Optimal lower bound of the average indeterminate length lossless quantum block encoding

Rabins Wosti (joint work with George Androulakis)

University of South Carolina

arXiv:2305.18748

Outline

Classical Source

Consider a **classical source** $\mathcal{B}$ that emits one symbol at each discrete time step from the **symbol set** $\mathcal{S} = \{s_i\}_{i=1}^{4}$ with the following probabilities. Assume that the emissions at each time step are independent and identically distributed (i.i.d.).

Symbol	*Probability*
s_1	$p(1) = 0.5$
s_2	$p(2) = 0.25$
s_3	$p(3) = 0.125$
s_4	$p(4) = 0.125$

Fixed-length classical encoding scheme

Let $A = \{0, 1\}$ be the **binary alphabet**. An example of a **fixed-length encoding scheme** $\phi = \{E : \mathcal{S} \to A^*, D : A^* \to \mathcal{S}\}$:

Symbol	*Probability*	**Codeword**	**Codeword Length**
s_1	$p(1) = 0.5$	$E(s_1) = 00$	$length(E(s_1)) = 2$
s_2	$p(2) = 0.25$	$E(s_2) = 01$	$length(E(s_2)) = 2$
s_3	$p(3) = 0.125$	$E(s_3) = 10$	$length(E(s_3)) = 2$
s_4	$p(4) = 0.125$	$E(s_4) = 11$	$length(E(s_4)) = 2$

Average codeword length $= \sum_{i=1}^{4} p(i) * length(E(s_i)) = 2$ bits/symbol. In general, the average codeword length $= \lceil \log_2 N \rceil$ for a symbol set of size N.

Variable-length classical encoding scheme

An example of a **variable-length encoding scheme** $\phi' = \{E', D'\}$:

Symbol	*Probability*	*Codeword*	*Codeword Length*
s_1	$p(1) = 0.5$	$E'(s_1) = 0$	$length(E(s_1)) = 1$
s_2	$p(2) = 0.25$	$E'(s_2) = 10$	$length(E(s_2)) = 2$
s_3	$p(3) = 0.18$	$E'(s_3) = 110$	$length(E(s_3)) = 3$
s_4	$p(4) = 0.07$	$E'(s_4) = 111$	$length(E(s_4)) = 3$

Average codeword length =
$\sum_{i=1}^{4} p(i) * length(E'(s_i)) = 1.75$ bits/symbol.
Called **Huffman code**.

Uniquely-decodable codes (Lossless codes)

Symbol	*Probability*	*Huffman Codeword*	*Alternate Codeword*
s_1	$p(1) = 0.5$	$E(s_1) = 0$	$E'(s_1) = 0$
s_2	$p(2) = 0.25$	$E(s_2) = 10$	$E'(s_2) = 10$
s_3	$p(3) = 0.125$	$E(s_3) = 110$	$E'(s_3) = 100$
s_4	$p(4) = 0.125$	$E(s_4) = 111$	$E'(s_4) = 111$

A sequence of codeword 100 can be decoded in two ways: $s_2 s_1$ and s_3.

Classical Kraft-McMillan Inequality:
Assume that a uniquely decodable classical encoding scheme over a binary alphabet encodes a set of D-many symbols into codewords of lengths $\{\ell_i \in \mathbb{N}\}_{i=1}^{D}$, then the codeword lengths must satisfy the following inequality

$$\sum_{i=1}^{D} 2^{-\ell_i} \leq 1$$

Conversely, if there exists a set of lengths $\{\ell_i\}_{i=1}^{D}$ that satisfy the above inequality, then there exists a uniquely decodable classical encoding scheme with those codeword lengths.

Classical Block Encoding

Fix block size $l = 2$. Consider 2-extension of the classical source $\mathcal{B}^2$ that emits one symbol from the symbol set

$$\mathcal{S}^2 = \{s_{i_1} s_{i_2}\}_{i_1, i_2 = 1}^{4}$$

at each discrete time step. Each symbol $s_{i_1} s_{i_2}$ is emitted with the probability $p(i_1)p(i_2)$ for $i_1, i_2 \in \{1, \ldots, 4\}$.

Average codeword length of Huffman code that encodes symbols from the source $\mathcal{B}^2$ is 3.46 bits/block. That is, 1.73 bits/symbol. Also called **code rate**.

