Quantum Information Theory - 67749
Recitation 2, May 13, 2025

1 Recitation Overview - Quantum States as Computational Resources.

In the last lectures, we saw that quantum states can be considered as resources, presenting the quantum AEP concept and the compression theorem. In this recitation, we continue that line. We'll start by completing the compression theorem proof. Then we will present (and represent) simple applications of quantum states, including going over superdense decoding and quantum teleportation again, and presenting gate teleportation and quantum money. Then we will strengthen our understanding of mixed states by presenting and proving Uhlmann's theorem. Finally, we will discuss EPR distillation (entanglement purification), a process by which one tries to extract from a noisy EPR mixture much better EPR pairs.

2 Asymptotic Equipartition Property and The Compression Theorem[1].

Analogy to the classical case, the quantum AEP states that only a subspace of the outcomes is likely to be drawn from a product distribution. In contrast to the classical case, here the distribution is induced by taking copies of a mixed state $\rho^{\otimes n}$, and we understand the term 'likely outcomes' by asking what would be the most expected states to measure when we measure in the diagonalized basis of $\rho^{\otimes n}$. The theorem formalizes into:

Theorem 2.1 (Quantum AEP)**.** *Let $\varepsilon, \delta > 0$. Let P_T be the projection over the subspace spanned by the (classical) typical set given by measuring ρ^n in its diagonalized basis. Then:*

[1] Taken from Prof. Kochman notes.

1. ***Direct.*** *For any n large enough, the typical subspace T with underlying ϵ satisfies:*

$$\mathbf{Tr}\left(\rho^{\otimes n} P_T\right) \geq 1 - \delta.$$

2. ***Converse.*** *For every subspace $\tilde{T} = \tilde{T}(n, \varepsilon, \delta)$ in $\mathcal{H}_d^{\otimes n}$ of dimension $2^{n(S(\rho)-\epsilon)}$, there exists n large enough, such that:*

$$\mathbf{Tr}\left(\rho^{\otimes n} P_{\tilde{T}}\right) < \delta.$$

Having typical behavior, we can quantify how many (q)bits one has to pay to compress states drawn from given distributions. Recall that we define the fidelity of the compression protocol as the expected fidelity over the drawn state and the mixed state yielded by a decoding-encoding round. That gives us the compression theorem [Sch95]:

Theorem 2.2 (The Quantum Coding Theorem, Schumacher 1995)**.** *Let ψ be a quantum source with density matrix ρ and let $\epsilon, \delta > 0$.*

1. ***Direct:*** *For every large enough n, there exists a coding scheme of rate $R < S(\rho) + \epsilon$, and fidelity $F > 1 - \delta$.*

2. ***Converse:*** *For every large enough n, the fidelity of any coding scheme with $R < S(\rho) - \epsilon$ satisfies $F < 1 - \delta$.*

Remark 2.1. *Observe that the fidelity of the compression is not identical to the fidelity between two mixed states.*

2.1 Proof of the Converse's part.

Proof. We shall restrict our attention to unitary decoders only, since the proof for the general case is much more complicated. Denote by $\Omega_a \in \Delta(\mathcal{H}_{2^{nR}})$ the density matrix of the state $|a\rangle$, of the source A_n, after encoding:

$$\begin{aligned} &\Omega_a = E(|a\rangle) \in \Delta\left(\mathcal{H}_{2^{nR}}\right), \\ &\dim \Omega_a \leq 2^{nR}. \end{aligned}$$

Unitary Decoder: Denote by ω_a the reconstructed state at the decoder:

$$\omega_a = D(\Omega_a) = U\left(\underbrace{\Omega_a}_{\dim \leq 2^{nR}} \otimes |0\rangle\langle 0|\right) U^\dagger,$$

for some unitary operator U. Note that $\dim(\omega_a) = \dim(\Omega_a) \leq 2^{nR}$, and therefore there exists some subspace Λ_n of $H_d^{\otimes n}$, of dimension 2^{nR}, such that:

$$\forall\, |a\rangle \in A_n : \quad \text{support}(\omega_a) \subseteq \Lambda_n\,, \qquad \dim(\Lambda_n) = 2^{nR}\,.$$

