma) = \mathrm{Tr}(\rho^{1/2}\sigma\rho^{1/2}). \tag{3}$$

If $\sigma = |\varphi\rangle\langle\varphi|$ is a pure state, the definition simplifies to

$$F(\rho, |\varphi\rangle\langle\varphi|) = \langle\varphi\,|\rho|\,\varphi\rangle \tag{4}$$

In the special case that $|\,\varphi\rangle = \Psi_N$ is the maximally entangled state, we call the fidelity of ρ and $|\,\varphi\rangle$ the *fidelity of state* ρ, and the definition simplifies to:

$$F(\rho) = \langle\Psi_N\,|\rho|\,\Psi_N\rangle \tag{5}$$

The fidelity is linear with respect to ensembles.

Claim 1 *Let ρ be the density matrix for a mixed state that is an ensemble $\{p_i,\ |\,\phi_i\rangle\}$. The fidelity of ρ is the weighted averages of the qualities of the pure states:*

$$F(\rho) = \sum_i p_i \cdot F(|\phi_i\rangle\langle\phi_i|)$$

This linearity is used in several proofs in this paper.

[3] Notice that we are using a different definition from some other literature, including [18].

3 General Entanglement Purification Protocols

3.1 The General Setting

Alice and Bob are given a state in $\mathcal{H}_N^A \otimes \mathcal{H}_N^B$. They are also given an auxiliary input $\Psi_K \in \mathcal{H}_K^A \otimes \mathcal{H}_K^B$. Alice can perform unitary transformations on her part of the state ($\mathcal{H}_N^A \otimes \mathcal{H}_K^A$) and Bob can perform unitary transformations on his part ($\mathcal{H}_N^B \otimes \mathcal{H}_K^B$). Since those transformations only affect one part of the state, they are called *local operations*. Alice and Bob are also allowed to communicate classical bits but not quantum bits. This model is called *LOCC (local operations and classical communication)* [4, 17].

If the starting state is disentangled, applying LOCC operations keeps the state disentangled [4]. Thus, LOCC operations cannot create entanglement but they can be used to extract the entanglement that already exists in the state.

We use the letter $\mathcal{P}$ to denote protocols for extracting entanglement by LOCC operations. At the end of a protocol $\mathcal{P}$, Alice and Bob have two options:

1. They can abort and claim failure by outputting a special symbol FAIL. We denote this by $\mathcal{P}(\rho) = \mathsf{FAIL}$.

2. They can output a (possibly mixed) state σ in $\mathcal{H}_M^A \otimes \mathcal{H}_M^B$. We denote this by $\mathcal{P}(\rho) = \sigma$.

We now define the error model. We first give an unsuccessful definition to illuminate some of difficulties that we face and to explain the reasons behind our final definition.

3.2 Extracting Entanglement From an Arbitrary State

Ideally, we would like to have a protocol that takes any entangled state in $\mathcal{H}_N^A \otimes \mathcal{H}_N^B$ with at least a certain amount of entanglement and extracts a state close to Ψ_M for some $M < N$. This would correspond the definition of extractors transforming any probability distribution with min-entropy at least m into a probability distribution close to uniform.

Unfortunately, this is not possible, even if we restrict ourselves to starting states with the maximum possible entanglement. Unlike in the classical world where there is just one probability distribution over N elements with entropy $\log N$ (the uniform distribution), there are infinitely many quantum states with entanglement $\log N$. Namely, any quantum state of the form

$$|\,\phi\rangle = \sum_{i=0}^{N-1} \alpha_i |\,i\rangle |\,i\rangle \qquad (6)$$

with $|\alpha_i|^2 = 1/N$ for all $i \in \{0, \dots, N-1\}$ has entanglement $\log N$. In particular, this includes

$$|\,\phi_a\rangle = \sum_{b=0}^{N-1} \frac{1}{\sqrt{N}} e^{2i\,ab\pi/N} |\,b\rangle |\,b\rangle$$

for $a \in \{1, \dots, N\}$. Assume that we have a protocol that extracts Ψ_M from any $|\,\phi_a\rangle$. This means that, given $|\,\phi_a\rangle$, the protocol ends with the final state of the form $\Psi_M \otimes |\,\phi'_a\rangle$. We consider running this protocol on the mixed state ρ that is $|\,\phi_0\rangle$ with probability $1/N$, $|\,\phi_1\rangle$ with probability $1/N$, ..., $|\,\phi_{N-1}\rangle$ with probability $1/N$. Then, the final state is of the form $\Psi_M \otimes \rho'$ where ρ' is some mixed state.

