Quantum Capacity of the Depolarizing Channel

Martin Leslie

Department of Mathematics
University of Arizona

February 21, 2012

Some resources

- Quantum Computation and Quantum Information, Nielsen and Chuang.
- Lecture notes for physics 229: Quantum information and computation, J. Preskill, `http://theory.caltech.edu/~preskill/ph229/`
- Universal optimal cloning of qubits and quantum registers, Buzek and Hillery, `http://arxiv.org/abs/quant-ph/9801009`
- Quantum Error-Correcting Codes Need Not Completely Reveal the Error Syndrome, Shor and Smolin, `http://arxiv.org/abs/quant-ph/9604006`

Qubits

- Let $|0\rangle = [1,0]^T$ and $|1\rangle = [0,1]^T$ be the *computational basis* for $\mathbb{C}^2$.
- The state of a *qubit* is described by a vector in $\mathcal{H}_1 = \{|\psi\rangle = a|0\rangle + b|1\rangle : a, b \in \mathbb{C} \text{ with } |a|^2 + |b|^2 = 1\}$.
- For n qubits, the state is described by an element of $\mathcal{H}_n = \mathcal{H}_1^{\otimes n}$.
- So for $|\psi\rangle \in \mathcal{H}_n$ we can write

$$|\psi\rangle = \sum_{i \in \mathbb{F}_2^n} a_i |i\rangle$$

where $|i\rangle = |i_1\rangle \otimes \ldots \otimes |i_n\rangle$ and $\sum_i |a_i|^2 = 1$.

Qubits

- Let $|0\rangle = [1,0]^T$ and $|1\rangle = [0,1]^T$ be the *computational basis* for $\mathbb{C}^2$.
- The state of a *qubit* is described by a vector in $\mathcal{H}_1 = \{|\psi\rangle = a|0\rangle + b|1\rangle : a, b \in \mathbb{C} \text{ with } |a|^2 + |b|^2 = 1\}$.
- For n qubits, the state is described by an element of $\mathcal{H}_n = \mathcal{H}_1^{\otimes n}$.
- So for $|\psi\rangle \in \mathcal{H}_n$ we can write

$$|\psi\rangle = \sum_{i \in \mathbb{F}_2^n} a_i |i\rangle$$

  where $|i\rangle = |i_1\rangle \otimes \ldots \otimes |i_n\rangle$ and $\sum_i |a_i|^2 = 1$.
- Actually, we're really in projective Hilbert space

$$\mathbb{P}(\mathcal{H}_n) = \mathcal{H}_n / \sim$$

  where $|\psi\rangle \sim \lambda|\psi\rangle$ (we don't distinguish between states differing only by a *global phase factor*).

Linear algebra in quantum mechanics notation

- Define $A^\dagger = (A^*)^T$ to be the conjugate transpose of A. Matrix A is *Hermitian* if $A^\dagger = A$ and *unitary* if $A^\dagger A = I$.
- If we take (column) vectors $|\psi\rangle, |\varphi\rangle \in \mathcal{H}_n$ then $|\psi\rangle^\dagger$ is a row vector and $|\psi\rangle^\dagger|\varphi\rangle$ is the complex dot product of $|\psi\rangle$ and $|\varphi\rangle$.
- For this reason define $\langle\psi| = |\psi\rangle^\dagger$.
- Then the inner product is $\langle\psi|\varphi\rangle = \langle\psi||\varphi\rangle$.

Quantum mechanics on qubits

- The evolution of the state $|\psi\rangle$ of a closed quantum system from one time to another is given by a unitary matrix U via

$$|\psi\rangle \mapsto U|\psi\rangle.$$

- A measurement in the computational basis will always give a result $|i\rangle$ for some $i \in \mathbb{F}_2^n$. If

$$|\psi\rangle = \sum_{i \in \mathbb{F}_2^n} a_i |i\rangle$$

then the probability of result i is $p(i) = |a_i^2|$ and if i is measured then the state of the qubit after the measurement is $|\psi'\rangle = |i\rangle$.