What if block size $(l) \to \infty$?

How much can you compress the information from an i.i.d. classical source with negligible loss of information?

Shannon's source coding theorem (Informally)

Consider an i.i.d. classical source $\mathcal{B} = \{p(i), s_i\}_{i=1}^N$ with **Shannon entropy** $H(\mathcal{B}) = -\sum_{i=1}^N p(i) \log_2 p(i)$. Also, let $\mathcal{S}^n = \{s_{i_1}, \ldots, s_{i_n}\}$ for $i_1, \ldots, i_n \in \{1, \ldots, N\}$ be the set consisting of all n-sequences of symbols from the source. For any $\epsilon \geq 0$, let $R = H(\mathcal{B}) + \epsilon$. Then there exists a reliable compression scheme $\phi^n = \{E^n, D^n\}$ of code rate R for the n-extension of the source. That is,

$$Pr(s_1, \ldots, s_n \ : \ D^n \circ E^n(s_1, \ldots, s_n) = (s_1, \ldots, s_n)) \to 1 \quad \text{as } n \to \infty$$

Conversely, any compression scheme $\phi^n = \{E^n, D^n\}$ with code rate $R < H(\mathcal{B})$ is not reliable for the n-extension of the source. That is,

$$Pr(s_1, \ldots, s_n \ : \ D^n \circ E^n(s_1, \ldots, s_n) = (s_1, \ldots, s_n)) \to 0 \quad \text{as } n \to \infty$$

Based on Asymptotic Equipartition Property (AEP) of typical sequences.
$nH(\mathcal{B}) \leq$ Average codeword length of Huffman code for $\mathcal{B}^n \leq nH(\mathcal{B}) + 1$.

Schumacher's quantum source coding theorem (Informally)

Consider an i.i.d. quantum source $\mathcal{B} = \{p(i), |s_i\rangle\}_{i=1}^{N}$ where each $|s_i\rangle \in \mathcal{H}$ of dimension D. Let $\rho = \sum_{i=1}^{N} p(i) |s_i\rangle\langle s_i|$ be the **ensemble state**.
Consider the spectral decomposition of $\rho = \sum_{i=1}^{D} \lambda_i |\lambda_i\rangle\langle\lambda_i|$.
Then, the **von-Neumann entropy** of the source is given by $\mathcal{S}(\rho) = -\sum_{i=1}^{D} \lambda_i \log_2 \lambda_i$. Also, let $\mathcal{S}^n = \{s_{i_1}, \ldots, s_{i_n}\}$ for $i_1, \ldots, i_n \in \{1, \ldots, N\}$ be the set consisting of all n-sequences of symbols from the source. For any $\epsilon \geq 0$, let $R = \mathcal{S}(\rho) + \epsilon$. Then there exists a reliable compression scheme $\phi^n = \{E^n, D^n\}$ of code rate R for the n-extension of the source. That is,

$$tr(\Pi_{\Lambda_n} \rho^{\otimes n}) \to 1 \text{ as } n \to \infty$$

where $\Lambda_n \subset \mathcal{H}^{\otimes n}$ is the typical subspace and Π_{Λ_n} is the projector from $\mathcal{H}^{\otimes n}$ onto Λ_n. Conversely, any compression scheme $\phi^n = \{E^n, D^n\}$ with code rate $R < \mathcal{S}(\rho)$ is not reliable for the n-extension of the source. That is,

$$tr(\Pi_{\Lambda_n} \rho^{\otimes n}) \to 0 \text{ as } n \to \infty$$

Based on AEP of typical subspaces.

