FRAMES IN QUANTUM AND CLASSICAL INFORMATION THEORY

Emina Soljanin
Mathematical Sciences Research Center, Bell Labs

April 16, 2003

A FRAME

- A sequence $\{x_i\}$ of vectors in a Hilbert space with the property that there are constants $A, B \geq 0$ such that

$$A\|x\|^2 \leq \sum_i |\langle x, x_i \rangle|^2 \leq B\|x\|^2$$

for all x in the Hilbert space.

- Examples?

A SOURCE OF INFORMATION

Classical

- Discrete: produces sequences of letters.
- Letters belong to a finite alphabet $\mathcal{X}$.

A SOURCE OF INFORMATION
Classical

- Discrete: produces sequences of letters.
- Letters belong to a finite alphabet $\mathcal{X}$.
- Memoryless: each letter is produced independently.
- Probability of letter a is P_a.

A SOURCE OF INFORMATION
Classical

- Discrete: produces sequences of letters.
- Letters belong to a finite alphabet $\mathcal{X}$.
- Memoryless: each letter is produced independently.
- Probability of letter a is P_a.
- Example: coin tossing with $\mathcal{X} = \{H, T\}$.

A SOURCE OF INFORMATION
Quantum

- Letters are transmitted as d-dimensional unit-length vectors.

A SOURCE OF INFORMATION

Quantum

- Letters are transmitted as d-dimensional unit-length vectors.

- $|e_0\rangle$ and $|e_1\rangle$ are the basis vectors of 2D space $\mathcal{H}_2$:

$$|e_0\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad |e_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- A qubit is a vector in $\mathcal{H}_2$: $|\psi\rangle = \alpha|e_0\rangle + \beta|e_1\rangle$

A SOURCE OF INFORMATION
Quantum

- Letters are transmitted as d-dimensional unit-length vectors.
- $|e_0\rangle$ and $|e_1\rangle$ are the basis vectors of 2D space $\mathcal{H}_2$:

$$|e_0\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad |e_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- A qubit is a vector in $\mathcal{H}_2$: $|\psi\rangle = \alpha|e_0\rangle + \beta|e_1\rangle$
- Example: $\mathcal{X} = \{0, 1, 2, 3\}$,

$$|\psi_0\rangle = \alpha_0|e_0\rangle + \beta_0|e_1\rangle \quad |\psi_1\rangle = \alpha_1|e_0\rangle + \beta_1|e_1\rangle$$
$$|\psi_2\rangle = \alpha_2|e_0\rangle + \beta_2|e_1\rangle \quad |\psi_3\rangle = \alpha_3|e_0\rangle + \beta_3|e_1\rangle.$$

QUANTUM DISCRETE MEMORYLESS SOURCE

The Density Matrix and Von Neumann Entropy

- Source density matrix:

$$\rho = \sum_{a \in \mathcal{X}} P_a |\psi_a\rangle\langle\psi_a|.$$

QUANTUM DISCRETE MEMORYLESS SOURCE

The Density Matrix and Von Neumann Entropy

- Source density matrix:

$$\rho = \sum_{a \in \mathcal{X}} P_a |\psi_a\rangle\langle\psi_a|.$$

- Von Neumann entropy of the source:

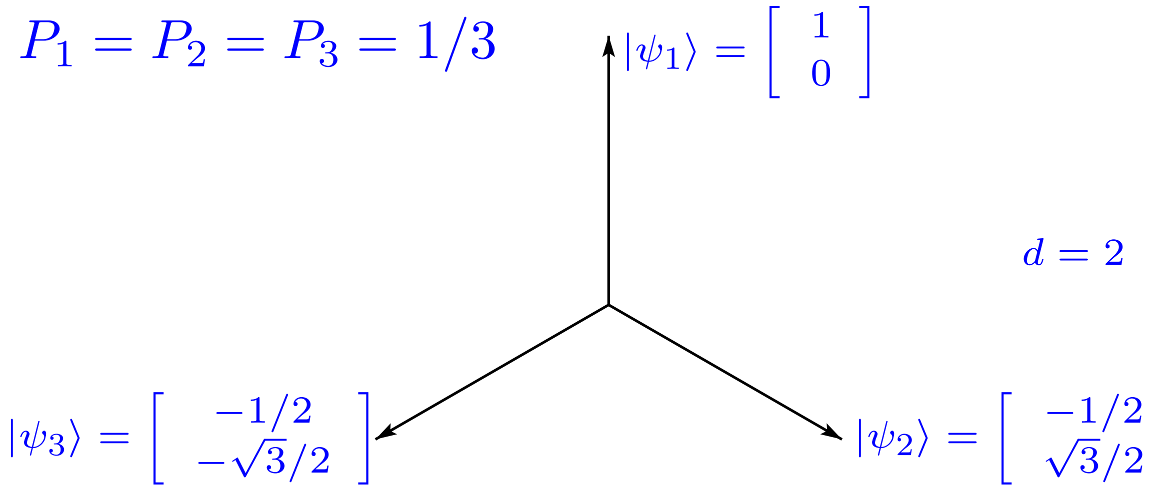
$$\begin{aligned} S(\rho) &= -\operatorname{Tr} \rho \log \rho \\ &= -\sum_i \lambda_i \log \lambda_i, \end{aligned}$$

where λ_i are the eigenvalues of ρ.

QUANTUM DISCRETE MEMORYLESS SOURCE

MB Example

$$\mathcal{X} = \{1, 2, 3\} \quad P_1 = P_2 = P_3 = 1/3$$



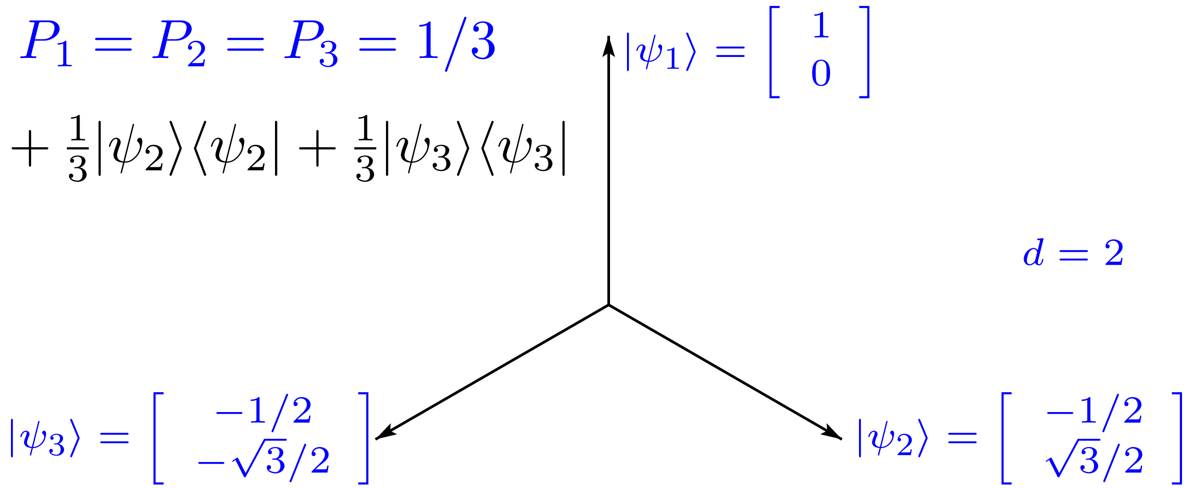
$$d = 2$$

QUANTUM DISCRETE MEMORYLESS SOURCE

MB Example

$$\mathcal{X} = \{1,2,3\} \quad P_1 = P_2 = P_3 = 1/3$$

$$\begin{aligned} \rho &= \tfrac{1}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{1}{3}|\psi_3\rangle\langle\psi_3| \\ &= \tfrac{1}{2}I \end{aligned}$$