The problem is that ρ is equivalent to the mixed state that is $|\,0\rangle|\,0\rangle$ with probability $1/N$, $|\,1\rangle|\,1\rangle$ with probability $1/N$, ..., $|\,N-1\rangle|\,N-1\rangle$ with probability $1/N$. (This equivalence can be verified by writing out the density matrices of both states.) None of the states $|\,i\rangle|\,i\rangle$ is entangled, so the mixed state obtained by combining them is also not entangled. Yet, since this mixed state is equivalent to ρ, it gets transformed into $\Psi_M \otimes \rho'$, which is entangled.

We have constructed a protocol that transforms a disentangled starting state into entangled end state without quantum communication. Since this is impossible [4], our assumption is wrong and there is no protocol that extracts any Ψ_M from an arbitrary $|\,\phi_a\rangle$.

The argument described above is still valid if we relax the requirement to extracting a state close to Ψ_M and if we allow to use a perfect auxiliary state Ψ_K. In the second case, we can get the perfect Ψ_K back but cannot get an entangled state of higher dimension.

This is a clear distinction between the situation of classical randomness extraction and quantum entanglement extraction. In the classical case, all the probabilities are nonnegative real numbers, and the min entropy of a random distribution already characterizes the distribution well. In the quantum case, the magnitudes are complex numbers, and the entanglement alone isn't good enough to describe the state. Even more interestingly, since one has the freedom to switch bases in quantum, one can build a mixed state which is a mixture of maximally entangled states, yet the mixed state itself is completely disentangled. This phenomenon doesn't seem to have a counterpart in classical probability.

3.3 Extracting From a State Close to Ψ_N

The reason for the problem in the previous section is that there are multiple maximally entangled states and combining them into a mixed state can cancel the entanglement and create a state with no entanglement. To be able to extract entanglement, we have to restrict ourselves to states that are close to a fixed highly entangled state (rather than some highly entangled state). Therefore, we assume that the starting state is close to Ψ_N[4].

A common way to measure the closeness to Ψ_N is the fidelity (section 2.2). This gives the following definitions.

Definition 1 (Absolutely Successful GEPP) *A General Entanglement Purification Protocol $\mathcal{P}$ is* absolutely successful *(AS) with parameters $\langle N, K, M, \epsilon, \delta\rangle$, if for all states ρ such that $F(\rho) \geq 1-\epsilon$,*

$$\textsf{Prob}\,[\mathcal{P}(\rho) = \textsf{FAIL}] = 0$$

and

$$\textsf{Prob}\,[F(\mathcal{P}(\rho)) \geq 1-\delta] = 1$$

Definition 2 (CS Protocol) *A General Entanglement Purification Protocol $\mathcal{P}$ is* conditionally successful *(CS) with parameters $\langle N, K, M, \epsilon, \delta, p\rangle$ if for all input states ρ such that $F(\rho) = 1-\epsilon$, we have* $\textsf{Prob}\,[\mathcal{P}(\rho) = \textsf{FAIL}] \leq p$ *and*

$$\mathbb{E}_{\mathcal{P}}\left[F(\mathcal{P}(\rho)) \mid \mathcal{P}(\rho) \neq \textsf{FAIL}\right] \geq 1-\delta,$$

where $\mathbb{E}_{\mathcal{P}}$ denotes the expectation taken over the classical communication in the protocol $\mathcal{P}$.

Note that in the previous definition, we only require that the *average* fidelity be high when the protocol succeeds. Although good enough in many cases, there are situations where a stronger condition is desired. Consider the following adversarial setting: Alice and Bob try to extract EPR pairs through LOCC, and Eve can see the classical communication between Alice and Bob. The previous definition doesn't rule out the possibility that when Alice and Bob don't obtain EPR pairs of high fidelity (which can happen with small probability), Eve knows about it and can attack Alice and Bob. This situation is undesirable since Eve has the knowledge about the fidelity. A stronger, more desirable definition would imply that the classical communication Eve sees is *oblivious* to the fidelity of the EPR pairs. In other words, Alice and Bob should always output EPR pairs with high fidelity, regardless the classical messages they send. We call protocols satisfying this stronger definition *deterministically*(conditionally) successful:

