QI Lecture 19

Quantum Networks

Entanglement Distillation

Entanglement is a resource in the sense that many, k, weakly entangled, E, qubits can be distilled into fewer, k', qubits that are strongly entangled, E', such that

$$kE \geq k'E'$$

meaning that after the distillation process, the total average entropy after distillation will not increase. The loss depends on efficiencies and the technique chosen. A full treatment of the optimal methods of distillation require a bit more time than we reasonably have, so we'll consider one (sub-optimal) entanglement distillation process: "**The Procrustean Method**".[1]

The Procrustean Method

Alice and Bob are assumed to begin with the state

$$|\eta\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle \in A \otimes B$$

where $\theta \in \left[0, \frac{\pi}{4}\right]$, just so that we can say (without loss of generalization) that $\cos(\theta) \geq \sin(\theta)$. The entanglement of $\rho = |\eta\rangle\langle\eta|$ is

$$\begin{aligned}
E &= S\left[\rho_a\right] \\
&= S\left[\begin{bmatrix} \cos^2(\theta) & 0 \\ 0 & \sin^2(\theta) \end{bmatrix}\right] \\
&= H\left[\left\{\cos^2(\theta), \sin^2(\theta)\right\}\right] \\
&= -\cos^2(\theta)\log_2\left(\cos^2(\theta)\right) - \sin^2(\theta)\log_2\left(\sin^2(\theta)\right)
\end{aligned}$$

So $0 \leq E < 1$ increases from 0 to 1, as θ ranges from 0 to $\frac{\pi}{2}$. Clearly, $|\eta\rangle$ is entangled, but not *maximally* entangled.

[1] Procrustes was a demigod son of Poseidon who spent his time tricking people into sleeping on his bed. If they were too tall for the bed he'd cut off their feet and if they were too short he'd rack them until they weren't. Nobody ever perfectly fit the bed, not even Procrustes himself, so he's considered more of "a jerk" than "a perfectionist".

In order to (hopefully) increase the entanglement, Alice and Bob will set up a POVM. Bob will add an ancilla qubit, $|0\rangle_c \in C$, so that the state of the system is

$$|\eta\rangle_{ab}|0\rangle_c = \cos(\theta)|000\rangle + \sin(\theta)|110\rangle \in A \otimes B \otimes C$$

and then applies a *local* unitary operation,

$$\begin{aligned} U_{bc}|00\rangle_{bc} &= \tan(\theta)|00\rangle_{bc} + \sqrt{1-\tan^2(\theta)}|01\rangle_{bc} \\ U_{bc}|10\rangle_{bc} &= |10\rangle_{bc} \end{aligned}$$

There are a lot of different unitary operations that will do this, most simply and in the $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ basis with

$$U_{bc} \sim \begin{bmatrix} \color{red}{\tan(\theta)} & -\sqrt{1-\tan^2(\theta)} & 0 & 0 \\ \color{red}{\sqrt{1-\tan^2(\theta)}} & \tan(\theta) & 0 & 0 \\ 0 & 0 & \color{red}{1} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where the **red** terms indicate the terms we need and the other terms are just "made up" to fit and make this operator unitary.[2]

The effect of this operator is

$$\begin{aligned} U_{bc}|\eta\rangle_{ab}|0\rangle_c &= U_{bc}\left[\cos(\theta)|000\rangle + \sin(\theta)|110\rangle\right] \\ &= \cos(\theta)\tan(\theta)|000\rangle + \cos(\theta)\sqrt{1-\tan^2(\theta)}|001\rangle + \sin(\theta)|110\rangle \\ &= \sin(\theta)|000\rangle + \sin(\theta)|110\rangle + \sqrt{\cos^2(\theta)-\sin^2(\theta)}|011\rangle \\ &= \sqrt{2}\sin(\theta)\left(\tfrac{|00\rangle+|11\rangle}{\sqrt{2}}\right)|0\rangle + \sqrt{\cos(2\theta)}|011\rangle \end{aligned}$$

