

Quantum Information and the Representation Theory of the Symmetric Group

Información Cuántica y la Teoría de Representación del Grupo
Simétrico

ALONSO BOTERO¹

¹Universidad de los Andes, Bogotá, Colombia

ABSTRACT. A number of important results in quantum information theory can be connected quite elegantly to the representation theory of the symmetric group through a quantum analogue of the classical information-theoretic “method of types” that arises naturally from the Schur-Weyl duality. We will give a brief introduction to this connection and briefly discuss some of the results that follow from it, such as quantum source compression rates, entanglement concentration rates, quantum entropy inequalities, and the admissible spectra of partial density matrices from pure, multipartite entangled states.

Key words and phrases. Representation Theory, Quantum Information Theory, Schur-Weyl duality, Quantum Shannon theorem, Entanglement concentration.

2010 Mathematics Subject Classification. 20G05, 94A15, 81P45, 60F10.

RESUMEN. Un gran número de resultados importantes en la teoría de la información cuántica se pueden conectar con la teoría de la representación del grupo simétrico, a través de un análogo cuántico del llamado *método de tipos* que emerge de manera natural de la dualidad de Schur-Weyl. En este artículo daremos una breve introducción a esta conexión y discutiremos algunos resultados que emergen de la misma, como son las tasas de compresión de a fuente cuántica, tasas de concentración de enredamiento, desigualdades de la entropía cuántica, y condiciones sobre los espectros admisibles de matrices parciales de densidad provenientes de un estado cuántico puro multipartito.

Palabras y frases clave. Teoría de la representación, Teoría de la información cuántica, dualidad de Schur-Weyl, Teorema de Shannon cuántico, Concentración de enredamiento.

1. Introduction

Many classical information-theoretic quantities, such as the Shannon entropy, the Kullback-Liebler entropy, or the mutual information, are asymptotic rates with a combinatorial origin, usually the size of certain sets of permutationally-equivalent sequences with some given constraints. Thus, there is a close relationship between information theory and the symmetric (or permutation) group S_n . In recent years, it has become increasingly clear that the symmetric group also plays a fundamental role in quantum information theory (QIT), particularly in asymptotic rate problems involving n copies of some quantum resource. Such problems include source compression [15], spectrum estimation [1, 17], and entanglement concentration [14], to name a few. Moreover, a number of unexpected connections between quantum information and combinatorics, as well as other areas as diverse as geometric complexity theory, have been gradually emerging from applying the representation theory of the symmetric group to quantum information problems. Among these are the unexpected connection between the Kronecker coefficients and the admissible local spectra of reduced density matrices [4]

Here, I would like to give a very brief introduction to some of the basic results that establish the link between quantum information and the representation theory of the symmetric group, and to briefly mention a few of the interesting consequences that have emerged from this connection, which have to do with the Kronecker coefficients and multipartite entanglement. The presentation will be rather informal, aiming more at conveying key ideas and developments; rigorous derivations can be found in the references provided.

2. Rate Exponents and Types

Most relevant quantities in information theory are defined operationally with respect to so-called *extended resources*; that is, the repeated (and generally independent) use of an information-theoretic resource. The quantities are usually *rate exponents* characterizing an asymptotic behavior of numbers of combinatorial origin associated with these extended resources, numbers that generally grow exponentially with the extension parameter n (i.e., the number of repetitions). For instance, given a certain channel Q , the extended channel $Q^{\times n}$ refers to a composite channel made of n independent copies of Q , and the *channel capacity* of Q is defined operationally in terms of how many perfectly distinguishable messages can be transmitted through the extended channel, in the asymptotic limit $n \rightarrow \infty$. To say that the channel Q has a channel capacity C is to say that the maximum number of perfectly distinguishable messages that can be transmitted with the extended channel $Q^{\times n}$ grows like $\sim e^{nC}$ as $n \rightarrow \infty$. Note that in practice, it is customary to characterize the growth in base 2 (e.g., as $\sim 2^{nC}$), not e , but for our purposes it will be more convenient to work with e ; however, in most cases, rate exponents involve a logarithm, so

e^{nr} can also be interpreted as 2^{nr} if the logarithm in r is interpreted as being in base 2 (in which case we say the rate r is given in “bits”).

To make terminology more precise, let $f(x, y, n)$ be a real positive function of n , which may be implicit, and two sets of real parameters: an *extensive* set $x = (x_1, x_2, \dots, x_p)$, where the x_i scale with n , and an *intensive* set $y = (y_1, y_2, \dots, y_q)$, where the y_i scale independently of n . Consider the sequence of scaled values $x_i^{(n)} \equiv n\bar{x}_i$ (or if f is only defined for integer x_i , $x_i^{(n)} \equiv \lceil n\bar{x}_i \rceil$), where \bar{x}_i is real, and let $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p)$. If the limit

$$r_f(\bar{x}, y) = \lim_{n \rightarrow \infty} \frac{1}{n} \log f(x^{(n)}, y, n) \tag{1}$$

exists and is finite, we say that f exhibits *large deviation behavior* (or is *exponential* in n) with *rate function* (or Cramér function) $r_f(\bar{x}, y)$ (see e.g. [30]). For simplicity, we shall henceforth use the notation

$$f(x, y, n) \sim e^{nr(\bar{x}, y)}, \tag{2}$$

to indicate that f has large deviation behavior with rate function $r_f(\bar{x}, y)$.