Hence, ω_a has an orthonormal decomposition with a basis $\left\{|\xi_1^{(a)}\rangle\,, ..., |\xi_{2^{nR}}^{(a)}\rangle\right\}$ laying within the subspace Λ_n:

$$\omega_a = \sum_{j=1}^{2^{nR}} q_j^{(a)}\, |\xi_j^{(a)}\rangle\, \langle \xi_j^{(a)}|\,,$$

where $\{q_j^{(a)}\}$ are non-negative and sum to 1. The fidelity of $|a\rangle$ and its corresponding 'quantized' state ω_a is

$$\begin{aligned} F\left(|a\rangle\,, \omega_a\right) &= \langle a|\omega_a|a\rangle = \sum_{j=1}^{2^{nR}} q_j^{(a)}\, \langle a|\xi_j^{(a)}\rangle\, \langle \xi_j^{(a)}|a\rangle = \sum_{j=1}^{2^{nR}} q_j^{(a)}\, \langle \xi_j^{(a)}|a\rangle\, \langle a|\xi_j^{(a)}\rangle \\ &\leq \sum_{j=1}^{2^{nR}} \langle \xi_j^{(a)}|\left(|a\rangle\, \langle a|\right)|\xi_j^{(a)}\rangle = \sum_{j=1}^{2^{nR}} \langle \xi_j^{(a)}|\, P_a\, |\xi_j^{(a)}\rangle = \sum_{j=1}^{2^{nR}} \mathbf{Tr}\left(P_a\, |\xi_j^{(a)}\rangle\, \langle \xi_j^{(a)}|\right) \\ &= \mathbf{Tr}\left(P_a \sum_{j=1}^{2^{nR}} |\xi_j^{(a)}\rangle\, \langle \xi_j^{(a)}|\right) = \mathbf{Tr}(P_a P_{\Lambda_n})\,, \end{aligned}$$

where the inequality holds true since $q_j^{(a)} \leq 1$, and P_a and P_{Λ_n} are the projectors on $|a\rangle$ and the subspace Λ_n, respectively.

Remark 2.2. *$|a\rangle$ does not necessarily lay inside Λ_n.*

The fidelity of the coding scheme can be upper bounded by:

$$\begin{aligned} F = &\sum_{(p_a,|a\rangle)\in A_n} p_a\, \langle a|\omega_a|a\rangle \leq \sum_{(p_a,|a\rangle)\in A_n} p_a \mathbf{Tr}(P_a P_{\Lambda_n}) \\ =&\mathbf{Tr}\left(P_{\Lambda_n} \sum_{(p_a,|a\rangle)\in A_n} p_a\, |a\rangle\, \langle a|\right) = \mathbf{Tr}\left(P_{\Lambda_n} \rho^{\otimes n}\right)\,. \end{aligned}$$

But by the converse to the quantum AEP, this must be low. □

3 Quantum Money.

First application of the non-cloning theorem is quantum money, which is a scheme for banks to print money that will be hard to forge. In the protocol, any banknote (coin) is associated with a number s, and the bank computes $f(s)$ according to some chosen function f which is hard to invert. Then it initializes a quantum state $|\psi_s\rangle$ such that the i-th qubit is:

1. $|0\rangle$ if $s_{2i}s_{2i+1} = 00$
2. $|1\rangle$ if $s_{2i}s_{2i+1} = 01$
3. $|+\rangle$ if $s_{2i}s_{2i+1} = 10$
4. $|-\rangle$ if $s_{2i}s_{2i+1} = 11$

Now the bank gives the client the pair $(f(s), |\psi_s\rangle)$ and keeps a table which, for any key $f(s)$, stores the value s.

