Lecture 12: Distance Measures for Quantum Information

Static Metrics, Uhlmann's Theorem, and Channel Fidelity

Moody T. Chu

North Carolina State University

Spring 2026

Lecture Overview

1. Classical Distance Foundations: probability distribution similarity.
2. Quantum Trace Distance: geometric interpretation and contractivity.
3. Quantum Fidelity: overlap measures and Uhlmann's Theorem.
4. Bures Geometry: the angle and metric derived from fidelity.
5. Properties of Fidelity: monotonicity and strong concavity.
6. Distinguishability Relations: the Fuchs–van de Graaf inequalities.
7. Dynamic Distance Measures: entanglement fidelity and channel preservation.

Classical Distance Measures: Motivation

- In information theory, sources are modeled by probability distributions $\{p_x\}$ over a common index set. Quantifying the difference between two distributions requires a metric.
  - ◇ **Classical Trace Distance** (L_1 or Kolmogorov distance):

$$D(p_x, q_x) := \frac{1}{2}\sum_x |p_x - q_x|. \tag{1}$$

    - Satisfies all metric axioms (non-negativity, symmetry, triangle inequality).
    - Note: (p_x, q_x) on the left refers to the entire distributions, not individual values.
  - ◇ **Classical Fidelity** (Bhattacharyya coefficient):

$$F(p_x, q_x) := \sum_x \sqrt{p_x q_x}. \tag{2}$$

    - Interpreted as the inner product $\langle \vec{u}, \vec{v} \rangle$ between unit vectors $\vec{u} = (\sqrt{p_x})$ and $\vec{v} = (\sqrt{q_x})$ lying on a unit hypersphere.
    - $F \in [0, 1]$, with $F = 1$ iff $p = q$.

Operational Meaning of Classical Trace Distance

Theorem

The trace distance satisfies:

$$D(p_x, q_x) = \max_S \left| \sum_{x \in S} p_x - \sum_{x \in S} q_x \right|$$

where the maximum is taken over all subsets S of the sample space.

- The trace distance equals the maximum probability difference any single event can have under p versus q — the *optimal distinguishability* under the best possible test.
- **Proof:** Partition the sample space into $V_+ := \{x : p_x \geq q_x\}$ and $V_- := \{x : p_x < q_x\}$.
  - ◇ Since probabilities sum to 1, $\sum_{x \in V_+}(p_x - q_x) = \sum_{x \in V_-}(q_x - p_x)$, so both equal $D(p, q)$.
  - ◇ For any S: $\sum_{x \in S}(p_x - q_x) \leq \sum_{x \in V_+}(p_x - q_x) = D(p, q)$, with equality at $S = V_+$.

Classical Dynamic Distance: Evaluating a Channel

- Consider a random variable X sent through a noisy channel, producing output Y via conditional probability $p(Y|X)$.
  - ◇ The natural dynamic measure of noise is the error probability $p(X \neq Y)$.
  - ◇ How do we evaluate how well a channel preserves information using static tools?
- **Key idea:** Introduce a *reference copy* $\tilde{X} = X$ to convert the dynamic question into a static distance comparison.
  - ◇ **Ideal state:** $(\tilde{X}, X)$ — perfectly correlated (identical).
  - ◇ **Noisy state:** $(\tilde{X}, Y)$ — $\tilde{X}$ is the original, Y is the channel output.
- The trace distance of these two joint distributions directly recovers the error probability (shown on the next slide).
- This reference-system trick is exactly the idea behind *entanglement fidelity* for quantum channels.

Classical Dynamic Distance: via Static Trace Distance

- Compute the static trace distance of the joint distributions:

$$
\begin{aligned}
&D\Big((\tilde{X}, X), (\tilde{X}, Y)\Big) \\
&= \frac{1}{2}\sum_{x,x'} \Big|\delta_{x,x'}\, p(X = x) - p(\tilde{X} = x, Y = x')\Big| \\
&= \frac{1}{2}\sum_{x \neq x'} p(\tilde{X} = x, Y = x') + \frac{1}{2}\sum_{x}\big(p(X = x) - p(\tilde{X} = x, Y = x)\big) \\
&= \frac{1}{2} p(X \neq Y) + \frac{1}{2}\big(1 - p(X = Y)\big) = p(X \neq Y).
\end{aligned}
$$

  - ◇ The trace distance between the ideal and noisy joint distributions *equals* the error probability.
  - ◇ Tracking how well a channel preserves external correlation provides a rigorous dynamic quality measure — a principle we will generalize to quantum channels via entanglement fidelity.

The Quantum Trace Distance

- Generalize the classical L_1 distance to density operators on Hilbert spaces.

Definition: Quantum Trace Distance

The quantum trace distance between two quantum states ρ and σ is:

$$D(\rho, \sigma) := \frac{1}{2} \operatorname{tr} |\rho - \sigma|, \tag{3}$$

where $|A| := \sqrt{A^\dagger A}$ is the unique positive square root (operator absolute value).