The asymptotic behavior of multinomial coefficients is the origin of many information-theoretic rates, such as Shannon’s source-encoding rate. Consider sequences of length n , $\vec{x} = (x_1 x_2, \dots, x_n)$, where $x_i \in \mathbb{Z}_d = \{1, 2, \dots, d - 1, d \equiv 0\}$. The *weight* or *type* of the sequence is a frequency distribution $w = \{w_1, w_2, \dots, w_d\}$, where w_x gives the number of times that $x \in \mathbb{Z}_d$ appears in the sequence. The set of all sequences of a given type w is known as the *type class* of w , and all sequences in a given type class can be obtained from each other by a permutation; in other words, type classes are equivalence classes under the action of the symmetric group. The number of sequences in the type class of w is then the multinomial coefficient

$$\binom{n}{w} \equiv \frac{n!}{\prod_{x \in \mathbb{Z}_d} w_x!}. \tag{3}$$

Asymptotically, the multinomial coefficient shows large deviation behavior

$$\binom{n}{w} \sim e^{nH(\bar{w})}, \tag{4}$$

where the rate function $H(\bar{w})$ is the *Shannon entropy* [7]

$$H(\bar{w}) = - \sum_{x \in \mathbb{Z}_d} \bar{w}_x \log \bar{w}_x \tag{5}$$

of the *reduced type*, or relative frequency distribution, $\bar{w} = w/n$ defined by the type w .

It will be useful to discuss another function with large deviation behavior: the multinomial probability distribution for generalized Bernoulli trials. If the above sequences are obtained from n i.i.d. samples from the sampling distribution $p = \{p_1, p_1, \dots, p_d\}$, the multinomial distribution gives the probability of obtaining a sequence in the type class w ; in other words, a sequence with the empirical distribution of symbols w :

$$P(w) = \binom{n}{w} \prod_{x \in \mathbb{Z}_d} p_x^{w_x} = \binom{n}{w} e^{n \sum_{x \in \mathbb{Z}_d} w_x \log p_x}. \quad (6)$$

The multinomial thus has large deviation behavior $P(w) \sim e^{-nD(\bar{w}|p)}$ where (minus) the rate function is the so-called relative (or Kullback-Leibler) entropy [22]

$$D(\bar{w}|p) = \sum_{x \in \mathbb{Z}_d} w_x \log \frac{\bar{w}_x}{p_x}. \quad (7)$$

Though not strictly a distance due to its lack of symmetry, the relative entropy is nevertheless a useful means of quantifying proximity between distributions. It is a convex function of both w and p , attaining its minimum value of zero when $\bar{w} = p$; that is, when the empirical and sampling distributions coincide. Close to this point, it can be expanded as

$$D(\bar{w}|p) \simeq \frac{1}{2} \sum_{i \in \mathbb{Z}_d} \frac{(w_i - p_i)^2}{p_i}, \quad (8)$$

accounting for the asymptotic normality of empirical distributions around the sampling distribution in the large n limit, with $O(1/\sqrt{n})$ fluctuations. Given an information source emitting symbols $x \in \mathbb{Z}_d$, with probabilities $p(x)$, then as $n \rightarrow \infty$, the empirical frequency distribution \bar{w}_x of an i.i.d. sequence \vec{x} from the source will almost surely coincide with the sampling distribution $p(x)$. We can then talk of *typical sequences* as being those for which $\bar{w} \simeq p$ in some sense to be defined. This notion of typicality underlies almost all important results in Information Theory, for instance Shannon's Source compression Theorem.

3. Source compression and the method of types

Shannon's Source compression theorem [29] provides the canonical application of the so-called *method of types* [8], which broadly speaking is the method by which information-theoretic rates are connected to the size of the most probable type class in the given context. Strictly speaking, Shannon's Source compression theorem states that for any given $\delta, \epsilon > 0$, there exists a sufficiently large n_o such that for $n > n_o$, the smallest set S_δ of sequences with probability $P(S_\delta) > 1 - \delta$ has a size $|S_\delta|$ satisfying [22]

$$\left| \frac{1}{n} \log |S_\delta| - H(p) \right| < \epsilon. \quad (9)$$

However, the key insight underlying the theorem is that with probability approaching unity as $n \rightarrow \infty$, an emitted sequence will belong to a *typical set* T of sequences (with type $\bar{w} \simeq p$). Since the number of sequences of type w shows large deviation behavior with rate exponent $H(\bar{w})$, the number of typical sequences satisfies

$$|T| \sim e^{nH(p)}. \tag{10}$$

Interpreted in base-2, this expression means that the sequences in the typical set T can be labeled with a binary word of $\simeq nH(p)$ bits, which amounts to $H(p)$ bits per emitted character; thus, transmitting the binary code of a typical sequence affords a compression ratio of $H(p)/\log_2 d$ with respect to a naive binary encoding of all sequences of length n (of which there are $d^n = 2^{n \log_2 d}$ of them).

Most proofs of Shannon’s theorem therefore rely on some characterization of a typical set consistent with $\bar{w} \simeq p$ and then establishing bounds on the probability and size of the set; these can then be used as bounds for the smallest possible set to which the theorem ultimately refers. By far the most efficient route is the one employed in the so-called Asymptotic Equipartition Theorem (AEP)[7], in which an typical set T_ϵ is defined as the set of sequences \vec{x} such that

$$\left| \frac{1}{n} \log \frac{1}{P(\vec{x})} - H(p) \right| < \epsilon, \tag{11}$$

a property that captures the fact that a sequence with $w \simeq p$ will occur with probability $P(\vec{x}) \simeq \prod_x p(x)^{np(x)} = e^{-nH(p)}$. With this definition, it is easy to bound the size of the set and its probability, resulting in the bounds $\left| \frac{1}{n} \log |T_\epsilon| - H(p) \right| < \epsilon + O(1/n)$, with $1 - P(T_\epsilon) = O(1/n)$. A definition of the typical set that is more faithful to the essence of the method of types is what is known as *strong typicality*[8], in which the typical set T_ϵ is defined as the set of sequences such that their reduced type \bar{w} satisfy

$$|\bar{w} - p| < \epsilon, \tag{12}$$

in Euclidean distance. Counting and calculating the probability of this set is technically more difficult, but the results are nevertheless consistent with those obtained from AEP.

