Technical University of Crete
Department of Electrical and Computer Engineering

Επαναληπτικός Αλγόριθμος Συμπίεσης Κβαντικού Εναγκαλισμού

Iterative Compression Algorithm of Quantum Entanglement

BY
Konstantinos Vardakis

Supervisor:
Ellinas Demosthenes

Thesis Committee:
Angelakis Dimitris
Zervakis Michalis

JULY 2019

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία ασχολείται με ένα κεντρικό ερώτημα στο πεδίο της επιστήμης της Κβαντικής Πληροφορίας και Τεχνολογίας. Το αντικείμενο ενδιαφέροντος αποτελεί ο κβαντικός εναγκαλισμός (με άλλα λόγια κβαντικές συσχετίσεις) μεταξύ δύο τμημάτων ενός διμερούς κβαντικού συστήματος, καθένα από τα οποία περιγράφεται μαθηματικά από ένα σύνολο καταστατικών διανυσμάτων, το όποιο εμπεριέχεται σε ένα πραγματικό διανυσματικό χώρο. Έπειτα, το ερώτημα διατυπώνεται ως εξής: «Είναι δυνατόν να μειωθεί ο αριθμός των διανυσμάτων που περιγράφουν κάθε υποσύστημα του διμερούς κβαντικού συστήματος ενώ θα διατηρηθεί το ποσό του συνολικού κβαντικού εναγκαλισμού;» Για το λόγο αυτό, το ζητούμενο, εν συντομία *συμπίεση εναγκαλισμού*, περιλαμβάνει τη διαστατική ελάττωση των τοπικών υποδιαστημάτων, που απαρτίζουν το διμερές σύστημα, υπό τον περιορισμό της διατήρησης του αρχικού εναγκαλισμού. Με περιγραφικούς όρους το πρόβλημα μπορεί να διατυπωθεί με την εξής μορφή: «*μπορώ να έχω το ίδιο με λιγότερα;*» Δεδομένου ότι ο κβαντικός εναγκαλισμός θεωρείται ο περιζήτητος νέος τύπος πόρων που απαιτείται στη κβαντική τεχνολογία και το γεγονός ότι τα υποσυστήματα, που υπόκεινται τη διαστατική ελάττωση, αποτελούνται από πολλαπλά qubits, το ερώτημα συνεπώς ανάγεται στο βέλτιστο χειρισμό των πόρων. Βασισμένη σε προηγούμενες εργασίες και πρόσφατη έρευνα, η διπλωματική εργασία προχωρεί εκκινώντας με μία ευρετική ιδέα. Με την προϋπόθεση ότι η σύζευξη των υπό μελέτη υποσυστημάτων είναι μαθηματικά προσδιορισμένη από ένα πίνακα συντελεστών που καθορίζει το συνολικό καταστατικό διάνυσμα του διμερούς, γίνεται η εξής επιλογή: ως πίνακας συντελεστών επιλέγεται ένας οποιοσδήποτε πίνακας που περιγράφει μια grayscale ψηφιακή εικόνα. Ψηφιακές εικόνες όπως π.χ. Schroedinger, Lena χρησιμοποιούνται για την κατασκευή καταστατικών

διανυσμάτων διμερών κβαντικών συστημάτων πολλαπλών qubits. Αυτή η κατάσταση ευνοεί την εκμετάλλευση τεχνικών προσέγγισης πίνακα χαμηλής τάξης από το πεδίο της συμπίεσης εικόνας στο πεδίο της εφαρμογής της συμπίεσης εναγκαλισμού. Επιλέγοντας ως μέτρο ποσοτικοποίησης του εναγκαλισμού την κβαντική εντροπία Rényi του περιθώριου (ελαττωμένου) πίνακα πυκνότητας του διμερούς κβαντικού συστήματος, ο στόχος επαναδιατυπώνεται ως η επίτευξη της διαστατικής ελάττωσης του συνολικού διμερούς καταστατικού διανύσματος ενώ διατηρείται (ή βελτιστοποιείται) η Rényi εντροπία ενός κβαντικού υποσυστήματος. Η διπλωματική εργασία δείχνει ότι το πρώτο ζητούμενο, αυτό της διαστατικής ελάττωσης, επιλύεται μέσω της προσέγγισης χαμηλής τάξης της Διάσπασης Ιδιαζουσών Τιμών (SVD) του πίνακα καταστατικού διανύσματος εικόνας. Προς τούτο παρουσιάζεται μια κβαντική αλγοριθμική υλοποίηση της κλασσικής τεχνικής συμπίεσης στο κβαντικό πεδίο. Το δεύτερο ζητούμενο της συμπίεσης, αυτό της διατήρησης εναγκαλισμού, επιλύεται μέσω ενός αλγορίθμου με κβαντικό κύκλωμα παρόμοιο εκείνου της παραγωγής καταστάσεων Bell, γενικευμένο για την περίπτωση συστημάτων πολλαπλών qubits. Αυτό το τμήμα της διαδικασίας συμπίεσης εναγκαλισμού κάνει χρήση εργαλείων από την ανάλυση πινάκων, όπως το σημειακό γινόμενο στοιχείων πινάκων και σχετικών ανισοτήτων, προκειμένου να καταδείξει ότι υπάρχει μια ενάντια αναλογία μεταξύ των δύο ζητουμένων, συγκεκριμένα της διαστατικής ελάττωσης και της διατήρησης του εναγκαλισμού (εντροπίας). Αυτό οδηγεί στην ανάπτυξη μιας επαναληπτικής διαδικασίας, που περιλαμβάνει γιουνίταρι πύλες σε συνδυασμό με μαντειο-οδηγούμενες προβολές υψηλότερων διαστάσεων πάνω σε διανύσματα πολλαπλών qubit. Τέλος αναπτύσσεται μια λεπτομεριακή αριθμητική διερεύνηση, βασισμένη σε παραδειγματικές περιπτώσεις εικονο-καταστάσεων, η οποία επιβεβαιώνει την υλοποιησιμότητα και την αποδοτικότητα της επαναληπτικής διαδικασίας συμπίεσης κβαντικού εναγκαλισμού.

ABSTRACT

This thesis addresses a central question in the field of Quantum Information Science and Technology. The object of concern is the quantum entanglement (i.e. quantum correlations) between two parts of a bipartite quantum system, each of which is mathematically described by a set of state vectors, all lying in a real vector space. The question then is formulated as follows: "Is it possible to reduce the number of vectors describing each subsystem of the bipartite quantum system and still have the same amount of total quantum entanglement ?" Hence the question, abbreviated to the name *entanglement compression*, is that of dimensional reduction of local sub-spaces composing the bipartite system, under the constraint of preserving the initial entanglement. In descriptive terms, the problem may be cast in the form: "*can I have the same with less*?" Given that quantum entanglement is treated par excellence as the new type of resources required by quantum technology, and the fact that the sub-systems, aimed to be dimensionally reduced, are composed by multiple qubits, the question then amounts to one of optimal handling of resources. Building upon previous works and recent developments, the thesis then proceeds to exploit a heuristic idea. Given that the coupling of sub-systems under study is mathematically determined by a coefficient-matrix, specifying the multi-tensor state vector of the bi-partite, a choice is made: use for coefficient-matrix any matrix describing a grayscale digital image. Digital images of e.g. Schroedinger, Lena are employed to build state vectors of multi-qubit bi-partite quantum systems. This situation motivates the exploitation of low rank matrix approximation techniques from image compression within the context of entanglement compression. Employing as quantitative measure of entanglement the quantum Rényi entropy of the marginal (reduced) density matrix of the bipartite system, the aim becomes to achieve dimensional reduction of the total bipartite state vector, while preserving (or optimizing on) the Rényi entropy of a quantum subsystem. The thesis shows that the first task of dimensional reduction is achieved via low rank approximation in the Singular Value Decomposition (SVD) of the image-state-vector matrix. A quantum algorithmic implementation of the classical technique to the quantum context is provided. The second task of the compression, that of entanglement preservation, is achieved via an algorithm, akin to Bell state generation quantum circuit, generalized to the context of multi-qubit systems. This part of the entanglement compression procedure utilizes tools from matrix analysis, such as pair-wise Hadamard product of matrices and related inequalities, to show that there is a trading between the two tasks, namely dimensional reduction and entanglement (entropy) preservation. This leads to building an iterative procedure, which involves unitary gates combined with higher dimensional oracle-driven projections acting on multi-qubit vectors. A thorough numerical investigation, based on exemplary cases of image-states, confirms the feasibility and the efficiency of quantum entanglement compression iterative procedure.

Acknowledgements

First of all, I would like to express my sincere gratitude to my thesis supervisor, professor Demosthenes Ellinas, for the continuous support of my thesis as well as his patience and enthousiasm. He consistently steered me in the right direction whenever he thought I needed it.

In addition , I would like to express my deepest appreciation to associate professor Dimitris Angelakis and professor Michalis Zervakis for their participation in my thesis committee.

To my friends and roommates, thank you for encouraging me through this entire process.

Finally, I am greatly indebted to my parents, who supported me emotionally and financially over the course of my studies. This thesis stands as a testament to your unconditional love and constant motivation.

Contents

List of Figures

1 Quantum Information

1.1 Introduction

It all began with the "bit", the fundamental representation of information in the digital computer system. This became the cornerstone of computer science as a whole, with information theory and conventional computation methods benefiting greatly from its deterministic nature.

However, technological advancements have led to the development of the quantum computer system, which may exploit quantum effects to compute in ways that are faster or more efficient than, or even impossible, on conventional computers. Various quantum properties like supersposition and entanglement are the main assets of such systems, resulting in the exponential amount of parallelism in computation [2]. It should also be noted that quantum computation is not possible with the deterministic "bit", but with the probabilistic quantum "bit" or "qubit".

1.2 The qubit

The qubit (short for quantum bit) is the basic unit of quantum information. It is used to decribe the state of a quantum system and, as it turns out, quantum states behave mathematically in an analogous way to physical vectors [3]. For this reason, every qubit exists in a column vector form in a complex domain called Hilbert space $\mathcal{H}$. In this space, such vectors are called "kets", with the symbol $|\cdot\rangle$ according to the Dirac notation.

While a classical bit may be either in the state 0 or in the state 1, the qubit $|\psi\rangle$ may exist in the state $|0\rangle$ or in the state $|1\rangle$, but may also exist in a superposition state $a\,|0\rangle + b\,|1\rangle$, with a,b being complex numbers. In this way, the values of such qubits do not exist in a real discrete domain prior to measurement, but in a continuous one in the form of $\mathbb{C}^2$. However, when the qubit $|\psi\rangle$ is measured, it is going to be found either in the state $|0\rangle$ or in the state $|1\rangle$. According to the laws of quantum mechanics, the modulus squared of a,b corresponds to the probability of finding the qubit in state $|0\rangle$ or $|1\rangle$ respectively [3], meaning that:

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \tag{1}$$

with $p_{|0\rangle} = |a|^2$, $p_{|1\rangle} = |b|^2$.

Since the states $|0\rangle$, $|1\rangle$ are the only possible outcomes after the measurement, the coefficients a,b are constrained by the requirement that $|a|^2 + |b|^2 = 1$,

which is atributed to the completeness theorem. In general, if a qubit $|\psi\rangle$ may be found in one out of n possible states $|u_0\rangle$, $|u_1\rangle$,...,$|u_n\rangle$ after measurement, then:

$$|\psi\rangle = c_0 |u_0\rangle + c_1 |u_1\rangle + ... + c_n |u_n\rangle = \sum_{i=1}^{n} c_i |u_i\rangle$$

with $\sum_{i=1}^{n} |c_i|^2 = 1$.

In this case, it can be seen that the qubit $|\psi\rangle$ is expressed in terms of the states $|0\rangle$, $|1\rangle$. These state vectors form a two-dimensional orthonormal basis, with their vector representations being $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$.As a result, the equivalent vector representation for $|\psi\rangle$ (eq. 1) is:

$$|\psi\rangle = a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}a\\0\end{pmatrix} + \begin{pmatrix}0\\b\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix}$$

Even though there are numerous orthonormal basis, the basis $\{|0\rangle, |1\rangle\}$ is exclusively used in this thesis.

1.3 The Hilbert space

Being a vector domain, the Hilbert space $\mathcal{H}$ has its dual vector domain called $\mathcal{H}^*$, whose elements are row vectors and are called "bras", with the symbol $\langle\cdot|$. It can be seen that there is a $1-1$ correspondence between those two vectors spaces, resulting in the fact that, for every ket $|\psi\rangle \in \mathcal{H}$, there is the bra $\langle\psi| \in \mathcal{H}^*$, with $\langle\psi| = (|\psi\rangle)^\dagger = ((|\psi\rangle)^*)^T$ [16].

Several properties regarding Hilbert space $\mathcal{H}$ are shown below:

·*Linearity*

For any state vectors $|a\rangle$, $|b\rangle$ of the Hilbert space $\mathcal{H}$, every linear combination $c_1|a\rangle + c_2|b\rangle$, with c_1,c_2 being complex numbers, also belongs in the same Hilbert space $\mathcal{H}$.

·*Inner Product*

The inner product of the state vectors $|a\rangle$, $|b\rangle$ is a complex number and is known with the symbol $\langle a|\,|b\rangle$, called bra-ket.

Some useful properties are the following:

→ $\langle a|\,|b\rangle = \langle b|\,|a\rangle^*$

→ $\langle a|\,|\lambda b\rangle = \lambda\,\langle a|\,|b\rangle$ for every complex number λ

→ $\langle \lambda a|\,|b\rangle = \lambda^*\,\langle a|\,|b\rangle$ for every complex number λ

→ $\langle a + c|\,|b\rangle = \langle a|\,|b\rangle + \langle c|\,|b\rangle$

$\rightarrow \langle a|\,|b+c\rangle = \langle a|\,|b\rangle + \langle a|\,|c\rangle$
$\rightarrow \langle a|\,|a\rangle \geq 0$ (refers to the norm of the vector a: $\langle a|\,|a\rangle = \|a\|$)
$\rightarrow \langle a|\,|a\rangle = 0 \Longleftrightarrow |a\rangle = 0$

·*Vector Orthogonality*

Two state vectors $|a\rangle$ and $|b\rangle$ of the Hilbert space $\mathcal{H}$ are orthogonal iff $\langle a|\,|b\rangle = 0$.

·*Orthonormal basis*

A set of state vectors $\{|e_i\rangle\}$ forms an orthonormal basis of Hilbert space $\mathcal{H}$, iff they are all unit vectors (meaning $\|e_i\| = 1$ for every i) and they are orthogonal to each other. Every vector $|a\rangle$ with $|a\rangle \in \mathcal{H}$ may be expressed as a linear combination of the basis state vectors of that Hilbert space $\mathcal{H}$, as shown below:

$$|a\rangle = \sum_{i=1}^{\infty} c_i\,|e_i\rangle$$

·*First postulate of quantum information*

Every property of a quantum system can be described by a normalized vector $|\psi\rangle$ of the Hilbert space $\mathcal{H}$, meaning that $\langle\psi|\,|\psi\rangle = \|\psi\| = 1$.

·*Superposition of states*

Let $|a\rangle$, $|b\rangle$ be two linearly independent normalized state vectors in $\mathcal{H}$. Every linear combination $|\psi\rangle = c_1\,|a\rangle + c_2\,|b\rangle$, with c_1, $c_2 \in \mathbb{C}$,

results in another normalized state vector of the Hilbert space $\mathcal{H}$.

1.4 Multiple qubits

Much like classical information, the need arises for the study of more than one qubits at a single time in the field of quantum information. So far, a two-dimensional Hilbert space $\mathcal{H}$ is enough for the description of quantum states $|\psi\rangle$ of a single qubit. However, in order to describe composite quantum systems of two (or more) qubits, a more complex Hilbert space $\mathcal{H}$ of a greater dimension is needed. Such systems may be expressed in terms of their components with the use of the Kronecker product $\otimes$ (known as tensor product in quantum information). This type of product has the following properties for random state vectors $|x\rangle$, $|y\rangle$, $|p\rangle$, $|q\rangle$:

- $(|x\rangle \otimes |y\rangle) \cdot (|p\rangle \otimes |q\rangle) = (|x\rangle \cdot |p\rangle) \otimes (|y\rangle \cdot |q\rangle)$
- $|x\rangle \otimes (|y\rangle + |p\rangle) = |x\rangle \otimes |y\rangle + |x\rangle \otimes |p\rangle$
- $(|x\rangle + |y\rangle) \otimes |p\rangle = |x\rangle \otimes |p\rangle + |y\rangle \otimes |p\rangle$

- $(a\,|x\rangle) \otimes |y\rangle = |x\rangle \otimes (a\,|y\rangle) = a(|x\rangle \otimes |y\rangle)$ for any scalar $a \in \mathbb{C}$

The same properties apply to any choice of matrices A, B, C, D, even combined with vectors. For example: $(A \otimes B) \cdot (|p\rangle \otimes |q\rangle) = (A\,|p\rangle) \otimes (B\,|q\rangle)$

So, the combination of two (or more) subsystems into a composite one $|\psi\rangle$ is primarily attributed to the tensor product. However, the expression of the state vector $|\psi\rangle$ may be either factorized or un-factorized with respect to its subsystems. These two cases are presented below:

Case #1: *Factorized/ Uncorrelated/ Un-entangled quantum states*

Let there be a composite quantum system $|\psi\rangle$ that consists of two qubits $|\psi_1\rangle = a_1\,|0\rangle + b_1\,|1\rangle = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$, $|\psi_2\rangle = a_2\,|0\rangle + b_2\,|1\rangle = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$ with a_1, b_1, a_2, $b_2 \in \mathbb{C}$. Then:

$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$
$|\psi\rangle = (a_1\,|0\rangle + b_1\,|1\rangle) \otimes (a_2\,|0\rangle + b_2\,|1\rangle)$
$|\psi\rangle = a_1a_2(|0\rangle \otimes |0\rangle) + a_1b_2(|0\rangle \otimes |1\rangle) + b_1a_2(|1\rangle \otimes |0\rangle) + b_1b_2(|1\rangle \otimes |1\rangle)$
$|\psi\rangle = a_1a_2\,|00\rangle + a_1b_2\,|01\rangle + b_1a_2\,|10\rangle + b_1b_2\,|11\rangle$

It should be noted that the term $|0\rangle \otimes |1\rangle$ is the same as $|01\rangle$. This substitution takes place on the grounds of abbreviation. In general: $|i\rangle \otimes |j\rangle = |i\rangle\,|j\rangle = |ij\rangle$ for every (i, j).

Using their respective vector forms, it can be seen that:

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \otimes \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 \\ a_1b_2 \\ b_1a_2 \\ b_1b_2 \end{pmatrix}$$

Case #2: *Un-factorized/ Correlated/ Entangled quantum states*

In this case, $|\psi\rangle \neq |\psi_1\rangle \otimes |\psi_2\rangle$. However, certain states of $|\psi_1\rangle$ may be correlated with specific states of $|\psi_2\rangle$, as shown in the examples below:

e.g. $|\psi\rangle = a\,|0\rangle \otimes |0\rangle + b\,|1\rangle \otimes |1\rangle$ with $a, b \in \mathbb{C}$

$$|\psi\rangle = a[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}] + b[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}]$$

$$|\psi\rangle = a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \\ b \end{pmatrix}$$

or

e.g. $|\psi\rangle = a\,|0\rangle \otimes |1\rangle + b\,|1\rangle \otimes |0\rangle$ with $a, b \in \mathbb{C}$

$$|\psi\rangle = a[\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix}] + b[\begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix}]$$

$$|\psi\rangle = a\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + b\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \begin{pmatrix}0\\a\\b\\0\end{pmatrix}$$

In any case, the state vector $|\psi\rangle$ describes a quantum system and its squared coefficients should correspond to the probabilities of finding the system in one of these states upon measurement. So:

case #1: $\|a_1a_2\|^2 + \|a_1b_2\|^2 + \|b_1a_2\|^2 + \|b_1b_2\|^2 = 1$

case #2: $\|a\|^2 + \|b\|^2 = 1$

As for the respective Hilbert space $\mathcal{H}$ of $|\psi\rangle$, it is the outcome of the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ as well. In these cases, both of the component Hilbert spaces involved are two-dimensional ones, while the orthonormal basis for $|\psi\rangle$ is clearly the four-dimensional one, whose basis vectors are presented below:

$$|00\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}; |01\rangle = |0\rangle \otimes |1\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}$$

$$|10\rangle = |1\rangle \otimes |0\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}; |11\rangle = |1\rangle \otimes |1\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

In general, a quantum state of n two-dimensional qubits exists in a Hilbert space $\mathcal{H}$ of 2^n dimensions, with the basis vectors needed being also 2^n. This exponential growth of the domain dimensions, as a result of the linear increase in the number of qubits, is the main reason for the exponential increase in speed of quantum computations [16].

It can be seen that the distinction between factorized and un-factorized quantum states $|\psi\rangle$ is inextricably linked to the degree of correlation between possible states of different subsystems. This property of un-factorized systems is better known as quantum entanglement. The qualitative assessment of the entanglement of a system involves the determinant of the square matrix containing the coefficients of every possible state the system may be found in. If the determinant is equal to zero, then the system is factorized (un-entangled), whereas a non-zero determinant implies that the system is entangled to a certain extent. This specific procedure, regarding the abovementioned cases, is presented below:

Case #1: Un-entangled system

$|\psi\rangle = a_1a_2\,|00\rangle + a_1b_2\,|01\rangle + b_1a_2\,|10\rangle + b_1b_2\,|11\rangle$

In this case, the square matrix of the coefficients of $|\psi\rangle = c_{00}\,|00\rangle + c_{01}\,|01\rangle + c_{10}\,|10\rangle + c_{11}\,|11\rangle$ is:

$$A = \begin{bmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{bmatrix} = \begin{bmatrix} a_1 a_2 & a_1 b_2 \\ b_1 a_2 & b_1 b_2 \end{bmatrix}$$

The determinant of this matrix is:
$\det(A) = c_{00}c_{11} - c_{01}c_{10}$
$\det(A) = a_1 a_2 b_1 b_2 - a_1 b_2 b_1 a_2$

$$\det(A) = 0$$

Case #2: Entangled system

$|\psi\rangle = a\,|00\rangle + b\,|11\rangle$
The square matrix of the coefficients of $|\psi\rangle = c_{00}\,|00\rangle + c_{01}\,|01\rangle + c_{10}\,|10\rangle + c_{11}\,|11\rangle$ in this case is:

$$A = \begin{bmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

So, the determinant is obviously: $\det(A) = a \cdot b \neq 0$ when $a, b \neq 0$.

In cases of correlated systems, the amount of entanglement depends on the values of the coefficients *a*,*b*. However, the issue of the quantitative assessment of the entanglement will be addressed later.

1.5 Bipartite quantum system

In the previous section, it was seen that composite quantum systems $|\psi\rangle$ may consist of more than one qubits. The simplest example of such a system has two qubits and it is known as a bipartite quantum system. In particular, the expression of a bipartite quantum system $|\psi\rangle$ with reference to its two distinct qubit components $|\phi\rangle$ and $|x\rangle$ is the following:

$$|\psi\rangle = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} c_{ij}\,|\phi_i\rangle \otimes |x_j\rangle$$

with $c_{ij} \in \mathbb{C}$ and $\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} |c_{ij}|^2 = 1$.

However, this definition may be extended to quantum systems that have more than two qubits (multipartite systems). In this case, a bipartite system $|\psi\rangle$ consists of two quantum subsystems $|\phi\rangle$ and $|x\rangle$, each one being a system of two or more qubits. There are a lot of different interpretations of the same

multipartite system as a bipartite one, due to the various scenarios of two complementary sets of qubits. Such consideration is beneficial to the study of the correlation (thus entanglement) between these two sets of qubits.

1.6 Transformations and operators on quantum states

An operator is a mathematical rule that can be applied to a function to transform it into another function [3]. Technically, an operator A is a matrix that acts on a state vector $|\psi\rangle$ of the Hilbert space $\mathcal{H}$, resulting in another state vector $\left|\psi'\right\rangle$ of the same domain:

$$\left|\psi'\right\rangle = A\,|\psi\rangle$$

with $\left|\psi'\right\rangle$ also in $\mathcal{H}$.

Operators can also act on bras, with the result being another bra:

$$\left\langle\psi'\right| = \langle\psi|\,B$$

with $\langle\psi|$, $\left\langle\psi'\right|$ in $\mathcal{H}^*$.