The verification of a banknote is as follows: the bank asks for the coin $(f(s), |\psi_s\rangle)$. Then, by looking at the table, it restores s and measures the ith qubit in the basis according to s_{2i} and compares the result to s_{2i+1}. If one of the measurements doesn't match the expected, then it returns failure.

If the chosen f were a pseudorandom function, then even when $f(s)$ is given, the ith qubit of $|\psi_s\rangle$ appears to a polynomial observer as the fully mixed state, or in other words, having $f(s)$ doesn't help in preparing another copy of $|\psi_s\rangle$. Combined with the fact that $|\psi_s\rangle$ can't be cloned, we get that the banknote can't be duplicated.

4 Superdense Encoding.

The idea behind superdense coding is that a local operation over $|\mathbf{EPR}\rangle = |\beta_{00}\rangle$, namely an operation on the partition held by Alice, sends it to any of the other Bell states. For example, $X \otimes I_{\text{Bob}} |\mathbf{EPR}\rangle \to \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) = |\beta_{10}\rangle$. Thus, Alice can change the complete encoded state shared between Alice and Bob by acting only over her local system. Yet, as we have seen, for Bob to infer what that state is, he has to hold both qubits, otherwise, his local view is the uniformly mixed state over a single qubit. So the protocol works as follows:

1. We assume that Alice and Bob share an $|\mathbf{EPR}\rangle$, each holding a single qubit.
2. Alice applies one of I, X, Z, XZ on her qubit to transform the shared state to one of the following $\beta_{00}, \beta_{10}, \beta_{01}, \beta_{11}$ and sends it to Bob.
3. Bob decodes the qubit pair by rotating the state into the computational basis and then measuring.

In total, we have a protocol that, by sending a single qubit through the channel, passes two classical bits. Notice that any attempt to listen to the channel[2] would reveal no information since, again, when squeezing the view into a single qubit, it seems uniformly mixed.

[2]The assumption that Alice and Bob initially held a valid **EPR** pair is important, yet much weaker than assuming they have a private key.

5 Quantum Teleportation.

Quantum teleportation is a gadget (method) to transfer a quantum state from one location to another, without the physical transfer of the underlying particles themselves. This is achieved through the consumption of **EPR** pairs.

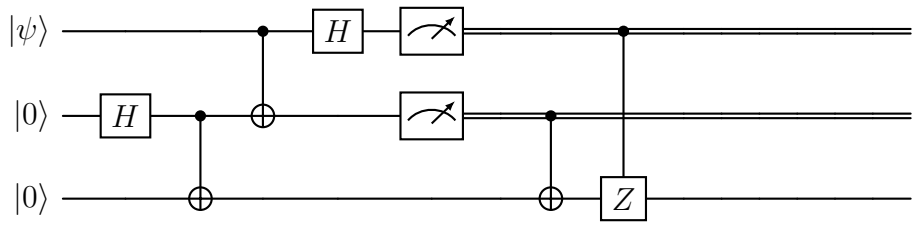

Figure 1: The teleportation gate. Note that at the beginning, the circuit initializes the $|\mathbf{EPR}\rangle$ state. This is not standard, and usually, teleportation diagrams start by acting on an **EPR**.

For understanding the protocol, let's first examine a much simpler circuit. Consider the state given by the partial circuit obtained by first omitting the phase correction, and second by pulling the XORing of $|\psi\rangle$ and Alice's part into a third ancilla. The circuit is presented in Figure 2.

Let's assume the measurement result is zero. Hence, we know that in the computational basis, Alice's **EPR** part agrees with $|\psi\rangle$, or in other words, if we measure the entire system in the computational basis, then the $|\psi\rangle$ bit would be equal to Alice's bit, and therefore would also be equal to Bob's bit. From here, we understand that the probability of Bob's bit being measured as zero equals the probability of $|\psi\rangle$ being measured as zero.

What happens if the result measurement is 1? In that case, we find that $|\psi\rangle$ disagrees with Alice's part and hence also disagrees with Bob's part. We can fix it by applying a Pauli correction, and that is the reason for the controlled XOR at the end of the teleportation.

