# Quantum Information Theory

Yonina Eldar

6.962, Week 12

May 5, 2001

## Physics Of Information

Information is encoded in a state of a physical system

- Classical information is represented by two well separated values of a continuous parameter such as voltage.
- Quantum information is encoded in the state of a quantum system such as the electron spin or the polarization of a photon.

# Motivation

- Theoretically intriguing

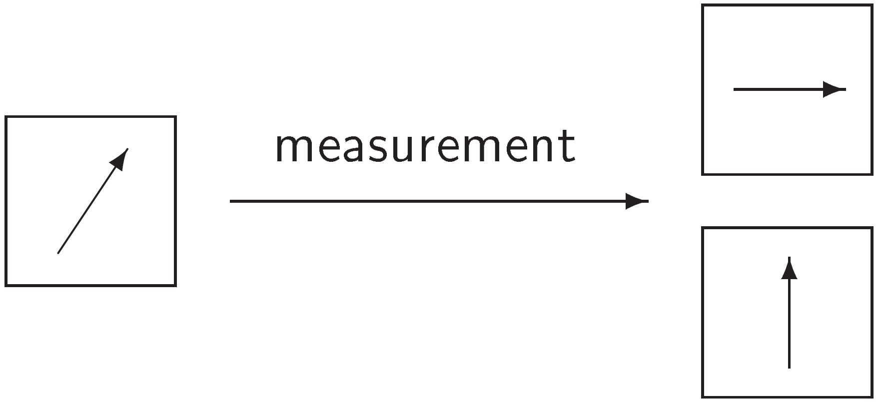
  *"Anyone who can contemplate quantum mechanics without getting dizzy hasn't properly understood it"*, Niels Bohr

- Engineering perspective
  - New ways of encoding classical information
  - Quantum information is the basis for understanding and implementing

    Quantum cryptography

    Quantum computing

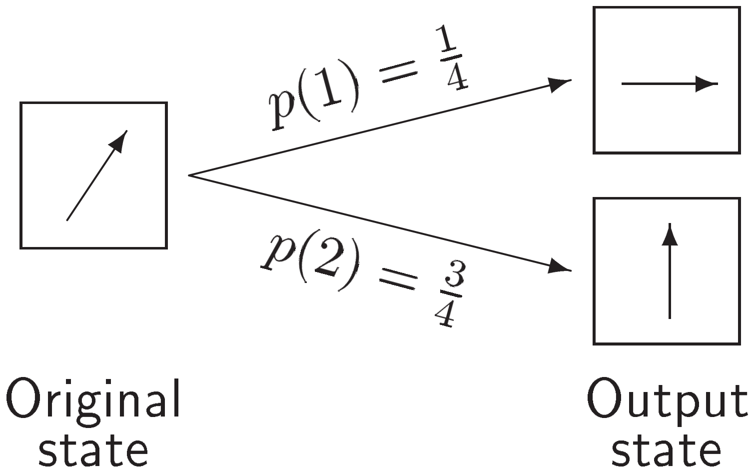
# What Is A Quantum State?

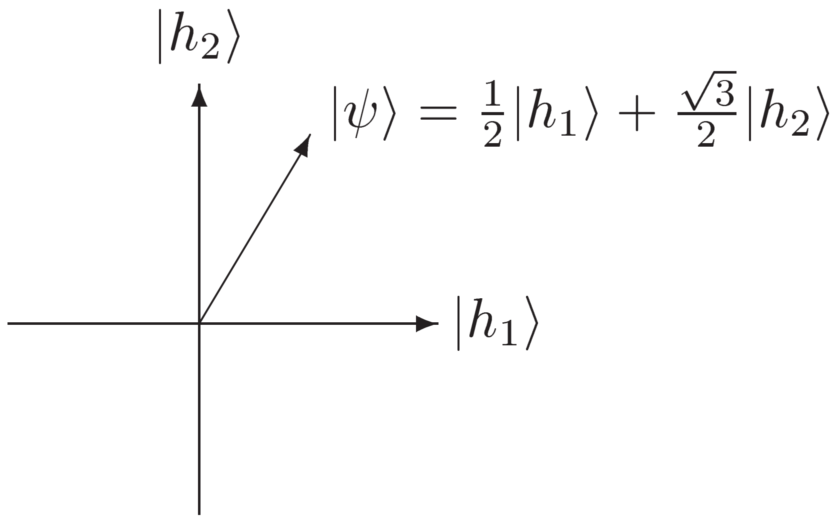



- A state is a normalized vector in an inner-product space.
- Determines the probabilities of the measurement outcomes.