Generalization of Schumacher's compression

- Universal compression scheme: Josza et. al.[1] for i.i.d. quantum sources, Kaltchenko & Yang [2, 3] for stationary ergodic sources
- Variable-length (indeterminate-length) compression scheme: Schumacher & Westmoreland [4], and Hayashi & Matsumoto [5, 6] for i.i.d. sources
- quantum source: Kaltchenko & Yang [2, 3] and Bjelaković & Igor [7] for stationary ergodic sources, all others for i.i.d. sources

No-go theorems for (completely) lossless quantum data compression introduced by Boström & Felbinger [8].
We are interested in:

- completely lossless
- source-dependent
- indeterminate-length
- quantum stochastic source
- encode tensor product of pure states in blocks of equal size

Quantum code

Let $\mathcal{H}$ be a Hilbert space of dimension D. A **quantum code** on $\mathcal{H}$ is a linear isometry $U : \mathcal{H} \to (\mathbb{C}^2)^{\oplus}$. Thus, for every quantum code U on $\mathcal{H}$, the dimension of its range is equal to D, and if $(|\psi_i\rangle)_{i=1}^{D}$ is any orthonormal sequence in its range, then U has the form

$$U = \sum_{i=1}^{D} |\psi_i\rangle\langle e_i| \,,$$

where $(|e_i\rangle)_{i=1}^{D}$ is an orthonormal basis of $\mathcal{H}$. For any pure state $|s\rangle = \sum_{i=1}^{D} \alpha_i |e_i\rangle$, the quantum state $|\sigma\rangle = U|s\rangle = \sum_{i=1}^{D} \alpha_i |\psi_i\rangle$ obtained by applying U to a state $|s\rangle \in \mathcal{H}$ is called **codeword**.
Let each $|\psi_i\rangle \in (\mathbb{C}^2)^{\otimes \ell_i}$ for some $\ell_i \in \mathbb{N}$. If $\ell_i = \ell \; \forall i$, then $|\sigma\rangle \in (\mathbb{C}^2)^{\otimes \ell}$ and is a **length state**, otherwise it is an **indeterminate-length state**.
Define length observable $\Lambda = \sum_{\ell=0}^{\infty} \ell \Pi_\ell$. Indeterminate-length of a codeword $|\sigma\rangle$ is given by $Tr(|\sigma\rangle\langle\sigma| \Lambda) = \langle\sigma| \Lambda |\sigma\rangle = \sum_{i=1}^{D} |\alpha_i|^2 \ell_i$.

Concatenation of quantum codewords

Consider a linear map $U^m = U \circ \cdots \circ U : \mathcal{H}^{\otimes m} \to (\mathbb{C}^2)^{\oplus}$ defined such that

$$(U \circ \cdots \circ U) |s_1\rangle \otimes \cdots \otimes |s_m\rangle = U|s_1\rangle \circ \cdots \circ U|s_m\rangle = |\sigma_i\rangle \circ \cdots \circ |\sigma_m\rangle$$

where each $|s_i\rangle = \sum_{j=1}^{D} \alpha_i^j |e_j\rangle$ such that $\sum_{j=1}^{D} |\alpha_i^j|^2 = 1$. Concatenation of normalized indeterminate-length states can result in unnormalized states. (Check $\frac{1}{\sqrt{2}}(|0\rangle + |00\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |00\rangle!)$ [9]
So, $(U \circ \cdots \circ U) |s_1\rangle \otimes \cdots \otimes |s_m\rangle$

$$= \sum_{j_1,\ldots,j_m=1}^{D} \alpha_1^{j_1} \cdots \alpha_m^{j_m} |\psi_{j_1}\rangle \circ \cdots \circ |\psi_{j_m}\rangle$$

U^m is an isometry $\iff \{|\psi_{j_1} \circ \cdots \circ \psi_{j_m}\rangle\}$ is an orthonormal set.
U is uniquely decodable
$\iff$ U^m is an isometry for every $m \in \mathbb{N}$
$\iff$ $\{|\psi_{j_1} \circ \cdots \circ \psi_{j_m}\rangle : (j_1, \ldots, j_m) \in \{1, \ldots, D\}^m\}$ is an orthonormal set for every $m \in \mathbb{N}$. Such a sequence $(|\psi_1\rangle)_{i=1}^{D}$ is called jointly orthonormal sequence of length codewords.