Definition 3 (DCS Protocol) *A General Entanglement Purification Protocol $\mathcal{P}$ is* deterministically conditionally successful *with parameters $\langle N, K, M, \epsilon, \delta, p\rangle$, if for all input states ρ such that $F(\rho) = 1-\epsilon$,* $\textsf{Prob}\,[\mathcal{P}(\rho) = \textsf{FAIL}] \leq p$ *and*

$$\textsf{Prob}\,[F(\mathcal{P}(\rho)) \geq 1-\delta \mid \mathcal{P}(\rho) \neq \textsf{FAIL}] = 1$$

Clearly, any DCS protocol is also CS. In the converse direction, there is a very simple way to convert any CS protocol to a DCS one by encrypting the classical communication with a one-time pad. In this way, the communication between Alice and Bob will be totally oblivious to Eve. Alice and Bob can then erase all their private (classical) memories, and then the protocol becomes DCS. To set the one-time pad, Alice and Bob can start by sharing c perfect EPR pairs, if the classical communication complexity is c.

Proposition 1 *A CS protocol with parameters $\langle N, M, K, \epsilon, \delta, p\rangle$ which uses c bits of communication can be converted to a DCS protocol with parameters $\langle N, M, 2^cK, \epsilon, \delta, p\rangle$.*

Finally, we say a GEPP is *efficient* if it can be implemented by quantum circuits of size $O(poly(\log N + K))$.

4 Summary of Results

4.1 Impossibility Result for Absolutely Successful Protocols

We first prove that there don't exist absolutely successful protocols with "interesting" parameters.

Theorem 1

(a) *For all absolutely successful GEPPs with parameters $\langle N, K, M, \epsilon, \delta\rangle$, we have the following inequality:*

$$\delta \geq \frac{M-K}{M}\frac{N}{N-1}\epsilon.$$

(b) *The bound in (a) is tight. For any integers N, M, K such that NK/M and M/K are both integers, there exists an absolutely successful GEPP with parameters $\langle N, K, M, \epsilon, \frac{M-K}{M}\frac{N}{N-1}\epsilon\rangle$.*

This shows that absolutely successful protocols are quite weak. If we just want to extract the auxiliary state Ψ_K and t more EPR pairs, then $M = 2^tK$ and we can achieve fidelity at most $1-\frac{2^t-1}{2^t}\frac{N}{N-1}\epsilon < 1-(1-\frac{1}{2^t})\epsilon$ which is hardly better than $1-\epsilon$ that we had at the beginning. If we want to get Ψ_K plus a linear number of EPR pairs, the improvement in fidelity is an exponentially small fraction of ϵ.

4.2 Constructions of CS Protocols

We show that conditionally successful GEPPs can be constructed.

Theorem 2 *For any integer n and any $s \in \{1, 2, ..., n\}$:*

[4] The protocols can be modified to use any other fixed state of the form (6) instead of Ψ_N.

(a) *there exist efficient CS protocols with parameters*

$$\langle 2^n, 1, 2^{n-s}, \epsilon, \frac{2^{-s}}{(1-\epsilon)}, \epsilon \rangle$$

(b) *there exist efficient DCS protocols with parameters*

$$\langle 2^n, 2^{2s+1}, 2^{n-s}, \epsilon, \frac{2^{-s+\log n}}{(1-\epsilon)}, \epsilon \rangle.$$

Notice that our results are near-optimal in terms of fidelity: by Theorem 1 , if the input state has fidelity $1-\epsilon$, then the *overall* fidelity of the output of a protocol cannot be significantly higher than $1-\epsilon$. Therefore, if a CS protocol has a very high fidelity in its output when it doesn't fail, then this protocol must fail with probability at least about ϵ. Both protocols in Theorem 2 fail with probability ϵ, and when they don't fail, the fidelity their output can be made arbitrarily close to 1. Also, the CS protocol in part **(a)** of Theorem 2 is optimal in the usage of additional EPR pairs: it doesn't use any at all. This is interesting since in classical randomness extraction, one has to invest logarithmically number of perfect random bits in order to extract high-quality random bits. However, in the case of quantum entanglement extraction, no perfect EPR pairs are needed.

We will discuss the construction of the 2 protocols in the next section.

5 Impossibility for Absolutely Successful GEPPs

We prove Theorem 1 in this section.

We first study a simpler problem. Suppose Alice and Bob share a maximally entangled state Ψ_K and some private ancillary bits, initialized to $|\,0\rangle$. We describe this shared state by

$$|\,\phi\rangle = (|\,Z_N\rangle^A \otimes |\,Z_N\rangle^B) \otimes \Psi_K$$

The fidelity of this state is K/M by a simple computation.