- ◇ Since $\rho - \sigma$ is Hermitian ($A = A^\dagger$), its singular values equal the absolute values of its eigenvalues $\{\lambda_i\}$.
- ◇ Thus $D(\rho, \sigma) = \frac{1}{2} \sum_i |\lambda_i|$ where λ_i are the eigenvalues of $\rho - \sigma$. This is also called the *nuclear norm* of $\rho - \sigma$.
- ◇ Well-defined: $\rho - \sigma$ is traceless, so positive and negative eigenvalues balance, and $D(\rho, \sigma) \in [0, 1]$.

Motivating Trace Distance: I. Classical Generalization

- The quantum trace distance reduces to its classical counterpart when the states commute.
  - If $[\rho, \sigma] = 0$, they share a simultaneous eigenbasis $\{|i\rangle\}$:

$$\rho = \sum_i p_i |i\rangle\langle i|, \quad \sigma = \sum_i q_i |i\rangle\langle i|.$$

  - Since $\rho - \sigma = \sum_i (p_i - q_i) |i\rangle\langle i|$ is already diagonal, the operator absolute value acts directly on eigenvalues:

$$|\rho - \sigma| = \sum_i |p_i - q_i| \, |i\rangle\langle i|.$$

  - Taking the trace:

$$D(\rho, \sigma) = \frac{1}{2} \operatorname{tr} |\rho - \sigma| = \frac{1}{2} \sum_i |p_i - q_i| = D_{\text{classical}}(p, q).$$

- **Takeaway:** The quantum definition is a genuine generalization — it coincides with the classical formula in the classical (commuting) limit.

Motivating Trace Distance: II. Bloch Sphere Geometry

- For single-qubit states, the trace distance has a clean geometric meaning.
  - ◇ Let ρ and σ be parametrized by Bloch vectors $\vec{r}$ and $\vec{s}$:

$$\rho = \frac{I + \vec{r} \cdot \vec{\sigma}}{2}, \quad \sigma = \frac{I + \vec{s} \cdot \vec{\sigma}}{2},$$

    so $\rho - \sigma = \frac{1}{2}(\vec{r} - \vec{s}) \cdot \vec{\sigma}$.
  - ◇ The matrix $(\vec{r} - \vec{s}) \cdot \vec{\sigma}$ is traceless Hermitian. Its square is:

$$\left[(\vec{r} - \vec{s}) \cdot \vec{\sigma}\right]^2 = |\vec{r} - \vec{s}|^2 \, I,$$

    so its eigenvalues are $\pm|\vec{r} - \vec{s}|$, giving $\rho - \sigma$ eigenvalues $\pm\frac{1}{2}|\vec{r} - \vec{s}|$.
  - ◇ Therefore:

$$D(\rho, \sigma) = \frac{1}{2} \operatorname{tr} |\rho - \sigma| = \frac{1}{2} \cdot 2 \cdot \frac{|\vec{r} - \vec{s}|}{2} = \frac{1}{2}|\vec{r} - \vec{s}|.$$

- The quantum trace distance is exactly half the Euclidean distance between Bloch vectors.

Operator Meaning of Quantum Trace Distance

Theorem

The quantum trace distance satisfies:

$$D(\rho, \sigma) = \max_P \operatorname{tr}\big(P(\rho - \sigma)\big),$$

where the maximum is taken over all orthogonal projectors $0 \le P \le I$.

- **Proof:** Write $\rho - \sigma = Q - S$ via the spectral decomposition, where $Q, S \ge 0$ with $QS = 0$ (orthogonal supports).
  - Then $|\rho - \sigma| = Q + S$, and since $\operatorname{tr}(Q) = \operatorname{tr}(S)$ (both sides of the traceless decomposition are equal), $D(\rho, \sigma) = \operatorname{tr}(Q)$.
  - Let P be the projector onto the range of Q. Then $\operatorname{tr}(P(\rho - \sigma)) = \operatorname{tr}(PQ) - \operatorname{tr}(PS) = \operatorname{tr}(Q) = D(\rho, \sigma)$. ✓
  - For any other projector P': since $P' \le I$ and $S \ge 0$,

$$\operatorname{tr}(P'(\rho - \sigma)) = \operatorname{tr}(P'Q) - \operatorname{tr}(P'S) \le \operatorname{tr}(P'Q) \le \operatorname{tr}(IQ) = D(\rho, \sigma).$$

Operator Meaning of Trace Distance: via POVM

Theorem

Let $\{E_m\}$ be a POVM with $p_m = \mathrm{tr}(\rho E_m)$, $q_m = \mathrm{tr}(\sigma E_m)$. Then $D(\rho,\sigma) = \max_{\{E_m\}} D(p_m, q_m)$.

- The quantum trace distance equals the *maximum classical trace distance* over all measurements — the ultimate statistical distinguishability of ρ versus σ.
- **Proof:** Using $\rho - \sigma = Q - S$ (positive-negative spectral decomposition):

$$D(p_m, q_m) = \frac{1}{2}\sum_m |\,\mathrm{tr}(E_m(\rho-\sigma))| \le \frac{1}{2}\sum_m \mathrm{tr}(E_m(Q+S))$$

$$= \frac{1}{2}\mathrm{tr}\Big[\Big(\sum_m E_m\Big)|\rho-\sigma|\Big] = D(\rho,\sigma).$$

Equality is achieved by $\{P_Q,\, P_S,\, I - P_Q - P_S\}$, where P_Q, P_S project onto the supports of Q and S respectively.