Quantum stochastic source

A **quantum stochastic source** $\mathcal{B}$ consists of a set of pure states $\{|s_n\rangle\}_{n=1}^N$ of a Hilbert space $\mathcal{H}$, and a stochastic process $X = (X_n)_{n=1}^{\infty}$, where each X_n is a random variable which takes values in $\{1, 2, \ldots, N\}$. At every positive integer time n the state $|s_{X_n}\rangle$ is emitted from the quantum source. If p denotes the probability distribution of the stochastic process X, then for every $k \in \mathbb{N}$ and $(n_1, \ldots, n_k) \in \{1, \ldots, N\}^k$, we have that

$$p(n_1, \ldots, n_k) = \mathbb{P}(X_1 = n_1, \ldots, X_k = n_k).$$

Also, the conditional probability distribution of the stochastic process X is defined for any $k \geq 2$ and $(n_1, \ldots, n_k) \in \{1, \ldots, N\}^k$ by

$$p(n_k | n_{k-1}, \ldots, n_1) = \mathbb{P}(X_k = n_k | X_{k-1} = n_{k-1}, \ldots, X_1 = n_1).$$

Case for 2 blocks

Fix $r \in \mathbb{N}$ as the block size and $m = 2$ as the number of blocks. Consider the ensemble state of two block ρ_{2r}

$$
\begin{aligned}
&= \sum_{i_1,\ldots,i_{2r}=1}^{N} p(i_1,\ldots,i_{2r}) \left|s_{i_1}\cdots s_{i_{2r}}\right\rangle\!\left\langle s_{i_1}\cdots s_{i_{2r}}\right| \\
&= \sum_{i_1,\ldots,i_r=1}^{N} p(i_1,\ldots,i_r) \left|s_{i_1}\cdots s_{i_r}\right\rangle\!\left\langle s_{i_1}\cdots s_{i_r}\right| \otimes \\
&\qquad \sum_{i_{r+1},\ldots,i_{2r}=1}^{N} p(i_{r+1},\ldots,i_{2r}|i_1,\ldots,i_r) \left|s_{i_{r+1}}\cdots s_{i_{2r}}\right\rangle\!\left\langle s_{i_{r+1}}\cdots s_{i_{2r}}\right| \\
&= \sum_{i_1,\ldots,i_r=1}^{N} p(i_1,\ldots,i_r) \left|s_{i_1}\cdots s_{i_r}\right\rangle\!\left\langle s_{i_1}\cdots s_{i_r}\right| \otimes \rho^{i_1,\ldots,i_r}
\end{aligned}
$$

where $\rho^{i_1,\ldots,i_r}$ is second block ensemble state given that $\left|s_{i_1}\cdots s_{i_r}\right\rangle$ is emitted in the first block.

Consider a linear map $U_1 \circ \cdots \circ U_m : \mathcal{H}^{\otimes m} \to (\mathbb{C}^2)^{\oplus}$ defined such that

$$(U_1 \circ \cdots \circ U_m) |s_1\rangle \otimes \cdots \otimes |s_m\rangle = U_1 |s_1\rangle \circ \cdots \circ U_m |s_m\rangle$$

and each $U_j = \sum_{i=1}^{D} \left|\psi_i^j\right\rangle\!\left\langle e_i^j\right|$.
$U_1 \circ \cdots \circ U_m$ is uniquely decodable
$\iff$ $U_1 \circ \cdots \circ U_m$ is an isometry
$\iff$ $\left|\psi_{j_1}^1 \circ \cdots \circ \psi_{j_m}^m\right\rangle : (j_1, \ldots, j_m) \in \{1, \ldots, D\}^m\}$ is an orthonormal set.
Fix a jointly orthonormal sequence of length codewords $(|\psi_i\rangle)_{i=1}^{D}$. Setting $(\left|\psi_i^j\right\rangle)_{i=1}^{D} = (|\psi_i\rangle)_{i=1}^{D}$ $\forall j$ makes $U_1 \circ \cdots \circ U_m$ uniquely decodable.

Special block codes

Consider a quantum stochastic source $\mathcal{S}$ which contains an alphabet of N many pure states $(|s_i\rangle)_{i=1}^{N}$ that span a Hilbert space $\mathcal{H}$ of dimension D. Let $r, m \in \mathbb{N}$ where r denotes the block size, and m denotes the number of blocks. A **special block code** is a family of isometries

$$\mathcal{U} = \left\{ U^{n_1,\dots,n_{(k-1)r}} : 1 \le k \le m,\ n_1,\dots,n_{(k-1)r} \in \{1,\dots,N\} \right\},$$