Now if Bob measures C in the computational basis, then with $p(0) = 2\sin^2(\theta)$ the initial state will be projected into a maximally entangled state

$$U_{bc}|\eta\rangle_{ab}|0\rangle_c \quad \longrightarrow \quad |\eta_0\rangle = \frac{\langle 0|_c U_{bc}|\eta\rangle_{ab}|0\rangle_c}{\sqrt{p(0)}} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

and with probability $p(1) = \cos(2\theta) = 1 - 2\sin^2(\theta)$ the state will be projected into

[2]When you're only writing down a few relevant terms from a matrix you have tremendous freedom of choice. This is dangerous because the few terms you write down may not always be consistent with a unitary operation. This matrix is explicitly written out to remind you that sometimes you should double check that you really are talking about "some unitary operation".

$$U_{bc}|\eta\rangle_{ab}|0\rangle_c \quad \longrightarrow \quad |\eta_1\rangle = \frac{\langle 1|_c U_{bc}|\eta\rangle_{ab}|0\rangle_c}{\sqrt{p(0)}} = |01\rangle$$

Using their density matrices

$$\rho^{(0)} = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \qquad \rho^{(1)} = |01\rangle\langle 01|$$

we find that their entanglements are

$$E^{(0)} = S\left[\rho_a^{(0)}\right] = S\left[\frac{1}{2}|0\rangle_a\langle 0|_a + \frac{1}{2}|1\rangle_a\langle 1|_a\right] = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$$

$$E^{(1)} = S\left[\rho_a^{(1)}\right] = S\left[|0\rangle_a\langle 0|_a\right] = -1\log_2(1) = 0$$

The entanglement of the original state, $|\eta\rangle$, is E and the average entropy after the Procrustean Method is applied is

$$p(0)E^{(0)} + p(1)E^{(1)} = 2\sin^2(\theta)\cdot 1 + \cos(2\theta)\cdot 0 = 2\sin^2(\theta)$$

As we expect, the average entropy drops. Although it isn't obvious,

$$-\cos^2(\theta)\log_2\left(\cos^2(\theta)\right) - \sin^2(\theta)\log_2\left(\sin^2(\theta)\right) = E \geq p(0)E^{(0)} + p(1)E^{(1)} = 2\sin^2(\theta)$$

Even though the average entanglement drops, the result of Bob's measurement tells he and Alice (if he tells her) whether the entanglement is $1 = E^{(0)}$ or $0 = E^{(1)}$. In other words, you don't get to choose when $\rho = |\eta\rangle\langle\eta|$ will be distilled into $\rho^{(0)}$. Each individual instance has a probability $p = 2\sin^2(\theta)$ of being distilled.

So we should consider the aggregate behavior. If Alice and Bob share N copies of $|\eta\rangle_{ab}$, then they can distill them down to

$$Np = 2\sin^2(\theta)N$$

copies of $|\Phi_+\rangle$. Or, to use needlessly mathy language

$$\left(|\eta\rangle\langle\eta|\right)^{\otimes N} \quad \longrightarrow \quad \left(|\Phi_+\rangle\langle\Phi_+|\right)^{\otimes Np}$$

meaning that, <u>on average</u>, you can turn N copies of $|\eta\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle$ into $Np = 2\sin^2(\theta)N$ copies of $|\Phi_+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$.

Notice that $p = 2\sin^2(\theta) = 1$ when $\theta = \frac{\pi}{4}$ and $|\eta\rangle = \cos\left(\frac{\pi}{4}\right)|00\rangle + \sin\left(\frac{\pi}{4}\right)|11\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$. In other words: you can distill N maximally entangled states from N maximally entangled states.

On the other end of the scale, $p = 2\sin^2(\theta) = 0$ when $\theta = 0$ and $|\eta\rangle = \cos(0)|00\rangle + \sin(0)|11\rangle = |00\rangle$. In other words: if you have N separable states, then you can't distill any entangled states.

