Introduction to quantum information processing
Fidelity and other distance metrics

Brad Lackey

3 November 2016
OUTLINE

1. Noise channels
2. Fidelity
3. Trace distance
Mixed states on the Bloch sphere:

\[ \rho = \frac{1}{2}(1 + r_x X + r_y Y + r_z Z) \leftrightarrow (r_x, r_y, r_z). \]

We can “purify” a mixed state to have \( \rho = \text{tr}_B(|\psi\rangle\langle\psi|) \).

Quantum channels are generally given in Kraus form:

\[ C(\rho) = \sum_j E_j \rho E_j^\dagger. \]
1. Noise channels
2. Fidelity
3. Trace distance
The **bit flip** channel:
- flips the qubit state, $|0\rangle \leftrightarrow |1\rangle$, with probability $p$, and
- leaves the qubit alone with probability $(1 - p)$.

Therefore $\mathcal{B}_p(\rho) = (1 - p)\rho + pX\rho X$. On the Bloch sphere:

$$(r_x, r_y, r_z) \mapsto (r_x, (1 - 2p)r_y, (1 - 2p)r_z).$$

The **phase flip** channel:
- flips the phase, $|0\rangle \leftrightarrow |0\rangle$ and $|1\rangle \leftrightarrow -|1\rangle$, with probability $p$, and
- leaves the qubit alone with probability $(1 - p)$.

Therefore $\mathcal{P}_p(\rho) = (1 - p)\rho + pZ\rho Z$. On the Bloch sphere:

$$(r_x, r_y, r_z) \mapsto ((1 - 2p)r_x, (1 - 2p)r_y, r_z).$$
**THE DEPOLARIZING CHANNEL**

The *depolarizing* channel is a very population error channel to study:
- with probability $p$ it applies one of the Pauli operators $\{X, Y, Z\}$ (each equally likely), and
- with probability $p$ it leaves the qubit alone.

Therefore $\mathcal{D}_p(\rho) = (1 - p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z)$.

However we showed last time that $\rho + X\rho X + Y\rho Y + Z\rho Z = 2 \cdot \mathbb{1}$. So:

$$\mathcal{D}_p(\rho) = \left(1 - \frac{4p}{3}\right)\rho + \frac{p}{3}(\rho + X\rho X + Y\rho Y + Z\rho Z) = \left(1 - \frac{4p}{3}\right)\rho + \frac{2p}{3}\mathbb{1}.$$  

On the Bloch sphere, this is easy to understand:

$$(r_x, r_y, r_z) \mapsto \left(1 - \frac{4p}{3}\right) (r_x, r_y, r_z).$$

So it simply shrinks the Bloch sphere vector, well at least when $p < \frac{3}{4}$. 
The amplitude and phase damping channels

The *amplitude damping* channel models loss of energy to the environment.

- It has two Kraus operators, which in basis $(|0\rangle, |1\rangle)$ are
  
  $$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \; \text{and} \; E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$  

  - And: $(r_x, r_y, r_z) \mapsto (r_x \sqrt{1-\gamma}, r_y \sqrt{1-\gamma}, \gamma + r_z (1-\gamma))$.

The *phase damping* channel models coherence loss (quantum information).

- It also has two Kraus operators, which in basis $(|0\rangle, |1\rangle)$ are
  
  $$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix} \; \text{and} \; E_1 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{pmatrix}.$$  

  - And: $(r_x, r_y, r_z) \mapsto (r_x \sqrt{1-\gamma}, r_y \sqrt{1-\gamma}, r_z)$. 

FIDELITY AND OTHER DISTANCE METRICS
1. Noise channels
2. Fidelity
3. Trace distance
Fidelity between states

The classical fidelity, or Bhattacharyya coefficient, of $p$ and $q$ is

- $F(p, q) = \sum_i \sqrt{p_i q_i}$.
- A high fidelity indicates the distributions are close.
- It is related to the so-called “cosine” metric for distributions.

Then to make a notion of quantum fidelity $F(\rho, \sigma)$ it makes sense to:

- note that for any POVM $\{E_i\}$, we have $p_i = \text{tr}(E_i \rho)$ and $q_i = (E_i \sigma)$,
- choose the POVM that distinguishes these distributions best:

$$F(\rho, \sigma) = \min_{\{E_i\} \text{ POVM}} F(\{\text{tr}(E_i \rho)\}, \{\text{tr}(E_i \sigma)\}).$$

This has a simple formula $F(\rho, \sigma) = \text{tr}\sqrt{\rho^{1/2} \sigma \rho^{1/2}}$ (we won’t prove this).

