A New Algorithm for Calculating Squashed Entanglement and a Python Implementation Thereof

Robert R. Tucci
tucci@ar-tiste.com

October 10, 2019

Abstract

In this paper, we propose an algorithm for calculating the squashed entanglement of any pure or mixed quantum state. Our algorithm is a generalization of the Arimoto-Blahut algorithm of classical information theory, where it is used to calculate channel capacities. The author of this paper published a previous paper proposing a previous algorithm for calculating squashed entanglement. Even though that previous algorithm was also based on the Arimoto-Blahut algorithm, the algorithm presented in this paper is substantially different and performs much better. This paper is being released at the same time that a Python software library called "Entanglish" is being released at GitHub as open source. Entanglish implements the algorithm for calculating squashed entanglement proposed in this paper and also several other algorithms related to quantum entanglement. In order to test the values of squashed entanglement that our algorithm yields, we introduce a family of pure states of multiple qubits that we call the symmetrized N-up states. We give and prove an analytical formula for the squashed entanglement of this family of states. This paper also proposes a method for doing density matrix perturbation theory that we call "bootstrapped perturbation theory".

1 Introduction

In this paper, we propose an algorithm for calculating the squashed entanglement (Ref.[1]) of any pure or mixed quantum state. Our algorithm is a generalization of the Arimoto-Blahut algorithm (Ref.[2]) of classical information theory, where it is used to calculate channel capacities.

This paper is being released at the same time that a Python library called "Entanglish" is being released at GitHub as open source. Entanglish implements the algorithm for calculating squashed entanglement proposed in this paper and many other algorithms related to quantum entanglement.

In order to test the values of squashed entanglement that our algorithm yields, we introduce a family of pure states of multiple qubits that we call the symmetrized N-up states. We give and prove an analytical formula for the squashed entanglement of this family of states.

One method of executing our algorithm for finding the squashed entanglement of a quantum state ρ requires finding the eigen-decompositions of various Hermitian matrices of the same dimension as the matrix ρ. If the dimension of ρ is small, finding those eigen-decompositions, with acceptable speed and accuracy, is feasible using standard matrix digitalization methods. But for dimension of ρ bigger than a certain value, these standard methods will not work adequately. We present an algorithm that we call "bootstrapped perturbation theory" which can handle better than the standard matrix digitalization methods those ρ's with a large dimension.

For a review of the history of squashed entanglement, see Ref.[1]. The property of strong subadditivity of quantum entropy (i.e., that $S(\underline{x} : \underline{y}|\underline{\alpha}) \geq 0$) which guarantees that squashed entanglement is non-negative, was first proven by Lieb and Ruskai in Ref.[3]. The definition of squashed entanglement was first proposed and some of its properties were first discovered by Tucci in a series of six papers Refs.[4]-[9] during the years 1999 to 2002.[1] One of those six papers, Ref.[7], proposed a method for calculating squashed entanglement that is different to the one proposed in this paper, but also takes inspiration from the Arimoto-Blahut algorithm. The algorithm for calculating squashed entanglement proposed in this paper is substantially better and performs much better than its predecessor. The first paper to study squashed entanglement post Tucci's work appears to be Ref.[10], by Christandl and Winter in 2003. They proved that squashed entanglement satisfies other interesting properties not touched upon in Tucci's work. The fact that the squashed entanglement of ρ is "faithful", meaning that it is zero iff ρ is "separable" was conjectured in the work of Tucci (see Ref.[5]) and was later proven more rigorously in Refs.[11] and [12].

[1] In his original 6 papers, Tucci didn't give any particular name to his entanglement measure. The name "squashed entanglement" was given to it later by Christandl and Winter in Ref.[10]. Some people prefer to call it "CMI entanglement". CMI, pronounced "see-me", stands for Conditional Mutual Information.

2 Notation and Preliminaries

In this section, we will discuss some notation and standard definitions that are used throughout the paper.

Let $\theta(\mathcal{S})$ be the so called truth or indicator function. It equals 1 if statement $\mathcal{S}$ is true and zero otherwise. For example, $\theta(x > 1)$ equals 1 when $x > 1$ and zero otherwise. $\delta(a, b) = \delta_b^a = \theta(a = b)$ will denote the Kronecker delta function.

We will use Z_I to denote the set of integers contained in an interval I, where I is an interval of real numbers whose left side can be either open or closed, and the same for its right side. For example, $Z_{[0,4)} = \{0, 1, 2, 3\}$ and $Z_{[0,4]} = \{0, 1, 2, 3, 4\}$.

For any positive integers n, k with $k \leq n$, let

$$n^{\downarrow k} = n(n-1)(n-2)\dots(n-k+1) \; . \tag{1}$$

For example, $n^{\downarrow 3} = n(n-1)(n-2)$. $n^{\downarrow k}$ has some properties in common with n^k, the k'th power of n. "n choose k", i.e., the number of combinations of n objects taken k at a time, can be expressed using $n^{\downarrow k}$. Indeed, one has

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n^{\downarrow k}}{k!} \; . \tag{2}$$

For any square matrix M, let $\| M \|_2 = \sqrt{\mathrm{tr}(MM^\dagger)}$, where $M^\dagger$ is the Hermitian conjugate of M. This is called the 2-norm or Frobenius norm of M. If all the entries of M are put into a column vector v, this is just the magnitude $\sqrt{v^\dagger v}$ of the vector.

Given any Hermitian matrix ρ, one can express ρ as an eigen-decomposition $\rho = UDU^\dagger$, where U is a unitary matrix and $D = \mathrm{diag}(\lambda_0, \lambda_1, \dots \lambda_{n_{\mathrm{eva}}-1})$ is a diagonal matrix that contains the eigenvalues λ_j of ρ along its diagonal. Let $\pi_1(\rho)$ represent the matrix one obtains from the eigen-decomposition of ρ if one replaces each λ_j in the matrix D by $\lambda_j \theta(|\lambda_j| > \epsilon)$ for some small number $\epsilon > 0$. Let $\pi_0(\rho) = 1 - \pi_1(\rho)$, so in $\pi_0(\rho)$, each λ_j in D is replaced by $\lambda_j \theta(|\lambda_j| \leq \epsilon)$. $\pi_1(\rho)$ (resp., $\pi_0(\rho)$) are the projectors onto the vector space of non-zero (resp., zero) eigenvalues of ρ. One can define the so called "Penrose pseudo-inverse" of ρ as:

$$\mathrm{pinv}(\rho) = \pi_1(\rho)\rho^{-1}\pi_1(\rho) \; . \tag{3}$$

What this means is that, to arrive at $\mathrm{pinv}(\rho)$, one replaces in the eigen-decomposition of ρ, each eigenvalue λ_j in D by $1/\lambda_j$ if $|\lambda_j| > \epsilon$ and by 0 otherwise.