$d = 2$

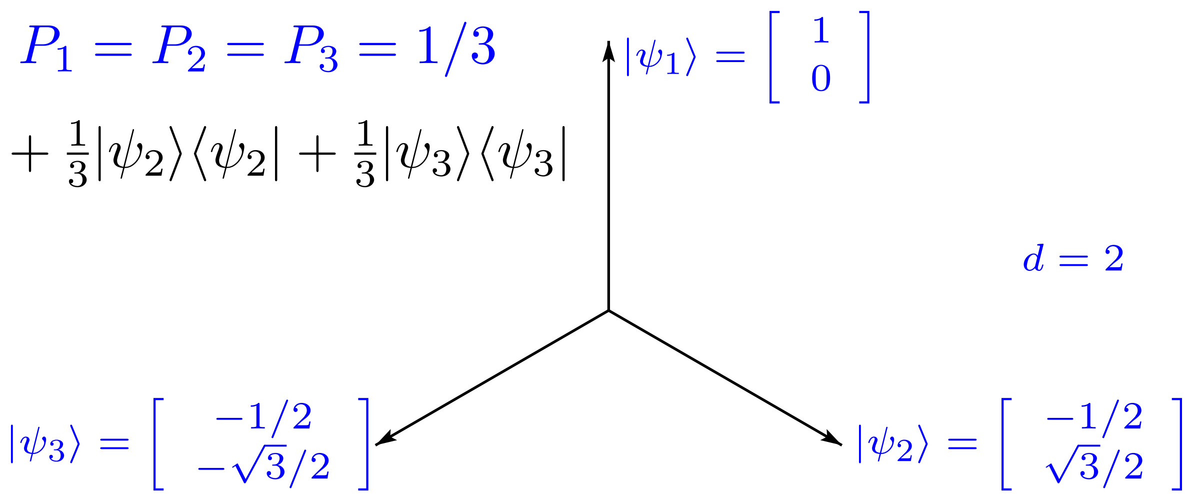
QUANTUM DISCRETE MEMORYLESS SOURCE

MB Example

$$\mathcal{X} = \{1,2,3\} \quad P_1 = P_2 = P_3 = 1/3$$

$$\begin{aligned} \rho &= \tfrac{1}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{1}{3}|\psi_3\rangle\langle\psi_3| \\ &= \tfrac{1}{2}I \end{aligned}$$

$$S(\rho) = 1$$



$d = 2$

QUANTUM DISCRETE MEMORYLESS SOURCE

Vector Sequences

- Sequences of length n are d^n-dimensional vectors.
- Source vector-sequence (state):

$$|\Psi_{\boldsymbol{x}}\rangle = |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle, \qquad x_i \in \mathcal{X}.$$

QUANTUM DISCRETE MEMORYLESS SOURCE

Vector Sequences

- Sequences of length n are d^n-dimensional vectors.

- Source vector-sequence (state):

$$|\Psi_{\boldsymbol{x}}\rangle = |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle, \qquad x_i \in \mathcal{X}.$$

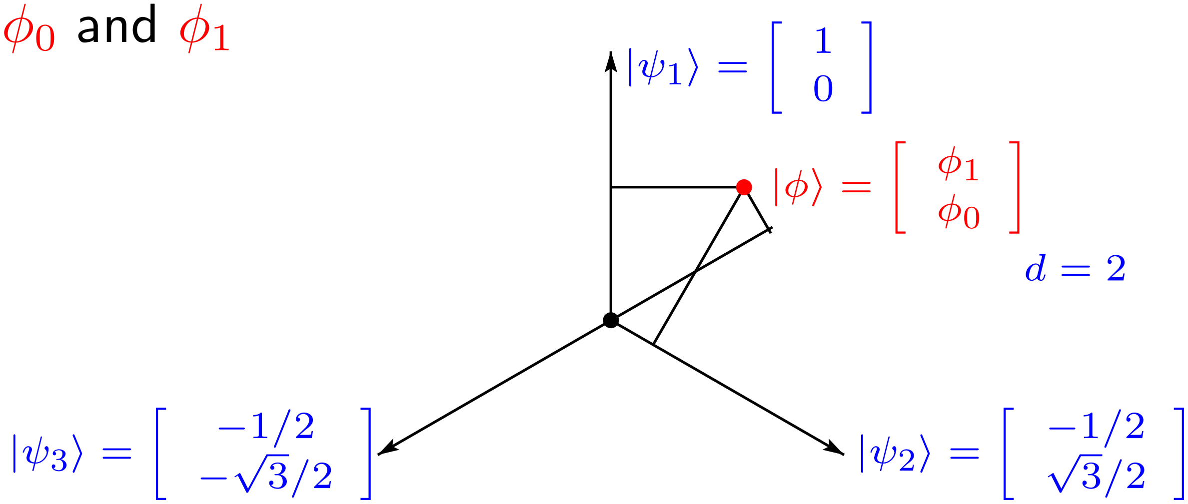
- Among all states that come from the source, we can distinguish

$$2^{n(S(\rho)-\varepsilon_n)}$$

reliably.

SENDING PACKETS OVER LOSSY NETWORKS
MB Example

Send $|\phi\rangle$ by sending ϕ_0 and ϕ_1

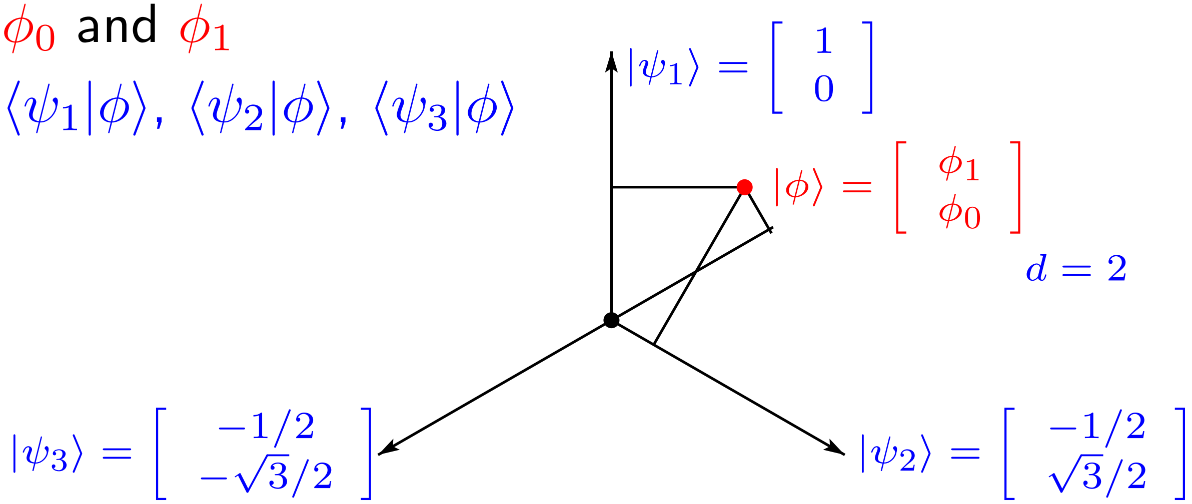


SENDING PACKETS OVER LOSSY NETWORKS

MB Example

Send $|\phi\rangle$ by sending ϕ_0 and ϕ_1

Send $|\phi\rangle$ by sending $\langle\psi_1|\phi\rangle$, $\langle\psi_2|\phi\rangle$, $\langle\psi_3|\phi\rangle$

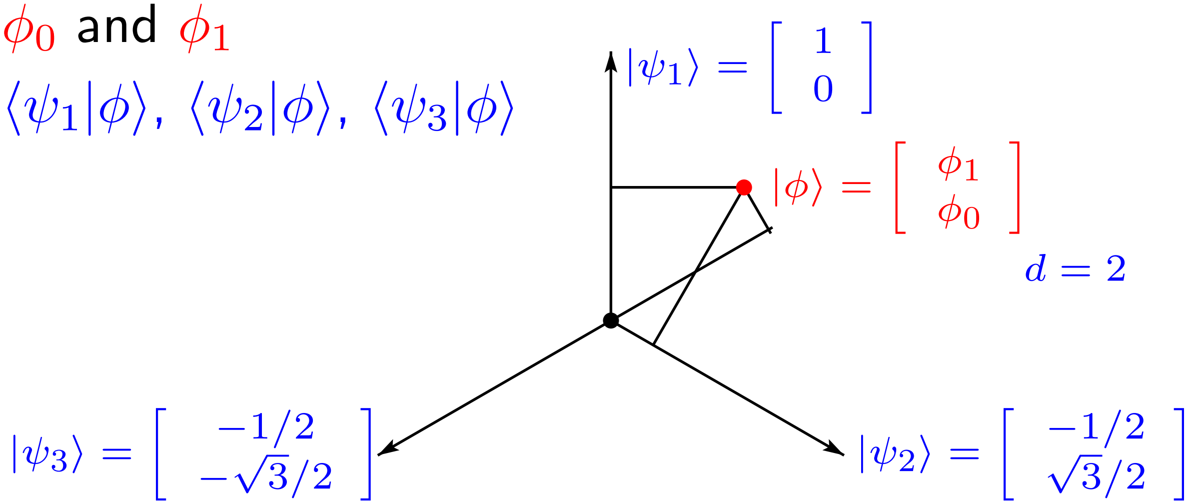


SENDING PACKETS OVER LOSSY NETWORKS

MB Example

Send $|\phi\rangle$ by sending ϕ_0 and ϕ_1

Send $|\phi\rangle$ by sending $\langle\psi_1|\phi\rangle$, $\langle\psi_2|\phi\rangle$, $\langle\psi_3|\phi\rangle$