Alice and Bob try to convert state $|\,\phi\rangle$ as close to Ψ_M as possible by LOCC. How close can they get? If M/K is an integer, Alice and Bob just trace out a subsystem of their ancillary bits to bring the dimension of each their subsystem to M, then they obtain a state

$$|\,\psi_0\rangle = (|\,Z_{M/K}\rangle^A \otimes |\,Z_{M/K}\rangle^B) \otimes \Psi_K$$

which has fidelity K/M by a straightforward computation. In fact, this is the best Alice and Bob can do:

Lemma 1 *Let $|\,\phi\rangle = (|\,Z_N\rangle^A \otimes |\,Z_N\rangle^B) \otimes \Psi_K$ be a state in a bipartite system $\mathcal{H}^A_{NK} \otimes \mathcal{H}^B_{NK}$ shared between Alice and Bob. Let σ be the state Alice and Bob output after performing LOCC operations. Suppose that σ is in the subspace $\mathcal{H}^A_M \otimes \mathcal{H}^B_M$. We have $F(\sigma) \leq \frac{K}{M}$.* ■

This lemma is a direct corollary of a result by Vidal, Jonathan, and Nielsen [25].

Proof: [Proof of Theorem 1, part (a)]

We prove part (a) of the theorem by demonstrating a particular mixed state ρ such that ρ has a fidelity $1-\epsilon$, and no LOCC can increase its fidelity to more than $1-\frac{M-K}{M}\frac{N}{N-1}\epsilon$.

Let $\epsilon' = \frac{N}{N-1}\epsilon$. We define the state ρ to be

$$\rho = (1-\epsilon')\cdot|\Psi_N\rangle\langle\Psi_N| + \epsilon'\cdot|Z^A_N \otimes Z^B_N\rangle\langle Z^A_N \otimes Z^B_N|$$

In fact, ρ is the maximally entangled state Ψ_M with probability $(1-\epsilon')$ and the totally disentangled state $Z^A_N \otimes Z^B_N$ with probability ϵ'.

It is easy to verify that $F(\rho) = 1-\epsilon$, since $\langle\Psi_N \mid Z^A_N \otimes Z^B_N\rangle = 1/\sqrt{N}$ and, therefore,

$$\begin{aligned} F(\rho) &= (1-\epsilon')F(|\Psi_N\rangle\langle\Psi_N|) \\ &\quad +\epsilon' F(|Z^A_N \otimes Z^B_N\rangle\langle Z^A_N \otimes Z^B_N|) \\ &= (1-\epsilon') + \frac{1}{N}\epsilon' = 1-(1-\frac{1}{N})\epsilon' = 1-\epsilon. \end{aligned}$$

For an arbitrary GEPP $\mathcal{P}$ that never fails, we define

$$f_1 = F(\mathcal{P}(|\Psi_N\rangle\langle\Psi_N|)),$$

$$f_2 = F(\mathcal{P}(|Z^A_N \otimes Z^B_N\rangle\langle Z^A_N \otimes Z^B_N|))$$

Then we have $f_1 \leq 1$ and by Lemma 1, $f_2 \leq K/M$.

By the linearity of fidelity of quantum operations,

$$F(\mathcal{P}(\rho)) = (1-\epsilon')f_1 + \epsilon' f_2 \leq 1 - \frac{M-K}{M}\epsilon'$$

$$= 1 - \frac{M-K}{M}\frac{N}{N-1}\epsilon.$$

■

The part (b) of Theorem 1 is proven by the following protocol.

The input to the protocol is a state ρ in $\mathcal{H}^A_N \otimes \mathcal{H}^B_N$. The output is a state in $\mathcal{H}^A_M \otimes \mathcal{H}^B_M$ where $M < N$ and M divides N. There is no auxiliary state used, i.e., $K = 1$.

Construction 1 (Random Permutation Protocol)

1. *Alice generates a uniformly random permutation π on N elements using classical randomness and transmits the permutation to Bob.*

2. *Alice applies permutation π on $\mathcal{H}^A_N$, mapping $|\,i\rangle$ to $|\,\pi(i)\rangle$, Bob does the same on $\mathcal{H}^B_N$.*

3. *Alice and Bob decompose $\mathcal{H}_N$ as $\mathcal{H}_M \otimes \mathcal{H}_L$, $L = N/M$ and measure the $\mathcal{H}_L$ part.*

4. *Alice sends the result of her measurement to Bob, Bob sends his result to Alice.*

5. *They compare the results. If the results are the same, they output the state that they have in $\mathcal{H}_M^A \otimes \mathcal{H}_M^B$. If the results are different, they output $|\, Z_M\rangle \otimes |\, Z_M\rangle$.*

If the input state ρ has fidelity $1-\epsilon$, then the output state has fidelity at least $1 - \frac{M-K}{M}\frac{N}{N-1}\epsilon$, matching the lower bound of part (a). The proof is straightforward, and is omitted due to space constraints.