Optimal Entanglement Distillation and Dilution

First a quick note about typical sequences. The probability of event from a random variable, $x \in X$, is $p(x)$ and if every trial is independent, then the probability of a particular string of events is

$$p(x_1 x_2 \ldots x_n) = \prod_{j=1}^{n} p(x_j)$$

where every x_j is randomly selected from X. By the law of large numbers

$$\lim_{n\to\infty} -\frac{1}{n}\log_2(p(x_1\ldots x_n)) = \lim_{n\to\infty} -\frac{1}{n}\sum_{j=1}^{n}\log_2(p(x_j)) = E[-\log_2(p)] = H[X]$$

or written another way (and using the definition of the limit), for large enough values of n

$$\left|-\frac{1}{n}\log_2(p(x_1\ldots x_n)) - H[X]\right| \leq \epsilon$$

$$|\log_2(p(x_1\ldots x_n)) + nH[X]| \leq n\epsilon$$

$$-n\epsilon \leq \log_2(p(x_1\ldots x_n)) + nH[X] \leq n\epsilon$$

$$-nH[X] - n\epsilon \leq \log_2(p(x_1\ldots x_n)) \leq -nH[X] + n\epsilon$$

For any given ϵ, an "ϵ-**typical sequence**", $x_1 x_2 \ldots x_n$, has a probability between

$$2^{-n(H[X]+\epsilon)} \leq p(x_1 x_2 \ldots x_n) \leq 2^{-n(H[X]-\epsilon)}$$

and for any given δ and sufficiently large n, the number of ϵ-typical sequences, T, satisfies

$$(1-\delta)2^{n(H[X]-\epsilon)} \leq T \leq 2^{n(H[X]+\epsilon)}$$

and the probability of picking a sequence at random and finding that it's ϵ-typical is $1-\delta$ for large enough n.

In short,

For large values of n, we can assume that all sequences are typical: there are only $2^{nH[X]}$ sequences and they're all equally probable with $p = 2^{-nH[X]}$.

This is the "asymptotic equipartition principle". It says that, for example, if you roll a bucket of dice, you are almost guaranteed to see that all six numbers show up about equally often. Since these results are typical, they're the only one's you'll ever see.

If you have N Bell states, $|\Phi_+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}} \in A \otimes B$, and you find that you can produce m copies[3] of

$$|\eta\rangle = \sum_j \sqrt{p(j)}|v_j\rangle_a|w_j\rangle_b \in A \otimes B$$

then $|\eta\rangle$'s "**Entanglement of Formation**" is the limit $\frac{N}{m} \to E_f$. $\frac{N}{m}$ is the average number of ebits needed to create $|\eta\rangle$.

If you have m' copies of $|\eta\rangle$ and find that you can make N copies of $|\Phi_+\rangle$, then $|\eta\rangle$'s "**Distillable Entanglement**" is $\frac{N}{m'} \to E_d$. $\frac{N}{m'}$ is the average number of ebits recoverable when destroying $|\eta\rangle$.

For pure states (and perfect equipment) $m = m'$ and $E_f = E_d$!

As with classical communication over a noisy channel, quantum communication is optimal over many uses of entangled states. In order to establish m copies of the entangled state $|\eta\rangle \in A \otimes B$ with Bob, Alice can first prepare many complete copies of $|\eta\rangle \in A \otimes A'$ and then teleport A' to Bob all at once, possibly using an N-qubit teleportation scheme.[4] We've seen something similar to this before[5] when we considered data compression. On a bit-by-bit basis compression doesn't make much sense, but long strings of low entropy bits (e.g., 000000000000100000...) can be compressed into short strings of high-entropy bits in such a way that the total entropy stays the same. The idea here is the same; the total entanglement stays the same and distillation becomes effectively optimal for large collections of states.

Define $H[X] = -\sum_j p(j)\log_2(p(j))$ using the squared amplitudes from $|\eta\rangle = \sum_j \sqrt{p(j)}|v_j\rangle_a|w_j\rangle_b$. Notice that the entanglement, E, of $|\eta\rangle$ is

$$E = S\left[Tr_b\left[|\eta\rangle\langle\eta|\right]\right] = S\left[\sum_j p(j)|v_j\rangle_a\langle v_j|_a\right] = H[X]$$

Alice begins with m copies

[3]This is the "Schmidt Decomposition" for pure states in composite spaces, introduced in lecture 10.
[4]For example, using the protocol from the homework in lecture 13.
[5]See lecture 7.