Every operator has a matrix representation in respect to a given orthonormal basis, thus making its action on a state vector a simple matter of matrix multiplication. For the purposes of this thesis, the matrix representation of every operator is expressed in terms of the basis $\{|0\rangle, |1\rangle\}$.

It should be mentioned that the entirety of the operators used in this project are linear operators. An operator A is linear, if the following relationship holds, given complex numbers a_i and the state vectors $|u_i\rangle$:

$$A(\sum_{i=1}^{n} a_i\,|u_i\rangle) = \sum_{i=1}^{n} a_i(A\,|u_i\rangle)$$

with $a_i \in \mathbb{C}$.

Some basic linear operators are shown below:

·*Zero operator* $\mathbb{O}$
$\mathbb{O}\,|a\rangle = |a\rangle$ for every $|a\rangle \in \mathcal{H}$

·*Identity operator* $\mathbb{I}$
$\mathbb{I}\,|a\rangle = |a\rangle$ for every $|a\rangle \in \mathcal{H}$

·*Inverse operator* T^{-1}
$T\,|a\rangle = |b\rangle$ then $T^{-1}\,|b\rangle = |a\rangle$ for every $|a\rangle$,$|b\rangle \in \mathcal{H}$

It can be seen that the application of the inverse operator T^{-1} neutralizes the effect of T:

$$T^{-1}T\,|a\rangle = T^{-1}\,|b\rangle = |a\rangle$$

and

$$TT^{-1}\,|b\rangle = T\,|a\rangle = b$$

So the concecutive application of T^{-1} and T is equivalent to the application of the identity operator:

$$T^{-1}T = TT^{-1} = \mathbb{I}$$

·*Unitary operator U*

An operator U is unitary iff its inverse is equal to its conjugate tranpose, meaning that $U^{-1} = (U^*)^T = U^\dagger$. A few notable properties of the unitary operators include the following:

– The product of unitary operators is a unitary operator.
– The tranpose of a unitary operator is a unitary operator.
– The inverse of a unitary operator is a unitary operator.
– Every unitary operator preserves the norm of the vector it acts on.

$$\||a\rangle\| = \|U\,|a\rangle\|$$

·Hermitian operator T

An operator T is hermitian iff it is equal to its conjugate tranpose one, meaning $T = (T^*)^T = T^\dagger$.

Some useful properties are:

– The addition of two hermitian operators is hermitian.
– The product of hermitian operators is hermitian.
– The eigenvalues of a hermitian operator are real numbers.

·Projection operator P

Suppose that a given vector space has n dimensions and a basis given by orthonormal state vectors $|i\rangle$ with $1 \leq i \leq n$. The projection operator $P = \sum_{i=1}^{m} |i\rangle\langle i|$ (with $m < n$) projects any state vector $|\psi\rangle$ onto the subspace spanned by the set $|i\rangle$ with $1 \leq i \leq m$.

Properties of the projection operator include the following:

– For every projection operator P, it stands $P^2 = P$.
– Every projection operator has eigenvalues 0 and 1.

Let two complementary projection operators $P_\uparrow = \sum_{i=1}^{m} |i\rangle\langle i|$ (with $m < n$) and $P_\downarrow = \sum_{i=m+1}^{n} |i\rangle\langle i|$. It can be seen that:

$$\rightarrow P_{\uparrow} + P_{\downarrow} = \sum_{i=1}^{m} |i\rangle \langle i| + \sum_{i=m+1}^{n} |i\rangle \langle i| = \sum_{i=1}^{n} |i\rangle \langle i| = \mathbb{I}$$
$$\rightarrow P_{\uparrow} \cdot P_{\downarrow} = 0$$

The complementary projection operators of basis $\{|0\rangle, |1\rangle\}$ are:

$$\begin{aligned} P_0 &= |0\rangle \langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ P_1 &= |1\rangle \langle 1| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

In addition, an operator A may be applied in an "adjoint" way $Ad(A)$ on a system B, meaning that $Ad(A)B = ABA^{\dagger}$. This type of action serves only abbreviation purposes. A useful property of this notation is:

$$Ad(A)Ad(B)C = Ad(AB)C$$

According to the second postulate of quantum mechanics, every dynamical variable A that is a physically measurable quantity has an operator A in the Hilbert space $\mathcal{H}$ corresponding to it. The set of eigenvectors of such an operator form an orthonormal basis and its eigenvalues are real numbers. That is the reason for the Hermitian operators being used exclusively in the field of quantum information.

1.7 Quantum gates: 1-qubit and 2-qubit cases

In quantum information, gates acting on qubits are represented by hermitian and unitary operators. Unlike classical logic gates, quantum gates need to be reversible, guaranteed by its hermiticity, and they need to preserve the norm of their input quantum state vector $|\psi\rangle$, leading to the prerequisite of unitarity. Quantum gates may act on a single qubit, or more qubits at once.

A set of quantum single-qubit gates that have fundamental importance in quantum computation are the Pauli gates, named after the physicist Wolfgang Pauli. The matrix representation of these gates are expressed with reference to the orthonormal basis $\{|0\rangle, |1\rangle\}$ as:

$\sigma_1 = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\sigma_2 = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$,$\sigma_3 = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

The X gate is the quantum NOT gate, leading to the reverse of any quantum state $|\psi\rangle = a|0\rangle + b|1\rangle$ as shown below:

$X|\psi\rangle = X(a|0\rangle + b|1\rangle) = X(a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix}) = X(\begin{pmatrix}a\\0\end{pmatrix} + \begin{pmatrix}0\\b\end{pmatrix}) = X\begin{pmatrix}a\\b\end{pmatrix}$

$X|\psi\rangle = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}b\\a\end{pmatrix} = \begin{pmatrix}b\\0\end{pmatrix} + \begin{pmatrix}0\\a\end{pmatrix} = b\begin{pmatrix}1\\0\end{pmatrix} + a\begin{pmatrix}0\\1\end{pmatrix}$

So:

$$X(a|0\rangle + b|1\rangle) = b|0\rangle + a|1\rangle$$

Another important single-qubit gate is called the Hadamard operator, whose matrix representation in the basis $\{|0\rangle, |1\rangle\}$ is:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$$

This gate transforms the basis state vector $|0\rangle$ and $|1\rangle$ into the superposition states $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$ respectively.

Also, the two-qubit gate that is the cornerstone of every quantum computer system is the "controlled-not" gate U_{cn}. As its name suggests, its operation is actually a controlled application of the X gate over the second Hilbert space, depending on the value of the target qubit of the first Hilbert space.

$$U_{cn}(|x\rangle \otimes |y\rangle) = |x\rangle \otimes |x \oplus_2 y\rangle$$

In particular, this two-qubit gate "flips" the second qubit, if the first qubit is equal to $|1\rangle$, and it does nothing to the second qubit, if the first one is equal to $|0\rangle$. This procedure may be expressed in terms of unitary operators acting on the respective Hilbert space as shown below:

$U_{cn} = P_0 \otimes \mathbb{I} + P_1 \otimes X$

$U_{cn} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} \otimes \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} \otimes \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$

$$U_{cn} = \begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 1\\0 & 0 & 1 & 0\end{bmatrix}$$

This is the matrix representation of the U_{cn} operator with reference to the orthonormal basis $\{|0\rangle, |1\rangle\}$. The explicit formula for the matrix representation of this gate acting on the linear combination of the two-qubits $|x\rangle$ and $|\psi\rangle$ is:

$$U_{cn} = \sum_{x=0}^{1}\sum_{y=0}^{1}(|x\rangle \otimes |x \oplus_2 y\rangle)(\langle x| \otimes \langle y|) = \sum_{x=0}^{1}\sum_{y=0}^{1}|x\rangle\langle x| \otimes |x \oplus_2 y\rangle\langle y|$$

The "controlled-not" gate (or quantum XOR-gate) can be used for many practical tasks of quantum information processing such as quantum state swapping, entangling quantum states, performing Bell measurements, dense coding, quantum teleportation and more.

1.8 Density operator

In many cases, the need arises for the study of a composite quantum system that is comprised of a number of different systems. Such a collection is known as an *ensemble*. Each member of the ensemble may be found in one of two or more different quantum states, according to their given probabilities. However, there is also the probability of the composite system being found in one of the members of this ensemble. In the latter level, the use of probability is acting on a classical way, thus reflecting a simple statistical mixture of quantum systems.

In such cases, the calculation of usual quantities for this composite quantum system becomes more complex, since the mathematical rules mentioned so far are not enough to bypass the restrictions imposed by the two levels of probabilities. The operator used for the in-depth study of any quantum system is known as the density operator.

So, let there be a quantum system $|\psi\rangle$ in the Hilbert space $\mathcal{H}$ that may be found in one of the states $|\psi_i\rangle$ with respective probabilities p_i. The density operator ρ is defined as shown below:

$$\rho = \sum_i p_i \, |\psi_i\rangle \, \langle\psi_i|$$

The density operator ρ provides a greater insight to the quantum system, since it makes its study focus around the application of operators, whose use is more practical than that of individual state vectors.

So, an operator ρ is the density operator of an ensemble $\{|\psi_i\rangle, p_i\}$, iff the following requirements are met:

– A density operator ρ is Hermitian, meaning that $\rho = \rho^\dagger$.

– A density operator ρ is Positive, meaning that $\langle\psi| \, \rho \, |\psi\rangle \geq 0$ for any state vector $|\psi\rangle$.

– $Tr(\rho) = 1$. This property is attributed to the completeness theorem.

It should be mentioned that, for the purposes of this thesis, the term "density matrix" is used instead of "density operator".

Based on their formation, an ensemble may be put into one of two categories:

– The ensemble with a single quantum system (described by $|\psi\rangle$) does not have the probabilistic feature of the statistical mixture mentioned above. In these cases, it is said that the composite system is in a pure state and the respective density matrix ρ is expressed as $\rho = |\psi\rangle \, \langle\psi|$.

– The ensemble with more than one quantum system is the general case of a composite system. In these cases, it is said that the system is in a mixed state.

It can be seen that, for systems in a pure state, their respective density matrix ρ has the additional property of $\rho^2 = \rho$, thus becoming a projection operator. This makes the identification of such systems much easier, since $\rho^2 \neq \rho$ for any other quantum system.

Several important properties of the density matrix include the following:

- If ρ is a density matrix, then ρ^T is also a density matrix.

– If ρ_a is a density matrix on a Hilbert space $\mathcal{H}_a$ and ρ_b is a density matrix on a Hilbert space $\mathcal{H}_b$, then $\rho_a \otimes \rho_b$ is also a density matrix on the Hilbert space $\mathcal{H}_a \otimes \mathcal{H}_b$.

- The action of a unitary operator U on an ensemble of systems may be expressed in terms of the respective density matrix ρ as: $\rho \rightarrow U\rho U^\dagger$.

1.8.1 Reduced density operator

Perhaps the deepest application of the density operator is as a descriptive tool for subsystems of a composite system. Such a description is provided by the reduced density operator [1].

Let there be a bipartite system with density matrix:

$$\rho = \sum_x \sum_y \sum_k \sum_l \rho_{xykl} \left|x\right\rangle \left\langle y\right| \otimes \left|k\right\rangle \left\langle l\right|$$

A reduced density matrix ρ_i is produced by the application of trace over all of the basis states of the other Hilbert space alone. So, in this case, this process may yield two reduced density matrices ρ_1 and ρ_2, depending on the perspective, as shown below:

$$\begin{aligned} \rho_1 &\equiv Tr_2(\rho) \\ \rho_2 &\equiv Tr_1(\rho) \end{aligned}$$

The partial tracing involved in the calculation of ρ_1 is presented below:

$\rho_1 = Tr_2(\rho)$

$\rho_1 = \sum_x \sum_y \sum_k \sum_l \rho_{xykl} \left|x\right\rangle \left\langle y\right| \cdot Tr(\left|k\right\rangle \left\langle l\right|)$

$\rho_1 = \sum_x \sum_y \sum_k \sum_l \rho_{xykl} \left|x\right\rangle \left\langle y\right| \cdot \left\langle l\right| \left|k\right\rangle$

$$\rho_1 = \sum_x \sum_y \left|x\right\rangle \left\langle y\right| \cdot (\sum_k \rho_{xykk})$$

As for the reduced density matrix ρ_2:

$$\rho_2 = \sum_k \sum_l \left|k\right\rangle \left\langle l\right| \cdot (\sum_x \rho_{xxkl})$$

The reduced density operator is virtually indispensable in the analysis of composite quantum systems. The partial trace operation involved is the unique operation that gives rise to the correct description of observable quantities for subsystems of a composite system [1].

1.9 Quantum entanglement measures

From qualitative to quantitative measures:

One of the most unusual and fascinating aspects of quantum mechanics is the fact that particles or systems can become entangled. The fact that two quantum systems (say A and B) are entangled, means that the values of certain properties of system A are correlated with the values that those properties will assume for system B [3]. This effect is preserved even when the two systems are spatially separated, no matter the distance. That is why this phenomenon was initially perceived as "spooky action at a distance".

In an entangled quantum system, any measurement upon one of the component subspaces provides accurate information about the state of the other subsystem, without the need for any measurements upon it. It should be noted that the concept of entanglement is of paramount importance to quantum information theory, since an entangled quantum system is capable of providing more information than its subsystems combined. Due to this feature, even greater transmission rates and faster information processing in this field are possible [16].

In a previous section, it was seen that a quantum system is considered to be entangled as long as its state vector $|\psi\rangle$ cannot be expressed as the tensor product of the state vectors of its subsystems. The "factoriz-ability" of a given system (known as separability) may be assessed directly by the value of the determinant of the matrix of its coefficents. In particular, a determinant that is equal to zero implies that the system is un-entangled (separable), whereas a non-zero determinant is attributed to an entangled system. However, the exact value of the non-zero determinant corresponds to the degree of factorization of the composite quantum system $|\psi\rangle$, which is inextricably linked to the amount of entanglement this system has. The assessment of the amount of entanglement in any system is attributed to the concept of quantum entropy.

1.10 Quantum entropy

The entropy characterizes a degree of disorder in systems with fluctuating physical observables [18]. The interpretation of order is related to the statistical properties of the system. Such properties of classical and quantum systems are described within the framework of the probability-distribution formalism in classical domain and the density matrices in quantum domain. For example, the Shannon entropy determines the characteristics of classical states based on a probability distribution, while the von Neumann entropy represents the amount of entanglement that is associated with a quantum-state density matrix. In particular, the formula for the von Neumann entropy is:

$$S(\rho) = -Tr(\rho \log \rho)$$

The association of the density matrix ρ with the concept of quantum entropy is attributed to the fact that its diagonal elements form a probability distribution, regarding the possible states that the respective system may be found in. The same distribution is also formed by its respective eigenvalues λ_i that correspond to its eigendecomposition:

$$\rho = UDU^\dagger = U \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} U^\dagger$$

with U being the unitary matrix of its eigenvectors. Therefore, the von Neumann quantum entropy may as well be expressed as:

$$S(\rho) = -\sum_{i=1}^{n} \lambda_i \log \lambda_i$$

It should be mentioned that several generalizations, regarding the mathematical formulas used for the calculation of quantum entropy, have been proposed over the years. One of them is the quantum Rényi Entropy, whose formula is used thoughout this project.

1.10.1 Rényi entropy

The Rényi Entropy, developed by Hungarian mathematician Alfred Rényi, was introduced on axiomatic grounds as a generalization of Shannon entropy. For a discrete probability distribution $P = (p_1, ..., p_n)$, the Rényi Entropy of order a ($a \geq 0$) is defined as:

$$H_a(P) = \frac{1}{1-a} \log(\sum_{i=1}^{n} p_i^a)$$

This definition may be extended to continuous random variables X by

$$H_a(X) = \frac{1}{1-a} \log(\int f_X(x)^a dx)$$

Unlike the discrete case, the value of the Rényi Entropy may be negative for continuous random variables. Therefore, the Rényi Entropy is typically only used for cases of discrete variables.

Several special cases based on its order a are mentioned below:

$\rightarrow a = 0$

The Rényi Entropy H_0 of order $a = 0$ is known as the max-entropy or Hartley Entropy. It is obvious that is is equal to the logarithm of the number of probabilities regarding the given classical system.

$\rightarrow a = 1$

In this case, the expression of the Rényi entropy is perceived as the limit of a approaching 1. So:

$$H_1(P) = \lim_{a \to 1} (\frac{1}{1-a} \log(\sum_{i=1}^{n} p_i^a))$$

It should be noted that the Rényi Entropy of order $a = 1$ results in the Shannon Entropy, as:

$$H_1(P) = \lim_{a \to 1} (\frac{1}{1-a} \log(\sum_{i=1}^{n} p_i^a)) = -(\sum_{i=1}^{n} p_i \log p_i)$$

So, the Rényi Entropy may be thought of as a more fundamental concept of which Shannon Entropy is just an important special case.

$\rightarrow a = \infty$

As in the case of $a = 1$, the Rényi Entropy of order $a = \infty$ is perceived as the limit of a approaching ∞. So:

$$H_\infty(P) = \lim_{a \to \infty} (\frac{1}{1-a} \log(\sum_{i=1}^{n} p_i^a)) = -\log(\max\{p_i\})$$

This case is known as the min-entropy, since it converges to the negative logarithm of the probability of the most probable outcome.

A few notable theorems regarding the classical Rényi Entropy are presented below:

·"The Rényi Entropy $H_a(P)$ is a continuous function is P for $a > 1$ and discontinuous for $a \leq 1$"

·"Let $a \in (0, 1]$ and let P be a probability distribution over $\mathbb{Z}_+$.Then there exists a sequence of distributions P_n converging to P with respect to the total variation distance, such that, for arbitrary $r \in [0, \infty]$,

$$\lim_{n \to \infty} H_a(P_n) = H_a(P) + r$$

·"$H_a(P)$ is a convex function in P for $a \leq 1$ and is neither convex nor concave for $a > 1$."

In the quantum setting, the Rényi Entropy S of order a is given as:

$$S_a(\rho) = \frac{1}{1-a}\log(Tr(\rho^a))$$

with ρ being the density matrix of the system. The special cases of order a of this definition are:

$\rightarrow$ The quantum Rényi Entropy of order $a = 1$ converts to the von Neumann entropy as:

$$S_1(\rho) = \lim_{a\to 1}(\frac{1}{1-a}\log(Tr(\rho^a))) = -Tr(\rho\log\rho)$$

$\rightarrow$The Rényi Entropy of order $a = \infty$ results in the min-entropy as well, since:

$$S_\infty(\rho) = \lim_{a\to\infty}(\frac{1}{1-a}\log(Tr(\rho^a))) = -\log(\|\rho\|)$$

with $\|\cdot\|$ denoting the operator norm.

A complication regarding the quantum domain is that the values for the entropy must usually range between 0 (pure state) and 1(maximally entangled). However, the dimension n of the reduced density matrix ρ_1 may be greater than 2, resulting in the entropy becoming greater than 1. Therefore, the base of the logarithm used in the entropy formula is n, thus limiting the value of the Rényi Entropy between 0 and 1.

It should be mentioned that the phenomenon of entanglement is a specific particularity of strong quantum correlations, some aspects of which can be characterized by the values of entropies of the subsystem states [18]. Therefore, the calculation of entropy and the assessment of the amount of entanglement of a bipartite quantum system occurs from the perspective of a chosen subsystem with the use of a reduced density matrix.

1.11 Schmidt decomposition of quantum states

Considering any orthonormal basis $\{|k\rangle\}_{k=1}^{d_1}$ of $\mathcal{H}_1$ and any orthonormal basis $\{|l\rangle\}_{l=1}^{d_2}$ of $\mathcal{H}_2$, a bipartite quantum system $|\psi\rangle$ in $\mathcal{H} = \mathcal{H}_1\otimes\mathcal{H}_2$ may be expressed as:

$$|\psi\rangle = \sum_{k=1}^{d_1}\sum_{l=1}^{d_2} c_{kl}\,|k\rangle\otimes|l\rangle$$

with $c_{kl} \in \mathbb{C}$ being the elements of the matrix of coefficients C. It is obvious that the size of this matrix is $d_1 \cdot d_2$. The singular value decomposition $UDV^\dagger$ of matrix C results in $c_{kl} = \sum_{i=1}^{r} U_{ki} D_{ii} (V^\dagger)_{il}$, where D is a diagonal matrix with non-negative elements, U,V are unitary matrices and r is the rank of matrix C. So, by substitution:

$$|\psi\rangle = \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} \sum_{i=1}^{r} U_{ki} D_{ii} (V^\dagger)_{il} |k\rangle \otimes |l\rangle$$

$$|\psi\rangle = \sum_{i=1}^{r} D_{ii} (\sum_{k=1}^{d_1} U_{ki} |k\rangle) \otimes (\sum_{l=1}^{d_2} (V^\dagger)_{il} |l\rangle)$$

Define $|i_1\rangle \equiv \sum_{k=1}^{d_1} U_{ki} |k\rangle$, $|i_2\rangle \equiv \sum_{l=1}^{d_2} (V^\dagger)_{il} |l\rangle$ and $\lambda_i \equiv D_{ii}$. Then:

$$|\psi\rangle = \sum_{i=1}^{r} \lambda_i |i_1\rangle \otimes |i_2\rangle$$

This expression is the Schmidt decomposition of the bipartite quantum system $|\psi\rangle$. The bases $\{|i_1\rangle\}_{i=1}^{d_1}$ and $\{|i_2\rangle\}_{i=1}^{d_2}$ are also orthonormal bases of the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively (known as Schmidt bases) and $r = \min\{d_1, d_2\}$. The coefficients λ_i are non-negative real numbers satisfying $\sum_{i=1}^{r} \lambda_i^2 = 1$ and are known as Schmidt coefficients. Since $\lambda_i \equiv D_{ii}$, the Schmidt coefficients λ_i correspond to the singular values of the matrix representation C of quantum system $|\psi\rangle$.

The number of the non-zero values λ_i is called the Schmidt number for the state $|\psi\rangle$. The Schmidt number is an important property of a composite quantum system, which in some sense quantifies the amount of entanglement between the two subsystems of system $|\psi\rangle$ [1]. In particular, this number indicates the separability of the given system, since:

– If a system is separable, the Schmidt number is 1.

– If a system is entangled, the Schmidt number is greater than 1. A pure state for which all Schmidt coefficients λ_i are equal to $\frac{1}{\sqrt{r}}$ is called a maximally entangled state.

It should be noted that the Schmidt coefficients λ_i are invariant under local unitary transformations $U_1 \otimes U_2$ applied to the bipartite quantum state $|\psi\rangle$. Therefore, the Schmidt number is preserved under such local transformations.

In addition, the Schmidt decomposition implies that both partial traces of any bipartite pure state $|\psi\rangle$ have the same nonzero part of the spectrum:

$$\rho_1 = Tr_2(|\psi\rangle\langle\psi|) = \sum_{i=1}^{r} \lambda_i^2 |i_1\rangle\langle i_1|$$
$$\rho_2 = Tr_1(|\psi\rangle\langle\psi|) = \sum_{i=1}^{r} \lambda_i^2 |i_2\rangle\langle i_2|$$

As a result, the eigenvalues of ρ_1 and ρ_2 are identical, namely λ_i^2, for both density matrices. Many important properties of quantum systems are completely determined by the eigenvalues of the reduced density matrix of the system, so, for a pure state of a composite system, such properties will be the same for both systems [1].

1.12 The "double-wedge" notation

It is known that any bipartite quantum system may be described by the state vector $|\psi\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}|\phi_i\rangle\otimes|x_j\rangle$, with $\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}|c_{ij}|^2 = 1$. Consider the following state vector $|\psi\rangle$ of two qubits in $\{|0\rangle, |1\rangle\}$:

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle = \begin{pmatrix}|0\rangle & |1\rangle\end{pmatrix}\begin{bmatrix}c_{00} & c_{01}\\ c_{10} & c_{11}\end{bmatrix}\begin{pmatrix}|0\rangle\\ |1\rangle\end{pmatrix}$$

There is an equivalent expression for bipartite systems called "double-wedge" ket vector $|A\rangle\rangle$. In particular, A is the matrix representation of a quantum system $|\psi\rangle$, whose element in position (i, j) is the respective coefficient c_{ij} of $|\psi\rangle$. So, in this case:

$$\begin{pmatrix}|0\rangle & |1\rangle\end{pmatrix}\begin{bmatrix}c_{00} & c_{01}\\ c_{10} & c_{11}\end{bmatrix}\begin{pmatrix}|0\rangle\\ |1\rangle\end{pmatrix} \equiv \left|\begin{bmatrix}c_{00} & c_{01}\\ c_{10} & c_{11}\end{bmatrix}\right\rangle\rangle = |A\rangle\rangle$$

Due to this system description, a double-wedge ket $|A\rangle\rangle$ must also be normalized as a prerequisite of the completeness theorem, meaning that $\||A\rangle\rangle\|_2 = \|A\|_F = 1$.