Metric Properties & Unitary Invariance

- The quantum trace distance satisfies all metric axioms:
  1. **Positive definiteness:** $D(\rho, \sigma) = 0 \iff \rho = \sigma$.
  2. **Symmetry:** $D(\rho, \sigma) = \frac{1}{2} \operatorname{tr} |\rho - \sigma| = \frac{1}{2} \operatorname{tr} |\sigma - \rho| = D(\sigma, \rho)$.
  3. **Triangle inequality:** Pick P such that $D(\rho, \tau) = \operatorname{tr}(P(\rho - \tau))$. Then:

$$D(\rho, \tau) = \operatorname{tr}(P(\rho - \sigma)) + \operatorname{tr}(P(\sigma - \tau)) \leq D(\rho, \sigma) + D(\sigma, \tau).$$

- The quantum trace distance is unitarily invariant: for any unitary U,

$$D(U\rho U^\dagger, U\sigma U^\dagger) = \frac{1}{2} \operatorname{tr} |U(\rho - \sigma)U^\dagger| = D(\rho, \sigma),$$

using $|UAU^\dagger| = U|A|U^\dagger$ for Hermitian A, and cyclicity of trace.

Contractivity under Quantum Operations

Theorem

Let $\mathcal{E}$ be a trace-preserving quantum operation. For all states ρ, σ:

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D(\rho, \sigma).$$

- **Physical meaning:** No physical process can make two quantum states more distinguishable. Noise only destroys information.
- **Proof:** Write $\rho - \sigma = Q - S$ with $Q, S \geq 0$, $\mathrm{tr}(Q) = \mathrm{tr}(S) = D(\rho, \sigma)$. Let P achieve $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) = \mathrm{tr}(P(\mathcal{E}(\rho) - \mathcal{E}(\sigma)))$. Then:

$$\begin{aligned} D(\rho, \sigma) &= \mathrm{tr}(Q) = \mathrm{tr}(\mathcal{E}(Q)) && \text{(trace preservation)} \\ &\geq \mathrm{tr}(P\,\mathcal{E}(Q)) \geq \mathrm{tr}(P(\mathcal{E}(Q) - \mathcal{E}(S))) && (P \leq I,\ \mathcal{E}(Q) \geq 0) \\ &= \mathrm{tr}(P(\mathcal{E}(\rho) - \mathcal{E}(\sigma))) = D(\mathcal{E}(\rho), \mathcal{E}(\sigma)). && \blacksquare \end{aligned}$$

Contractivity: Partial Trace Implications

- The partial trace $\mathcal{E}(\rho) := \text{tr}_B(\rho)$ is itself a trace-preserving quantum operation. Contractivity therefore gives:

$$D(\text{tr}_B(\rho), \text{tr}_B(\sigma)) \leq D(\rho, \sigma).$$

  - ◇ Discarding a subsystem *reduces* our ability to distinguish two global states.
  - ◇ **Intuition:** Suppose ρ and σ differ only in a correlation between subsystems A and B. Once B is discarded, the reduced states on A may be identical, making the two distributions look the same even though the global states were different. Destroying information makes things harder to tell apart.
  - ◇ Conversely, to *maximize* distinguishability, one should use *all* available degrees of freedom — never discard a subsystem unnecessarily.

Strong Convexity of Trace Distance

Theorem (Strong Convexity)

Let $\{p_i\}$ and $\{q_i\}$ be probability distributions over the same index set. Then:

$$D\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \leq D(p_i, q_i) + \sum_i p_i\, D(\rho_i, \sigma_i).$$

- The overall distinguishability is bounded by two additive terms:
  - *Label distinguishability:* $D(p_i, q_i)$ — how different are the mixing weights?
  - *State distinguishability:* $\sum_i p_i D(\rho_i, \sigma_i)$ — how different are the individual states on average?
- In particular, when $p_i = q_i$ (same mixing weights), D reduces to the average state distance — matching the ordinary convexity inequality.
- Proof on the next slide.

Strong Convexity: Proof

- **Proof:** Let P be the projector maximizing $D(\sum_i p_i\rho_i, \sum_i q_i\sigma_i) = \mathrm{tr}[P(\sum_i p_i\rho_i - \sum_i q_i\sigma_i)]$.
- Add and subtract $\sum_i p_i\sigma_i$ to split:

$$\mathrm{tr}\Big[P\Big(\sum_i p_i\rho_i - \sum_i q_i\sigma_i\Big)\Big]$$
$$= \underbrace{\sum_i (p_i - q_i)\,\mathrm{tr}(P\sigma_i)}_{\le D(p_i, q_i)} + \underbrace{\sum_i p_i\,\mathrm{tr}(P(\rho_i - \sigma_i))}_{\le \sum_i p_i D(\rho_i, \sigma_i)}.$$