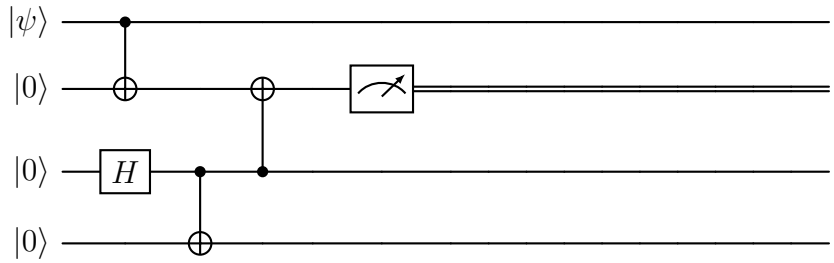

Figure 2: Simple circuit to undersatnad teleportation concept.

In the teleportation protocol, the state progresses up to measurements as:

$$\begin{aligned}
&|\psi\rangle \otimes |\mathbf{EPR}\rangle \left(\alpha |0\rangle + \beta |1\rangle\right) \\
&\quad \rightarrow \frac{1}{\sqrt{2}}\alpha |000\rangle + \frac{1}{\sqrt{2}}\beta |110\rangle \frac{1}{\sqrt{2}} + \alpha |011\rangle + \frac{1}{\sqrt{2}}\beta |101\rangle \\
&\quad \rightarrow \frac{1}{2}\alpha |000\rangle + \frac{1}{2}\alpha |100\rangle + \frac{1}{2}\beta |010\rangle - \frac{1}{2}\beta |110\rangle \\
&\qquad + \frac{1}{2}\alpha |011\rangle + \frac{1}{2}\alpha |111\rangle + \frac{1}{2}\beta |001\rangle - \frac{1}{2}\beta |101\rangle \\
&\quad = \frac{1}{2} |00\rangle |\psi\rangle + \frac{1}{2} |01\rangle X |\psi\rangle + \frac{1}{2} |10\rangle Z |\psi\rangle + \frac{1}{2} |11\rangle XZ |\psi\rangle
\end{aligned}$$

Where the first arrow stands for the Xoring operation, and the second stands for the Hadamard.

6 Gate Teleportation.

Gate teleportation is a method to 'encode' operations by quantum states. At a high level, given a precomputed state, it allows one to apply an operation (gate) using (probably) simpler gates. The precomputed states are called **Magic States**.

6.1 Leading Example: T-Teleportation.

Recall that the Clifford[3] + T is a universal quantum gate set. The Clifford group alone is considered, from the computer science point of view, a simple/weak computational class since it can be classically simulated[4]. Yet, we will see that given access to the magic $|T\rangle = T|+\rangle$, one can simulate the T gate using only Clifford gates and measurements.

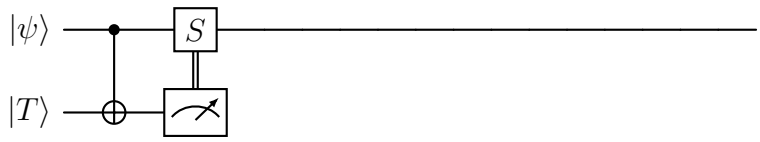

Figure 3: Simulating the T-gate, using the $|T\rangle$ magic state.

In Figure 3, we have the T-teleportation gadget. The state progresses up to the

[3]Generated by H, S and CX

[4]And conjectured to be strictly weaker than **P**

measurement as follows:

$$\left(\sum_x \alpha_x |x\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle\right) \overset{\mathbf{CX}}{\mapsto} \sum_{x,y} \frac{1}{\sqrt{2}} \alpha_x |x\rangle |x \oplus y\rangle e^{i\frac{\pi}{4}y}$$

$$\mapsto \begin{cases} \sum_x \alpha_x |x\rangle e^{i\frac{\pi}{4}x} = T|\psi\rangle & \text{measured 0} \\ \sum_x \alpha_x |x\rangle e^{i\frac{\pi}{4}\bar{x}} & \text{measured 1} \end{cases}$$