4. Some Quantum Analogs

As an illustration of the method of types, the Source Compression Theorem shows how a classical information-theoretic rate (e.g., the Shannon entropy $H(p)$) is the rate exponent of the size (e.g., the multinomial coefficient) of the most probable type class ($\bar{w} = p$) in a given context (i.i.d. sequences from the sampling distribution p). What we would like to show in this review is how there is a quantum analog of the classical method of types, which follows quite naturally from the representation theory of the symmetric group. With the

quantum method of types (QMT)[13], we should then be able to show how certain quantum Information-theoretic rates can be related, in the context of extended quantum resources, to the rate of exponent of the *dimension* of a typical subspace that transforms irreducibly under the action of the symmetric group, in analogy with type classes in the classical method of types. Before jumping into technicalities, it may prove useful to have in mind some concrete examples to later understand from the perspective of the QMT. Two well-know quantum information rates are the Quantum Source Compression rate, the analog of the Classical Source compression rate, and the Entanglement Concentration rate.

4.1. Quantum Source Compression

The setting for Quantum Source compression is a *quantum source*, a device that outputs some unspecified quantum system with Hilbert-space \mathcal{H} of dimension d , in one of several possible, generally non-orthogonal, quantum states $|\psi_x\rangle \in \mathcal{H}$, with probability p_x . In an extended setting, the source produces n copies of the quantum system in a tensor product state $|\Psi(\vec{x})\rangle = |\psi_{x_1}\rangle|\psi_{x_2}\rangle \dots |\psi_{x_n}\rangle$ with probability $P(\vec{x}) = \prod_i p(x_i)$. The extended source can then be characterized by the mixed quantum state $\rho^{\otimes n} = \rho \otimes \rho \dots \otimes \rho$ (n -times), where ρ is the density matrix of the source

$$\rho = \sum_x p_x |\psi_x\rangle\langle\psi_x|. \quad (13)$$

The idea of compression is to find a Hilbert space \mathcal{V} of dimension e^{nR} , that is hopefully smaller than the extended Hilbert space $\mathcal{H}^{\otimes n}$ (that is, with $R < \log d$) in which the quantum information in $\rho^{\otimes n}$ can be stored faithfully. More precisely, we seek encoding and decoding completely positive maps \mathcal{E} and \mathcal{D} , from the linear operators on $\mathcal{H}^{\otimes n}$ through linear operators on \mathcal{V} ,

$$\mathcal{L}(\mathcal{H}^{\otimes n}) \xrightarrow{\mathcal{E}} \mathcal{L}(\mathcal{V}) \xrightarrow{\mathcal{D}} \mathcal{L}(\mathcal{H}^{\otimes n}), \quad (14)$$

such that $\mathcal{D}(\mathcal{E}(\rho^{\otimes n}))$ is close to $\rho^{\otimes n}$ (for instance in the trace-distance sense $\|\mathcal{D}(\mathcal{E}(\rho^{\otimes n})) - \rho^{\otimes n}\|_1 < \epsilon$) [25]. Schumacher's source compression theorem [28] states that for sufficiently large n , all rates above the von Neumann entropy of ρ ,

$$S(\rho) = -\text{Tr}(\rho \log \rho) \quad (15)$$

are achievable with trace-distance of at most ϵ .

4.2. Entanglement Concentration

Closely related to Quantum Source Compression is Entanglement Concentration. Here, the idea is to take n copies of a certain entangled (but not maximally entangled) bipartite state $|\psi\rangle_{AB}$ shared between two parties A and B that is

entangled and concentrate the entanglement into a certain number of maximally entangled states $|\Gamma\rangle_{AB}$ between two Hilbert spaces, each of dimension d , shared by parties A and B , assuming the protocol uses only local resources and, possibly, classical communication between the parties. As shown by Popescu and Rohrlich [26], asymptotically, an optimal protocol is achievable for the reversible interconversion

$$|\psi\rangle_{AB}^{\otimes n} \leftrightarrow |\Gamma\rangle_{AB}^{\otimes m}, \quad m = \frac{nE(\psi)}{\log d}, \tag{16}$$

where $E(\psi)$ is the so-called entanglement entropy, which is also the von Neumann entropy of the reduced density matrices ρ_A and ρ_B (they are equal for pure states),

$$E(\psi) = S(\rho_A) = S(\rho_B). \tag{17}$$

where

$$\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|), \quad \rho_B = \text{Tr}_A(|\psi\rangle\langle\psi|). \tag{18}$$

In short, the idea behind these optimal rates is that the relevant quantum information in $|\psi\rangle_{AB}^{\otimes n}$ can be projected on a maximally entangled state between Hilbert spaces of dimension $e^{nE(\psi)}$, which in turn can be broken up into $\frac{nE(\psi)}{\log d}$ maximally entangled states between spaces of dimension d .

5. Partitions and the Representation Theory of S_n and $GL(d)$

Quantum states of a d -level system are represented by vectors $|\phi\rangle \in \mathcal{H}$, where $\mathcal{H} = \mathbb{C}^d$. The Hilbert space \mathcal{H} is therefore the carrier space for the defining representation of the general linear group $GL(d, \mathbb{C})$, the group of invertible $d \times d$ complex valued matrices. Irreducible representations of $GL(d, \mathbb{C})$ arise naturally when considering the description of a composite system made up of n copies of the d - level system, the quantum states of which are represented as vectors $|\Phi\rangle \in \mathcal{H}^{\otimes n}$, in the product space $\mathcal{H}^{\otimes n} = \mathbb{C}^d \otimes \mathbb{C}^d \otimes \dots \otimes \mathbb{C}^d$ (n times). The product space is then a carrier space for a reducible representation of G , under the action

$$R(g) |\Phi\rangle = g^{\otimes n} |\Phi\rangle, \tag{19}$$

with the same element g acting on each Hilbert space.

Similarly, the product space $\mathcal{H}^{\otimes n} = \mathbb{C}^d \otimes \mathbb{C}^d \otimes \dots \otimes \mathbb{C}^d$ (n times) is also the carrier space for a reducible representation of the symmetric group S_n of permutations of n elements. To define this action, note that the product space $\mathcal{H}^{\otimes n}$ inherits from its construction a natural product basis built from some standard basis $|1\rangle, |2\rangle \dots |d\rangle$ of the single-copy Hilbert space \mathcal{H} . We refer to this basis as the *computational basis* in analogy with the case $d = 2$. The elements of this basis are labelled by sequences $\mathbf{s} \in (\mathbb{Z}_d)^n$ where

$$|\mathbf{s}\rangle = |s_1\rangle |s_2\rangle \dots |s_n\rangle. \tag{20}$$

The action of the permutation group on $\mathcal{H}^{\otimes n}$ is then best summarized by in terms of its action on the computational basis, namely

$$U(\pi) |s\rangle = |\pi s\rangle, \tag{21}$$

where πs is the corresponding permutation of the elements of the sequence s as defined earlier.