In this paper, random variables will be represented by underlined letters instead of by the more common convention of using capital letters to represent them. For example $P_{\underline{x},\underline{y}}(x, y) = P(\underline{x} = x, \underline{y} = y) = P(x, y)$ will denote a joint distribution of the random variables $\underline{x}$ and $\underline{y}$.

Let $\mathcal{H}_{\underline{x}}$ denote a Hilbert (or vector) space with an orthonormal basis $\{|x\rangle | x = 0, 1, 2, \dots, n_{\underline{x}} - 1\}$. We will refer to $n_{\underline{x}}$ as the size (or number of possible values, or number of possible states) of $\underline{x}$. Let $\mathcal{L}(\mathcal{H}_{\underline{x}})$ denote the set of all linear operators

acting on $\mathcal{H}_{\underline{x}}$. Let $\mathcal{H}_{\underline{x},\underline{y}}$ denote the tensor product of $\mathcal{H}_{\underline{x}}$ and $\mathcal{H}_{\underline{y}}$. If $\rho_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{x},\underline{y}})$ is a density matrix, then we define its partial traces as $\rho_{\underline{x}} = \mathrm{tr}_{\underline{y}}\rho_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{x}})$ and $\rho_{\underline{y}} = \mathrm{tr}_{\underline{x}}\rho_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{y}})$.

Next, we review a few standard definitions used in classical and quantum information theory.

In classical information theory, one defines the entropy of a classical probability distribution $P_{\underline{x}}(x)$ by

$$H(\underline{x}) = -\sum_x P(x)\log P(x)\ , \tag{4}$$

where the sum is over all values that $\underline{x}$ can assume. Then one defines the conditional information (CI) by

$$\begin{aligned} H(\underline{x}|\underline{y}) &= \sum_{x,y} P(x,y)\log\frac{1}{P(x|y)} &(5)\\ &= H(\underline{x},\underline{y}) - H(\underline{y})\ , &(6)\end{aligned}$$

and the conditional mutual information (CMI) by

$$\begin{aligned} H(\underline{x}:\underline{y}|\underline{\alpha}) &= \sum_{x,y,\alpha} P(x,y,\alpha)\log\frac{P(x,y,|\alpha)}{P(x|\alpha)P(y|\alpha)} &(7)\\ &= -H(\underline{x},\underline{y}|\underline{\alpha}) + H(\underline{x}|\underline{\alpha}) + H(\underline{y}|\underline{\alpha})\ . &(8)\end{aligned}$$

(The mutual information (MI) is the CMI without the $\underline{\alpha}$).

In quantum information theory, one defines the von Neumann entropy of a density matrix $\rho_{\underline{x}}$ by

$$S(\underline{x}) = S(\rho_{\underline{x}}) = -\mathrm{tr}_{\underline{x}}[\rho_{\underline{x}}\log\rho_{\underline{x}}]\ , \tag{9}$$

where the rows (and columns) of $\rho_{\underline{x}}$ are labelled by the values that $\underline{x}$ can assume. Then one defines the conditional information (CI) by

$$S(\underline{x}|\underline{y}) = S(\underline{x},\underline{y}) - S(\underline{y})\ , \tag{10}$$

and the conditional mutual information (CMI) by

$$S(\underline{x}:\underline{y}|\underline{\alpha}) = -S(\underline{x},\underline{y}|\underline{\alpha}) + S(\underline{x}|\underline{\alpha}) + S(\underline{y}|\underline{\alpha})\ . \tag{11}$$

(The mutual information (MI) is the CMI without the $\underline{\alpha}$).

A maximally entangled pure state with two parts $\underline{x}$ and $\underline{y}$ is defined as a pure state whose density matrix $\rho_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{x},\underline{y}})$ satisfies

$$S(\rho_{\underline{x}}) = S(\rho_{\underline{y}}) = \log n_{min}\ , \tag{12}$$

where $n_{min} = \min(n_{\underline{x}}, n_{\underline{y}})$ and $n_{\underline{x}}, n_{\underline{y}}$ are the sizes of spaces $\mathcal{H}_{\underline{x}}, \mathcal{H}_{\underline{y}}$, respectively. Let

$$|\psi^{\max-\text{entan}}_{\underline{x},\underline{y}}\rangle = \sum_{n=0}^{n_{min}} |n\rangle_{\underline{x}} |n\rangle_{\underline{y}} \, , \tag{13}$$

and

$$\rho^{\max-\text{entan}}_{\underline{x},\underline{y}} = |\psi^{\max-\text{entan}}_{\underline{x},\underline{y}}\rangle\langle\psi^{\max-\text{entan}}_{\underline{x},\underline{y}}| \, . \tag{14}$$

Then it is easy to show that state $\rho^{\max-\text{entan}}_{\underline{x},\underline{y}}$ is a maximally entangled state with parts $\underline{x}$ and $\underline{y}$.

3 Generalization of Arimoto Blahut Algorithm

In this section, we will propose an algorithm for calculating the squashed entanglement of any pure or mixed state.

Consider a bipartite system consisting of two parts labelled by the random variables $\underline{x}$ and $\underline{y}$, and described by a density matrix $\rho_{\underline{x},\underline{y}}$. The squashed entanglement of such a system is defined as

$$E_{\underline{x},\underline{y}}(\rho_{\underline{x},\underline{y}}) = \frac{1}{2} \min S(\underline{x} : \underline{y}|\underline{\alpha}) \, . \tag{15}$$

The min()—or infimum() if one wishes to be more mathematically precise[2]–is over all density matrices $\rho_{\underline{x},\underline{y},\underline{\alpha}}$ such that $\text{tr}_{\underline{\alpha}} \, \rho_{\underline{x},\underline{y},\underline{\alpha}} = \rho_{\underline{x},\underline{y}}$ with $\rho_{\underline{x},\underline{y}}$ held fixed. This definition of $E_{\underline{x},\underline{y}}$ has many desirable properties. For example,

- $E_{\underline{x},\underline{y}} = E_{\underline{y},\underline{x}} \geq 0$. This is true because of the so called "strong subadditivity of quantum entropy", which states that $S(\underline{x} : \underline{y}|\underline{\alpha}) \geq 0$ for all density matrices $\rho_{\underline{x},\underline{y},\underline{z}}$. See Ref.[3] for the first ever proof of $S(\underline{x} : \underline{y}|\underline{\alpha}) \geq 0$. The classical analogue $H(\underline{x} : \underline{y}|\underline{\alpha}) \geq 0$ was known to be true long before Ref.[3].

