QUANTUM DATA COMPRESSION

Emina Soljanin

Bell Labs

MSRI, February 2002

LOSSLESS CLASSICAL DATA COMPRESSION

The Source

- 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 $\mathcal{X} = \{H, T\}$.

LOSSLESS CLASSICAL DATA COMPRESSION

Sequences and Compression Fidelity

- Source sequence $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ is in $\mathcal{X}^n$.
- Sequence $\boldsymbol{x}$ is compressed and then reproduced as $\hat{\boldsymbol{x}}$.
- Fidelity between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$:

$$F(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \begin{cases} 1 & \text{if } \boldsymbol{x} = \hat{\boldsymbol{x}} \\ 0 & \text{otherwise.} \end{cases}$$

- Compression fidelity

$$\bar{F} = \sum_{\boldsymbol{x} \in \mathcal{X}^n} P(\boldsymbol{x}) F(\boldsymbol{x}, \hat{\boldsymbol{x}})$$

should approach 1 as $n \to \infty$.

LOSSLESS CLASSICAL DATA COMPRESSION

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:

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

LOSSLESS CLASSICAL DATA COMPRESSION

Shannon Entropy and Compression Rate

- Compress with perfect fidelity only the typical sequences.
- Compression fidelity: $\bar{F} \geq P^n(\mathsf{T}^n_{P,\delta}) \geq 1 - \epsilon_n$.
- Set $\mathsf{T}^n_{P,\delta}$ contains approximately $2^{nH(P)}$ sequences.
- Shannon entropy: $H(P) = -\sum_{a \in \mathcal{X}} P(a) \log P(a)$.
- Instead of $n \log |\mathcal{X}|$, we can transmit only $nH(P)$ bits.
- Compression rate?

LOSSLESS QUANTUM DATA COMPRESSION
Discrete Memoryless Source

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

$$|\leftrightarrow\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad |\updownarrow\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- A qubit is a vector in $\mathcal{H}_2$:

$$|\psi\rangle = \alpha|\leftrightarrow\rangle + \beta|\updownarrow\rangle$$

- Example: $\mathcal{X} = \{0, 1, 2, 3\}$,

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

LOSSLESS QUANTUM DATA COMPRESSION

Vector Sequences

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

$$|\leftrightarrow\rangle \otimes |\leftrightarrow\rangle \quad |\leftrightarrow\rangle \otimes |\updownarrow\rangle \quad |\updownarrow\rangle \otimes |\leftrightarrow\rangle \quad |\updownarrow\rangle \otimes |\updownarrow\rangle$$

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

- Source vector-sequence (state):

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

- vector in $\mathcal{H}^{2^n}$ v.s. n qubits

LOSSLESS QUANTUM DATA COMPRESSION

Sequences and Fidelity

- Source vector-sequence $|\Psi_{\boldsymbol{x}}\rangle$ is in $\mathcal{H}^{2^n}$, ($\boldsymbol{x} \in \mathcal{X}^n$)

- Vector $|\Psi_{\boldsymbol{x}}\rangle$ is compressed and then reproduced as $|\widehat{\Psi_{\boldsymbol{x}}}\rangle$.

- Fidelity between $|\Psi_{\boldsymbol{x}}\rangle$ and $|\widehat{\Psi_{\boldsymbol{x}}}\rangle$:

$$F(|\Psi_{\boldsymbol{x}}\rangle, |\widehat{\Psi_{\boldsymbol{x}}}\rangle) = |\langle\Psi_{\boldsymbol{x}}|\widehat{\Psi_{\boldsymbol{x}}}\rangle|^2$$

- Compression fidelity

$$\bar{F} = \sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) F(|\Psi_{\boldsymbol{x}}\rangle, |\widehat{\Psi_{\boldsymbol{x}}}\rangle)$$

  should approach 1 as $n \to \infty$.