The reducible actions of $GL(d, \mathbb{C})$ and S_n , as defined above can be seen to commute. Moreover, the operator algebras generated by the elements of the respective representations in $\mathcal{H}^{\otimes n}$ are each other's commutant; thus, if A is an operator acting on $\mathcal{H}^{\otimes n}$ with the property $A = U(\pi)AU^\dagger(\pi)$, the A can be expressed as a linear combination of the group elements $g^{\otimes n}$ for $g \in GL(d, \mathbb{C})$; similarly, any $GL(d, \mathbb{C})$ -invariant operator B (under the above action), can be written as a linear combination of the operators $U(\pi)$, for $\pi \in S_n$. This mutual commutancy relation between the group algebras is known as the *Schur-Weyl duality*[13, 23]. A straightforward consequence of this duality is a tight coordination in the breakup of the tensor space $H^{\otimes n}$ into irreducible sectors under $GL(d, C)$ and S_n ; in particular, irreducible representations (IRRs) of both groups are labeled by *partitions* λ of n , and denoting by V_λ and $[\lambda]$ the carrier spaces for the $GL(d, C)$ and S_n IRRs labeled by λ , the decomposition of $H^{\otimes n}$ is such that the multiplicity space for the IRR $[\lambda]$ of S_n is precisely the the IRR space \mathcal{V}_λ of $GL(d, \mathbb{C})$, and vice versa. Formally, one has

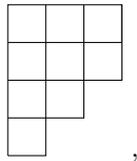
$$\mathcal{H}^{\otimes n} = \bigoplus_{\lambda \vdash_d n} \mathcal{V}_\lambda \otimes [\lambda], \tag{22}$$

where $\lambda \vdash_d n$ means that λ is a *partition* of at most d rows.

Let us then discuss some important aspects of the relationship between the IRRs of the symmetric group and the general linear group and partitions, including the dimensions of the IRRs, which are given by the size of certain sets of *Young tableaux* [10], which are closely tied to the graphical representation of partitions. A partition λ is a weakly decreasing sequence of nonnegative integers, λ_i , also known as parts:

$$\lambda = (\lambda_1 \lambda_2 \dots), \quad \lambda_1 \geq \lambda_2 \geq \lambda_3 \dots, \quad \lambda_i \geq 0, \tag{23}$$

and where $|\lambda| = \sum_i \lambda_i$ is known as the size of the partition. Partitions are conventionally represented by *Young frames*, which are diagrams of the form



with $|\lambda|$ boxes, and where the parts are represented by the rows in descending order, so that λ_1 gives the number of boxes in the first (uppermost) row, λ_2 the number of boxes in the second row, etc; it is understood that all parts beyond the last row are zero. A partition of n , denoted as $\lambda \vdash n$, is a partition in which $|\lambda| = n$. We also use the notation $\lambda \vdash_d n$ when λ is a partition of n with at most d rows.

A Young tableaux is a filling of the boxes in a Young frame with numbers according to certain prescriptions. The irreducible representations of the symmetric group S_n and the general linear group $GL(d, \mathbb{C})$ are respectively connected to two types of such fillings, the so-called *standard Young tableaux* (SYT) and the *semi-standard Young tableaux* (SSYT). We begin with the definition of the latter:

Tensors that transform irreducibly under $GL(d, \mathbb{C})$ satisfy certain symmetry conditions encoded in the partition λ that labels the corresponding irreducible representation, henceforth denoted as V_λ . Specifically, each box in a Young frame λ is understood to stand for one of the indices of the tensor, and the tensor is understood to be antisymmetric under permutation of indices that are common to a given column and symmetric under permutation of indices that are common to a given row[11]. Each independent component of such a tensor is labeled by a possible semi-standard Young tableau (SSYT) t of shape λ , which is a filling of the Young frame λ with the integers $1, 2, \dots, d$ and satisfying the condition that the numbers are non-decreasing along the rows and strictly increasing along columns.

The dimension of V_λ , is then the set of all SSYT of shape λ filled with numbers $\in \{1, 2, \dots, d\}$, and is given by the formula

$$\dim(V_\lambda) = \frac{\prod_{1 \leq i < j \leq d} (\lambda_i - \lambda_j - i + j)}{1!2! \dots (d-1)!}. \tag{24}$$

which is a straightforward consequence of the Weyl character formula [12] for highest-weight representations. For fixed d and $\bar{\lambda}_i$, this number asymptotically scales polynomially with n , typically as $\sim n^{d(d-1)/2}$. In fact, it satisfies the bound[4]

$$\dim(V_\lambda) \leq (n+1)^{d(d-1)/2}. \tag{25}$$

In the case of the symmetric group, we follow the notation of [16] and denote by $[\lambda]$ the irreducible representation corresponding to the partition λ . The basis vectors of $[\lambda]$ are in correspondence with the independent ways in which the indices of the tensor can be made to satisfy the symmetry rules of the GL irreducible tensor of symmetry class λ . A basis vector is hence encoded in a *standard Young tableau* (SYT) τ of shape λ , which is a filling of the Young frame λ with the integers from 1 to $n = |\lambda|$, and satisfying the condition that the numbers are strictly increasing along rows and along columns. Thus, for the shape $\lambda = (3, 2)$, the following are all possible SYT:

1	2	3
4	5	

1	2	4
3	5	

1	2	5
3	4	

1	3	4
2	5	

1	3	5
2	4	

The dimension of the S_n irreducible representation $[\lambda]$, henceforth denoted by f_λ , is given by the so-called *hook length formula* of Frame, Robinson and Thrall [9]:

$$f_\lambda \equiv \dim([\lambda]) = \frac{n!}{\prod_{i,j} h_{i,j}}, \tag{26}$$

where the product runs over all boxes (i, j) in the frame, where i and j are row and column coordinates respectively, and $h_{i,j}$ is the so-called *hook length* of the frame (i, j) ; namely, $h_{i,j}$ is one more than the number of boxes to the right or below the box (i, j) . As opposed to GL irreducible representations, the dimensions of the S_n irreducible representations grow exponentially with n . To see this, we use the following bounds [4]:

$$\frac{n!}{\nu_1! \nu_2! \dots \nu_d!} \leq f_\lambda \leq \frac{n!}{\lambda_1! \lambda_2! \dots \lambda_d!}, \quad \nu_i = \lambda_i - l(\lambda) - i, \tag{27}$$

where $l(\lambda)$ is the number of rows in λ . Using Stirling’s formula and taking the limit $n \rightarrow \infty$ with d and the ratios $\frac{\lambda_i}{n}$ fixed, it is seen that the rate exponent is given by the Shannon entropy $H(\bar{\lambda})$ of the so called *reduced partition* $\bar{\lambda}$:

$$f_\lambda \sim e^{nH(\bar{\lambda})}, \quad \bar{\lambda} = \left(\frac{\lambda_1}{n}, \frac{\lambda_2}{n}, \dots, \frac{\lambda_d}{n} \right). \tag{28}$$

Given the Schur-Weyl decomposition (22), an orthonormal basis for $\mathcal{H}^{\otimes n}$ adapted to this decomposition will comprise basis elements $|\lambda, t, \tau\rangle$ (or when written in product form, $|t\rangle_\lambda |\tau\rangle_\lambda$), which are naturally labeled by three objects: the partition $\lambda \vdash_d n$, an Semi-standard Young Tableau t , and a Standard Young Tableau, τ . While the labelling scheme (t, τ) can be applied to any basis that one may choose for $\mathcal{U}_\lambda \otimes \mathcal{V}_\lambda$, it is especially adapted to canonical choices of basis in which the tableaux t and τ have a combinatorial meaning relative to the computational basis, such as the so-called Gelfand-Tsetlin basis for V_λ (see e.g., [21]), and the Young-Yamanouchi basis for S_n [6].

In this decomposition, a permutation $\pi \in S_n$ is represented by

$$U(\pi) \equiv \bigoplus_{\lambda \vdash_d n} 1_{V_\lambda} \otimes S_\lambda(\pi), \tag{29}$$

where $S_\lambda(\pi)$ is the IRR matrix for π in the representation $[\lambda]$. Similarly, for the action of $GL(d)$, the n -fold tensor operator $g^{\otimes n}$ is represented by

$$g^{\otimes n} = \bigoplus_{\lambda \vdash_d n} D_\lambda(g) \otimes 1_{[\lambda]}, \tag{30}$$

where $D_\lambda(g)$ is the corresponding V_λ representation matrix. The $GL(d)$ characters will be important in what follows. For a given $g \in GL(d)$, let $\gamma = \text{spec}(g)$; then

$$\text{Tr}(D_\lambda(g)) = s_\lambda(\gamma_1, \gamma_2, \dots, \gamma_d), \tag{31}$$

where $s_\lambda(x_1, x_2, \dots)$ are the Schur polynomials, which can be defined in terms of a sum over semi-standard Young Tableaux of shape λ :

$$s_\lambda(x_1, x_2, \dots) = \sum_{t \in SSYT(\lambda)} x_1^{w_1(t)} x_2^{w_2(t)} \dots, \tag{32}$$

where $w_i(t)$ denotes the weight of the integer i in the tableau t .

6. The Keyl-Werner Theorem and Quantum Source Compression

The previous section introduced the main ingredients that are necessary to establish a connection between quantum information rates and large-deviation behavior of the dimensions of irreducible representations of S_n , following an analogy with the classical method of types. It is then time to flesh out this connection. Our first step will be to establish the quantum concept that is analogous to the classical concept of ‘‘type’’ (as in sequence type). In the quantum case, this will be played by the reduced partition $\bar{\lambda} = \lambda/n$ associated with a given irreducible sector. As in the classical case, the reduced partition is a distribution normalized to unity, although the meaning of the individual parts $\bar{\lambda}_i = \lambda_i/n$ may seem to be rather abstract at first. The Keyl-Werner theorem provides the necessary typicality connection with properties of a quantum source in the asymptotic regime; namely, in the same way that classically, typical sequences are those with reduced weights \bar{w} close to the probability p distribution of the source, quantum-mechanically, the typical reduced partitions are those that are close to the spectrum of the density matrix ρ describing the quantum source, when the eigenvalues are arranged in non-decreasing order.

To see this, let us consider the product $\rho^{\otimes n}$ of a density matrix ρ acting on \mathcal{H} . According to (30), we can write this as

$$\rho^{\otimes n} = \sum_{\lambda \vdash dn} D_\lambda(\rho) \otimes 1_{[\lambda]}. \tag{33}$$

Thus, in a measurement of projectors onto the different λ sectors, the probability of obtaining the result λ is

$$P(\lambda) = \text{Tr}(D_\lambda(\rho))\text{Tr}(1_{[\lambda]}) = s_\lambda(r_1, r_2, \dots, r_d) f_\lambda, \tag{34}$$

where r_1, r_2, \dots, r_d are the eigenvalues of ρ . Now, assume that the eigenvalues have been arranged in non-increasing order, so that $r_1 \geq r_2 \geq \dots \geq r_d$. Then, using the definition (32), the Schur polynomial can be bounded by

$$r_1^{\lambda_1} r_2^{\lambda_2} \dots r_d^{\lambda_d} \leq s_\lambda(r_1, r_2, \dots, r_d) \leq \dim(V_\lambda) r_1^{\lambda_1} r_2^{\lambda_2} \dots r_d^{\lambda_d}. \tag{35}$$

Using the bounds (25) and (27), it is then straightforward to show that $P(\lambda)$ satisfies a large deviation behavior that is similar to that of the multinomial distribution in the classical case; namely,

$$P(\lambda) \sim e^{-nD(\bar{\lambda}||r_{\downarrow})}, \quad (36)$$

where r_{\downarrow} is the spectrum of ρ ordered non-increasingly. Thus, typical diagrams are those such that

$$\bar{\lambda} \simeq r_{\downarrow}, \quad (37)$$

with $o(n^{-1/2})$ statistical fluctuations. The Keyl-Werner theorem[17] is a statement of the convergence of $\bar{\lambda}$ to r_{\downarrow} ; it states that for all continuous functions g on the probability simplex,

$$\lim_{n \rightarrow \infty} \sum_{\lambda \vdash n} g(\bar{\lambda}) P(\lambda) = g(r_{\downarrow}), \quad (38)$$

uniformly.