## Quantum Measurement (Simplified)

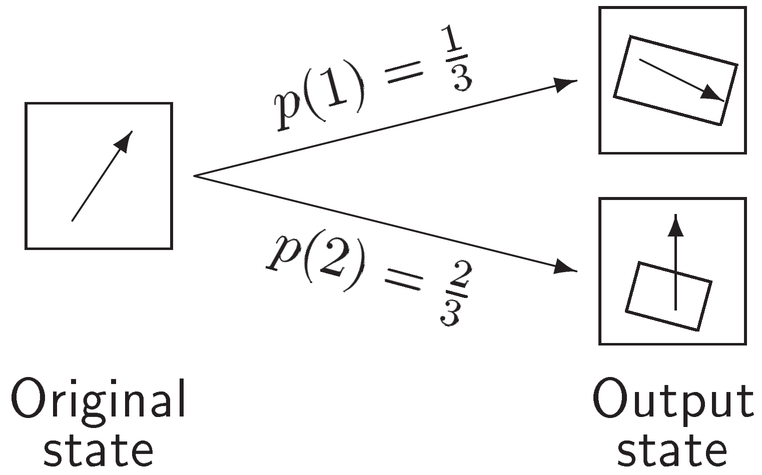
- A measurement selects one of a set of orthonormal vectors $|h_k\rangle$
- The probability of obtaining outcome $k$ given the state $|\psi\rangle$ is $|\langle\psi|h_k\rangle|^2$
- The state after the measurement collapses onto a state consistent with the measurement outcome

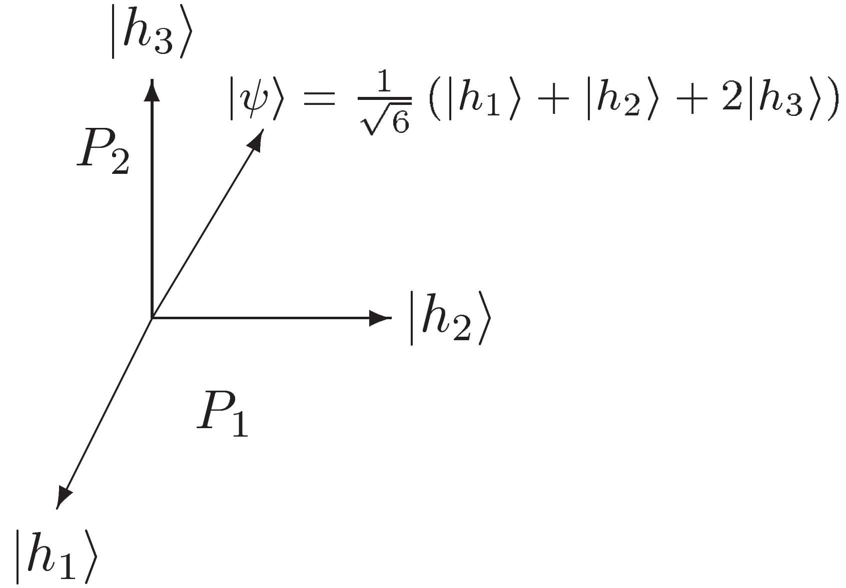


# Quantum Measurement (More General)

- A measurement selects a projector out of a complete set of orthogonal projectors $P_k = \sum_m |h_m\rangle\langle h_m|$
- The prob. of outcome $k$ given the state $|\psi\rangle$ is $\langle\psi|P_k\psi\rangle = \sum_m |\langle\psi|h_m\rangle|^2$
- The state after the measurement collapses onto a state consistent with the measurement outcome