6 Constructing Conditionally Successful GEPPs

Purity-testing Protocols We construct conditionally successful GEPP's based on the "purity testing" protocols of [2]. A purity testing protocol is an LOCC protocol where the input is joint state shared by Alice and Bob which they think might be the EPR state $|\, \Phi^{(n)}\rangle = |\, \Phi^{+}\rangle^{\otimes n}$. Alice and Bob want to test if their shared state is indeed $|\, \Phi^{(n)}\rangle$, while sacrificing the least number of EPR pairs.

Definition 4 ([2]) *A purity testing protocol with error α is a LOCC super-operator $\mathcal{T}$ which maps $2n$ qubits (half held by Alice and half held by Bob) to $2m+1$ qubits (m of which are held by Bob) and satisfies the following two conditions:*

- Completeness: $\mathcal{T}(|\, \Phi^{(n)}\rangle) = |\, \Phi^{(m)}\rangle \otimes |\, \text{ACC}\rangle$
- Soundness: *Let P be the projection on the subspace spanned by $|\, \Phi^{(m)}\rangle \otimes |\, \text{ACC}\rangle$ and $|\, \psi\rangle \otimes |\, \text{REJ}\rangle$ for all $|\, \psi\rangle$. Then $\mathcal{T}$ is* sound *if for all ρ,*
$$\text{Tr}\,(P\mathcal{T}(\rho)) \geq 1-\alpha.$$

It's convenient to think of purity testing as approximating the measurement given by the projector onto $|\, \Phi^{(m)}\rangle$ and its orthogonal complement.

A particularly simple purity-testing protocol consists of picking a random stabilizer code of dimension 2^{n-s}, having Alice and Bob both measure the syndrome of the code, and then extracting the encoded state if both measurement results are the same.

Lemma 2 (Random hashing) *There exist purity testing protocols such that $m = n-s$, $\alpha \leq 2^{-s}$ and which use $ns+s+1$ bits of (classical) communication.*

This lemma actually follows from the observation that the set of *all* stabilizer codes [9] of dimension 2^{n-s} is a purity-testing code family with error $\alpha \leq 2^{s}$ in the sense of [2]. However, we give a direct proof with an explicit protocol description in Section 6.2 below.

Barnum et al. provide a construction which achieves better communication complexity at the cost of increasing the error by a factor of (roughly) $n = \log N$.

Fact 1 ([2]) *There exist purity testing protocols such that $m = n-s$, $\alpha \leq 2^{-s+\log n}$ and which use $2s+1$ bits of (classical) communication.*

6.1 GEPP's from Purity-Testing Protocols

Suppose we now use a purity testing protocol as a GEPP. That is, we set $N = 2^n$, $M = 2^m$ and $K = 1$, and just run the purity-testing protocol, outputting FAIL when the purity testing rejects the input. Suppose at the end of the protocol we trace out everything except the $2m$ output qubits and the qubit indicating accept/reject. Consider the three projectors:

$$\begin{aligned} P_1 &= |\, \Phi^{(m)}\rangle\langle\Phi^{(m)}\,| \otimes |\, \text{ACC}\rangle\langle\text{ACC}\,| \\ P_2 &= (I_{\mathcal{M}} - |\, \Phi^{(m)}\rangle\langle\Phi^{(m)}\,|) \otimes |\, \text{ACC}\rangle\langle\text{ACC}\,| \\ P_3 &= I_{\mathcal{M}} \otimes |\, \text{REJ}\rangle\langle\text{REJ}\,| \end{aligned}$$

And define $\gamma_i = \text{Tr}[P_i\rho']$ where ρ' is the final state.

If the input to the system had fidelity $1-\epsilon$, then the completeness of the purity-testing protocol implies that the fidelity of the output to $|\, \Phi^{(m)}\rangle|\, \text{ACC}\rangle$ must be $1-\epsilon$, and so $\gamma_1 \geq 1-\epsilon$. If the purity-testing protocol has soundness error α, then the soundness condition implies $\gamma_2 \leq \alpha$.