$$|\eta\rangle^{\otimes m} = \sum_{x_1x_2\ldots x_m} \sqrt{p(x_1)p(x_2)\ldots p(x_m)}|v_{x_1}v_{x_2}\ldots v_{x_m}\rangle_a|w_{x_1}w_{x_2}\ldots w_{x_m}\rangle_{a'}$$

which is very closely approximated by

$$|\varphi\rangle = \kappa \sum_{typical} \sqrt{p(x_1)p(x_2)\ldots p(x_m)}|v_{x_1}v_{x_2}\ldots v_{x_m}\rangle_a|w_{x_1}w_{x_2}\ldots w_{x_m}\rangle_{a'}$$

which is the same sum with all of the non-typical states removed and renormalized with $\kappa \approx 1$. These are very nearly the same state in the sense that $\langle\eta|\varphi\rangle \approx 1$.[6] Define $\rho_\eta = Tr_b[|\eta\rangle\langle\eta|^{\otimes m}]$ and $\rho_\varphi = Tr_b[|\varphi\rangle\langle\varphi|]$.

Tracing over A', we find that

$$\rho_\varphi = Tr_{a'}\left[|\varphi\rangle\langle\varphi|^{\otimes m}\right] = \sum_{typical} p(x_1)p(x_2)\ldots p(x_m)|v_{x_1}v_{x_2}\ldots v_{x_m}\rangle_a\langle v_{x_1}v_{x_2}\ldots v_{x_m}|_a$$

Since $\{|v_{x_1}v_{x_2}\ldots v_{x_m}\rangle_a\rangle\}$ are an orthonormal basis, we can find the entropy (and thus entanglement) of this easily. Remembering that $p(x_1,x_2,\ldots,x_m) \approx 2^{-mH[X]}$ and there are approximately $2^{mH[X]}$ terms in the sum,

$$\begin{aligned}
S[\rho_\varphi] &= -\textstyle\sum_{x_1x_2\ldots x_m} p(x_1)p(x_2)\ldots p(x_m)\log_2\left(p(x_1)p(x_2)\ldots p(x_m)\right) \\
&\approx -\textstyle\sum_{x_1x_2\ldots x_m} 2^{-mH[X]}\log_2\left(2^{-mH[X]}\right) \\
&\approx -2^{mH[X]}\left[2^{-mH[X]}\log_2\left(2^{-mH[X]}\right)\right] \\
&= -\log_2\left(2^{-mH[X]}\right) \\
&= mH[X] \\
&= mE
\end{aligned}$$

and therefore the total entanglement is approximately $mH[X]$, which follows nicely from the entanglement of an individual state[7]

$$S[\rho_\eta] = mS\left[Tr_{a'}\left[|\eta\rangle\langle\eta|\right]\right] = -m\sum_j p(j)\log_2(p(j)) = mH[X] = mE$$

Before seeing why typical sequences and the whole approach above is useful, we need to introduce one more idea: "majorization". Given two vectors, $\vec{v},\vec{w}$, written in decreasing order such that $v_i \geq v_{i+1}$ for all i, we say that "$\vec{w}$ majorizes $\vec{v}$", written and defined

[6] There are several different ways to describe how similar two states are. "Fidelity", $F[\rho,\sigma] \equiv Tr\left[\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}\right]$, ranges from 0 to 1 and is a useful measure for density matrices.

[7] And the additivity of the entanglement of independent systems (see lecture 10).

$$\vec{v} \prec \vec{w} \qquad \Leftrightarrow \qquad \sum_{k=1}^{d} v_k \leq \sum_{k=1}^{d} w_k,\ \forall d$$

Theorem. *Assume the entangled states $|\nu\rangle, |\phi\rangle \in A \otimes B$, have reduced density matrices $\rho_\nu, \rho_\phi \in A$, and that those density matrices have eigenvalues $\vec{\lambda}_\nu, \vec{\lambda}_\phi$, where these vectors are the eigenvalues listed in decreasing order, $\lambda_{\nu j} \geq \lambda_{\nu(j+1)}$, so that the largest eigenvalue is first.*

Using only LOCC, we can transform, $|\nu\rangle \to |\phi\rangle$, if and only if $\vec{\lambda}_\nu \prec \vec{\lambda}_\phi$, the eigenvalues of ρ_ϕ majorizes the eigenvalues of ρ_ν.