This "double-wedge" notation exploits the correspondence between quantum state vectors in $\mathcal{H}_1\otimes\mathcal{H}_2$ and $n_1\cdot n_2$ matrices, where n_1 and n_2 are the dimensions of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Due to the isomorphism of matrices $A \in \mathbb{C}^{n_1,n_2}$ and vectors $|A\rangle\rangle \in \mathbb{C}^{n_1\cdot n_2}$, it can be seen that:

$$A = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} A_{ij}\,|i\rangle\,\langle j| \leftrightarrow |\,A\rangle\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} A_{ij}|i\rangle \otimes |j\rangle$$

Similar to the vector representation, there is also a "double-wedge" bra vector:

$$\langle\langle A\,| = (|\,A\rangle\rangle)^\dagger = \left\langle\left\langle \begin{bmatrix} c_{00}^* & c_{01}^* \\ c_{10}^* & c_{11}^* \end{bmatrix} \right|$$

Some notable properties of the double-wedge ket notation include the following (*Appendix D - Proposition 1*):

- $(C \otimes K)\,|\,A\rangle\rangle = \left|\,CAK^T\right\rangle\rangle$
- $(C \otimes \mathbb{I})\,|\,A\rangle\rangle = |\,CA\rangle\rangle$
- $(\mathbb{I} \otimes K)\,|\,A\rangle\rangle = \left|\,AK^T\right\rangle\rangle$
- $\|\,|\,A\rangle\rangle\|_2 = \sqrt{\langle\langle A\,|\,|\,A\rangle\rangle} = \sqrt{Tr(A^\dagger A)} = \|A\|_F$
- $\langle\langle A\,|\,|\,B\rangle\rangle = Tr(A^\dagger B)$
- $|\,A + B\rangle\rangle = |\,A\rangle\rangle + |\,B\rangle\rangle$

Several notable examples of double-wedge ket systems are presented below:

$|\,X\rangle\rangle = \left|\,\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\rangle\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$

$|\,iY\rangle\rangle = \left|\,\frac{i}{\sqrt{2}}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\right\rangle\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$

$|\,Z\rangle\rangle = \left|\,\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right\rangle\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$

$|\,\mathbb{I}\rangle\rangle = \left|\,\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\rangle\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

Considering that a bipartite system is in a pure state, its respective density matrix ρ is equal to $|\psi\rangle\,\langle\psi|$, which is equivalent to $|\,A\rangle\rangle\,\langle\langle A\,|$ with reference to its double-wedge ket notation. In addition, the calculation of the reduced density matrices is done according to the following formulas (*Appendix D - Proposition 2*):

$$\rho_1 = AA^\dagger \tag{2}$$

$$\rho_2 = A^T A^* \tag{3}$$

1.13 The Hadamard product map

Consider the "controlled-not" gate U_{cn} mentioned in a previous section. The action of this quantum XOR-gate onto a chosen set of basis states $\{|i\rangle\}$ with $i \in \{0, 1\}$ of the Hilbert space of each qubit has been defined as:

$$U_{cn} |i\rangle |j\rangle = |i\rangle |i \oplus_2 j\rangle$$

This two-qubit transformation has the following properties:
– it is unitary and thus reversible,
– it is Hermitian,
– $i \oplus_2 j = 0$ if and only if $i = j$.

However, for many practical tasks of quantum information processing, it is often desirable to extend the basic notion of such a quantum XOR-operation to higher-dimensional Hilbert spaces. The desired generalized quantum XOR-gate (U_{xor}-gate) should act on two n-dimensional quantum systems. In principle, this U_{xor} gate could be defined in a straightforward way, with the exception of performing $i \oplus_n j$, as shown below:

$$U_{xor} |i\rangle |j\rangle = |i\rangle |i \oplus_n j\rangle$$

Despite the fact that this process is also unitary, it is is not Hermitian for $n > 2$, thus becoming irreversible. Therefore, the inverse U_{xor}-gate has to be obtained from this U_{xor}-gate by iteration, meaning that

$$U_{xor}^{-1} = U_{xor}^{n-1} = U_{xor}^{\dagger} \neq U_{xor}$$

As it turns out, all these inconvenient properties of this initial approach can be removed by the following definition [5]:

$$U_{xor} |i\rangle |j\rangle = |i\rangle |i \ominus_n j\rangle$$

The symbol $\ominus_n$ denotes the difference $i - j$ modulo n. This definition preserves all the necessary properties for arbitrary values of n, namely:
– it is unitary,
– it is Hermitian,
– $i \ominus_n j = 0$ if and only if $i = j$ (recall that $i, j \in [0, n - 1]$).

An interesting class of nonlinear quantum maps can be implemented with the help of this U_{xor}-gate. Together with filtering measurements acting on a target qubit system, the U_{xor}-gate induces nonlinear transformations of quantum states of a control system [5].

Consider the n-dimensional quantum state vectors:

$$|\psi\rangle = \sum_{i=0}^{n-1} a_i |i\rangle = \begin{pmatrix} a_0 \\ a_1 \\ ... \\ a_{n-1} \end{pmatrix}$$
$$|\phi\rangle = \sum_{j=0}^{n-1} b_j |j\rangle = \begin{pmatrix} b_0 \\ b_1 \\ ... \\ b_{n-1} \end{pmatrix}$$

with $a_i, b_j \in \mathbb{C}$ for every (i, j).

Recalling the generalized U_{xor}-gate mentioned earlier, the following operation is defined:

$$\mathcal{V} \equiv (\mathbb{I} \otimes P_0) U_{xor} \tag{4}$$

with P_0 corresponding to the projector of state $|0\rangle$ in the n-dimensional vector domain. It should be noted that this transformation is not unitary and so it does not preserve the norm of the system it acts on. However, it is indeed possible to realize this transformation in the lab; it is only a question of technological difficulties and therefore time [6].

The application of this $\mathcal{V}$ process on the composite system $|\psi\rangle \otimes |\phi\rangle$ is:

$$\mathcal{V}(|\psi\rangle \otimes |\phi\rangle) = (\mathbb{I} \otimes P_0) U_{xor}((|\psi\rangle \otimes |\phi\rangle)) \rightarrow |\psi \circ \phi\rangle \otimes |0\rangle \tag{5}$$

with

$$|\psi \circ \phi\rangle = \sum_{i=0}^{n-1} (a_i \circ b_i) |i\rangle = \begin{pmatrix} a_0 b_0 \\ a_1 b_1 \\ ... \\ a_{n-1} b_{n-1} \end{pmatrix}$$

So, this non-linear operation generates the elementwise multiplication (also known as the Hadamard Product) of the state coefficients of the n-dimensional quantum state vectors involved. The fact that the norm of the state is not preserved poses a problem of minimum significance, since it is possible to normalize this output if necessary [6].

This Hadamard Product map may also be applied to the combination of two square complex matrices A, B of the same dimension n in a slightly different way. In particular, the adjoint application of process $\mathcal{V}$ (as $Ad(\mathcal{V}) = Ad((\mathbb{I} \otimes P_0) U_{xor})$) on $(A \otimes B)$ yields the following result (*Appendix D - Proposition 3*):

$$Ad(\mathcal{V})(A \otimes B) = (\mathbb{I} \otimes P_0) U_{xor} (A \otimes B) U_{xor}^{\dagger} (\mathbb{I} \otimes P_0) \rightarrow (A \circ B) \otimes P_0 \tag{6}$$

with $\circ$ denoting the Hadamard Product of the matrices involved.

1.14 Oracle algebra

Given a set with N elements, it is possible to find $1 \leq k \ll N$ marked elements from this set, via a "black box" (known as oracle) that answers queries [8]. This "black box" can be described by an oracle boolean function $f : \{1, 2, ..., N\} \rightarrow \{0, 1\}$, introduced as the characteristic function of subset $I \subset \{1, 2, ..., N\}$ of marked items,with $f(i) = 1$ for $i \in I$ and $f(i) = 0$ otherwise.

Let a boolean function $f : \{1, 2, ..., N\} \rightarrow \{0, 1\}$ and the orthogonal vectors $|x\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{N} f(i) |i\rangle$ and $\left|x^{\perp}\right\rangle = \frac{1}{\sqrt{n^{\perp}}} \sum_{i=1}^{N} (1 - f(i)) |i\rangle$, with $n = \sum_{i=1}^{N} f(i)$ and $n^{\perp} = \sum_{i=1}^{N} (1 - f(i))$,which generate the Hilbert space $\mathcal{H}_2 = V_x = span\{|x\rangle, \left|x^{\perp}\right\rangle\}$ and the unit element $S_0 = |x\rangle \langle x| + \left|x^{\perp}\right\rangle \left\langle x^{\perp}\right|$. The oracle algebra is defined as the vector space $\mathcal{A}_f = \{M \in \mathbb{C}^{NxN}; MS_0M^{\dagger} = S_0\}$, generated by the elements $S_1 = |x\rangle \left\langle x^{\perp}\right| + \left|x^{\perp}\right\rangle \langle x|$, $S_2 = -i |x\rangle \left\langle x^{\perp}\right| + i \left|x^{\perp}\right\rangle \langle x|$ and $S_3 = |x\rangle \langle x| - \left|x^{\perp}\right\rangle \left\langle x^{\perp}\right|$.

The matrices S_i are Hermitian and satisfy the commutation relations $[S_a, S_b] = 2iS_c$ (cyclically) and $[S_0, everything] = 0$. It can be seen that the set $\{S_0, S_1, S_2, S_3\}$ is analogous to the set of Pauli matrices and $\mathcal{A}_f \approx u(2)$, i.e. oracle algebra is isomorphic to $u(2)$ matrix algebra [10].

2 Iterative Entanglement Compression Algorithm

2.1 Introduction

It is a well known fact that every matrix A of digital (classical) information has a Singular Value Decomposition $A = USV^\dagger$. As mentioned in the appendix, this process transforms a matrix of dimensions $n \cdot m$ into a sum of rank-one matrices, whose coefficients are the singular values σ_i of the matrix:

$$A = USV^\dagger = \sum_{i=1}^{r} \sigma_i u_i v_i^\dagger$$

with r being the rank of the matrix A.

Through the Singular Value Decomposition, the information included in a matrix A may be compressed by omitting rank-one matrices, whose coefficients σ_i are considered to be very small compared to the rest. The "small" criteria vary between the various applications of this method, depending on their contribution to certain matrix properties. So, the matrix is composed by a linear combination of fewer rank-one matrices, thus reducing its rank while maintaining the "lion's share" of the initial information.

This compression method is a vital mathematical tool in the field of digital image processing. After all, every digital image may be expressed as a two dimensional matrix, whose elements correspond to the pixels of the image. The properties of each pixel, such as brightness (the intensity of light) and color, are described by real numbers, that are stored in the respective position in the two dimensional image matrix. The simplest example is the grayscale digital image. The information contained in this image is considered to be the changes of brightness between pixels (as shades of gray), with no color information whatsoever. Therefore, the elements of its matrix representation are just single real positive integer values, ranging from pure black (minimum value 0) to pure white (maximum value). It should be mentioned that the maximum pixel value depends on the amount of binary digits (bits) used to describe image brightness (also known as "bits-per-pixel"). For 8 bits per pixel, the maximum pixel value is $2^8 - 1 = 255$.

With the use of its Singular Value Decomposition, a grayscale digital image A may be approximated by a digital image of a lesser rank $1 < k \ll r$, as shown below:

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^\dagger \rightarrow A' = \sum_{i=1}^{k} \sigma_i u_i v_i^\dagger$$

This Low Rank Matrix Approximation is considered to be a lossy compression method of images, since any extra information corresponding to the rank-one matrices of the last $r - k$ singular values is lost. However, several

image features are preserved to an extent depending on the amount of "small" singular value components omitted. The greater the amount, the fewer image details are preserved in the result, leading to a higher compression ratio.

In this thesis, the matrix representation of a square grayscale digital image A of dimension n is considered to be the coefficient matrix of a bipartite quantum system $|A\rangle\rangle$ as well. This way, the normalized pixel values of the image correspond to the coefficients of all the possible states that the system may be found in, as presented in the following example (with black (■), gray (◊) and white(□) pixels):

$$\left|\begin{pmatrix} ■ & ■ & ◊ & ■ \\ ■ & □ & □ & ■ \\ ■ & ■ & ◊ & □ \\ ■ & ◊ & □ & ■ \end{pmatrix}\right\rangle\rangle = \begin{pmatrix} |00\rangle & |01\rangle & |10\rangle & |11\rangle \end{pmatrix} \begin{pmatrix} ■ & ■ & ◊ & ■ \\ ■ & □ & □ & ■ \\ ■ & ■ & ◊ & □ \\ ■ & ◊ & □ & ■ \end{pmatrix} \begin{pmatrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{pmatrix}$$

.

In particular, such an image is viewed as the entangling factor between two pairs of $\log_2 n$ qubits from the perspective of quantum information, with each subsystem having n possible states, as shown below (here image $256x256$):

$$\begin{pmatrix} |a_0b_0c_0d_0e_0f_0g_0h_0\rangle & \dots & |a_1b_1c_1d_1e_1f_1g_1h_1\rangle \end{pmatrix} \begin{bmatrix} \\ \text{  } \\ \text{Lena} \\ \\ \end{bmatrix} \begin{pmatrix} |j_0k_0l_0m_0n_0p_0q_0t_0\rangle \\ \dots \\ |j_1k_1l_1m_1n_1p_1q_1t_1\rangle \end{pmatrix}$$

The goal of this iterative algorithm is the compression of the information contained in an input image A_{in}, in terms of the amount of entanglement of the respective bipartite system $|A_{in}\rangle\rangle$. This process involves two main stages: the dimensional reduction of the input image down to its Low Rank Approximated Matrix A_l of a certain rank k and the iterative application of the Hadamard product for a number of iterations, on the grounds of the optimum restoration of the initial entanglement. In other words, the output quantum system $|A_{out}\rangle\rangle$ of this algorithm is a better approximation of the original system $|A_{in}\rangle\rangle$ than the Low Rank Approximated system $|A_l\rangle\rangle$ in terms of entropy.

2.2 Preliminary material

In order for this algorithm to have any effect, a main prerequisite have to be met regarding the consideration of the digital image A_{in} as the double-wedge ket vector $|A_{in}\rangle\rangle$. Also, two mathematical processes, that are the cornerstone of the proposed compression scheme, are presented in this section.

2.2.1 Quantum prerequisite

Consider the Singular Value Decomposition (SVD) of the matrix representation of a given grayscale image A. This decomposition is $A = UD_AV^\dagger$, with U,V being unitary matrices and D_A diagonal matrix, the elements of the latter being the singular values of matrix A. It should be noted that this decomposition exists for every complex matrix $A \in M^{m,n}$.

In this thesis, the focus lies on real square matrices of dimension 2^n ($n \in \mathbb{N}^*$), rather than complex ones of random dimension. The reasons for such criteria include the consideration of these matrices to be both digital images (hence real) and double-wedge kets describing the entanglement between two pairs of n qubits (hence square of dimension 2^n). In the special case that $A \in \mathbb{R}$, then U,V may be taken to be real orthogonal matrices [13]. As a result, the following decomposition is considered:

$$A = UD_AV^T$$

However, for any matrix A to be a valid representation of a quantum system, its norm must be equal to one as a prerequisite attributed to the completeness theorem. So, in order that the double-wedge ket $|A_{in}\rangle\rangle$ is to be valid, its argument matrix A must be normalized in terms of its Frobenius norm since $\||A_{in}\rangle\rangle\|_2 = \|A_{in}\|_F = 1$. The norm of initial matrix A is:

$$\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2} = \|D_A\|_F$$

So, the normalized version of A is:

$$A_{in} = \frac{A}{\|A\|_F}$$

By substitution of A with its respective SVD representation, it can be seen that:

$$|A_{in}\rangle\rangle = (U \otimes V)\,|D\rangle\rangle \tag{7}$$

with D being the normalized diagonal matrix of singular values.

Due to the isomorphism of matrices D and double-wedge ket vectors $|\,D\rangle\rangle$, the following correlation is valid:

$$D = \sum_{i=1}^{r} \sigma_i \,|i\rangle \,\langle i| \leftrightarrow |\,D\rangle\rangle = \sum_{i=1}^{r} \sigma_i |i\rangle \otimes |i\rangle \tag{8}$$

with r being the rank of matrix A.

So, equation (7) becomes:

$$|\,A_{in}\rangle\rangle = \sum_{i=1}^{r} \sigma_i U|i\rangle \otimes V|i\rangle$$

$$|\,A_{in}\rangle\rangle = \sum_{i=1}^{r} \sigma_i \,|i'\rangle \otimes |i''\rangle$$

with

$$\begin{aligned} |i'\rangle &\equiv U\,|i\rangle \\ |i''\rangle &\equiv V\,|i\rangle \end{aligned}$$

and

$$\|A_{in}\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2} = 1$$

2.2.2 Orthogonal decomposition

Given the set of r singular values in D, it is possible to mark $1 \leq k \ll r$ of them via a "black box" (known as oracle) that answers queries. This "black box" can be described by an oracle boolean function $f : \{1, 2, ..., r\} \rightarrow \{0, 1\}$. For the purposes of this thesis, the oracle function f decomposes the set of singular values σ_i of the input matrix A_{in} into two orthogonal sets $x_{large} = \{\sigma_i : f(i) = 1\}$ and $x_{small} = \{\sigma_i : f(i) = 0\}$. The set x_{large} contains the k largest singular values, while the set x_{small} contains the $r - k$ smallest ones. It should be noted that this type of partitioning occurs in the "black box" for a fixed value k over different inputs A_{in}, without any prior knowledge regarding the way it is performed. Also, it is impossible to interfere with this partitioning process.

The singular values σ corresponding to each of the sets x_{large}, x_{small} are then assigned as the diagonal elements of the matrices D_{large} and D_{small}, in a way that $D = D_{large} + D_{small}$. Due to the non-increasing order of the singular values in D (*Appendix B*), the matrix D_{large} has k consecutive elements σ in the diagonal positions 1 through k, whereas the matrix D_{small} has $r-k$ consecutive elements σ in the diagonal positions $k+1$ through r.

Following this assignment, the following matrices are defined:

$$A_{large} \equiv U D_{large} V^T = \sum_{i=1}^{k} \sigma_i \left|i'\right\rangle \left\langle i''\right| \tag{9}$$

and

$$A_{small} \equiv U D_{small} V^T = \sum_{i=k+1}^{r} \sigma_i \left|i'\right\rangle \left\langle i''\right| \tag{10}$$

with $\left|i'\right\rangle = U \left|i\right\rangle$ and $\left|i''\right\rangle = V \left|i\right\rangle$.

While $\langle A_{large}, A_{small}\rangle = \langle A_{small}, A_{large}\rangle = 0$, this corresponds to the orthogonal decomposition of $A_{in} = A_{large} + A_{small}$.

Concerning the respective double-wedge ket form, it is obvious that:
$\left|\, A_{in}\right\rangle\rangle = \left|\, A_{large} + A_{small}\right\rangle\rangle$
$\left|\, A_{in}\right\rangle\rangle = \left|\, A_{large}\right\rangle\rangle + \left|\, A_{small}\right\rangle\rangle$

Since the initial matrix A is normalized into $\left|\, A_{in}\right\rangle\rangle$, the addition of $\left|\, A_{large}\right\rangle\rangle +$ $\left|\, A_{small}\right\rangle\rangle$ is normalized as well. However, each of the terms A_{large}, A_{small} is not. So, they are normalized into A_l, A_s respectively, as shown below:

$$A_l = \frac{A_{large}}{c_l} \tag{11}$$

with $c_l \equiv \left\|A_{large}\right\|_F = \left\|D_{large}\right\|_F = \sqrt{\sum_{i=1}^{k} \sigma_i^2}$

and

$$A_s = \frac{A_{small}}{c_s} \tag{12}$$

with $c_s \equiv \left\|A_{small}\right\|_F = \left\|D_{small}\right\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2}$

By substitution, it can be seen that:

$$\left|\, A_{in}\right\rangle\rangle = c_l \left|\, A_l\right\rangle\rangle + c_s \left|\, A_s\right\rangle\rangle \tag{13}$$

In an analogous way, the diagonal matrices corresponding to the "large" and "small" components of A_{in} are normalized as:

$$\begin{aligned} D_l &\equiv \frac{D_{large}}{c_l} \\ D_s &\equiv \frac{D_{small}}{c_s} \end{aligned}$$

2.2.3 Generating correlation

With the use of an auxiliary qubit (say $|0\rangle$), it is possible to build-up correlation between the two orthonormal parts $|A_l\rangle\rangle, |A_s\rangle\rangle$ and the states of that qubit. This process includes the creation of superposition states through the Hadamard operator, followed by proper projections over the auxilliary qubit. It should be noted that the availability of the initial state $|X_0\rangle = |0\rangle \otimes |A_{in}\rangle\rangle$ for further processing is required.

At first, with the use of the Hadamard Operator over the first Hilbert space, the creation of superposition states between two instances of $|A_{in}\rangle\rangle$ takes place as shown below:

$|X_1\rangle = (H \otimes \mathbb{I}\ ) |X_0\rangle$
$|X_1\rangle = (H \otimes \mathbb{I})(|0\rangle \otimes |A_{in}\rangle\rangle)$
$|X_1\rangle = H|0\rangle \otimes \mathbb{I}|A_{in}\rangle\rangle$
$|X_1\rangle = (\frac{|0\rangle+|1\rangle}{\sqrt{2}}) \otimes |A_{in}\rangle\rangle$

$$|X_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |A_{in}\rangle\rangle + |1\rangle \otimes |A_{in}\rangle\rangle) \tag{14}$$

As mentioned above, it is possible to establish correlation between the systems $|A_l\rangle\rangle, |A_s\rangle\rangle$ and the states of the auxiliary qubit through proper projections. In particular, such projections include the transformation of each instance of $|A_{in}\rangle\rangle$ into its respective components$|A_l\rangle\rangle$ and $|A_s\rangle\rangle$. For this purpose, the following projector operators are defined:

$$P_{large} \equiv |A_l\rangle\rangle \langle\langle A_l| \tag{15}$$

and

$$P_{small} \equiv |A_s\rangle\rangle \langle\langle A_s| \tag{16}$$

These projection operators have the following applications:

$$P_{large}\,|\,A_{in}\rangle\rangle = c_l\,|\,A_l\rangle\rangle \tag{17}$$

$$P_{small}\,|\,A_{in}\rangle\rangle = c_s\,|\,A_s\rangle\rangle \tag{18}$$

Such projections are performed with the use of a "controlled" projection operator over the state $|X_1\rangle$, targeting each instance of $|\,A_{in}\rangle\rangle$ and projecting it into $c_l\,|\,A_l\rangle\rangle\,, c_s\,|\,A_s\rangle\rangle$ with reference to the state of the auxilliary qubit $|0\rangle\,, |1\rangle$ respectively. This "controlled" projection operator is shown below:

$$P_{sel} \equiv P_0 \otimes P_{large} + P_1 \otimes P_{small} \tag{19}$$

Its application upon the state $|X_1\rangle$ yields the following result :
$|X_2\rangle = P_{sel}\,|X_1\rangle$
$|X_2\rangle = (P_0 \otimes P_{large} + P_1 \otimes P_{small})(\frac{1}{\sqrt{2}}(|0\rangle\,|\,A_{in}\rangle\rangle + |1\rangle\,|\,A_{in}\rangle\rangle))$

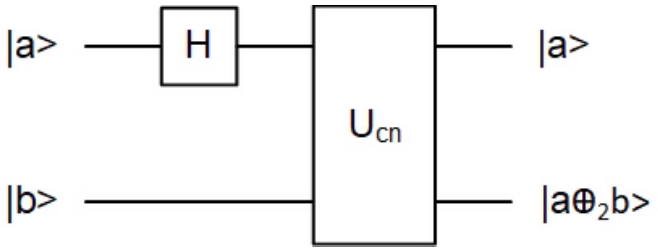
$$|X_2\rangle = \frac{1}{\sqrt{2}}[(P_0\,|0\rangle \otimes P_{large}\,|\,A_{in}\rangle\rangle) + (P_0\,|1\rangle \otimes P_{large}\,|\,A_{in}\rangle\rangle) + \\ + (P_1\,|0\rangle \otimes P_{small}\,|\,A_{in}\rangle\rangle) + (P_1\,|1\rangle \otimes P_{small}\,|\,A_{in}\rangle\rangle)]$$

Recall that: $P_0\,|0\rangle = |0\rangle$, $P_1\,|1\rangle = |1\rangle$ and $P_0\,|1\rangle = 0$, $P_1\,|0\rangle = 0$

As a result:
$|X_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes c_l\,|\,A_l\rangle\rangle + 0 + 0 + |1\rangle \otimes c_s\,|\,A_s\rangle\rangle)$

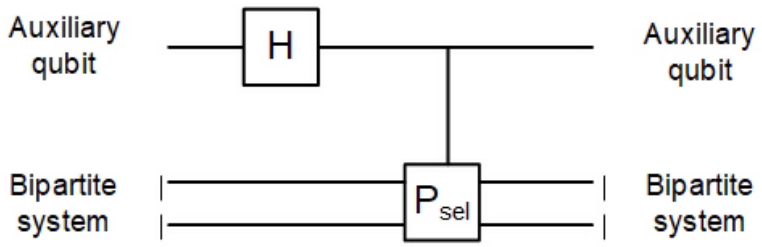
$$|X_2\rangle = \frac{1}{\sqrt{2}}(c_l\,|0\rangle \otimes |\,A_l\rangle\rangle + c_s\,|1\rangle \otimes |\,A_s\rangle\rangle) \tag{20}$$

It should be mentioned that this specific correlating process is inspired by the Bell State correlation mechanism $|\beta_{ab}\rangle = (H \otimes \mathbb{I})U_{cn}(|a\rangle \otimes |b\rangle)$ for arbitrary qubits $|a\rangle\,, |b\rangle$ presented below:



Bell state correlation

While this mechanism results in the correlation between the values of the two qubits involved (such as $|\beta_{00}\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$), the proposed process establishes correlation between the values of an auxiliary qubit and the components of a bipartite system.