such that every isometry used in the family $\mathcal{U}$ has a common sequence of jointly orthonormal length codewords. Thus more explicitly, there exists a jointly orthonormal sequence of length codewords $(|\psi_i\rangle)_{i=1}^{D^r} \subseteq (\mathbb{C}^2)^{\oplus}$, and an orthonormal sequence $\left(\left|e_i^{n_1,\dots,n_{(k-1)r}}\right\rangle\right)_{i=1}^{D^r}$ for $1 \le k \le m$ and $n_1,\dots,n_{(k-1)r} \in \{1,\dots,N\}$ such that

$$U^{n_1,\dots,n_{(k-1)r}} = \sum_{i=1}^{D^r} \left|\psi_i\right\rangle\!\left\langle e_i^{n_1,\dots,n_{(k-1)r}}\right|.$$

Average codeword length of special block code (for 2 blocks)

We will denote by $\boldsymbol{ACL(\mathcal{U})}$ the **average codeword length of the special block code** $\boldsymbol{\mathcal{U}}$, which is defined to be equal to

$$\sum_{n_1,\dots,n_{ml}=1}^{N} p(n_1,\dots,n_{mr}) \operatorname{Tr}\Bigg(\Big| U(s_{n_1}\cdots s_{n_r}) \circ U^{n_1,\dots,n_r}(s_{n_{r+1}}\cdots s_{n_{2r}}) \circ$$

$$\cdots \circ U^{n_1,\dots,n_{(m-1)r}}(s_{n_{(m-1)r+1}}\cdots s_{n_{mr}}) \Big\rangle$$

$$\Big\langle U(s_{n_1}\cdots s_{n_r}) \circ U^{n_1,\dots,n_{n_r}}(s_{n_{r+1}}\cdots s_{n_{2r}}) \circ$$

$$\cdots \circ U^{n_1,\dots,n_{(m-1)r}}(s_{n_{(m-1)r+1}}\cdots s_{n_{mr}}) \Big| \Lambda \Bigg)$$

Main result

Consider a quantum stochastic source $\mathcal{B}$ consisting of an alphabet of N many pure states spanning a D-dimensional Hilbert space $\mathcal{H}$, and a stochastic process X having mass function p as defined before. Fix $m, r \in \mathbb{N}$. Let

$$LB(\mathcal{B}, m, r)$$

denote the infimum of the set containing $ACL(\mathcal{U})$ for every special block code $\mathcal{U}$ that is used to encode mr many states emitted by $\mathcal{B}$ into m blocks each of size r. Then $LB(\mathcal{B}, m, r)$ can be computed as follows: For each $k = 1, \ldots, m$, and a sequence $n_1, \ldots, n_{(k-1)r}$ of integers chosen from the set $\{1, \ldots, N\}$, let $\left(\lambda_i^{n_1,\ldots,n_{(k-1)r}}\right)_{i=1}^{D^r}$ be the eigenvalues of the k^{th} block conditional ensemble state $\rho^{n_1,\ldots,n_{(k-1)r}}$, arranged in decreasing order, and $\left(\left|\lambda_i^{n_1,\ldots,n_{(k-1)r}}\right\rangle\right)_{i=1}^{D^r}$ be the corresponding eigenvectors.

Let

$$\mathfrak{L} = \Big\{(\ell_1, \dots, \ell_{D^r}) : \ell_i \in \mathbb{N} \cup \{0\} \text{ for all } i,\ \ell_1 \leq \ell_2 \leq \cdots \leq \ell_{D^r},\ \text{and} \sum_{i=1}^{D^r} 2^{-\ell_i} \leq 1\Big\}.$$

Define a function $L : \mathfrak{L} \to [0, \infty)$ by

$$L((\ell_i)_{i=1}^{D^r}) := \sum_{j=2}^{m} \left(\sum_{n_1,\dots,n_{(j-1)r}=1}^{N} p(n_1, \dots, n_{(j-1)r}) \sum_{i=1}^{D^r} \lambda_i^{n_1,\dots,n_{(j-1)r}} \ell_i \right) + \sum_{i=1}^{D^r} \lambda_i \ell_i.$$

Then,

$$LB(\mathcal{S}, m, r) = \min\{L((\ell_i)_{i=1}^{D^r}) : (\ell_i)_{i=1}^{D^r} \in \mathfrak{L}\}.$$