No-cloning theorem

No perfect qubit copying machine can be built to copy arbitrary qubits

Proof.

Assume unitary U is our copying machine and that we can copy qubits $|\psi\rangle$ and $|\varphi\rangle$. Thus

$$U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle \text{ and } U|\varphi\rangle|0\rangle = |\varphi\rangle|\varphi\rangle.$$

Take the inner product of the two equalities to get

$$\langle\psi|\langle 0|U^\dagger U|\varphi\rangle|0\rangle = \langle\psi|\langle\psi||\varphi\rangle|\varphi\rangle.$$

This simplifies to $\langle\psi|\varphi\rangle = \langle\psi|\varphi\rangle^2$. Thus the two qubits must be equal or orthogonal. □

Some quantum gates

- The following Hermitian, unitary operators act on one qubit:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

- We think of X as a bit flip

$$X(a|0\rangle + b|1\rangle) = a|1\rangle + b|0\rangle$$

Z as a phase flip

$$Z(a|0\rangle + b|1\rangle) = a|0\rangle - b|1\rangle$$

and $Y = iXZ$ as both. The Hadamard gate H takes $|0\rangle$ and $|1\rangle$ to the *dual basis*

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \text{ and } |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

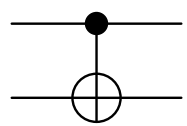
A 2-qubit gate

- The CNOT gate (controlled not) acts by

$$|00\rangle \mapsto |00\rangle$$
$$|01\rangle \mapsto |01\rangle$$
$$|10\rangle \mapsto |11\rangle$$
$$|11\rangle \mapsto |10\rangle$$

  extended linearly.

- We represent it by the following circuit where the top wire is the control and the bottom wire is the target



Density Matrices

- To describe ensembles of states $\{p_i, |\psi_i\rangle\}$ define the *density matrix*

$$\rho = \sum p_i |\psi_i\rangle\langle\psi_i|.$$

- Then unitary evolution U occurs by

$$\rho \mapsto U\rho U^\dagger.$$

- For measurements in the computational basis, outcome i occurs with probability

$$p(i) = \langle i|\rho|i\rangle = \rho_{ii}$$

and if outcome i occurs then the new density matrix is

$$\rho' = |i\rangle\langle i|.$$

Pure states vs mixed states - an example

- Consider the pure state $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, with density matrix

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

- Alternatively, a mixed state prepared with equal probability of being $|0\rangle$ or $|1\rangle$ has density matrix

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

- If we measure these systems both give outcomes 0 or 1 each with probability half.
- However if instead we apply a Hadamard gate and then measure the two systems, the first system will always be measured as a 0, the second system still gives 0 and 1 with probability half.

What do we know quantum computers can do?

- Factor large numbers in time polynomial in the number of digits (Shor).
- Search an unsorted database in time square root of the number of entries (Grover).
- Simulate quantum systems.

What's the problem?

- Quantum systems have noise: unwanted unitary evolution or unwanted measurements (decoherence).
- Circuits built out of gates are one possible way to build a quantum computer. The fault tolerance theorem promises that if we get the probability of decoherence of each gate down to $\sim 10^{-4}$ then we can use error correction to build a reliable quantum computer.
- We will not discuss this, instead using a model more similar to classical communication.

The depolarizing channel

- This channel acts on density matrices ρ by not altering the state with probability $1-q$ and replacing the state with the maximally mixed state $I/2$ with probability q

$$\rho \mapsto (1-q)\rho + q\frac{I}{2}.$$

- For any density matrix ρ

$$\frac{I}{2} = \frac{I\rho I + X\rho X + Y\rho Y + Z\rho Z}{4}$$

so the depolarizing channel can also be thought of as

$$\rho \mapsto (1-p)\,\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z)$$

where $p = 3q/4$ is the probability that input state $|\psi\rangle$ is replaced by $X|\psi\rangle$, $Y|\psi\rangle$ or $Z|\psi\rangle$.