LOSSLESS QUANTUM DATA COMPRESSION

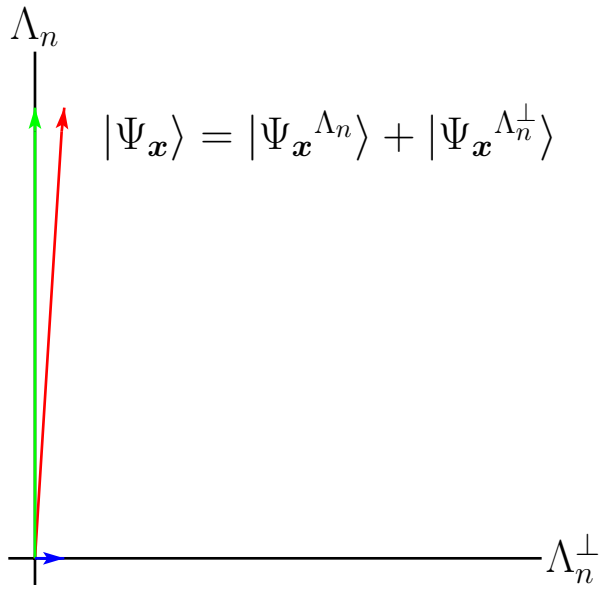
Typical States and Visible Compression

- The encoder Alice knows sequence $\boldsymbol{x}$.
- She can compress with perfect fidelity the typical states.
- Instead of n qubits, she can transmit $nH(P)$ bits.
- The decoder Bob prepares $|\Psi_{\boldsymbol{x}}\rangle$ as $|\widehat{\Psi_{\boldsymbol{x}}}\rangle$ for typical $\boldsymbol{x}$.
- Is this strategy optimal?

LOSSLESS QUANTUM DATA COMPRESSION

Typical Subspace and Blind Compression

- Source vector-sequence $|\Psi_{\boldsymbol{x}}\rangle$ is in $\mathcal{H}^{2^n}$.
- Typical subspace Λ_n of $\mathcal{H}^{2^n}$:



- Compression fidelity

$$\bar{F} = \sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) |\langle \Psi_{\boldsymbol{x}} | \Psi_{\boldsymbol{x}}{}^{\Lambda_n} \rangle|^2.$$

- vector in $\mathcal{H}^{2^k}$ v.s. k qubits

LOSSLESS QUANTUM DATA COMPRESSION

Von Neumann Entropy and Compression Rate

- 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.

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

LOSSLESS QUANTUM DATA COMPRESSION

The Typical Subspace Λ_n

- We represent the source density matrix

$$\rho = \sum_{a \in \mathcal{X}} P(a) |\psi_a\rangle\langle\psi_a|$$

in terms of its eigenvectors and eigenvalues:

$$\rho = \lambda_0 |\varphi_0\rangle\langle\varphi_0| + \lambda_1 |\varphi_1\rangle\langle\varphi_1|.$$

- Note that $\lambda = \{\lambda_0, \lambda_1\}$ is a PD on $\{0, 1\}$ and $\langle\varphi_0|\varphi_1\rangle = 0$.

- Consider

$$\rho^{\otimes n} = \sum_{\boldsymbol{x} \in \mathcal{X}^n} P(\boldsymbol{x}) |\Psi_{\boldsymbol{x}}\rangle\langle\Psi_{\boldsymbol{x}}| = \sum_{\boldsymbol{z} \in \{0,1\}^n} \lambda(\boldsymbol{z}) |\Phi_{\boldsymbol{z}}\rangle\langle\Phi_{\boldsymbol{z}}|.$$

LOSSLESS QUANTUM DATA COMPRESSION

The Typical Subspace Λ_n

- Consider all sequences $\boldsymbol{z} \in \{0,1\}^n$ for which

$$\left|\frac{1}{n}N(0|\boldsymbol{z}) - \lambda_0\right| \leq \delta \text{ and } \left|\frac{1}{n}N(1|\boldsymbol{z}) - \lambda_1\right| \leq \delta.$$

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

- Λ_n is the subspace spanned by $|\Phi_{\boldsymbol{z}}\rangle$, $\boldsymbol{z} \in \mathsf{T}^n_{\lambda,\delta}$.

- $\Pi = \sum_{\boldsymbol{z}\in\mathsf{T}^n_{\lambda,\delta}} |\Phi_{\boldsymbol{z}}\rangle\langle\Phi_{\boldsymbol{z}}|$ is the projector to Λ_n.

- $\Pi^{\perp} = \sum_{\boldsymbol{z}\in\{0,1\}^n\setminus\mathsf{T}^n_{\lambda,\delta}} |\Phi_{\boldsymbol{z}}\rangle\langle\Phi_{\boldsymbol{z}}|$ is the projector to $\Lambda_n^{\perp}$.