Proposed correlation process

This correlation is of paramount importance, since the measurement upon the auxiliary qubit provides accurate information about the state of the bipartite system, without the need for further measurements upon the latter.

2.3 Compression process

The process of quantum compression refers to the dimensional reduction of a bipartite quantum system $|A_{in}\rangle\rangle$, while preserving most of the amount of entanglement contained in the system. The dimensional reduction implies the omission of the components of the Schmidt decomposition of the given bipartite system, whose coefficients are relatively small compared to the rest. This is the equivalent process of extracting the respective quantum system corresponding to the Low Rank Approximation (LRA) A_l, of a certain rank k, out of the initial A_{in}. A quantum process is then applied to this system, in an attempt to restore a significant amount of the entropy of the initial system upon it.

In this thesis, the proposed compression algorithm of quantum information focuses on the optimum restoration of the initial entropy through the iterative application of the Hadamard Product between the "large" component of the previous output $(A_{out})_l$ and the "large" component of the initial input A_l. The main idea involves the indirect multiplication of the singular values of these components, thus inducing an exponential reduction in their size, since their values are positive numbers between 0 and 1. In this way, the contributions of the various components to the entropy of the current output bipartite quantum system become more disproportionate, allowing the omission of several components with no relative effect on the general structure of the system.

The stages for this proposed compression method are analyzed below:

2.3.1 Compression process - Stage 0: Input state

The first step of this process requires that the quantum state $|X_0\rangle = |0\rangle \otimes |0\rangle \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle$ be available for processing. It can be seen that this state consists of 2 auxilliary qubits $|0\rangle$, the double-wedge ket $|A_{in}\rangle\rangle$ as well as its "transpose" system $\left|A_{in}^T\right\rangle\rangle$.

At this point, the use of a boolean function f with respect to oracle matrix algebra decomposes the set of the singular values of A_{in} into two subsets, one containing the largest k singular values in x_{large} and the other one containing the rest, meaning the $r-k$ smallest singular values in x_{small}. As mentioned in a previous section, these sets may correspond to the complementary diagonal matrices D_{large} and D_{small}, with $D = D_{large} + D_{small}$. This leads to the orthogonal decomposition of matrix $A_{in} = A_{large} + A_{small}$ with:

$$\begin{aligned} A_{large} &\equiv UD_{large}V^T \\ A_{small} &\equiv UD_{small}V^T \end{aligned}$$

Since the initial matrix A is normalized into $|A_{in}\rangle\rangle$, the addition of $|A_{large}\rangle\rangle + |A_{small}\rangle\rangle$ is normalized as well. However, each of the terms A_{large}, A_{small} is not. So, they are normalized into A_l, A_s respectively, as shown below:

$$\begin{aligned} A_l &= \frac{A_{large}}{c_l} \\ A_s &= \frac{A_{small}}{c_s} \end{aligned}$$

with $c_l \equiv \|A_{large}\|_F$ and $c_s \equiv \|A_{small}\|_F$.

In an analogous way, the diagonal matrices corresponding to the "large" and "small" components of A_{in} are normalized as:

$$\begin{aligned} D_l &\equiv \frac{D_{large}}{c_l} \\ D_s &\equiv \frac{D_{small}}{c_s} \end{aligned}$$

By substitution, it can be seen that:

$$|\,A_{in}\rangle\rangle = c_l\,|\,A_l\rangle\rangle + c_s\,|\,A_s\rangle\rangle$$

Considering that the transpose matrix A_{in}^T has the same singular values as A_{in}, the use of the same boolean function f on A_{in}^T yields the following orthogonal decomposition:

$$\begin{aligned} VD_{large}U^T &= (UD_{large}V^T)^T = (A_{large})^T \\ VD_{small}U^T &= (UD_{small}V^T)^T = (A_{small})^T \end{aligned}$$

It can be easily seen that the norms of the orthogonal components $(A_{large})^T, (A_{small})^T$ are c_l, c_s respectively. As a result, the bipartite quantum system corresponding to A_{in}^T may be expressed as:

$$\left|\,A_{in}^T\right\rangle\rangle = c_l\left|\,A_l^T\right\rangle\rangle + c_s\left|\,A_s^T\right\rangle\rangle$$

2.3.2 Compression process - Stage 1: Generating correlation

As described thoroughly in a previous section, it is possible to establish correlation between the orthonormal parts $|\,A_l\rangle\rangle, |\,A_s\rangle\rangle$ of $|\,A_{in}\rangle\rangle$ and the values of an auxiliary qubit with the use of a Hadamard Operator and proper projections over that qubit. In this case, such operations are extended over two auxiliary qubits, thus requiring the assignment of each auxiliary qubit to its respective set of Hilbert spaces regarding the double-wedge ket vectors. Despite the fact that it is binding for the entirety of this algorithm, the choice concerning the correlated sets of Hilbert spaces is arbitrary and different choices yield the same results. For the purposes of this thesis, the auxiliary qubit 1 corresponds to the bipartite system $\left|\,A_{in}^T\right\rangle\rangle$ (position 3) and the auxiliary qubit 2 corresponds to the bipartite system $|\,A_{in}\rangle\rangle$ (position 4). These results involve the composite correlation between pairs of the orthonormal components $|\,A_l\rangle\rangle, |\,A_s\rangle\rangle, \left|\,A_l^T\right\rangle\rangle, \left|\,A_s^T\right\rangle\rangle$ of $|\,A_{in}\rangle\rangle$, $\left|\,A_{in}^T\right\rangle\rangle$ and the values of the auxiliary qubits $\{|0\rangle, |1\rangle\}$.

At first, the action of the Hadamard Operator occurs over the first two Hilbert spaces, corresponding to the two auxiliary qubits. This application serves as a means to build-up the correlation mentioned above.

$(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle = (H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})(|0\rangle \otimes |0\rangle \otimes \left|\,A_{in}^T\right\rangle\rangle \otimes |\,A_{in}\rangle\rangle)$

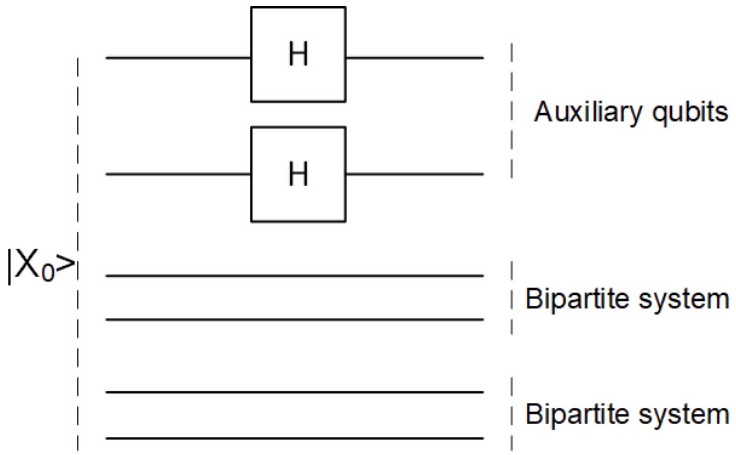
$(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle = H\,|0\rangle \otimes H\,|0\rangle \otimes \mathbb{I}\left|\,A_{in}^T\right\rangle\rangle \otimes \mathbb{I}\,|\,A_{in}\rangle\rangle$

$(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle = (\frac{|0\rangle+|1\rangle}{\sqrt{2}}) \otimes (\frac{|0\rangle+|1\rangle}{\sqrt{2}}) \otimes \left|\,A_{in}^T\right\rangle\rangle \otimes |\,A_{in}\rangle\rangle$

$$(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle = \tfrac{1}{2}[(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle]$$

$$(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle = \frac{1}{2}(|00\rangle \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle + |01\rangle \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle + \\ + |10\rangle \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle + |11\rangle \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle)$$

The quantum circuit corresponding to this application is shown below:



Building-up correlation

Up to this point, the quantum state includes the superposition states consisting of pairs of $\left|A_{in}^T\right\rangle\rangle, |A_{in}\rangle\rangle$. As mentioned in a previous section, the application of a "controlled projection" can establish the correlation between the orthonormal parts of $|A_{in}\rangle\rangle$ to the respective auxiliary qubit. In this case, the "controlled projection" is extended, while the results of its application include the correlation between pairs of the orthonormal components of $|A_{in}\rangle\rangle, \left|A_{in}^T\right\rangle\rangle$ to the bipartite system of the auxiliary qubits.

Recall that the orthogonal decomposition of the quantum systems $|A\rangle\rangle, \left|A^T\right\rangle\rangle$ is based on the distinction between large and small singular values in sets, which is attributed to the use of the oracle boolean function f. Since their respective matrix representations have the same singular values, it is obvious that the partitioning process results in the same sets x_{large} and x_{small}. Due to the fact that there is no prior knowldge about the actual "large" and "small" components generated by the oracle function, each of the orthonormal components $|A_l\rangle\rangle, |A_s\rangle\rangle, \left|A_l^T\right\rangle\rangle, \left|A_s^T\right\rangle\rangle$ may as well be assigned to one of the abstract double-wedge ket vectors $|large\rangle\rangle$ and $|small\rangle\rangle$, depending on the set of the singular values they contain. With this notation in mind, the following matrices are defined:

$$P_{large} = |\,large\rangle\rangle\,\langle\langle large\,| \tag{21}$$

and

$$P_{small} = |\,small\rangle\rangle\,\langle\langle small\,| \tag{22}$$

Provided that the double-wedge kets $|\,large\rangle\rangle$ and $|\,small\rangle\rangle$ correspond to normalized state vectors, it can be seen that these matrices are indeed Hermitian ($P^{\dagger}_{large} = P_{large}, P^{\dagger}_{small} = P_{small}$) and $P^2_{large} = P_{large}, P^2_{small} = P_{small}$, thus being in fact projector operators. The form of the actual large and small projectors depends on the matrix, whose singular values are split by the oracle function f in each case. Nevertheless, the use of the $|\,large\rangle\rangle$ and $|\,small\rangle\rangle$ notations provides the following general results:

$$P_{large}\,|\,A_{in}\rangle\rangle = c_l\,|\,A_l\rangle\rangle \tag{23}$$

$$P_{small}\,|\,A_{in}\rangle\rangle = c_s\,|\,A_s\rangle\rangle \tag{24}$$

$$P_{large}\,\big|\,A^T_{in}\big\rangle\big\rangle = c_l\,\big|\,A^T_l\big\rangle\big\rangle \tag{25}$$

$$P_{small}\,\big|\,A^T_{in}\big\rangle\big\rangle = c_s\,\big|\,A^T_s\big\rangle\big\rangle \tag{26}$$

Taking the "controlled" projection process P_{sel} presented earlier into account (eq. 19), the "controlled" projection of this scenario has the following form:

$$\begin{aligned} P_{sel} \equiv P_0 \otimes P_0 \otimes P_{large} \otimes P_{large} + P_0 \otimes P_1 \otimes P_{large} \otimes P_{small} + \\ + P_1 \otimes P_0 \otimes P_{small} \otimes P_{large} + P_1 \otimes P_1 \otimes P_{small} \otimes P_{small} \end{aligned} \tag{27}$$

The application of this "controlled" projection may as well be split into two separate projections, in accordance with the correlated Hilbert spaces of the composite system (here $1-3$ and $2-4$). So, the following "controlled" sub-projections are defined:

$$P_{13} \equiv P_0 \otimes \mathbb{I} \otimes P_{large} \otimes \mathbb{I} + P_1 \otimes \mathbb{I} \otimes P_{small} \otimes \mathbb{I} \tag{28}$$

$$P_{24} \equiv \mathbb{I} \otimes P_0 \otimes \mathbb{I} \otimes P_{large} + \mathbb{I} \otimes P_1 \otimes \mathbb{I} \otimes P_{small} \tag{29}$$

with

$$P_{sel} = P_{13} \cdot P_{24} = P_{24} \cdot P_{13} \tag{30}$$

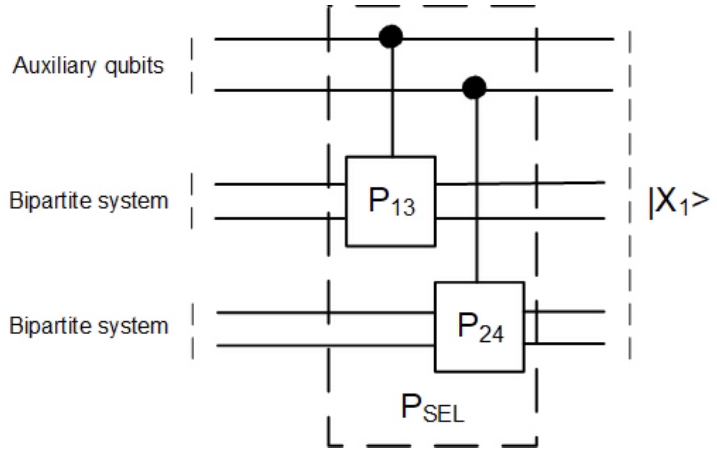
At first, the application of the "controlled projection" over the Hilbert spaces $1-3$ results in (*Appendix D - Proposition 4*):

$$P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle = \frac{1}{2}(c_l\,|00\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + c_l\,|01\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + \\ + c_s\,|10\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + c_s\,|11\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle) \quad (31)$$

In addition, the application of the other "controlled" sub-projection over the Hilbert spaces $2-4$ of the previous state is shown below (*Appendix D - Proposition 5*):

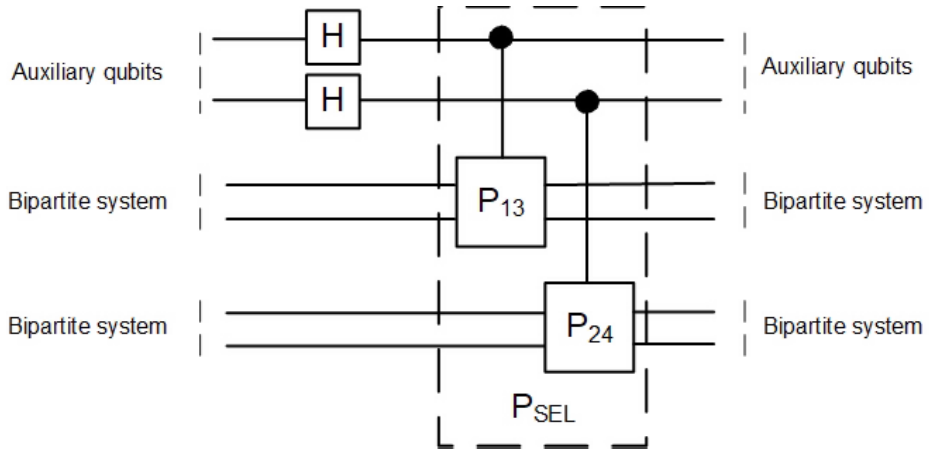
$$|X_1\rangle = \frac{1}{2}(c_l^2\,|00\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_l\rangle\rangle + c_lc_s\,|10\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_l\rangle\rangle + \\ + c_lc_s\,|01\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_s\rangle\rangle + c_s^2\,|11\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_s\rangle\rangle) \quad (32)$$

This quantum process has the following circuit representation:



Controlled projection

With reference to the Bell State correlation mechanism mentioned earlier, the corresponding quantum circuit of this extended correlating process is the following:



Extended correlating process

2.3.3 Compression process - Stage 2: Hadamard product

So far, a correlation between pairs of the orthonormal components of $|A_{in}\rangle\rangle$, $|A_{in}^T\rangle\rangle$ and the values of the auxiliary qubits has been established in the quantum state $|X_1\rangle$.

After further processing, the state $|X_1\rangle$ may be expressed as:

$$|X_1\rangle = \frac{1}{2}(c_l^2\,|00\rangle\otimes(\mathbb{I}\otimes A_l)\,|\mathbb{I}\rangle\rangle\otimes(A_l\otimes\mathbb{I})\,|\mathbb{I}\rangle\rangle + c_l c_s\,|10\rangle\otimes(\mathbb{I}\otimes A_s)\,|\mathbb{I}\rangle\rangle\otimes(A_l\otimes\mathbb{I})\,|\mathbb{I}\rangle\rangle + \\ + c_l c_s\,|01\rangle\otimes(\mathbb{I}\otimes A_l)\,|\mathbb{I}\rangle\rangle\otimes(A_s\otimes\mathbb{I})\,|\mathbb{I}\rangle\rangle + c_s^2\,|11\rangle\otimes(\mathbb{I}\otimes A_s)\,|\mathbb{I}\rangle\rangle\otimes(A_s\otimes\mathbb{I})\,|\mathbb{I}\rangle\rangle)$$

which is equivalent to:

$$|X_1\rangle = \frac{1}{2}\cdot(c_l^2\,|00\rangle\otimes(\mathbb{I}\otimes(A_l\otimes A_l)\otimes\mathbb{I})(|\mathbb{I}\rangle\rangle\otimes|\mathbb{I}\rangle\rangle) + c_l c_s\,|10\rangle\otimes(\mathbb{I}\otimes(A_s\otimes A_l)\otimes\mathbb{I})(|\mathbb{I}\rangle\rangle\otimes|\mathbb{I}\rangle\rangle) + \\ + c_l c_s\,|01\rangle\otimes(\mathbb{I}\otimes(A_l\otimes A_s)\otimes\mathbb{I})(|\mathbb{I}\rangle\rangle\otimes|\mathbb{I}\rangle\rangle) + c_s^2\,|11\rangle\otimes(\mathbb{I}\otimes(A_s\otimes A_s)\otimes\mathbb{I})(|\mathbb{I}\rangle\rangle\otimes|\mathbb{I}\rangle\rangle))$$

As mentioned above, there is a quantum process (called $\mathcal{V}$ in this thesis), whose adjoint action $Ad(\mathcal{V})$ upon a combination of matrices $A \otimes B$ results in their Hadamard Product and the matrix P_0:

$$Ad(\mathcal{V})(A \otimes B) = (A \circ B) \otimes P_0$$

This quantum process is properly extended into a new one $\mathcal{V}' = \mathbb{I} \otimes Ad(\mathcal{V}) \otimes \mathbb{I}$, which is suitable for application upon the combination of two bipartite quantum systems in the form of $\left|\, A^T \right\rangle\!\rangle \otimes |\, B\rangle\!\rangle$, yielding the following result:

$$\mathcal{V}'(\left|\, A^T \right\rangle\!\rangle \otimes |\, B\rangle\!\rangle) = \left|\, (A \circ B)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle$$

The explicit formula corresponding to the extended Hadamard Product map $\mathcal{V}'$ is (*Appendix D - Proposition 6*):

$$\mathcal{V}' = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} |j\rangle \langle j| \otimes |i\rangle \langle i| \otimes |0\rangle \langle i| \otimes |0\rangle \langle j| \tag{33}$$

With the application of the quantum process $\mathbb{I}_2 \otimes \mathcal{V}'$ upon the state $|X_1\rangle$, the extended process $\mathcal{V}'$ acts on the respective Hilbert spaces corresponding to the pairs of the orthonormal components $A_{l,s}$ of A_{in}, resulting in their Hadamard product $\left|\, (A_{l,s} \circ A_{l,s})^T \right\rangle\!\rangle$ and the system described by $|\, P_0\rangle\!\rangle$, as shown below (*Appendix D - Proposition 7*):

$$\begin{aligned} |X_2\rangle = \frac{1}{2}(c_l^2 \,|00\rangle \otimes \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle + c_l c_s \,|10\rangle \otimes \left|\, (A_s \circ A_l)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle + \\ + c_l c_s \,|01\rangle \otimes \left|\, (A_l \circ A_s)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle + c_s^2 \,|11\rangle \otimes \left|\, (A_s \circ A_s)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle) \end{aligned} \tag{34}$$

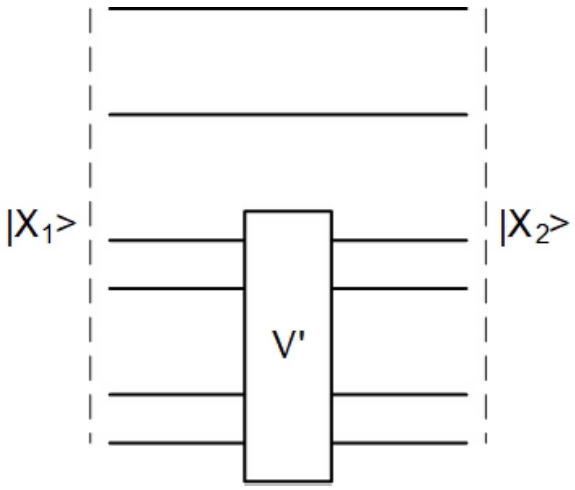
Due to the commutative property of the Hadamard Product, it can be seen that: $(A_s \circ A_l) = (A_l \circ A_s) \Rightarrow \left|\, (A_s \circ A_l)^T \right\rangle\!\rangle = \left|\, (A_l \circ A_s)^T \right\rangle\!\rangle$. So:

$$\begin{aligned} |X_2\rangle = \frac{1}{2}[c_l^2 \,|00\rangle \otimes \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle + c_l c_s (|01\rangle + |10\rangle) \otimes \left|\, (A_l \circ A_s)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle + \\ + c_s^2 \,|11\rangle \otimes \left|\, (A_s \circ A_s)^T \right\rangle\!\rangle \otimes |\, P_0\rangle\!\rangle] \end{aligned} \tag{35}$$

As a result of the application of the Hadamard product map, the third set of Hilbert spaces is now occupied by the possible Hadamard Products between the orthonormal parts of the matrix A_{in}, whereas the fourth set of Hilbert spaces includes a bipartite quantum system described by $|\, P_0\rangle\!\rangle$. Since this quantum system is a common one along the entire linear combination of components, it may be set aside, resulting in the following equivalent expression for the quantum state vector $|X_2\rangle$:

$$\begin{aligned} |X_2\rangle = \frac{1}{2}[c_l^2 \,|00\rangle \otimes \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle + c_l c_s (|01\rangle + |10\rangle) \otimes \left|\, (A_l \circ A_s)^T \right\rangle\!\rangle + \\ + c_s^2 \,|11\rangle \otimes \left|\, (A_s \circ A_s)^T \right\rangle\!\rangle] \otimes |\, P_0\rangle\!\rangle \end{aligned} \tag{36}$$

This quantum process has the following circuit implementation:



Elementwise multiplication

It should be mentioned that there is no guarantee that the quantum system $|X_2\rangle$ is a normalized one. This is attributed to the fact that the double-wedge kets of the third set of Hilbert spaces describe completely different systems now, without any useful property about the norm of the Hadamard Product. Nevertheless, for the purposes of this thesis, the normalization process of the result takes place right before the end of the next stage of the algorithm, ensuring the validity of the output.

