A most compendious and facile quantum de Finetti theorem

Robert König[*] and Graeme Mitchison[†]

Centre for Quantum Computation, DAMTP, University of Cambridge, Cambridge CB3 0WA, UK

In its most basic form, the finite quantum de Finetti theorem states that the reduced k-partite density operator of an n-partite symmetric state can be approximated by a convex combination of k-fold product states. Variations of this result include Renner's "exponential" approximation by "almost-product" states, a theorem which deals with certain triples of representations of the unitary group, and D'Cruz et al.'s result for infinite-dimensional systems. We show how these theorems follow from a single, general de Finetti theorem for representations of symmetry groups, each instance corresponding to a particular choice of symmetry group and representation of that group. This gives some insight into the nature of the set of approximating states, and leads to some new results, including an exponential theorem for infinite-dimensional systems.

I. INTRODUCTION

Edmund Halley, of the comet, used the opening words of our title to describe a method for calculating logarithms, making it clear that his principal claim was to an increased ease and scope [1]. This is our claim too: our main theorem gathers together the currently available examples of quantum de Finetti theorems and gives them a larger setting and what we hope is an illuminating and accessible proof.

Our theorem is of the general type first proved in [2] and recently reviewed in [3]. Recall that the standard de Finetti theorem says that the state obtained by tracing out $n-k$ parts from a symmetric n-partite state can be approximated by a convex sum of product states of the form $\sigma^{\otimes k}$ [4, 5], the error in the approximation being of size $O(k/n)$ for a fixed dimension of the individual subsystems [5]. This theorem has many uses [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16], but for some purposes one needs the type of theorem mentioned above, where the class of approximating states is broadened and the convergence is much more rapid [2, 6]. The approximating states in question are *almost-product states* [2], that are sums of terms of the form $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_k\rangle$, where all but a few of the $|\psi_i\rangle$ are identical. The convergence now has an exponential character, with an error of order $e^{-\gamma n}$ for some $\gamma > 0$ (where γ depends on k/n and the fraction of $|\psi_i\rangle$ that differ; see Corollary V.2). We shall therefore refer to this as an *exponential* theorem; it implies that, with almost product states, far fewer subsystems have to be traced out to get a close approximation. For many physical questions, the fact that a few subsystems differ from the rest has only a small effect; for instance, most thermodynamic measures will register little difference between product states and almost-product states [3].

The setting for our theorem is representation theory, and the class of approximating states is determined by some subspace $\mathcal{X}$ of the representation. One case of interest is where there is a weight space structure, and where $\mathcal{X}$ consists of high weight vectors. Such vectors are closely related to almost-product states. To see how this works, consider the representation of $\mathcal{SU}(d)$ on the symmetric subspace $\mathsf{Sym}^k(\mathbb{C}^d)$ of $(\mathbb{C}^d)^{\otimes k}$. The highest weight vector is $|1\rangle^{\otimes k}$, and any product state $|\psi\rangle^{\otimes k}$ can be obtained by applying some element g of $\mathcal{SU}(d)$ to this vector. Thus a product state is a rotated highest weight vector. There is an ordering of weights, the next highest weight vector being $(|211\cdots11\rangle + |121\cdots11\rangle + \cdots + |111\cdots12\rangle)/\sqrt{k}$, and a symmetric almost-product state with all but one of its factors identical can be obtained by applying some $\mathsf{g} \in \mathcal{SU}(d)$ to this vector.

The original exponential theorem in [2] can be obtained from the symmetric representation $\mathsf{Sym}^k(\mathbb{C}^d)$ in this way. The theorem for unitary representations in [5] is another corollary, in a more general form that allows all types of representations and extends the class of approximating states. A further corollary shows that, for certain representations, the de Finetti approximation becomes exact. Finally, using representations of the Heisenberg group, we obtain the de Finetti theorem recently proved for coherent states of infinite-dimensional systems [17] together with a new exponential version of it.

Having laid claim to an all-embracing theorem, we must acknowledge that there are some special de Finetti results that do not lie within our conspectus. One is the theorem for unitarily-invariant, symmetric mixed states proved in [5]. Though the theorem in the present paper applies to pure states, it can be extended by symmetric purification to mixed states. However, this does not yield the special form of the approximation given in [5]. Nor does our theorem

apply to the unitarily-invariant states considered in [18]; but here unitary invariance takes the place of symmetry, and the result is of a very different character from all other de Finetti theorems.

II. BACKGROUND ON REPRESENTATIONS

We first recall some basic facts about representations (for more details, see e.g., [19, 20, 21, 22, 23, 24]). Section II A deals with the group of special unitary matrices, whereas Section II B discusses the Heisenberg group. Throughout, we will only consider representations on complex Hilbert spaces $\mathcal{H}$ which are unitary, i.e. the group elements act as unitary operators on $\mathcal{H}$.

A. The special unitary group $\mathcal{SU}(d)$ and its representations

Let $\mathcal{SU}(d)$ denote the group of unitary matrices on $\mathbb{C}^d$ with determinant 1, and let $|1\rangle, \ldots |d\rangle$ be a chosen basis for $\mathbb{C}^d$. Let H denote the subgroup of diagonal matrices with respect to this basis. Let $\mathbf{w} = (w_1, \ldots, w_d)$ be a vector of integers with $w_i \geq 0$, and let $\mathcal{R}$ be a representation of $\mathcal{SU}(d)$. Then a *weight vector* of $\mathcal{R}$ with weight $\mathbf{w}$ is a vector $|\mathbf{w}\rangle \in \mathcal{R}$ satisfying $h|\mathbf{w}\rangle = \prod h_i^{w_i}|\mathbf{w}\rangle$, where h is a diagonal matrix in H with diagonal entries $h_1, \ldots, h_d$. For instance, the weight $(n, 0, \ldots, 0)$ corresponds to the weight vector $|1\rangle^{\otimes n}$ in the symmetric representation $\mathsf{Sym}^n(\mathbb{C}^d) \subset (\mathbb{C}^d)^{\otimes n}$. In general, there may be several weight vectors for a given weight. In the sequel, we will write $|\mathbf{w}\rangle$ whenever we refer to a normalised vector of weight $\mathbf{w}$. These should not be confused with basis vectors $|i\rangle$, which are labelled by $i \in \{1, \ldots, d\}$. We will also distinguish different vectors of the same weight $\mathbf{w}$ by superscripts $|\mathbf{w}\rangle^i$.

A special part is played by irreducible representations of $\mathcal{SU}(d)$, i.e. representations that cannot be written as a direct sum of two subrepresentations. Because of the correspondence between Lie groups and their Lie algebras (see Section IV B), the study of these representations is equivalent to the study of irreducible representations of the (real) Lie algebra $\mathfrak{su}_d$ consisting of anti-hermitian traceless matrices. Irreducible representations of $\mathfrak{su}_d$ are in turn in one-to-one-correspondence with irreducible representations of its complexification $\mathfrak{sl}_d(\mathbb{C})$, which consists of all traceless matrices. Thus it suffices to consider the Lie algebra $\mathfrak{sl}_d(\mathbb{C})$ instead of the Lie group $\mathcal{SU}(d)$, which often turns out to be simpler.

An irreducible representation of $\mathcal{SU}(d)$ (or equivalently $\mathfrak{sl}_d(\mathbb{C})$) has a unique highest weight (we will discuss the ordering of weights below), and two irreducible representations are equivalent if and only if they have the same highest weight. The weight space corresponding to the highest weight is one-dimensional. We write $\mathcal{R}_\lambda$ for the irreducible representation of $\mathcal{SU}(d)$ with highest weight λ (Greek letters are conventionally used), and the normalised highest weight vector of $\mathcal{R}_\lambda$ will be denoted by $|\lambda\rangle$. The highest weights have the property that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d = 0$. We will often omit λ_i that are zero, for instance writing the highest weight of the symmetric representation $\mathsf{Sym}^n(\mathbb{C}^d)$ as (n) rather than $(n, \underbrace{0, \ldots, 0}_{d-1})$. We will also sometimes write $\mathcal{R}_\lambda$ for an irreducible representations with highest weight $\lambda = (\lambda_1, \ldots, \lambda_d)$ where $\lambda_d \neq 0$; note that this is equivalent to an irreducible representation $\mathcal{R}_{\lambda'}$ where λ' is obtained from λ by subtracting $\lambda_d \cdot (1, \ldots, 1)$.

There is a special set of weights, of the form $\alpha_i = (0, \ldots, 0, 1, -1, 0, \ldots, 0)$ where the i-th entry is 1. These are a set of so-called *simple roots* of the Lie algebra $\mathfrak{sl}_d(\mathbb{C})$. Any weight occurring in the irreducible representation $\mathcal{R}_\lambda$ can be obtained by subtracting integer combinations of the α_i from the highest weight vector. This allows us to define a notion of the *height* of a weight $\mathbf{w}$ in the irreducible representation $\mathcal{R}_\lambda$ as $ht_\lambda^\downarrow(\mathbf{w}) := \max_i |n_i|$, where $\mathbf{w} = \lambda - \sum n_i \alpha_i$. More generally, we will extend this definition to the space $\mathbb{R}^n$ of all n-tuples $\mathbf{w} = (w_1, \ldots, w_n)$ (not necessarily weights); this is well-defined because the n_i for a given $\mathbf{w}$ are uniquely determined, as the simple roots are a basis of $\mathbb{R}^n$. There is an analogous formulation using the lowest weight vector λ_* of $\mathcal{R}_\lambda$. Any weight in $\mathcal{R}_\lambda$ can be obtained by adding integer combinations of the α_i to λ_*, so we can define $ht_\lambda^\uparrow(\mathbf{w}) := \max_i |m_i|$, where $\mathbf{w} = \lambda_* + \sum m_i \alpha_i$.

There is a partial ordering on the weights, with $\mathbf{w} \preceq \mathbf{w}'$ if the difference $\mathbf{w}' - \mathbf{w}$ is a combination of the simple roots α_i with nonnegative coefficients, which is equivalent to the condition $\sum_{i=1}^\ell w_i \leq \sum_{i=1}^\ell w_i'$ for all $i = 1, \ldots, \ell$. The set of weights in a representation $\mathcal{R}$ will be denoted by $\mathcal{W}_\mathcal{R}$. Given some subset $\mathcal{W} \subset \mathcal{W}_\mathcal{R}$, the subspace of $\mathcal{R}$ that it generates will be denoted by $\mathcal{R}^\mathcal{W}$. For a finite-dimensional representation, the weight vectors generate the whole representation, so we have $\mathcal{R}^{\mathcal{W}_\mathcal{R}} = \mathcal{R}$.

Given two representations $\mathcal{A}$ and $\mathcal{B}$, not necessarily irreducible, we can define the tensor product representation $\mathcal{A} \otimes \mathcal{B}$ on which any $\mathsf{g} \in \mathcal{SU}(d)$ acts by the tensor product of its actions on $\mathcal{A}$ and $\mathcal{B}$. Since the weight vectors of $\mathcal{A}$ and $\mathcal{B}$ span each representation, a weight vector $|\mathbf{w}\rangle_{\mathcal{A} \otimes \mathcal{B}}$ of the tensor product can be written as a sum of products

of weight vectors as

$$|\mathbf{w}\rangle_{\mathcal{A}\otimes\mathcal{B}} = \sum_{\mathbf{w}_{\mathcal{A}},\mathbf{w}_{\mathcal{B}},i,j} \gamma_{ij}^{\mathbf{w}_{\mathcal{A}}\mathbf{w}_{\mathcal{B}}} |\mathbf{w}_{\mathcal{A}}\rangle_{\mathcal{A}}^{i} \otimes |\mathbf{w}_{\mathcal{B}}\rangle_{\mathcal{B}}^{j} , \tag{1}$$

and it follows from the definition that the weights occuring with non-zero coefficent $\gamma_{ij}^{\mathbf{w}_{\mathcal{A}}\mathbf{w}_{\mathcal{B}}}$ satisfy

$$\mathbf{w}_{\mathcal{A}} + \mathbf{w}_{\mathcal{B}} = \mathbf{w} . \tag{2}$$

The tensor product of two irreducible representations $\mathcal{R}_{\mu}$ and $\mathcal{R}_{\nu}$ is in general reducible and decomposes as

$$\mathcal{R}_{\mu} \otimes \mathcal{R}_{\nu} \cong \bigoplus_{\lambda} c_{\mu\nu}^{\lambda} \mathcal{R}_{\lambda}, \tag{3}$$

where the multiplicities $c_{\mu\nu}^{\lambda}$ are the *Littlewood-Richardson coefficients*. The subrepresentations on the right-hand side include $\mathcal{R}_{\mu+\nu}$, which occurs with multiplicity 1.