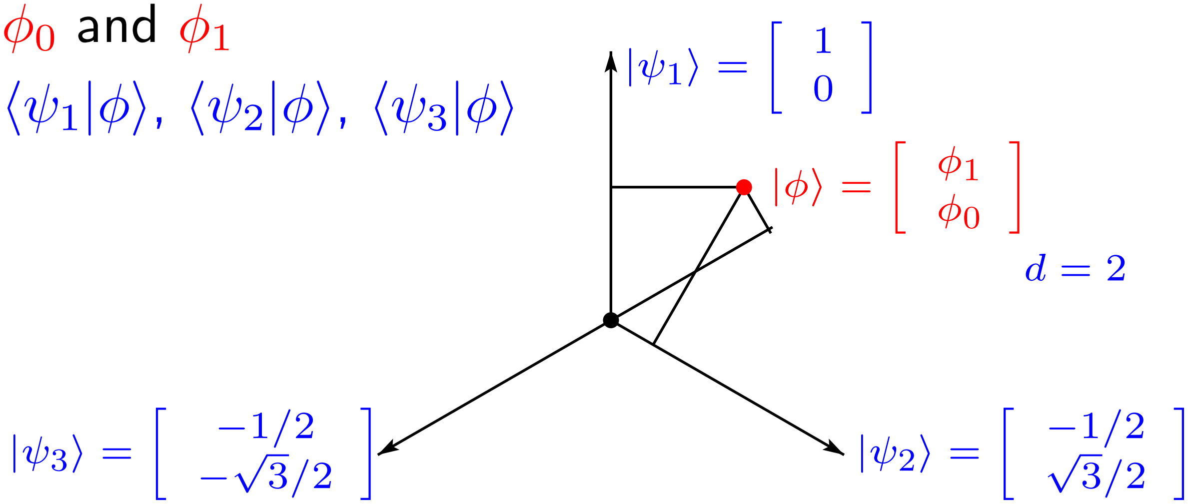
$$|\phi\rangle = \frac{2}{3}(\langle\psi_1|\phi\rangle|\psi_1\rangle + \langle\psi_2|\phi\rangle|\psi_2\rangle + \langle\psi_3|\phi\rangle|\psi_3\rangle)$$

SENDING PACKETS OVER LOSSY NETWORKS

MB Example

Send $|\phi\rangle$ by sending ϕ_0 and ϕ_1

Send $|\phi\rangle$ by sending $\langle\psi_1|\phi\rangle$, $\langle\psi_2|\phi\rangle$, $\langle\psi_3|\phi\rangle$



$$|\phi\rangle = \frac{2}{3}(\langle\psi_1|\phi\rangle|\psi_1\rangle + \langle\psi_2|\phi\rangle|\psi_2\rangle + \langle\psi_3|\phi\rangle|\psi_3\rangle)$$

$$|\phi\rangle = \frac{2}{3}(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + |\psi_3\rangle\langle\psi_3|)|\phi\rangle$$

SYNCHRONOUS CDMA SYSTEMS

K users and processing gain N

- Each user has a signature $N \times 1$ length-$\sqrt{N}$ complex vector.
- Let $\boldsymbol{s_i}$ be the signature and p_i the power of user i for $1 \leq i \leq K$.

SYNCHRONOUS CDMA SYSTEMS

K users and processing gain N

- Each user has a signature $N \times 1$ length-$\sqrt{N}$ complex vector.
- Let $\boldsymbol{s_i}$ be the signature and p_i the power of user i for $1 \leq i \leq K$.
- The received vector is given by

$$\boldsymbol{r} = \sum_{i=1}^{K} \sqrt{p_i} b_i \boldsymbol{s_i} + \boldsymbol{n}$$

where

– b_i is the information symbol, for user i, $E[b_i] = 0$, $E[b_i^2] = 1$;
– $\boldsymbol{n}$ is the (Gaussian) noise vector; $E[\boldsymbol{n}] = \mathbf{0}$, $E[\boldsymbol{n}\boldsymbol{n}^\dagger] = \sigma^2 I_N$.

SYNCHRONOUS CDMA SYSTEMS

The Sum Capacity

- Let user signatures and powers be given:

$$S = [\boldsymbol{s}_{\mathbf{1}}, \ldots \boldsymbol{s}_K] \text{ and } P = \text{diag}\{p_1, \ldots, p_K\}$$

- The sum capacity:

$$C_{\text{sum}} = \frac{1}{2}\log[\det(I_N + \sigma^{-2} S P S^{\dagger})]$$

A COMMON MODEL

The Object of Interest

- An ensemble:
  - K d-dimensional unit-length vectors $|\psi_i\rangle$

A COMMON MODEL

The Object of Interest

- An ensemble:
  - K d-dimensional unit-length vectors $|\psi_i\rangle$
  - K real numbers p_i such that $p_1 + \cdots + p_K = 1$.

A COMMON MODEL

The Object of Interest

- An ensemble:
  - K d-dimensional unit-length vectors $|\psi_i\rangle$
  - K real numbers p_i such that $p_1 + \cdots + p_K = 1$.

- Two matrices:
  - F is the $d \times K$ matrix whose columns are $\sqrt{p_i}|\psi_i\rangle$. Thus

$$FF^\dagger = \sum_{i=1}^{K} p_i |\psi_i\rangle\langle\psi_i|.$$

A COMMON MODEL
The Object of Interest

- An ensemble:
  - K d-dimensional unit-length vectors $|\psi_i\rangle$
  - K real numbers p_i such that $p_1 + \cdots + p_K = 1$.

- Two matrices:
  - F is the $d \times K$ matrix whose columns are $\sqrt{p_i}|\psi_i\rangle$. Thus

$$FF^\dagger = \sum_{i=1}^{K} p_i|\psi_i\rangle\langle\psi_i|.$$

  - $FF^\dagger$ is the density matrix (frame operator)
  - $F^\dagger F$ is the Gram matrix.

A FRAME

- An ensemble $\{\sqrt{p_i}|\psi_i\rangle\}$ of vectors in a Hilbert space with the property that there are constants $A, B \geq 0$ such that

$$A\langle\varphi|\varphi\rangle \leq \sum_i p_i|\langle\varphi|\psi_i\rangle|^2 \leq B\langle\varphi|\varphi\rangle$$

for all $|\varphi\rangle$ in the Hilbert space.

A FRAME

- An ensemble $\{\sqrt{p_i}|\psi_i\rangle\}$ of vectors in a Hilbert space with the property that there are constants $A, B \geq 0$ such that

$$A\langle\varphi|\varphi\rangle \leq \sum_i p_i |\langle\varphi|\psi_i\rangle|^2 \leq B\langle\varphi|\varphi\rangle$$

for all $|\varphi\rangle$ in the Hilbert space.

- Equivalently,

$$AI_d \leq FF^\dagger \leq BI_d$$

A COMMON MODEL
Information Measures

- The Von Neumann entropy:

$$S = -\operatorname{Tr} FF^\dagger \log FF^\dagger$$

- The sum capacity:

$$C_{\text{sum}} = \frac{1}{2}\log[\det(I_d + d\sigma^{-2}FF^\dagger)]$$

A COMMON MODEL
Information Measures

- The Von Neumann entropy:

$$S = -\operatorname{Tr} FF^\dagger \log FF^\dagger$$

- The sum capacity:

$$C_{\text{sum}} = \frac{1}{2} \log[\det(I_d + d\sigma^{-2} FF^\dagger)]$$

- Both are maximized by

$$FF^\dagger =$$

A COMMON MODEL

Information Measures

- The Von Neumann entropy:

$$S = -\operatorname{Tr} FF^\dagger \log FF^\dagger$$

- The sum capacity:

$$C_{\text{sum}} = \frac{1}{2}\log[\det(I_d + d\sigma^{-2}FF^\dagger)]$$

- Both are maximized by

$$FF^\dagger = \frac{1}{d}\, I_d.$$

HOW TO MIX A DENSITY MATRIX

$$FF^\dagger = \sum_{i=1}^{K} p_i |\psi_i\rangle\langle\psi_i|$$

- Two problems:
  - classification of ensembles having a given density matrix,

HOW TO MIX A DENSITY MATRIX

$$FF^\dagger = \sum_{i=1}^{K} p_i |\psi_i\rangle\langle\psi_i|$$

- Two problems:
  - classification of ensembles having a given density matrix,
  - characterization of PDs consistent with a given density matrix.