2.3.4 Compression process - Stage 3: Output state

Up to this point, the correlation that was initially established between pairs of the orthonormal components of $|A_{in}\rangle\rangle$, $|A_{in}^T\rangle\rangle$ and the respective values of the auxiliary qubits has been passed on to their respective Hadamard Products. This stage involves the assessment of the contribution of each Hadamard Product to the entropy of the composite system, allowing the omission of components whose contribution seems insignificant. In this way, the "lion's share" of the entanglement is preserved, while compressing the entire system in terms of dimensional reduction.

The density matrix ρ of the composite quantum system $|X_2\rangle$ is defined as:

$$\rho = |X_2\rangle \langle X_2|$$

This quantum state spans over 6 Hilbert spaces, 2 corresponding to the auxilliary qubits, 2 corresponding to the bipartite system of the Hadamard Products and 2 corresponding to $|P_0\rangle\rangle$. At this point, the omission of the last subsystem regarding $|P_0\rangle\rangle$ takes place with the use of partial tracing, thus focusing more on the study of the remaining components that contain the processed information about the initial image A_{in}. The application of the proper partial trace over the density matrix ρ is:

$$Tr_{5,6}(\rho) = Tr_{5,6}(|X_2\rangle \langle X_2|)$$

This reduced density matrix $Tr_{5,6}(\rho)$ describes the correlated system consisting of the auxiliary qubits and the Hadamard Product systems $\left|(A_l \circ A_l)^T\right\rangle\rangle$, $\left|(A_l \circ A_s)^T\right\rangle\rangle$, $\left|(A_s \circ A_s)^T\right\rangle\rangle$. However, from a rather abstract point of view, the system described by $\left|(A_l \circ A_l)^T\right\rangle\rangle$ is relatively more interesting than the rest, since it may as well be considered as the optimum approximation of the composite system $|X_2\rangle$ in the same way that the matrix A_l is to the initial matrix A_{in}. Therefore, for the purposes of this thesis, it is assumed that the "lion's share" of the quantum information of the composite system is contained in the respective component of the system $\left|(A_l \circ A_l)^T\right\rangle\rangle$, which allows the omission of the rest components. This is feasible with the direct application of the proper projection regarding the auxiliary qubits upon the reduced density matrix, as shown below:

$$(P_0 \otimes P_0 \otimes \mathbb{I} \otimes \mathbb{I}) Tr_{5,6}(\rho) (P_0 \otimes P_0 \otimes \mathbb{I} \otimes \mathbb{I})^\dagger$$

The result of this process is the following density matrix:

$$\frac{1}{4} c_l^4 (|00\rangle \langle 00| \otimes \left|(A_l \circ A_l)^T\right\rangle\rangle \left\langle\left\langle (A_l \circ A_l)^T \right|\right)$$

This density matrix corresponds to bipartite quantum system spanning over 4 Hilbert spaces, 2 regarding the auxiliary qubits and 2 regarding the double-wedge system $\left|(A_l \circ A_l)^T\right\rangle\rangle$. Like the subsystem $|P_0\rangle\rangle$ before, the first one corresponding to the auxiliary bipartite system serves no purpose whatsoever and it is bound to be omitted, since there is no need for correlation between the various components of the quantum system $|A\rangle\rangle$ anymore. The result of this action is the following reduced density matrix:

$\rho' = Tr_{1,2}(\frac{1}{4} c_l^4 (|00\rangle \langle 00| \otimes \left|(A_l \circ A_l)^T\right\rangle\rangle \left\langle\left\langle (A_l \circ A_l)^T \right|\right))$
$\rho' = \frac{1}{4} c_l^4 \cdot Tr_{1,2}(|00\rangle \langle 00| \otimes \left|(A_l \circ A_l)^T\right\rangle\rangle \left\langle\left\langle (A_l \circ A_l)^T \right|\right)$
$\rho' = \frac{1}{4} c_l^4 \cdot Tr(|00\rangle \langle 00|) \cdot \left|(A_l \circ A_l)^T\right\rangle\rangle \left\langle\left\langle (A_l \circ A_l)^T \right|\right.$
$\rho' = \frac{1}{4} c_l^4 \cdot \langle 00|\, |00\rangle \cdot \left|(A_l \circ A_l)^T\right\rangle\rangle \left\langle\left\langle (A_l \circ A_l)^T \right|\right.$

$\rho' = \frac{1}{4}c_l^4 \cdot 1 \cdot \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \left\langle\!\left\langle (A_l \circ A_l)^T \,\right|$

$$\rho' = \frac{1}{4}c_l^4 \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \left\langle\!\left\langle (A_l \circ A_l)^T \,\right|$$

It should be mentioned that the bipartite quantum system described by this density matrix is also not normalized. The normalization of this state implies that its density matrix meets the requirement $Tr(\rho') = 1$ as an equivalent interpretation of the completeness theorem. So, the trace of this density matrix is:

$Tr(\rho') = Tr(\frac{1}{4}c_l^4 \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \left\langle\!\left\langle (A_l \circ A_l)^T \,\right|)$
$Tr(\rho') = \frac{1}{4}c_l^4 \cdot Tr_{1,2}(\left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \left\langle\!\left\langle (A_l \circ A_l)^T \,\right|)$
$Tr(\rho') = \frac{1}{4}c_l^4 \cdot Tr[(A_l \circ A_l)^T \cdot ((A_l \circ A_l)^T)^\dagger]$
$Tr(\rho') = \frac{1}{4}c_l^4 \cdot Tr[(A_l \circ A_l)^T \cdot ((A_l \circ A_l)^T)^T]$ (since $(A_l \circ A_l)^T$ is real)
$Tr(\rho') = \frac{1}{4}c_l^4 \cdot Tr[(A_l \circ A_l)^T \cdot (A_l \circ A_l)]$

$$Tr(\rho') = \frac{1}{4}c_l^4 \cdot \left\|(A_l \circ A_l)^T\right\|_F^2$$

In order for $Tr(\rho') = 1$, then ρ' becomes ρ_{out}:

$\rho_{out} = \frac{\rho'}{Tr(\rho')}$
$\rho_{out} = \frac{\frac{1}{4}c_l^4 \left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \left\langle\!\left\langle (A_l \circ A_l)^T \,\right|}{\frac{1}{4}c_l^4 \cdot \|(A_l \circ A_l)^T\|_F^2}$

$$\rho_{out} = \frac{\left|\, (A_l \circ A_l)^T \right\rangle\!\rangle \left\langle\!\left\langle (A_l \circ A_l)^T \,\right|}{\|(A_l \circ A_l)^T\|_F^2}$$

This density matrix corresponds to the normalized quantum system:

$$|\, A_{out}\rangle\!\rangle = \frac{\left|\, (A_l \circ A_l)^T \right\rangle\!\rangle}{\|(A_l \circ A_l)^T\|_F}$$

This bipartite quantum system is considered to be the output of this algorithm after one iteration. It should be noted that the rank r_out of the matrix $(A_l \circ A_l)^T$ is bounded between 1 and k^2 , since:

$$rank(A \circ B) \leq rank(A) rank(B) \Rightarrow rank(A_l \circ A_l) \leq rank(A_l) rank(A_l) \Rightarrow r\_out \leq k^2$$

Provided that $k^2 < r\_in$, with r_in being the rank of the initial matrix A_{in}, the dimensional reduction between the Schmidt decompositions of $|\, A_{in}\rangle\!\rangle$ and $|\, A_{out}\rangle\!\rangle$ is obvious. A simple table schematic regarding the corresponding Singular Value Decompositions $A_{in} = U_{in} D_{in} V_{in}^T$, $A_{out} = U_{out} D_{out} V_{out}^T$ is presented below (considering dimension n):

$$|A_{in}\rangle\rangle = \big(\log_2 \mathrm{n} \qquad \text{qubits}\big)\, U_{in} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & ... & 0 & 0 & 0 & 0 \\ 0 & 0 & ... & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{r\_in} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0... & 0 \\ 0 & 0 & 0 & 0 & 0 & ...0 \end{bmatrix} V_{in}^T \begin{pmatrix} \log_2 \mathrm{n} \\ \\ \text{qubits} \end{pmatrix}$$

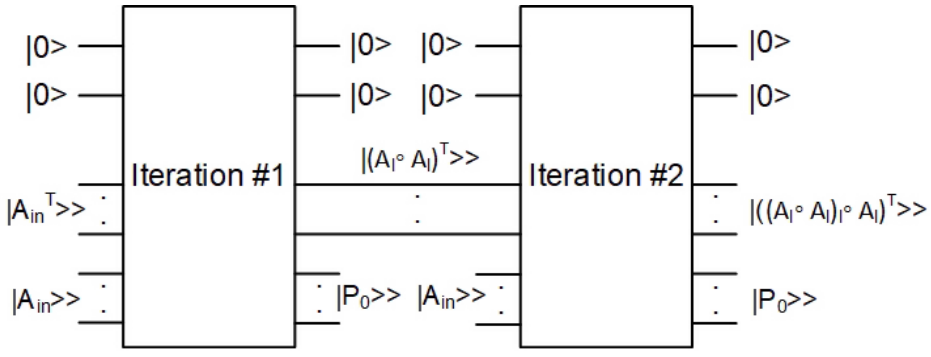
$$|A_{out}\rangle\rangle = \big(\log_2 \mathrm{n} \qquad \text{qubits}\big)\, U_{out} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & ... & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{r\_out} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0... & 0 \\ 0 & 0 & 0 & 0 & 0 & ...0 \end{bmatrix} V_{out}^T \begin{pmatrix} \log_2 \mathrm{n} \\ \\ \text{qubits} \end{pmatrix}$$

2.3.5 Compression process - Iterations

This compression process may be applied iteratively to the input image A_{in}. In particular, each iteration results in an output quantum state vector in the form of $|A_{out}\rangle\rangle = \left|((A_l \circ ... \circ A_l)_l \circ A_l)^T\right\rangle\rangle$, which, in turn, becomes the input quantum system to the next iteration of the algorithm as:

$$|X_0\rangle = |0\rangle \otimes |0\rangle \otimes |A_{out}\rangle\rangle \otimes |A\rangle\rangle$$

The physical implementation of an iterative quantum algorithm may be split into consecutive levels, each corresponding to an iteration of the algorithm, as shown below:



Iterative quantum circuit

It should be mentioned that the oracle value k chosen remains fixed for every iteration. The iterative process of this figure has the following stages:

Iteration #1:

→At first, the algorithm has two inputs, $|A_{in}\rangle\rangle$ and $\left|A_{in}^T\right\rangle\rangle$. Along with two auxiliary qubits set to $|0\rangle$,they form the quantum state $|X_0\rangle = |0\rangle \otimes |0\rangle \otimes \left|A_{in}^T\right\rangle\rangle \otimes |A_{in}\rangle\rangle$.

→Using a boolean function with the parameter k, an oracle decomposes each of these inputs orthonormally into $|A_l\rangle\rangle$,$|A_s\rangle\rangle$ and $\left|A_l^T\right\rangle\rangle$,$\left|A_s^T\right\rangle\rangle$ respectively. A "controlled projective" process then establishes correlation between the these components and the auxiliary qubits.

→The application of the process $\mathcal{V}'$ induces the Hadamard Products between the orthonormal terms.

→ During the output stage, the subsystem containing the Hadamard Product term $\underbrace{\left|(A_l \circ A_l)^T\right\rangle\rangle}$ is extracted.

Iteration #2:

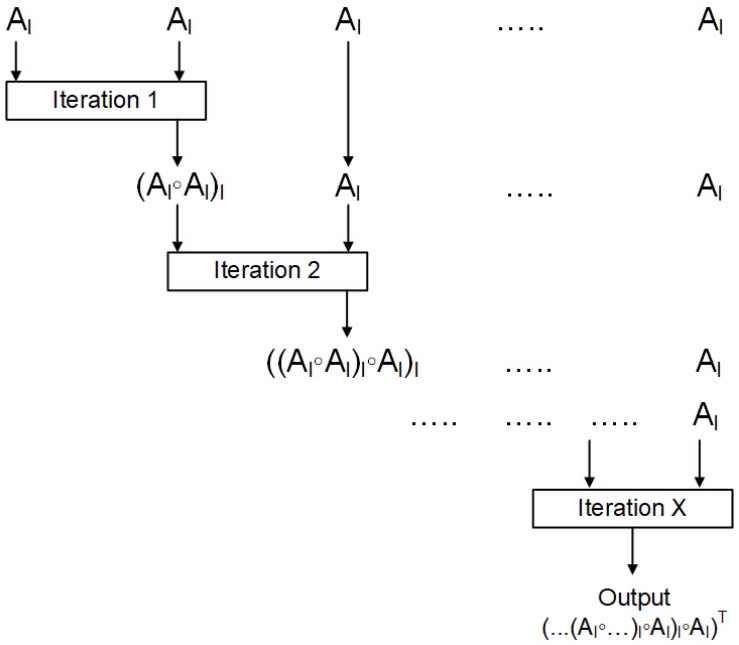
→This time, the inputs of the algorithm are $|A_{in}\rangle\rangle$ and $\underbrace{\left|(A_l \circ A_l)^T\right\rangle\rangle}$. Along with two auxilliary qubits set to $|0\rangle$,they form the quantum state $|X_0\rangle = |0\rangle \otimes |0\rangle \otimes \underbrace{\left|(A_l \circ A_l)^T\right\rangle\rangle} \otimes |A_{in}\rangle\rangle$.

→Using a boolean function with the parameter k, an oracle decomposes each of these inputs orthonormally into $|A_l\rangle\rangle$,$|A_s\rangle\rangle$ and $\left|(A_l \circ A_l)_l^T\right\rangle\rangle$,$\left|(A_l \circ A_l)_s^T\right\rangle\rangle$ respectively. A "controlled projective" process then establishes correlation between the these components and the auxiliary qubits.

→The application of the process $\mathcal{V}'$ induces the Hadamard Products between the orthonormal terms.

→ During the output stage, the subsystem containing the Hadamard Product term $\underbrace{\left|((A_l \circ A_l)_l \circ A_l)^T\right\rangle\rangle}$ is extracted. It can be seen that the matrix $((A_l \circ A_l)_l \circ A_l)^T$ is formed as the Hadamard product of the "large" component of $A_l \circ A_l$, whose rank is equal to k due to the oracle boolean function f, and the "large" component A_l, also of rank k.

In general, the output of this algorithm after a number of iterations is composed of the components shown below:



Output components

For example, the outputs of this algorithm for the first 3 iterations are:
$|A_{out}\rangle\rangle_1 = (A_l \circ A_l)^T$
$|A_{out}\rangle\rangle_2 = ((A_l \circ A_l)_l \circ A_l)^T$
$|A_{out}\rangle\rangle_3 = ((((A_l \circ A_l)_l \circ A_l))_l \circ A_l)^T$

It should be mentioned that this iterative feature of the algorithm causes a great demand in quantum information resources (such as auxiliary qubits).

2.4 Results

After extensive experimentation using the MATLAB interface, a few notable results are presented below:

No.1: "Lena"

The INITIAL image
Dimensions : 512x512
Rank : 507
Entropy : 0.382491

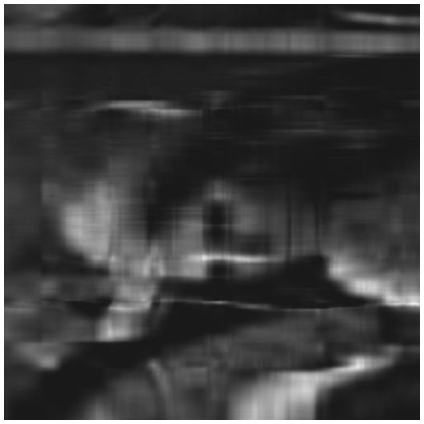
Lena->INITIAL image

The LRA image for $k = 10$
Rank : 10
Entropy : 0.176076
Entropy Percentage of initial : 46.03%

Lena->LRA (k=10)

a BETTER image
Rank : 100
Iterations : 3
Entropy : 0.319821
Entropy Percentage of initial : 83.62%

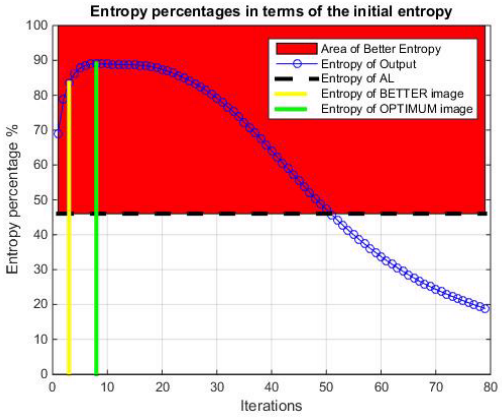
OPTIMUM image
Rank : 100
Iterations : 8
Entropy : 0.340687
Entropy Percentage of initial : 89.07%



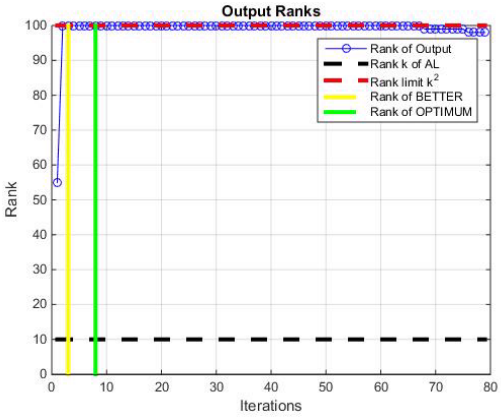
Lena->BETTER image

Lena->OPTIMUM image

The graphs for the entropy percentages with respect to the initial system and the ranks of output vs the number of iterations for the value $k = 10$ are shown below:



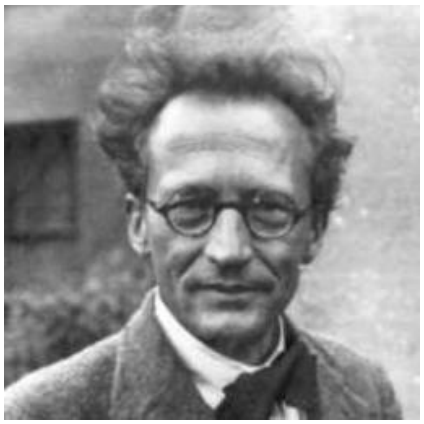
Lena->Entropy Graph (k=10)



Lena->Output RANKS (k=10)

No.2: "Schroedinger"

The INITIAL image
Dimensions : 256x256
Rank : 256
Entropy : 0.299594



Schroedinger->INITIAL image

The LRA image for $k = 6$
Rank : 6
Entropy : 0.142830
Entropy Percentage of initial : 47.67%

Schroedinger->LRA (k=6)

The LRA image for $k = 7$
Rank : 7
Entropy : 0.156270
Entropy Percentage of initial : 52.16%

Schroedinger->LRA (k=7)

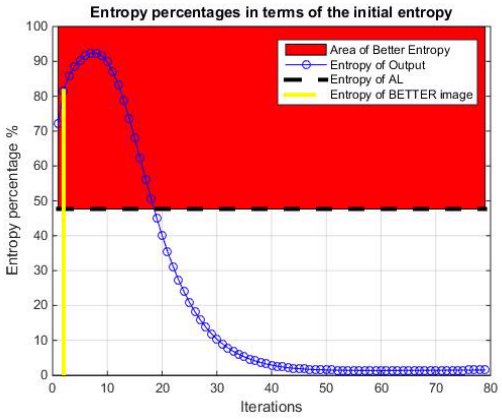
a BETTER image ($k = 6$) ...
Rank : 36
Iterations : 2
Entropy : 0.244263
Entropy Percentage of initial : 81.53%

a BETTER image ($k = 7$) ...
Rank : 49
Iterations : 2
Entropy : 0.263798
Entropy Percentage of initial : 88.05%

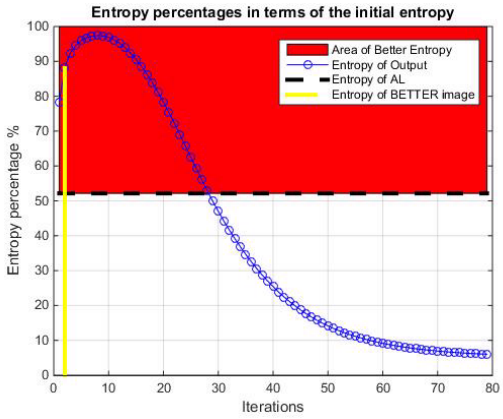
Schroedinger->BETTER(6)

Schroedinger->BETTER(7)

The graphs for the entropy percentages with respect to the initial system vs the number of iterations for the values $k = 6$, $k = 7$ are shown below:



Schroedinger->Entropy Graph (k=6)



Schroedinger->Entropy Graph (k=7)

Some worthmentioning comments regarding the abovementioned results are the following:

→The rank of any output system $|A_{out}\rangle\rangle$ (no matter the number of iterations) is limited by the value of k. Since any output matrix $((A_l \circ ... \circ A_l)_l \circ A_l)^T$ is formed as the Hadamard product of the "large" component $(...(A_l \circ ...)_l \circ A_l)_l$ (whose rank is equal to k due to the oracle boolean function) and the "large" component A_l (also of rank k), the upper bound for its rank is set to k^2, due to the property regarding the Hadamard Product:

$$rank(A \circ B) \leq rank(A)rank(B)$$

→While every image A_{in} is considered to be the bipartite quantum system $|A_{in}\rangle\rangle$, the entropy in question is calculated with respect to its first subsystem, with the use of the reduced density matrix ρ_1. The specific formula used for the purposes of this thesis is the Rényi Entropy of order $a = \frac{1}{2}$. The reason for the assignment of this value to the order a becomes clear in the proof of the following equation.

Consider the Singular Value Decomposition of a real matrix A describing a bipartite quantum system as $A = U_A D_A V_A^T$, with U_A, V_A real orthogonal and D_A diagonal matrices. The Rényi Entropy of order $a = \frac{1}{2}$ with respect to its first subsystem is (*Appendix D - Proposition 8*):

$$S_R(\rho_1) = 2\log_n(Tr(D_A)) = 2\log_n(\sum_{i=1}^{r} \sigma_i) \tag{37}$$

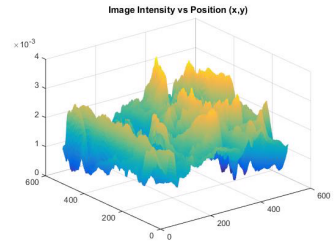
with r being the rank of matrix A and σ_i its (normalized) singular values. This specific formula is used for the calculation of the entropy of the initial,the LRA and every output quantum systems of this algorithm. The base of the logarithm used in the entropy formula is n (the dimension of the image A_{in}), thus limiting the value of the Rényi Entropy between 0 and 1.

→The upper bound for the entropy of a quantum system of dimension n, composed with just k singular values is equal to $\log_n k$, in terms of any Rényi entropy of order a. In general, the entropy of any rank-k approximation of a quantum system is bounded by (*Appendix D - Proposition 9*):

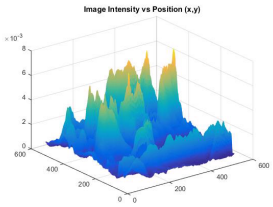
$$0 \leq S_R(\rho_1) \leq \log_n k \tag{38}$$

→The goal for the optimum entropy restoration is achieved at the expense of image quality. This is attributed to the iterative application of the Hadamard Product (aka elementwise multiplication) upon the respective Low Rank Approximation components of $(A_l \circ ... \circ A_l)_l$ and A_l, while preserving the unit norm of the result in the process. The more iterations are needed, the greater the gap between "high" and "low" pixel values becomes. Consider the following

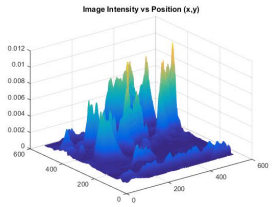
3D representations of intensity (z axis) versus image position (x,y) of the Lena images shown above:



Lena->LRA 3D



Lena->BETTER 3D



Lena->OPTIMUM 3D

The Hadamard Product application results in more pixels being considered as black, with only the relatively high value pixels remaining intact. This way, the fewer visible edges are preserved and so, the more distorted the output image becomes.