Now the output fidelity conditioned on acceptance is

$$\frac{\gamma_1}{\gamma_1+\gamma_2} = 1 - \frac{\gamma_2}{\gamma_1+\gamma_2} \geq 1 - \frac{\alpha}{1-\epsilon+\alpha}$$

Choosing the soundness error to be small enough yields very good average fidelity of the output. The only problem here is that conditioned on the communication in the channel (which the adversary will presumably be able to see), the fidelity may be quite low. Thus, we do not obtain a DCS protocol, but a CS protocol with parameters:

$$\left\langle 2^n, 1, 2^{n-s}, \epsilon, \frac{\alpha}{1-\epsilon}, \epsilon \right\rangle$$

Applying the reasoning above to Lemma 2 and Fact 1 completes the proof of Theorem 2.

6.2 A Simple Random Hashing Protocol

Without loss of generality, we describe the protocol in terms of purifying the state $|\, \Psi^{-}\rangle$[5]. We describe a protocol with $m = n-1$ and error $\alpha = \frac{1}{2}$. Repeating the protocol s times yields $m = n-s$ and $\alpha = 2^{-s}$.

Construction 2 (Simple Random Hashing Protocol)

1. *Alice picks $2n$ random bits $x_1, ..., x_n, z_1, ..., z_n$ such that not all the bits are 0.*

2. *Alice will measure the operator given by $X^{x_1}Z^{z_1} \otimes \cdots \otimes X^{x_n}Z^{z_n}$. To do this Alice:*

   (a) *Considers only qubits where $(x_i, z_i) \neq (0,0)$. Say there remain ℓ qubits.*

[5] For example, Bob can perform a "phase-shift" (Z) followed by a "bit-flip" (X) to every qubit he possesses. This will transform $|\, \Phi^{+}\rangle$ to $|\, \Psi^{-}\rangle$.

(b) *On qubit j, applies either*
$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ *if* $(x_j, z_j) = (0, 1)$,
$B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ *if* $(x_j, z_j) = (1, 1)$,
the identity if $(x_j, z_j) = (1, 0)$.

(c) *Applies C-NOT from each of the first $\ell - 1$ qubits onto the last.*

(d) *Measures the last in the computational basis.*

(e) *Applies the inverse transformation to the remaining qubits.*

3. *Alice sends $x_1, ..., x_n, z_1, ..., z_n$ and her measurement result to Bob.*

4. *Bob performs the same measurement and sends back the result.*

5. *Alice and Bob accept if the two results are different and reject otherwise.*

Proof: (Sketch of proof of Lemma 2)

As argued in [2], for protocols of this form it is sufficient to consider the performance of the protocol on states of the form $X^{\vec{a}}Z^{\vec{b}}|\Psi^-\rangle^{\otimes n}$, where $X^{\vec{a}}$ denotes $X^{a_1} \otimes \cdots \otimes X^{a_n}$ when $\vec{a} = (a_1, ..., a_n) \in \{0, 1\}^n$. WLOG we assume all the error operators are applied to Alice's share of the EPR pairs.

The reduction to these Bell states is via a "quantum-to-classical reduction", as used in [14] for key distribution. The reduction works because ultimately, the accept/reject decision is diagonal in the Bell basis, and moreover if the input to the protocol can be described as $X^{\vec{a}}Z^{\vec{b}}|\Psi^-\rangle^{\otimes n}$, the the output can be written $X^{\vec{a}'}Z^{\vec{b}'}|\Psi^-\rangle^{\otimes m}$.

The idea is that measuring the operator $X^{x_1}Z^{z_1} \otimes \cdots \otimes X^{x_n}Z^{z_n}$ on both Alice and Bob's shares and comparing the results is equivalent to measuring the bit $\vec{a} \odot \vec{x} + \vec{b} \odot \vec{z}$, i.e. a random linear function of the vector $(\vec{x}, \vec{z})$. To see this, first observe that $HX^aZ^b = (-1)^{ab}X^bZ^bH$ and $BX^aZ^b = i^bX^{a+b}Z^bB$. Moreover, both $B \otimes B$ and $H \otimes H$ have $|\Psi^-\rangle$ as an eigenvector. Thus, in each position we will end up with a state proportional to $X^{x_ja_j+z_jb_j}Z^c|\Psi^-\rangle$ after Alice and Bob have applied their transformations and before they measure, where c is a bit. Measuring both halves in the computational basis and comparing results allows one to compute $x_ja_j + z_jb_j$. Similarly, the protocol computes $\vec{x} \odot \vec{a} + \vec{b} \odot \vec{z}$.