HOW TO MIX A DENSITY MATRIX

$$FF^{\dagger} = \sum_{i=1}^{K} p_i |\psi_i\rangle\langle\psi_i|$$

- Two problems:
  - classification of ensembles having a given density matrix,
  - characterization of PDs consistent with a given density matrix.
- Let $\{p_i\}$ be a PD, and $p_1 \geq p_2 \geq \cdots \geq p_K$.

HOW TO MIX A DENSITY MATRIX

$FF^\dagger = \sum_{i=1}^{K} p_i |\psi_i\rangle\langle\psi_i|$

- Two problems:
  - classification of ensembles having a given density matrix,
  - characterization of PDs consistent with a given density matrix.
- Let $\{p_i\}$ be a PD, and $p_1 \geq p_2 \geq \cdots \geq p_K$.
- Let ρ be a density matrix, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ its eigenvalues.
- There exist vectors $|\psi_i\rangle$ such that $\rho = \sum_{i=1}^{K} p_i |\psi_i\rangle\langle\psi_i|$ iff

$$\sum_{i=1}^{n} p_i \leq \sum_{i=1}^{n} \lambda_i \quad \text{for all } n < d.$$

HOW TO MIX A DENSITY MATRIX

$$FF^\dagger = \frac{1}{d} I_d$$

- For $\rho = \frac{1}{d} I_d$, condition

$$\sum_{i=1}^{n} p_i \leq \sum_{i=1}^{n} \lambda_i \text{ for all } n < d.$$

becomes $p_1 \leq 1/d$.

HOW TO MIX A DENSITY MATRIX

$$FF^\dagger = \frac{1}{d} I_d$$

- For $\rho = \frac{1}{d} I_d$, condition

$$\sum_{i=1}^{n} p_i \leq \sum_{i=1}^{n} \lambda_i \ \text{ for all } n < d.$$

  becomes $p_1 \leq 1/d$.

- In CDMA, user i is said to be oversized if

$$p_i > \frac{\sum_{j=i+1}^{K} p_j}{d - i}$$

INTERFERENCE MEASURE IN CDMA
Total Square Correlation (TSC)

- The Welch's lower bound to TSC (frame potential):

$$\sum_{i=1}^{K}\sum_{j=1}^{K} p_i p_j |\langle\psi_i|\psi_j\rangle|^2 \geq \frac{1}{d}$$

INTERFERENCE MEASURE IN CDMA
Total Square Correlation (TSC)

- The Welch's lower bound to TSC (frame potential):

$$\sum_{i=1}^{K}\sum_{j=1}^{K} p_i p_j |\langle\psi_i|\psi_j\rangle|^2 \geq \frac{1}{d}$$

with equality iff $\sum_{i=1}^{K} p_i|\psi_i\rangle\langle\psi_i| = \frac{1}{d} I_d$.

- WBE sequences
  - minimize the TSC
  - maximize the sum capacity and Von Neumann entropy

INTERFERENCE MEASURE IN CDMA
Total Square Correlation (TSC)

- The Welch's lower bound to TSC (frame potential):

$$\sum_{i=1}^{K}\sum_{j=1}^{K} p_i p_j |\langle\psi_i|\psi_j\rangle|^2 \geq \frac{1}{d}$$

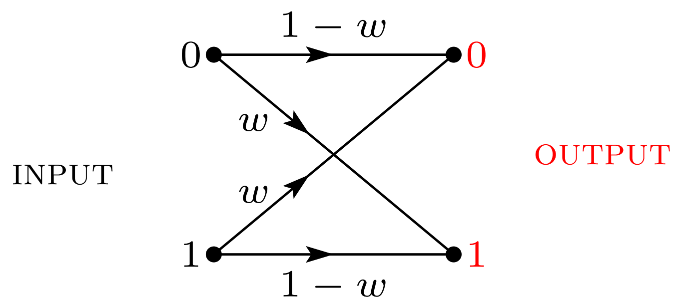
  with equality iff $\sum_{i=1}^{K} p_i|\psi_i\rangle\langle\psi_i| = \frac{1}{d} I_d$.

- WBE sequences
  - minimize the TSC
  - maximize the sum capacity and Von Neumann entropy

- What does reducing TSC mean?

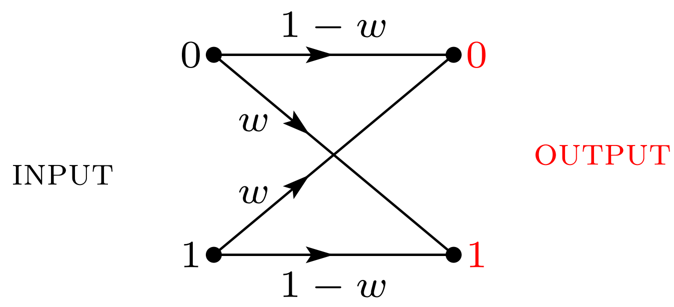
SOME COMMUNICATION CHANNELS

- Binary Symmetric Channel:

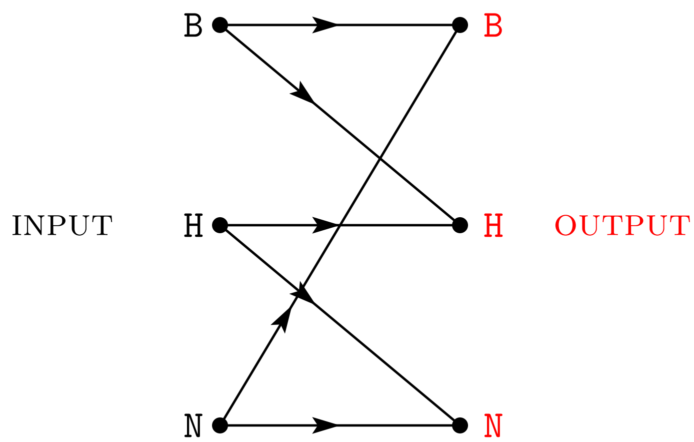


SOME COMMUNICATION CHANNELS

- Binary Symmetric Channel:

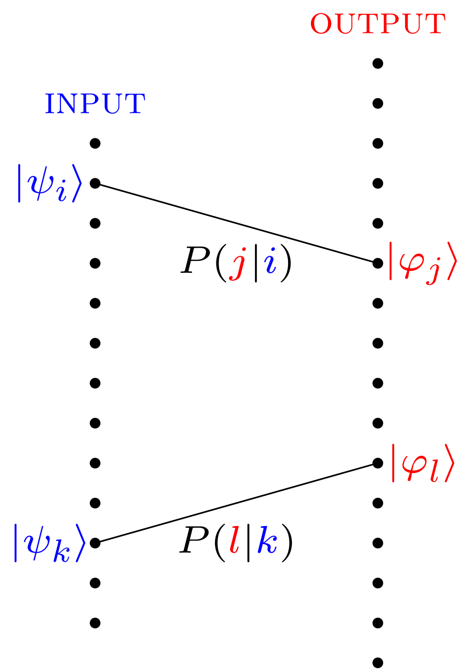


- Noisy Typewriter:



A CQ COMMUNICATION CHANNEL
A Probabilistic Device

- Inputs: vectors $|\psi_i\rangle$, $i \in \mathcal{X}$.
- Outputs: vectors $|\varphi_j\rangle$ determined by the chosen measurement.
- Transition probabilities determined by the chosen measurement.



QUANTUM MEASUREMENT

Von Neumann's Measurement

- A set of pairwise orthogonal projection operators $\{\Pi_i\}$.
- They form a complete resolution of the identity: $\sum_i \Pi_i = I$.

QUANTUM MEASUREMENT

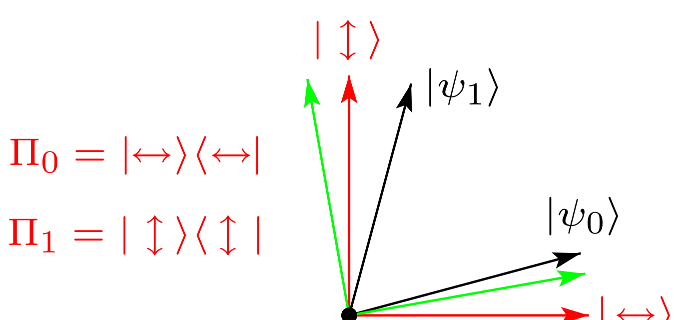
Von Neumann's Measurement

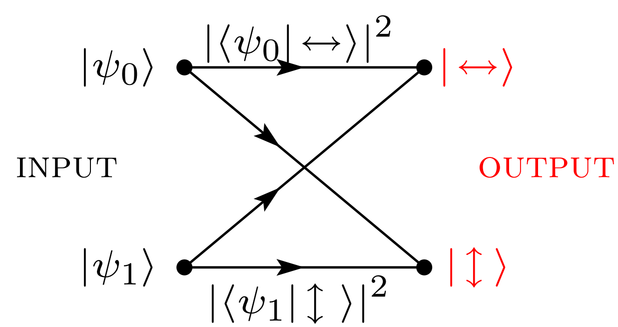
- A set of pairwise orthogonal projection operators $\{\Pi_i\}$.
- They form a complete resolution of the identity: $\sum_i \Pi_i = I$.
- For input $|\psi_j\rangle$, output $\Pi_i|\psi_j\rangle$ happens with probability $\langle\psi_j|\Pi_i|\psi_j\rangle$.