For $d = 2$, the multiplicities $c_{\mu\nu}^{\lambda}$ are all 1 and furthermore there is a unique weight vector for each weight. Thus equation (1) can be written more simply as

$$|\mathbf{w}\rangle_{\lambda} = \sum_{\mathbf{w}'} a_{\mathbf{w},\mathbf{w}'} |\mathbf{w}'\rangle_{\mu} \otimes |\mathbf{w} - \mathbf{w}'\rangle_{\nu} , \tag{4}$$

and the coefficients $a_{\mathbf{w},\mathbf{w}'}$ define the inclusion map $\mathcal{R}_{\lambda} \to \mathcal{R}_{\mu} \otimes \mathcal{R}_{\nu}$ in (3). These are the *Clebsch-Gordan coefficients*, and there seems to be no consensus on notation for them. Here we write $a_{\mathbf{w},\mathbf{w}'}$ as $\langle j_1 m_1 j_2 m_2 | jm \rangle$, where $j_1 = (\mu_1 - \mu_2)/2$, $j_2 = (\nu_1 - \nu_2)/2$, $j = (\lambda_1 - \lambda_2)/2$, $m = (w_1 - w_2)/2$, $m_1 = (w_1' - w_2')/2$, and $m_2 = m - m_1$. In other words $\{|jm\rangle\}_{m=-j}^{j}$ is a basis of weight vectors of $\mathcal{R}_{(2j)}$, where $|jm\rangle$ has weight $(m, -m)$. Thus (4) can equivalently be written as

$$|jm\rangle = \sum_{m_1 + m_2 = m} \langle j_1 m_1 j_2 m_2 | jm \rangle |j_1 m_1\rangle \otimes |j_2 m_2\rangle . \tag{5}$$

Here $j_1, j_2 \in \{0, 1/2, 1, \ldots, \}$, and $\mathcal{R}_{(2j)} \subset \mathcal{R}_{(2j_1)} \otimes \mathcal{R}_{(2j_2)}$ with multiplicity 1 for all $j \in \{|j_1 - j_2|, |j_1 - j_2| + 1/2, \ldots, j_1 + j_2\}$.

B.   Representations of the Heisenberg group

The Heisenberg group H is $\mathbb{C} \times \mathbb{R}$ with multiplication

$$(\alpha; t)(\beta; t') = (\alpha + \beta; t + t' + \Im(\alpha\bar{\beta})) ,$$

where $\Im(\gamma)$ denotes the imaginary part of $\gamma \in \mathbb{C}$. The centre of H is $Z = \{(0; t) \mid t \in \mathbb{R}\}$, and the quotient group H/Z is isomorphic to the abelian group $\mathbb{R}^2$. Elements of H/Z will be denoted simply by a complex number $\alpha \in \mathbb{C}$. The irreducible infinite-dimensional unitary representations of H are determined by a real number $\lambda \neq 0$. The action of Z in the irreducible representation $\mathcal{H}_{\lambda}$ is then given by

$$(0; t)|\psi\rangle = e^{i\lambda t}|\psi\rangle \qquad \text{for every } |\psi\rangle \in \mathcal{H}_{\lambda} . \tag{6}$$

The representation $\mathcal{H}_1$ can be obtained from an *annihilation operator* a and its conjugate *creation operator* $a^{\dagger}$ acting on some (infinite-dimensional) Hilbert space $\mathcal{H}$ satisfying the canonical commutation relation $[a, a^{\dagger}] = 1$ and a "vacuum state" $|0\rangle \in \mathcal{H}$ satisfying $a|0\rangle = 0$. The space $\mathcal{H}_1$ is then spanned by the orthonormal vectors $\{|n\rangle\}_{n \in \mathbb{N}_0}$, where $|n\rangle := \frac{(a^{\dagger})^n}{\sqrt{n!}}|0\rangle$ for every nonnegative integer n. To define the action of H on this space, we introduce the *displacement operators*

$$D(\alpha) = \exp(\alpha a^{\dagger} - \bar{\alpha}a) = e^{-|\alpha|^2/2} \exp(\alpha a^{\dagger}) \exp(-\bar{\alpha}a) \tag{7}$$

for $\alpha \in \mathbb{C}$. (We will sometimes write $D_a(\alpha)$ to clarify what the operators a and $a^{\dagger}$ in this definition are.) It is straightforward to verify that these operators satisfy the relation

$$D(\alpha)D(\beta) = e^{i\Im(\alpha\bar{\beta})}D(\alpha + \beta) . \tag{8}$$

This identity implies that an action of H on $\mathcal{H}_1$ is defined by

$$(\alpha; t)|\psi\rangle := e^{it}D(\alpha)|\psi\rangle \qquad \text{for all } |\psi\rangle \in \mathcal{H}_1 . \tag{9}$$

This completes the description of $\mathcal{H}_1$. Note that the creation and annihilation operators satisfy

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \qquad \text{and} \qquad a|n\rangle = \sqrt{n}|n-1\rangle , \tag{10}$$

and the state $|n\rangle$ is an eigenstate of the number operator $a^\dagger a$ with eigenvalue n. The subspace spanned by a set $\mathcal{N}$ of number states plays the same role as the subspaces spanned by weights $\mathcal{W}$ in the case of representations of $\mathcal{SU}(d)$. The number state $|0\rangle$ is analogous to a lowest weight vector, and the corresponding displaced states $D(\alpha)|0\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^\infty \frac{\alpha^n}{\sqrt{n!}}|n\rangle$ for $\alpha \in \mathbb{C}$ are commonly called *coherent states*.

To give an explicit construction of $\mathcal{H}_\lambda$ for $\lambda \neq 0$, we choose the space $\mathcal{H}_\lambda$ as the span of $\{|n\rangle \mid n \in \mathbb{N}_0\}$ and define

$$(\alpha; t)|\psi\rangle := e^{i\lambda t}D(\sqrt{\lambda}\alpha)|\psi\rangle \qquad \text{for all } |\psi\rangle \in \mathcal{H}_\lambda , \tag{11}$$

where the states $|n\rangle$ and the operators $D(\alpha)$ are defined as before. Again using (8), it is straightforward to check that this defines a representation, and it is irreducible as a consequence of the fact that $\mathcal{H}_1$ is irreducible.

A few subtleties arise when integrating over the group. The quotient group H/Z is a *unimodular group*, meaning that there is a left- and right-invariant Haar measure μ which assigns finite mass $\mu(K)$ to every compact set $K \subset \mathsf{H}/Z$. This Haar measure is given by the standard Lebesque measure on $\mathbb{C}^2$; unlike the case of a compact group $\mathcal{K}$, this can not be normalised so that $\mu(\mathcal{K}) = 1$. Thus the Haar measure on a unimodular group is only fixed up to a constant. For reasons that will become clearer below, we will choose the measure on H/Z as $d\mu(\alpha) = \frac{1}{\pi}d\Re(\alpha)d\Im(\alpha)$ on H/Z.

A version of Schur's lemma which applies to general unimodular groups $\mathcal{G}$ involves the notion of a *square-integrable representation* on a Hilbert space $\mathcal{H}$. Such a representation has the property that

$$\int_{\mathcal{G}} |\langle\psi|\mathbf{g}|\varphi\rangle|^2 d\mu(\mathbf{g}) < \infty \qquad \text{for all } |\psi\rangle, |\varphi\rangle \in \mathcal{H} .$$

It is known [25, p. 439, Proposition 29] that if $\mathcal{G}$ is unimodular and the representation $\mathcal{H}$ irreducible and square-integrable, then there is a constant $d_\mathcal{H}$ such that

$$\int_{\mathcal{G}} \langle\beta|\mathbf{g}^\dagger|\alpha\rangle\langle\gamma|\mathbf{g}|\delta\rangle d\mu(\mathbf{g}) = \frac{\langle\gamma|\alpha\rangle\langle\beta|\delta\rangle}{d_\mathcal{H}} \tag{12}$$

for all $|\alpha\rangle, |\beta\rangle, |\gamma\rangle, |\delta\rangle \in \mathcal{H}$. The quantity $d_\mathcal{H}$ is called the *formal degree* of $\mathcal{H}$. Note that it depends on the intitial choice of the Haar measure μ. For a finite-dimensional representation $\mathcal{H}$ and a compact group $\mathcal{G}$, the formal degree $d_\mathcal{H}$ is equal to the dimension of $\mathcal{H}$ if the Haar measure is normalised so that $\int_{\mathcal{G}} d\mu(\mathbf{g}) = 1$.

Note that the representation $\mathcal{H}_\lambda$ of the Heisenberg group H defined by (11) gives an irreducible representation of the unimodular quotient group H/Z. For the chosen Haar measure, the formal degree of this representation can easily be computed using the fact that $|\langle 0|D(\alpha)|0\rangle|^2 = e^{-|\alpha|^2}$. One finds $d_{\mathcal{H}_\lambda} = \lambda$. This is all we need for the Heisenberg group; we refer the reader to the literature (e.g., [19]) for more details.

III.   THE MAIN THEOREM

We consider irreducible representations $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ of a general unimodular group $\mathcal{G}$ satisfying $\mathcal{C} \subset \mathcal{A} \otimes \mathcal{B}$. The trace $\mathrm{tr}_\mathcal{B}|\Psi\rangle\langle\Psi|$ of a state $|\Psi\rangle \in \mathcal{C}$ is now well-defined, and our aim is to show that this trace can be approximated by a convex sum of a special class of states given by:

Definition III.1. *Given a subset $\mathcal{X}$ of $\mathcal{A}$, an $\mathcal{X}$-state is a state of the form $\mathbf{g}|\varphi\rangle$, where $\mathbf{g} \in \mathcal{G}$ and $|\varphi\rangle \in \mathcal{X}$.*

Now suppose we have a triple of representations $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C} \subset \mathcal{A} \otimes \mathcal{B}$ with $\mathcal{B}$ and $\mathcal{C}$ square-integrable (see Section II B).

Definition III.2. *Let $|\psi\rangle \in \mathcal{B}$ be arbitary, and let*

$$\delta_{|\psi\rangle}(\mathcal{X}) = \frac{d_\mathcal{B}}{d_\mathcal{C}}\mathrm{tr}[P_\mathcal{C}(P_\mathcal{X} \otimes |\psi\rangle\langle\psi|)] ,$$

where $P_\mathcal{C}$ and $P_\mathcal{X}$ are the projectors onto the subspaces $\mathcal{C} \subset \mathcal{A} \otimes \mathcal{B}$ and $\mathcal{X} \subset \mathcal{A}$, and $d_\mathcal{B}$, $d_\mathcal{C}$ are the formal degrees of $\mathcal{B}$ and $\mathcal{C}$, respectively. Finally, let

$$\delta(\mathcal{X}) = \sup_{|\psi\rangle \in \mathcal{B}} \delta_{|\psi\rangle}(\mathcal{X}),$$

where the maximisation is over all normalised pure states on $\mathcal{B}$.

We will discuss two basic properties of this definition below (Lemma III.5). The following theorem, which is our main result, shows that $\delta(\mathcal{X})$ is a useful measure for the error in the de Finetti approximation.