A random linear function will detect a non-zero vector with probability $\frac{1}{2}$. Thus, the overall error probability of the one-step protocol is bounded by $\frac{1}{2}$. Repeating the protocol s times lowers this error to 2^{-s}. ■

7 Conclusions and Open Problems

We investigated the problem of quantum entanglement extraction by Alice and Bob via LOCC. We used a general model of the imperfect EPR pairs: the only information Alice and Bob have is a lower bound on the fidelity of the input state. We also argued that an apparently more "general" model, where Alice and Bob only know the *entanglement* of the shared state, doesn't work: there don't exist LOCC protocols that output near-perfect EPR pairs on any input of a certain entanglement.

We defined 3 types of General Entanglement Purification Protocols. Absolutely successful (AS) GEPPs never fail, and they always output states of high fidelity. Conditionally successful (CS) protocols are allowed to fail with a small probability, but conditioned on that they don't fail, the *expected* fidelity of their output is high. We can strengthen a CS protocol to a Deterministic (conditionally) successful (DCS) protocols, which are *guaranteed* to output a state of high fidelity with probability 1 when they don't fail.

We proved a negative result that there don't exist AS protocols of interesting parameters. Therefore, on average, the ability of Alice and Bob to enhance the fidelity is very limited. However, Alice and Bob can "concentrate" the fidelity into some cases, while fail in others. We used "purity-testing protocols" [2] to obtain efficient CS and DCS protocols that are nearly optimal in terms of fidelity: they can achieve arbitrarily high fidelity when they succeed, and their failure probability is about the minimal possible. In addition, the CS protocol doesn't use any perfect EPR pairs as auxiliary input. This shows a stark contrast to the case of classical randomness extraction, where additional perfect EPR pairs are necessary.

Our study shows some interesting facts about the comparison of classical randomness extraction and quantum entanglement extraction. There have been a lot of apparent similarities between the two, the most obvious one being the protocol to extract perfect EPR pairs from identical copies of imperfect EPR pairs [4] and the algorithm for extracting perfect random bits from a biased coin [16]. However, there are some clear distinctions when we move to more general paradigms. A classical extractor works with any random source of enough min-entropy, regardless its distribution. There are no such counterparts in the quantum world: no LOCC protocol can work with any bipartite state of certain entanglement and produce EPR pairs. An extractor needs at least logarithmic number of perfect random bits as its auxiliary input, while we showed that there exists a CS protocol that doesn't consume any additional EPR pairs at all.

There are still many open problems remaining.

1. **Complete Characterization of "Extractable Entanglement"**

In classical randomness extraction, the min entropy completely characterizes the amount the "extractable randomness" of the source: there is an upper bound in terms of min entropy on how many high-quality random bits one can extract from this source, and there are constructions that almost matches the bound. However, such a complete characterization for extractable quantum entanglement still evades us. Different models of imperfect EPR pairs have been proposed, each with its own particular characterization, which are not comparable to each other. We have argued in our paper that the entanglement of a bipartite state by itself isn't a good characterization, and some "closeness" condition to a pre-defined maximally entangled state seems necessary. We feel that finding such a complete characterization will help understanding entanglement extraction greatly.

2. **Optimality**

   In the classical case of randomness extraction, where have a very good understanding on the optimal efficiency of an extractor: how many bits it can output, and how many perfect random bits it has to invest. In the quantum case, the question of optimality is much less clearly understood. The fact that we are studying a protocol (which involves communication) rather than an algorithm further complicates the situation. There are many different features and resource usages one can optimize:

   (a) **Fidelity and Failure Probability.** We want to maximize the fidelity of the output state of a protocol and minimized the probability a protocol fails. Our CS and DCS protocols constructed in this paper are nearly optimal under the "general error" model. What about other models of imperfect EPR pairs?

   (b) **Yield.** We want to maximize the number of high-quality EPR pairs a protocol outputs. It is not clear what the optimal value is, and we don't know if our protocol is optimal in this sense.

   (c) **Perfect EPR Pairs Invested.** We want to minimize the number of perfect EPR pairs a protocol invests as an auxiliary input. The CS protocol in our paper doesn't use any perfect EPR pairs and is thus optimal. The DCS protocol, however, uses $(2s+1)$ perfect EPR pairs. Is that optimal? Can we eliminate the needs for additional EPR pairs totally?