- Each bound follows from the projector maximization characterization of D: $\mathrm{tr}(PA) \le D(\cdot,\cdot)$ for any normalized A and projector P. ■

Unnormalized States and Post-Selected Measurements

- When a projector P filters a state ρ, the operation $\mathcal{E}(\rho) = P\rho P^\dagger$ is *not* trace-preserving:

$$\mathrm{tr}(\mathcal{E}(\rho)) = \mathrm{tr}(P^\dagger P\,\rho) = \mathrm{tr}(P\rho) = p(\text{success}).$$

  The trace of the unnormalized output *equals* the success probability.
- The normalized post-selected state is:

$$\rho_{\text{post}} = \frac{\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))} = \frac{P\rho P^\dagger}{\mathrm{tr}(P\rho)}.$$

- **Post-selected distinguishability bound:** If $p = \mathrm{tr}(\mathcal{E}(\rho))$ and $q = \mathrm{tr}(\mathcal{E}(\sigma))$, then:

$$D\left(\frac{\mathcal{E}(\rho)}{p}, \frac{\mathcal{E}(\sigma)}{q}\right) \leq \frac{D(\rho,\sigma)}{\max(p,q)}.$$

  - A selective filter can *amplify* apparent distinguishability by at most a factor of $1/\max(p,q)$ — the rarer the event, the more amplification is possible, but it is always finite.

Quantum Fidelity: Definition

- Fidelity measures closeness from an *overlap* perspective rather than a distance metric perspective.

Definition: Quantum Fidelity

The fidelity between density operators ρ and σ is:

$$F(\rho,\sigma) \equiv \mathrm{tr}\sqrt{\rho^{1/2}\,\sigma\,\rho^{1/2}}.$$

- ◇ $F(\rho,\sigma) \in [0,1]$, with $F = 1$ iff $\rho = \sigma$.
- ◇ Despite the asymmetric appearance, F is symmetric: $F(\rho,\sigma) = F(\sigma,\rho)$.
  - ◦ **Proof of symmetry:** Let $A = \sqrt{\rho}\sqrt{\sigma}$. Then:

$$A^\dagger A = \sqrt{\sigma}\,\rho\,\sqrt{\sigma} \implies \mathrm{tr}\,|A| = \mathrm{tr}\sqrt{A^\dagger A} = F(\sigma,\rho),$$
$$AA^\dagger = \sqrt{\rho}\,\sigma\,\sqrt{\rho} \implies \mathrm{tr}\,|A^\dagger| = \mathrm{tr}\sqrt{AA^\dagger} = F(\rho,\sigma).$$

Since A and $A^\dagger$ share the same singular values, $\mathrm{tr}\,|A| = \mathrm{tr}\,|A^\dagger|$, so $F(\sigma,\rho) = F(\rho,\sigma)$.

Fidelity Special Case: Pure State Overlap

- If $\rho = |\psi\rangle\langle\psi|$ is pure, then $\rho^{1/2} = |\psi\rangle\langle\psi|$ as well. Substituting:

$$\rho^{1/2}\,\sigma\,\rho^{1/2} = |\psi\rangle\underbrace{\langle\psi|\,\sigma\,|\psi\rangle}_{\text{scalar}}\langle\psi| = \langle\psi|\,\sigma\,|\psi\rangle \cdot |\psi\rangle\langle\psi|\,.$$

Taking the square root and then the trace:

$$F(|\psi\rangle, \sigma) = \sqrt{\langle\psi|\,\sigma\,|\psi\rangle}.$$

- **Two sub-cases:**
  - ◇ If $\sigma = |\phi\rangle\langle\phi|$ is also pure, then $\langle\psi|\,\sigma\,|\psi\rangle = |\langle\psi|\phi|\psi|\phi\rangle|^2$, so:

$$F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi|\psi|\phi\rangle|.$$

    Fidelity equals the absolute inner product between the two state vectors.
  - ◇ If $\sigma = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ is mixed:

$$F(|\psi\rangle, \sigma) = \sqrt{\sum_i p_i |\langle\psi|\phi_i|\psi|\phi_i\rangle|^2}.$$

Fidelity Special Case: Commuting States

- If $[\rho, \sigma] = 0$, they share an eigenbasis $\{|i\rangle\}$ with eigenvalues r_i and s_i:

$$\rho^{1/2}\,\sigma\,\rho^{1/2} = \sum_i r_i s_i \, |i\rangle\langle i| \implies \sqrt{\rho^{1/2}\sigma\rho^{1/2}} = \sum_i \sqrt{r_i s_i}\, |i\rangle\langle i| .$$

- Taking the trace:

$$F(\rho, \sigma) = \sum_i \sqrt{r_i s_i} = F_{\text{classical}}(r, s).$$

  - ◇ Fidelity reduces to the classical Bhattacharyya coefficient between the eigenvalue distributions.
- **Summary of special cases** (in order of generality):
  - ◇ Two pure states: $F = |\langle\psi|\phi|\psi|\phi\rangle|$.
  - ◇ Pure vs. mixed: $F = \sqrt{\langle\psi|\,\sigma\,|\psi\rangle}$.
  - ◇ Two commuting mixed: $F = \sum_i \sqrt{r_i s_i}$.
  - ◇ Two general mixed: $F = \operatorname{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$.