- This isn’t too bad if one state is pure: $F(\rho, |\psi\rangle\langle\psi|) = \sqrt{\langle\psi|\rho|\psi\rangle}$.
- It’s even better if both are pure: $F(|\phi\rangle\langle\phi|, |\psi\rangle\langle\psi|) = |\langle\psi|\phi\rangle|$. 

Fidelity and other distance metrics
**Uhlmann’s theorem**

**Theorem (Uhlmann (1976))**

Let \( \rho \) and \( \sigma \) be densities on \( \mathcal{H} \), with \( \dim \mathcal{H} = n \). Then

\[
F(\rho, \sigma) = \max_{|\psi\rangle, |\phi\rangle} |\langle \psi | \phi \rangle|
\]

where the maximum is over all purifications \( |\psi\rangle \) of \( \rho \) and \( |\phi\rangle \) of \( \sigma \) on \( \mathcal{H} \otimes \mathbb{C}^n \).

This really can’t be used to compute the fidelity, but we can derive nice facts:

- Since always \( 0 \leq |\langle \psi | \phi \rangle| \leq 1 \), we must have \( 0 \leq F(\rho, \sigma) \leq 1 \).
- If \( F(\rho, \sigma) = 1 \) then \( \rho = \sigma \) (a purification has \( |\psi\rangle \propto |\phi\rangle \)).
- If \( F(\rho, \sigma) = 0 \) then \( \rho \) and \( \sigma \) are distinguishable (from \( \langle \psi | \phi \rangle = 0 \)).
MONOTONICITY OF FIDELITY

A unitary transformation, $U$, does not change the fidelity.

- From the spectral theorem $\rho^{1/2} = \sum_j \sqrt{\lambda_j} |\phi_j\rangle \langle \phi_j|.$
- So $(U \rho U^\dagger)^{1/2} = \sum_j \sqrt{\lambda_j} U |\phi_j\rangle \langle \phi_j| U^\dagger = U (\rho^{1/2}) U^\dagger.$

Therefore,

$$F(U \rho U^\dagger, U \sigma U^\dagger) = \text{tr} \sqrt{U \rho^{1/2} U^\dagger U \sigma U^\dagger U \rho^{1/2} U^\dagger} = \text{tr}(U \sqrt{\rho^{1/2} \sigma \rho^{1/2} U^\dagger}) = F(\rho, \sigma).$$

A quantum channel, $C$, never shrinks the fidelity (but may enlarge it).

- We need some facts about what we can do:
  - We can jointly purify $\rho = \text{tr}_B (|\psi\rangle \langle \psi|)$ and $\sigma = \text{tr}_B (|\phi\rangle \langle \phi|)$.
  - We can add more ancilla to dilate $C(\rho) = \text{tr}_E (U(\rho \otimes |0\rangle \langle 0|) U^\dagger)$.
  - We can combine these $C(\rho) = \text{tr}_{BE} (U (|\psi\rangle \langle \psi| \otimes |0\rangle \langle 0|) U^\dagger)$.
  - Then $U |\psi\rangle \otimes |0\rangle$ is a purification of $C(\rho)$.
  - Same for $U (|\phi\rangle \otimes |0\rangle)$ of $C(\sigma)$.

Therefore, using Uhlman’s theorem

$$F(\rho, \sigma) = |\langle \psi | \phi \rangle| = |(\langle \psi | \otimes \langle 0 |)(UU^\dagger)(|\phi\rangle \otimes |0\rangle)| \leq F(C(\rho), C(\sigma)).$$
**HOW WELL DO CHANNELS PRESERVE INFORMATION?**

E.g. consider the fidelity of the depolarizing channel on a pure state:

$$F(|\psi\rangle\langle\psi|, D_p |\psi\rangle\langle\psi|) = \sqrt{\langle\psi| \left( \frac{2p}{3} \mathbf{1} + \left( 1 - \frac{4p}{3} \right) |\psi\rangle\langle\psi| \right) |\psi\rangle}$$

$$= \sqrt{\frac{2p}{3} + \left( 1 - \frac{4p}{3} \right)} = \sqrt{1 - \frac{2p}{3}}.$$

Here the loss of fidelity is the same for any state.

