Ralph's quantum data compression

Schumacher's quantum data compression theorem

Schumacher's quantum data compression theorem closely parallels Shannon's noiseless channel (classical data compression) theorem. While Shannon's theorem indicates the compression rate is bounded by classical (Shannon) entropy $H(p)$, Schumacher's theorem indicates the quantum compression rate R is bounded by quantum (von Neumann) entropy $S(\rho)$. Shannon's proof argues compression of typical sequences only is sufficient while Schumacher's proof argues typical subspaces are sufficient to achieve high fidelity. Further, in the limit sequences or qubits are almost all typical. A key connection is characters employed in messages can be represented by qubits analogous to their representation by bits classically.

The compression rate R is the number of qubits transmitted per qubit in the source message. Hence, smaller R is greater compression. For instance, if $S(\rho) = \frac{1}{2}$ and the source employs four qubit coding for characters than the maximum effective compression $R = \frac{1}{2}$ implies sending two qubits per (four qubit) character produces high fidelity and $R < \frac{1}{2}$ (say, transmitting one qubit per character) fails to provide high fidelity.

Quantum fidelity

Quantum fidelity is a measure of closeness between two density operators ρ and σ.

$$F(\rho,\sigma) = F(\sigma,\rho) = Tr\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 = Tr\left(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2$$

If $\rho = |\psi_\rho\rangle\langle\psi_\rho|$ is a pure state then $F(\rho,\sigma) = \langle\psi_\rho|\sigma|\psi_\rho\rangle$.

Example

Ralph wishes to send a three qubit message but only has capacity for two qubits (qubit processing is costly). Ralph's characters are created from single qubits drawn from the ensemble

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with equal probability.[1] Hence, the density operator for the source is

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| = \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

[1] An ensemble or mixed state provides sufficient variation to distinguish characters. This mixed state may have arisen from tracing out the second qubit of a two qubit entangled state, say, $\frac{\sqrt{3}}{2}|00\rangle + \frac{1}{2\sqrt{3}}|10\rangle + \frac{1}{\sqrt{6}}|11\rangle$.

Spectral decomposition of ρ is

$$\rho = Q\Lambda Q^H = \begin{bmatrix} \cos\frac{\pi}{8} & \sin\frac{\pi}{8} \\ \sin\frac{\pi}{8} & -\cos\frac{\pi}{8} \end{bmatrix} \begin{bmatrix} \cos^2\frac{\pi}{8} & 0 \\ 0 & \sin^2\frac{\pi}{8} \end{bmatrix} \begin{bmatrix} \cos\frac{\pi}{8} & \sin\frac{\pi}{8} \\ \sin\frac{\pi}{8} & -\cos\frac{\pi}{8} \end{bmatrix}$$

This leads to von Neumann entropy equal to

$$\begin{aligned} S(\rho) = -Tr(\rho \log_2 \rho) = -\sum \lambda_i \log_2 \lambda_i \\ = -\cos^2\frac{\pi}{8}\log_2\cos^2\frac{\pi}{8} - \sin^2\frac{\pi}{8}\log_2\sin^2\frac{\pi}{8} = 0.60088 \end{aligned}$$

Schumacher's theorem indicates two qubit compression of three qubit messages produces high fidelity. Ralph is anxious to discover the fidelity of this compression strategy.

Denote the eigenstates $|0'\rangle = \begin{bmatrix} \cos\frac{\pi}{8} \\ \sin\frac{\pi}{8} \end{bmatrix}$ and $|1'\rangle = \begin{bmatrix} \sin\frac{\pi}{8} \\ -\cos\frac{\pi}{8} \end{bmatrix}$ then the full complement of three qubit messages are

$$|0'0'0'\rangle, |0'0'1'\rangle, |0'1'0'\rangle, |1'0'0'\rangle, |0'1'1'\rangle, |1'0'1'\rangle, |1'1'0'\rangle, |1'1'1'\rangle$$

The first four messages are highly probable and comprise the typical subspace while the last four are much less likely and are orthogonal to the typical subspace.[2]

$$\begin{aligned} \Pr(0'0'0') = Tr\left(|0'0'0'\rangle\langle 0'0'0'|\,\rho^{\otimes 3}\right) = 0.6219 \\ \Pr(0'0'1') = \Pr(0'1'0') = \Pr(1'0'0') = 0.1067 \\ \Pr(0'1'1') = \Pr(1'0'1') = \Pr(1'1'0') = 0.0183 \\ \Pr(1'1'1') = 0.0031 \end{aligned}$$