Uhlmann's Theorem

Uhlmann's theorem gives an operational definition of mixed-state fidelity as a geometric optimization over purifications.

Theorem (Uhlmann's Theorem)

Let ρ and σ be density operators on system Q. Fix a purification $|\psi\rangle$ of ρ into a joint system RQ. Then:

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

where the maximum runs over all purifications $|\phi\rangle$ of σ into RQ.

- **Significance:** Mixed-state fidelity equals the maximum inner product over purifications. This lifts the computation to pure states on a larger space, where geometry is transparent and proofs are concise.
- Monotonicity and concavity of fidelity (next two slides) follow almost immediately as corollaries.

Proof of Uhlmann's Theorem: Setup

- Fix the purification $|\psi\rangle = \sum_i \left(\sqrt{\rho}\,|i\rangle_Q\right) \otimes |i\rangle_R$ of ρ.
- Any purification $|\phi\rangle$ of σ can be written as $(I_Q \otimes V_R) \sum_j \sqrt{\sigma}\,|j\rangle_Q\,|j\rangle_R$ for some partial isometry V on R. Compute the inner product:

$$\langle\psi|\phi|\psi|\phi\rangle = \sum_{i,j} \left(\langle i|_Q \langle i|_R \sqrt{\rho}\right)\left(\sqrt{\sigma}\,|j\rangle_Q\, V_R\,|j\rangle_R\right)$$
$$= \sum_{i,j} \langle i|\sqrt{\rho}\sqrt{\sigma}\,|j\rangle_Q \cdot \langle i|\,V\,|j\rangle_R = \mathrm{tr}\Big(\sqrt{\rho}\sqrt{\sigma}\,V^\top\Big).$$

- The problem reduces to: maximize $\left|\mathrm{tr}(A\,U)\right|$ over all unitaries $U \equiv V^\top$, where $A = \sqrt{\rho}\sqrt{\sigma}$.

Proof of Uhlmann's Theorem: Maximization

- Apply the (left) polar decomposition to $A = \sqrt{\rho}\sqrt{\sigma}$:

$$\sqrt{\rho}\sqrt{\sigma} = \sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}\;U_0,$$

  where U_0 is a unitary and $|A| = \sqrt{A^\dagger A} = \sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}$.
- Choose $U = U_0^\dagger$. Then:

$$\operatorname{tr}(A\,U_0^\dagger) = \operatorname{tr}\left(\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}\;U_0 U_0^\dagger\right) = \operatorname{tr}\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}} = F(\rho,\sigma).$$

- For any other unitary U', the **von Neumann trace inequality** gives $|\operatorname{tr}(A\,U')| \leq \operatorname{tr}|A|$, so $F(\rho,\sigma)$ is the global maximum. $\blacksquare$
  - ◇ *Von Neumann trace inequality:* $|\operatorname{tr}(AB)| \leq \sum_k \sigma_k(A)\,\sigma_k(B)$, where σ_k denote singular values arranged in decreasing order. In particular, $|\operatorname{tr}(AU')| \leq \operatorname{tr}|A| \cdot \operatorname{tr}|U'|/n \leq \operatorname{tr}|A|$.

The Bures Angle and Metric

- Fidelity itself is not a metric (it does not satisfy the triangle inequality). However, valid metrics can be built from it.

Definition: Bures Angle

The angle between two states ρ and σ is:

$$A(\rho, \sigma) \equiv \arccos F(\rho, \sigma), \quad A \in \left[0, \frac{\pi}{2}\right].$$

- ◇ By Uhlmann's Theorem, $A(\rho, \sigma)$ is the *minimum angle* between any purification $|\psi\rangle$ of ρ and any purification $|\phi\rangle$ of σ on a hypersphere.
- ◇ For pure states: $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi|\psi|\phi\rangle|$, so

$$A(|\psi\rangle, |\phi\rangle) = \arccos|\langle\psi|\phi|\psi|\phi\rangle| = \text{geodesic angle}.$$

- ◇ Spherical distance satisfies the triangle inequality, so $A(\rho, \sigma)$ is a metric:

$$A(\rho, \tau) \leq A(\rho, \sigma) + A(\sigma, \tau).$$

The Bures Distance

- An equivalent Euclidean-type metric is the **Bures distance**: $d_B(\rho, \sigma) \equiv \sqrt{2(1 - F(\rho, \sigma))}$.
- **Relationship to Bures angle $\theta = A(\rho, \sigma)$:** Via the chord-length identity,

$$d_B = \sqrt{2(1 - \cos\theta)} = 2\sin\left(\frac{\theta}{2}\right).$$

  This equals the Euclidean chord between two unit vectors at angle θ on a hypersphere.
- **Metric verification sketch:** Since $d_B = 2\sin(A/2)$ and A is a metric, the triangle inequality for d_B holds by sub-additivity of sine on $[0, \pi]$:

$$2\sin\left(\tfrac{A(\rho,\tau)}{2}\right) \leq 2\sin\left(\tfrac{A(\rho,\sigma)}{2}\right) + 2\sin\left(\tfrac{A(\sigma,\tau)}{2}\right).$$

- **Pure states:** $F = |\langle\psi|\phi|\psi|\phi\rangle|$, so $d_B = 2\sqrt{1 - |\langle\psi|\phi|\psi|\phi\rangle|^2} =$ twice the sine of the geodesic angle.
- **Exercise:** Verify positive definiteness and symmetry directly. (Assigned.)