Theorem III.3 (Approximation by $\mathcal{X}$-states). *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C} \subset \mathcal{A} \otimes \mathcal{B}$ be irreducible representations of a unimodular group $\mathcal{G}$, where $\mathcal{B}$ and $\mathcal{C}$ are square-integrable. Let $\mathcal{X}$ be a finite-dimensional subspace of $\mathcal{A}$. Then $\mathrm{tr}_\mathcal{B}|\Psi\rangle\langle\Psi|$ for every $|\Psi\rangle \in \mathcal{C}$ can be approximated by a convex combination of $\mathcal{X}$-states with error $2\sqrt{1 - \delta(\mathcal{X})}$. That is, there is a probability measure m on $\mathcal{G}$ and states $|\chi_{(\mathsf{g})}\rangle \in \mathsf{g}\mathcal{X}$ for $\mathsf{g} \in \mathcal{G}$ such that*

$$\left\| \mathrm{tr}_\mathcal{B}|\Psi\rangle\langle\Psi| - \int |\chi_{(\mathsf{g})}\rangle\langle\chi_{(\mathsf{g})}| dm(\mathsf{g}) \right\| \leq 2\sqrt{1 - \delta(\mathcal{X})} . \tag{13}$$

In this expression, the trace norm *is defined as $\|A\| = \frac{1}{2}\mathrm{tr}(\sqrt{A^\dagger A})$ for any operator A.*

Remark III.4. *In certain cases, it is possible to improve the bound (13) by adapting the proof technique introduced in [5] and applied in [3, 7, 17]. For example, if the representation $\mathcal{C}$ appears with multiplicity 1 in the tensor product $\mathcal{A} \otimes \mathcal{B}$, then the rhs. of (13) can be replaced by $2(1 - \delta(\mathcal{X}))$. Also, if $\mathcal{X}$ has an orthonormal basis $\{|\varphi_1\rangle, \ldots, |\varphi_\ell\rangle\}$ and there is a vector $|\psi\rangle \in \mathcal{B}$ such that $|\varphi_i\rangle \otimes |\psi\rangle \in \mathcal{C}$ for all i, then we can replace the rhs. of (13) by $2(1 - \delta_{|\psi\rangle}(\mathcal{X}))$. This improvement is made possible by the fact that twirling the operator $P_\mathcal{X} \otimes |\psi\rangle\langle\psi|$, where $P_\mathcal{X}$ is the projector onto $\mathcal{X}$, gives an operator which is proportional to the identity on $\mathcal{C}$. We do not elaborate on this improvement any further, as it is a straightforward consequence of the technique presented in [5].*

Proof. Let $|\psi\rangle \in \mathcal{B}$ be arbitrary and define $|\psi_\mathsf{g}\rangle = \mathsf{g}|\psi\rangle$ for $\mathsf{g} \in \mathcal{G}$. Schur's Lemma in the form (12) tells us that the operator $d_\mathcal{B} \int |\psi_\mathsf{g}\rangle\langle\psi_\mathsf{g}| d\mathsf{g}$ acts as the identity on $\mathcal{B}$ where integration is over the normalised Haar measure on $\mathcal{G}$. In particular, for $|\Psi\rangle \in \mathcal{C}$, we have

$$\begin{aligned}
\mathrm{tr}_\mathcal{B}|\Psi\rangle\langle\Psi| &= d_\mathcal{B} \, \mathrm{tr}_\mathcal{B} \left[\int (P_\mathcal{A} \otimes |\psi_\mathsf{g}\rangle\langle\psi_\mathsf{g}|) |\Psi\rangle\langle\Psi| d\mathsf{g} \right] \\
&= \int |\tilde{\chi}_{(\mathsf{g})}\rangle\langle\tilde{\chi}_{(\mathsf{g})}| dm(\mathsf{g}) ,
\end{aligned} \tag{14}$$

where the normalised states $|\tilde{\chi}_{(\mathsf{g})}\rangle$ and the probability measure m on $\mathcal{G}$ are defined by

$$|\tilde{\chi}_{(\mathsf{g})}\rangle\langle\tilde{\chi}_{(\mathsf{g})}| dm(\mathsf{g}) = d_\mathcal{B} \, \mathrm{tr}_\mathcal{B}\big((P_\mathcal{A} \otimes |\psi_\mathsf{g}\rangle\langle\psi_\mathsf{g}|) |\Psi\rangle\langle\Psi|\big) d\mathsf{g} \tag{15}$$

and where $P_\mathcal{A}$ is the identity on $\mathcal{A}$. (To see that $|\tilde{\chi}_\mathsf{g}\rangle$ is indeed a pure state, observe that the rhs. of (15) is of the form $d_\mathcal{B}\mathrm{tr}_\mathcal{B}|\Phi_{(\mathsf{g})}\rangle\langle\Phi_{(\mathsf{g})}|$, where $|\Phi_{(\mathsf{g})}\rangle = (P_\mathcal{A} \otimes |\psi_\mathsf{g}\rangle\langle\psi_\mathsf{g}|)|\Psi\rangle$ is a product of pure states.)

Let P be the projector onto $\mathcal{X}$, and let $P_\mathsf{g} = \mathsf{g}P\mathsf{g}^\dagger$ be the projector onto $\mathsf{g}\mathcal{X}$. We claim that $|\tilde{\chi}_{(\mathsf{g})}\rangle$ is on average close to the projected state

$$|\chi_{(\mathsf{g})}\rangle := \frac{P_\mathsf{g}|\tilde{\chi}_{(\mathsf{g})}\rangle}{\langle\tilde{\chi}_{(\mathsf{g})}|P_\mathsf{g}|\tilde{\chi}_{(\mathsf{g})}\rangle} .$$

For this purpose, we use the gentle measurement lemma (see e.g., [26, 27]), which implies with the triangle inequality that

$$\left\| |\tilde{\chi}_{(\mathsf{g})}\rangle\langle\tilde{\chi}_{(\mathsf{g})}| - |\chi_{(\mathsf{g})}\rangle\langle\chi_{(\mathsf{g})}| \right\| \leq 2\sqrt{1 - \langle\chi_{(\mathsf{g})}|P_\mathsf{g}|\chi_{(\mathsf{g})}\rangle} .$$

We thus obtain

$$\begin{aligned}
\left\| \mathrm{tr}_\mathcal{B}|\Psi\rangle\langle\Psi| - \int |\chi_{(\mathsf{g})}\rangle\langle\chi_{(\mathsf{g})}| dm(\mathsf{g}) \right\| &\leq 2 \int \sqrt{1 - \langle\tilde{\chi}_{(\mathsf{g})}|P_\mathsf{g}|\tilde{\chi}_{(\mathsf{g})}\rangle} dm(\mathsf{g}) \\
&\leq 2\sqrt{1 - \kappa} ,
\end{aligned} \tag{16}$$

using the convexity of the trace distance and the square root, where

$$\kappa = \int \langle \tilde{\chi}_{(g)} | P_g | \tilde{\chi}_{(g)} \rangle dm(g) \ . \tag{17}$$

Because of (15) and the cyclicity of the trace, we have

$$\langle \tilde{\chi}_{(g)} | P_g | \tilde{\chi}_{(g)} \rangle dm(g) = d_{\mathcal{B}} \ \mathsf{tr}\big((P_g \otimes |\psi_g\rangle\langle\psi_g|)|\Psi\rangle\langle\Psi|\big) dg$$
$$= d_{\mathcal{B}} \ \mathsf{tr}\big((P \otimes |\psi\rangle\langle\psi|)|\Psi_{g^\dagger}\rangle\langle\Psi_{g^\dagger}|\big) dg \ ,$$

and therefore by the linearity and (again) the cyclicity of the trace

$$\kappa = d_{\mathcal{B}} \ \mathsf{tr}\big((P \otimes |\psi\rangle\langle\psi|) \int |\Psi_{g^\dagger}\rangle\langle\Psi_{g^\dagger}| dg\big)$$
$$= d_{\mathcal{B}} \ \int \langle\Psi|gP_{\mathcal{C}}(P \otimes |\psi\rangle\langle\psi|)P_C g^\dagger|\Psi\rangle dg \ .$$

Schur's Lemma (cf. (12)) immediately implies that $\kappa = \delta_{|\psi\rangle}(\mathcal{X})$. Since $|\psi\rangle$ was arbitrary, we may take $\kappa = \delta(\mathcal{X})$ in (16), which concludes the proof. $\qquad\square$

For later reference, we point out the following properties of the quantity δ.

Lemma III.5. *Let $\mathcal{C} \subset \mathcal{A} \otimes \mathcal{B}$, and let δ be as in Definition III.2. Then*

(i) $\delta(\mathcal{X}) \leq \delta(\mathcal{X}')$ if $\mathcal{X} \subset \mathcal{X}'$.

(ii) If $\mathcal{C}$ is finite-dimensional, $\delta(\mathcal{X}) \leq 1$ with equality if $\mathcal{X} = \mathcal{A}$.

Proof. Inequality (i) follows directly from the definition of $\delta(\mathcal{X})$. Using the fact that the projectors $P_{\mathcal{C}}$ and $P_{\mathcal{A}}$ are invariant, i.e., $gP_{\mathcal{C}}g^\dagger = P_{\mathcal{C}}$ and $gP_{\mathcal{A}}g^\dagger = P_{\mathcal{A}}$, we obtain using the cyclicity of the trace

$$\mathsf{tr}[P_{\mathcal{C}}(P_{\mathcal{A}} \otimes |\psi\rangle\langle\psi|)] = \mathsf{tr}[g^\dagger P_{\mathcal{C}} g(P_{\mathcal{A}} \otimes |\psi\rangle\langle\psi|)]$$
$$= \mathsf{tr}[P_{\mathcal{C}}(gP_{\mathcal{A}}g^\dagger \otimes |\psi_g\rangle\langle\psi_g|)]$$
$$= \mathsf{tr}[P_{\mathcal{C}}(P_{\mathcal{A}} \otimes |\psi_g\rangle\langle\psi_g|)]$$

for all $|\psi\rangle \in \mathcal{B}$ and $g \in \mathcal{G}$. By linearity and Schur's lemma, we get

$$\mathsf{tr}(P_{\mathcal{C}}(P_{\mathcal{A}} \otimes |\psi\rangle\langle\psi|)) = \frac{\mathsf{tr}(P_{\mathcal{C}}(P_{\mathcal{A}} \otimes P_{\mathcal{B}}))}{d_{\mathcal{B}}} \tag{18}$$

for any $|\psi\rangle$, which implies $\delta(\mathcal{A}) = 1$ by the definition of $\delta(\mathcal{X})$. This proves (ii) when combined with (i). $\qquad\square$

IV. EXACT EXPRESSIONS

As a first application of the our theorem, we gives some examples where $\delta(\mathcal{X}) = 1$. This implies that the approximation by $\mathcal{X}$-states in equation (13) is exact. The subspaces $\mathcal{X}$ in question are direct sums of weight spaces for a particular range of weight values; we call these $\mathcal{W}^r$-states (see definition below). In this section, therefore, we confine attention to semi-simple Lie groups, that have a weight space structure.

A. $\mathcal{W}^r$-states for $\mathcal{SU}(d)$

Recall first from Section II A that $ht_\lambda^\downarrow(\mathbf{w})$ is defined as $\max(n_i)$, where $\lambda - \mathbf{w} = \sum_i n_i \alpha_i$ and α_i are the simple roots. Similarly, $ht_\lambda^\uparrow(\mathbf{w})$ is defined as $\max(m_i)$, where $\mathbf{w} - \lambda_* = \sum_i m_i \alpha_i$, and λ_* is the lowest weight of $\mathcal{R}_\lambda$. We will partition the set of weights according to their height as follows.

Definition IV.1. *Let $\mathcal{W}_\lambda$ be the set of weights occuring in the irreducible representation $\mathcal{R}_\lambda$ of $\mathcal{SU}(d)$. We define $\mathcal{W}_\lambda^{\downarrow r}$ for $r \geq 0$ to be the set of weights $\mathbf{w} \in \mathcal{W}_\lambda$ satisfying $ht_\lambda^\downarrow(\mathbf{w}) \leq r$. Similarly, $\mathcal{W}_\lambda^{\uparrow r}$ is the set of weights $\mathbf{w} \in \mathcal{W}_\lambda$ satisfying $ht_\lambda^\uparrow(\mathbf{w}) \leq r$. A $\mathcal{W}_\lambda^r$-state is a state of the form $g|\varphi\rangle$, where $g \in \mathcal{SU}(d)$ and $|\varphi\rangle$ is supported on the weight space corresponding to $\mathcal{W}_\lambda^{\uparrow r}$.*

Thus $\mathcal{W}_\lambda^{\downarrow r}$ consists of weights that lie within distance r of the highest weight, and $\mathcal{W}_\lambda^{\uparrow r}$ of those within distance r of the lowest weight. Note that the same set of states is generated by $\mathcal{W}_\lambda^{\uparrow r}$ and $\mathcal{W}_\lambda^{\downarrow r}$. This is because $\mathcal{SU}(d)$ includes any permutation of the basis vectors. Thus $\mathcal{W}_\lambda^r$ could equivalently have been defined in terms of the high weights $\mathcal{W}_\lambda^{\downarrow r}$.

We now consider states that can be expressed *exactly* as a convex sum of $\mathcal{W}^r$-states (we will often omit the representation label when clear from the context).