Stabilizer Codes

- Let the *Pauli group* be

$$G_n = \left\{ c \bigotimes_{i=1}^{n} A_i \mid c \in \{\pm 1, \pm i\}, A_i \in \{I, X, Y, Z\} \right\}.$$

- Notice that elements of this group commute iff they have an even number of places with different non-identity matrices. If they do not commute then they anti-commute.
- For a *stabilizer group* $S \leq G_n$ define the *stabilizer code* to be the subspace of $\mathcal{H}_n$ given by

$$V_S = \{|\psi\rangle \colon s|\psi\rangle = |\psi\rangle \text{ for all } s \in S\}$$

- For $V_S \neq 0$ we need all elements of S to commute and $-I \notin S$.

The 5 qubit code

- Let $S = \langle XZZXI, IXZZX, XIXZZ, ZXIXZ \rangle$. We can write this as a matrix

$$A = \begin{bmatrix} X & Z & Z & X & I \\ I & X & Z & Z & X \\ X & I & X & Z & Z \\ Z & X & I & X & Z \end{bmatrix}$$

- Now consider the set of errors $\{I, X_i, Y_i, Z_i : i = 1, \ldots, 5\}$. Define the *syndrome* of an error to be a vector of ± 1's: $+1$ if it commutes with a stabilizer generator and -1 if it anti-commutes.
- For example X_2 has syndrome $(-1, +1, +1, +1)$.
- If you check through you will see that each of the errors we are considering has a different syndrome.

Decoding the 5 qubit code

- Begin with a state $|\psi\rangle \in V_S$.
- An error $E \in G_n$ occurs, changing the state to $E|\psi\rangle$.
- Now for each stabilizer generator g_l, find $g_l E|\psi\rangle$. We get

$$g_l E|\psi\rangle = \pm E g_l|\psi\rangle = \pm E|\psi\rangle$$

  depending on the syndrome of E.
- Looking at the syndrome, find the unique error E that generates that syndrome.
- Apply $E^\dagger$ to fix it: $E^\dagger E|\psi\rangle = |\psi\rangle$.

Decoding stabilizer codes in general

- We can decode a collection of errors $\{E_j\}$ if all the errors in the collection with the same syndrome differ by an element of the stabilizer.
- To see this: if E_j and E_k have the same syndrome but $E_k^\dagger E_j = s \in S$ then if E_j occurs we can still fix the error by doing $E_k^\dagger$:

$$E_k^\dagger E_j |\psi\rangle = s|\psi\rangle = |\psi\rangle.$$

- Codes which use this property (i.e. the set of decodable errors has some overlapping syndromes) are called *degenerate*.

Quantum capacity

- Define the *quantum capacity* of a channel to be the largest number Q such that for any $R < Q$ and $\epsilon > 0$, there is a quantum code with rate at least R such that for any $|\psi\rangle$ in the code the state ρ recovered after passing through the channel has fidelity $F^2 = \langle\psi|\rho|\psi\rangle > 1 - \epsilon$.
- In what follows, let $Q(p)$ be the quantum capacity of the depolarizing channel with parameter p.

A no-cloning upper bound

- Assume an eavesdropper steals your qubit with probability $q \geq 1/2$ and replaces it with the state $I/2$.
- If you could reliably recover your qubit under this scenario then using the same protocol so could the eavesdropper.
- Thus by the no-cloning theorem reliable quantum communication over the depolarizing channel is impossible if $p = 3q/4 > 3/8$.
- In fact we can improve this result: the optimal approximate cloner (by Bužek and Hillery, 1998) creates two copies of a given qubit equivalent to sending both qubits through a depolarizing channel with $p = 1/4$.
- So $Q(p) = 0$ for $p \geq 1/4$.