QUANTUM MEASUREMENT

Von Neumann's Measurement

- A set of pairwise orthogonal projection operators $\{\Pi_i\}$.
- They form a complete resolution of the identity: $\sum_i \Pi_i = I$.
- For input $|\psi_j\rangle$, output $\Pi_i|\psi_j\rangle$ happens with probability $\langle\psi_j|\Pi_i|\psi_j\rangle$.
- Example:





QUANTUM MEASUREMENT

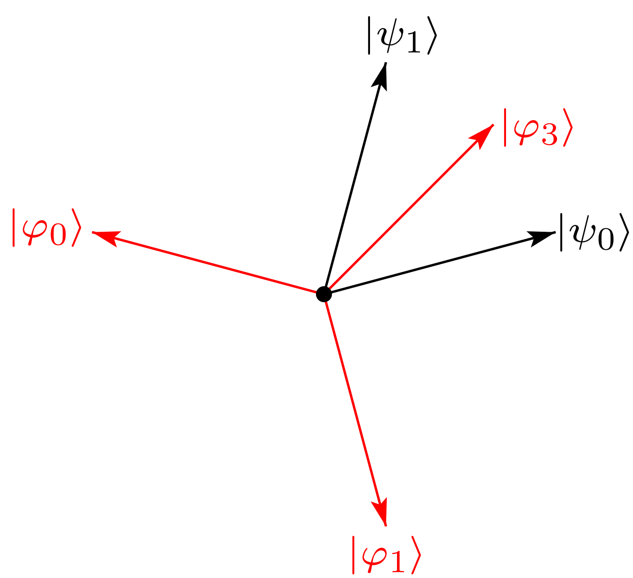
Positive Operator-Valued Measure

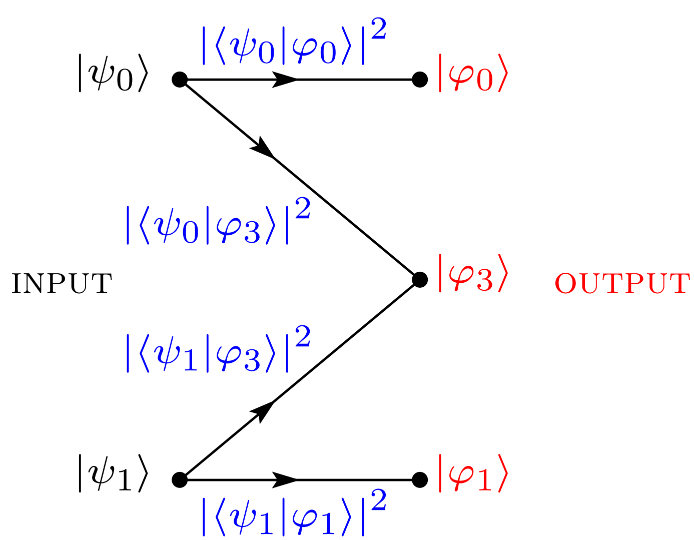
- Any set of positive-semidefinite operators $\{E_i\}$.
- They form a complete resolution of the identity: $\sum_i E_i = I$.
- For input $|\psi_j\rangle$, output $E_i|\psi_j\rangle$ happens with probability $\langle\psi_j|E_i|\psi_j\rangle$.

QUANTUM MEASUREMENT

Positive Operator-Valued Measure

- Any set of positive-semidefinite operators $\{E_i\}$.
- They form a complete resolution of the identity: $\sum_i E_i = I$.
- For input $|\psi_j\rangle$, output $E_i|\psi_j\rangle$ happens with probability $\langle\psi_j|E_i|\psi_j\rangle$.
- Example:





OPTIMAL QUANTUM MEASUREMENTS

Some Open Problems

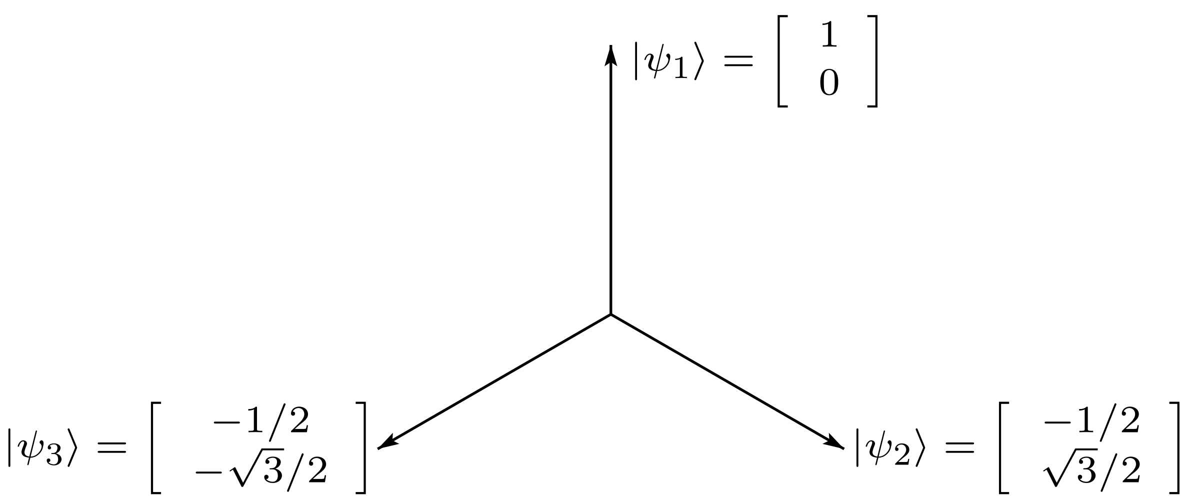
- POVMs minimizing the detection error-probability.

OPTIMAL QUANTUM MEASUREMENTS
Some Open Problems

- POVMs minimizing the detection error-probability.
- POVMs attaining the accessible information (number of elements)

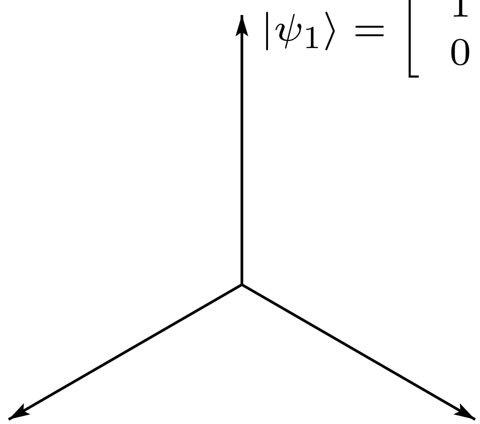
OPTIMAL QUANTUM MEASUREMENTS
Some Open Problems

- POVMs minimizing the detection error-probability.
- POVMs attaining the accessible information (number of elements).
- Example:



OPTIMAL QUANTUM MEASUREMENTS
Some Open Problems

- POVMs minimizing the detection error-probability.
- POVMs attaining the accessible information (number of elements).
- Example:



$$|\widetilde{\psi_1}\rangle = \begin{bmatrix} \sqrt{1-\alpha} \\ 0 \\ \sqrt{\alpha} \end{bmatrix}$$

$$|\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$|\widetilde{\psi_3}\rangle = \begin{bmatrix} -\sqrt{1-\alpha}/2 \\ -\sqrt{3}\sqrt{1-\alpha}/2 \\ \sqrt{\alpha} \end{bmatrix}$$

$$|\psi_3\rangle = \begin{bmatrix} -1/2 \\ -\sqrt{3}/2 \end{bmatrix}$$

$$|\widetilde{\psi_2}\rangle = \begin{bmatrix} -\sqrt{1-\alpha}/2 \\ \sqrt{3}\sqrt{1-\alpha}/2 \\ \sqrt{\alpha} \end{bmatrix}$$

$$|\psi_2\rangle = \begin{bmatrix} -1/2 \\ \sqrt{3}/2 \end{bmatrix}$$

A SOURCE OF INFORMATION
Classical

- Discrete: produces sequences of letters.
- Letters belong to a finite alphabet $\mathcal{X}$.
- Memoryless: each letter is produced independently.
- Probability of letter a is P_a.
- Example: coin tossing with $\mathcal{X} = \{H, T\}$.