- If $\rho_{\underline{x},\underline{y}}$ is pure (which is true iff $\text{tr}(\rho^2_{\underline{x},\underline{y}} - \rho_{\underline{x},\underline{y}}) = 0$), then, as Tucci pointed out in his very first paper on squashed entanglement (this was one of his original motivations for defining squashed entanglement)[3]

$$E_{\underline{x},\underline{y}} = S(\underline{x}) = S(\underline{y}) \, . \tag{16}$$

[2]The author of this paper is a physicist. He is often lax in Bourbaki style rigor, as most physicists are wont to be, with good reason. Physicists often skip excessive rigor, because it's not the goal of physics, as it is of mathematics, where it is essential.

[3]The factor of 1/2 in the definition of squashed entanglement is necessary to make this true. Contrary to what Ref.[12] claims, this factor of 1/2 appears in all of Tucci's papers about squashed entanglement, from the very earliest.

- $E_{\underline{y},\underline{x}}(\rho_{\underline{x},\underline{y}}) = 0$ iff $\rho_{\underline{x},\underline{y}}$ is conditionally separable, by which we mean that $\rho_{\underline{x},\underline{y}}$ can be expressed as[4]

$$\rho_{\underline{x},\underline{y}} = \sum_{\alpha} w^{\alpha} \rho^{\alpha}_{\underline{x}} \rho^{\alpha}_{\underline{y}} \ , \tag{17}$$

  where w^{α} for all α is a probability distribution, and $\rho^{\alpha}_{\underline{x}}, \rho^{\alpha}_{\underline{y}}$ are density matrices. This property is enunciated by some by saying that "squashed entanglement is faithful". In particular, this implies in the classical case when $\rho_{\underline{x},\underline{y}}$ is diagonal with diagonal $P(x,y)$, that squashed entanglement is zero iff $P(x,y)$ is expressible as

$$P(x,y) = \sum_{\alpha} P(\alpha) P(x|\alpha) P(y|\alpha) \ . \tag{18}$$

  This last equation expresses conditional separabilty in the classical case and is equivalent to the classical Bayesian network $\underline{x} \leftarrow \underline{\alpha} \rightarrow \underline{y}$.[5]

  We will assume that $\rho_{\underline{x},\underline{y},\underline{\alpha}}$ is of the form

$$\rho_{\underline{x},\underline{y},\underline{\alpha}} = \sum_{\alpha} w^{\alpha} |\alpha\rangle\langle\alpha| \rho^{\alpha}_{\underline{x},\underline{y}} \ , \tag{19}$$

where w_{α} for $\alpha \in Z_{[0,n_{\alpha})}$ is a probability distribution , where $\rho^{\alpha}_{\underline{x},\underline{y}}$ for each α is a density matrix, and where the states $|\alpha\rangle$ for all α are orthonormal. Therefore,

$$\rho_{\underline{x},\underline{y}} = \mathrm{tr}_{\underline{\alpha}} \ \rho_{\underline{x},\underline{y},\underline{\alpha}} = \sum_{\alpha} w^{\alpha} \rho^{\alpha}_{\underline{x},\underline{y}} \ . \tag{20}$$

  Define

$$E^{\alpha}_{\underline{x},\underline{y}} = \frac{1}{2} \mathrm{tr}_{\underline{x},\underline{y}} \left\{ \rho^{\alpha}_{\underline{x},\underline{y}} [\log \rho^{\alpha}_{\underline{x},\underline{y}} - \log \rho^{\alpha}_{\underline{x}} \rho^{\alpha}_{\underline{y}}] \right\} \tag{21}$$

so that the squashed entanglement can be expressed as

$$E_{\underline{x},\underline{y}} = \min \sum_{\alpha} w^{\alpha} E^{\alpha}_{\underline{x},\underline{y}} \ . \tag{22}$$

  It is also convenient to define

$$K^{\alpha}_{\underline{x},\underline{y}} = \rho^{\alpha}_{\underline{x},\underline{y}} w^{\alpha} \ . \tag{23}$$

[4]It is more conventional in the literature to call property Eq.(17) simply "separability" instead of conditional separability, but I prefer to use the word "separable" to refer to case that $\rho_{\underline{x},\underline{y}} = \rho_{\underline{x}} \rho_{\underline{y}}$.

[5]This fact, which can be enunciated by saying that squashed entanglement is faithful in the classical case, was first pointed out in Tucci's original 6 papers on squashed entanglement and was one of his original motivations for defining squashed entanglement.

Note that

$$w^\alpha = \mathrm{tr}_{\underline{x},\underline{y}} K^\alpha_{\underline{x},\underline{y}}, \quad \rho^\alpha_{\underline{x},\underline{y}} = \frac{K^\alpha_{\underline{x},\underline{y}}}{\mathrm{tr}_{\underline{x},\underline{y}} K^\alpha_{\underline{x},\underline{y}}} \ . \tag{24}$$

Hence, one can always translate back and forth between the pair $(w^\alpha, \rho^\alpha_{\underline{x},\underline{y}})$ and $K^\alpha_{\underline{x},\underline{y}}$ so one can think of these two constructs as 2 equivalent representations of the same thing.