Extending this bound linearly

- If we send n bits knowing that $n(1-p_2/p_1)$ specified qubits will arrive safely and the other np_2/p_1 qubits will go through a depolarizing channel with parameter p_1 then the capacity of this channel is

$$Q \leq 1 - \frac{p_2}{p_1} + \frac{p_2}{p_1} Q(p_1).$$

- But this channel is effectively a depolarizing channel with parameter p_2 except we know some of the qubits are invulnerable to error. This information can only improve the capacity so we must have $Q(p_2) \leq Q$.
- Now take $p_1 = 1/4$, $p_2 = p \leq 1/4$. Then $Q(p_1) = 0$ so our lemma above gives us the bound

$$Q(p) \leq 1 - 4p \quad \text{for } p \leq 1/4.$$

A random-coding lower bound

- Create a random stabilizer code by choosing $n-k$ independent, commuting generators for S sequentially.
- Let X be the random process generating the depolarizing channel errors. This has entropy

$$H(X) = H(1-p, p/3, p/3, p/3) = h_2(p) + p\log_2(3).$$

- The typical set of errors is

$$A_\epsilon^{(n)} = \left\{ x \in \{I, X, Y, Z\}^{\otimes n} : \left| -\frac{1}{n}\log p(x_1, \ldots, x_n) - H(X) \right| < \epsilon \right\}$$

  with $P(A_\epsilon^{(n)}) > 1-\epsilon$ for n sufficiently large and $|A_\epsilon^{(n)}| \leq 2^{n(H(X)+\epsilon)}$.
- To decode we measure the syndrome of the generators for S then decode if there is a unique typical error operator with syndrome equal to the syndrome measured.

Probability estimation

- The probability of this decoder failing is

$$\begin{aligned}
&P_{\text{fail}} \\
&\leq P(x \notin A_\epsilon^{(n)}) + P(x \in A_\epsilon^{(n)}, \exists x' \text{ typical, same syndrome as } x) \\
&< \epsilon + 2^{n(H(X)+\epsilon)} P(x' \text{ has same syndrome as } x) \\
&= \epsilon + 2^{n(H(X)+\epsilon)} (1/2)^{n-k} \\
&= \epsilon + 2^{n(H(X)-1+k/n+\epsilon)}.
\end{aligned}$$

- This probability can be made arbitrarily small when $k/n < 1 - H(X)$. This gives us the 'hashing bound' or 'random coding bound'

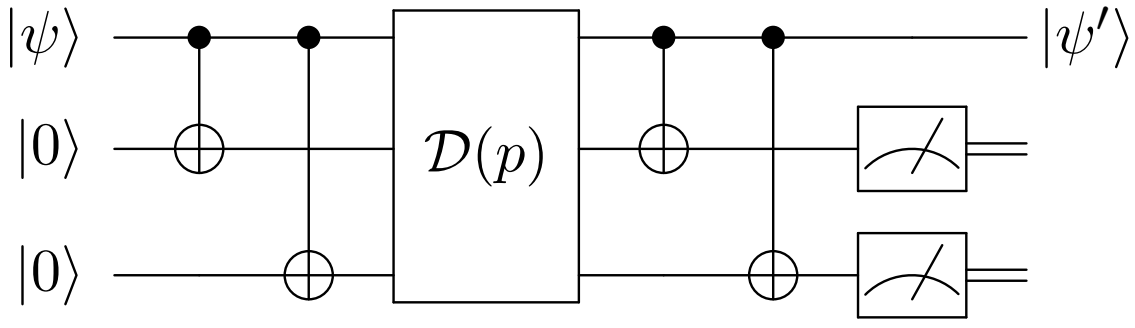
$$Q(p) \geq 1 - h_2(p) - p \log_2(3).$$

- For a more general *Pauli channel* with varied probability of acting like I, X, Y, Z this same argument gives

$$Q(p) \geq 1 - H(p_I, p_X, p_Y, p_Z).$$

Degenerate repetition codes

- Following Shor and Smolin (1996), consider a k–repetition code designed to correct bit-flip errors (a circuit for $k = 3$ is below).



