Concentrating Partial Entanglement by Local Operations

Reid Hayes 20457389
for PHYS 467

March 26, 2017

1   Abstract

Two observers sharing n partially entangled pairs of qubits, with identical entanglement entropy, may through local operations only, generate a smaller number of maximally entangled qubits. The method for doing so by Bennett et al. [3] approaches perfect yield in the limit of large n.

2   Introduction

Quantum teleportation and superdense coding both presuppose that Alice and Bob share a maximally entangled qubit [1, 2]. This requirement was an an initial blow to hopes of using these protocols for communication. Bennett admits in the 1992 paper that introduced superdense coding that

> The communication of two bits via two particles, one of which remains fixed while the other makes a round trip, is no more efficient in number of particles or number of transmissions than the obvious scheme of directly encoding each bit in one transmitted particle.   [2]

Bennett argues that there is still one advantage of superdense coding: the Bell states may be sent during 'off-peak' hours, before the message has been prepared [2].

There is hope of doing even better. Although entanglement is a precious resource for anyone hoping to build a quantum computer or channel, nature has an over-abundance of it. Any two systems that share a causal history, and are capable of interacting, almost certainly share some small degree of entanglement. This is what gives rise to the high dimensionality that makes quantum systems notoriously difficult to simulate and what inspired pioneers like Richard Feynman to first propose quantum computing [4]. This natural entanglement could in principle be harnessed for quantum communication.

Any progress towards creating maximally entangled qubits from less idealized entangled states would be a step in this direction. This would also open up the possibility of sending Bell states over lower fidelity channels.

In this report, we review the progress made by Bennett et al. to solve this problem. They detail a procedure for converting a collection of partially entangled pairs into fewer maximally entangled qubits [3]. This procedure uses only local operations [3]. Therefore the two systems may remain spacelike separated for the duration of the procedure. Furthermore, in the limit of large n this procedure conserves expected entanglement entropy [3]. This means that no entanglement goes to waste. Furthermore, this conservation elevates entanglement entropy — not shared Bell states — as the fundamental resource for quantum communication; at least within the limitations of this method.

Bennett et al.'s method is not the final word in extracting useful entanglement from our messy real world. It has shortcomings, which mean that, at least for now, cheap entanglement is still illusive. Bennett et al. assume that

1. the initial state, from which we concentrate and distill Bell states, possesses only pairwise entanglement. That is, that there is no tripartite or higher order entanglement.

2. all pairs are equally entangled. That is, the entanglement entropy of every entangled pair is the same. For two state systems, this is equivalent to every pair having the same Schmidt coefficients.

3. there is detailed knowledge of the initial state, which must be used to decide which measurements and unitary operations should be applied throughout the procedure. That is, that the initial state is pure, using the Bayesian definition of probability.

For certain physical systems these assumptions may hold. For example with entanglement formed by ions interacting in a highly symmetric crystal lattice, whose ground-state is well known. It is more likely however, that this procedure would be applied to using low fidelity channels to transmit Bell states. Since the message, not initial the character of the entangled states,

is of cryptographic importance, Alice and Bob could communicate over un-secured classical channels and perform projective measurements until they were confident that their shared state satisfied all these assumptions.

We explain Bennett's et al.'s method in this report. It consists of two stages. In the first, partially entangled qubits are converted into a maximally entan-gled system in a smaller Hilbert space. In the second, any means for creating maximally entangled systems of unconstrained dimension is used to generate Bell states. Finally, we will explain Bennett et al.'s analysis of the efficiency of this method: demonstrating how this method approaches perfect yield as the size of the initial system increases.

3 The Entanglement Concentration and Standardization Protocol

3.1 Concentrating Entanglement

The first step towards creating Bell states is transforming the initial state into a state that is maximally entangled in some subspace of the original Hilbert space [3]. Bennett et al. call this part of the procedure *Schmidt projection* because it involves projecting the initial state into a subspace that corresponds to a particular binomial combination of the Schmidt coefficents.