$$\overset{\mathbf{CS}}{\mapsto} \begin{cases} T|\psi\rangle \\ \sum_x \alpha_x |x\rangle e^{i\left(\frac{\pi}{4}\bar{x}+\frac{\pi}{2}x\right)} = \sum_x \alpha_x |x\rangle e^{i\frac{\pi}{4}} e^{i\left(\frac{\pi}{4}\bar{x}+\frac{\pi}{4}x\right)} \end{cases}$$

$$= \begin{cases} T|\psi\rangle \\ e^{i\frac{\pi}{4}} \sum_x \alpha_x |x\rangle e^{i\frac{\pi}{4}} = e^{i\frac{\pi}{4}} T|\psi\rangle \end{cases}$$

7 Uhlmann's Theorem

Fidelity of pure quantum states is defined as the absolute square of the inner product between two quantum states. It measures the similarity or overlap between two quantum states, indicating how close they are to being identical. We define the fidelity of a pure state $|\psi\rangle$ with a mixed one ρ as the expectation of the fidelity of the pure state drawn from the distribution induced by ρ. In this section we are about to define the fidelity of two mixed states σ and ρ.

Definition 7.1. *The fidelity of ρ and σ is the maximal fidelity between chosen purifications of them.*

Theorem 7.1 (Uhlmann's Theorem)**.** *Let ρ, σ be mixed states, and let's denote by $|\psi_\rho\rangle$, $|\psi_\sigma\rangle$ arbitrary purifications of ρ, σ. Then their fidelity equals:*

$$F(\sigma,\rho) = \max_{|\psi_\rho\rangle,|\psi_\sigma\rangle} |\langle\psi_\rho|\psi_\sigma\rangle|^2 = \left|\mathbf{Tr}\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}\right|^2$$

For proving it, we will need the following definition and claims:

Definition 7.2. *Denote by $|\Omega\rangle$ the sate: $|\Omega\rangle = \sum_i |i,i\rangle$*

Claim 7.1. *For any matrices A, B (with matched dimensions), it holds that:*

$$\langle\Omega|A \otimes B|\Omega\rangle = \mathbf{Tr}AB^\dagger$$

Proof.

$$\langle\Omega|A \otimes B|\Omega\rangle = \sum_{ij} \langle i,i|AB|j,j\rangle = \sum_{ij} \langle i|A|j\rangle \langle i|B|j\rangle = \sum_{ij} \langle i|A|j\rangle \langle j|B^\dagger|i\rangle$$

$$= \sum_i \langle i|AB^\dagger|i\rangle = \mathbf{Tr}AB^\dagger$$

□

Claim 7.2. *For any square matrix A:*

$$\max_{U\in\mathcal{U}} \mathbf{Tr}AU = \mathbf{Tr}\sqrt{A^\dagger A}$$

The proof is left as an exercise.

7.1 Uhlmann's Theorem Proof.

Let: $\{|\psi_i\rangle\}$ and $\{|\psi'_i\rangle\}$ be the eigenstates of ρ and σ. Any purifications can be written as follows:

$$|\psi_\rho\rangle = \sum_i \left(\rho^{\frac{1}{2}} |\psi_i\rangle\right) |i\rangle = \sum_i \left(\rho^{\frac{1}{2}} U_\rho |i\rangle\right) |i\rangle = \left(\rho^{\frac{1}{2}} U_\rho\right) \otimes I |\Omega\rangle$$
$$|\psi_\sigma\rangle = \sum_i \left(\sigma^{\frac{1}{2}} |\psi'_i\rangle\right) |i'\rangle = \sum_i \left(\sigma^{\frac{1}{2}} U_\sigma |i\rangle\right) V |i\rangle = \left(\sigma^{\frac{1}{2}} U_\sigma\right) \otimes V |\Omega\rangle$$