It should be mentioned that the pixel value range of the output image A_{out} is properly adjusted to that of the initial image A using a custom MATLAB function.

→The red highlighted areas in the entropy graphs represent the entropy percentage values between those of the Low Rank Approximated system $|A_l\rangle\rangle$ (black lower bound) and the initial image (always 100% - upper bound). It can be seen that, as the rank k of the LRA system increases, the width of this band decreases (see Schroedinger->Entropy Graphs). This is attributed to the increase of the entropy of the LRA system as its rank rises, gradually "bridging the gap" with the entropy of the initial system and practically "closing it" after a certain rank, depending on the image properties and singular value distribution. As a result, the Hadamard Product map as well as its iterative application become more redundant as the value k rises, since the "lion's share" of the initial entropy is already restored in the approximated quantum system $|A_l\rangle\rangle$.

3 Conclusions

In conclusion, it is feasible to compress most of the amount of entanglement contained in a bipartite quantum system in fewer components of its Schmidt decomposition. In the special cases of digital grayscale images A_{in} being the matrix representation of this bipartite system, the proposed algorithm compresses the entropy of the system $|A_{in}\rangle\rangle$, while preserving its main image characteristics as well. This property is attributed to two factors:

- the Schmidt coefficients of a given system $|A_{in}\rangle\rangle$ corresponds to the Singular values of its respective image matrix A_{in}. So, the quality of the Low Rank Approximation of the image A_l is analogous to the quantity of entropy contained in the respective approximation of the initial system $|A_l\rangle\rangle$ for a given rank k.

- the process used iteratively on the grounds of entropy restoration is the Hadamard Product. This type of product preserves a substantial amount of the initial image details, which is proportionate to the number of its iterative applications.

Based on the given results, the parameter k, corresponding to the amount of Schmidt coefficients of the approximated system $|A_l\rangle\rangle$, is of paramount importance to the performance of this algorithm, with a trade-off constantly looming between the image quality of the output and the probability of needing a solution for a better preservation of the entanglement after a few iterations. In particular, small values for k have a higher chance of yielding one or more better solutions regarding the restoration of the initial entropy, with the quality of the respective images being rather poor. On the other hand, large values for k results in fewer chances for a need of a better solution entropy-wise, but with a respective image representation of much greater standards. In fact, the quality of the output image depends greatly on the number of consecutive applications (iterations) of the Hadamard Product $((A_l \circ ... \circ A_l)_l \circ A_l)^T$ needed for a better (or the optimum) ratio of entropy preservation.

It should be mentioned that there is no guarantee for a better solution for every possible value k for any given image A_{in} of rank r. While $k < r$, the rank of the output bounded by k^2 may actually attain an exponentially higher value, even r, thus violating the dimensional reduction criterion of this compression algorithm. As a consequence, only relatively small values for $k << r$ are recommended, so that $k^2 < r$ as well.

Still, the application of this algorithm may be extended to systems described by images of random dimensions. Such double-wedge kets may be the entangling factor between two sets of random number of qubits. Also, the order a of the Rényi Entropy may be fixed ($a = \frac{1}{2}$) for the purposes of this thesis, but it might as well be a contributing factor to the entropy restoration scheme.

Appendix A: The Hadamard product

The Hadamard Product is simply the entrywise multiplication of two or more matrices, according to the following definition:

"Let A and B be $m \cdot n$ matrices with entries in $\mathbb{C}$. The Hadamard product of A and B is defined as $(A \circ B)_{ij} = A_{ij}B_{ij}$ for all $1 \leq i \leq m$, $1 \leq j \leq n$."

The Hadamard Product inherits the same benefits (and restrictions) of multiplication in $\mathbb{C}$ [15]. It is obvious that the matrices involved in a Hadamard Product need to be the same size, but not necessarily square.

Some useful basic properties of the Hadamard Product are the following:

→"Let A and B be $m \cdot n$ matrices with entries in $\mathbb{C}$. Then $A \circ B = B \circ A$."

Proof: This follows directly from the fact that the multiplication in $\mathbb{C}$ is commutative. For every element of the matrices, it can be seen that:

$(A \circ B)_{ij} = A_{ij}B_{ij} = B_{ij}A_{ij} = (B \circ A)_{ij}$, therefore $A \circ B = B \circ A$.

→"The identity matrix under Hadamard Product is the $m \cdot n$ matrix will all entries equal to 1, denoted as J_{mn}. That is $[J_{mn}]_{ij} = 1$ for all $1 \leq i \leq m$, $1 \leq j \leq n$."

Proof: Take any $m \cdot n$ matrix with entries in $\mathbb{C}$. Then $[J_{mn} \circ A]_{ij} = [J_{mn}]_{ij} \cdot [A]_{ij} = 1 \cdot [A]_{ij} = [A]_{ij}$, and so $J_{mn} \circ A = A$. Due to the commutative property of the Hadamard Product, $J_{mn} \circ A = A \circ J_{mn} = A$. Therefore, the matrix J_{mn} is defined as the identity matrix under the Hadamard Product.

→" Let A be an $m \cdot n$ matrix. Then A has a Hadamard inverse, denoted $A^{'}$, if and only if $A_{ij} \neq 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$. Furthermore, $A^{'}_{ij} = (A_{ij})^{-1}$."

Proof: Let A be an $m \cdot n$ matrix with Hadamard inverse A'. This means that $A \circ A' = J_{mn}$. That is $[A \circ A']_{ij} = [A]_{ij} \cdot [A']_{ij} = 1 \Rightarrow [A']_{ij} = \frac{1}{[A]_{ij}} = (A_{ij})^{-1}$, which is only possible when all entries of A are invertible. In other words, $[A]_{ij} \neq 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$.

Also, take any $m \cdot n$ matrix A such that $[A]_{ij} \neq 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$. Then, there exists $(A_{ij})^{-1}$ for all entries i,j. This implies that $A_{ij} \cdot (A_{ij})^{-1} = (A_{ij})^{-1} \cdot A_{ij} = 1$, and so A has a Hadamard inverse A' defined by $A^{'}_{ij} = (A_{ij})^{-1}$ for all i,j.

→(Linearity)"Let $a \in \mathbb{C}$ and A,B,C are $m \cdot n$ matrices. Then $C \circ (A + B) = C \circ A + C \circ B$ and $a(A \circ B) = (aA) \circ B = A \circ (aB)$"

Proof:

As for the first part:

$[C \circ (A + B)]_{ij} = [C]_{ij}[A + B]_{ij}$

$[C \circ (A + B)]_{ij} = [C]_{ij}([A]_{ij} + [B]_{ij})$

$[C \circ (A + B)]_{ij} = [C]_{ij}[A]_{ij} + [C]_{ij}[B]_{ij}$

$[C \circ (A + B)]_{ij} = [C \circ A]_{ij} + [C \circ B]_{ij}$

$[C \circ (A + B)]_{ij} = [C \circ A + C \circ B]_{ij}$

As for the second part:
$[a(A \circ B)]_{ij} = a[A \circ B]_{ij}$
$[a(A \circ B)]_{ij} = a[A]_{ij}[B]_{ij}$
$[a(A \circ B)]_{ij} = [aA]_{ij}[B]_{ij}$
$[a(A \circ B)]_{ij} = [aA \circ B]_{ij}$
and
$[a(A \circ B)]_{ij} = a[A]_{ij}[B]_{ij}$
$[a(A \circ B)]_{ij} = [A]_{ij}a[B]_{ij}$
$[a(A \circ B)]_{ij} = [A]_{ij}[aB]_{ij}$
$[a(A \circ B)]_{ij} = [A \circ aB]_{ij}$

$\rightarrow$"Let A and B be square matrices of size n. Then $rank(A \circ B) \leq rank(A)rank(B)$"
Proof:

For the purposes of this proof, the rank-one decompositions of A and B are required. Suppose A has rank ρ_1 with singular values λ_k, $1 \leq k \leq n$, and B has rank ρ_2 with singular values λ_l, $1 \leq l \leq n$. Then, according to the Singular Value Decomposition:

$$A = \sum_{k=1}^{\rho_1} \lambda_k x_k y_k^T \text{ and } B = \sum_{l=1}^{\rho_2} \lambda_l v_l w_l^T$$

where x_k,v_l the left singular vectors and y_k,w_l the right singular vectors.
Then:
$[A \circ B]_{ij} = [A]_{ij}[B]_{ij}$
$[A \circ B]_{ij} = [\sum_{k=1}^{\rho_1} \lambda_k x_k y_k^T]_{ij}[\sum_{l=1}^{\rho_2} \lambda_l v_l w_l^T]_{ij}$
$[A \circ B]_{ij} = \sum_{k=1}^{\rho_1}\sum_{l=1}^{\rho_2}[\lambda_k x_k y_k^T]_{ij}[\lambda_l v_l w_l^T]_{ij}$
$[A \circ B]_{ij} = \sum_{k=1}^{\rho_1}\sum_{l=1}^{\rho_2}[((\lambda_k x_k) \circ (\lambda_l v_l))(y_k \circ w_l)^T]_{ij}$
$[A \circ B]_{ij} = [\sum_{k=1}^{\rho_1}\sum_{l=1}^{\rho_2}((\lambda_k x_k) \circ (\lambda_l v_l))(y_k \circ w_l)^T]_{ij}$
So, $A \circ B$ has at most rank $\rho_1\rho_2 = rank(A)rank(B)$.

Appendix B: The Singular value decomposition

The Singular Value Decomposition (SVD) is matrix decomposition that was defined many times from independent mathematicians through the ages. It was finally published in its latest form by Eugenio Beltrami in 1873, Gamille Jordan in 1875, J.Sylvester in 1889, L. Autonne in 1913 and C. Eckart and G. Young in 1936. The definition proposed by C.D.Meyer is the following [14]:

"For every matrix $A \in \mathbb{C}^{m \cdot n}$ of order r, there are unitary matrices U of dimension $m \cdot m$, V of dimension $n \cdot n$ and real diagonal matrix D of dimension $r \cdot r$, with $D = diag(\sigma_1, \sigma_2, ..., \sigma_{r-1}, \sigma_r)$ such that

$$A = U \begin{bmatrix} D_{rxr} & 0 \\ 0 & 0 \end{bmatrix} V^{\dagger}$$

with $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_{r-1} \geq \sigma_r$."

The values σ_i are called the singular values of A and are real non-zero values. Also, the column vectors of the unitary matrices U,V are called left- and right-singular vectors respectively.

Low rank matrix approximation

Let the singular value decomposition of a random matrix A be:

$$A = U \begin{bmatrix} D_{rxr} & 0 \\ 0 & 0 \end{bmatrix} V^{\dagger}$$

with D_{rxr} being the diagonal matrix containing the singular values σ_i in a descending order along its main diagonal

$$D_{rxr} = \begin{bmatrix} \sigma_1 & 0 & ... & 0 & 0 \\ 0 & \sigma_2 & ... & 0 & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & \sigma_{r-1} & 0 \\ 0 & 0 & ... & 0 & \sigma_r \end{bmatrix}$$

Theorem: The matrix A may be represented as the sum of r rank-one matrices, as shown below:

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^{\dagger}$$

with the vectors u_j,v_j being the column vectors of the unitary matrices U,V respectively. This theorem implies that every matrix A may be approximated by the sum of rank-one matrices of equal dimension.

Theorem: For every $0 \leq k \leq r$, the following matrix A_k is defined:

$$A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^\dagger$$

This approximation is called Low Rank Matrix Approximation. So, this process leads to a matrix A_k, which is a matrix of rank $k \leq r$. As it turns out, every other matrix B of rank k has a greater $L2-norm$ difference from A than that between A_k and A, meaning that:

$$\|A - A_k\|_2 < \|A - B\|_2$$

A useful property is that the $L2-norm$ of the difference between A and A_k is equal to the $k+1$-th singular value of A, as:

$$\|A - A_k\|_2 = \sigma_{k+1}$$

In the special case that $k = r = \min\{m, n\}$, then $\sigma_{k+1} = 0$, leading to $\|A - A_k\|_2 = \sigma_{k+1} = 0$. In other words, if the approximation process takes into account every rank-one component matrix of A, then the matrix A_k is identical to the matrix A, resulting in a perfect approximation.

Several theorems regarding the Singular Value Decomposition in general are presented below, some of which are accompanied by their proof.

Theorem:
"The non-zero singular values of a matrix A are the sqaure roots of the non-zero eigenvalues of the matrices $AA^\dagger$ and $A^\dagger A$."

Proof:
Let the singular value decomposition of A be $A = UDV^\dagger$, with D having the singular values of A along its main diagonal. Then:
$AA^\dagger = UDV^\dagger(UDV^\dagger)^\dagger$
$AA^\dagger = UDV^\dagger VD^\dagger U^\dagger$
$AA^\dagger = UDD^\dagger U^\dagger$
However, the matrix D is a real diagonal matrix, so:

$$AA^\dagger = UD^2U^\dagger$$

This way, the unitary U resembles a similarity transformation, thus the eigenvalues of $AA^\dagger$ are included in D^2. If λ_i are the eigenvalues of $AA^\dagger$, then

$$\sigma_i^2 = \lambda_i \implies \sigma_i = \sqrt{\lambda_i}$$

Theorem:

"The rank r of a matrix is equal to the number of its non-zero singular values."

Proof:

It is known that the rank of a diagonal matrix is equal to the number of its non-zero elements. Regarding the singular value decomposition of A into $A = UDV^{\dagger}$, it can be said that the matrices U,V are unitary, thus being full rank. The matrix D is diagonal, so the rank of A is the same with the rank of the matrix D, which is equal to the number of its non-zero singular values.

Theorem:

"The Frobenius norm of a matrix A is equal to the square root of the sum of the squared singular values"

$$\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2}$$

Proof:

Let the singular value decomposition of A be $A = UDV^{\dagger}$.Then, its Frobenius norm is equal to:

$$\|A\|_F = \left\|UDV^{\dagger}\right\|_F = \|D\|_F ,$$

since the Frobenius norm is similarity invariant. So:

$$\|A\|_F = \|D\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2}$$

Theorem:

"For every k with $0 \leq k \leq r$, the Frobenius norm of the difference of the matrices A and A_k is equal to the square root of the sum of the squared singular values σ_i for $i \geq k+1$. This means that:

$$\|A - A_k\|_F = \sqrt{\sigma_{k+1}^2 + \sigma_{k+2}^2 + ... + \sigma_r^2}$$

Proof:

Let the singular value decomposition of A be $A = UDV^{\dagger}$. Then, according to the Low Rank Matrix Approximation theorem, the matrices A and A_k may be expressed as:

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^{\dagger} \quad \text{and } A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^{\dagger}$$

with

$$A - A_k = \sum_{j=1}^{r} \sigma_j u_j v_j^\dagger - \sum_{j=1}^{k} \sigma_j u_j v_j^\dagger = \sum_{j=k+1}^{r} \sigma_j u_j v_j^\dagger$$

So, taking the previous theorem into account, the Frobenius norm of the difference between the matrices A and A_k is equal to:

$$\|A - A_k\|_F = \sqrt{\sigma_{k+1}^2 + \sigma_{k+2}^2 + ... + \sigma_r^2}$$

Theorem:

" For every matrix $A \in \mathbb{C}^{NxN}$, the absolute value of its determinant is equal to the product of its singular values."

$$|\det(A)| = \prod_{j=1}^{N} \sigma_j$$

Proof:

Let the singular value decomposition of A be $A = UDV^\dagger$. Then its determinant can be expressed as:

$|\det(A)| = \left|\det(UDV^\dagger)\right| = |\det(U)| \cdot |\det(D)| \cdot \left|\det(V^\dagger)\right|$

Since unitary matrices have determinants equal to one, it can be seen that:

$$|\det(A)| = |\det(D)| = \prod_{j=1}^{N} \sigma_j$$

Theorem:

" Every matrix $A \in \mathbb{C}^{MxN}$ has a singular value decomposition. The singular values σ_i are uniquely defined."

Applications of SVD

The singular value decomposition, along with the low rank matrix approximation theorem, are considered mathematical rules of paramount importance, that may be put to a variety of uses. The two basic features that make them so important include the relatively easy calculation of the singular value decomposition as well as the trimming of the smaller singular values in low rank matrix approximation, while still maintaining the best possible rank k matrix. Besides their contribution on the theoretical field, such properties become extremely useful when dealing with matrices of higher dimensions (e.g. dimensions being in the thousands).

The solution of systems of equations, search engines, data compression-image processing are just a few of the many applications of SVD and Low Rank Approximation. In particular, the application concerning the data compression-image processing is analyzed below:

Suppose that there is an image A that needs compression. It is known that every image A may be expressed as an $m \cdot n$ matrix, whose element values are determined in terms of the respective pixels forming the image. So, in a way, an image is nothing more than a data matrix. The elements of this matrix are real numbers and their values depend on properties of the pixels themselves, such as brightness or color. In the special case of a grayscale image, the matrix element values range from 0 (black) to 255(white), if that image is encoded according to the rule of "8-bit-per-pixel" (leading to $2^8 = 256$ variations of gray).

It can be seen that such a matrix needs $m \cdot n$ real values for its complete definition. However, if its dimension become even greater, the number $m \cdot n$ becomes huge, resulting in the storage and processing of such information being too difficult. The goal of the compression is the approximation of such matrices with fewer elements than $m \cdot n$, thus making its storage and processing much easier.

Let A be a data matrix of dimensions $m \cdot n$, with the following singular value decomposition:

$$A = UDV^{\dagger}$$

with the matrix D including all of the singular values of A, even the zero ones as shown below:

$$D = diag(\sigma_1, \sigma_2, ..., \sigma_{r-1}, \sigma_r, 0, ..., 0)$$

Then, according to the Low Rank Approximation theorem, the matrix A may be approximated by the matrix A_k, that is the sum of rank-one matrices, as shown below:

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^{\dagger} \quad \text{and } A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^{\dagger}$$

with $1 \leq k \leq r$.

It can be seen that the values needed for this approximation include k singular values and $k + k = 2k$ singular vectors. So, recalling that every singular vector involved in a rank-one matrix has m and n values, it is obvious that $km + kn + k = (m + n + 1)k$ values are required for the representation of the low rank approximation A_k. It is quite clear that the number $(m + n + 1)k$ is less than $m \cdot n$, with the difference rising, depending on the choice for k.

While this is a lossy approximation of A, the approximation error is defined as the relative information lost during the process. In particular, the following matrix is defined:

$$S_b = \sum_{j=k+1}^{r} \sigma_j u_j v_j^{\dagger}$$

This matrix represents the information lost to the low rank approximation of a matrix. So, an relative error variable err is defined as shown [16]:

$$err = \frac{\|S_b\|_F}{\|A\|_F} = \sqrt{\frac{\sum_{j=k+1}^{r} \sigma_j^2}{\sum_{i=1}^{r} \sigma_i^2}}$$

So, depending on the margin for error of a specific application, the optimum value for k is chosen.

Appendix C: Entropy-optimum low rank approximation

For the purposes of this section, let A be a real matrix of dimension $n = 2^x$ with singular value decomposition $A = UDV^T$. The reasons for such arbitrary requirements include the consideration of this matrix A as a digital grayscale image (with each element corresponding to a variation of gray) as well as the assumption that this digital image describes a bipartite quantum system, in terms of a double-wedge state vector $|\,A\rangle\rangle$.

It is a well known fact that the Low Rank Matrix Approximation A_k of any matrix A is considered to be the best approximation of A of rank k, with reference to the norm of the difference between A and A_k $\|A - A_k\|$. However, nothing can be deduced about the entropy of the bipartite quantum system, described by the double-wedge ket $|\,A\rangle\rangle$. In the special case of the rank of the approximation being $k = 2$, there is a quantum algorithm that increases the entropy of the quantum system $|\,A_k\rangle\rangle$ with the iterative application of a certain non-local process. Its goal is the increase of the entropy restoration ratio of the approximation as close to 1 as possible. That is why this algorithm results in the entropy-optimum low rank matrix approximation $|\,A_{opt}\rangle\rangle$.

Let the low rank matrix approximation of the matrix A be $A_2 = UD_2V^T = \sum_{i=1}^{2} \sigma_i u_i v_i^T$. This matrix has rank $k = 2$. For such a matrix to correspond to the bipartite quantum system $|\,A_2\rangle\rangle$, its normalization is required, as shown below:

$$\||\,A_2\rangle\rangle\|_2 = \|A_2\|_F = 1 \implies \sqrt{\sum_{i=1}^{2} \sigma_i} = 1 \implies \sigma_1^2 + \sigma_2^2 = 1$$

Such values for σ_1, σ_2 may as well be expressed in terms of sinusoidal functions, since $\sigma_1^2 + \sigma_2^2 = 1 = \cos^2(\theta) + \sin^2(\theta)$ for every $\theta \in [0, 2\pi)$. In particular, since the prerequisite $\sigma_1 \geq \sigma_2$ stands, it is defined that $\sigma_1 \equiv \cos(\theta)$ and $\sigma_2 \equiv \sin(\theta)$, with $\theta \in (0, \frac{\pi}{4}]$, where $\cos(\theta) > \sin(\theta)$ in that value range. So, the bipartite state vector $|\,A_2\rangle\rangle$ may be written as:

$|\,A_2\rangle\rangle = \left|UD_2V^T\right\rangle\rangle$

$|\,A_2\rangle\rangle = (U \otimes V)\,|\,D_2\rangle\rangle$

$$|\,A_2\rangle\rangle = (U \otimes V) \left|\left|\begin{bmatrix} \cos(\theta) & 0 & 0 & ... & 0 \\ 0 & \sin(\theta) & 0 & ... & 0 \\ 0 & 0 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 \end{bmatrix}\right\rangle\right\rangle$$

It turns out that the amount of entanglement included in a quantum system depends on the distribution of its singular values. In particular, the more it resembles the uniform distribution, the higher the level of "uncertainty" of the

system rises, thus the greater its entropy becomes. Using the above representation for the quantum system $|A_2\rangle\rangle$, it can be seen that the increase of the argument θ of the sinusoidal function , up to $\theta_0 = \frac{\pi}{4}$, causes the shrinking of the first singular value $\sigma_1 \equiv \cos(\theta)$ and the increase of the second one $\sigma_2 \equiv \sin(\theta)$, while preserving the various initial restrictions such as $\sigma_1^2 + \sigma_2^2 = 1 \implies \|A_2\|_F = 1$ and $\sigma_1 \geq \sigma_2$. In this way, the distribution of the two singular values resembles the uniform one even more, which leads to the increase in the entropy of the quantum system.

So, for a given angle ϕ_0, the following process may be applied to the system $|A_2\rangle\rangle$,in order to increase its entropy:

$$U_{shift\_gen} = (U \otimes V)(\cos(\phi_0)\mathbb{I}^{\otimes 2\log_2 n} - \sin(\phi_0)(Z \otimes \mathbb{I}^{\otimes 2\log_2 n - 1})X^{\otimes 2\log_2 n})(U^T \otimes V^T)$$

with n being the dimension of the square matrix A ($n = 2^x$). It is obvious that this non-local process is hermitian, but not unitary.