Moreover, the infimum defining $LB(\mathcal{S}, m, r)$ is actually a minimum, i.e., there exists a special block code

$$\min\{L((\ell_i)_{i=1}^{D^r}) : (\ell_i)_{i=1}^{D^r} \in \mathfrak{L}\} = ACL(\mathcal{V}).$$

The minimizer $\mathcal{V}$ is given as follows: Assume that L achieves its minimum on $\mathfrak{L}$ at the point $(\ell_i)_{i=1}^{D^r} \in \mathfrak{L}$. Since the sequence $(\ell_i)_{i=1}^{D^r}$ satisfies the classical Kraft-McMillan inequality, (which is the last condition in the definition of $\mathfrak{L}$), there exists a classical uniquely decodable sequence $(\omega_i)_{i=1}^{D^r}$ of codewords with corresponding lengths $(\ell_i)_{i=1}^{D^r}$. Let $(|\omega_i\rangle)_{i=1}^{D^r}$ be the corresponding sequence of qubit strings in the Fock space $(\mathbb{C}^2)^{\oplus}$. For each $k \in \{1, \ldots, m\}$, and string $n_1, \ldots, n_{(k-1)r} \in \{1, \ldots, N\}$, define

$$V^{n_1,\ldots,n_{(k-1)r}} : \mathcal{H}^{\otimes r} \to (\mathbb{C}^2)^{\oplus},$$

by

$$V^{n_1,\ldots,n_{(k-1)r}} = \sum_{i=1}^{D^r} \left|\omega_i\right\rangle\left\langle\lambda_i^{n_1,\ldots,n_{(k-1)r}}\right|.$$

Concatenation of rank-1 operators

Consider a collection of normalized states $\{|\phi_i\rangle\}_{i=1}^{N}$ such that each $|\phi_i\rangle$ is in the linear span of a jointly orthonormal sequence of length codewords $(|\psi_j\rangle)_{j=1}^{D}$. Then, $|\phi_i \circ \cdots \circ \phi_N\rangle$ is a normalized state. So, one can define the concatenation of rank-1 operators $(|\phi_i\rangle\langle\phi_i|)_{i=1}^{N}$ as

$$|\phi_1\rangle\langle\phi_1| \circ \cdots \circ |\phi_N\rangle\langle\phi_N| = |\phi_1 \circ \cdots \circ \phi_N\rangle\langle\phi_1 \circ \cdots \circ \phi_N|$$

Supporting lemma

For $z \in \mathbb{N}$ consider a non-increasing sequence of positive real numbers $Q_1 \geq Q_2 \geq \cdots \geq Q_z \geq 0$. Further, consider another arbitrary sequence of positive real numbers $l_1, l_2, \ldots, l_z$ and its non-decreasing enumeration $l'_1 \leq l'_2 \leq \cdots \leq l'_z$. Then,

$$\sum_{i=1}^{z} Q_i l'_i \leq \sum_{i=1}^{z} Q_i l_i.$$

Sketch of the proof for 2 blocks

Consider the quantum stochastic source as described above, and fix the number of blocks, m=2. The ensemble state for two blocks ρ_{2r} can be written as

$$= \sum_{n_1,\ldots,n_{2r}=1}^{D^r} p(n_1,\ldots,n_{2r}) \, |s_{n_1},\cdots,s_{n_{2r}}\rangle\langle s_{n_1},\cdots,s_{n_{2r}}|$$

The average codeword length of our encoding for two blocks is given by

$$= \sum_{n_1,\ldots,n_{2r}=1}^{N} p(n_1,\ldots,n_{2r}) \, \mathrm{Tr}\Bigg(\Big| U(s_{n_1}\cdots s_{n_r}) \circ U^{n_1,\ldots,n_{2r}}(s_{n_{r+1}}\cdots s_{n_{2r}}) \Big\rangle \Big\langle U(s_{n_1}\cdots s_{n_r}) \circ U^{n_1,\ldots,n_r}(s_{n_{r+1}}\cdots s_{n_{2r}}) \Big| \Lambda \Bigg)$$

$$= \sum_{n_1,\ldots,n_{2r}=1}^{N} p(n_1,\ldots,n_{2r}) \operatorname{Tr}\Bigg(U \left|s_{n_1}\cdots s_{n_r}\right\rangle\!\left\langle s_{n_1}\cdots s_{n_r}\right| U^\dagger \circ$$