This theorem is why typical sequences are so important. The reduced density matrix of $|\Phi_+\rangle^{\otimes N}$ is

$$Tr_b\left[|\Phi_+\rangle\langle\Phi_+|^{\otimes N}\right] = 2^{-N} \sum_{x=0}^{2^N-1} |x\rangle_a \langle x|_a$$

Notice that the only eigenvalue is 2^{-N}, and it's repeated 2^N times. If we have that[8]

$$2^{-mH[X]} < 2^{-N}$$

then padding the vector of eigenvalues of $Tr_b\left[|\Phi_+\rangle\langle\Phi_+|^{\otimes N}\right]$ with zeros as needed,

$$\{2^{-mH[X]}, 2^{-mH[X]}, \ldots, 2^{-mH[X]}\} \prec \{2^{-N}, 2^{-N}, \ldots, 2^{-N}, 0, \ldots, 0\}$$

and therefore we can transform m copies of $|\eta\rangle$ into N copies of $|\Phi_+\rangle$ where $N < mH[X]$. More succinctly

$$N < mE$$

where we asymptotically approach equality as $m, N \to \infty$. N here is quietly being multiplied by 1, the entanglement of a Bell state.

Distillation is a very important part of building a quantum network. Weakly entangled states are bad at teleportation, so if you want to teleport a qubit between two locations, you'd started by distilling out a few maximally entangled ebits.[9] Once a maximally entangled state has been established, we can use it as a noiseless quantum channel.

[8] The inequality is strict, because we need 2^{-N} to be larger than the largest eigenvalue of the "typical state" $|\varphi\rangle$, which is a little larger than $2^{-mH[X]}$.

[9] "ebit" = "entangled qubit", the amount of entanglement that a Bell state has.

Quantum Repeaters

Quantum repeaters and networks solve two problems: how do you set up quantum channels over distances greater than you can physically send quantum information[10] and how do you send quantum information to multiple destinations when you have access to only one quantum channel?

If you want to teleport a qubit over an even greater range, you have a couple of choices: you can teleport the qubit from A to B then B to C, or you can chain entanglements AB and BC to create a new entanglement AC.

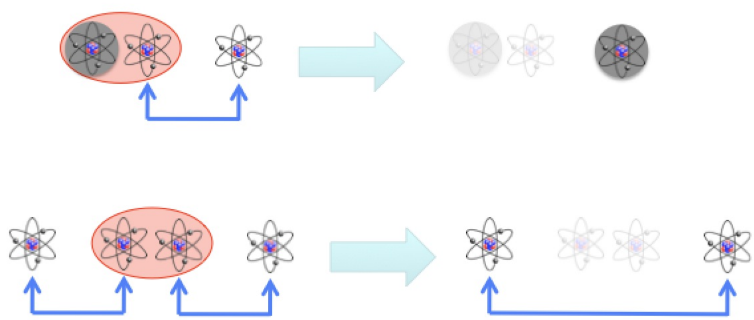

Figure 1: Red ovals are Bell basis measurements and blue arrows are entanglement. Top: With entangled pairs, we can teleport a qubit. Bottom: With two entangled pairs we can "chain" the entanglement, by teleporting one entangled qubit using the other pair. You can picture this as being like connecting a pair of hoses to make one long hose.

Here we'll look at a very reasonable question: what happens when you teleport an entangled qubit?

Example In exactly the same way that we can teleport an ordinary qubit, we can teleport an entangled qubit. In other words, if Alice and Bob share $|\Psi_-\rangle_{ab}$ and Bob and Carol share $|\Phi_+\rangle_{cd}$, then Bob can use qubit C, to teleport qubit B to D. Alice and Carol will then share $|\Psi_-\rangle_{ad}$. The initial state of the pair of qubits is

$$|\Psi_-\rangle_{ab} \otimes |\Phi_+\rangle_{cd} = \left(\frac{|0\rangle_a|1\rangle_b - |1\rangle_a|0\rangle_b}{\sqrt{2}}\right)\left(\frac{|0\rangle_c|0\rangle_d + |1\rangle_c|1\rangle_d}{\sqrt{2}}\right) = \frac{|0100\rangle + |0111\rangle - |1000\rangle - |1011\rangle}{2}$$

Alice is in control of qubit A, Bob has qubits B and C, and Carol has qubit D.