Corollary IV.2 (Exact expression by $\mathcal{W}^r$-states)**.** *Let $\mathcal{R}_\mu$ and $\mathcal{R}_\nu$ be irreducible representations of $\mathcal{SU}(d)$, and let $\mathcal{R}_\lambda \subset \mathcal{R}_\mu \otimes \mathcal{R}_\nu$ be a subrepresentation of their tensor product. Then the partial trace $\mathrm{tr}_{\mathcal{R}_\nu}|\Psi\rangle\langle\Psi|$ of every state $|\Psi\rangle \in \mathcal{R}_\lambda$ is a convex sum of pure $\mathcal{W}_\mu^r$-states on $\mathcal{R}_\mu$, with $r = ht_\mu^\uparrow(\lambda - \nu)$. That is, there is a probability measure m on $\mathcal{SU}(d)$ and states $|\chi_{(\mathsf{g})}\rangle = \mathsf{g}|\varphi_{(\mathsf{g})}\rangle$ with $\mathsf{g} \in \mathcal{SU}(d)$ and $|\varphi_{(\mathsf{g})}\rangle$ supported on the weight space $\mathcal{R}_\mu^{\mathcal{W}^{\uparrow r}}$ such that*

$$\mathrm{tr}_{\mathcal{R}_\nu}|\Psi\rangle\langle\Psi| = \int |\chi_{(\mathsf{g})}\rangle\langle\chi_{(\mathsf{g})}|dm(\mathsf{g}) \ . \tag{19}$$

Proof. We will derive this from Theorem III.3 by showing that $\delta_{|\nu\rangle}(\mathcal{W}^{\uparrow r}) = 1$ for $r = ht_\mu^\uparrow(\lambda - \nu)$, where $|\nu\rangle$ is the highest weight vector in $\mathcal{R}_\nu$. Let us define the set of weights

$$\Lambda^\uparrow = \{\mathbf{w} \in \mathcal{W}_\mu| \ \mathbf{w} + \nu \in \mathcal{W}_\lambda\} \ . \tag{20}$$

Note that if $\mathbf{w} \in \Lambda^\uparrow$, then $\mathbf{w} + \nu \preceq \lambda$ since λ is the highest weight, and this is equivalent to $\mathbf{w} \preceq \lambda - \nu$. We conclude that $\Lambda^\uparrow \subset \mathcal{W}^{\uparrow r}$, for $r = ht_\mu^\uparrow(\lambda - \nu)$ since $ht_\mu^\uparrow(\mathbf{w}) \le ht_\mu^\uparrow(\mathbf{w}')$ whenever $\mathbf{w} \preceq \mathbf{w}'$. By Lemma III.5 (i), it thus suffices to show that

$$\delta_{|\nu\rangle}(\Lambda^\uparrow) = 1 \ . \tag{21}$$

Suppose first that $|\Psi\rangle \in \mathcal{R}_\lambda$ is a weight vector of weight $\mathbf{w}_\lambda \in \mathcal{W}_{\mathcal{R}_\lambda}$. Then $|\Psi\rangle$ has the form

$$|\Psi\rangle = \sum_{\mathbf{w}_\mu,\mathbf{w}_\nu,j,k} \gamma_{jk}^{\mathbf{w}_\mu\mathbf{w}_\nu}|\mathbf{w}_\mu\rangle^j|\mathbf{w}_\nu\rangle^k$$

where $|\mathbf{w}_\mu\rangle^j$ are weights vectors in $\mathcal{R}_\mu$, $|\mathbf{w}_\nu\rangle^k$ are weight vectors in $\mathcal{R}_\nu$, $|\mathbf{w}_\nu\rangle^0 = |\nu\rangle$ is the highest weight vector in $\mathcal{R}_\nu$, and $\mathbf{w}_\mu + \mathbf{w}_\nu = \mathbf{w}_\lambda$ for all nonzero terms in the sum, by (1) and (2). It is straightforward to check that

$$(P \otimes |\nu\rangle\langle\nu|)|\Psi\rangle = (P_\mu \otimes |\nu\rangle\langle\nu|)|\Psi\rangle \ , \tag{22}$$

where P is the projector onto the weightspace $\mathcal{R}_\mu^{\Lambda^\uparrow}$, and where P_μ is the identity on $\mathcal{R}_\mu$. But (22) holds for all $|\Psi\rangle \in \mathcal{R}_\lambda$, since $|\Psi\rangle$ can be expanded in terms of weight vectors. This implies that

$$\mathrm{tr}\big(|\Psi\rangle\langle\Psi|(P \otimes |\nu\rangle\langle\nu|)\big) = \mathrm{tr}\big(|\Psi\rangle\langle\Psi|(P_\mu \otimes |\nu\rangle\langle\nu|)\big) \ .$$

for all $|\Psi\rangle \in \mathcal{R}_\lambda$. In particular, we have

$$\delta_{|\nu\rangle}(\Lambda^\uparrow) = \frac{d_\nu}{d_\lambda}\mathrm{tr}\big(P_\lambda(P_\mu \otimes |\nu\rangle\langle\nu|)\big)$$

$$= \frac{1}{d_\lambda}\mathrm{tr}\big(P_\lambda(P_\mu \otimes P_\nu)\big) \ ,$$

where we used the invariance of P_λ, the cyclicity of the trace and Schur's Lemma. Because $\mathcal{R}_\lambda \subset \mathcal{R}_\mu \otimes \mathcal{R}_\nu$, this gives (21), as desired. $\qquad\blacksquare$

The following lemma shows that it is natural to bound $\Lambda^\uparrow$ by the set of weights $\mathcal{W}^{\uparrow r}$:

Lemma IV.3. *The set $\Lambda^\uparrow$ (cf. (20)) includes the lowest weight, μ_*, of $\mathcal{R}_\mu$.*

Proof. (Compare [24], Proposition 9.72.) Write any weight vector $|\mathbf{w}_\lambda\rangle$ of $\mathcal{R}_\lambda$ as $|\mathbf{w}_\lambda\rangle = \sum_j \mu_j|\mathbf{w}_{\mu,j}\rangle|\mathbf{w}_{\nu,j}\rangle$ where the sum is over pairs of weights $(\mathbf{w}_{\mu,j}, \mathbf{w}_{\nu,j})$ of $\mathcal{R}_\mu$ and $\mathcal{R}_\nu$, respectively, satisfying $\mathbf{w}_{\mu,j} + \mathbf{w}_{\nu,j} = \mathbf{w}_\lambda$ (abusing notation, in that we allow repetitions, if necessary), and $|\mathbf{w}_{\nu,j}\rangle$ are weight vectors of $\mathcal{R}_\nu$ and the $|\mathbf{w}_{\mu,j}\rangle$ are mutually orthogonal weight vectors of $\mathcal{R}_\mu$. Suppose some weight $\mathbf{w}_{\nu,j}$ is maximal, in the sense that no $\mathbf{w}_{\nu,j'}$ with $\mathbf{w}_{\nu,j'} \succ \mathbf{w}_{\nu,j}$ occurs in the expansion. Apply the raising operator $E_{rs} \in \mathfrak{sl}_d(\mathbb{C})$, i.e. the matrix that is zero except for a 1 in the r-th row and s-th column, with $r < s$. Then, $E_{rs}(|\mathbf{w}_{\mu,j}\rangle|\mathbf{w}_{\nu,j}\rangle) = (E_{rs}|\mathbf{w}_{\mu,j}\rangle)|\mathbf{w}_{\nu,j}\rangle + |\mathbf{w}_{\mu,j}\rangle(E_{rs}|\mathbf{w}_{\nu,j}\rangle)$, and the term

$|\mathbf{w}_{\mu,j}\rangle(E_{rs}|\mathbf{w}_{\nu,j}\rangle)$ cannot be cancelled by some $|\mathbf{w}_{\mu,j'}\rangle(E_{rs}|\mathbf{w}_{\nu,j'}\rangle)$ because of the orthogonality of the $|\mathbf{w}_{\mu,j}\rangle$s, nor by some $(E_{rs}|\mathbf{w}_{\mu,j'}\rangle)|\mathbf{w}_{\nu,j'}\rangle$ because, if $E_{rs}|\mathbf{w}_{\mu,j'}\rangle$ has weight $\mathbf{w}_{\mu,j}$, then $\mathbf{w}_{\mu,j'} \prec \mathbf{w}_{\mu,j}$, and hence $\mathbf{w}_{\nu,j'} \succ \mathbf{w}_{\nu,j}$, in contradiction to the maximality of $\mathbf{w}_{\nu,j}$. Thus the term $|\mathbf{w}_{\mu,j}\rangle(E_{rs}|\mathbf{w}_{\nu,k}\rangle)$ can only vanish if killed by E_{rs}, and it can only be killed by all raising operators if it is the highest weight vector $|\nu\rangle$ ([24], Theorem 5.5). So we must eventually, after a finite number of repeated raisings, reach $|\nu\rangle$, showing that there is a weight vector of $\mathcal{R}_\lambda$ whose weight is $\mathbf{w}+\nu$ for some $\mathbf{w} \in \mathcal{W}_\mu$. Starting with this vector and applying the same procedure using lowering operators, and with the roles of μ and ν interchanged, we must reach a weight vector with weight $\mu_* + \nu$, and hence $\mu_* \in \Lambda^\uparrow$. $\qquad\square$

Because of this Lemma, it makes sense to measure the distance r from the lowest weight μ_*, as one implicitly does in defining $r = ht_\mu^\uparrow(\lambda - \nu)$. Note that, if one defines

$$\Lambda^\downarrow = \{\mathbf{w} \in \mathcal{W}_\mu | \; \mathbf{w} + \nu_* \in \mathcal{W}_\lambda\} \;,$$

and sets $r = ht_\mu^\downarrow(\lambda_* - \nu_*)$, one obtains the same value of r and the same approximating states as those given by the theorem. This is because the map $w_i \leftrightarrow w_{d-i}$ on the components of a weight $\mathbf{w}$ interchanges the two sets of definitions.

We conclude with a simple example:

Example IV.4. *Take $d = 2$, $\mu = (k)$, $\nu = (n-k)$, where $k \leq n-k$ and $\lambda = (n-\ell, \ell) \cong (n-2\ell)$. Note that $k \geq \ell$ by the Littlewood-Richardson-rule. As the lowest weight μ_* is $(0, k)$, we have*

$$\lambda - \nu = (k - \ell, \ell) = \mu_* + (k-\ell)\alpha_1 \;,$$

so $r = k - \ell$.

Rephrased in terms of angular momentum, we have $(k) = (2j_1)$, $(n-k) = 2j_2$ and $(n-2\ell) = (2j)$ and thus $ht_\mu^\uparrow(\lambda - \nu) = j_1 - j_2 + j$. In particular, if $j \approx j_2 - j_1$, then only a small number r is needed to obtain an exact expression in terms of $\mathcal{W}^r$-states (cf. Figure 2).

B. $\mathcal{W}^r$-states and exact expression for semi-simple Lie groups

The results of the preceding section can be extended, almost without change, to a simply-connected semi-simple Lie group $\mathcal{G}$. Let $\mathfrak{g}$ be the Lie algebra corresponding to $\mathcal{G}$. A representation of $\mathfrak{g}$ is a map from $\mathfrak{g}$ into $\mathfrak{gl}(\mathcal{V})$ that preserves the Lie bracket. For a simply-connected Lie group, the representations of $\mathcal{G}$ and $\mathfrak{g}$ are in one-to-one correspondence via the differential map. This allows one to deal with the algebra $\mathfrak{g}$ rather than the group $\mathcal{G}$, which is often more convenient.

Let $\mathfrak{h}$ be the Cartan subalgebra of $\mathfrak{g}$, and suppose $\mathcal{R}$ is a representation of $\mathfrak{g}$. Then a weight vector $|\mathbf{w}\rangle \in \mathcal{R}$ is a vector with the property that $h|\mathbf{w}\rangle = \mathbf{w}(h)|\mathbf{w}\rangle$ for all $h \in \mathfrak{h}$ where the *weight* $\mathbf{w} : \mathfrak{h} \to \mathbb{C}$ is a linear functional. The *adjoint representation*, $ad : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, is defined by $ad(g)(h) = [g, h]$ for all $h \in \mathfrak{g}$. The weights of the adjoint representation are called *roots*, and a subset $\{\alpha_1, \ldots, \alpha_d\}$ of these is called a set of *simple roots* if every root β can be written uniquely as $\beta = \sum_{i=1}^d n_i \alpha_i$, with integers $(n_1, \ldots, n_d)$ which are either all nonnegative or all nonpositive. A set of simple roots is a basis of the space $\mathfrak{h}^*$ of linear functionals on $\mathfrak{h}$. The set of weights is ordered by the rule that $\mathbf{w} \preceq \mathbf{w}'$ if and only if $\mathbf{w}' - \mathbf{w} = \sum_{i=1}^d n_i \alpha_i$ for nonnegative $n_i \in \mathbb{R}$. With this rule, every irreducible representation $\mathcal{R}_\lambda$ is characterized by its highest weight λ.