Fidelity Properties: I. Monotonicity under Quantum Operations

Theorem

Let $\mathcal{E}$ be a trace-preserving quantum operation. Then:

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma).$$

- **Physical meaning:** Fidelity is *non-decreasing* under channels — noise can only bring states closer together in the fidelity sense, never push them further apart.
- **Proof (via Uhlmann):**
  - Let $|\psi\rangle, |\phi\rangle$ be purifications of ρ, σ achieving $F(\rho, \sigma) = |\langle\psi|\phi|\psi|\phi\rangle|$.
  - Model $\mathcal{E}$ via a unitary U on system + environment initialized to $|0\rangle_E$.
  - Then $U|\psi\rangle|0\rangle_E$ and $U|\phi\rangle|0\rangle_E$ are valid purifications of $\mathcal{E}(\rho)$ and $\mathcal{E}(\sigma)$ on the enlarged space.
  - By Uhlmann's Theorem:

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq \left|\langle\psi|\langle 0| U^\dagger U |\phi\rangle|0\rangle\right| = |\langle\psi|\phi|\psi|\phi\rangle| = F(\rho, \sigma). \blacksquare$$

Fidelity Properties: II. Strong Concavity

Theorem

Let $\{p_i\}$ and $\{q_i\}$ be probability distributions. Then:

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geq \sum_i \sqrt{p_i q_i}\, F(\rho_i, \sigma_i).$$

- Fidelity is *concave* in each ensemble when $p_i = q_i$, and more generally increases with mixing.
- **Proof (via Uhlmann):** Let $|\psi_i\rangle, |\phi_i\rangle$ purify ρ_i, σ_i with $\langle\psi_i|\phi_i|\psi_i|\phi_i\rangle = F(\rho_i, \sigma_i) \geq 0$. Introduce an ancilla $\{|i\rangle\}$ and define $|\Psi\rangle = \sum_i \sqrt{p_i}\,|\psi_i\rangle\,|i\rangle$ and $|\Phi\rangle = \sum_i \sqrt{q_i}\,|\phi_i\rangle\,|i\rangle$, which purify $\sum_i p_i\rho_i$ and $\sum_i q_i\sigma_i$. By Uhlmann's Theorem:

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geq |\langle\Psi|\Phi|\Psi|\Phi\rangle| = \sum_i \sqrt{p_i q_i}\, F(\rho_i, \sigma_i). \blacksquare$$

Relationships Between Distance Measures

Theorem (Fuchs–van de Graaf Inequalities)

For any two density operators ρ and σ:

$$1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}.$$

- Trace distance and fidelity are *qualitatively equivalent*: closeness in one measure implies closeness in the other.
  - The **lower bound** says: if $D(\rho, \sigma) \to 0$ then $F(\rho, \sigma) \to 1$.
  - The **upper bound** says: if $F(\rho, \sigma) \to 1$ then $D(\rho, \sigma) \to 0$.
- These bounds allow us to freely switch between the two distance measures in proofs, using whichever is more convenient.
- For pure states $\rho = |\psi\rangle\langle\psi|$, $\sigma = |\phi\rangle\langle\phi|$, with $F = |\langle\psi|\phi|\psi|\phi\rangle|$:

$$D = \sqrt{1 - F^2},$$

so the upper bound is *tight* for pure states.

Proof of Fuchs–van de Graaf: Upper Bound

- **Goal:** Show $D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}$.
- Let $|\psi\rangle$ and $|\phi\rangle$ be purifications of ρ and σ chosen such that $F(\rho, \sigma) = |\langle\psi|\phi|\psi|\phi\rangle|$ (achievable by Uhlmann's Theorem).
- By contractivity of trace distance under the partial trace over the reference system R:

$$D(\rho, \sigma) = D(\mathrm{tr}_R |\psi\rangle\langle\psi|, \mathrm{tr}_R |\phi\rangle\langle\phi|) \leq D(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|).$$

- Compute the pure-state trace distance directly. The eigenvalues of $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$ are $\pm\sqrt{1 - |\langle\psi|\phi|\psi|\phi\rangle|^2}$ (and zeros), so:

$$D(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = \sqrt{1 - |\langle\psi|\phi|\psi|\phi\rangle|^2} = \sqrt{1 - F(\rho, \sigma)^2}.$$