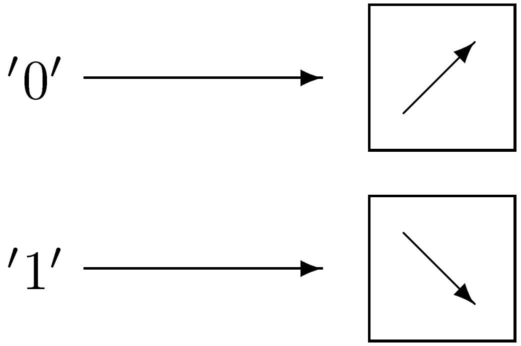


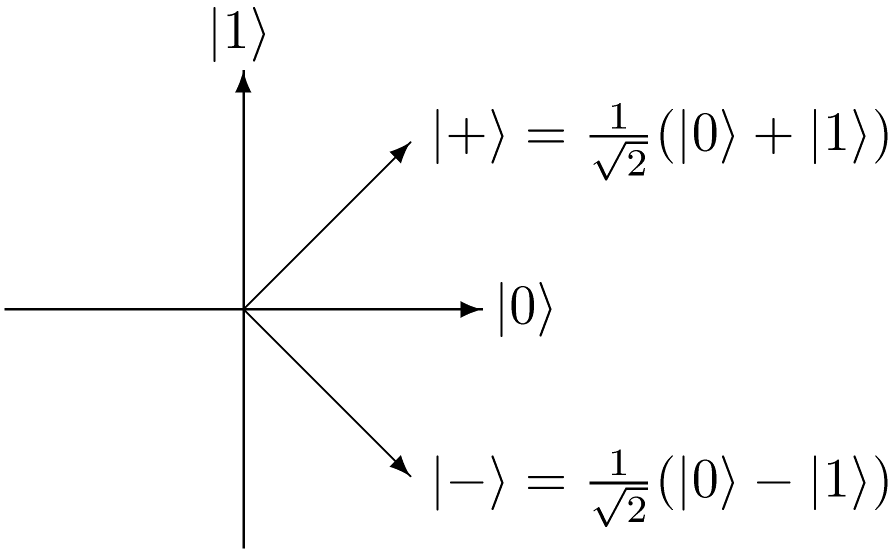



## Fundamental Tasks

- Using quantum systems to send **classical** information
  - How do we detect the information?
  - How many bits can we send?
- Using quantum systems to send **quantum** information
  - Can we compress the quantum state (without knowing it)?

# Classical Information via Quantum States



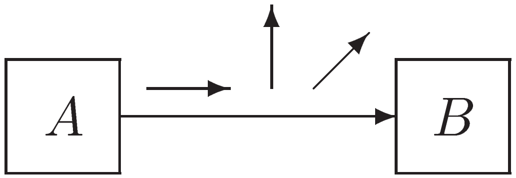

Measure $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$:

$$p(0|+) = |\langle +|0\rangle|^2 = 1/2, \; p(1|+) = |\langle +|1\rangle|^2 = 1/2.$$
$$p(0|-) = |\langle -|0\rangle|^2 = 1/2, \; p(1|-) = |\langle -|1\rangle|^2 = 1/2.$$

Measure $P_+ = |+\rangle\langle+|$ and $P_- = |-\rangle\langle-|$:

$$p(+|+) = |\langle+|+\rangle|^2 = 1,\ p(-|+) = |\langle+|-\rangle|^2 = 0.$$
$$p(-|-) = |\langle-|-\rangle|^2 = 1,\ p(+|-) = |\langle-|+\rangle|^2 = 0.$$

## Quantum Detection Problem




- $A$ sends one of the states $|\psi_1\rangle, \ldots, |\psi_m\rangle$ to $B$
- $B$ performs a measurement to detect which state was received
- $B$'s problem is to construct a 'good' measurement