As with $\mathcal{SU}(d)$, every weight $\mathbf{w}$ of $\mathcal{R}_\lambda$ can be written as $\mathbf{w} = \lambda - \sum n_i \alpha_i$, and this enables us to define $ht_\lambda^\downarrow(\mathbf{w}) := \max_i |n_i|$. Similarly, we can define $ht_\lambda^\uparrow(\mathbf{w}) := \max_i |m_i|$, where $\mathbf{w} = \lambda_* + \sum m_i \alpha_i$. We can now define $\mathcal{W}^r$-states exactly as in Definition IV.1, with $\mathcal{G}$ replacing $\mathcal{SU}(d)$, and we have:

Corollary IV.5 (Exact expression by $\mathcal{W}^r$-states for Lie groups). *Let $\mathcal{R}_\mu$ and $\mathcal{R}_\nu$ be irreducible representations of a simply-connected semi-simple Lie group $\mathcal{G}$, and let $\mathcal{R}_\lambda \subset \mathcal{R}_\mu \otimes \mathcal{R}_\nu$ be a subrepresentation of their tensor product. Then the partial trace $\text{tr}_{\mathcal{R}_\nu}|\Psi\rangle\langle\Psi|$ of every state $|\Psi\rangle \in \mathcal{R}_\lambda$ is a convex sum of pure $\mathcal{W}_\mu^r$-states on $\mathcal{R}_\mu$, with $r = ht_\mu^\uparrow(\lambda - \nu)$.*

Lemma IV.3 also carries over, if one replaces the raising (lowering) operator E_{rs} by a positive (negative) root of $\mathfrak{g}$.

V. APPROXIMATION USING WEIGHT SPACES

A. The case $\lambda = \mu + \nu$ for $\mathcal{SU}(d)$, and some numerical examples for $\mathcal{SU}(2)$

Let us apply Theorem III.3 to rederive the following known result ([5, Theorem II.2]), which can be used to prove a standard de Finetti theorem (cf. [5]).
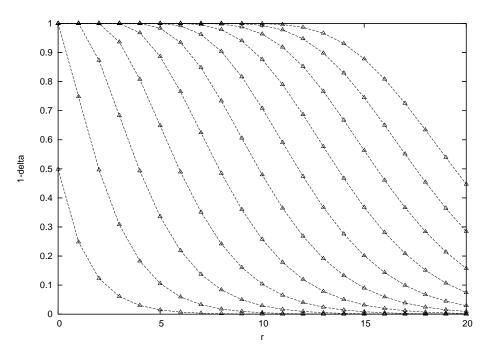
FIG. 1: This figure shows the quantity $1 - \delta_{|j_2 j_2\rangle}(\mathcal{W}^{\downarrow r})$ as a function of r. By Corollary III.3, this quantity is a measure of the error made when approximating $\mathrm{tr}_{j_2}|\Psi\rangle\langle\Psi|$ for $|\Psi\rangle \in \mathcal{R}_{(2j)} \subset \mathcal{R}_{(2j_1)} \otimes \mathcal{R}_{(2j_2)}$ by states with angular momentum $m_1 \geq j_1 - r$ in some direction. The different curves correspond to different values of j in the regime where $j \approx j_1 + j_2$. The solid innermost curve is for $j = j_1 + j_2$, and the dashed outermost curve corresponds to total angular momentum $j = j_1 + j_2 - 10$. In this example, $j_1 = j_2 = 100$. The point with $r = 0$ and $1 - \delta \approx 0.5$ of the innermost line corresponds to the result of [5] discussed at the end of Section III, where the highest weight vector $|(j_1 + j_2)(j_1 + j_2)\rangle = |j_1 j_1\rangle \otimes |j_2 j_2\rangle$ is a product. This no longer holds for general j. We thus expect the approximation to become worse for smaller j. The curves for smaller total angular momentum j show that we can nevertheless obtain a small approximation error when we include high weights (i.e., $\mathcal{W}^r$-states for $r > 0$).

Corollary V.1 (The case $\lambda = \mu + \nu$). *Let $\mathcal{R}_\lambda \subset \mathcal{R}_\mu \otimes \mathcal{R}_\nu$ be irreducible representations of $\mathcal{SU}(d)$, with $\lambda = \mu + \nu$, let $|\mu\rangle$ be the highest weight vector in $\mathcal{R}_\mu$ and let $|\Psi\rangle \in \mathcal{R}_\lambda$. Then there is a probability measure m on $\mathcal{SU}(d)$ such that*

$$\left\| \mathrm{tr}_{\mathcal{R}_\nu}|\Psi\rangle\langle\Psi| - \int \mathsf{g}|\mu\rangle\langle\mu|\mathsf{g}^\dagger dm(\mathsf{g}) \right\| \leq 2\left(1 - \frac{d_{\mathcal{R}_\nu}}{d_{\mathcal{R}_\lambda}}\right) .$$

Proof. This is directly obtained from Theorem III.3 by computing $\delta_{|\nu\rangle}(\mathcal{W}^{\downarrow 0})$, where $|\nu\rangle$ is the highest weight vector in $\mathcal{R}_\nu$ because the weight space corresponding to $\mathcal{W}^{\downarrow 0}$ is spanned by $|\mu\rangle$. Since $|\mu\rangle \otimes |\nu\rangle$ is the highest-weight vector in λ, we get $\delta_{|\nu\rangle}(\mathcal{W}^{\downarrow 0}) = d_{\mathcal{R}_\nu}/d_{\mathcal{R}_\lambda}$ which concludes the proof. $\square$

Our theorem also allows us to extend this to representations with highest weights $\lambda \neq \mu + \nu$ occuring in the tensor product $\mathcal{R}_\mu \otimes \mathcal{R}_\nu$, and include more weights in the approximation, i.e., use $\mathcal{W}^r$-states for $r > 0$.

Next we apply Theorem III.3 to the case of $\mathcal{SU}(2)$, where we can obtain bounds on the error made in the approximation by using the Racah formula for the Clebsch-Gordan coefficients to calculate $\delta(\mathcal{W})$. That is, we would like to approximate the partial trace $\mathrm{tr}_{\mathcal{R}_{(2j_2)}}|\Psi\rangle\langle\Psi|$ of a state $|\Psi\rangle \in \mathcal{R}_{(2j)} \subset \mathcal{R}_{(2j_1)} \otimes \mathcal{R}_{(2j_2)}$ with total angular momentum $|j_1 - j_2| \leq j \leq j_1 + j_2$. Note that each of these representations occurs with multiplicity 1 and the case $\lambda = \mu + \nu$ corresponds to $j = j_1 + j_2$. It is straightforward to check that the weight space corresponding to the extremal weights $\mathcal{W}^{\downarrow r}$ is spanned by $|j_1 m_1\rangle$ with $m_1 \geq j_1 - r$. These correspond to states with high angular momentum in the Z-direction.

With Theorem III.3, we conclude that there is a probability measure m on $\mathcal{SU}(2)$ and states $|\varphi_{(\mathsf{g})}\rangle$ supported on eigenstates of the angular momentum in the Z-direction with eigenvalue at least $j_1 - r$ such that

$$\left\| \mathrm{tr}_{j_2}|\Psi\rangle\langle\Psi| - \int \mathsf{g}|\varphi_{(\mathsf{g})}\rangle\langle\varphi_{(\mathsf{g})}|\mathsf{g}^\dagger dm(\mathsf{g}) \right\| \leq 2\left(1 - \delta_{|j_2 m_2\rangle}(\mathcal{W}^{\downarrow r})\right)$$

where $m_2 \in \{-j, -j+1, \ldots, j\}$ is arbitrary and

$$\delta_{|j_2 m_2\rangle}(\mathcal{W}^{\downarrow r}) = \frac{2j_2 + 1}{2j + 1} \sum_{m_1 = j_1 - r}^{j_1} |\langle j(m_1 + m_2)|j_1 m_1 j_2 m_2\rangle|^2 , \tag{23}$$
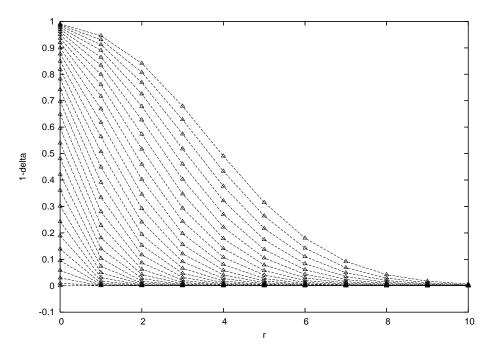
FIG. 2: This figure shows $1 - \delta_{|j_2 j_2\rangle}(\mathcal{W}^{\uparrow r})$ as a function of r in the regime where $j \approx |j_1 - j_2|$. More precisely, we set $j_1 = j_2 = 100$ and $j \in \{0, \ldots, 30\}$ (the curve corresponding to $j = 0$ is the innermost one). In accordance with Corollary IV.2, we can see from the figure that the error made in the approximation by $\mathcal{W}^r$-states (cf. Corollary III.3) vanishes as soon as sufficiently many high weights are used ($r \geq j$; see Example IV.4). Observe, however, that the error becomes exceedingly small for r significantly below the value j; e.g. for $j = 30$ (right-most curve), where $1 - \delta$ is close to zero for $r = 10$.

as a straightforward calculation shows. An equivalent statement holds for $\mathcal{W}^{\uparrow r}$. We give two numerical examples of this type of approximation in Figures 1 and 2.

B. Symmetric States: the exponential de Finetti Theorem

We now focus on symmetric states $|\Psi\rangle \in \mathsf{Sym}^n(\mathbb{C}^d)$. Here, the *symmetric subspace* $\mathsf{Sym}^n(\mathbb{C}^d) \subset (\mathbb{C}^d)^{\otimes n}$ is the subspace of vectors which are invariant under permutations of the systems, i.e., under the action of the symmetric group S_n on $(\mathbb{C}^d)^{\otimes n}$ defined by

$$\pi(|\varphi_1\rangle \otimes \cdots \otimes |\varphi_n\rangle) := |\varphi_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |\varphi_{\pi^{-1}(n)}\rangle \tag{24}$$

for all vectors $|\varphi_1\rangle, \ldots, |\varphi_n\rangle \in \mathbb{C}^d$ and $\pi \in S_n$. The group $\mathcal{SU}(d)$ acts on the same space simply by the n-fold tensor product $\mathsf{g}^{\otimes n}$ for every $\mathsf{g} \in \mathcal{SU}(d)$, i.e., its action is given by

$$\mathsf{g}(|\varphi_1\rangle \otimes \cdots \otimes |\varphi_n\rangle) := (\mathsf{g}|\varphi_1\rangle) \otimes \cdots \otimes (\mathsf{g}|\varphi_n\rangle) \tag{25}$$

for all vectors $|\varphi_1\rangle, \ldots, |\varphi_n\rangle \in \mathbb{C}^d$ and $\mathsf{g} \in \mathcal{SU}(d)$. The actions (24) and (25) commute, and $\mathsf{Sym}^n(\mathbb{C}^d)$ is well known to be an irreducible subspace with respect to the action of $\mathcal{SU}(d)$. Its dimension is $\dim \mathsf{Sym}^n(\mathbb{C}^d) = \binom{n+d-1}{n}$.