We will now explain this procedure in more explicit detail than originally given in the paper: relating these subspaces a Hamming weight over the indices of the Schmidt vectors. This perspective gives a more detailed and clear picture of how the Bennett et al.'s method works.

As a warm-up before considering the initial state, consider the general Schmidt decomposition for a bipartite two-state system:

$$\cos\theta \left|\alpha_0\right\rangle \left|\beta_0\right\rangle + \sin\theta \left|\alpha_1\right\rangle \left|\beta_1\right\rangle$$

Following convention, the Schmidt coefficients are non-decreasing: so $\theta \in [0, \pi/4)$.

Recall that entanglement entropy is defined as the von Neumann entropy of the (either Alice's or Bob's) partial trace of a pure state density matrix. For a state in Schmidt form, this is just the Shannon entropy of the Schmidt coefficients squared. Clearly then, $\theta = 0$ corresponds to zero entanglement, while $\theta = \pi/4$ corresponds to maximal entanglement.

Now we move on to considering the initial state provided for this protocol. This state consists of n pairs of two state systems: each shared between Alice and Bob. As discussed in the in the introduction, we assume that we are provided with an initial state where entanglement occurs only pairwise, and that each pair be entangled to the same degree. Since entanglement entropy is a non-decreasing function of θ (which follows from its correspondence to Shannon entropy), each pair must therefore have the same Schmidt coefficients.

Therefore, under the given assumptions, the Schmidt decomposition of the ith pair is
$$\cos\theta \, |\alpha_0(i)\rangle \, |\beta_0(i)\rangle + \sin\theta \, |\alpha_1(i)\rangle \, |\beta_1(i)\rangle$$

And the total state of the system is

$$
\begin{aligned}
|\psi\rangle &= \bigotimes_{i=1}^{n} \left(\cos\theta \, |\alpha_0(i)\rangle \, |\beta_0(i)\rangle + \sin\theta \, |\alpha_1(i)\rangle \, |\beta_1(i)\rangle \right) \\
&= \sum_{k=0}^{n} \cos^{n-k}\theta \sin^k\theta \sum_{x \in H(k)} \bigotimes_{i=1}^{n} |\alpha_{x_i}(i)\rangle \, |\beta_{x_i}(i)\rangle \\
&= \sum_{k=0}^{n} \cos^{n-k}\theta \sin^k\theta \sum_{x \in H(k)} |\phi_x\rangle
\end{aligned}
$$

Where
$$H(k) = \{ x \in \mathbb{Z}_2^n : \text{the Hamming Weight of } x \text{ is } k \}$$
and
$$|\phi_x\rangle = \bigotimes_{i=1}^{n} |\alpha_{x_i}(i)\rangle \, |\beta_{x_i}(i)\rangle$$

In these steps, we have simply expanded the product and grouped by coefficients. In doing so we have realized that each coefficient corresponds to

vectors of a particular Hamming weight if we make the correspondence between states such as

$$|\alpha_1(1)\rangle\,|\beta_1(1)\rangle\,|\alpha_0(2)\rangle\,|\beta_0(2)\rangle\,|\alpha_0(3)\rangle\,|\beta_0(3)\rangle\,|\alpha_1(4)\rangle\,|\beta_1(4)\rangle$$

and the binary string $x = 1001 = 9$ for example.

Since we have used the Schmidt decomposition $\langle\alpha_j(i)|\alpha_k(i)\rangle = \langle\beta_j(i)|\beta_k(i)\rangle = \delta_{jk}$, and therefore $\langle\phi_x|\phi_y\rangle = \delta_{xy}$. This is useful, because it means that $V_k = \mathrm{span}(|\phi_x\rangle : x \in H(k))$ is a subspace of the full Hilbert space with dimension $|H(k)| = \binom{n}{k}$, and furthermore that V_k and $V_{k'}$ $(k \neq k')$ are orthogonal subspaces. Since they are orthogonal subspaces, a projective measurement may be performed that projects the system into exactly one of them.