Suppose $M^\alpha_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{x},\underline{y}})$ for all $\alpha \in Z_{[0,n_{\underline{\alpha}})}$. Define the function regulate() by

$$\mathrm{regulate}(M^\alpha_{\underline{x},\underline{y}}) = \frac{[M^\alpha_{\underline{x},\underline{y}}]_{\mathrm{pos.part}}}{\sum_\alpha \mathrm{tr}_{\underline{x},\underline{y}}(numerator)} \ . \tag{25}$$

Here $[M]_{\mathrm{pos.part}}$ is the positive part of the operator M, which is obtained by replacing all negative eigenvalues of M by zero and leaving non-negative ones the same. Note that the regulate() function maps $K^\alpha_{\underline{x},\underline{y}}$ to itself:

$$\mathrm{regulate}(K^\alpha_{\underline{x},\underline{y}}) = K^\alpha_{\underline{x},\underline{y}} \ . \tag{26}$$

Define the Lagrangian

$$\mathcal{L} = \sum_\alpha w^\alpha E^\alpha_{\underline{x},\underline{y}} + \mu(\sum_\alpha w^\alpha - 1) + \mathrm{tr}_{\underline{x},\underline{y}} \left[\frac{\lambda'_{\underline{x},\underline{y}}}{2} (\sum_\alpha w^\alpha \rho^\alpha_{\underline{x},\underline{y}} - \rho_{\underline{x},\underline{y}}) \right] \ . \tag{27}$$

To calculate $E_{\underline{x},\underline{y}}$, we will set to zero the variation of $\mathcal{L}$ when we vary $w^\alpha \in \mathbb{R}$, $\rho^\alpha_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{x},\underline{y}})$ and the Lagrange multipliers $\mu \in \mathbb{R}$, $\lambda'_{\underline{x},\underline{y}} \in \mathcal{L}(\mathcal{H}_{\underline{x},\underline{y}})$.

To calculate the optimum $\rho^\alpha_{\underline{x},\underline{y}}$, note that

$$\delta\mathcal{L} = \left\{ \begin{array}{l} \sum_\alpha \frac{w^\alpha}{2} \mathrm{tr}_{\underline{x},\underline{y}} \left[\delta\rho^\alpha_{\underline{x},\underline{y}} (\log \rho^\alpha_{\underline{x},\underline{y}} + 1 - \log \rho^\alpha_{\underline{x}} \rho^\alpha_{\underline{y}} + \lambda'_{\underline{x},\underline{y}}) \right] \\ + \sum_\alpha \delta w^\alpha \left[E^\alpha_{\underline{x},\underline{y}} + \mathrm{tr}_{\underline{x},\underline{y}} \frac{\lambda'_{\underline{x},\underline{y}}}{2} \rho^\alpha_{\underline{x},\underline{y}} + \mu \right] \end{array} \right. \ . \tag{28}$$

The two independent variations with respect to $\rho^\alpha_{\underline{x},\underline{y}}$ and w^α are zero iff

$$\left\{ \begin{array}{l} \forall\alpha, \lambda_{\underline{x},\underline{y}} = -\lambda'_{\underline{x},\underline{y}} - 1 = \log \rho^\alpha_{\underline{x},\underline{y}} - \log \rho^\alpha_{\underline{x}} \rho^\alpha_{\underline{y}} \\ \mu = \frac{1}{2} \end{array} \right. \ . \tag{29}$$

The above results suggest the following recursion to go from $K^\alpha_{\underline{x},\underline{y}}(k)$ to $K^\alpha_{\underline{x},\underline{y}}(k+1)$ for $k = 0, 1, 2, \ldots$.

For $k = 0$, define[6]

[6]Defining $K^\alpha_{\underline{x},\underline{y}}(0) = \frac{1}{n_{\underline{\alpha}}} \rho_{\underline{x},\underline{y}}$ for all α, does not work because this is a stationary point of the recursion. Try it.

$$K^{\alpha}_{\underline{x},\underline{y}}(0) = \begin{cases} \text{regulate}\left[\rho_{\underline{x},\underline{y}} + (-1 + \frac{1}{n_{\underline{\alpha}}})\rho^{\text{diag.part}}_{\underline{x},\underline{y}}\right] & \text{if } \alpha = 0 \\ \text{regulate}\left[\frac{1}{n_{\underline{\alpha}}}\rho^{\text{diag.part}}_{\underline{x},\underline{y}}\right] & \text{otherwise} \end{cases} \tag{30}$$

where $\rho^{\text{diag.part}}_{\underline{x},\underline{y}}$ is obtained from $\rho_{\underline{x},\underline{y}}$ by replacing its non-diagonal elements by zero.

For $k = 0, 1, 2, ..$, given $K^{\alpha}_{\underline{x},\underline{y}}(k)$ for all α, choose $K^{\alpha}_{\underline{x},\underline{y}}(k+1)$ for all α as follows. Let $\alpha_{rand}(k)$ be a value of $\underline{\alpha}$ chosen uniformly at random from all possible values of $\underline{\alpha}$. Set

$$\lambda_{\underline{x},\underline{y}}(k) = \log \rho^{\alpha_{rand}(k)}_{\underline{x},\underline{y}}(k) - \log \rho^{\alpha_{rand}(k)}_{\underline{x}}(k)\rho^{\alpha_{rand}(k)}_{\underline{y}}(k) \ , \tag{31}$$

$$M^{\alpha}_{\underline{x},\underline{y}}(k) = \exp\left[\lambda_{\underline{x},\underline{y}}(k) + \log \rho^{\alpha}_{\underline{x}}(k)\rho^{\alpha}_{\underline{y}}(k)\right] w^{\alpha} \ , \tag{32}$$

$$K^{\alpha}_{\underline{x},\underline{y}}(k+1) = \text{regulate}\left(\frac{1}{2}\rho_{\underline{x},\underline{y}}[\sum_{\alpha} M^{\alpha}_{\underline{x},\underline{y}}(k)]^{-1} M^{\alpha}_{\underline{x},\underline{y}}(k) + h.c.\right) \ , \tag{33}$$

$$E_{\underline{x},\underline{y}}(k) = \frac{1}{2}\text{tr}_{\underline{x},\underline{y}}\left[\rho_{\underline{x},\underline{y}}\lambda_{\underline{x},\underline{y}}(k)\right] \ . \tag{34}$$

We claim but will not prove rigorously that

$$E_{\underline{x},\underline{y}}(k) \to E_{\underline{x},\underline{y}} \tag{35}$$

as $k \to \infty$.

4 Bootstrapped 2nd Order Perturbation Theory

In this section, we propose a method for calculating perturbatively any function $f(\rho)$ of ρ expressible as a power series of ρ, for any quantum density matrix ρ. We are especially interested in using this method to evaluate any function $f(\cdot)$ that arises in the calculation of squashed entanglement for pure or mixed states.