The set of weights $\mathcal{W}_{\mathsf{Sym}^n(\mathbb{C}^d)}$ occurring in the symmetric representation is given by the set of all d-tuples $\mathbf{w} = (w_1, \ldots, w_d)$ of nonnegative integers summing to n. The weight space corresponding to a weight $\mathbf{w} = (w_1, \ldots, w_d) \in \mathcal{W}_{\mathsf{Sym}^n(\mathbb{C}^d)}$ is one-dimensional and spanned by the vector

$$|\mathbf{w}\rangle = \frac{1}{\sqrt{|T^{\mathbf{w}}|}} \sum_{(v_1, \ldots, v_n) \in T^{\mathbf{w}}} |v_1\rangle \otimes \cdots \otimes |v_n\rangle . \tag{26}$$

In this expression, the *type class* $T^{\mathbf{w}}$ is defined as the set of all n-tuples $\mathbf{w}$ where i occurs w_i times, for $1 \leq i \leq d$. Thus the highest weight vector $|(n)\rangle = |1\rangle^{\otimes n}$ is of product form.

For the symmetric representation, writing the weights as vectors of integers, we have $\lambda = (n, 0, \ldots, 0)$, $\mathbf{w} = (w_1, \ldots, w_d)$, and, since $\alpha_i = (0, 0, \ldots, 1, -1, 0)$ with '1' in the i-th position, we can write $n_i = n - \sum_{j=1}^{i} w_j$. So

$ht^{\downarrow}_{\lambda}(\mathbf{w})$ is just $n - w_1$, and our distance measure just counts the number of times that a '1' in the highest weight vector is replaced by some other number. We conclude that $\mathcal{W}^{\downarrow r}$ of $\mathsf{Sym}^n(\mathbb{C}^d)$ is given by

$$\mathcal{W}^{\downarrow r} = \{\mathbf{w} \in \mathcal{W}_{\mathsf{Sym}^n(\mathbb{C}^d)} \mid w_1 \geq n - r\} . \tag{27}$$

Together with expression (26), this tells us that the weight space corresponding to $\mathcal{W}^{\downarrow r}$ is spanned by states which are of the form $|1\rangle^{\otimes n-r} \otimes |\varphi\rangle$ for (arbitrary) states $|\varphi\rangle \in (\mathbb{C}^d)^{\otimes r}$ up to permutations of the subsystems. Such almost-product states were called "$\binom{n}{n-r}$-i.i.d. with prototype $|1\rangle$" in [3]. In this terminology, a $\mathcal{W}^r$-state $|\chi\rangle$ is a symmetric state in the subspace spanned by states which are $\binom{n}{n-r}$-i.i.d. with prototype $\mathbf{g}|1\rangle$, where $\mathbf{g} \in \mathcal{SU}(d)$ is determined by $|\chi\rangle$.

We can now state the main result of this section, which tells us about the approximation of the partial trace of a symmetric state $|\Psi\rangle \in \mathsf{Sym}^n(\mathbb{C}^d)$ by $\mathcal{W}^r$-states. A theorem of this type was first proved in [2], and is further discussed and given a somewhat simpler proof in [3].

Corollary V.2 (Exponential approximation by $\mathcal{W}^r$-states). *Let $0 \leq r \leq k \leq n$. Then the partial trace $\mathrm{tr}_{n-k}|\Psi\rangle\langle\Psi|$ of every state $|\Psi\rangle \in \mathsf{Sym}^n(\mathbb{C}^d)$ can be approximated by a convex combination of $\mathcal{W}^r$-states in $\mathsf{Sym}^k(\mathbb{C}^d)$ with error*

$$\varepsilon_{n,k,r,d} := 2\frac{\dim \mathsf{Sym}^{n-k}(\mathbb{C}^d)}{\dim \mathsf{Sym}^n(\mathbb{C}^d)} \sum_{i=r+1}^{k} \frac{\binom{k}{i}}{\binom{n}{i}}\binom{i+d-2}{i} . \tag{28}$$

That is, there is a probability measure m on $\mathcal{SU}(d)$ and states $|\chi_{(\mathbf{g})}\rangle = \mathbf{g}|\varphi_{(\mathbf{g})}\rangle$, where $\mathbf{g} \in \mathcal{SU}(d)$ and $|\varphi_{(\mathbf{g})}\rangle$ is supported on the weight space corresponding to $\mathcal{W}^{\downarrow r}$, such that

$$\left\|\mathrm{tr}_{n-k}|\Psi\rangle\langle\Psi| - \int |\chi_{(\mathbf{g})}\rangle\langle\chi_{(\mathbf{g})}|dm(\mathbf{g})\right\| \leq \varepsilon_{n,k,r,d} . \tag{29}$$

In particular,

(i) if $r = 0$, then $\varepsilon_{n,k,0,d} = 2(1 - \frac{\dim \mathsf{Sym}^{n-k}(\mathbb{C}^d)}{\dim \mathsf{Sym}^n(\mathbb{C}^d)})$.

(ii) if $d \leq \min\{k, n - k\}$, then $\varepsilon_{n,k,r,d} \leq 2\frac{e^{3d}}{(d-2)!}\left(\frac{k}{n-r}\right)^{r+1}\left(\frac{k(n-k)}{n-r}\right)^{d-2}$.

Proof. We derive this result by applying Theorem III.3 to $\mathsf{Sym}^n(\mathbb{C}^d) \subset \mathsf{Sym}^k(\mathbb{C}^d) \otimes \mathsf{Sym}^{n-k}(\mathbb{C}^d)$, and bounding $\delta_{|\psi\rangle}(\mathcal{W}^{\downarrow r})$ for the highest weight vector $|\psi\rangle = |1\rangle^{\otimes n-k}$.

Consider a weight $\mathbf{w} \in \mathcal{W}_{\mathsf{Sym}^k(\mathbb{C}^d)}$. It is easy to verify for every permutation $\pi \in S_n$ and any two $v, v' \in T^{\mathbf{w}}$, $\langle v|\langle 1|^{\otimes n-k}\pi|v'\rangle|1\rangle^{\otimes n-k}$ is either 1 or 0, and the number of permutations that map $|v\rangle|1\rangle^{\otimes n-k}$ to $|v'\rangle|1\rangle^{\otimes n-k}$ is equal to $(w_1 + n - k)! \prod_{i=2}^{d} w_i!$, for any v and v' (Here, we write $|v\rangle$ for $|v_1\rangle \otimes \cdots \otimes |v_k\rangle$). In particular, since $P_{\mathsf{Sym}^n(\mathbb{C}^d)} = \frac{1}{n!}\sum_{\pi \in S_n} \pi$, we conclude

$$\mathrm{tr}\big(P_{\mathsf{Sym}^n(\mathbb{C}^d)}(|v'\rangle\langle v| \otimes |1\rangle\langle 1|^{\otimes n-k})\big) = \frac{1}{n!}(w_1 + n - k)!\prod_{i=2}^{d} w_i! .$$

Using this identity and inserting the definition (26) of $|\mathbf{w}\rangle$, we obtain

$$\mathrm{tr}\big(P_{\mathsf{Sym}^n(\mathbb{C}^d)}(|\mathbf{w}\rangle\langle\mathbf{w}| \otimes |1\rangle\langle 1|^{\otimes n-k})\big) = \frac{|T^{\mathbf{w}}|}{n!}(w_1 + n - k)!\prod_{i=2}^{d} w_i!$$

$$= \frac{k!}{n!}\frac{(w_1 + n - k)!}{w_1!} , \tag{30}$$

where we inserted the cardinality $|T^{\mathbf{w}}| = \frac{k!}{\prod_{i=1}^{d} w_i!}$ of the type class. This implies that for any set of weights $\mathcal{W} \subset \mathcal{W}_{\mathsf{Sym}^k(\mathbb{C}^d)}$, we have

$$\delta_{|\psi\rangle}(\mathcal{W}) = \frac{\dim \mathsf{Sym}^{n-k}(\mathbb{C}^d)}{\dim \mathsf{Sym}^n(\mathbb{C}^d)}\frac{k!}{n!}\sum_{i=0}^{k} \frac{(n - k + i)!}{i!}f_i(\mathcal{W}) , \tag{31}$$

where $f_i(\mathcal{W})$ is the number of weights of $\mathcal{W}$ with $w_1 = i$. Taking $\mathcal{W} = \mathcal{W}^r$, (27) tells us that $f_i(\mathcal{W}^r) = 0$ unless $i \geq k - r$, and for any such i, $f_i(\mathcal{W}^r)$ counts all possible weights with $w_1 = i$. This is just the number of weights with symbols in $\{2, \ldots, d\}$, or, equivalently, the dimension of $\mathsf{Sym}^{k-i}(d-1)$, which is $\binom{k+d-i-2}{k-i}$. Thus we obtain

$$\delta_{|\psi\rangle}(\mathcal{W}^r) = \frac{\dim \mathsf{Sym}^{n-k}(\mathbb{C}^d)}{\dim \mathsf{Sym}^n(\mathbb{C}^d)} \frac{k!}{n!} \sum_{i=k-r}^{k} \frac{(n-k+i)!}{i!} \binom{k+d-i-2}{k-i} . \tag{32}$$

However, we also know by Lemma III.5 (ii) that the sum on the r.h.s., if extended to all $0 \leq i \leq k$, is equal to 1, since this is the value $\delta_{|\psi\rangle}(\mathcal{W}_{\mathsf{Sym}^k(\mathbb{C}^d)})$. Combining this with (32) and substituting i by $k-i$ then shows that the expression (28) given in the theorem is equal to $2(1 - \delta_{|\psi\rangle}(\mathcal{W}^r))$. The first claim of the Theorem therefore follows from Theorem III.3. (Note that we can apply Remark III.4 because the symmetric representation occurs with multiplicity 1 in the tensor product $\mathsf{Sym}^k(\mathbb{C}^d) \otimes \mathsf{Sym}^{n-k}(\mathbb{C}^d)$.)

We have already shown the bound (i) for $r = 0$ in Corollary V.1. The bound (ii) follows from (28) after some further algebra that is deferred to Appendix 1. □

When $r = 0$, only the highest weight is used, resulting in an approximation by a convex combination of product states; this is the standard de Finetti theorem. The bound we obtain in this case is given as statement (i). It implies the $2\frac{dk}{n}$ bound stated as Corollary II.3 in [5] and shown there to be optimal in the number of systems. The special case where $k = 2$ was recently treated in [28]. Note that the version of the theorem in [2, 3] does not yield a useful bound for the $r = 0$ case; this is because some algebraic steps lose precision.

When $r > 0$, the approximation is by almost product states. The bound (ii) explains why we call this an exponential approximation. Consider for example the case where d is fixed, with $k = \alpha n$ and $r = \beta k$ for some constants $\alpha, \beta \in [0, 1]$. In the limit as $n \to \infty$, the bound has an asymptotic behaviour of the form $O(n^{d-2}(e^{-\alpha\beta \log(\frac{1-\alpha\beta}{\alpha})})^n)$. This tends to zero exponentially fast with n for suitable parameters α and β. Note that the bound $3(n-k)^d e^{-r(\frac{n-k}{n})}$ given in [2] translates into $O(n^d(e^{-\alpha\beta(1-\alpha)})^n)$, and thus the convergence of our bound (ii) is better when $\beta \leq 1/\alpha - e^{1-\alpha}$, which is the case whenever $\alpha \lesssim 0.34$.

By the arguments in [2, 5] we can extend this theorem from pure states $|\Psi\rangle \in \mathsf{Sym}^n(\mathbb{C}^d)$ to general (mixed) states on $(\mathbb{C}^d)^{\otimes n}$ that are symmetric, satisfying $\pi \rho^n \pi^\dagger = \rho^n$ for all $\pi \in S_n$. We first purify ρ^n in a symmetric way; this gives a symmetric pure state $|\Psi\rangle \in (\mathbb{C}^d \otimes \mathbb{C}^d)^{\otimes n}$ to which the theorem can be applied. The approximation error is then given by (28) with d replaced by d^2, and the approximating states are partial traces of pure almost product states.

VI. AN EXPONENTIAL THEOREM FOR THE HEISENBERG GROUP

We now turn to the Heisenberg group H. There is no weight space structure for the representations of H, as the Heisenberg algebra is nilpotent rather than semi-simple. However, the subspaces spanned by a particular range of number states play the role previously taken by weight spaces.

Consider the tensor product representation $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$, where $\mathcal{H}_\mu$ and $\mathcal{H}_\nu$ are irreducible representations with parameters μ, ν as in (6) of Section II B. For simplicity, we will henceforth assume that both μ and ν are positive, but we point out that our results can be extended to other cases.