CLASSICAL DISCRETE MEMORYLESS SOURCE

Sequences and "Large" Sets

- Source sequence $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ is in $\mathcal{X}^n$.
- $N(a|\boldsymbol{x})$ denotes the number occurrences of a in $\boldsymbol{x}$.

CLASSICAL DISCRETE MEMORYLESS SOURCE

Sequences and “Large” Sets

- Source sequence $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ is in $\mathcal{X}^n$.
- $N(a|\boldsymbol{x})$ denotes the number occurrences of a in $\boldsymbol{x}$.
- Consider all sequences $\boldsymbol{x}$ for which

$$\left|\frac{1}{n}N(a|\boldsymbol{x}) - P_a\right| \le \delta \text{ for every } a \in \mathcal{X}.$$

They form the set of typical sequences $\mathsf{T}^n_{P,\delta}$.

CLASSICAL DISCRETE MEMORYLESS SOURCE

Sequences and "Large" Sets

- Source sequence $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ is in $\mathcal{X}^n$.
- $N(a|\boldsymbol{x})$ denotes the number occurrences of a in $\boldsymbol{x}$.
- Consider all sequences $\boldsymbol{x}$ for which

$$\left|\frac{1}{n}N(a|\boldsymbol{x}) - P_a\right| \leq \delta \quad \text{for every} \quad a \in \mathcal{X}.$$

  They form the set of typical sequences $\mathsf{T}^n_{P,\delta}$.

- Set $\mathsf{T}^n_{P,\delta}$ is probabilistically large:

$$P^n(\mathsf{T}^n_{P,\delta}) \geq 1 - \epsilon_n.$$

DISCRETE MEMORYLESS SOURCE

Shannon Entropy

- $H(P) = -\sum_{a \in \mathcal{X}} P_a \log P_a$.

DISCRETE MEMORYLESS SOURCE
Shannon Entropy

- $H(P) = -\sum_{a \in \mathcal{X}} P_a \log P_a$.
- Set $\mathsf{T}^n_{P,\delta}$ contains approximately $2^{nH(P)}$ sequences:

$$2^{n[H(P)-\epsilon_n]} \leq |\mathsf{T}^n_{P,\delta}| \leq 2^{n[H(P)+\epsilon_n]}$$

DISCRETE MEMORYLESS SOURCE
Shannon Entropy

- $H(P) = -\sum_{a \in \mathcal{X}} P_a \log P_a$.
- Set $\mathsf{T}^n_{P,\delta}$ contains approximately $2^{nH(P)}$ sequences:

$$2^{n[H(P)-\epsilon_n]} \leq |\mathsf{T}^n_{P,\delta}| \leq 2^{n[H(P)+\epsilon_n]}$$

- The probability of typical sequences $\boldsymbol{x}$ is approximately $2^{-nH(P)}$:

$$2^{-n[H(P)+\epsilon'_n]} \leq P_{\boldsymbol{x}} \leq 2^{-n[H(P)-\epsilon'_n]}$$

QUANTUM DISCRETE MEMORYLESS SOURCE

Vector Sequences

- Sequences of length n are d^n-dimensional vectors:

$$|e_0\rangle \otimes |e_0\rangle \quad |e_0\rangle \otimes |e_1\rangle \quad |e_1\rangle \otimes |e_0\rangle \quad |e_1\rangle \otimes |e_1\rangle$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

QUANTUM DISCRETE MEMORYLESS SOURCE

Vector Sequences

- Sequences of length n are d^n-dimensional vectors:

$$|e_0\rangle \otimes |e_0\rangle \quad |e_0\rangle \otimes |e_1\rangle \quad |e_1\rangle \otimes |e_0\rangle \quad |e_1\rangle \otimes |e_1\rangle$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

- Source vector-sequence (state)

$$|\Psi_{\boldsymbol{x}}\rangle = |\psi_{x_1}\rangle \otimes |\psi_{x_2}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle, \qquad x_i \in \mathcal{X},$$

appears with probability $P_{\boldsymbol{x}} = P_{x_1} \cdot P_{x_2} \cdot \ldots \cdot P_{x_n}$.

QUANTUM DISCRETE MEMORYLESS SOURCE

Vector Sequences

- Source vector-sequence (state)

$$|\Psi_{\boldsymbol{x}}\rangle = |\psi_{x_1}\rangle \otimes |\psi_{x_2}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle, \qquad x_i \in \mathcal{X},$$

  appears with probability $P_{\boldsymbol{x}} = P_{x_1} \cdot P_{x_2} \cdot \ldots \cdot P_{x_n}$.

- Typical states $|\Psi_{\boldsymbol{x}}\rangle \in \mathcal{H}^{2^n}$ correspond to typical sequences $\boldsymbol{x}$.

- There are approximately $2^{nH(P)}$ typical states.

QUANTUM DISCRETE MEMORYLESS SOURCE

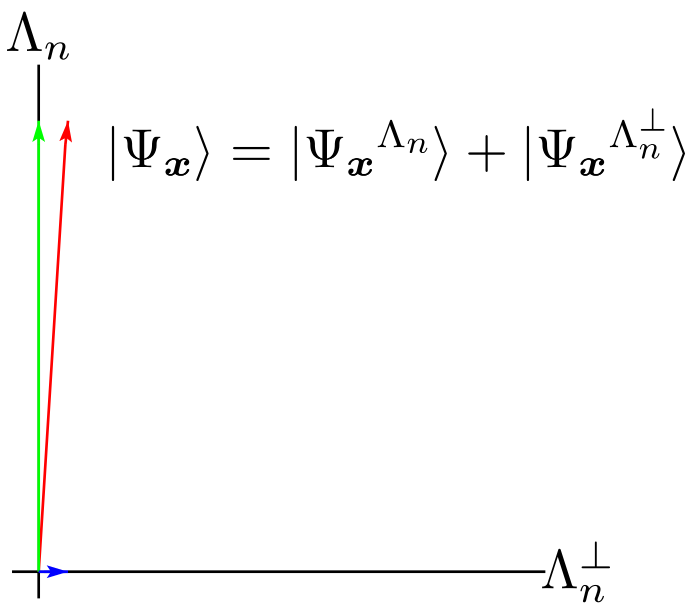
Typical Subspace

- Typical states $|\Psi_{\boldsymbol{x}}\rangle \in \mathcal{H}^{2^n}$ "live" in the typical subspace.

QUANTUM DISCRETE MEMORYLESS SOURCE

Typical Subspace

- Typical states $|\Psi_{\boldsymbol{x}}\rangle \in \mathcal{H}^{2^n}$ "live" in the typical subspace.
- Typical subspace Λ_n of $\mathcal{H}^{2^n}$:

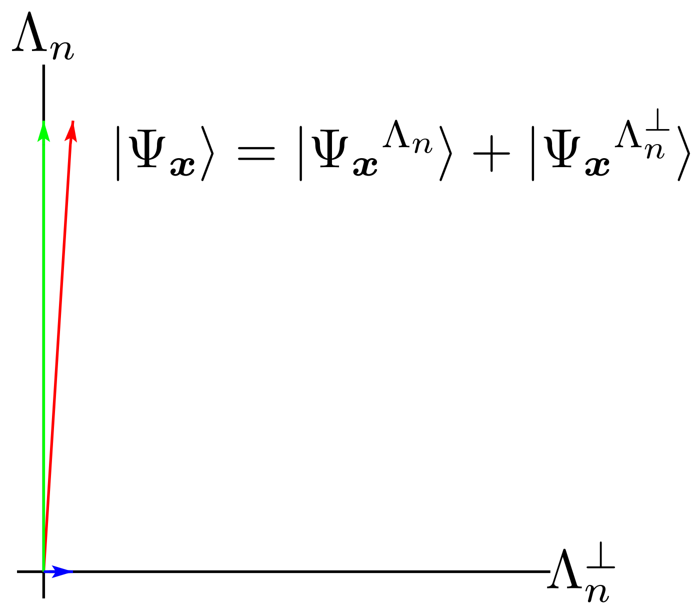


QUANTUM DISCRETE MEMORYLESS SOURCE

Typical Subspace

- Typical states $|\Psi_{\boldsymbol{x}}\rangle \in \mathcal{H}^{2^n}$ "live" in the typical subspace.

- Typical subspace Λ_n of $\mathcal{H}^{2^n}$:



- The dimension of Λ_n is approximately $2^{nS(\rho)}$.