# Maximize Probability of Detection

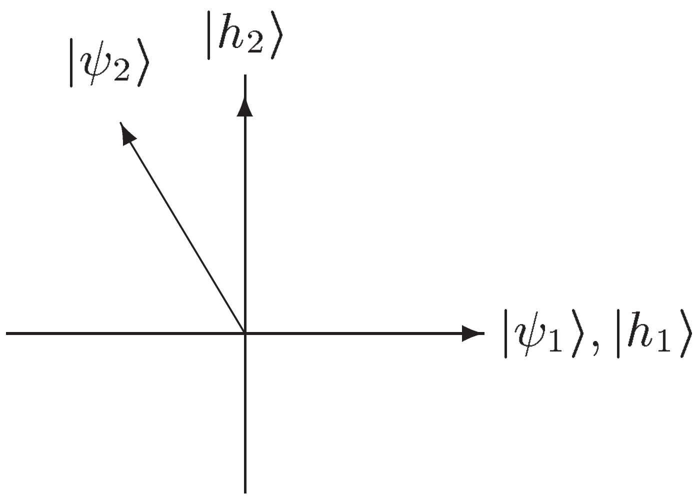
$$P_d = \sum_{k=1}^{m} p(\psi_k) p(h_k|\psi_k) = \frac{1}{m} \sum_{k=1}^{m} |\langle \psi_k | h_k \rangle|^2$$

**Difficult and unsolved problem**

- Iterative procedures
- *Ad-hoc* measurements

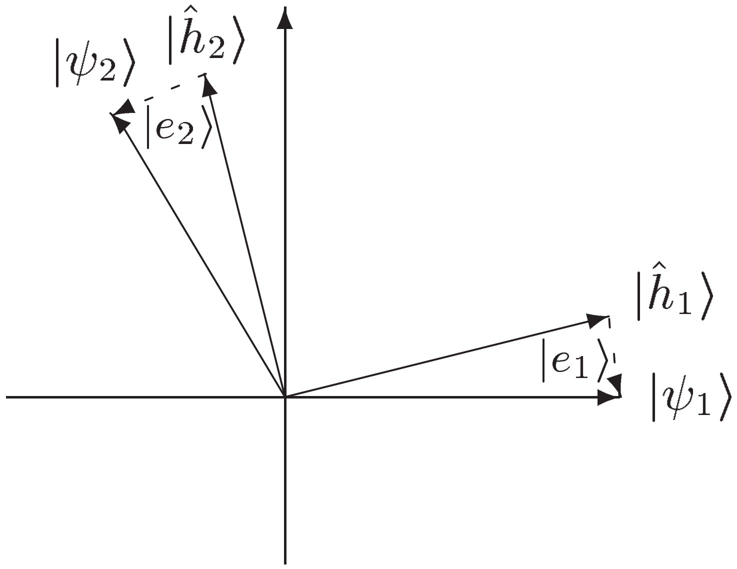
$$|\psi_1\rangle = [1\ 0]^T,\ |\psi_2\rangle = [-1/2\ \sqrt{3}/2]^T$$