However for the phase damping channel:

$$F(|\psi\rangle\langle\psi|, P_p |\psi\rangle\langle\psi|) = \sqrt{\langle\psi| \left( E_0 |\psi\rangle\langle\psi| E_0^\dagger + E_1 |\psi\rangle\langle\psi| E_1^\dagger \right) |\psi\rangle}$$

$$= \sqrt{1 - \frac{1}{2} (\sqrt{1 - \lambda} - 1) \sin^2 \theta}$$

where $$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle.$$

- Loss of fidelity depends on the “coherence” $$\frac{1}{2} \sin^2 \theta = \cos \frac{\theta}{2} \sin \frac{\theta}{2}.$$
OUTLINE

1. Noise channels
2. Fidelity
3. Trace distance
Trace distance

Another popular metric is the statistical distance: \( D(p, q) = \frac{1}{2} \sum_i |p_i - q_i| \).

The quantum analogue of this is the *trace distance*: \( D(\rho, \sigma) = \frac{1}{2} \text{tr}(|\rho - \sigma|) \).

Like fidelity, one can “optimize” statistical distance over POVMs:

\[
D(\rho, \sigma) = \max_{\{E_i\} \text{ POVM}} D(\{\text{tr}(E_i \rho)\}, \{\text{tr}(E_i \sigma)\}).
\]

Here’s a sketch of the proof: Using the spectral decomposition

1. \( \rho - \sigma = A_+ - A_- \) where \( A_+, A_- \geq 0 \) and \( A_- A_+ = A_+ A_- \).
2. The \( A_{\pm} \) commute so \( |\rho - \sigma| = |A_+ - A_-| = A_+ + A_- \).
3. Then \( D(\{\text{tr}(E_i \rho)\}, \{\text{tr}(E_i \sigma)\}) = \sum_i |\text{tr}(E_i (\rho - \sigma))| \) and this equals
   \[
   \sum_i |\text{tr}(E_i (A_+ - A_-))| \leq \sum_i \text{tr}(E_i (A_+ + A_-)) \quad \text{(since } |x - y| \leq |x| + |y|) \]
   \[= \sum_i \text{tr}(E_i |\rho - \sigma|) = D(\rho, \sigma).\]
4. To achieve the maximum take the projections onto \( \text{supp}(A_{\pm}) \).
**Contractivity of Channels**

Also like the fidelity, the trace distance is preserved under unitaries:

- Easily $U|\rho - \sigma|U^\dagger = |U(\rho - \sigma)U^\dagger|$.

Quantum channels contract trace distance: $D(\mathcal{C}(\rho), \mathcal{C}(\sigma)) \leq D(\rho, \sigma)$.

- We won’t prove this, but it’s not too difficult.
- Note $\rho \mapsto \text{tr}_B(\rho)$ is a channel, so $D(\text{tr}_B(\rho), \text{tr}_B(\sigma)) \leq D(\rho, \sigma)$.

A useful metric is the “gate error” $E(U, \mathcal{C}) = \max_\rho D(U\rho U^{-1}, \mathcal{C}(\rho))$.

- This bounds errors from approximating to $U$ with the channel $\mathcal{C}$.
- It has the nice identity $E(VU, \mathcal{D} \circ \mathcal{C}) \leq E(U, \mathcal{C}) + E(V, \mathcal{D})$.
- This follows from the triangle inequality and contractivity.
Trace distance versus fidelity

For pure states $|\psi\rangle$ and $|\phi\rangle$ (with angle $\theta$ between them):

- $D(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\sin \theta|$, and
- $F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\cos \theta|$. 
- Therefore $D(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = \sqrt{1 - F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|)^2}$.

Otherwise, by Uhlmann’s theorem $F(\rho, \sigma) = |\langle\psi|\phi\rangle| = F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|)$.

- Then $D(\rho, \sigma) = D(\text{tr}_B(|\psi\rangle\langle\psi|), \text{tr}_B(|\phi\rangle\langle\phi|)) \leq D(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|)$
- Therefore, $D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}$.

On the other hand using a POVM with $F(\rho, \sigma) = F(\{\text{tr}(E_j\rho)\}, \{\text{tr}(E_j\sigma)\})$.

- For distributions: $\sum_i (\sqrt{p_i} - \sqrt{q_i})^2 = 2 - 2\sum_i \sqrt{p_iq_i}$
- Also: $\sum_i (\sqrt{p_i} - \sqrt{q_i})^2 \leq \sum_i |p_i - q_i|$
- Therefore, $1 - F(\rho, \sigma) \leq \frac{1}{2} \sum_i |p_i - q_i| \leq D(\rho, \sigma)$.
Next time...

- Classical and quantum codes.