[10]The farthest that an uncorrupted photon can be sent through air or fibre optic cable seems to be on the order of 100 km. This isn't a fundamental limit, but an engineering limit (for now) and is a big part of the motivation behind satellite quantum communication.

The standard teleportation scheme when using $|\Phi_+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, is for Bob to measure the to-be-sent qubit and his half of the entangled pair in the Bell basis, and with result $|B_{jk}\rangle$ Carol applies Z^jX^k to her half of the entangled pair, D, so that it assumes the to-be-sent qubit.

In order to teleport B to Carol, Bob needs to make a measurement of $B\otimes C$ in the Bell basis. To make things easier to see, we'll write the middle two qubits in the Bell basis. This will be easier if we rearrange the order of the qubits, so it's important to be very careful about the subscripts.

$$
\begin{aligned}
|\Psi_-\rangle_{ab}\otimes|\Phi_+\rangle_{cd} &= \frac{|0100\rangle+|0111\rangle-|1000\rangle-|1011\rangle}{2} \\
&= \frac{1}{2}|10\rangle_{bc}|00\rangle_{ad}+\frac{1}{2}|11\rangle_{bc}|01\rangle_{ad}-\frac{1}{2}|00\rangle_{bc}|10\rangle_{ad}-\frac{1}{2}|01\rangle_{bc}|11\rangle_{ad} \\
&= \begin{cases} \frac{1}{2}\left[\frac{|B_{01}\rangle_{bc}-|B_{11}\rangle_{bc}}{\sqrt{2}}\right]|00\rangle_{ad}+\frac{1}{2}\left[\frac{|B_{00}\rangle_{bc}-|B_{10}\rangle_{bc}}{\sqrt{2}}\right]|01\rangle_{ad} \\ -\frac{1}{2}\left[\frac{|B_{00}\rangle_{bc}+|B_{10}\rangle_{bc}}{\sqrt{2}}\right]|10\rangle_{ad}-\frac{1}{2}\frac{|B_{01}\rangle_{bc}+|B_{11}\rangle_{bc}}{\sqrt{2}}|11\rangle_{ad} \end{cases} \\
&= \begin{cases} \frac{1}{2}|B_{00}\rangle_{bc}\left[\frac{|01\rangle_{ad}-|10\rangle_{ad}}{\sqrt{2}}\right]+\frac{1}{2}|B_{01}\rangle_{bc}\left[\frac{|00\rangle_{ad}-|11\rangle_{ad}}{\sqrt{2}}\right] \\ -\frac{1}{2}|B_{10}\rangle_{bc}\left[\frac{|01\rangle_{ad}+|10\rangle_{ad}}{\sqrt{2}}\right]-\frac{1}{2}|B_{11}\rangle_{bc}\left[\frac{|00\rangle_{ad}+|11\rangle_{ad}}{\sqrt{2}}\right] \end{cases}
\end{aligned}
$$

When Bob makes a measurement in $B\otimes C$ it leaves the qubits in $A\otimes D$ in a new state. Following the teleportation protocol exactly as normal, Carol applies the Pauli matrices, X_d and/or Z_d, as prescribed by the results of Bob's measurement.

$$
\begin{array}{c|lllllll}
|B_{00}\rangle & \frac{|01\rangle_{ad}-|10\rangle_{ad}}{\sqrt{2}} & = & |\Psi_-\rangle_{ad} & & & & \\
|B_{01}\rangle & \frac{|00\rangle_{ad}-|11\rangle_{ad}}{\sqrt{2}} & \xrightarrow{X_d} & \frac{|01\rangle_{ad}-|10\rangle_{ad}}{\sqrt{2}} & = & |\Psi_-\rangle_{ad} & & \\
|B_{10}\rangle & -\frac{|01\rangle_{ad}+|10\rangle_{ad}}{\sqrt{2}} & \xrightarrow{Z_d} & \frac{|01\rangle_{ad}-|10\rangle_{ad}}{\sqrt{2}} & = & |\Psi_-\rangle_{ad} & & \\
|B_{11}\rangle & -\frac{|00\rangle_{ad}+|11\rangle_{ad}}{\sqrt{2}} & \xrightarrow{X_d} & -\frac{|01\rangle_{ad}+|10\rangle_{ad}}{\sqrt{2}} & \xrightarrow{Z_d} & \frac{|01\rangle_{ad}-|10\rangle_{ad}}{\sqrt{2}} & = & |\Psi_-\rangle_{ad}
\end{array}
$$

By teleporting an entangled qubit to Carol, Carol now has an entangled qubit. In fact, Alice and Carol now share the exact entangled state, $|\Psi_-\rangle_{ad}$, that Alice and Bob had shared, $|\Psi_-\rangle_{ab}$.