Where U_ρ, U_σ and V are unitireis. Now using Claim 7.1 and Claim 7.2 we get:

$$\begin{aligned}
\max |\langle\psi_\rho|\psi_\sigma\rangle|^2 &= \max |\langle\Omega| \left(U_\rho^\dagger \rho^{\frac{1}{2}}\right) \otimes I \left(\sigma^{\frac{1}{2}} U_\sigma\right) \otimes V |\Omega\rangle|^2 \\
&= \max |\mathbf{Tr}\left[\left(U_\rho^\dagger \rho^{\frac{1}{2}} \sigma^{\frac{1}{2}} U_\sigma\right) V^\dagger\right]|^2 \\
&= \max |\mathbf{Tr}\left[\rho^{\frac{1}{2}} \sigma^{\frac{1}{2}} V^\dagger\right]|^2 \\
&\leq \left|\mathbf{Tr}\sqrt{\rho^{\frac{1}{2}} \sigma^{\frac{1}{2}} \sigma^{\frac{1}{2}} \rho^{\frac{1}{2}}}\right|^2 = \left|\mathbf{Tr}\sqrt{\rho^{\frac{1}{2}} \sigma \rho^{\frac{1}{2}}}\right|^2
\end{aligned}$$

7.2 Monotonicity of Fidelity.

Let $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then the fidelity is non-decreasing with respect to the partial trace:

$$F(\rho_{AB}, \sigma_{AB}) \leq F(\rho_A, \sigma_A),$$

where $\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}$ and $\sigma_A = \mathrm{Tr}_B\{\sigma_{AB}\}$.

Proof. Notice that any fixed purifications $|\psi\rangle_{RAB}$ and $|\phi\rangle_{RAB}$ of ρ_{AB} and σ_{AB}, respectively, also purify ρ_A and σ_A. Namely, the purifications over the AB-subsystem are a subset of the purifications over the A-subsystem. Therefore:

$$F(\rho_{AB}, \sigma_{AB}) = \max_{|\psi_{\rho_{AB}}\rangle, |\psi_{\sigma_{AB}}\rangle} |\langle\psi_{\rho_{AB}}|\psi_{\sigma_{AB}}\rangle|^2 \leq \max_{|\psi_{\rho_A}\rangle, |\psi_{\sigma_A}\rangle} |\langle\psi_{\rho_A}|\psi_{\sigma_A}\rangle|^2 = F(\rho_A, \sigma_A)$$

□

8 $|$EPR$\rangle$ Distillation / Entanglement Purification.

Distillation of entanglement is a process where two parties distill from a shared mixed entangled state, i.e., noisy EPR, a higher quality entanglement state. Formally, they take n copies of a mixed state ρ at fidelity p with the $|\mathbf{EPR}\rangle$ and output (with probability) k copies of the mixed state ρ' at fidelity $p' > p$ with $|\mathbf{EPR}\rangle$.

It's interesting to look for protocols that use only local operations and classical communication (LOCC), since they suit the realistic case in which remote entities try

to generate their shared entanglement state from a noisy source. Here we present one of the primary protocols [Ben+96][5].

It goes as follows: We consider two copies of ρ such that Alice holds the first qubit of each of the ρ qubit pairs. Alice xors her first qubit into her second and measures, and Bob does the same. Then Alice sends the result to Bob. If the results are equal, they accept, and the mixed state supported over their first shared pair has a higher fidelity.

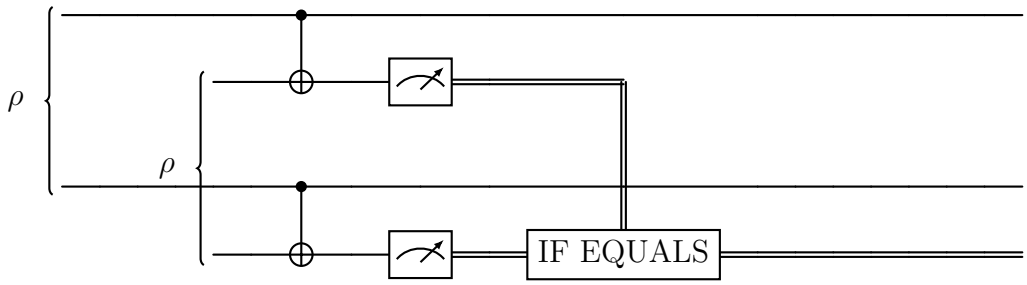

Figure 4: Bennett at el protocol for Entanglement Purification.