The application of this process is shown below, for the special case of $n = 2$:

$$\begin{aligned}
U_{shift_gen}\,|A_2\rangle\rangle &= [(U \otimes V)(\cos(\phi_0)\mathbb{I}^{\otimes 2} - \sin(\phi_0)(Z \otimes \mathbb{I})X^{\otimes 2})(U^T \otimes V^T)]\,\big|UD_2V^T\big\rangle\big\rangle \\
U_{shift_gen}\,|A_2\rangle\rangle &= [(U \otimes V)(\cos(\phi_0)\mathbb{I}^{\otimes 2} - \sin(\phi_0)(Z \otimes \mathbb{I})X^{\otimes 2})]\,\big|U^TUD_2V^TV\big\rangle\big\rangle \\
U_{shift_gen}\,|A_2\rangle\rangle &= [(U \otimes V)(\cos(\phi_0)\mathbb{I}^{\otimes 2} - \sin(\phi_0)(Z \otimes \mathbb{I})X^{\otimes 2})]\,|D_2\rangle\rangle \\
U_{shift_gen}\,|A_2\rangle\rangle &= (U \otimes V)[\cos(\phi_0)(\mathbb{I} \otimes \mathbb{I})\,|D_2\rangle\rangle - \sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle]
\end{aligned}$$

Below is the separate calculation of the components of the linear combination seen in the above application:

$$\begin{aligned}
\cos(\phi_0)(\mathbb{I} \otimes \mathbb{I})\,|D_2\rangle\rangle &= \cos(\phi_0)\,|\mathbb{I}D_2\mathbb{I}\rangle\rangle \\
\cos(\phi_0)(\mathbb{I} \otimes \mathbb{I})\,|D_2\rangle\rangle &= \cos(\phi_0)\,|D_2\rangle\rangle \\
\cos(\phi_0)(\mathbb{I} \otimes \mathbb{I})\,|D_2\rangle\rangle &= \cos(\phi_0)\left|\begin{bmatrix}\cos(\theta) & 0\\ 0 & \sin(\theta)\end{bmatrix}\right\rangle\right\rangle \\
\cos(\phi_0)(\mathbb{I} \otimes \mathbb{I})\,|D_2\rangle\rangle &= \left|\begin{bmatrix}\cos(\phi_0)\cos(\theta) & 0\\ 0 & \cos(\phi_0)\sin(\theta)\end{bmatrix}\right\rangle\right\rangle
\end{aligned}$$

Also:

$$\begin{aligned}
\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle &= \sin(\phi_0)(Z \otimes \mathbb{I})\,\big|XD_2X^T\big\rangle\big\rangle \\
\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle &= \sin(\phi_0)(Z \otimes \mathbb{I})\left|\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\cdot\begin{bmatrix}\cos(\theta) & 0\\ 0 & \sin(\theta)\end{bmatrix}\cdot\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\right\rangle\right\rangle \\
\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle &= \sin(\phi_0)(Z \otimes \mathbb{I})\left|\begin{bmatrix}\sin(\theta) & 0\\ 0 & \cos(\theta)\end{bmatrix}\right\rangle\right\rangle \\
\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle &= \sin(\phi_0)\left|Z\begin{bmatrix}\sin(\theta) & 0\\ 0 & \cos(\theta)\end{bmatrix}\mathbb{I}\right\rangle\right\rangle \\
\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle &= \sin(\phi_0)\left|\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\cdot\begin{bmatrix}\sin(\theta) & 0\\ 0 & \cos(\theta)\end{bmatrix}\cdot\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right\rangle\right\rangle \\
\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X)\,|D_2\rangle\rangle &= \sin(\phi_0)\left|\begin{bmatrix}\sin(\theta) & 0\\ 0 & -\cos(\theta)\end{bmatrix}\right\rangle\right\rangle
\end{aligned}$$

$$\sin(\phi_0)(Z \otimes \mathbb{I})(X \otimes X) \,|\, D_2\rangle\rangle = \left| \begin{bmatrix} \sin(\phi_0)\sin(\theta) & 0 \\ 0 & -\sin(\phi_0)\cos(\theta) \end{bmatrix} \right\rangle\rangle$$

So:

$$U_{shift\_gen} \,|\, A_2\rangle\rangle = (U \otimes V)[\left| \begin{bmatrix} \cos(\phi_0)\cos(\theta) & 0 \\ 0 & \cos(\phi_0)\sin(\theta) \end{bmatrix} \right\rangle\rangle - \left| \begin{bmatrix} \sin(\phi_0)\sin(\theta) & 0 \\ 0 & -\sin(\phi_0)\cos(\theta) \end{bmatrix} \right\rangle\rangle]$$

$$U_{shift\_gen} \,|\, A_2\rangle\rangle = (U \otimes V) \left| \begin{bmatrix} \cos(\phi_0)\cos(\theta) - \sin(\phi_0)\sin(\theta) & 0 \\ 0 & \cos(\phi_0)\sin(\theta) + \sin(\phi_0)\cos(\theta) \end{bmatrix} \right\rangle\rangle$$

Recall the following trigonometric identities:
$\cos(\phi_0)\cos(\theta) - \sin(\phi_0)\sin(\theta) = \cos(\theta + \phi_0)$
$\cos(\phi_0)\sin(\theta) + \sin(\phi_0)\cos(\theta) = \sin(\theta + \phi_0)$

So:

$$U_{shift\_gen} \,|\, A_2\rangle\rangle = (U \otimes V \left| \begin{bmatrix} \cos(\theta + \phi_0) & 0 \\ 0 & \sin(\theta + \phi_0) \end{bmatrix} \right\rangle\rangle$$

$$U_{shift\_gen} \,|\, A_2\rangle\rangle = (U \otimes V) \,|\, D'\rangle\rangle$$

$$|\, A_{opt}\rangle\rangle = U_{shift\_gen} \,|\, A_2\rangle\rangle = \left|\, UD'V^T \right\rangle\rangle$$

As result, the two singular values σ_1 and σ_2 were transformed from $\cos(\theta)$ and $\sin(\theta)$ to $\cos(\theta + \phi_0)$ and $\sin(\theta + \phi_0)$ respectively. As long as the argument $\theta + \phi_0$ belongs in $(0, \frac{\pi}{4}]$, the D' matrix is a valid singular value matrix, with increased entropy compared to D. It should be noted that this process may be applied to matrices of dimensions $r > 2$ as well, as seen in its general form above.

Moreover, this process is iterative, meaning that consecutive applications of this operator upon the double-wedge ket $|\, A_k\rangle\rangle$ are possible. Each iteration brings the two singular values σ_1 and σ_2 even closer, thus increasing the entropy of the given system. However, it should be noted that it is required that the respective argument $\theta + \phi_0$ be in range $(0, \frac{\pi}{4}]$ after every iteration.

This process results in the entropy-optimum low rank matrix approximation of rank $k = 2$ after a number of iterations, since it causes the entropy of the system to range between its initial approximated value and the maximum value this system may attain. In particular, the entropy of any given system of dimension n and rank $k = 2$ is limited between 0 and $\log_n 2$ with reference to the Rényi entropy formula (*Appendix D - Proposition 9*). So, there are two possible scenarios regarding the entropy of the system:

– If the entropy of the initial system $|\, A\rangle\rangle$ is less than $\log_n 2$, then, after a finite number of iterations, this iterative process will result in the optimum value for the entropy, with an error depending on the given value for ϕ_0 (the smaller ϕ_0 is, the more accurately the initial entropy is approximated).

– If the entropy of the initial system $|\, A\rangle\rangle$ is greater or equal to $\log_n 2$, then the optimum entropy value of the approximated system $|\, A_k\rangle\rangle$ will be assigned to $\log_n 2$ after a number of iterations, with the accuracy depending once again on the value for ϕ_0.

Appendix D: Proofs

The mathematical propositions regarding many defined parameters as well as a few of the more complex properties are presented in this section. Each of them is properly cited throughout this project. These propositions are listed below:

·Proposition 1: Properties of the double-wedge ket notation
·Proposition 2: $\rho_1 = AA^\dagger$ and $\rho_2 = A^T A^*$ for system $|A\rangle\rangle$ (eq.2, 3)
·Proposition 3: Hadamard Product map on matrices (eq. 6)
·Proposition 4: Application of P_{13} upon $(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle$ (eq.31)
·Proposition 5: Application of P_{24} upon $P_{13}(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I})\,|X_0\rangle$ (eq.32)
·Proposition 6: Equivalent formula of $\mathcal{V}$ for quantum systems (eq.33)
·Proposition 7: Application of Hadamard Product map $\mathcal{V}'$ (eq.34)
·Proposition 8: Expression of Rényi Entropy of order $a = \frac{1}{2}$ (eq.37)
·Proposition 9: Upper bound $\log_n k$ for Rényi Entropy (eq.38)

Proposition 1: Properties of the double-wedge ket notation

Consider the double-wedge ket vector $|A\rangle\rangle = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle \otimes |j\rangle$, with d_1,d_2 the dimensions of matrix A and $A_{ij} \in \mathbb{C}$. Also, let C,K be random matrices of size $n \times d_1$ and $m \times d_2$ respectively as shown below:

$$C = \sum_{x=1}^{n}\sum_{y=1}^{d_1} C_{xy}\,|x\rangle\langle y|,\; K = \sum_{p=1}^{m}\sum_{q=1}^{d_2} K_{pq}\,|p\rangle\langle q|$$

Proof that $(C \otimes K)\,|A\rangle\rangle = \left|CAK^T\right\rangle\rangle$:

$$(C \otimes K)\,|A\rangle\rangle = [(\sum_{x=1}^{n}\sum_{y=1}^{d_1} C_{xy}\,|x\rangle\langle y|) \otimes (\sum_{p=1}^{m}\sum_{q=1}^{d_2} K_{pq}\,|p\rangle\langle q|)](\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle \otimes |j\rangle)$$

$$(C \otimes K)\,|A\rangle\rangle = (\sum_{x=1}^{n}\sum_{y=1}^{d_1}\sum_{p=1}^{m}\sum_{q=1}^{d_2} C_{xy}K_{pq}\,|x\rangle\langle y| \otimes |p\rangle\langle q|)(\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle \otimes |j\rangle)$$

$$(C \otimes K)\,|A\rangle\rangle = \sum_{x=1}^{n}\sum_{y=1}^{d_1}\sum_{p=1}^{m}\sum_{q=1}^{d_2}\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} C_{xy}K_{pq}A_{ij}\,|x\rangle\langle y|\,|i\rangle \otimes |p\rangle\langle q|\,|j\rangle$$

The inner product $\langle y|\,|i\rangle$ is equal to 1 iff $y = i \in [1, d_1]$. Similarly, the inner product $\langle q|\,|j\rangle$ is equal to 1 iff $q = j \in [1, d_2]$. So:

$$(C \otimes K) \left| A \right\rangle\rangle = \sum_{x=1}^{n} \sum_{p=1}^{m} \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} C_{xi} K_{pj} A_{ij} \left| x \right\rangle \otimes \left| p \right\rangle$$

$$(C \otimes K) \left| A \right\rangle\rangle = \sum_{x=1}^{n} \sum_{p=1}^{m} \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} C_{xi} A_{ij} K_{jp}^T \left| x \right\rangle \otimes \left| p \right\rangle$$

$$(C \otimes K) \left| A \right\rangle\rangle = \sum_{x=1}^{n} \sum_{p=1}^{m} \sum_{i=1}^{d_1} C_{xi} \cdot (\sum_{j=1}^{d_2} A_{ij} K_{jp}^T) \left| x \right\rangle \otimes \left| p \right\rangle$$

$$(C \otimes K) \left| A \right\rangle\rangle = \sum_{x=1}^{n} \sum_{p=1}^{m} (\sum_{i=1}^{d_1} C_{xi} (AK^T)_{ip}) \left| x \right\rangle \otimes \left| p \right\rangle$$

$$(C \otimes K) \left| A \right\rangle\rangle = \sum_{x=1}^{n} \sum_{p=1}^{m} (CAK^T)_{xp} \left| x \right\rangle \otimes \left| p \right\rangle$$

$$(C \otimes K) \left| A \right\rangle\rangle = \left| CAK^T \right\rangle\rangle$$

For $K = \mathbb{I}$, then $(C \otimes \mathbb{I}) \left| A \right\rangle\rangle = \left| CA\mathbb{I}^T \right\rangle\rangle = \left| CA \right\rangle\rangle$.
For $C = \mathbb{I}$, then $(\mathbb{I} \otimes K) \left| A \right\rangle\rangle = \left| \mathbb{I}AK^T \right\rangle\rangle = \left| AK^T \right\rangle\rangle$.

Proof that $\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{\langle\langle A \left| \right| A \rangle\rangle} = \sqrt{Tr(A^\dagger A)} = \left\| A \right\|_F$:

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{\langle\langle A \left| \right| A \rangle\rangle}$$

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{(\sum_{i=1}^{d_1} \sum_{j=1}^{d_2} A_{ij} |i\rangle \otimes |j\rangle)^\dagger (\sum_{x=1}^{d_1} \sum_{y=1}^{d_2} A_{xy} |x\rangle \otimes |y\rangle)}$$

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{(\sum_{i=1}^{d_1} \sum_{j=1}^{d_2} A_{ij}^* \langle i| \otimes \langle j|)(\sum_{x=1}^{d_1} \sum_{y=1}^{d_2} A_{xy} |x\rangle \otimes |y\rangle)}$$

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{\sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \sum_{x=1}^{d_1} \sum_{y=1}^{d_2} A_{ij}^* A_{xy} \langle i| \left| x \right\rangle \cdot \langle j| \left| y \right\rangle}$$

The inner product $\langle i| \left| x \right\rangle$ is equal to 1 iff $x = i \in [1, d_1]$. Similarly, the inner product $\langle j| \left| y \right\rangle$ is equal to 1 iff $y = j \in [1, d_2]$. So:

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{\sum_{i=1}^{d_1} \sum_{j=1}^{d_2} A_{ij}^* A_{ij}}$$

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{\sum_{j=1}^{d_2} (\sum_{i=1}^{d_1} A_{ji}^\dagger A_{ij})}$$

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{\sum_{j=1}^{d_2} (A^\dagger A)_{jj}}$$

$$\left\| \left| A \right\rangle\rangle \right\|_2 = \sqrt{Tr(A^\dagger A)}$$

At this point, consider the Singular Value Decomposition of $A = UDV^\dagger$, with U,V unitary matrices and D diagonal matrix containing the singular values σ_i of A.

$\||\,A\rangle\rangle\|_2 = \sqrt{Tr[(UDV^\dagger)^\dagger(UDV^\dagger)]}$
$\||\,A\rangle\rangle\|_2 = \sqrt{Tr(VD^2V^\dagger)}$ (U unitary and D diagonal)
$\||\,A\rangle\rangle\|_2 = \sqrt{Tr(D^2)}$ (since trace is similarity-invariant)

$$\||\,A\rangle\rangle\|_2 = \sqrt{\sum_i \sigma_i^2} = \|A\|_F$$

Consider another double-wedge ket vector $|\,B\rangle\rangle = \sum_{k=1}^{d_1}\sum_{l=1}^{d_2} B_{kl}|k\rangle \otimes |l\rangle$, with d_1,d_2 the dimensions of matrix B and $B_{kl} \in \mathbb{C}$.

Proof that $\langle\langle A\,|\,B\rangle\rangle = Tr(A^\dagger B)$:

$$\langle\langle A\,|\,B\rangle\rangle = (\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle \otimes |j\rangle)^\dagger(\sum_{k=1}^{d_1}\sum_{l=1}^{d_2} B_{kl}|k\rangle \otimes |l\rangle)$$
$$\langle\langle A\,|\,B\rangle\rangle = (\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}^* \langle i| \otimes \langle j|)(\sum_{k=1}^{d_1}\sum_{l=1}^{d_2} B_{kl}|k\rangle \otimes |l\rangle)$$
$$\langle\langle A\,|\,B\rangle\rangle = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\sum_{k=1}^{d_1}\sum_{l=1}^{d_2} A_{ij}^* B_{kl} \langle i|\,|k\rangle \cdot \langle j|\,|l\rangle$$

The inner product $\langle i|\,|k\rangle$ is equal to 1 iff $k = i \in [1, d_1]$. Similarly, the inner product $\langle j|\,|l\rangle$ is equal to 1 iff $l = j \in [1, d_2]$. So:

$$\langle\langle A\,|\,B\rangle\rangle = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}^* B_{ij}$$
$$\langle\langle A\,|\,B\rangle\rangle = \sum_{j=1}^{d_2}(\sum_{i=1}^{d_1} A_{ji}^\dagger B_{ij})$$
$$\langle\langle A\,|\,B\rangle\rangle = \sum_{j=1}^{d_2}(A^\dagger B)_{jj}$$
$$\langle\langle A\,|\,B\rangle\rangle = Tr(A^\dagger B)$$

Finally, the property $|\,A+B\rangle\rangle = |\,A\rangle\rangle + |\,B\rangle\rangle$ is an obvious one as:

$$|\,A\rangle\rangle + |\,B\rangle\rangle = (\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle \otimes |j\rangle) + (\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} B_{ij}|i\rangle \otimes |j\rangle)$$
$$|\,A\rangle\rangle + |\,B\rangle\rangle = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}(A_{ij} + B_{ij})|i\rangle \otimes |j\rangle$$
$$|\,A\rangle\rangle + |\,B\rangle\rangle = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}(A+B)_{ij}|i\rangle \otimes |j\rangle = |\,A+B\rangle\rangle$$

■

Proposition 2: $\rho_1 = AA^\dagger$ and $\rho_2 = A^T A^*$ for system $|\,A\rangle\rangle$ (eq.2, 3)

Let the matrix $A = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}\,|i\rangle\langle j|$ of dimensions $d_1 \times d_2$ be the representation of a double-wedge ket of a bipartite quantum system. Due to the isomorphism of matrices $A \in \mathbb{C}^{d_1\times d_2}$ and vectors $|\,A\rangle\rangle \in \mathbb{C}^{d_1\times d_2}$:

$$|\,A\rangle\rangle = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle\otimes|j\rangle$$

As a result, any partial trace a of the density matrix ρ may be expressed as:

$$Tr_a(|\,A\rangle\rangle\langle\langle A\,|) = Tr_a[(\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle\otimes|j\rangle)(\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{xy}|x\rangle\otimes|y\rangle)^\dagger]$$

$$Tr_a(|\,A\rangle\rangle\langle\langle A\,|) = Tr_a[(\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} A_{ij}|i\rangle\otimes|j\rangle)(\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A^*_{xy}\langle x|\otimes\langle y|)]$$

$$Tr_a(|\,A\rangle\rangle\langle\langle A\,|) = Tr_a(\sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{ij}A^*_{xy}|i\rangle\langle x|\otimes|j\rangle\langle y|)$$

So:

$\rightarrow a = 2$

$$\rho_1 = Tr_2(|\,A\rangle\rangle\langle\langle A\,|)$$

$$\rho_1 = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{ij}A^*_{xy}|i\rangle\langle x|\cdot Tr(|j\rangle\langle y|)$$

$$\rho_1 = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{ij}A^*_{xy}|i\rangle\langle x|\cdot\langle y|\,|j\rangle$$

Since $y, j \in [1, d_2]$, the inned product $\langle y|\,|j\rangle$ is equal to 1 iff $j = y \in [1, d_2]$. So:

$$\rho_1 = \sum_{i=1}^{d_1}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{iy}A^*_{xy}|i\rangle\langle x|$$

$$\rho_1 = \sum_{i=1}^{d_1}\sum_{x=1}^{d_1}(\sum_{y=1}^{d_2} A_{iy}A^\dagger_{yx})|i\rangle\langle x|$$

$$\rho_1 = AA^\dagger$$

$\rightarrow a = 1$

$\rho_2 = Tr_1(|A\rangle\rangle \langle\langle A|)$

$\rho_2 = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{ij}A^*_{xy} \cdot Tr(|i\rangle \langle x|) \otimes |j\rangle \langle y|$

$\rho_2 = \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{ij}A^*_{xy} \langle x|i\rangle \otimes |j\rangle \langle y|$

Since $y, j \in [1, d_1]$, the inned product $\langle x|i\rangle$ is equal to 1 iff $i = x \in [1, d_1]$. So:

$\rho_2 = \sum_{j=1}^{d_2}\sum_{x=1}^{d_1}\sum_{y=1}^{d_2} A_{xj}A^*_{xy}|j\rangle \langle y|$

$\rho_2 = \sum_{j=1}^{d_2}\sum_{y=1}^{d_2}(\sum_{x=1}^{d_1} A^T_{jx}A^*_{xy})|j\rangle \langle y|$

$$\rho_2 = A^T A^*$$

■

Proposition 3: Hadamard Product map on matrices (eq. 6)

Consider the following matrices:

$$A = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy} |x\rangle \langle y|$$

$$B = \sum_{k=0}^{n-1}\sum_{l=0}^{n-1} B_{kl} |k\rangle \langle l|$$

with $A_{xy}, B_{kl} \in \mathbb{C}$ for every (x, y, k, l).

Recall that $Ad(\mathcal{V}) = Ad((\mathbb{I} \otimes P_0)U_{xor})$. The explicit form of each operation involved is presented below:

$\mathbb{I} \otimes P_0 = \sum_{q=0}^{n-1} |q\rangle \langle q| \otimes |0\rangle \langle 0|$, since $\mathbb{I} = \sum_{q=0}^{n-1} |q\rangle \langle q|$ and $P_0 = |0\rangle \langle 0|$

and

$U_{xor} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(|i\rangle \otimes |i \ominus_n j\rangle)(\langle i| \otimes \langle j|) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |i\rangle \langle i| \otimes |i \ominus_n j\rangle \langle j|$

So:

$\mathcal{V} = (\mathbb{I} \otimes P_0)U_{xor}$

$\mathcal{V} = (\sum_{q=0}^{n-1} |q\rangle \langle q| \otimes |0\rangle \langle 0|)(\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |i\rangle \langle i| \otimes |i \ominus_n j\rangle \langle j|)$

$$\mathcal{V} = \sum_{q=0}^{n-1}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |q\rangle\langle q|\,|i\rangle\langle i| \otimes |0\rangle\langle 0|\,|i \ominus_n j\rangle\langle j|$$

Since $q \in [0, n-1]$ and $i \in [0, n-1]$, the inner product $\langle q|\,|i\rangle$ equals 1 only for $q = i \in [0, n-1]$:

$$\mathcal{V} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |i\rangle\langle i| \otimes |0\rangle\langle 0|\,|i \ominus_n j\rangle\langle j|$$

Based on the third property of the generalized quantum XOR gate, $i \ominus_n j = 0$ if and only if $i = j$, since $i, j \in [0, n-1]$. So, the inner product $\langle 0|\,|i \ominus_n j\rangle$ equals 1 only for $i = j$:

$$\mathcal{V} = \sum_{i=0}^{n-1} |i\rangle\langle i| \otimes |0\rangle\langle i|$$

Also:

$$A \otimes B = (\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy}\,|x\rangle\langle y|) \otimes (\sum_{k=0}^{n-1}\sum_{l=0}^{n-1} B_{kl}\,|k\rangle\langle l|)$$

$$A \otimes B = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{kl}\,|x\rangle\langle y| \otimes |k\rangle\langle l|$$

In addition:

$$\mathcal{V}^\dagger = [(\mathbb{I} \otimes P_0)U_{xor}]^\dagger$$

$$\mathcal{V}^\dagger = [\sum_{i=0}^{n-1} |i\rangle\langle i| \otimes |0\rangle\langle i|]^\dagger$$

$$\mathcal{V}^\dagger = \sum_{i=0}^{n-1} |i\rangle\langle i| \otimes |i\rangle\langle 0|$$

Moreover:

$$\mathcal{V}(A \otimes B) = (\sum_{i=0}^{n-1} |i\rangle\langle i| \otimes |0\rangle\langle i|)(\sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{kl}\,|x\rangle\langle y| \otimes |k\rangle\langle l|)$$

$$\mathcal{V}(A \otimes B) = \sum_{i=0}^{n-1}\sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{kl}\,|i\rangle\langle i|\,|x\rangle\langle y| \otimes |0\rangle\langle i|\,|k\rangle\langle l|$$

Since $i \in [0, n-1]$ and $x \in [0, n-1]$, the inner product $\langle i|\,|x\rangle$ equals 1 only for $i = x \in [0, n-1]$:

$$\mathcal{V}(A \otimes B) = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{kl}\,|x\rangle\langle y| \otimes |0\rangle\langle x|\,|k\rangle\langle l|$$

Also, the inner product $\langle x|\,|k\rangle$ equals 1 only for $k = x \in [0, n-1]$:

$$\mathcal{V}(A \otimes B) = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{xl}\,|x\rangle\langle y| \otimes |0\rangle\langle l|$$

Finally:

$Ad(\mathcal{V})(A \otimes B) = \mathcal{V}(A \otimes B)\mathcal{V}^\dagger$

$Ad(\mathcal{V})(A \otimes B) = (\sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{xl} |x\rangle \langle y| \otimes |0\rangle \langle l|)(\sum_{p=0}^{n-1} |p\rangle \langle p| \otimes |p\rangle \langle 0|)$

$Ad(\mathcal{V})(A \otimes B) = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{l=0}^{n-1}\sum_{p=0}^{n-1} A_{xy}B_{xl} |x\rangle \langle y|p\rangle \langle p| \otimes |0\rangle \langle l|p\rangle \langle 0|$

Since $y \in [0, n-1]$ and $p \in [0, n-1]$, the inner product $\langle y|p\rangle$ equals 1 only for $p = y \in [0, n-1]$:

$Ad(\mathcal{V})(A \otimes B) = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1}\sum_{l=0}^{n-1} A_{xy}B_{xl} |x\rangle \langle y| \otimes |0\rangle \langle l|y\rangle \langle 0|$