First, let us review the standard theory, as used in Quantum Mechanics, of (stationary, time independent) perturbation theory for a Hamiltonian H. We will denote the eigen-system of H by

$$\text{esys}(H) = \{(\lambda_j, v_j)\}_{j=0,1,...,n_{\text{eva}}-1} \ , \tag{36}$$

where λ_j are the eigenvalues of H and the column vectors v_j are eigenvectors satisfying $Hv_j = \lambda_j v_j$. One can define a unitary matrix U by stacking side by side the column vectors v_j, and a diagonal matrix D by placing the eigenvalues of H along the diagonal of D. Then

$$U = [v_0, v_1, \ldots, v_{n_{\text{eva}}-1}], \quad D = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{n_{\text{eva}}-1}) \tag{37}$$

and

$$H = \sum_{j=0,1,\ldots,n_{\rm eva}-1} \lambda_j v_j v_j^\dagger = U D U^\dagger \ . \tag{38}$$

Suppose

$$H = H_0 + V \ , \tag{39}$$

where H_0 is a Hermitian matrix for which esys(H_0) is known. H_0 is called the unperturbed Hamiltonian and V is called the perturbation. Perturbation Theory is a method for finding an approximation, call it esys$^{\approx}(H)$, for the exact eigen-system esys(H) of H. Symbolically, perturbation theory takes us from esys(H_0) to esys$^{\approx}(H)$:

$$\text{esys}(H_0) \to \text{esys}^{\approx}(H) \ . \tag{40}$$

The terms of the approximation esys$^{\approx}(H)$ are grouped by their order (power) of V. Ref[13] is quite useful because it lists all terms up to fifth order! For completeness, we list them here next, up to 3rd order for the eigenvalues and up to 2nd order for the eigenvectors. To switch from our notation to the notation of Ref.[13], let $\lambda_j \to E_j$, $v_j \to |j\rangle$, $v_j^\dagger \to \langle j|$. Also let

$$V_{nm} = \langle n^{(0)}|V|m^{(0)}\rangle \ , \tag{41}$$

$$E_{nm} = E_n^{(0)} - E_m^{(0)} \ , \tag{42}$$

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \ldots \tag{43}$$

$$|n\rangle = |n^{(0)}\rangle + |n^{(1)}\rangle + |n^{(2)}\rangle + \ldots \ . \tag{44}$$

The eigenvalue approximations up to 3rd order are:

$$E_n^{(1)} = V_{nn} \tag{45}$$

$$E_n^{(2)} = \sum_{k_2 \neq n} \frac{|V_{nk_2}|^2}{E_{nk_2}} \tag{46}$$

$$E_n^{(3)} = \sum_{k_2 \neq n, k_3 \neq n} \frac{V_{nk_3} V_{k_3 k_2} V_{k_2 n}}{E_{nk_2} E_{nk_3}} - \sum_{k_3 \neq n} V_{nn} \frac{|V_{nk_3}|^2}{E_{nk_3}^2} \ . \tag{47}$$

The eigenvector approximations up to 2nd order are:

$$|n^{(1)}\rangle = \sum_{k_1 \neq n} \frac{V_{k_1 n}}{E_{nk_1}} |k_1^{(0)}\rangle \tag{48}$$

$$|n^{(2)}\rangle = \sum_{k_1 \neq n, k_2 \neq n} \frac{V_{k_1 k_2} V_{k_2 n}}{E_{nk_1} E_{nk_2}} |k_1^{(0)}\rangle - \sum_{k_1 \neq n} \left(\frac{V_{nn} V_{k_1 n}}{E_{nk_1}^2} |k_1^{(0)}\rangle + \frac{V_{nk_1} V_{k_1 n}}{2 E_{nk_1}^2} |n^{(0)}\rangle \right) \ . \tag{49}$$

The higher order terms in these perturbative approximations to the eigenvalues and eigenvectors diverge when some eigenvalues are degenerate, i.e., when at least one eigenvalue has an eigenspace that is of bigger dimension than 1. For those cases, one does what is called "degenerate perturbation theory". One does a similarity transformation (change of bases) in each degenerate eigenspace so as to diagonalize V within that eigenspace. After that change of basis, for any n, m with $n \neq m$, if $E_{nm} = 0$ then $V_{nm} = 0$ too. So now instead of getting expressions that have a finite numerator and a zero denominator, the only problematic expressions that we get are of the type $0/0$, which we set to zero.

Next we try to answer the question, how can we apply this machinery of perturbation theory for Hamiltonians H to do perturbation theory of density matrices ρ. Both H and ρ are Hermitian matrices, which insures that their eigen systems exist and the eigenvalues are real numbers. In the case of ρ, the eigenvalues are also guaranteed to be ≥ 0. Let us begin by expressing ρ as a sum of an unperturbed density matrix ρ_0 and a perturbation $\delta\rho$.

$$\rho = \rho_0 + \delta\rho \ . \tag{50}$$

Then the machinery of perturbation theory will allow us to go from $\text{esys}(\rho_0)$ to $\text{esys}^{\approx}(\rho)$.

$$\text{esys}(\rho_0) \to \text{esys}^{\approx}(\rho) \ . \tag{51}$$

It remains for us to define what is ρ_0. We want a ρ_0 that is easy to calculate from ρ, and such that $\text{esys}(\rho_0)$ is also easy to calculate, much easier than calculating $\text{esys}(\rho)$. Next we define a ρ_0 that satisfies this desiderata, and which is also a particularly natural choice for studying entanglement.

Before defining ρ_0, let us define certain terms describing the shape and partial traces of ρ. For this, we will use a vocabulary very close to the one used in the documentation of the Python library numpy.

We will use a fixed ρ in two different shapes: either as an $n_{\text{rows}} \times n_{\text{rows}}$ matrix, or as a tensor with $2n_{\text{ra}}$ axes (axes are the same as tensor indices). So

$$\rho \in \mathbb{C}^{n_{\text{rows}} \times n_{\text{rows}}} \simeq \mathbb{C}^{[n_0 \times n_1 \times \cdots \times (n_{\text{ra}}-1)]^2} \tag{52}$$

where

- row shape $= (n_0, n_1, \ldots, n_{\text{ra}} - 1)$
- n_{rows} = number of rows = product(row shape)
- n_{ra} = number of row axes = len(row shape)

In the language of quantum information, one can call say that ρ describes n_{ra} qudits where $d = n_j$ for the j'th qudit.