Our first aim is to identify irreducible subspaces in the tensor product $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$. To do so, we will use the realisation of $\mathcal{H}_\mu$ and $\mathcal{H}_\nu$ based on a pair of creation- and annihilation operators $(a, a^\dagger)$ described by (11). We will write $a_1 = a \otimes 1_{\mathcal{H}_\nu}$ and $a_2 = 1_{\mathcal{H}_\mu} \otimes a$ and similarly for $a_1^\dagger$ and $a_2^\dagger$. We then have commutation relations such as $[a_1, a_2] = 0$. By definition of the tensor product representation, the element $(\alpha; t) \in \mathsf{H}$ acts on $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$ as the operator

$$e^{i\mu t} D_{a_1}(\sqrt{\mu}\alpha) \otimes e^{i\nu t} D_{a_2}(\sqrt{\nu}\alpha) = e^{i(\mu+\nu)t} D_{a_{\mu\otimes\nu}}(\sqrt{\mu+\nu}\alpha) , \tag{33}$$

where we used the commutation relations and identity (7) and introduced the operators

$$a_{\mu\otimes\nu} = \frac{1}{\sqrt{\mu+\nu}}(\sqrt{\mu}a_1 + \sqrt{\nu}a_2) \qquad \text{and} \qquad a_{\mu\otimes\nu}^\dagger = \frac{1}{\sqrt{\mu+\nu}}(\sqrt{\mu}a_1^\dagger + \sqrt{\nu}a_2^\dagger) . \tag{34}$$

It is straightforward to check that $a_{\mu\otimes\nu}^\dagger$ and $a_{\mu\otimes\nu}$ satisfy canonical commutation relations. Combining (33) with (11) thus demonstrates part (i) of the following:

Lemma VI.1. *Let $\mu, \nu > 0$. Consider the tensor product representation $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$ of the Heisenberg group H, where $\mathcal{H}_\mu$ and $\mathcal{H}_\nu$ are the irreducible representations as described by (11), and let $a_{\mu\otimes\nu}^\dagger$ and $a_{\mu\otimes\nu}$ be defined by (34). Then the following holds.*

(i) Let $|0\rangle_{\mu\otimes\nu} \in \mathcal{H}_\mu\otimes\mathcal{H}_\nu$ be a normalised vector such that $a_{\mu\otimes\nu}|0\rangle_{\mu\otimes\nu} = 0$. Then the action of H on $|0\rangle_{\mu\otimes\nu}$ generates an irreducible subspace isomorphic to $\mathcal{H}_{\mu+\nu}$ with orthonormal basis $\{\frac{(a^\dagger_{\mu\otimes\nu})^n}{\sqrt{n!}}|0\rangle_{\mu\otimes\nu}\}_{n\in\mathbb{N}_0}$.

(ii) Let $|n\rangle := \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle$ and define

$$|\psi^\Delta\rangle = \sum_{\ell=0}^{\Delta}(-1)^\ell\sqrt{\alpha_\ell}|\Delta-\ell\rangle\otimes|\ell\rangle \qquad where \qquad \alpha_\ell = \frac{\binom{\Delta}{\ell}\mu^\ell\nu^{\Delta-\ell}}{(\mu+\nu)^\Delta}$$

for $\Delta \in \mathbb{N}_0$. Then $|\psi^\Delta\rangle$ is a normalised vector in $\mathcal{H}_\mu\otimes\mathcal{H}_\nu$ with $a_{\mu\otimes\nu}|\psi^\Delta\rangle = 0$.

(iii) Let $\mathcal{H}^\Delta_{\mu+\nu} \subset \mathcal{H}_\mu\otimes\mathcal{H}_\nu$ be the irreducible representation generated by $|\psi^\Delta\rangle$. Then $\mathcal{H}^\Delta_{\mu+\nu}$ is isomorphic to $\mathcal{H}_{\mu+\nu}$ and has formal degree $d_{\mathcal{H}^\Delta_{\mu+\nu}} = \mu+\nu$ as a representation of the quotient group H/Z.

(iv) The tensor product representation decomposes into a direct sum of the representations $\mathcal{H}^\Delta_{\mu+\nu}$; that is,

$$\mathcal{H}_\mu\otimes\mathcal{H}_\nu \cong \bigoplus_{\Delta\in\mathbb{N}_0} \mathcal{H}^\Delta_{\mu+\nu} \ .$$

Proof. Statement (ii) follows by straightforward computation using (34). Statement (iii) is a consequence of the fact that $d_{\mathcal{H}_{\mu+\nu}} = \mu+\nu$ (cf. end of the Section II B) and the fact that the representation $\mathcal{H}_{\mu+\nu}$ is identical to $\mathcal{H}^\Delta_{\mu+\nu}$ when the "vacuum state" $|0\rangle$ is identified with $|\psi^\Delta\rangle$ and the operators a, $a^\dagger$ are identified with (34). To prove (iv), we first prove that the subspaces $\mathcal{H}^\Delta_{\mu+\nu}$ and $\mathcal{H}^{\Delta'}_{\mu+\nu}$ are orthogonal for $\Delta\neq\Delta'$. By definition, this is equivalent to showing $\langle\varphi^\Delta_n|\varphi^{\Delta'}_{n'}\rangle = 0$ for all n and n', where $|\varphi^\Delta_n\rangle := (a^\dagger_{\mu\otimes\nu})^n|\psi^\Delta\rangle$. Because $|\varphi^\Delta_n\rangle$ is supported on the span of the states $\{|n_1\rangle\otimes|n_2\rangle \mid n_1+n_2 = n+\Delta\}$, this certainly holds when $\Delta+n \neq \Delta'+n'$. If on the other hand $\Delta+n = \Delta'+n'$, then the overlap between $|\varphi^\Delta_n\rangle$ and $|\varphi^{\Delta'}_{n'}\rangle$ must vanish since these two vectors are eigenvectors of the hermitian operator $a^\dagger_{\mu\otimes\nu}a_{\mu\otimes\nu}$ with distinct eigenvalues $n\neq n'$. Finally, to prove that these subspaces provide a complete decomposition of $\mathcal{H}_\mu\otimes\mathcal{H}_\nu$, it suffices to observe that for each $m\in\mathbb{N}_0$, the subspace spanned by $\{|k\rangle\otimes|\ell\rangle \mid k+\ell = m\}$ is identical to the span of $\{|\varphi^\Delta_n\rangle \mid \Delta+n = m\}$. $\qquad\square$

We will now study the partial trace of a state $|\Psi\rangle \in \mathcal{H}^\Delta_{\mu+\nu}$ and show that it can be approximated by rotated number states. For an irreducible representation $\mathcal{H}_\lambda$, we set $\mathcal{N}_{\mathcal{H}_\lambda} = \mathbb{N}_0$, and, for a set of numbers $\mathcal{N}\subset\mathcal{N}_{\mathcal{H}_\lambda}$, we will define the corresponding number subspace $\mathcal{H}^\mathcal{N}_\lambda$ as the span of the states $|n\rangle$, $n\in\mathcal{N}$ obtained by raising the vacuum state n times. By analogy with the set of extremal weights $\mathcal{W}^{\uparrow r}_\lambda$ we define, for $r\in\mathbb{N}_0$,

$$\mathcal{N}^r := \{n \mid n \leq r\} \subset \mathcal{H}_\mu \ .$$

The error term in the following corollary depends on the way the representation $\mathcal{H}_{\mu+\nu}$ is embedded into $\mathcal{H}_\mu\otimes\mathcal{H}_\nu$; different values of Δ lead to a different approximation error.

Corollary VI.2 (Number space approximation for the Heisenberg group). *Let $\mu,\nu > 0$ and let $\Delta \in \mathbb{N}_0$. Let $\mathcal{H}^\Delta_{\mu+\nu} \subset \mathcal{H}_\mu\otimes\mathcal{H}_\nu$ be irreducible representations of the Heisenberg group as described by Lemma VI.1. Then*

$$\delta_{|0\rangle}(\mathcal{N}^r) = \left(\frac{\nu}{\mu+\nu}\right)^{\Delta+1}\sum_{n=0}^{r-\Delta}\binom{n+\Delta}{\Delta}\left(\frac{\mu}{\mu+\nu}\right)^n \ . \tag{35}$$

In particular, for every $|\Psi\rangle \in \mathcal{H}^\Delta_{\mu+\nu}$, the partial trace $\mathrm{tr}_{\mathcal{H}_\nu}|\Psi\rangle\langle\Psi|$ can be approximated by a convex combination of $\mathcal{N}^r$-states. That is, there are states $|\chi_{(\alpha)}\rangle$ and a probability measure m on $\mathbb{C}$ such that

$$\left\|\mathrm{tr}_{\mathcal{U}_\nu}|\Psi\rangle\langle\Psi| - \int|\chi_{(\alpha)}\rangle\langle\chi_{(\alpha)}|dm(\alpha)\right\| \leq \varepsilon_{\mu,\nu,\Delta,r} \ ,$$

where $|\chi_{(\alpha)}\rangle$ is supported on $D_{a_1}(\alpha)\,\mathsf{span}\{|n\rangle \mid n \leq r\}$ and where

$$\varepsilon_{\mu,\nu,\Delta,r} = \begin{cases} 2\big(1-\delta_{|0\rangle}(\mathcal{N}^r)\big) & if\ \Delta=0\ and\ r=0 \\ 2\sqrt{1-\delta_{|0\rangle}(\mathcal{N}^r)} & otherwise \ . \end{cases}$$

In particular, if $\Delta = 0$, then

$$\varepsilon_{\mu,\nu,0,r} = \begin{cases} 2\big(\frac{\mu}{\mu+\nu}\big) & if\ r=0 \\ 2\big(\frac{\mu}{\mu+\nu}\big)^{(r+1)/2} & otherwise \ . \end{cases}$$
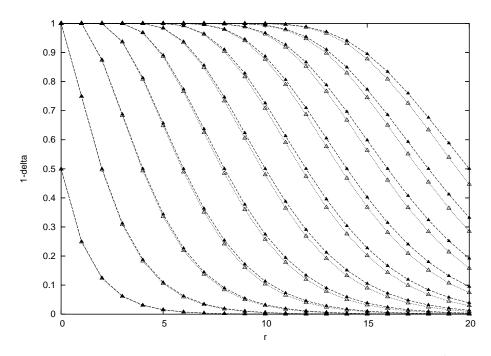
FIG. 3: We illustrate the result of Corollary VI.2 for the case of the irreducible representations $\mathcal{H}_n^\Delta \subset \mathcal{H}_k \otimes \mathcal{H}_{n-k}$ for $k = 50$ and $n = 100$. The figure shows the quantity $1 - \delta_{|0\rangle}(\mathcal{N}^r)$ as a function of r, where the different curves correspond to different values of Δ (dashed lines with filled triangles). The innermost curve corresponds to $\Delta = 0$, whereas the outermost curve corresponds to $\Delta = 10$. For comparison, the curves from Figure 1 are also shown (dotted lines with empty triangles). Note how close they are, even though they represent very different objects: For $\mathcal{SU}(2)$, the curves correspond to inequivalent irreducible representations, whereas for the Heisenberg group, they are associated with different copies of the same representation in the tensor product.

We use this corollary to produce Figure 3, which shows both the exponential decay of the approximation error when more approximating states are used (i.e., for varying r) and how different embeddings of the same representation $\mathcal{H}_{\mu+\nu}$ into the tensor product $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$ (i.e., different values of Δ) give rise to varying approximation errors.

Proof. We will prove identity (35). The result then is a direct consequence of Theorem III.3 and Remark III.4.