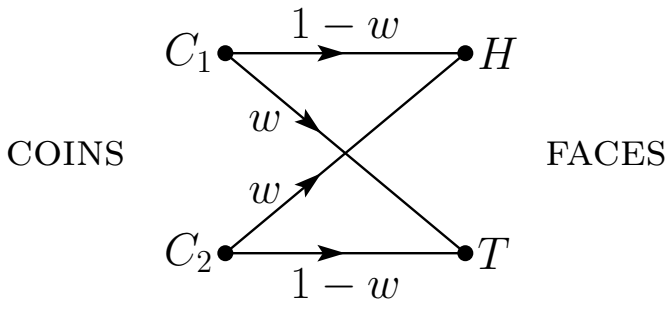
LOSSLESS QUANTUM DATA COMPRESSION
Encoding and Decoding

- Operations have to be allowed by quantum mechanics.
- Measurement: $\Pi + \Pi^{\perp} = I_{2^n}$.
- State after measurement:
  - $\Pi \cdot |\Psi_{\boldsymbol{x}}\rangle$ with probability $|\langle\Psi_{\boldsymbol{x}}|\Pi|\Psi_{\boldsymbol{x}}\rangle|^2$
  - $\Pi^{\perp} \cdot |\Psi_{\boldsymbol{x}}\rangle$ with probability $|\langle\Psi_{\boldsymbol{x}}|\Pi^{\perp}|\Psi_{\boldsymbol{x}}\rangle|^2$
- Expected probability of outcome $\Pi \cdot |\Psi_{\boldsymbol{x}}\rangle$:

$$\begin{aligned}\sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x})|\langle\Psi_{\boldsymbol{x}}|\Pi|\Psi_{\boldsymbol{x}}\rangle|^2 &\geq -1 + 2\,\mathrm{Tr}(\Pi\rho^{\otimes n}) \\ &= -1 + 2\,\mathrm{Tr}\Big\{\Big[\sum_{\boldsymbol{z}\in\mathsf{T}^n_{\lambda,\delta}} |\Phi_{\boldsymbol{z}}\rangle\langle\Phi_{\boldsymbol{z}}|\Big]\cdot\Big[\sum_{\boldsymbol{z}\in\{0,1\}^n} \lambda(\boldsymbol{z})|\Phi_{\boldsymbol{z}}\rangle\langle\Phi_{\boldsymbol{z}}|\Big]\Big\} \\ &= 1 - 2\epsilon_n\end{aligned}$$

SOURCES OF MIXED STATES

- Produce sequences of sources, $e.g.$, coins:



- A quantum example:

$$\begin{aligned}
\rho_1 &= \tfrac{2}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| \\
\rho_2 &= \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{2}{3}|\psi_3\rangle\langle\psi_3| \\
\rho &= \tfrac{1}{2}\rho_1 + \tfrac{1}{2}\rho_2 \\
&= \tfrac{1}{3}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{3}|\psi_2\rangle\langle\psi_2| + \tfrac{1}{3}|\psi_3\rangle\langle\psi_3| \\
&= \tfrac{1}{2}I
\end{aligned}$$

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

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

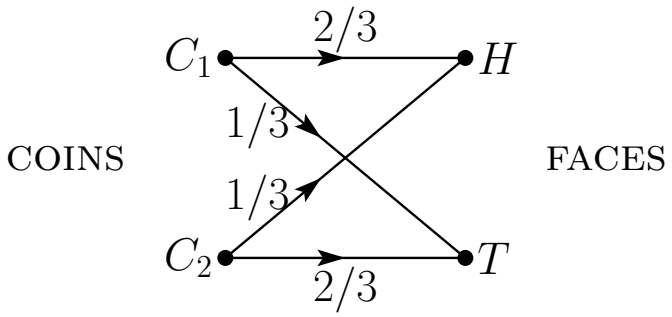
$$|\psi_2\rangle = \begin{bmatrix} -1/2 \\ \sqrt{3}/2 \end{bmatrix}$$

SOURCES OF SOURCES

Sequences of Sources

- Example:

$$P(C_1) = 1/2 \text{ and } P(C_2) = 1/2$$



$$Q(H) = 1/2 \text{ and } Q(T) = 1/2$$

- Sequences:

C_1	C_1	C_2	C_1	C_2	C_2
H	H	H	T	T	T
H	T	H	H	T	T
H	T	T	H	T	H

$\vdots$

SOURCES OF SOURCES
Sequences of Sources

- For each source (coin) a in $\mathcal{X}$, we have a probability distribution $W(\cdot|a)$ over alphabet $\mathcal{Y}$ of letters (faces).

- For sequence of coins $\boldsymbol{x}$ in $\mathcal{X}^n$, probability distribution $W^n(\cdot|\boldsymbol{x})$ over $\mathcal{Y}^n$.

- Original source Alice and reproduction source Bob.

- Alice has coins $\boldsymbol{x}$ and sends N bits to Bob.