$$U^{n_1,\ldots,n_r} \left|s_{n_{r+1}}\cdots s_{n_{2r}}\right\rangle\!\left\langle s_{n_{r+1}}\cdots s_{n_{2r}}\right| (U^{n_1,\ldots,n_r})^\dagger \Lambda \Bigg)$$

$$= \sum_{n_1,\ldots,n_r=1}^{N} \operatorname{Tr}\Bigg(p(n_1,\ldots,n_r) U \left|s_{n_1}\cdots s_{n_r}\right\rangle\!\left\langle s_{n_1}\cdots s_{n_r}\right| U^\dagger \circ$$

$$U^{n_1,\ldots,n_r} \rho^{n_1,\ldots,n_r} (U^{n_1,\ldots,n_r})^\dagger \Lambda \Bigg)$$

Substituting $U = \sum_{j=1}^{D^r} |\psi_j\rangle\!\langle e_j|$, $U^\dagger = \sum_{j'=1}^{D^r} \left|e_{j'}\right\rangle\!\left\langle \psi_{j'}\right|$, $U^{n_1,\ldots,n_r} = \sum_{k=1}^{D^r} \left|\psi_k\right\rangle\!\left\langle e_k^{n_1,\ldots,n_r}\right|$, $U^{n_1,\ldots,n_r} = \sum_{k'=1}^{D^r} \left|e_{k'}^{n_1,\ldots,n_r}\right\rangle\!\left\langle \psi_{k'}\right|$, and $\Lambda = \sum_{\ell=0}^{\infty} \ell \Pi_\ell$ into the above equation and applying some simplifications, we get

$$= \sum_{n_1,\ldots,n_r=1}^{D^r} p(n_1,\ldots,n_r) \sum_{j,j'=1}^{D^r} \langle e_j | s_{n_1} \cdots s_{n_r} \rangle \left\langle s_{n_1} \cdots s_{n_r} \middle| e_{j'} \right\rangle$$

$$\sum_{i=1}^{D^r} \lambda_i^{n_1,\ldots,n_r} \sum_{k,k'=1}^{D^r} \left\langle e_k^{n_1,\ldots,n_r} \middle| \lambda_i^{n_1,\ldots,n_r} \right\rangle \left\langle \lambda_i^{n_1,\ldots,n_r} \middle| e_{k'}^{n_1,\ldots,n_r} \right\rangle$$

$$\sum_{\ell=0}^{\infty} \ell \left\langle \psi_{j'}\psi_{k'} \middle| \Pi_\ell \left| \psi_j \psi_k \right\rangle \right.$$

Since the sequence $\{|\psi_r\rangle_{r=1}^{D^r}$ is a jointly orthonormal sequence, that causes $j = j'$ and $k = k'$. So, the average codeword length simplifies to

$$= \sum_{n_1,\ldots,n_r=1}^{D^r} p(n_1,\ldots,n_r) \sum_{j=1}^{D^r} \left| \langle e_j | s_{n_1} \cdots s_{n_r} \rangle \right|^2 \sum_{i=1}^{D^r} \lambda_i^{n_1,\ldots,n_r}$$

$$\sum_{k=1}^{D^r} \left| \left\langle e_k^{n_1\ldots,n_r} \middle| \lambda_i^{n_1,\ldots,n_r} \right\rangle \right|^2 (\ell_j + \ell_k)$$

$$= \sum_{n_1,\dots,n_r=1}^{D^r} p(n_1,\dots,n_r) \sum_{j=1}^{D^r} \ell_j \left|\langle e_j | s_{n_1}\cdots s_{n_r}\rangle\right|^2 +$$

$$\sum_{n_1,\dots,n_r=1}^{D^l} p(n_1,\dots,n_r) \sum_{j=1}^{D^l} \left|\langle e_j | s_{n_1}\cdots s_{n_r}\rangle\right|^2 \sum_{i=1}^{D^l} \lambda_i^{n_1,\dots,n_r}$$

$$\sum_{k=1}^{D^l} \ell_k \left|\left\langle e_k^{n_1,\dots,n_r} \middle| \lambda_i^{n_1,\dots,n_r}\right\rangle\right|^2$$