Code $\mathcal{C} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\} \subset \mathcal{X}^n$

- F is the $d^n \times M$ matrix whose columns are $|\psi_{\boldsymbol{x_i}}\rangle / \sqrt{M}$. Thus

$$FF^\dagger = \frac{1}{M} \sum_{i=1}^{M} |\Psi_{\boldsymbol{x}_i}\rangle\langle\Psi_{\boldsymbol{x}_i}|.$$

Code $\mathcal{C} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\} \subset \mathcal{X}^n$

- F is the $d^n \times M$ matrix whose columns are $|\psi_{\boldsymbol{x_i}}\rangle/\sqrt{M}$. Thus

$$FF^\dagger = \frac{1}{M}\sum_{i=1}^{M} |\Psi_{\boldsymbol{x}_i}\rangle\langle\Psi_{\boldsymbol{x}_i}|.$$

- $|\Psi_{\boldsymbol{x}_1}\rangle, \ldots, |\Psi_{\boldsymbol{x}_M}\rangle$ span $\mathcal{U}$, an r-dimensional subspace of $\mathcal{H}^{d^n}$.

Code $\mathcal{C} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\} \subset \mathcal{X}^n$

- F is the $d^n \times M$ matrix whose columns are $|\psi_{\boldsymbol{x_i}}\rangle / \sqrt{M}$. Thus

$$FF^\dagger = \frac{1}{M} \sum_{i=1}^{M} |\Psi_{\boldsymbol{x}_i}\rangle\langle\Psi_{\boldsymbol{x}_i}|.$$

- $|\Psi_{\boldsymbol{x}_1}\rangle, \ldots, |\Psi_{\boldsymbol{x}_M}\rangle$ span $\mathcal{U}$, an r-dimensional subspace of $\mathcal{H}^{d^n}$.

- Perform the SVD of F and define a scaled projection on $\mathcal{U}$:

$$F = \sum_{k=1}^{r} \sqrt{\lambda_k} |u_k\rangle\langle v_k|, \quad FF^\dagger = \sum_{k=1}^{r} \lambda_k |u_k\rangle\langle u_k|, \quad P_{\mathcal{U}} = \sum_{k=1}^{r} \frac{1}{r} |u_k\rangle\langle u_k|$$

Code $\mathcal{C} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\} \subset \mathcal{X}^n$

- F is the $d^n \times M$ matrix whose columns are $|\psi_{\boldsymbol{x_i}}\rangle/\sqrt{M}$. Thus

$$FF^\dagger = \frac{1}{M}\sum_{i=1}^{M} |\Psi_{\boldsymbol{x}_i}\rangle\langle\Psi_{\boldsymbol{x}_i}|.$$

- $|\Psi_{\boldsymbol{x}_1}\rangle, \ldots, |\Psi_{\boldsymbol{x}_M}\rangle$ span $\mathcal{U}$, an r-dimensional subspace of $\mathcal{H}^{d^n}$.

- Perform the SVD of F and define a scaled projection on $\mathcal{U}$:

$$F = \sum_{k=1}^{r} \sqrt{\lambda_k}|u_k\rangle\langle v_k|, \quad FF^\dagger = \sum_{k=1}^{r} \lambda_k|u_k\rangle\langle u_k|, \quad P_{\mathcal{U}} = \sum_{k=1}^{r} \frac{1}{r}|u_k\rangle\langle u_k|$$

- How far is $FF^\dagger$ from $P_{\mathcal{U}}$?

DISTANCES BETWEEN DENSITY MATRICES

ρ and σ

- Trace distance:

$$D(\sigma, \omega) = \frac{1}{2} \operatorname{Tr} |\sigma - \omega|,$$

$|A|$ denotes the positive square root of $A^\dagger A$.

DISTANCES BETWEEN DENSITY MATRICES ρ and σ

- Trace distance:

$$D(\sigma,\omega) = \frac{1}{2}\,\mathrm{Tr}\,|\sigma - \omega|,$$

$|A|$ denotes the positive square root of $A^\dagger A$.

- Uhlman Fidelity:

$$F(\sigma,\omega) = \left\{\mathrm{Tr}\left[(\sqrt{\sigma}\omega\sqrt{\sigma})^{1/2}\right]\right\}^2.$$

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

DISTANCES BETWEEN DENSITY MATRICES
ρ and σ

- Trace distance:

$$D(\sigma,\omega) = \frac{1}{2}\,\mathrm{Tr}\,|\sigma - \omega|,$$

  $|A|$ denotes the positive square root of $A^\dagger A$.

- Uhlman Fidelity:

$$F(\sigma,\omega) = \left\{\mathrm{Tr}\big[(\sqrt{\sigma}\omega\sqrt{\sigma})^{1/2}\big]\right\}^2.$$

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

- Frobenius (Hilbert-Schmidt)?

DISTANCES BETWEEN PD'S

An Example

- $\mathcal{A}_N = \{a_1, \ldots, a_N\}$
- $N = 2^K$ and $n = 2^k$, with $k/K = c < 1$.

DISTANCES BETWEEN PD'S

An Example

- $\mathcal{A}_N = \{a_1, \ldots, a_N\}$
- $N = 2^K$ and $n = 2^k$, with $k/K = c < 1$.
- Distributions P and Q:

$$P(a_i) = \frac{1}{N}$$

DISTANCES BETWEEN PD'S

An Example

- $\mathcal{A}_N = \{a_1, \ldots, a_N\}$
- $N = 2^K$ and $n = 2^k$, with $k/K = c < 1$.
- Distributions P and Q:

$$P(a_i) = \frac{1}{N} \text{ and } Q(a_i) = \begin{cases} 1/n, & 1 \leq i \leq n \\ 0 & n+1 \leq i \leq N \end{cases}$$

DISTANCES BETWEEN PD'S

An Example

- $\mathcal{A}_N = \{a_1, \ldots, a_N\}$

- $N = 2^K$ and $n = 2^k$, with $k/K = c < 1$.

- Distributions P and Q:

$$P(a_i) = \frac{1}{N} \text{ and } Q(a_i) = \begin{cases} 1/n, & 1 \leq i \leq n \\ 0 & n+1 \leq i \leq N \end{cases}$$

- $P(\{a_{n+1}, \ldots, a_N\}) \to 1$ as $k, K \to \infty$.

DISTANCES BETWEEN PD'S

An Example

- $\mathcal{A}_N = \{a_1, \ldots, a_N\}$
- $N = 2^K$ and $n = 2^k$, with $k/K = c < 1$.
- Distributions P and Q:

$$P(a_i) = \frac{1}{N} \text{ and } Q(a_i) = \begin{cases} 1/n, & 1 \leq i \leq n \\ 0 & n+1 \leq i \leq N \end{cases}$$

- $P(\{a_{n+1}, \ldots, a_N\}) \to 1$ as $k, K \to \infty$.
- $\frac{1}{2}\sum_i |P(a_i) - Q(a_i)| \to 1$ and $\sum_i |P(a_i) - Q(a_i)|^2 \to 0$.

A BASIS FOR Λ_n

- How far is $FF^\dagger$ from $P_{\mathcal{U}}$?

A BASIS FOR Λ_n

- How far is $FF^\dagger$ from $P_{\mathcal{U}}$?

$$\Big[\sum_{k=1}^{r} |\lambda_k - \frac{1}{r}|\Big]^2 \leq r\sum_{k=1}^{r}\Big(\lambda_k - \frac{1}{r}\Big) = r\Big(\sum_{k=1}^{r}\lambda_k^2\Big) - 1$$

A BASIS FOR Λ_n

- How far is $FF^\dagger$ from $P_{\mathcal{U}}$?

$$\Big[\sum_{k=1}^{r} |\lambda_k - \frac{1}{r}|\Big]^2 \leq r\sum_{k=1}^{r}\Big(\lambda_k - \frac{1}{r}\Big) = r\Big(\sum_{k=1}^{r}\lambda_k^2\Big) - 1$$

$$= \frac{r}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} |\langle\Psi_{\boldsymbol{x}_i}|\Psi_{\boldsymbol{x}_j}\rangle|^2 - 1$$

A BASIS FOR Λ_n

- How far is $FF^\dagger$ from $P_{\mathcal{U}}$?

$$\Big[\sum_{k=1}^{r} |\lambda_k - \frac{1}{r}|\Big]^2 \leq r \sum_{k=1}^{r}\Big(\lambda_k - \frac{1}{r}\Big) = r\Big(\sum_{k=1}^{r} \lambda_k^2\Big) - 1$$

$$= \frac{r}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} |\langle \Psi_{\boldsymbol{x}_i} | \Psi_{\boldsymbol{x}_j} \rangle|^2 - 1$$

$$\leq \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} |\langle \Psi_{\boldsymbol{x}_i} | \Psi_{\boldsymbol{x}_j} \rangle|^2$$

A BASIS FOR Λ_n

- How far is $FF^\dagger$ from $P_{\mathcal{U}}$?

$$\Big[\sum_{k=1}^{r} |\lambda_k - \frac{1}{r}|\Big]^2 \leq r \sum_{k=1}^{r} \Big(\lambda_k - \frac{1}{r}\Big) = r\Big(\sum_{k=1}^{r} \lambda_k^2\Big) - 1$$

$$= \frac{r}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} |\langle \Psi_{\boldsymbol{x}_i} | \Psi_{\boldsymbol{x}_j} \rangle|^2 - 1$$

$$\leq \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} |\langle \Psi_{\boldsymbol{x}_i} | \Psi_{\boldsymbol{x}_j} \rangle|^2$$

- Use the random coding argument!