Consider a ρ with $n_{\rm ra} = 4$ for definiteness. Such a ρ can be represented by all of the following equivalent notations:

$$\rho = \rho(0,1,2,3) = \rho_{\underline{x}_0,\underline{x}_1,\underline{x}_2,\underline{x}_3} = \rho^{j_0,j_1,j_2,j_3;j'_0,j'_1,j'_2,j'_3} = \rho^{J;K} \tag{53}$$

Partial traces of a density matrix are fundamental to a rigorous definition of entanglement. They are analogous to finding the marginals of a joint distribution in classical probability. As an example of how we define and express them in this paper, suppose we want to take the trace of $\rho(0,1,2,3)$ with respect to the axes $1,3$. Using the notation of this paper, we might write:

$$\begin{aligned} \rho(0,2) &= \rho_{\underline{x}_0,\underline{x}_2} && (54)\\ &= {\rm tr}_{\underline{x}_1,\underline{x}_3}\rho_{\underline{x}_0,\underline{x}_1,\underline{x}_2,\underline{x}_3} && (55)\\ &= \sum_{j_1,j'_1}\sum_{j_3,j'_3}\delta(j_1,j'_1)\delta(j_3,j'_3)\rho^{j_0,j_1,j_2,j_3;j'_0,j'_1,j'_2,j'_3}\ . && (56) \end{aligned}$$

Suppose that

$$\text{All Row Axes} = A = Z_{[0,n_{\rm ra})}\ . \tag{57}$$

Hence

$$\rho = \rho(A)\ . \tag{58}$$

We define

$$\rho(j) = {\rm tr}_{A-\{j\}}\ \ \rho \tag{59}$$

and

$$\rho_0 = \rho(0)\otimes\rho(1)\otimes\ldots\otimes\rho(n_{\rm ra}-1)\ . \tag{60}$$

We will refer to the density matrix ρ_0 as the separable approximation of ρ.

To find esys(ρ_0), we only need to find the esys$(\rho(j))$ for each qudit j. In fact, suppose

$$\rho(j) = U(j)D(j)U^\dagger(j)\ , \tag{61}$$

for $j \in Z_{[0,n_{\rm ra})}$ where $U(j)$ is a unitary matrix and $D(j)$ is a diagonal matrix. Then it follows that

$$\rho_0 = UDU^\dagger\ , \tag{62}$$

where

$$\begin{aligned} U &= U(0)\otimes U(1)\otimes\ldots\otimes U(n_{\rm ra}-1) && (63)\\ D &= D(0)\otimes D(1)\otimes\ldots\otimes D(n_{\rm ra}-1)\ . && (64) \end{aligned}$$

In the case that the row shape of ρ is $(2, 2, \ldots 2)$ with $n_{\rm ra}$ components, i.e, if all the qudits are qubits, each $\rho(j)$ is a 2-dim matrix which can be diagonalized trivially. So, as we stipulated in our desiderata, finding esys(ρ_0) is much easier than finding esys(ρ), especially if the dimension $n_{\rm rows}$ of ρ is large.

With this definition of ρ_0, there is no reason that $\delta\rho = \rho - \rho_0$ must satisfy $\| \delta\rho \|_2 << 1$. But we would like that inequality to be true to insure that the perturbation theory converges quickly. If we define

$$\delta'\rho = \frac{\delta\rho}{n_{\rm bs}}, \tag{65}$$

for some integer $n_{\rm bs}$, then, for large enough $n_{\rm bs}$, $\| \delta'\rho \|_2 << 1$ will be true. This suggests what we will refer to as "bootstrap" perturbation theory. The idea is to define a chain of density matrices, where adjacent density matrices in the chain are a distance $\delta'\rho << 1$ apart, even if the beginning and end of the chain are much farther apart. Let

$$\rho_j = \rho_0 + j\delta'\rho \quad \text{for} \;\; j = 0, 1, \ldots, n_{\rm bs} - 1 \; . \tag{66}$$

Then one can apply 2nd order perturbation theory to bootstrap ourselves from ρ_0 to ρ by advancing a small distance $\delta'\rho$ many times until we cover the full distance $\delta\rho$. Bootstrap perturbation can be represented symbolically by

$$\text{esys}(\rho_0) \to \text{esys}^{\approx}(\rho_1) \to \cdots \to \text{esys}^{\approx}(\rho_{n_{\rm bs}-1}) = \text{esys}^{\approx}(\rho) \; , \tag{67}$$

where approximations are indicated by a $\approx$ superscript.

Consider any function $f(\rho)$ of ρ expressible as a power series of ρ. $f(\rho)$ can be many things. For example, $\log(\rho)$, $\exp(\rho)$, $\sqrt{\rho}$, $\text{pinv}(\rho)$ $\rho \log \rho$, etc. One has

$$\begin{aligned} f(\rho) &= f(UDU^\dagger) \approx f(U^{\approx} D^{\approx} U^{\approx\dagger}) & (68) \\ &\approx U^{\approx} f(D^{\approx}) U^{\approx\dagger} \, , & (69) \end{aligned}$$

where, as before, $\rho = UDU^\dagger$, U is unitary, D is diagonal, and approximations are indicated by a $\approx$ superscript.

In conclusion, bootstrap perturbation allows us to calculate esys(ρ_0) easily, to bootstrap esys(ρ_0) to esys$^{\approx}(\rho)$, and to approximate $f(\rho)$ from that esys$^{\approx}(\rho)$. Finding such an approximation of $f(\rho)$ was the goal that we set out for ourselves at the beginning of this section.

A Appendix: Symmetrized N-up States

In order to test the numerical algorithms presented in this paper for calculating squashed entanglement, it is very useful to have a family of quantum states for which

the squashed entanglement is known analytically. In this appendix, we describe one such family of quantum states, what we call the symmetrized N-up states.

The symmetrized N-up states are pure states composed of multiple qubits (so no qudits with $d \neq 2$ in them). They are simply obtained by starting with a state that is a tensor product of n_1 states $|1\rangle$ and n_0 states $|0\rangle$ and applying a full symmetrization operator to that.

Let Sym(n) be the group of all permutations of n objects (often called the Symmetric Group on n letters). Then we define the full symmetrization operator S_n by

$$S_n = \frac{1}{n!} \sum_{\sigma \in \mathrm{Sym(n)}} \sigma \ . \tag{70}$$

For example, for $n = 3$, we have

$$S_3 = \frac{1}{3!}[1 + \sigma_{0,1} + \sigma_{1,2} + \sigma_{0,2} + \sigma_{0,1}\sigma_{1,2} + \sigma_{1,2}\sigma_{0,1}] \ , \tag{71}$$

where $\sigma_{\alpha,\beta}$ is the operator that swaps qubits α and β. A nice discussion of the properties and use of S_n in Group Theory can be found in Ref.[14] by Cvitanovic. Note that S_n is "idempotent" (it equals its square):

$$\begin{aligned} S_n^2 &= \left(\frac{1}{n!}\right)^2 \sum_{\sigma_1} \sigma_1 \sum_{\sigma_2} \sigma_2 & (72) \\ &= \frac{1}{n!} \sum_{\sigma} \sigma & (73) \\ &= S_n \ . & (74) \end{aligned}$$