Using Birkhoff-von Neumann theorem, it can be shown that the above equation is minimized when $\left|e_i^{n_1,\dots,n_r}\right\rangle = \left|\lambda_i^{n_1,\dots,n_r}\right\rangle$ upto an overall phase factor for $1 \leq i \leq D^r$. So, the above equation simplifies to

$$= \sum_{n_1,\dots,n_r=1}^{D^l} p(n_1,\dots,n_r) \sum_{j=1}^{D^r} \ell_j \left|\langle e_j | s_{n_1,\dots,n_r}\rangle\right|^2 \; +$$

$$\sum_{n_1,\dots,n_r=1}^{D^l} p(n_1,\dots,n_r) \sum_{i=1}^{D^r} \ell_i \lambda_i^{n_1,\dots,n_r}$$

$$= \sum_{j=1}^{D^r} \ell_j \left\langle e_j \right| \Bigg(\sum_{n_1,\dots,n_r=1}^{D^r} p(n_1,\dots,n_r) \left|s_{n_1}\cdots s_{n_r}\right\rangle\!\left\langle s_{n_1}\cdots s_{n_r}\right| \Bigg) \left|e_j\right\rangle +$$

$$\sum_{n_1,\dots,n_r=1}^{D^r} p(n_1,\dots,n_r) \sum_{i=1}^{D^r} \ell_i \lambda_i^{n_1,\dots,n_r}$$

$$= \sum_{j=1}^{D^l} \ell_j \left\langle e_j \right| \rho_r \left|e_j\right\rangle + \sum_{n_1,\dots,n_r=1}^{D^r} p(n_1,\dots,n_r) \sum_{i=1}^{D^r} \ell_i \lambda_i^{n_1,\dots,n_r}$$

$$= \sum_{j,k=1}^{D^l} \ell_j \lambda_k \left|\langle e_j|\lambda_k\rangle\right|^2 + \sum_{n_1,\ldots,n_r=1}^{D^l} p(n_1,\ldots,n_r) \sum_{i=1}^{D^l} \ell_i \lambda_i^{n_1,\ldots,n_r}$$

Again, using Birkhoff-von Neumann theorem, the above equation is minimized when $|e_j\rangle = |\lambda_j\rangle$ upto an overall phase factor for $1 \leq j \leq D^l$. Thus, the equation reduces to

$$\sum_{j=1}^{D^r} \ell_j \lambda_j + \sum_{n_1,\ldots,n_r=1}^{D^r} p(n_1,\ldots,n_r) \sum_{i=1}^{D^r} \ell_i \lambda_i^{n_1,\ldots,n_r}$$

In general, for $m \in \mathbb{N}$ blocks, the equation is given by

$$\sum_{j=2}^{m} \left(\sum_{n_1,\ldots,n_{(j-1)r}}^{N} p(n_1,\ldots,n_{(j-1)r}) \sum_{i=1}^{D^r} \lambda_i^{n_1,\ldots,n_{(j-1)r}} \ell_i \right) + \sum_{i=1}^{D^r} \lambda_i \ell_i$$

Hence, the minimum average codeword length is obtained by using the set of $\{\ell_i\}_{i=1}^{D^r}$ that minimizes the above equation and satisfies $\sum_{i=1}^{D^r} 2^{-\ell_i} \leq 1$.

References

- Jozsa, Richard, et al. "Universal quantum information compression." Physical review letters 81.8 (1998): 1714.
- Kaltchenko, Alexei and Yang, En-Hui. (2003). "Universal Compression of Classically Correlated Stationary Quantum Sources." Proc SPIE. 5105. 10.1117/12.485663.
- Kaltchenko, Alexei, and Yang En-Hui. "Universal compression of ergodic quantum sources." arXiv:quant-ph/0302174 (2003).
- Schumacher, Benjamin, and Westmoreland, Michael D. "Indeterminate-length quantum coding." Physical Review A 64.4 (2001): 042304.
- Hayashi, Masahito, and Keiji Matsumoto. "Quantum universal variable-length source coding." Physical Review A 66.2 (2002): 022311.
- Hayashi, Masahito, and Keiji Matsumoto. "Simple construction of