Also, for $l = y \in [0, n-1]$:

$Ad(\mathcal{V})(A \otimes B) = \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy}B_{xy} |x\rangle \langle y| \otimes |0\rangle \langle 0|$

$Ad(\mathcal{V})(A \otimes B) = (\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy}B_{xy} |x\rangle \langle y|) \otimes |0\rangle \langle 0|$

At this point, it can be noticed that:

$\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy}B_{xy} |x\rangle \langle y| = (\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy} |x\rangle \langle y|) \circ (\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} B_{xy} |x\rangle \langle y|)$

$\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} A_{xy}B_{xy} |x\rangle \langle y| = A \circ B$

with $\circ$ denoting the elementwise multiplication (or Hadamard Product) of the matrices involved. So:

$Ad(\mathcal{V})(A \otimes B) = (A \circ B) \otimes |0\rangle \langle 0|$

$$Ad(\mathcal{V})(A \otimes B) = (A \circ B) \otimes P_0$$

■

Proposition 4: Application of P_{13} upon $(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I}) |X_0\rangle$ (eq.31)

$P_{13}(H \otimes H \otimes \mathbb{I} \otimes \mathbb{I}) |X_0\rangle = (P_0 \otimes \mathbb{I} \otimes P_{large} \otimes \mathbb{I} + P_1 \otimes \mathbb{I} \otimes P_{small} \otimes \mathbb{I}) \cdot$

$\cdot(\frac{1}{2}(|00\rangle \otimes |A_{in}^T\rangle\rangle \otimes |A_{in}\rangle\rangle + |01\rangle \otimes |A_{in}^T\rangle\rangle \otimes |A_{in}\rangle\rangle + |10\rangle \otimes |A_{in}^T\rangle\rangle \otimes |A_{in}\rangle\rangle + |11\rangle \otimes |A_{in}^T\rangle\rangle \otimes |A_{in}\rangle\rangle))$

$$P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle = \frac{1}{2}(P_0\,|0\rangle\otimes\mathbb{I}\,|0\rangle\otimes P_{large}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle +$$
$$+P_0\,|0\rangle\otimes\mathbb{I}\,|1\rangle\otimes P_{large}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle + P_0\,|1\rangle\otimes\mathbb{I}\,|0\rangle\otimes P_{large}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle +$$
$$+P_0\,|1\rangle\otimes\mathbb{I}\,|1\rangle\otimes P_{large}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle + P_1\,|0\rangle\otimes\mathbb{I}\,|0\rangle\otimes P_{small}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle +$$
$$+P_1\,|0\rangle\otimes\mathbb{I}\,|1\rangle\otimes P_{small}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle + P_1\,|1\rangle\otimes\mathbb{I}\,|0\rangle\otimes P_{small}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle +$$
$$+ P_1\,|1\rangle\otimes\mathbb{I}\,|1\rangle\otimes P_{small}\,\big|\,A_{in}^T\big\rangle\big\rangle\otimes\mathbb{I}\,|\,A_{in}\rangle\rangle)$$

Recall that:
$P_0\,|0\rangle = |0\rangle$, $P_1\,|1\rangle = |1\rangle$, $P_0\,|1\rangle = 0$, $P_1\,|0\rangle = 0$
and
$P_{large}\,\big|\,A_{in}^T\big\rangle\big\rangle = c_l\,\big|\,A_l^T\big\rangle\big\rangle$, $P_{small}\,\big|\,A_{in}^T\big\rangle\big\rangle = c_s\,\big|\,A_s^T\big\rangle\big\rangle$
So:

$$P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle = \frac{1}{2}(c_l\,|00\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + c_l\,|01\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle +$$
$$+ c_s\,|10\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + c_s\,|11\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle)$$

Proposition 5: Application of P_{24} upon $P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle$ (eq.32)

$$P_{24}P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle = (\mathbb{I}\otimes P_0\otimes\mathbb{I}\otimes P_{large} + \mathbb{I}\otimes P_1\otimes\mathbb{I}\otimes P_{small})\cdot$$
$$\cdot\,[\frac{1}{2}(c_l\,|00\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + c_l\,|01\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle +$$
$$+ c_s\,|10\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle + c_s\,|11\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_{in}\rangle\rangle)]$$

$$P_{24}P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle = \frac{1}{2}(c_l\mathbb{I}\,|0\rangle\otimes P_0\,|0\rangle\otimes\mathbb{I}\,\big|\,A_l^T\big\rangle\big\rangle\otimes P_{large}\,|\,A_{in}\rangle\rangle +$$
$$+c_l\mathbb{I}\,|0\rangle\otimes P_0\,|1\rangle\otimes\mathbb{I}\,\big|\,A_l^T\big\rangle\big\rangle\otimes P_{large}\,|\,A_{in}\rangle\rangle + c_s\mathbb{I}\,|1\rangle\otimes P_0\,|0\rangle\otimes\mathbb{I}\,\big|\,A_s^T\big\rangle\big\rangle\otimes P_{large}\,|\,A_{in}\rangle\rangle +$$
$$+c_s\mathbb{I}\,|1\rangle\otimes P_0\,|1\rangle\otimes\mathbb{I}\,\big|\,A_s^T\big\rangle\big\rangle\otimes P_{large}\,|\,A_{in}\rangle\rangle + c_l\mathbb{I}\,|0\rangle\otimes P_1\,|0\rangle\otimes\mathbb{I}\,\big|\,A_l^T\big\rangle\big\rangle\otimes P_{small}\,|\,A_{in}\rangle\rangle +$$
$$+c_l\mathbb{I}\,|0\rangle\otimes P_1\,|1\rangle\otimes\mathbb{I}\,\big|\,A_l^T\big\rangle\big\rangle\otimes P_{small}\,|\,A_{in}\rangle\rangle + c_s\mathbb{I}\,|1\rangle\otimes P_1\,|0\rangle\otimes\mathbb{I}\,\big|\,A_s^T\big\rangle\big\rangle\otimes P_{small}\,|\,A_{in}\rangle\rangle +$$
$$+ c_s\mathbb{I}\,|1\rangle\otimes P_1\,|1\rangle\otimes\mathbb{I}\,\big|\,A_s^T\big\rangle\big\rangle\otimes P_{small}\,|\,A_{in}\rangle\rangle)$$

Recall that:
$P_0\,|0\rangle = |0\rangle$, $P_1\,|1\rangle = |1\rangle$, $P_0\,|1\rangle = 0$, $P_1\,|0\rangle = 0$

and
$P_{large}\,|\,A_{in}\rangle\rangle = c_l\,|\,A_l\rangle\rangle$, $P_{small}\,|\,A_{in}\rangle\rangle = c_s\,|\,A_s\rangle\rangle$
So:

$$P_{24}P_{13}(H\otimes H\otimes\mathbb{I}\otimes\mathbb{I})\,|X_0\rangle = \frac{1}{2}(c_l^2\,|00\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_l\rangle\rangle + c_l c_s\,|10\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_l\rangle\rangle + \\ + c_l c_s\,|01\rangle\otimes\big|\,A_l^T\big\rangle\big\rangle\otimes|\,A_s\rangle\rangle + c_s^2\,|11\rangle\otimes\big|\,A_s^T\big\rangle\big\rangle\otimes|\,A_s\rangle\rangle)$$

■

Proposition 6: Equivalent formula of $\mathcal{V}$ for quantum systems (eq.33)

The need arises for the calculation of an equivalent formula $\mathcal{V}'$ for the adjoint action of $\mathcal{V}$, that is suitable for application upon composite quantum systems in the form of

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle\otimes|\,B\rangle\rangle) = \mathcal{V}'(\mathbb{I}\otimes A\otimes B\otimes\mathbb{I})(|\,\mathbb{I}\rangle\rangle\otimes|\,\mathbb{I}\rangle\rangle)$$

with A, B square complex matrices of the same dimension n.

Recall that $Ad(\mathcal{V}) = Ad((\mathbb{I}\otimes P_0)U_{xor})$. The explicit form of each operation involved is presented below:

$\mathbb{I}\otimes P_0 = \sum_{q=0}^{n-1}|q\rangle\langle q|\otimes|0\rangle\langle 0|$, since $\mathbb{I} = \sum_{q=0}^{n-1}|q\rangle\langle q|$ and $P_0 = |0\rangle\langle 0|$

and

$U_{xor} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(|i\rangle\otimes|i\ominus_n j\rangle)(\langle i|\otimes\langle j|) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}|i\rangle\langle i|\otimes|i\ominus_n j\rangle\langle j|$

So:

$\mathcal{V} = (\mathbb{I}\otimes P_0)U_{xor}$

$\mathcal{V} = (\sum_{q=0}^{n-1}|q\rangle\langle q|\otimes|0\rangle\langle 0|)(\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}|i\rangle\langle i|\otimes|i\ominus_n j\rangle\langle j|)$

$\mathcal{V} = \sum_{q=0}^{n-1}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}|q\rangle\langle q|\,|i\rangle\langle i|\otimes|0\rangle\langle 0|\,|i\ominus_n j\rangle\langle j|$

Since $q\in[0,n-1]$ and $i\in[0,n-1]$, the inner product $\langle q|\,|i\rangle$ equals 1 only for $q=i\in[0,n-1]$:

$\mathcal{V} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}|i\rangle\langle i|\otimes|0\rangle\langle 0|\,|i\ominus_n j\rangle\langle j|$

Based on the third property of the generalized quantum XOR gate, $i\ominus_n j = 0$ if and only if $i=j$, since $i,j\in[0,n-1]$. So, the inner product $\langle 0|\,|i\ominus_n j\rangle$ equals 1 only for $i=j$:

$$\mathcal{V} = \sum_{i=0}^{n-1} |i\rangle \langle i| \otimes |0\rangle \langle i|$$

Also:

$$\mathcal{V}^\dagger = \sum_{j=0}^{n-1} |j\rangle \langle j| \otimes |j\rangle \langle 0|$$

It is obvious that:
$(\mathbb{I} \otimes Ad(\mathcal{V})(A \otimes B) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle) = (\mathbb{I} \otimes (A \circ B) \otimes P_0 \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)$
$(\mathbb{I} \otimes Ad(\mathcal{V})(A \otimes B) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle) = \big|\,(A \circ B)^T\big\rangle\big\rangle \otimes |\,P_0\rangle\rangle$

So, the explicit form of the action of $\mathcal{V}' = \mathbb{I} \otimes Ad(\mathcal{V}) \otimes \mathbb{I}$ upon the composite quantum system $\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle$ is:
$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = \mathcal{V}'(\mathbb{I} \otimes A \otimes B \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)$
$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = (\mathbb{I} \otimes Ad(\mathcal{V}) \otimes \mathbb{I})(\mathbb{I} \otimes A \otimes B \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)$
$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = (\mathbb{I} \otimes Ad(\mathcal{V})(A \otimes B) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)$
$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = (\mathbb{I} \otimes \mathcal{V}(A \otimes B)\mathcal{V}^\dagger \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = (\mathbb{I} \otimes (\sum_{i=0}^{n-1} |i\rangle \langle i| \otimes |0\rangle \langle i|)(A \otimes B) \cdot \cdot (\sum_{j=0}^{n-1} |j\rangle \langle j| \otimes |j\rangle \langle 0|) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle))$$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = (\mathbb{I} \otimes (\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |i\rangle \langle i|\,A\,|j\rangle \langle j| \otimes |0\rangle \langle i|\,B\,|j\rangle \langle 0|) \otimes \mathbb{I}) \cdot \cdot (|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle))$$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = (\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \mathbb{I} \otimes |i\rangle \langle i|\,A\,|j\rangle \langle j| \otimes |0\rangle \langle i|\,B\,|j\rangle \langle 0| \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle))$$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} (\mathbb{I} \otimes |i\rangle \langle i|\,A\,|j\rangle \langle j|)\,|\,\mathbb{I}\rangle\rangle \otimes (|0\rangle \langle i|\,B\,|j\rangle \langle 0| \otimes \mathbb{I})\,|\,\mathbb{I}\rangle\rangle$$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} [\big|\,(|i\rangle \langle i|\,A\,|j\rangle \langle j|)^T\big\rangle\big\rangle \otimes |\,|0\rangle \langle i|\,B\,|j\rangle \langle 0|\rangle\rangle]$$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} (\big|\,|j\rangle \langle j|\,A^T\,|i\rangle \langle i|\big\rangle\big\rangle \otimes |\,|0\rangle \langle i|\,B\,|j\rangle \langle 0|\rangle\rangle)$$

$$\mathcal{V}'(\big|\,A^T\big\rangle\big\rangle \otimes |\,B\rangle\rangle) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} (|j\rangle \langle j| \otimes (|i\rangle \langle i|)^T)\,\big|\,A^T\big\rangle\big\rangle \otimes (|0\rangle \langle i| \otimes (|j\rangle \langle 0|)^T)\,|\,B\rangle\rangle$$

$$\mathcal{V}'(\left|\, A^T \right\rangle\rangle \otimes |\, B\rangle\rangle) = (\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |j\rangle\langle j| \otimes |i\rangle\langle i| \otimes |0\rangle\langle i| \otimes |0\rangle\langle j|)(\left|\, A^T \right\rangle\rangle \otimes |\, B\rangle\rangle)$$

As a result, the formula for $\mathcal{V}'$ is:

$$\mathcal{V}' = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |j\rangle\langle j| \otimes |i\rangle\langle i| \otimes |0\rangle\langle i| \otimes |0\rangle\langle j|$$

■

Proposition 7: Application of Hadamard Product map $\mathcal{V}'$ (eq.34)

$|X_2\rangle = (\mathbb{I}_2 \otimes \mathcal{V}')\,|X_1\rangle$

$$|X_2\rangle = [\mathbb{I}_2 \otimes \mathcal{V}'][\frac{1}{2}(c_l^2\,|00\rangle \otimes (\mathbb{I} \otimes (A_l \otimes A_l) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+c_l c_s\,|10\rangle\otimes(\mathbb{I}\otimes(A_s\otimes A_l)\otimes\mathbb{I})(|\,\mathbb{I}\rangle\rangle\otimes|\,\mathbb{I}\rangle\rangle)+c_l c_s\,|01\rangle\otimes(\mathbb{I}\otimes(A_l\otimes A_s)\otimes\mathbb{I})(|\,\mathbb{I}\rangle\rangle\otimes|\,\mathbb{I}\rangle\rangle)+$$
$$+ c_s^2\,|11\rangle \otimes (\mathbb{I} \otimes (A_s \otimes A_s) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle))]$$

$$|X_2\rangle = \frac{1}{2}[c_l^2\mathbb{I}_2\,|00\rangle \otimes \mathcal{V}'(\mathbb{I} \otimes (A_l \otimes A_l) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+ c_l c_s\mathbb{I}\,|10\rangle \otimes \mathcal{V}'(\mathbb{I} \otimes (A_s \otimes A_l) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+ c_l c_s\mathbb{I}\,|01\rangle \otimes \mathcal{V}'(\mathbb{I} \otimes (A_l \otimes A_s) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+ c_s^2\mathbb{I}\,|11\rangle \otimes \mathcal{V}'(\mathbb{I} \otimes (A_s \otimes A_s) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle))]$$

$$|X_2\rangle = \frac{1}{2}[c_l^2\,|00\rangle \otimes (\mathbb{I} \otimes (A_l \circ A_l \otimes P_0) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+ c_l c_s\,|10\rangle \otimes (\mathbb{I} \otimes (A_s \circ A_l \otimes P_0) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+ c_l c_s\,|01\rangle \otimes (\mathbb{I} \otimes (A_l \circ A_s \otimes P_0) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle)+$$
$$+ c_s^2\,|11\rangle \otimes (\mathbb{I} \otimes (A_s \circ A_s \otimes P_0) \otimes \mathbb{I})(|\,\mathbb{I}\rangle\rangle \otimes |\,\mathbb{I}\rangle\rangle))]$$

$$|X_2\rangle = \frac{1}{2}(c_l^2\,|00\rangle \otimes \left|\,(A_l \circ A_l)^T\right\rangle\rangle \otimes |\,P_0\rangle\rangle + c_l c_s\,|10\rangle \otimes \left|\,(A_s \circ A_l)^T\right\rangle\rangle \otimes |\,P_0\rangle\rangle +$$
$$+ c_l c_s\,|01\rangle \otimes \left|\,(A_l \circ A_s)^T\right\rangle\rangle \otimes |\,P_0\rangle\rangle + c_s^2\,|11\rangle \otimes \left|\,(A_s \circ A_s)^T\right\rangle\rangle \otimes |\,P_0\rangle\rangle)$$

■

Proposition 8: Expression of Rényi Entropy of order $a = \frac{1}{2}$ (eq.37)

Recall that the density matrix ρ is equal to $\rho = |A\rangle\rangle \langle\langle A|$ for any bipartite quantum system $|A\rangle\rangle$. Subsequently, the reduced density matrix ρ_1 regarding the first subsystem is:
$\rho_1 = Tr_2(\rho)$
$\rho_1 = Tr_2(|A\rangle\rangle \langle\langle A|)$
$\rho_1 = AA^\dagger$

With the SVD of A being $A = U_A D_A V_A^T$, since A is a real matrix, then by substitution:
$\rho_1 = (U_A D_A V_A^T)(U_A D_A V_A^T)^\dagger$
$\rho_1 = (U_A D_A V_A^T)(U_A D_A V_A^T)^T$ (since A real)
$\rho_1 = U_A D_A V_A^T V_A D_A^T U_A^T$
$\rho_1 = U_A D_A V_A^T$ (since $V_A$real orthogonal and D_A diagonal)

The expression of the reduced density matrix ρ_1 raised to the power of a results in:
$\rho_1^a = (U_A D_A^2 U_A^T)^a$
$\rho_1^a = U_A D_A^{2a} U_A^T$ (since U_A is real orthogonal)

So, the trace of this expression is:
$Tr(\rho_1^a) = Tr(U_A D_A^{2a} U_A^T)$
$Tr(\rho_1^a) = Tr(D_A^{2a})$ (since trace is similarity-invariant)

As a result, the Rényi Entropy of order a is equal to:
$S_R(\rho_1) = \frac{1}{1-a} \log(Tr(\rho_1^a))$
$S_R(\rho_1) = \frac{1}{1-a} \log(Tr(D_A^{2a}))$

So, for the order of the Rényi Entropy $a = \frac{1}{2}$, the formula above changes into:
$S_R(\rho_1) = \frac{1}{1-\frac{1}{2}} \log(Tr(D_A^{2\frac{1}{2}}))$

$$S_R(\rho_1) = 2\log(Tr(D_A)) = 2\log(\sum_{i=1}^{r} \sigma_i)$$

■

Proposition 9: Upper bound $\log_n k$ for Rényi Entropy (eq.38)

Consider the system $|A_k\rangle\rangle$ that corresponds to the rank-k approximation of the n-dimensional system $|A\rangle\rangle$. This means that the singular value of the matrix A_k is:

$$A_k = UD_kV^\dagger = U \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & ... & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & ... & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & \sigma_k & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 \end{bmatrix} V^\dagger$$

with U,V unitary matrices and D_k the diagonal matrix containing its (normalized) singular values.

Since the entanglement contained in a system is inextricably linked to the probability distribution formed by its singular values, it is clear that the system $|A_k\rangle\rangle$ is maximally entangled, provided that its singular values are all equal to $\frac{1}{\sqrt{k}}$. This way, they form a uniform distribution while maintaining the unit norm of the system. In this case, the Rényi Entropy of order $a \neq 1$ is:

$S_R(\rho_1) = \frac{1}{1-a} \log_n(Tr(D_k^{2a}))$

$S_R(\rho_1) = \frac{1}{1-a} \log_n(\sum_{i=1}^{k} \sigma_i^{2a})$

$S_R(\rho_1) = \frac{1}{1-a} \log_n[\sum_{i=1}^{k} (\frac{1}{\sqrt{k}})^{2a}]$

$S_R(\rho_1) = \frac{1}{1-a} \log_n[\sum_{i=1}^{k} (\frac{1}{k})^{a}]$

$S_R(\rho_1) = \frac{1}{1-a} \log_n[k \cdot (\frac{1}{k})^{a}]$

$S_R(\rho_1) = \frac{1}{1-a} \log_n(\frac{k}{k^a})$

$S_R(\rho_1) = \frac{1}{1-a} \log_n(k^{1-a})$

$S_R(\rho_1) = (1-a)\frac{1}{1-a} \log_n k$

$S_R(\rho_1) = \log_n k$

So, the bounds for the entropy of any system with reference to the amount k of its singular values that are taken into consideration is:

$$0 \leq S_R(\rho_1) \leq \log_n k$$

■

References

[1] M. A. Nielsen, I. L. Chuang, *"Quantum Computation and Quantum Information"*, Cambridge University Press, 2000

[2] R. T.Perry, *"The Temple of Quantum Computing"*, Version 1.0, 2004

[3] D. McMahon, *"Quantum Computing Explained"*, published by John Wiley & Sons,Inc., Hoboken, New Jersey, 2008

[4] G. M. D'Ariano, P. Lo Presti, M. F. Sacchi, *"Bell Measurements and Observables"*, Phys. Lett. A 272, 32 (2000)

[5] G. Alber, A. Delgado, N. Gisin and I. Jex, *"Efficient bipartite quantum state purification in arbitrary dimensional Hilbert spaces"*, J. Phys. A: Math. Gen. 34(42), 8821-8833 (2001)

[6] H. Bechmann-Pasquinucci, B. Huttner, N. Gisin, *"Non-linear quantum state transformation of spin-1/2"*, Physics Letters A. 242. 198-204. 10.1016/S0375-9601 00189-3, (1998)

[7] D. Ellinas, Ch. Konstandakis, *"Matrix Algebra for Quantum Search Algorithm: Non Unitary Symmetries and Entanglement"*, in Proceedings of 10th Conference Quantum Communication, Measurement and Computing (QCMC), Brisbane, Australia, 19-23 July 2010, Eds. T. Ralph, P. K. Lam. AIP Conf. Proc. 1363

[8] D. Ellinas, Ch. Konstandakis, *"Parametric Quantum Search Algorithm by CP maps: algebraic, geometric and complexity aspects"*, Journal of Physics A: Mathematical & Theoretical, 46, 415303-41530, (2013)

[9] D. Ellinas, Ch. Konstandakis, *"Faster Together: Collective Quantum Search"*, Entropy, 17, 4838-4862, (2015)

[10] D. Ellinas, Ch. Konstandakis, *"Parametric Quantum Search Algorithm as Quantum Walk: A Quantum Simulation"*, Reports on Mathematical Physics 77 (1), 105-128, (2016)

[11] D. Ellinas, *"Quantum compressing algorithm for image multiqubits states"*, unpublished notes, (2019)

[12] D. Ellinas, Lectures notes on Quantum Information, p.70, (2019)

[13] Horn and Johnson, *"Topics in Matrix Analysis"*, published by the Press Syndicate of the University of Cambridge, (1991)

[14] C. D. Meyer, *"Matrix Analysis and Applied Linear Algebra"*, SIAM, (2000)

[15] E. Million, *"The Hadamard Product"*, 1 Introduction and Basic Results, (2007)

[16] I. Androulidakis, *"Algorithm for Dimensional Reduction of Entangled Bipartite Quantum Systems"*, MSc Thesis, Technical University of Crete, (Supervisor D. Ellinas)

[17] E. Varikou, *"On quantum channels: Special constructions of random and optimally unitary channel maps"*, MSc Thesis, Technical University of Crete, (Supervisor D. Ellinas), (2014)

[18] M. A. Man'ko, V. I. Man'ko and G. Marmo, *"Entropies and correlations in classical and quantum systems"*, Nuovo Cimento della Societa Italiana di Fisica C. 38. 10.1393/ncc/i2015-15167-1, (2015)

[19] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr and M. Tomamichel, *"On quantum Rényi entropies: a new definition and some properties"*, CoRR Journal, abs/1306.3142, (2013)

[20] M. Kovačević, *"Some properties of Rényi Entropy over countably infinite alphabets"*, Problems of Information Transmission Journal 49, 99-110, (2013)

[21] J. D. Cook, (2018, November 21), *"Rényi Entropy"*, Retrieved from https://www.johndcook.com/blog/2018/11/21/renyi-entropy/

[22] Rafael C. Gonzalez, Richard E.Woods, Digital Image Processing, 3rd edition, (2015)

[23] L. Fei-Fei, R. Fergus and P. Perona. *"Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories"*. IEEE. CVPR 2004, Workshop on Generative-Model Based Vision. (2004)