- Combining: $D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}$. ■

Proof of Fuchs–van de Graaf: Lower Bound

- **Goal:** Show $1 - F(\rho, \sigma) \leq D(\rho, \sigma)$.
- Choose a POVM $\{E_m\}$ that maximizes classical trace distance: $D(\rho, \sigma) = D(p_m, q_m)$, where $p_m = \mathrm{tr}(\rho E_m)$, $q_m = \mathrm{tr}(\sigma E_m)$.
- By monotonicity of fidelity under the POVM measurement (a quantum operation): $F(\rho, \sigma) \leq F(p_m, q_m) = \sum_m \sqrt{p_m q_m}$.
- From the algebraic identity $(\sqrt{p_m} - \sqrt{q_m})^2 = p_m + q_m - 2\sqrt{p_m q_m}$, summing over m:

$$2(1 - F(p_m, q_m)) = \sum_m \left(\sqrt{p_m} - \sqrt{q_m}\right)^2.$$

Since $(\sqrt{p_m} - \sqrt{q_m})^2 \leq (\sqrt{p_m} + \sqrt{q_m})|\sqrt{p_m} - \sqrt{q_m}| = |p_m - q_m|$:

$$1 - F(\rho, \sigma) \leq 1 - F(p_m, q_m) \leq \frac{1}{2} \sum_m |p_m - q_m| = D(\rho, \sigma). \blacksquare$$

Information Preservation in Channels

- We now ask: how well does a quantum channel $\mathcal{E}$ preserve a state?
  - ◇ For a known pure input $|\psi\rangle$, performance is quantified by the *output state fidelity*:

$$F = F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)) = \sqrt{\langle\psi|\,\mathcal{E}(|\psi\rangle\langle\psi|)\,|\psi\rangle}.$$

  - ◇ **Example — depolarizing channel** $\mathcal{E}(\rho) = (1-p)\rho + p\,\frac{I}{2}$. Evaluating at input $|\psi\rangle\langle\psi|$:

$$\begin{aligned}\langle\psi|\,\mathcal{E}(|\psi\rangle\langle\psi|)\,|\psi\rangle &= (1-p)\langle\psi|\psi\rangle\langle\psi|\psi\rangle + \frac{p}{2}\langle\psi|\psi\rangle \\ &= (1-p) + \frac{p}{2} = 1 - \frac{p}{2},\end{aligned}$$

    so $F = \sqrt{1 - p/2}$, independent of the choice of $|\psi\rangle$ by isotropy.
- **Limitation:** Output state fidelity depends on the specific input $|\psi\rangle$. A real source produces mixed states or states *entangled* with an environment. This motivates a universal dynamic measure: *entanglement fidelity*.

Entanglement Fidelity

- Let system Q be in state ρ, entangled with a fictitious reference system R, with purification $|\psi\rangle_{RQ}$ satisfying $\mathrm{tr}_R(|\psi\rangle\langle\psi|) = \rho$.
- The channel $\mathcal{E}$ acts **only** on Q (as $\mathcal{I}_R \otimes \mathcal{E}_Q$). How well is the global correlation preserved?

Definition: Entanglement Fidelity

$$F(\rho, \mathcal{E}) \equiv F(|\psi\rangle, (\mathcal{I}_R \otimes \mathcal{E}_Q)(|\psi\rangle\langle\psi|))^2 = \langle\psi| (\mathcal{I}_R \otimes \mathcal{E}_Q)(|\psi\rangle\langle\psi|) |\psi\rangle .$$

- The second equality uses $F(|\psi\rangle, \sigma)^2 = \langle\psi| \sigma |\psi\rangle$ for pure $|\psi\rangle$ (verified on the next slide).
- **Independence of purification:** Any two purifications of ρ are related by a unitary V on R alone. Since $\mathcal{E}$ acts only on Q, it commutes with $V \otimes I_Q$, so $F(\rho, \mathcal{E})$ is unchanged. It depends only on ρ and $\mathcal{E}$.

Simplifying Entanglement Fidelity

- Write $\Lambda = |\psi\rangle\langle\psi|$ and $\sigma = (\mathcal{I}_R \otimes \mathcal{E}_Q)(|\psi\rangle\langle\psi|)$.
  - Since Λ is a rank-one projector, $\sqrt{\Lambda} = \Lambda = |\psi\rangle\langle\psi|$.
  - Simplify the inner expression:

$$\sqrt{\Lambda}\,\sigma\,\sqrt{\Lambda} = |\psi\rangle\underbrace{\langle\psi|\,\sigma\,|\psi\rangle}_{\text{scalar}}\langle\psi| = \langle\psi|\,\sigma\,|\psi\rangle \cdot |\psi\rangle\langle\psi|\,.$$

  - Taking the square root and the trace:

$$F(\Lambda, \sigma) = \sqrt{\langle\psi|\,\sigma\,|\psi\rangle}.$$

  - Therefore:

$$F(\rho, \mathcal{E}) \equiv F(|\psi\rangle, \sigma)^2 = \langle\psi|\,\sigma\,|\psi\rangle = \langle\psi|\,(\mathcal{I}_R \otimes \mathcal{E}_Q)(|\psi\rangle\langle\psi|)\,|\psi\rangle\,.$$

- **Intuition:** Entanglement fidelity measures the overlap of the global joint state *before* and *after* the channel acts on Q. A value of 1 means the channel perfectly preserves the system's entanglement with the outside world.