A BASIS FOR Λ_n

The Random Coding Argument

- Averaging over all codes:

$$E\Big\{\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}|\langle\Psi_{\boldsymbol{x}_i}|\Psi_{\boldsymbol{x}_j}\rangle|^2\Big\} = M(M-1)\,\mathrm{Tr}(\rho^{\otimes n}\cdot\rho^{\otimes n})$$

A BASIS FOR Λ_n
The Random Coding Argument

- Averaging over all codes:

$$E\Big\{\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}|\langle\Psi_{\boldsymbol{x}_i}|\Psi_{\boldsymbol{x}_j}\rangle|^2\Big\} = M(M-1)\,\mathrm{Tr}(\rho^{\otimes n}\cdot\rho^{\otimes n}) \quad \Rightarrow$$

- There exists a code $\mathcal{C}$ with M codewords s.t.

$$\Big[\sum_{k=1}^{r}\Big|\lambda_k-\frac{1}{r}\Big|\Big]^2 \leq M\,\mathrm{Tr}(\rho^{\otimes n}\cdot\rho^{\otimes n})$$

A BASIS FOR Λ_n
The Random Coding Argument

- Averaging over all codes:

$$E\Big\{\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}|\langle\Psi_{\boldsymbol{x}_i}|\Psi_{\boldsymbol{x}_j}\rangle|^2\Big\} = M(M-1)\,\mathrm{Tr}(\rho^{\otimes n}\cdot\rho^{\otimes n}) \quad \Rightarrow$$

- There exists a code $\mathcal{C}$ with M codewords s.t.

$$\Big[\sum_{k=1}^{r}\big|\lambda_k-\frac{1}{r}\big|\Big]^2 \leq M\,\mathrm{Tr}(\rho^{\otimes n}\cdot\rho^{\otimes n}) \quad \Rightarrow$$

- There exists a code $\mathcal{C}$, $|\mathcal{C}| = 2^{nR}$ s.t. on Λ_n

$$\Big[\sum_{k=1}^{r}\big|\lambda_k-\frac{1}{r}\big|\Big]^2 \leq 2^{-n(S(\rho)-\varepsilon_n-R)}$$

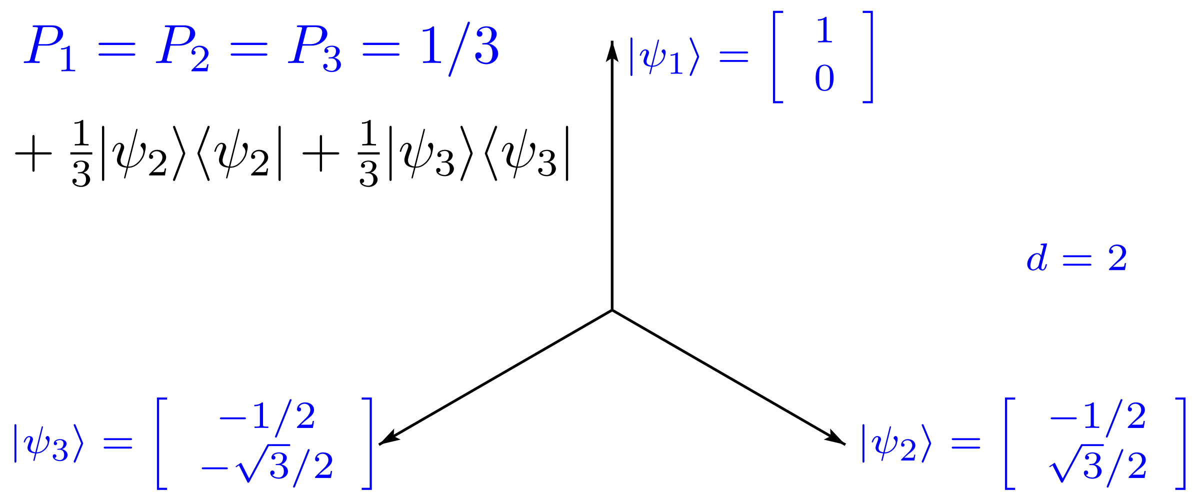
COMBINATORICS AND GEOMETRY

The MB Example

$$\mathcal{X} = \{1, 2, 3\} \quad P_1 = P_2 = P_3 = 1/3$$

$$\rho = \tfrac{1}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{1}{3}|\psi_3\rangle\langle\psi_3|$$
$$= \tfrac{1}{2}I$$

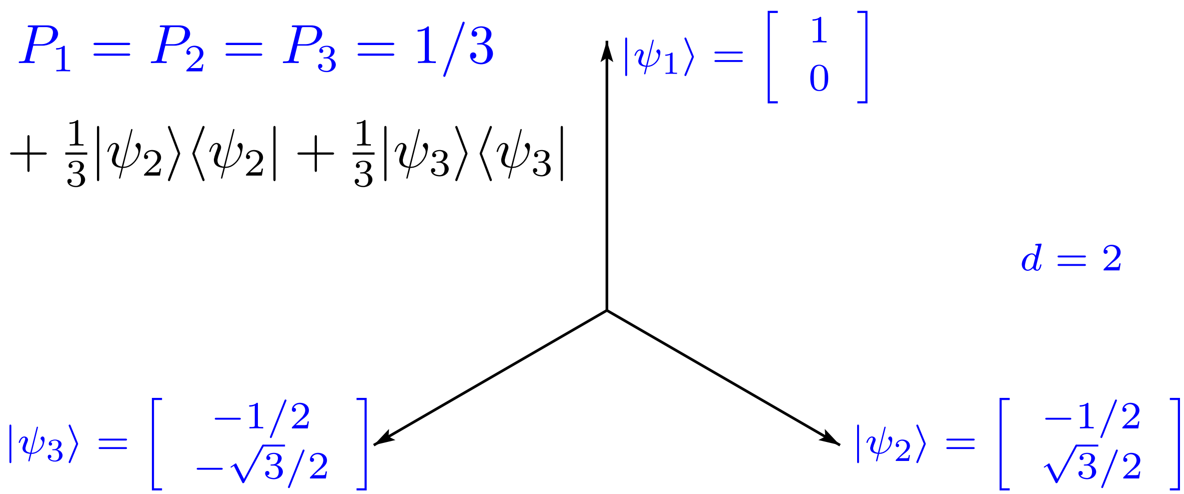
$$S(\rho) = 1$$



COMBINATORICS AND GEOMETRY

The MB Example

$$
\begin{aligned}
\mathcal{X} &= \{1,2,3\} \quad P_1 = P_2 = P_3 = 1/3 \\
\rho &= \tfrac{1}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{1}{3}|\psi_3\rangle\langle\psi_3| \\
&= \tfrac{1}{2}I \\
S(\rho) &= 1
\end{aligned}
$$



- There are 3^n typical vectors forming a frame in H^{2^n}.

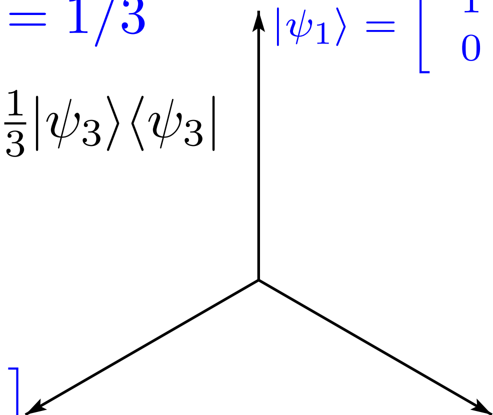
COMBINATORICS AND GEOMETRY

The MB Example

$$\mathcal{X} = \{1,2,3\} \quad P_1 = P_2 = P_3 = 1/3$$

$$\begin{aligned} \rho &= \tfrac{1}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{1}{3}|\psi_3\rangle\langle\psi_3| \\ &= \tfrac{1}{2}I \end{aligned}$$

$$S(\rho) = 1$$



- There are 3^n typical vectors forming a frame in H^{2^n}.
- About 2^n of those vectors form a basis of H^{2^n}.
- Which ones?