The typical subspace has probability equal to $0.6219 + 3(0.1067) = 0.9419$. Further, the base state for this system is

$$\begin{aligned} |\psi\rangle = \sqrt{0.6219}\,|0'0'0'\rangle + \sqrt{0.1067}\,|0'0'1'\rangle + \sqrt{0.1067}\,|0'1'0'\rangle + \sqrt{0.1067}\,|1'0'0'\rangle \\ +\sqrt{0.0183}\,|0'1'1'\rangle + \sqrt{0.0183}\,|1'0'1'\rangle + \sqrt{0.0183}\,|1'1'0'\rangle + \sqrt{0.0031}\,|1'1'1'\rangle \\ = \frac{1}{\sqrt{8}}\left(|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle\right) \end{aligned}$$

-

Data compression strategy

A simple strategy calls for transmitting the first two qubits of the coded message to the receiver and the receiver appending the highest probably third

[2]The typical subspace theorem indicates the dimension of the typical subspace is no greater than $2^{nS(\rho)}$. Since $S(\rho) = 0.60088$ and $n = 4$, dimension of the typical subspace is less than 5.3. However, the natural break is dimension four rather than five as messages five, six, and seven have the same probability (also, messages two, three, and four have the same probability).

qubit. The probability of the first two qubits is one and the probability for the third qubit is $\Pr(|0'\rangle) = \cos^2\frac{\pi}{8} = 0.8536$. Hence, the fidelity of this stategy is 0.8536. However, this is only a dimension one approach and there is a better strategy. Utilizing the dimension three subspace increases the fidelity as discussed next.

An improved data compression strategy follows. Transform the high probability qubits to $|\cdot\cdot 0\rangle$ and the low probability qubits to $|\cdot\cdot 1\rangle$ by unitary operator U. Then the sender measures the third qubit. If the third qubit is $|0\rangle$ the sender transmits the first two qubits. The receiver then appends $|0\rangle$ to the received qubits and applies the reverse transformation U^H leaving one of the high probability coded message. If measurement of the third qubit is $|1\rangle$ then the best strategy for the sender is to play the highest probability message (only typical subspaces are in play). In this case, whatever two qubits $|0'0'0'\rangle$ is transmitted as is sent and the receive applies the same process as above (appends $|0\rangle$ and applies U^H).

A unitary operator for this system is $T_{123}T_{231}C_{32}Q^{H\otimes 3}$ where T_{ijk} is a Toffoli gate with control qubits ij and target qubit k (iff $|ij\rangle = |11\rangle$ then $|k\rangle$ is bit flipped), C_{ij} is a controlled-not gate with i the control and j the target, and $Q^{H\otimes 3}$ is the Hermitian (conjugate transpose) of the matrix of eigenstates of ρ (tensored to accommodate three qubits).

Suggested:

1. Verify $Q^{H\otimes 3}|0'0'0'\rangle = |000\rangle$, $Q^{H\otimes 3}|0'0'1'\rangle = |001\rangle$, $Q^{H\otimes 3}|0'1'0'\rangle = |010\rangle$, $Q^{H\otimes 3}|1'0'0'\rangle = |100\rangle$, $Q^{H\otimes 3}|0'1'1'\rangle = |011\rangle$, $Q^{H\otimes 3}|1'0'1'\rangle = |101\rangle$, $Q^{H\otimes 3}|0'1'1'\rangle = |011\rangle$, and $Q^{H\otimes 3}|1'1'1'\rangle = |111\rangle$ (rotation to computational basis).

2. Verify $T_{123}T_{231}C_{32}$ takes the high probability states (after transforming to computational basis by $U^{H\otimes 3}$) to $|\cdot\cdot 0\rangle$ and the low probability states to $|\cdot\cdot 1\rangle$.

3. For each potential three qubit message indicate the two qubits transmitted to the receiver by this compression strategy.

4. Explain why this compression strategy produces

$$\rho' = E|\psi\rangle\langle\psi|E + |0'0'0'\rangle\langle\psi|(I-E)|\psi\rangle\langle 0'0'0'|$$

where

$$E = |0'0'0'\rangle\langle 0'0'0'| + |0'0'1'\rangle\langle 0'0'1'| + |0'1'0'\rangle\langle 0'1'0'| + |1'0'0'\rangle\langle 1'0'0'|$$

the projection into the high probability (typical) subspace.