Taking advantage of this, let Alice perform a von Neumann measurement that projects the state into one of these V_k subspaces. That is, let Alice to measure the observable A_H

$$A_H = \sum_{k=0}^{n} k \left(\sum_{x \in H(k)} \bigotimes_{i=1}^{n} |\alpha_{x_i}(i)\rangle\langle\alpha_{x_i}(i)| \right)$$

The resulting state is

$$|k\rangle = \binom{n}{k}^{-1/2} \sum_{x \in H(k)} |\phi_x\rangle \tag{1}$$

And it occurs with probability

$$p_k = \binom{n}{k} \cos^{2(n-k)}\theta \sin^{2k}\theta \tag{2}$$

Which is the familiar binomial distribution $B(n = n, p = \sin^2\theta)$.

For further use in the protocol described in section 3.2, Alice can communicate her result (k) to Bob over a classical channel, or she can let Bob perform

his own similar measurement: by virtue of the entanglement their results will always agree. It suffices to say that both Alice and Bob have access to this information, and that there is no reason to keep it as a secret.

This resulting state, $|k\rangle$, is a maximally entangled state in the $\binom{n}{k}^2$ dimensional subspace

$$\text{span}\left(\bigotimes_{i=1}^{n} |\alpha_{x_i}(i)\rangle |\beta_{y_i}(i)\rangle : x, y \in H(k)\right)$$

This is clear because the Schmidt coefficients (1) are uniform (recall that the entanglement entropy is the Shannon entropy of the Schmidt coefficients squared).

Using unitary operations, this state can be converted into any maximally entangled state between what are now considered to be two $\binom{n}{k}$ state systems. This is of little use in implementing conventional quantum algorithms and communication protocols however, since these are most often formulated in terms of maximally entangled 2 systems, i.e. Bell states.

Therefore for this procedure to be useful, we need a way of converting this maximally entangled $\binom{n}{k}$ state pair of systems into many maximally entangled two state systems. This is the subject of the following section.

3.2   Standardizing to Bell States

Bennett et al. outlines a strategy for converting the for converting a maximally entangled $\binom{n}{k}$ state system into component 2 state maximally entangled systems. In brief, it involves repeating the procedure of the previous section, combinatorially increasing the dimensionality of the the subspace in which the total state is maximally entangled, until the dimension is sufficiently close to a power of 2 [3]. Then one can project the system into a subspace with dimension 2^l with high probability [3]. The result is a maximally entangled state in a subspace of dimension 2^l, which can be decomposed into Bell states separate local unitary operations by Alice and Bob. In this section we explain this procedure in detail.

To begin, let us repeat the procedure of section 3.1 m times. The result is a sequence of 'k' values: $\vec{k} = (k_1, k_2, \ldots k_m)$; and the state $|K\rangle$

$$|K\rangle = \bigotimes_{\mu=1}^{m} |k_\mu\rangle$$

$$= D_m^{-1/2} \bigotimes_{\mu=1}^{m} \sum_{x^\mu \in H(k_\mu)} |\phi_{x^\mu}\rangle$$

$$= D_m^{-1/2} \sum_{x^1 \in H(k_1)} \cdots \sum_{x^m \in H(k_m)} \bigotimes_{\mu=1}^{m} |\phi_{x^\mu}\rangle$$

$$= D_m^{-1/2} \sum_{\vec{x} \in H(\vec{k})} |\xi_{\vec{x}}\rangle \tag{3}$$

Where we have defined $|\xi_{\vec{x}}\rangle = \bigotimes_{\mu=1}^{m} |\phi_{x^\mu}\rangle$ and $H(\vec{k}) = \{\vec{x} : x^\mu \in H(k_\mu)\}$. Since $\langle \phi_{x^\mu} | \phi_{y^\mu} \rangle = \delta_{x^\mu y^\mu}$, we know that $\langle \xi_{\vec{x}} | \xi_{\vec{y}} \rangle = \delta_{\vec{x}\vec{y}}$, and that (3) is the Schmidt decomposition of $|K\rangle$.