   (d) **Communication Complexity.** We want to minimize the communication complexity between Alice and Bob. How many (classical) bits does Alice need to send to Bob in order to extract entanglement? The CS and DCS protocols in this paper only involve one-way communication, save a single bit. Does two-way communication improve the communication complexity?

Acknowledgment

The authors would like to thank Bob Griffith and Steven Rudich for enlightening discussions. The authors thank John Langford for proof-reading the earlier version of this paper.

References

[1] C. H. Bennett and G. Brassard, *Quantum Cryptography: Public-key Distribution and Coin Tossing*, in Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, 1984 (IEEE Press, 1984), pp.175 - 179. See also C. H. Bennett and G. Brassard, *Quantum Public Key Distribution*, IBM Technical Disclosure Bulletin **28**, 3153-3163 (1985).

[2] H. Barnum, C. Crépeau, D. Gottesman, A. Smith and A. Tapp, *Authentication of Quantum Messages*, unpublished manuscript, 2001. Available from the Los Alamos e-print archive, or the authors.

[3] C. H. Bennett, G. Brassard, C. Crépeau, R. Josza, A. Peres, and W. K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett. **70**, 1895 (1993).

[4] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, *Concentrating partial entanglement by local operations*, In *Physical Review A*, vol. 53, No. 4, April 1996.

[5] C. H. Bennett, H. J. Bernstein, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, *Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels*, In *Physics Review Letters*, vol. 76, page 722, 1996.

[6] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed-state entanglement and quantum error correction*. In *Physical Review A*, vol. 54, No. 5, November 1996.

[7] C. H. Bennett and S. J. Wiesner, *Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states*, Phys. Rev. Lett. **69**, 2881 (1992).

[8] A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* Phys. Rev. **47**, 777 (1935) [reprinted in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. Z. Zurek, Princeton University Press, 1983].

[9] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*, Ph.D. thesis, California Institute of Technology, 1997.

[10] L. Hardy, *Method of areas for manipulating the entanglement properties of one copy of a two-particle pure entangled state*, Phys. Rev. A, **60**, 1912 (1999). also available at `quant-ph/9903001`.

[11] M. Horodecki, P. Horodecki, and R. Horodecki, *Distillability of Inseparable Quantum Systems*. In `quant-ph/9607009`.

[12] D. Jonathan and M. Plenio, *Minimal conditions for local pure-state entanglement manipulation*, Phys. Rev. Lett. **83**, 1455 (1999), also available at `quant-ph/9903054`.

[13] M. Luby, *Pseudorandomness and Cryptographic Applications*, Princeton University Press, 1996.

[14] Hoi-Kwong Lo and H.F. Chau, *Unconditional Security of Quantum Key Distribution Over Arbitrary Long Distances*, Science **283**, 2050-2056 (1999), also available at `quant-ph/9803006`.

[15] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, 1995.

[16] J. von Neumann, Various techniques used in connection with random digits. *Notes by G. E. Forsythe, National Bureau of Standards*, 1952, vol. 12, pages 36-38.

[17] M. Nielsen, *Conditions for a class of entanglement transformations*, Phys. Rev. Lett, 83(2):436:439, 1999.

[18] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[19] N. Nisan and A. Ta-Shma, *Extracting Randomness: A Survey and New Constructions*. JCSS 58(1): 148-173 (1999).

[20] Noam Nisan and David Zuckerman, *Randomness is Linear in Space*. JCSS 52(1): 43-52 (1996)

[21] Peter W. Shor and John Preskill *Simple Proof of Security of the BB84 Quantum key Distribution Protocol*, Phys.Rev.Lett. 85 (2000) 441-444, also available at `quant-ph/0003004`.

[22] A. Ta-Shma, C. Umans, D. Zuckerman. Loss-less condensers, unbalanced expanders and extractors. *Proceedings of STOC'01*, pp. 143-152.

[23] L. Trevisan. Construction of extractors using pseudorandom generators. *Proceedings of STOC'99*, pp. 141-148.

[24] G. Vidal, *Entanglement of pure states for a single copy*, Phys. Rev. Lett. 83 (1999) 1046-1049, quant-ph/9902033

[25] G. Vidal, D. Jonathan, and M. Nielsen *Approximation Transformations and Robust Manipulation of Bipartite Pure State Entanglement* Phys. Rev. A **62**, 012304 (2000) Also available at `quant-ph/9910099`.

[26] M. Wegman and J. Carter. *New Hash Functions and Their Use in Authentication and Set Equality*, In *Journal of Computer and System Sciences*, vol. 22, pp 265-279, 1981.