$$P_d = \tfrac{1}{2}(1 + \tfrac{3}{4}) = 0.875$$



# Least-Squares Measurement

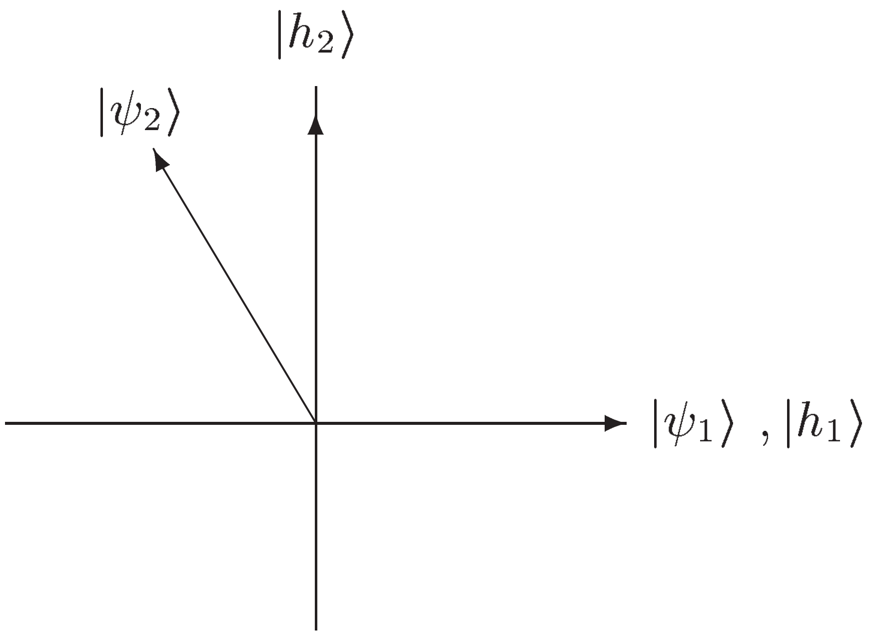
If you can't solve it – change it!




$$\widehat{H} = \Psi(\Psi^*\Psi)^{-1/2}$$

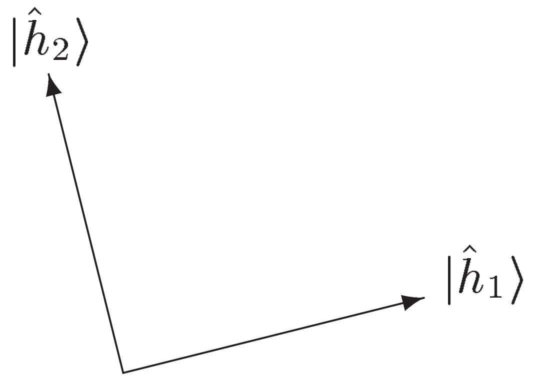
- Maximizes $P_d$ for state sets with certain symmetry properties.
- Asymptotically maximizes $P_d$.
- Approximately maximizes $P_d$ for nearly orthonormal states.

# Example




$$|\psi_1\rangle = [1\ 0]^T,\ |\psi_2\rangle = [-1/2\ \sqrt{3}/2]^T,\ |\langle\psi_1|\psi_2\rangle| = 1/2$$

$P_d$ using $|h_1\rangle, |h_2\rangle$: $0.875$




$P_d$ using LSM: $0.933$

## The Quantum Bit - "Qubit"

- A qubit is a normalized vector in a $2$-D space

$$|\psi\rangle = a|0\rangle + b|1\rangle, \quad |a|^2 + |b|^2 = 1$$

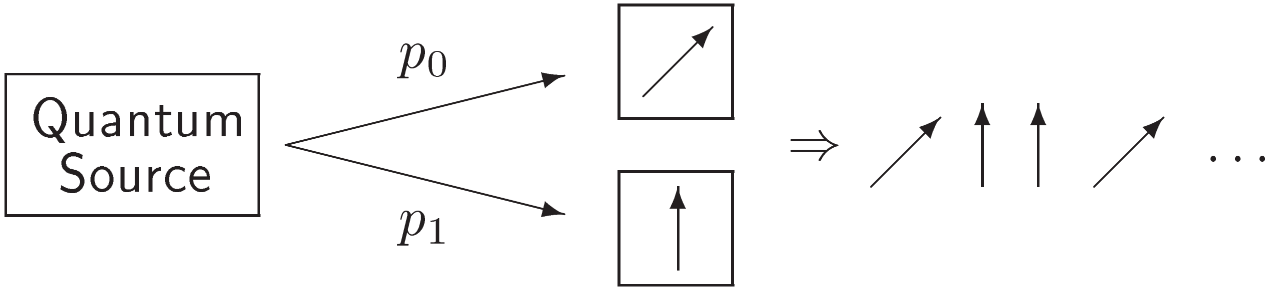
where $|0\rangle$ and $|1\rangle$ represent any orthonormal basis in this space.

- The state of two qubits is a normalized vector in a $4$-D space.

$\Rightarrow$ A pair of qubits can exist in superpositions of four bases states,

$|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle$.