As in section 3.1, we treat $|K\rangle$ as a member of the D_m^2 dimensional subspace W_K.

$$D_m = |H(\vec{k})| = \prod_{\mu=1}^{m} \binom{n}{k_\mu}$$

$$W_K = \text{span} \left(\bigotimes_{\mu=1}^{m} \bigotimes_{i=1}^{n} \left| \alpha_{x_i^\mu}(i) \right\rangle \left| \beta_{y_i^\mu}(i) \right\rangle : x^\mu, y^\mu \in H(k_\mu) \right)$$

Of course, in section 3.1, m was equal to 1.

Since (3) is the Schmidt decomposition of $|K\rangle$, we see that the Schmidt coefficients of $|K\rangle$ are uniform, and that therefore $|K\rangle$ is maximally entangled in W_K.

Notice that if $D_m = 2^l$ for some integer l then the appropriate unitary operator will leave us with l Bell states [3]. To see this consider the unitary

7

operators U_A and U_B such that

$$U_A : \left\{ \bigotimes_{\mu=1}^{m} \bigotimes_{i=1}^{n} \left| \alpha_{x_i^\mu}(i) \right\rangle : x^\mu \in H(k_\mu) \right\} \to \{ |0\rangle, \dots |2^l - 1\rangle \}$$

$$U_B : \left\{ \bigotimes_{\mu=1}^{m} \bigotimes_{i=1}^{n} \left| \beta_{x_i^\mu}(i) \right\rangle : x^\mu \in H(k_\mu) \right\} \to \{ |0\rangle, \dots |2^l - 1\rangle \}$$

Then

$$U_A U_B |K\rangle = 2^{-l/2} \sum_{i=0}^{2^l - 1} |i\rangle_A |i\rangle_B$$

$$= \bigotimes_{j=1}^{l} \frac{1}{\sqrt{2}} \left(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B \right)$$

$$= \bigotimes_{j=1}^{l} |\beta_{00}\rangle$$

If $D_m \neq 2^l$, we continue increasing m until $D_m \in \bigcup_{l \in \mathbb{Z}} [2^l, 2^l(1 + \epsilon))$ for some $\epsilon > 0$, which is our chosen precision. As m goes to increases the probability of D_m remaining outside this set goes to zero [3]. This can be seen by analysing the sequence $\{D_m\}_{m=1}^{\infty}$ as a random walk [3]. Since increasing m correspondingly increases our yield of Bell states (better than linearly) there is no harm in not meeting this criterion immediately.

Once a suitable m has been found we perform a measurement that projects the system into one of two subspaces: one of dimension 2^l and the other of dimension $D_m - 2^l < \epsilon$. For example, Alice can measure the observable

$$A_G = \sum_{\vec{x} \in H(\vec{k}) \backslash X} \left(\bigotimes_{\mu=1}^{m} \bigotimes_{i=1}^{n} \left| \alpha_{x_i^\mu}(i) \right\rangle \left\langle \alpha_{x_i^\mu}(i) \right| \right)$$

Where X is any subset of $H(\vec{k})$ with cardinality $D_m - 2^l$. There is a probability ϵ/D_m that this measurement will output a 0, and project the system into the subspace spanned by X. We consider this a failure. On the other

8

hand there is a $1 - \epsilon/D_m$ probability that the measurement outputs 1, and the resulting state is

$$2^{-l/2} \sum_{\vec{x} \in H(\vec{k}) \backslash X} |\xi_{\vec{x}}\rangle$$

which can be readily decomposed into Bell states using unitary operators U_A and U_B as previously described.

4 Efficiency

In this section, we will consider the efficiency of concentrating entanglement: explaining the analysis of Bennett et al. [3]. We will show that the entanglement entropy of the resulting maximally entangled state approaches that of the initial state for large n.

The expected entanglement entropy after the Schmidt projection is

$$\langle S_E^f \rangle = \sum_{k=0}^{n} p_k S_E[|k\rangle]$$

$$= \sum_{k=0}^{n} \underbrace{\binom{n}{k} \cos^{2(n-k)} \theta \sin^{2k} \theta}_{p_k} \underbrace{\log_2 \binom{n}{k}}_{S_E[|k\rangle]}$$

Since, as we have already shown, $|k\rangle$ is maximally entangled as a $\binom{n}{k}$ state bipartite system.