■

Evidently, when you teleport a qubit, you also teleport that qubit's entanglement. In this way we can chain entanglements together to create a single entanglement.

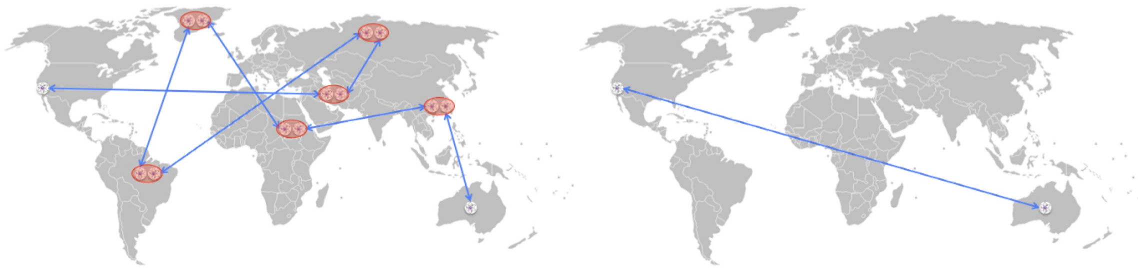

Figure 2: With a quantum network, where each "quantum server" has a collection of qubits entangled with several other servers, we can establish a simple entanglement between any two points on the network.

It turns out that a string of teleportations done together is easier than the sum of each individual teleportation. This is because each teleportation applies some combination of X and Z to the teleported qubit and the wonderful properties of the Pauli matrices (specifically, $X^2 = Z^2 = I$, $XZX = -X$, and $ZXZ = -Z$) mean that we can do a *single* correction for all of the teleportations when we're done.

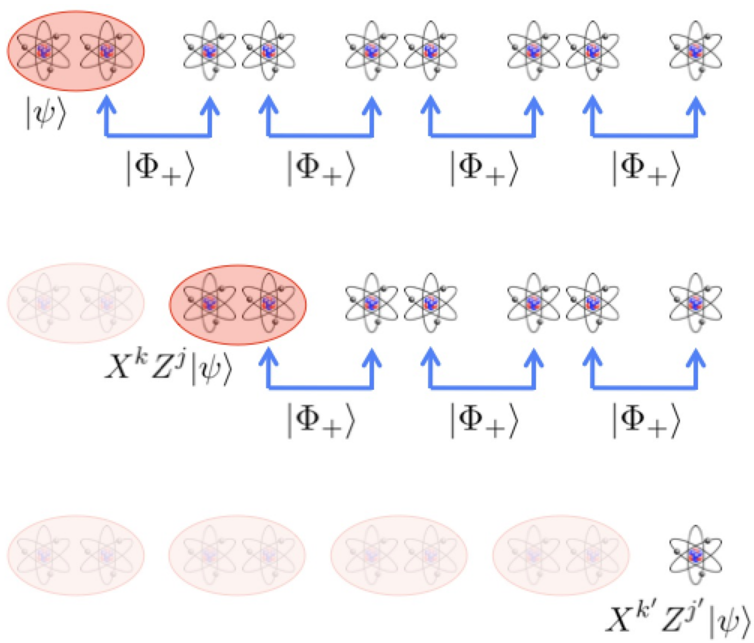

Figure 3: Every time a qubit is passed along in a "teleportation chain" it picks up some combination of new Pauli matrices. Ultimately, we don't need to deal with all of those Pauli matrices, just their aggregate effect.