# Quantum Source




- Prepare state $|\psi_i\rangle$ with probability $p_i$. Measure $P_0$, $P_1$:

$$p(0) = \sum_i p_i |\langle\psi_i|0\rangle|^2 = \sum_i p_i \mathsf{Tr}(|\psi_i\rangle\langle\psi_i|0\rangle\langle 0|) = \mathsf{Tr}(\rho P_0)$$

  where $\rho$ is the density matrix defined as $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$.

- State of $n$ independent outputs $\rho^n = \rho \otimes \rho \otimes \ldots \otimes \rho$.

**Properties:** $\rho$ is Hermitian, non-negative definite and Tr$(\rho) = 1$.

## Classical Noiseless Coding Theorem

**Objective:**

Represent a sequence of $n$ independent letters chosen from an ensemble $A = \{p_a, a\}$ by a sequence of bits with minimal length.

**For large $n$:**

- In a "typical" sequence of length $n$ the letter $a$ will appear $\sim np_a$ times.
- The probability of a "typical" sequence is approximately

$$\prod_a p_a^{np_a} = \prod_a 2^{np_a \log p_a} = 2^{n \sum_a p_a \log p_a} = 2^{-nH(A)}$$

where $H(A) = -\sum_a p_a \log p_a$ is the Shannon entropy of the ensemble.

## Typical Sequences

Divide the set of $n$-sequences into two:

- A subset of "typical" sequences where each sequences has probability $\sim 2^{-nH(A)}$ and contains $\sim 2^{nH(A)}$ elements.
- A subset of "atypical" sequences with very small total probability.

Coding strategy:

- Code typical sequences by binary strings of length $\sim nH(A)$.
- Code atypical sequences by $0000\ldots00$.

# Typical Probabilities

Typical prob.

$$0011\cdots \rightarrow \quad p_0 p_0 p_1 p_1 \cdots \sim 2^{-nH}$$
$$1010\cdots \rightarrow \quad p_1 p_0 p_1 p_0 \cdots \sim 2^{-nH}$$
$$\vdots$$
$$\sum \sim 1$$

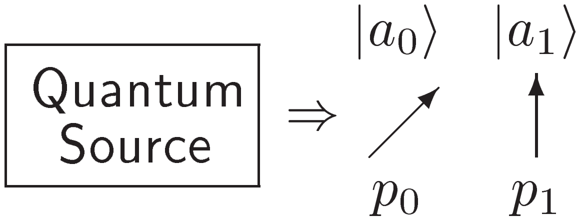
Atypical prob.

$$p_0 p_0 p_0 p_0 \cdots \quad \leftarrow 0000\cdots$$
$$p_1 p_1 p_1 p_1 \cdots \quad \leftarrow 1111\cdots$$
$$\vdots$$
$$\sum \sim 0$$

Divide the set of $n$-sequences of probabilities into two:

- A subset of $\sim 2^{nH(A)}$ "typical" probabilities with value $\sim 2^{-nH(A)}$
- A subset of "atypical" probabilities with very small sum

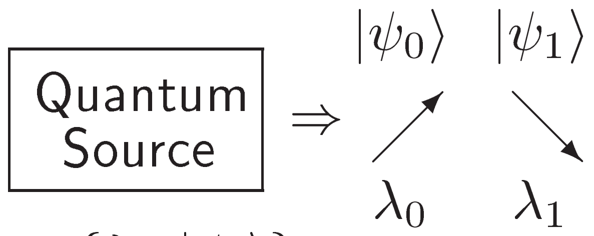
# von Neumann Entropy




$\varepsilon = \{p_i, |a_i\rangle\}$

$$\rho = \sum_i p_i |a_i\rangle\langle a_i|$$

**Measure of entropy:** $H(\{p_i\})$ ?




$\varepsilon = \{\lambda_i, |\psi_i\rangle\}$

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$$

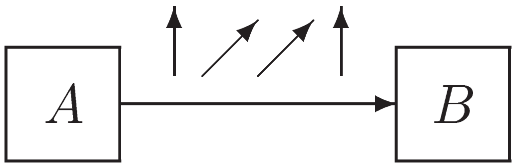
**von Neumann entropy:** $S(\rho) = H(\{\lambda_i\}) = -\mathsf{Tr}(\rho \log \rho)$