The connection with the quantum source compression rate is now straightforward. Defining a typical subspace by the projection operator $\Pi(T_{\epsilon}) = \bigoplus_{\bar{\lambda} \in T_{\epsilon}} \Pi_{\lambda}$, where T_{ϵ} is the set of typical partitions such that $|\bar{\lambda} - r_{\downarrow}| < \epsilon$, then, using the arguments from the classical method of types, we should have that for any $\epsilon > 0$, the ‘‘typical sector’’ $\Pi(T_{\epsilon})\rho^{\otimes n}\Pi(T_{\epsilon})$ should account for the quantum information in $\rho^{\otimes n}$:

$$\lim_{n \rightarrow \infty} \|\rho^{\otimes n} - \Pi(T_{\epsilon})\rho^{\otimes n}\Pi(T_{\epsilon})\|_1 = 1 \quad (39)$$

and that the size of this typical sector should scale exponentially with a rate exponent that is given by that of the dimension of the typical λ -subspace. Since $\dim(V_{\lambda})$ scales polynomially in n , this rate exponent is that of the IRR $[\lambda]$, which is the Shannon entropy. Evaluating this at the typical reduced partition $\bar{\lambda} = r_{\downarrow}$, the size of the typical subspace should then scale asymptotically as

$$\text{Tr}(\Pi(T_{\epsilon})) \sim e^{n(S(\rho)+c\epsilon)}, \quad (40)$$

where c is some positive constant independent of N and ϵ . Thus, the quantum information in $\rho^{\otimes n}$ can be faithfully transmitted using a smaller Hilbert space than $\mathcal{H}^{\otimes n}$, the dimension of which has an asymptotic rate of $S(\rho)$, the quantum source compression rate.

7. Entanglement Concentration

As suggested earlier, the Entanglement Concentration protocol is closely related to Quantum Source compression. In this case, the extended quantum resources is the n -fold tensor product $|\psi\rangle^{\otimes n}$ of an entangled bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. Without loss of generality, we can assume that the Hilbert spaces \mathcal{H}_A

and \mathcal{H}_B have the same dimension d . Therefore, $\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n}$ admits the Schur-Weyl decomposition

$$\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n} = \bigoplus_{\lambda, \mu \vdash dn} V_\lambda \otimes V_\mu \otimes [\lambda] \otimes [\mu]. \tag{41}$$

However, since the state $|\psi\rangle^{\otimes n}$ is a product of the same state, it is invariant under the simultaneous action of the permutation group on $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$, i.e.,

$$U_A(\pi) \otimes U_B(\phi) |\psi\rangle^{\otimes n} = |\psi\rangle^{\otimes n}. \tag{42}$$

This means that $|\psi\rangle^{\otimes n}$ can only be supported on the (λ, μ) subspaces of (41) such that there exists an S_n invariant subspace in $[\lambda] \otimes [\mu]$. In fact, the only possibility for this to happen is that $\lambda = \mu$, and in the tensor product $[\lambda] \otimes [\lambda]$, the invariant subspaces are one-dimensional, with the invariant normalized vector given by[24]

$$|\Gamma_\lambda\rangle = f_\lambda^{-1/2} \sum_{\tau \in SYT(\lambda)} |\tau\rangle_\lambda |\tau\rangle_\lambda, \tag{43}$$

a maximally entangled state. Hence, we can write

$$|\psi\rangle^{\otimes n} = \sum_\lambda \sqrt{P(\lambda)} |R_\lambda(\psi)\rangle |\Gamma_\lambda\rangle \tag{44}$$

where $|R_\lambda(\psi)\rangle$ is a normalized vector in $V_\lambda \otimes V_\lambda$, the coefficients of which depend on ψ , and $P(\lambda)$ turns out to be the same probability of equation (34), with r_\downarrow equal to the spectrum of the partial density matrices ρ_A and ρ_B of the state $|\psi\rangle\langle\psi|$. This follows from the fact that the partial density matrices of $|\psi\rangle^{\otimes n}$ are $\rho_A^{\otimes n}$ and $\rho_B^{\otimes n}$. Next, we imagine that a projective measurement of λ is performed on any side. Then the state collapses to a state within a given λ -sector:

$$|\psi\rangle^{\otimes n} \rightarrow |R_\lambda(\psi)\rangle |\Gamma_\lambda\rangle. \tag{45}$$

Given that the resulting state is in product form, we can then concentrate only on the maximally-entangled state $|\Gamma_\lambda\rangle$, which has an entanglement entropy $E(\Gamma_\lambda) = \log f_\lambda \simeq nH(\bar{\lambda})$. Moreover, from the Keyl-Werner theorem, we know that as $n \rightarrow \infty$, the λ -measurement will almost surely yield the value of λ such that $\bar{\lambda} = r_\downarrow$. Hence, as $n \rightarrow \infty$, we obtain a maximally entangled state with entanglement entropy, per copy, given by

$$\lim_{n \rightarrow \infty} \frac{1}{n} E(\Gamma_\lambda) = S(\rho_A); \tag{46}$$

in other words, the entanglement entropy $E(\psi)$ of the state $|\psi\rangle$.