- Consider the 'super channel' which takes $|\psi\rangle$ to $|\psi'\rangle$. For a classical output of the measurements $i \in \mathbb{F}_2^{k-1}$ this is a Pauli channel.
- Define the average entropy of the super channel to be

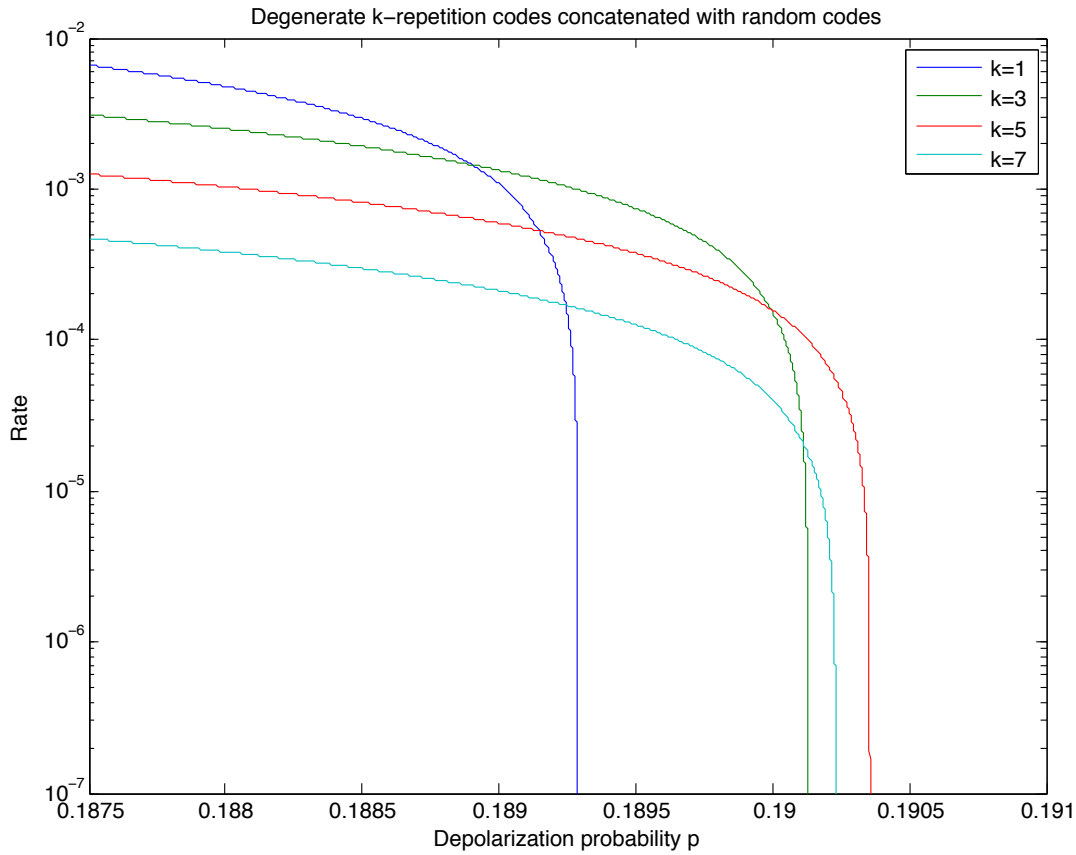
$$S(k) = \sum_{i \in \mathbb{F}_2^{k-1}} P(i) H\left(p_I^{(i)}, p_X^{(i)}, p_Y^{(i)}, p_Z^{(i)}\right).$$

Repetition codes concatenated with random codes

- A random code for the super channel has rate

$$R = \frac{1 - S(k)}{k}.$$

- Evaluating the $S(k)$ recursively gives us the following plot.



Final comparison

- The plots for $k = 2, 4, 6$ are similar but have worse rates. Thus we see that positive rate is achieved for the largest p when $k = 5$.
- This $k = 5$ code concatenated with random coding provides a very small advantage over random coding only when we are at high depolarizing probability.