- Bob prepares faces $\boldsymbol{y}$ with probability $\widehat{W}^n(\boldsymbol{y}|\boldsymbol{x})$.

SOURCES OF MIXED STATES

Compression by Sending Classical Information

- $W^n(\cdot|\boldsymbol{x})$ is reproduced as $\widehat{W}^n(\cdot|\boldsymbol{x})$.

- Fidelity between $W^n(\cdot|\boldsymbol{x})$ and $\widehat{W}^n(\cdot|\boldsymbol{x})$

$$F\big(W^n(\cdot|\boldsymbol{x}), \widehat{W}^n(\cdot|\boldsymbol{x})\big) = \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sqrt{W^n(\boldsymbol{y}|\boldsymbol{x}) \cdot \widehat{W}^n(\boldsymbol{y}|\boldsymbol{x})}$$

  is known as the Bhattacharyya-Wooters overlap between two probability distributions.

- Compression fidelity

$$\sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) F\big(W^n(\cdot|\boldsymbol{x}), \widehat{W}^n(\cdot|\boldsymbol{x})\big)$$

  should approach 1 as $n \to \infty$.

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:

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

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

SOURCES OF MIXED STATES

Compression Options

- Compress $\boldsymbol{x}$ or $\boldsymbol{y}$?

- Source of faces:

$$Q(b) = \sum_{a \in \mathcal{X}} P(a) W(b|a).$$

- Classical information measures:

$$H(Q) = -\sum_{b \in \mathcal{Y}} Q(b) \log Q(b)$$

$$H(W/P) = -\sum_{a \in \mathcal{X}} P(a) \sum_{b \in \mathcal{Y}} W(b|a) \log W(b|a)$$

$$I(P, W) = H(Q) - H(W/P)$$

SOURCES OF MIXED STATES
On the Idea and Proof

- $W^n(\cdot|\boldsymbol{x})$ is roughly a uniform distribution over $\boldsymbol{y}$s that are conditionally typical given $\boldsymbol{x}$.
- For a typical $\boldsymbol{x}$, there are about $2^{nH(W/P)}$ such $\boldsymbol{y}$s.
- These $\boldsymbol{y}$s are typical.
- There are about $2^{nH(Q)}$ typical $\boldsymbol{y}$s.
- A randomly chosen $\boldsymbol{y}$ will be conditionally typical with respect to any typical $\boldsymbol{x}$ with probability of about

$$\frac{2^{nH(W/P)}}{2^{nH(Q)}} = \frac{1}{2^{nI(P,W)}}.$$

SOURCES OF MIXED STATES
Compression Algorithm for a typical x

- Alice and Bob
  - have identical random number generators
  - which they use to form a list of N_l typical ys.

- If $N_l > 2^{nI(P,W)}$, then with high probability there will be at least one y on the list which is conditionally typical with respect to x.

- Alice sends $\log N_l$ bits to Bob identifying y.

- Compression rate $\log N_l/n$ approaches $I(P,W)$.

SOURCES OF MIXED STATES

Coins v.s. Mixed States

- Classical information measures:

$$H(Q) = -\sum_{b\in\mathcal{Y}} Q(b)\log Q(b)$$

$$H(W/P) = -\sum_{a\in\mathcal{X}} P(a)\sum_{b\in\mathcal{Y}} W(b|a)\log W(b|a)$$

$$I(P,W) = H(Q) - H(W/P)$$

- Quantum information measures:

$$S(\rho) = -\operatorname{Tr}\rho\log\rho$$

$$\bar{S} = \sum_{a\in\mathcal{X}} P(a)S(\rho_a)$$

$$\chi = S(\rho) - \bar{S}$$

QUANTUM DATA COMPRESSION
More Problems

- Type covering:

$$\bar{F} = \sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) F_{\mathcal{X}\times\mathcal{Y}}\big(P_{\boldsymbol{x}}W(\cdot|\cdot), P_{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x})}\big)$$

$$\begin{aligned} F_{\mathcal{X}\times\mathcal{Y}}\big(P_{\boldsymbol{x}}W(\cdot|\cdot), P_{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x})}\big) &= \Big[\sum_{(a,b)\in\mathcal{X}\times\mathcal{Y}} \sqrt{P_{\boldsymbol{x}}(a)W(b|a)\cdot P_{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x})}(a,b)}\Big]^2 \\ &= \Big[\sum_{(a,b)\in\mathcal{X}\times\mathcal{Y}} \frac{1}{n}\sqrt{N(a|\boldsymbol{x})W(b|a)\cdot N(a,b|\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))}\Big]^2. \end{aligned}$$

- Sources with memory?