Recall the entanglement entropy of the initial state was

$$S_E^i = -n \left(\cos^2 \theta \log_2 \left(\cos^2 \theta \right) + \sin^2 \theta \log_2 \left(\sin^2 \theta \right) \right)$$

To demonstrate the efficiency of our method, we wish to show that $\lim_{n \to \infty} \langle S_E^f \rangle / S_E^i = 1$. This could be achieve with pure mathematical analysis, but there more straightforward and physical approach.

Before taking her k measurement Alice sees her half of system as a mixed

state (tracing over Bob's half)

$$\rho_A = \sum_{k=0}^{n} \cos^{2(n-k)} \theta \sin^{2k} \theta \sum_{x \in H(k)} \bigotimes_{i=1}^{n} |\alpha_{x_i}(i)\rangle\langle\alpha_{x_i}(i)|$$

Therefore the probability distribution for x is classical as long as we stick to Alice's local perspective, and may be analysed with the tools of classical probability theory. Claude Shannon's theory of information tells us that Alice can reduce the entropy of x's probability distribution by on average no more than $S_H = -\sum_k p_k \log_2 p_k$ by determining the value of k that corresponds to x [5]. A simple example of this is determining the value of one out of ten coin-flips at most reduces the entropy from by 1 bit, from 10 bits to 9 bits, if the coins are uniformly random. The reader is referred to Shannon's seminal work for greater detail [5].

Notice that since ρ_A is diagonal, the entropy of entanglement is exactly the entropy of the probability distribution for x. Therefore our result from information theory is that

$$\langle S_E^f \rangle \geq S_E^i - S_H \Rightarrow \langle S_E^f \rangle / S_E^i \geq 1 - S_H / S_E^i$$

Recall from (2) that the the probability distribution for p_k is binomial: $B(n, p = sin^2\theta)$. It is well known that the entropy of the binomial distribution is

$$S_H = 1/2 \log_2 (2\pi enp(1-p)) + \mathcal{O}(1/n)$$

And therefore

$$\langle S_E^f \rangle / S_E^i = 1 - \mathcal{O}\left(\frac{\log(n)}{n}\right)$$

This results means that this procedure approaches a perfect yield of maximally entangled states as the system size increases.

5   Conclusions

We have explained how, in the context of common quantum communication protocols, the need for maximally entangled qubits arises. This requirement can be substituted by the weaker requirement of only partially entangled qubits, each with constant element entropy. We have explained the procedure proposed by Bennett et al. for doing so: including its merits, limitations, implementation and asymptotic efficiency.

This procedure can be done with only local operations by the two parties, although a classical communication channel can be used to reduce the required number of measurements by 1/2.

It is limited to concentrating entanglement on initially pure states with only monogamous entanglement, and constant pairwise entanglement entropy. We have described the implementation outline in Bennett et al.'s paper in new detail and clarity.

Finally we have explained how, using an information theoretic argument, Bennett et al. proved that this approach asymptotically conserves entanglement entropy as the initial system size increases. In total, Bennett et at.'s work represents a significant contribution towards practical quantum communication.

References

[1] MD Barrett et al. "Deterministic quantum teleportation of atomic qubits". In: *Nature* 429.6993 (2004), pp. 737–739.

[2] Charles H Bennett and Stephen J Wiesner. "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states". In: *Physical review letters* 69.20 (1992), p. 2881.

[3] Charles H Bennett et al. "Concentrating partial entanglement by local operations". In: *Physical Review A* 53.4 (1996), p. 2046.

[4] Richard P Feynman. "Simulating physics with computers". In: *International journal of theoretical physics* 21.6 (1982), pp. 467–488.

[5] Claude Elwood Shannon. "A mathematical theory of communication". In: *ACM SIGMOBILE Mobile Computing and Communications Review* 5.1 (2001), pp. 3–55.