Let

$$n = n_0 + n_1 \ , \tag{75}$$

where $n_0, n_1 \in Z_{[0,\infty)}$. Let

$$\begin{matrix} |0^{n_0}\rangle \\ |1^{n_1}\rangle \end{matrix} = \underbrace{|0\rangle \otimes |0\rangle \ldots |0\rangle}_{n_0 \text{ factors}} \otimes \underbrace{|1\rangle \otimes |1\rangle \ldots |1\rangle}_{n_1 \text{ factors}} \ . \tag{76}$$

We define the "symmetrized n_1-up state" as the following n qubit state:

$$|\psi_{n_1}^n\rangle = \sqrt{\binom{n}{n_1}} S_n \begin{matrix} |0^{n_0}\rangle \\ |1^{n_1}\rangle \end{matrix} \ . \tag{77}$$

The normalization of these states has been chosen so that they satisfy the following condition:

Claim 1

$$\langle \psi^n_{n_1} | \psi^n_{n'_1} \rangle = \delta(n_1, n'_1) \; . \tag{78}$$

proof:

$$\begin{aligned} \langle \psi^n_{n_1} | \psi^n_{n_1} \rangle &= \binom{n}{n_1} \begin{matrix} \langle 0^{n_0}| \\ \langle 1^{n_1}| \end{matrix} S_n \begin{matrix} |0^{n_0}\rangle \\ |1^{n_1}\rangle \end{matrix} & (79) \\ &= 1 \; . & (80) \end{aligned}$$

QED

Claim 2

$$\begin{matrix} \langle 0^{n_0}| \\ \langle 1^{n_1}| \end{matrix} |\psi^n_{n_1}\rangle = \begin{matrix} \langle 1^{n_1}| \\ \langle 0^{n_0}| \end{matrix} |\psi^n_{n_1}\rangle = \frac{1}{\sqrt{\binom{n}{n_1}}} \; . \tag{81}$$

proof:

$$\begin{aligned} \begin{matrix} \langle 0^{n_0}| \\ \langle 1^{n_1}| \end{matrix} |\psi^n_{n_1}\rangle &= \sqrt{\binom{n}{n_1}} \begin{matrix} \langle 0^{n_0}| \\ \langle 1^{n_1}| \end{matrix} S_n \begin{matrix} |0^{n_0}\rangle \\ |1^{n_1}\rangle \end{matrix} & (82) \\ &= \frac{1}{\sqrt{\binom{n}{n_1}}} \; . & (83) \end{aligned}$$

QED

Claim 3

$$|\psi^n_{n_1}\rangle = \sqrt{\frac{n_0}{n}} \begin{matrix} |0\rangle \\ |\psi^{n-1}_{n_1}\rangle \end{matrix} + \sqrt{\frac{n_1}{n}} \begin{matrix} |1\rangle \\ |\psi^{n-1}_{n_1-1}\rangle \end{matrix} \; . \tag{84}$$

proof:

By symmetry, there have to exist $a, b \in \mathbb{C}$ so that

$$|\psi^n_{n_1}\rangle = a \begin{matrix} |0\rangle \\ |\psi^{n-1}_{n_1}\rangle \end{matrix} + b \begin{matrix} |1\rangle \\ |\psi^{n-1}_{n_1-1}\rangle \end{matrix} \; . \tag{85}$$

To find a and b, one notes that

$$\begin{matrix} \langle 0^{n_0}| \\ \langle 1^{n_1}| \end{matrix} |\psi^n_{n_1}\rangle = a \begin{matrix} \langle 0^{n_0-1}| \\ \langle 1^{n_1}| \end{matrix} |\psi^{n-1}_{n_1}\rangle \; , \tag{86}$$

so

$$a = \sqrt{\frac{\binom{n-1}{n_1}}{\binom{n}{n_1}}} = \sqrt{\frac{n_0}{n}} \ . \tag{87}$$

Likewise, one notes that

$$\begin{matrix} \langle 1^{n_1}| \\ \langle 0^{n_0}| \end{matrix} \ |\psi^n_{n_1}\rangle = b \ \begin{matrix} \langle 1^{n_1-1}| \\ \langle 0^{n_0}| \end{matrix} \ |\psi^{n-1}_{n_1-1}\rangle \ , \tag{88}$$

so

$$b = \sqrt{\frac{\binom{n-1}{n_1-1}}{\binom{n}{n_1}}} = \sqrt{\frac{n_1}{n}} \ . \tag{89}$$

QED

Let Γ be a subset of $Z_{[0,n)}$ and let Γ^c be $Z_{[0,n)} - \Gamma$, i.e., the complement of Γ in $Z_{[0,n)}$. We will denote the number of elements (cardinality) of these two sets by $\gamma = |\Gamma|$, $\gamma^c = |\Gamma^c|$. Let

$$\rho^n_{n_1} = \rho^n_{n_1}(Z_{[0,n)}) = |\psi^n_{n_1}\rangle\langle\psi^n_{n_1}| \ . \tag{90}$$

We define $\rho^n_{n_1}(\Gamma)$ by

$$\rho^n_{n_1}(\Gamma) = \mathrm{tr}_{\Gamma^c}\rho^n_{n_1} \ . \tag{91}$$

The squashed entanglement is defined for this case as simply the von Neumann entropy of the trace over Γ^c of $\rho^n_{n_1}$:

$$E_{\Gamma,\Gamma^c} = S[\rho^n_{n_1}(\Gamma)] \ . \tag{92}$$

The goal of the rest of this appendix will be to find an analytical expression for the right hand side of Eq.(92).

Because the states $|\psi^n_{n_1}\rangle$ from which $\rho^n_{n_1}(\Gamma)$ is built are fully symmetrized, $\rho^n_{n_1}(\Gamma)$ is the same for all Γ with a fixed cardinality. Thus, we will assume from here on, without loss of generality, that

$$\Gamma^c = Z_{[0,\gamma^c)}, \ \ \Gamma = Z_{[\gamma^c,n)} \ . \tag{93}$$

Thus, in Eq.(91), we will take a trace over the first (top) γ^c qudits of $|\psi^n_{n_1}\rangle\langle\psi^n_{n_1}|$.