Entanglement Fidelity: Kraus Operator Formula

Theorem

Let $\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$. Then $F(\rho, \mathcal{E}) = \sum_k |\operatorname{tr}(\rho E_k)|^2$.

- **Proof:** Let $|\psi\rangle = \sum_i \sqrt{p_i}\, |i\rangle_R\, |i\rangle_Q$ purify $\rho = \sum_i p_i\, |i\rangle\langle i|$. Since $(I_R \otimes E_k)\, |\psi\rangle = \sum_j \sqrt{p_j}\, |j\rangle_R\, (E_k\, |j\rangle_Q)$,

$$\langle\psi|\,(I_R \otimes E_k)\,|\psi\rangle = \sum_{i,j} \sqrt{p_i p_j}\, \underbrace{\langle i|j|i|j\rangle_R}_{\delta_{ij}}\, \langle i|_Q\, E_k\, |j\rangle_Q$$

$$= \sum_i p_i\, \langle i|\, E_k\, |i\rangle = \operatorname{tr}(\rho\, E_k).$$

  Squaring and summing over k gives $F(\rho, \mathcal{E}) = \sum_k |\operatorname{tr}(\rho\, E_k)|^2$. ■

- This formula is computationally accessible: no purification or matrix square root is required, only traces of ρ against each Kraus operator.

Properties of Entanglement Fidelity

- **Boundedness:** $0 \leq F(\rho, \mathcal{E}) \leq 1$.
- **Channel linearity:** $F(\rho, \mathcal{E})$ is linear in the channel. If $\mathcal{E} = \sum_j p_j \mathcal{E}_j$ is a probabilistic mixture of channels, then:

$$F(\rho, \mathcal{E}) = \sum_j p_j F(\rho, \mathcal{E}_j).$$

- **Pure input:** For $\rho = |\psi\rangle \langle\psi|$, entanglement fidelity reduces to output state fidelity squared:

$$F(|\psi\rangle, \mathcal{E}) = \langle\psi| \mathcal{E}(|\psi\rangle \langle\psi|) |\psi\rangle .$$

- **Rigidity:** $F(\rho, \mathcal{E}) = 1$ if and only if $\mathcal{E}(|\psi\rangle \langle\psi|) = |\psi\rangle \langle\psi|$ for every pure state $|\psi\rangle$ in the support of ρ — the channel acts as the identity on the code subspace.
- **Relationship to error correction:** $F(\rho, \mathcal{E})$ is the primary figure of merit for quantum error-correcting codes; maximizing it over all possible recovery operations determines the best achievable recovery.

Summary of This Lecture

- **Classical foundations.** Trace distance $D(p,q) = \frac{1}{2}\sum_x |p_x - q_x|$ is the maximum-event distinguishability; fidelity $F(p,q) = \sum_x \sqrt{p_x q_x}$ is a vector inner product. Introducing a reference copy converts dynamic channel performance into a static distance.
- **Quantum trace distance** $D(\rho,\sigma) = \frac{1}{2}\operatorname{tr}|\rho - \sigma|$: equals $\frac{1}{2}|\vec{r} - \vec{s}|$ on the Bloch sphere; is *contractive* under all channels; equals max classical D over all POVMs.
- **Quantum fidelity** $F(\rho,\sigma) = \operatorname{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$: symmetric; reduces to $|\langle\psi|\phi|\psi|\phi\rangle|$ for pure states; is the *maximum purification overlap* (Uhlmann); is *monotone* (non-decreasing) under channels.
- **Fuchs–van de Graaf:** $1 - F \leq D \leq \sqrt{1 - F^2}$. The two measures are qualitatively interchangeable; equality $D = \sqrt{1 - F^2}$ holds for pure states.
- **Entanglement fidelity** $F(\rho,\mathcal{E}) = \sum_k |\operatorname{tr}(\rho E_k)|^2$: measures how faithfully a channel $\mathcal{E}$ preserves entanglement with a reference; $F = 1$ iff $\mathcal{E}$ acts as the identity on the support of ρ.

Look Ahead: Lecture 13

- We now have tools to *quantify* how noise damages quantum information. The final lecture asks: **can we actively fight back?**
- **Lecture 13: Quantum Error Correction**
  - **Why QEC is hard:** three obstacles — no-cloning, continuity of errors, and measurement disturbance.
  - **Basic codes:** 3-qubit bit-flip and phase-flip codes; syndrome diagnosis via ancilla parity measurements that extract error information without disturbing logical superpositions.
  - **Shor 9-qubit code:** the first code correcting an arbitrary single-qubit error, built by concatenating bit-flip and phase-flip protection.
  - **Knill–Laflamme conditions:** $PE_i^\dagger E_j P = \alpha_{ij} P$ — the algebraic condition for a recovery map to exist; continuous errors are discretized by syndrome measurement.
  - **Stabilizer formalism:** the Pauli group, stabilizer and CSS codes, Steane $[7, 1, 3]$ code, and the perfect $[[5, 1, 3]]$ code.
  - **Fault tolerance:** transversal gates, the Eastin–Knill theorem, magic state distillation, and the threshold theorem.