8. Tripartite Entanglement and Kronecker Coefficients

The extension of the previous results to the tripartite case, brings out a deeper connection between the representation theory of S_n and quantum information. In this case we consider the n -fold tensor product $|\psi\rangle^{\otimes n}$ of an entangled state $|\psi\rangle$ shared between parties A, B and C . As in the bipartite case, the extended Hilbert space decomposes into local $GL(d)$ and S_n irreducible subspaces as

$$\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_C^{\otimes n} = \bigoplus_{\lambda, \mu, \nu \vdash n} V_\lambda \otimes V_\mu \otimes V_\nu \otimes [\lambda] \otimes [\mu] \otimes [\nu], \quad (47)$$

and due to permutation symmetry, $|\psi\rangle^{\otimes n}$ is only supported on the subspaces labeled by triplets (λ, μ, ν) for which the tensor product $[\lambda] \otimes [\mu] \otimes [\nu]$ admits an invariant subspace. The dimension of these invariant subspaces is given by the so-called Kronecker coefficients $g_{\lambda\mu\nu}$

$$g_{\lambda\mu\nu} = \dim([\lambda] \otimes [\mu] \otimes [\nu])_{S_n} = \frac{1}{n!} \sum_{\pi \in S_n} \chi_\lambda(\pi) \chi_\mu(\pi) \chi_\nu(\pi), \quad (48)$$

where $\chi_\lambda(\pi)$ are the S_n -characters. The Kronecker coefficients can also be interpreted as the multiplicity of the S_n IRR $[\nu]$ in the Clebsch-Gordan expansion of the tensor product $[\lambda] \otimes [\mu]$, i.e.,

$$[\lambda] \otimes [\mu] = \sum_{\nu \vdash n} g_{\lambda\mu\nu} [\nu]. \quad (49)$$

So far, no general combinatorial formula for the Kronecker coefficients exists, and moreover, it is known that their computation is of $\#P$ complexity[2].

A decomposition of $|\psi\rangle^{\otimes n}$ analogous to that of (44)

$$|\psi\rangle^{\otimes n} = \sum_{(\lambda, \mu, \nu): g_{\lambda\mu\nu} > 0} \sqrt{P(\lambda\mu\nu)} \sum_{i=1}^{g_{\lambda\mu\nu}} |R_{\lambda\mu\nu}; i\rangle |K_{\lambda\mu\nu}; i\rangle, \quad (50)$$

where the vectors $|K_{\lambda\mu\nu}; i\rangle$ span the invariant subspace $([\lambda] \otimes [\mu] \otimes [\nu])^{S_n}$, the vectors $|R_{\lambda\mu\nu}; i\rangle \in V_\lambda \otimes V_\mu \otimes V_\nu$ are relative vectors to the $|K_{\lambda\mu\nu}; i\rangle$, and $P(\lambda, \mu, \nu)$ is the joint probability of obtaining the triplet (λ, μ, ν) in a joint measurement of Young subspaces.

As of now, no closed-form expression is known for the joint probability $P(\lambda, \mu, \nu)$. Nonetheless, the marginal probability distributions $P(\lambda) = \sum_{\mu, \nu} P(\lambda, \mu, \nu)$, etc., are those given by equation (34), in terms of the spectra r_A, r_B and r_C of partial density matrices

$$\rho_A = \text{tr}_{BC}(\rho_{ABC}), \quad \rho_B = \text{tr}_{AC}(\rho_{ABC}), \quad \rho_C = \text{tr}_{AB}(\rho_{ABC}), \quad (51)$$

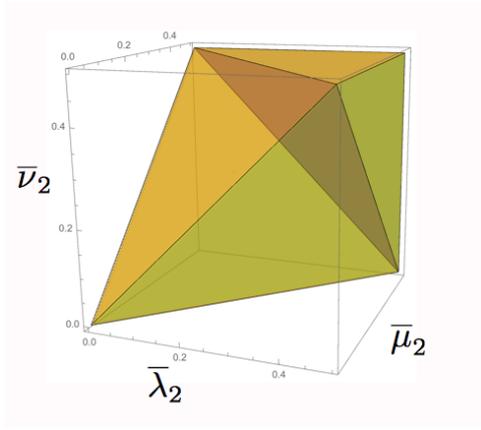


FIGURE 1. Closure of the Kron(2) polytope, which is equivalent to the polytope of admissible marginal spectral triplets for three-qubit pure states. The coordinates $(\bar{\lambda}_2, \bar{\mu}_2, \bar{\nu}_2)$ are the second parts of the reduced partitions $(\bar{\lambda}, \bar{\mu}, \bar{\nu})$.

which henceforth we will assume are give in non-decreasing order. This means that for large n , the joint distribution should become sharply peaked about the reduced triplet $(\bar{\lambda}, \bar{\mu}, \bar{\nu})$ that corresponds to the spectral triplet (r_A, r_B, r_C) ; explicitly,

$$\lim_{n \rightarrow \infty} (\bar{\lambda}, \bar{\mu}, \bar{\nu}) = (r_A, r_B, r_C) \tag{52}$$

almost surely [3]. This result suggests that there is a close relationship between sets of admissible marginal spectra of tripartite entangled states and the Kronecker coefficients. In fact, let $\text{Kron}(d)$ be the set of all reduced, Kronecker-compatible partitions of at most d -rows:

$$\text{Kron}(d) = \left\{ \frac{1}{n}(\lambda, \mu, \nu) \mid n \in \mathbb{N}, g_{\lambda, \mu, \nu} > 0 \right\}. \tag{53}$$

Then, as shown by Christandl et al [3], the set of admissible spectral triplets (r_A, r_B, r_C) is the closure of $\text{Kron}(d)$. Figure 1 shows the spectral polytope obtained from $\overline{\text{Kron}}(2)$, which can be computed from known closed-form formulas for the Kronecker coefficients for two-row partitions [27]. This somewhat unexpected result complements several other recent and interconnected results connecting the spectral properties of (density) matrices with linearization coefficients of polynomial group IRRs, such as the complete characterization of the so-called Horn inequalities for the sum of three matrices [19], and the complete solution to the one-body n -representability problem [18].