We will assume that the noise is uniformly distributed over $(|\beta_{00}\rangle)^{\perp}$ (there is a way to ensure that without decreasing the fidelity). Thus, ρ has the following structure:

$$\rho = p\,|\beta_{00}\rangle\,\langle\beta_{00}| + \frac{1-p}{3}\sum_{j\neq 00}|\beta_j\rangle\,\langle\beta_j|$$

Claim 8.1. *The protocol's success probability is greater than $\frac{1}{2}p^2$. In particular, if $p > \frac{1}{2}$, then the success probability is greater than $\frac{1}{8}$.*

Proof. With probability p^2, the state induced over the qubits is $|\beta_{00}\rangle^{\otimes 2}$. Thus, the density matrices over each of the qubits are the fully mixed state $\frac{1}{2}I$, i.e., a fair coin. So each of the xoring results is distributed as $\sim \mathrm{Bin}(\frac{1}{2})$, and therefore the probability for equality is $\frac{1}{2}$, and in total the success probability $\geq \frac{1}{2}p^2$. □

Now, upon a successful measurement, the state that is left is the projection of ρ into the space in which both the bits of Alice and Bob are equal, namely:

$$\frac{1}{4}((|00\rangle \pm |11\rangle) \otimes (|00\rangle \pm |11\rangle)) \to (|0000\rangle + |1111\rangle)$$
$$\frac{1}{4}((|00\rangle \pm |11\rangle) \otimes (|00\rangle \mp |11\rangle)) \to (|0000\rangle - |1111\rangle)$$
$$\frac{1}{4}((|00\rangle \pm |11\rangle) \otimes (|01\rangle \pm |01\rangle)) \to \emptyset$$
$$\frac{1}{4}((|01\rangle \pm |10\rangle) \otimes (|01\rangle \pm |10\rangle)) \to (|0101\rangle + |0101\rangle)$$
$$\frac{1}{4}((|01\rangle \pm |10\rangle) \otimes (|01\rangle \mp |10\rangle)) \to (|0101\rangle - |0101\rangle)$$

[5]The original protocol starts with a twirling stage to ensure that the noise is uniformly distributed over the $|\beta_{j\neq 00}\rangle$ states.

And the normalization factor is obtained by summing up the weight of each non-canceled term.

$$p^2 + 2\frac{1}{3}p(1-p) + \frac{1}{9}(1-p)^2 + 4\frac{1}{9}(1-p)^2$$

Where the red stands for the states wights which, after being projected, are transformed into $|\mathbf{EPR}\rangle$. So in total the fidelity of ρ' becomes:

$$p' \leftarrow \frac{p^2 + \frac{1}{9}(1-p)^2}{p^2 + 2\frac{1}{3}p(1-p) + \frac{1}{9}(1-p)^2 + 4\frac{1}{9}(1-p)^2}$$

References

[Sch95] Benjamin Schumacher. "Quantum coding". In: *Phys. Rev. A* 51 (4 Apr. 1995), pp. 2738–2747. DOI: 10.1103/PhysRevA.51.2738. URL: https://link.aps.org/doi/10.1103/PhysRevA.51.2738.

[Ben+96] Charles H. Bennett et al. "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels". In: *Physical Review Letters* 76.5 (Jan. 1996), pp. 722–725. ISSN: 1079-7114. DOI: 10.1103/physrevlett.76.722. URL: http://dx.doi.org/10.1103/PhysRevLett.76.722.