Multiplying Eq.(84) times its Hermitian conjugate, and taking the trace of this product over the top qudit, one finds that

$$\rho^n_{n_1}(\Gamma) = \frac{n_0}{n}\rho^{n-1}_{n_1}(\Gamma) + \frac{n_1}{n}\rho^{n-1}_{n_1-1}(\Gamma) \ . \tag{94}$$

Next, we will apply Eq.(94) recursively. Each recursion will "get rid of" the trace over the current top qudit. After $\gamma^c - 1$ recursions, we will have gotten rid of the trace of the top γ^c qudits of the original state $|\psi^n_{n_1}\rangle\langle\psi^n_{n_1}|$.

Here is the result of doing 1 recursion of Eq.(94):

$$\begin{aligned}
\rho^n_{n_1}(\Gamma) &= \frac{n_0}{n}\rho^{n-1}_{n_1}(\Gamma) + \frac{n_1}{n}\rho^{n-1}_{n_1-1}(\Gamma) && (95)\\
&= \frac{n_0}{n}\left[\frac{n_0-1}{n-1}\rho^{n-2}_{n_1}(\Gamma) + \frac{n_1}{n-1}\rho^{n-2}_{n_1-1}(\Gamma)\right] && (96)\\
&\quad + \frac{n_1}{n}\left[\frac{n_0}{n-1}\rho^{n-2}_{n_1-1}(\Gamma) + \frac{n_1-1}{n-1}\rho^{n-2}_{n_1-2}(\Gamma)\right] && (97)\\
&= \frac{n_0^{\downarrow 2}}{n^{\downarrow 2}}\rho^{n-2}_{n_1}(\Gamma) + \frac{2n_0n_1}{n^{\downarrow 2}}\rho^{n-2}_{n_1-1}(\Gamma) + \frac{n_1^{\downarrow 2}}{n^{\downarrow 2}}\rho^{n-2}_{n_1-2}(\Gamma)\ . && (98)
\end{aligned}$$

Here is the result of doing 2 recursions of Eq.(94):

$$\rho^n_{n_1}(\Gamma) = \frac{n_0^{\downarrow 3}}{n^{\downarrow 3}}\rho^{n-3}_{n_1}(\Gamma) + 3\frac{n_0^{\downarrow 2}n_1}{n^{\downarrow 3}}\rho^{n-3}_{n_1-1}(\Gamma) + 3\frac{n_0n_1^{\downarrow 2}}{n^{\downarrow 3}}\rho^{n-3}_{n_1-2}(\Gamma) + \frac{n_1^{\downarrow 3}}{n^{\downarrow 2}}\rho^{n-3}_{n_1-3}(\Gamma)\ . \qquad (99)$$

Note that in Eq.(99),

$$\text{coefficient of } \rho^{n-3}_{n_1-k} = \left(\frac{3!}{n^{\downarrow 3}}\right)\left(\frac{n_0^{\downarrow k}}{k!}\right)\left(\frac{n_1^{\downarrow 3-k}}{(3-k)!}\right) = \frac{\binom{n_0}{k}\binom{n_1}{3-k}}{\binom{n}{3}} \qquad (100)$$

for $k = 0, 1, 2, 3$.

We claim at this point that it is possible to prove by induction that if one goes the full way and recurses Eq.(94) $\gamma^c - 1$ times, then Eq.(100) is still valid but with 3 replaced by γ^c.

The right hand side of Eq.(100) is easy to recognize as the hypergeometric probability distribution. In fact, under the replacements

$$\begin{aligned}
k &\leftrightarrow k && (101)\\
K &\leftrightarrow n_0 && (102)\\
N &\leftrightarrow n && (103)\\
n &\leftrightarrow \gamma^c\ , && (104)
\end{aligned}$$

one gets

$$P(k|K,N,n) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}\ , \qquad (105)$$

which is the definition of the hypergeometric distribution in the notation of Ref.[15]. $P(k)$ for $k \in Z_{[0,n_0]}$ is bell shaped with mean at $\gamma^c\ \frac{n_0}{n}$.

Finally, note that if $\gamma^c = 3$, then in Eq.(99), the states $\rho^{n-3}_{n_1-k}$ on the right hand side are mutually orthogonal pure states. Hence, in the appropriate basis, Eq.(99) is a diagonal matrix. The von Neumann entropy of a diagonal density matrix is just the classical entropy of the probability distribution that appears on the diagonal of that diagonal density matrix.

References

[1] https://en.wikipedia.org/wiki/Squashed_entanglement

[2] Richard E. Blahut, *Principles and Practice of Information Theory* (1987, Addison-Wesley)

[3] Elliott H. Lieb, Mary Beth Ruskai, "Proof of the Strong Subadditivity of Quantum-Mechanical Entropy", Journal of Mathematical Physics 14 (1973) 1938-1941.

[4] Tucci, Robert R. (1999). "Quantum Entanglement and Conditional Information Transmission". arXiv:quant-ph/9909041.

[5] Tucci, Robert R. (2000). "Separability of Density Matrices and Conditional Information Transmission". arXiv:quant-ph/0005119.

[6] Tucci, Robert R. (2000). "Entanglement of Formation and Conditional Information Transmission". arXiv:quant-ph/0010041.

[7] Tucci, Robert R. (2001). "Relaxation Method for Calculating Quantum Entanglement". arXiv:quant-ph/0101123.

[8] Tucci, Robert R. (2001). "Entanglement of Bell Mixtures of Two Qubits". arXiv:quant-ph/0103040.

[9] Tucci, Robert R. (2002). "Entanglement of Distillation and Conditional Mutual Information". arXiv:quant-ph/0202144

[10] Matthias Christandl, Andreas Winter (2003). ""Squashed Entanglement": An Additive Entanglement Measure". Journal of Mathematical Physics. 45 (3): 829–840. arXiv:quant-ph/0308088

[11] P. Hayden, R. Jozsa, D. Petz, and A. Winter. "Structure of states which satisfy strong subadditivity of quantum entropy with equality", Comm. Math. Phys. , 246:359, 2004

[12] Brandao, F., Christandl, M., Yard, J. (September 2011). "Faithful Squashed Entanglement". Communications in Mathematical Physics. 306 (3): 805–830. arXiv:1010.1750

[13] https://en.wikipedia.org/wiki/Perturbation_theory_(quantum_mechanics)

[14] P. Cvitanović, *Group Theory: Birdtracks, Lie's, and Exceptional Groups*, Princeton University Press, Princeton 2008, available online at http://birdtracks.eu/

[15] https://en.wikipedia.org/wiki/Hypergeometric_distribution