Example $|\psi\rangle$, the state of a qubit, is teleported five times through five $|\Phi_+\rangle$ states. The results of the measurements are (in order) $|B_{11}\rangle, |B_{00}\rangle, |B_{01}\rangle, |B_{11}\rangle, |B_{10}\rangle$ and therefore the state of the teleported qubit at the end of its journey is

$$(Z)(XZ)(X)(I)(XZ)|\psi\rangle = ZXZXXZ|\psi\rangle$$

Since $XZ = -ZX$, we can rearrange these operators into groups

$$\begin{aligned} ZXZXXZ|\psi\rangle &= -ZZXXXZ|\psi\rangle \\ &= ZZXXZX|\psi\rangle \\ &= -ZZXZXX|\psi\rangle \\ &= ZZZXXX|\psi\rangle \end{aligned}$$

and since $X^2 = Z^2 = I$,

$$\begin{aligned} ZXZXXZ|\psi\rangle &= ZZZXXX|\psi\rangle \\ &= ZXXX|\psi\rangle \\ &= ZX|\psi\rangle \end{aligned}$$

So applying Z and then X will restore the original qubit and complete the teleportation. In fact, since $XZ = -ZX$, applying X and then Z will also restore the qubit, at the (free) cost of a global phase.

What we've just discovered is that

> The only thing we need is the parity of the number of measurements inducing an X or Z.

The parity of a number is its odd/even-ness, and in this case since X and Z both appeared an odd number of times (doesn't matter how many total), we know that the end state must be

$$\pm XZ|\psi\rangle$$

■

Exercises

1) Quantum Forwarding

Normally in order to teleport a qubit, the sender measures it and their entangled qubit in the Bell basis (getting $|B_{jk}\rangle$), sends the result (j,k) to the receiver, and then the receiver restores the original state (by applying $Z^j X^k$'s to their qubit). But in this case $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is teleported $A \to B \to C$ without being restored. That is, although the sender makes measurements, the receiver does not alter the state with Pauli matrices.

A copy of $|B_{00}\rangle = |\Phi_+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ is shared between Alice and Bob and another between Bob and Carol. In other words, the initial state is

$$|\Omega\rangle = |\psi\rangle_p |\Phi_+\rangle_{qr} |\Phi_+\rangle_{st}$$

where p, q belong to Alice, r, s belong to Bob, and t belongs to Carol. Alice's measurement of pq results in $|B_{11}\rangle$ and Bob's measurement of rs results in $|B_{10}\rangle$.

a) What is the state of the qubit after the (unrestored) teleportation from $A \to B$? After the second (unrestored) teleportation from $B \to C$?

b) Assuming that Carol receives both measurement results, what corrective operations need to be applied to the qubit?

c) It is discovered that Bob did his measurement early, before Alice. Find the state of the full system after Bob's measurement resulting in $|B_{10}\rangle_{rs}$. Find the state of the full system after Alice's result of $|B_{11}\rangle_{pq}$.

d) How does the order of these measurements affect the state that Carol receives and the corrections she needs to make?

2) Braiding Rope

There are two pairs of qubits, ρ and ρ', with each pair in the mixed state

$$\rho = p|\Phi_+\rangle\langle\Phi_+| + (1-p)|\Phi_-\rangle\langle\Phi_-|$$

where $p > \frac{1}{2}$ and $|\Phi_\pm\rangle = \frac{|00\rangle\pm|11\rangle}{\sqrt{2}}$. Alice and Bob each have a qubit from each pair, A and B from one pair and A' and B' from the other.

a) Alice applies $e^{i\frac{\pi}{4}X} = e^{i\frac{\pi}{4}}|+\rangle\langle+| + e^{-i\frac{\pi}{4}}|-\rangle\langle-|$ to both A and A' individually and Bob applies $e^{-i\frac{\pi}{4}X} = e^{-i\frac{\pi}{4}}|+\rangle\langle+| + e^{i\frac{\pi}{4}}|-\rangle\langle-|$ to both B and B' individually.

After applying these operations, what are the density matrices ρ and ρ'?

b) Alice sends her qubits through a CNOT gate, with A as the control and A' as the target. Bob does the same; B is the control and B' is the target.

After applying these CNOTs, what is the state that remains?

c) In the computational basis, Alice measures A' and Bob measures B'. They keep A and B if their results agree.

What is the state that remains if Alice and Bob both see 0? If they both see 1?

d) Show that if the results agree, then p, the probability of ρ being $|\Phi_+\rangle\langle\Phi_+|$, increases.