Perhaps one of the most elegant consequences of the connection between Kronecker coefficients, spectral properties of entangled states, and the Keyl-Werner theorem described earlier is a remarkably simple proof of the entropy subadditivity inequality for the marginal density matrices [4], i.e., :

$$S(\rho_C) \leq S(\rho_A) + S(\rho_B), \quad \& \text{ cyc. perm.}, \quad (54)$$

which is a straightforward consequence of the IRR dimension inequality for compatible triplets

$$\dim([\lambda])\dim([\mu]) \geq \dim([\nu]), \quad \& \text{ cyc. perm.}, \quad (55)$$

and which, through purification, implies the subadditivity of the von Neumann entropy:

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B). \quad (56)$$

It is worth noting that more recently, a closely-related representation-theoretic proof of the celebrated strong subadditivity inequality [20]

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC}), \quad (57)$$

has also been obtained in terms of the asymptotic behavior of the so-called recoupling coefficients [5].

9. Conclusion

I would like to thank the organizers for inviting me to present this very brief introduction to some of the most important developments in the connection between quantum information theory and the representation theory of symmetric group. I hope I have been able to convey some the excitement surrounding the currently very active field, which promises to yield new and surprising results connecting several diverse areas.

References

- [1] Robert Alicki, Slawomir Rudnicki, and Slawomir Sadowski, *Symmetry properties of product states for the system of n -level atoms*, Journal of Mathematical Physics **29** (1988), no. 5, 1158–1162.
- [2] P. Bürgisser and Christian Ikenmeyer, *The complexity of computing Kronecker coefficients*, DMTCS Proceedings (2008), 357–368.
- [3] Matthias Christandl, Aram W. Harrow, and Graeme Mitchison, *Nonzero Kronecker Coefficients and What They Tell us about Spectra*, Commun. Math. Phys. **270** (2007), no. 3, 575–585 (en).

- [4] Matthias Christandl and Graeme Mitchison, *The Spectra of Quantum States and the Kronecker Coefficients of the Symmetric Group*, Commun. Math. Phys. **261** (2005), no. 3, 789–797 (en).
- [5] Matthias Christandl, Mehmet Burak Sahinoglu, and Michael Walter, *Recoupling Coefficients and Quantum Entropies*, arXiv:1210.0463 [math-ph, physics:quant-ph] (2012), arXiv: 1210.0463.
- [6] Peter Dal Cin, *Groups, Systems and Many-Body Physics*, Springer Science & Business Media, April 2013 (en).
- [7] T. M. Cover and Joy A. Thomas, *Elements of information theory*, 2nd ed ed., Wiley-Interscience, Hoboken, N.J, 2006.
- [8] Imre Csiszar, *The method of types*, Information Theory, IEEE Transactions on **44** (1998), no. 6, 2505–2523.
- [9] J. S. Frame, G. de B. Robinson, and R. M. Thrall, *The hook graphs of the symmetric group*, Canadian Journal of Mathematics **6** (1954), no. 0, 316–324 (en).
- [10] William Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry*, Cambridge University Press, 1997 (en).
- [11] William Fulton and Joe Harris, *Representation Theory: A First Course*, Springer Science & Business Media, December 2013 (en).
- [12] Brian C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer, May 2015 (en).
- [13] Aram W. Harrow, *Applications of coherent classical communication and the Schur transform to quantum information theory*, arXiv:quant-ph/0512255 (2005), arXiv: quant-ph/0512255.
- [14] Masahito Hayashi, Masato Koashi, Keiji Matsumoto, Fumiaki Morikoshi, and Andreas Winter, *Error exponents for entanglement concentration*, J. Phys. A: Math. Gen. **36** (2003), no. 2, 527 (en).
- [15] Masahito Hayashi and Keiji Matsumoto, *Quantum universal variable-length source coding*, Phys. Rev. A **66** (2002), no. 2, 022311.
- [16] James, *The Representation Theory of the Symmetric Group*, Cambridge University Press, March 2009 (en).
- [17] M. Keyl and R. F. Werner, *Estimating the spectrum of a density operator*, Phys. Rev. A **64** (2001), no. 5, 052311.

- [18] Alexander A. Klyachko, *Quantum marginal problem and N -representability*, Journal of Physics: Conference Series **36** (2006), 72–86.
- [19] Allen Knutson and Terence Tao, *Honeycombs and sums of Hermitian matrices*, Notices Amer. Math. Soc **48** (2001), no. 2.
- [20] Elliott H. Lieb and Mary Beth Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, Journal of Mathematical Physics **14** (1973), no. 12, 1938–1941.
- [21] James D. Louck, *Unitary Symmetry and Combinatorics*, World Scientific, 2008 (en).
- [22] David Mackay, *Information Theory, Inference And Learning Algorithms*, Cambridge University Press, April 2005 (en).
- [23] Iman Marvian and Robert W. Spekkens, *A Generalization of Schur-Weyl Duality with Applications in Quantum Estimation*, Commun. Math. Phys. **331** (2014), no. 2, 431–475 (en).
- [24] Keiji Matsumoto and Masahito Hayashi, *Universal distortion-free entanglement concentration*, Phys. Rev. A **75** (2007), no. 6, 062338.
- [25] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, December 2010 (en).
- [26] Sandu Popescu and Daniel Rohrlich, *Thermodynamics and the measure of entanglement*, Phys. Rev. A **56** (1997), no. 5, R3319–R3321.
- [27] Mercedes H. Rosas, *The Kronecker Product of Schur Functions Indexed by Two-Row Shapes or Hook Shapes*, Journal of Algebraic Combinatorics **14** (2001), no. 2, 153–173 (en).
- [28] Benjamin Schumacher, *Quantum coding*, Phys. Rev. A **51** (1995), no. 4, 2738–2747.
- [29] Claude Elwood Shannon and Warren Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, 1998 (en).
- [30] S. R. S. Varadhan, *Large Deviations and Applications*, SIAM, 1984 (en).

(Recibido en julio de 2016. Aceptado en noviembre de 2016)

DEPARTAMENTO DE FÍSICA
UNIVERSIDAD DE LOS ANDES
CARRERA 1A NO. 18A-12
BOGOTÁ, COLOMBIA
e-mail: abotero@uniandes.edu.co