5. Verify the fidelity of this compression strategy is $\langle\psi|\rho'|\psi\rangle = 0.9234 \leq Tr\left(E\rho^{\otimes 3}\right) = 0.9419$.

Suggested responses:

1.

$$Q^H|0'\rangle = \begin{bmatrix} \cos\frac{\pi}{8} & \sin\frac{\pi}{8} \\ \sin\frac{\pi}{8} & -\cos\frac{\pi}{8} \end{bmatrix} \begin{bmatrix} \cos\frac{\pi}{8} \\ \sin\frac{\pi}{8} \end{bmatrix} = \begin{bmatrix} \cos^2\frac{\pi}{8} + \sin^2\frac{\pi}{8} \\ \cos\frac{\pi}{8}\sin\frac{\pi}{8} - \cos\frac{\pi}{8}\sin\frac{\pi}{8} \end{bmatrix} = |0\rangle$$

and

$$Q^H|1'\rangle = \begin{bmatrix} \cos\frac{\pi}{8} & \sin\frac{\pi}{8} \\ \sin\frac{\pi}{8} & -\cos\frac{\pi}{8} \end{bmatrix} \begin{bmatrix} \sin\frac{\pi}{8} \\ -\cos\frac{\pi}{8} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{8}\sin\frac{\pi}{8} - \cos\frac{\pi}{8}\sin\frac{\pi}{8} \\ \cos^2\frac{\pi}{8} + \sin^2\frac{\pi}{8} \end{bmatrix} = |1\rangle$$

Each $Q^H|\cdot\rangle$ is independent, hence transformation to computational basis is established.

2. The following table applies the remaining unitary operations in order (down the rows).

	$\|000\rangle$	$\|001\rangle$	$\|010\rangle$	$\|100\rangle$	$\|011\rangle$	$\|101\rangle$	$\|110\rangle$	$\|111\rangle$
C_{32}	$\|000\rangle$	$\|011\rangle$	$\|010\rangle$	$\|100\rangle$	$\|001\rangle$	$\|111\rangle$	$\|110\rangle$	$\|101\rangle$
T_{231}	$\|000\rangle$	$\|111\rangle$	$\|010\rangle$	$\|100\rangle$	$\|001\rangle$	$\|011\rangle$	$\|110\rangle$	$\|101\rangle$
T_{123}	$\|000\rangle$	$\|110\rangle$	$\|010\rangle$	$\|100\rangle$	$\|001\rangle$	$\|011\rangle$	$\|111\rangle$	$\|101\rangle$

3. The first two qubits in the last row of the above table are transmitted for each of the high probability states and the first column result $|00\rangle$ is transmitted if measurement of the third qubit produces $|1\rangle$.

4. The first term, $E|\psi\rangle\langle\psi|E$, is the projection of the base density operator, $|\psi\rangle\langle\psi|$ into the high probability subspace producing $|\psi_h\rangle\langle\psi_h|$ where

$$|\psi_h\rangle = \sqrt{0.6219}\,|0'0'0'\rangle + \sqrt{0.1067}\,|0'0'1'\rangle + \sqrt{0.1067}\,|0'1'0'\rangle + \sqrt{0.1067}\,|1'0'0'\rangle$$

The second term, $|0'0'0'\rangle\langle\psi|(I-E)|\psi\rangle\langle 0'0'0'|$, is the density operator associated with the highest probability state $|0'0'0'\rangle$ weighted by the probability associated with the low probability subspace $\langle\psi|(I-E)|\psi\rangle$.

5. Since the base state $|\psi\rangle$ is a pure state, fidelity $F(\psi,\rho')$ simplifies as $\langle\psi|\rho'|\psi\rangle$.

$$\begin{aligned}\langle\psi|\rho'|\psi\rangle &= \langle\psi|E|\psi\rangle\langle\psi|E|\psi\rangle + \langle\psi|0'0'0'\rangle\langle\psi|(I-E)|\psi\rangle\langle 0'0'0'|\psi\rangle \\ &= (\langle\psi|E|\psi\rangle)^2 + \langle\psi|(I-E)|\psi\rangle(\langle\psi|0'0'0'\rangle)^2 \\ &= (0.9419)^2 + (0.0581)(0.6219) = 0.9234\end{aligned}$$