$S(\rho) \leq H(\{p_i\})$ with equality iff the ensemble states are orthogonal

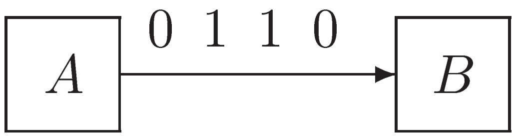
# Quantum Coding

**Objective:**

Represent a sequence of $n$ independent qubits chosen from an ensemble $\varepsilon = \{p_i, |a_i\rangle\}$ by a sequence of qubits with minimal length.




$A$ can prepare a message by drawing a sequence of numbers from $\{p_i, i\}$ and then send this sequence to $B$ using $H(\{p_i\})$ orthogonal qubits.




**Can we do better?**

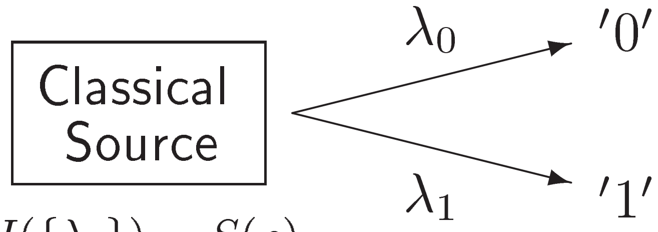
## Proof of Quantum Coding

**Observations:**

- The qubit state is $\rho = \sum_i p_i |a_i\rangle\langle a_i|$
- $\rho$ has eigenvalues satisfying $\lambda_0, \lambda_1 \geq 0$ and $\lambda_0 + \lambda_1 = 1$.
- The state of $n$ independent qubits is $\rho^n = \rho \otimes \ldots \otimes \rho$.
- $\rho^n$ has $2^n$ orthonormal eigenvectors and $2^n$ eigenvalues of the form $\lambda_{k_1} \lambda_{k_2} \ldots \lambda_{k_n},\ k_i = 0, 1$.

# Typical Subspaces

**Key idea:**




$H(\{\lambda_i\}) = S(\rho)$

$$\Rightarrow \quad \begin{array}{llll} 0 & 1 & 1 & 0\cdots \\ \lambda_0 & \lambda_1 & \lambda_1 & \lambda_0\cdots \end{array}$$

For large $n$ the eigenvalues $\lambda_{k_1}\lambda_{k_2}\ldots\lambda_{k_n}$ of $\rho^n$ satisfy:

- $\sim 2^{nS(\rho)}$ "typical" eigenvalues with value $\sim 2^{-nS(\rho)}$
- The sum of the "atypical" eigenvalues is very small

**Typical subspace $\Lambda$:**

Spanned by the eigenvectors with "typical" eigenvalues $\Rightarrow \dim(\Lambda) \sim 2^{nS(\rho)}$

## Coding Strategy

Given a sequence of $n$ qubits:

- Perform the measurement $P_\Lambda$, $P_{\Lambda_\perp}$.

  The probability of projecting onto $\Lambda$ is $\text{Tr}(\rho^n P_\Lambda) = \sum_{\text{typical}} \lambda \sim 1$.

- The projected state can be encoded using $\sim nS(\rho)$ qubits.
- If we projected onto $\Lambda^\perp$ then encode the message as $|0\rangle|0\rangle \ldots |0\rangle$.

The decoded state is very close to the original state with high probability

## Conclusion

- Encoding classical information in quantum states
  - Constructing a "good" measurement.
  - Can convey $S(\rho)$ bits.
- Quantum data compression
  - Clever use of (classical) typical sequences.
  - $S(\rho)$ is the compression limit.
- von Neumann entropy
  - Quantifies both the number of qubits needed to encode an ensemble state and the number of bits that can be conveyed by the ensemble.