The vectors $|\psi_n^\Delta\rangle = \frac{(a_{\mu\otimes\nu}^\dagger)^n}{\sqrt{n!}}|\psi^\Delta\rangle$ are a basis of the subspace $\mathcal{H}_{\mu+\nu}^\Delta$ generated by $|\psi^\Delta\rangle$ according to Lemma VI.1. Because of the commutativity of $a_1^\dagger$ and $a_2^\dagger$, we have

$$(a_{\mu\otimes\nu}^\dagger)^n = \left(\frac{1}{\sqrt{\mu+\nu}}\right)^n \sum_{s=0}^n \binom{n}{s} \sqrt{\mu}^s \sqrt{\nu}^{n-s} (a_1^\dagger)^s (a_2^\dagger)^{n-s} \ .$$

Note also that $(a^\dagger)^s|\ell\rangle = \frac{\sqrt{(s+\ell)!}}{\sqrt{\ell!}}|s+\ell\rangle$. In particular, (by setting $s = n$ and $\ell = 0$) we obtain

$$|\psi_n^\Delta\rangle = \sqrt{\alpha_{\Delta,n}}|\Delta+n\rangle \otimes |0\rangle + \dots \ ,$$

where

$$\alpha_{\Delta,n} = \left(\frac{\nu}{\mu+\nu}\right)^\Delta \binom{n+\Delta}{\Delta} \left(\frac{\mu}{\mu+\nu}\right)^n \tag{36}$$

and the remaining terms are supported on $\mathcal{H}_\mu \otimes \mathsf{span}\{|n\rangle \mid n \geq 1\}$. We can then compute

$$\sum_{n=0}^\infty \sum_{n'=0}^r \mathsf{tr}\big(|\psi_n^\Delta\rangle\langle\psi_n^\Delta|(|n'\rangle\langle n'| \otimes |0\rangle\langle 0|)\big) = \sum_{n=0}^{r-\Delta} \alpha_{\Delta,n} \ . \tag{37}$$

Identity (35) now follows from (36) and (37) using the definition of $\delta_{|0\rangle}(\mathcal{N}^r)$ and the formal degrees $d_{\mathcal{H}_\nu} = \nu$ (cf. discussion at the end of Section II B) and $d_{\mathcal{H}_{\mu+\nu}}^\Delta = \mu + \nu$ (Lemma VI.1). $\qquad\square$

Corollary VI.2 turns into the result of [17] when specialised to integer μ and ν, and $r = \Delta = 0$. We will now explain why this is the case, and how the subspace spanned by k-fold coherent states considered in [17] is related to representations of the Heisenberg group.

We begin by taking the k-fold tensor product of the irreducible representation $\mathcal{H}_1$. As before, let $a_i^\dagger$ and a_i be the creation and annihilation operators acting on the i-th factor in this tensor product. As with (34), it is then easy to see that the operators

$$a = \frac{1}{\sqrt{k}} \sum_{i=1}^k a_i \qquad \text{and} \qquad a^\dagger = \frac{1}{\sqrt{k}} \sum_{i=1}^k a_i^\dagger \ . \tag{38}$$

satisfy canonical commutation relations, and the vector $|0\rangle^{\otimes k}$ is annihilated by a. If we let $(\alpha; t)$ act as $e^{it} D_a(\alpha)$ on this vector, this generates an irreducible representation isomorphic to $\mathcal{H}_1$ with orthonormal basis vectors

$$|n\rangle_k = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle^{\otimes k} = k^{-n/2} \sum_{(f_1, \ldots, f_k): \sum_i f_i = n} \sqrt{\frac{n!}{f_1! \cdots f_k!}} |f_1\rangle \otimes \cdots \otimes |f_k\rangle \ , \tag{39}$$

where the sum is over all k-tuples $(f_1, \ldots, f_k)$ of nonnegative integers. This is the space spanned by vectors of the form $\{|\varphi_\alpha\rangle^{\otimes k}\}_{\alpha \in \mathbb{C}}$, where $|\varphi_\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^\infty \frac{\alpha^n}{\sqrt{n!}} |n\rangle$ is a *coherent* state. By the recipe described in Section II B, we can use this to define an irreducible representation isomorphic to $\mathcal{H}_k$. This is done by letting the element $(\alpha; t)$ act on the same space as $e^{ikt} D_a(\sqrt{k}\alpha)$. If we call the resulting representation $\mathcal{H}^{(k)}$, then clearly $\mathcal{H}^{(n)} \subset \mathcal{H}^{(k)} \otimes \mathcal{H}^{(n-k)}$ as representations, and this corresponds to the irreducible representation $\mathcal{H}_n^0 \subset \mathcal{H}_k \otimes \mathcal{H}_{n-k}$ identified in Lemma VI.1. With Corollary VI.2, we thus obtain the following generalised form of the statement of [17].

Corollary VI.3. *Let a and $a^\dagger$ be operators satisfying $[a, a^\dagger] = 1$ and let $|0\rangle$ be such that $a|0\rangle = 0$. Define the number states $|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle$ for $n \in \mathbb{N}_0$ and the displacement operators $D(\alpha) = \exp(\alpha a^\dagger - \bar{\alpha} a)$ for $\alpha \in \mathbb{C}$. Finally, let $|\Psi\rangle \in \mathsf{Sym}^n(\mathcal{H}) \subset \mathcal{H}^{\otimes n}$ be a pure state in the span of the set of states of the form $(D(\alpha)|0\rangle)^{\otimes n}$. Then there are states $|\chi_{(\alpha)}\rangle$ and a probability measure m on $\mathbb{C}$ such that*

$$\left\| \mathsf{tr}_{n-k} |\Psi\rangle\langle\Psi| - \int |\chi_{(\alpha)}\rangle\langle\chi_{(\alpha)}| dm(\alpha) \right\| \leq \begin{cases} 2\frac{k}{n} & \text{if } r = 0 \\ 2\left(\frac{k}{n}\right)^{(r+1)/2} & \text{otherwise,} \end{cases}$$

where $|\chi_{(\alpha)}\rangle$ is supported on $D(\alpha)^{\otimes k} \mathsf{span}\{|n\rangle_k \mid n \leq r\}$, with $|n\rangle_k$ defined by (39). In particular, for $r = 0$ we have $|\chi_{(\alpha)}\rangle = (D(\alpha)|0\rangle)^{\otimes k}$.

VII. CONCLUSIONS

It is striking that representations lend themselves so well to studying de Finetti theorems. Schur's lemma is the essential representation-theoretic tool in proving our main theorem. In addition, for $\mathcal{SU}(d)$, and more generally for semi-simple groups, the weight space structure provides a natural family of approximating states that we call $\mathcal{W}^r$-states; these amount, in the case of a symmetric representation, to almost-product states. For the Heisenberg group, subspaces $\mathcal{N}^r$ spanned by sets of number states play the same role.

Our theorem assumes that we have irreducible representations $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ satisfying $\mathcal{C} \subset \mathcal{A} \otimes \mathcal{B}$, and then tells us how well the trace $\mathsf{tr}_\mathcal{B} |\Psi\rangle\langle\Psi|$ of a state $|\Psi\rangle$ in $\mathcal{C}$ can be approximated by $\mathcal{W}^r$-states in $\mathcal{A}$ (or $\mathcal{N}^r$-states for the Heisenberg group). The quality of this approximation is determined by the way $\mathcal{C}$ is embedded in the product $\mathcal{A} \otimes \mathcal{B}$. This is captured by the number δ, given by Definition III.2. We give several examples of explicit calculations of this number:

(i) For the embedding of symmetric subspaces $\mathcal{R}_{(n)} \subset \mathcal{R}_{(k)} \otimes \mathcal{R}_{(n-k)}$, $\delta(\mathcal{W}^r)$ is given by equation (32). This corresponds to the exponential theorem for symmetric states proved in [2, 3], and we obtain bounds from (32) that reproduce, and in fact slightly improve upon, the results in those papers.

(ii) In the case of the Heisenberg group, there is an irreducible representation $\mathcal{H}_\mu$ for every non-zero parameter $\mu \in \mathbb{R}$, but only representations isomorphic to $\mathcal{H}_{\mu+\nu}$ occur in the tensor product $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$. The corresponding $\delta(\mathcal{N}^r)$ is given by equation (35) and depends on how this representation is embedded (see Corollary VI.2). This allows us to prove an exponential theorem (Corollary VI.3) that generalises the result in [17].

(iii) The theorem for representations of the unitary group proved in [5] corresponds to $\mathcal{R}_{\mu+\nu} \subset \mathcal{R}_\mu \otimes \mathcal{R}_\nu$, where each $\mathcal{R}$ is a representation of $\mathcal{SU}(d)$, and in this case $\delta(\mathcal{W}^r) = \dim \mathcal{R}_\nu / \dim \mathcal{R}_{\mu+\nu}$ for $r = 0$ (see Lemma III.5), giving the bound for Theorem 2.1 in [5].

(iv) For $\mathcal{SU}(2)$, the mapping of weight vectors in $\mathcal{C}$ into those of $\mathcal{A} \otimes \mathcal{B}$ is given by the Clebsch-Gordan coefficients, and in that case the explicit expression for $\delta(\mathcal{W}^r)$ in equation (23) enables us to give give bounds for the de Finetti approximation by $\mathcal{W}^r$-states for a range of values of angular momenta and r (see Figure 1).

(v) Given representations of $\mathcal{SU}(d)$ satisfying $\mathcal{R}_\lambda \subset \mathcal{R}_\mu \otimes \mathcal{R}_\nu$, we find $\delta(\mathcal{W}^r) = 1$ when $r = ht_\mu^\uparrow(\lambda - \nu)$, where $ht_\mu^\uparrow(\lambda - \nu)$ measures the distance from the lowest weight μ_* to $\lambda - \nu$ (see Section IV). For these values of r, the traced-out state is *exactly* given by a convex sum of $\mathcal{W}^r$-states.

Acknowledgments

1. Combinatorics and the upper bound in Corollary V.2

In this appendix, we show that

$$\frac{f_{n-k,d}}{f_{n,d}} \sum_{i=r+1}^k \frac{\binom{k}{i}}{\binom{n}{i}} \binom{i+d-2}{i} \leq \frac{e^{3d}}{(d-2)!} \left(\frac{k}{n-r}\right)^{r+1} \left(\frac{k(n-k)}{n-r}\right)^{d-2} \tag{40}$$

if $d \leq \min\{k, n-k\}$, where $f_{n,d} = \dim \mathsf{Sym}^n(\mathbb{C}^d) = \binom{n+d-1}{n}$.

Proof. We use the identity

$$\sum_{i=r+1}^k \frac{\binom{k}{i}}{\binom{n}{i}} = \frac{k!(n-r)!}{(n-k+1)n!(k-r-1)!} \tag{41}$$

which can be proven by observing that the identity holds trivially for $r = k-1$ and by checking that both sides satisfy the recursion relation

$$a_r = a_{r-1} - \frac{\binom{k}{r}}{\binom{n}{r}}$$

as a function of r. We thus obtain

$$\frac{f_{n-k,d}}{f_{n,d}} \sum_{i=r+1}^k \frac{\binom{k}{i}}{\binom{n}{i}} \binom{i+d-2}{i}$$

$$\leq \frac{f_{n-k,d}}{f_{n,d}} \binom{k+d-2}{k} \sum_{i=r+1}^k \frac{\binom{k}{i}}{\binom{n}{i}} \tag{42}$$

$$= \frac{1}{(d-2)!} \frac{(k+d-2)!(n-k+d-1)!(n-r)!}{(k-r-1)!(n-k+1)!(n+d-1)!} . \tag{43}$$

We can further bound (43) as follows. Note that

$$\frac{(k+d-2)!}{(k-(r+1))!} \leq k^{d+r-1} (1 + \frac{d-2}{k})^{d+r-1}$$

and thus with the inequality $(1+x)^y \leq e^{xy}$,

$$\frac{(k+d-2)!}{(k-(r+1))!} \leq k^{d-1+r} e^{\frac{(d-1)^2}{k}} e^{\frac{(d-1)r}{k}} . \tag{44}$$

Similarly, we have

$$\frac{(n-k+d-1)!}{(n-k+1)!} \leq (n-k)^{d-2} e^{\frac{(d-1)^2}{n-k}} \tag{45}$$

$$\frac{(n-r)!}{(n+d-1)!} \leq (n-r)^{-(d-1+r)} . \tag{46}$$

Combining (43) with (44), (45) and (46), we conclude that

$$\frac{1}{(d-2)!}\Big(\frac{k}{n-r}\Big)^{d-1+r}(n-k)^{d-2} e^{\frac{(d-1)r}{k}+\frac{(d-1)^2}{k}+\frac{(d-1)^2}{n-k}} \tag{47}$$

is an upper bound on the quantity (40). The claim then follows from this. $\square$

It is interesting to note that for $d=2$, inequality (42) is tight, and the error (28) in Corollary V.2 is given by

$$2\Big(\frac{k!}{(k-r-1)!} \cdot \frac{(n-r)!}{(n+1)!}\Big) \qquad \text{for } d=2 .$$

[1] E. Halley, *A Most Compendious and Facile Method for Constructing the Logarithms, Exemplified and Demonstrated from the Nature of Numbers, without any Regard to the Hyperbola, with a Speedy Method for Finding the Number from the Logarithm Given*, Philosophical Transaction Series I, vol. 19, p. 58–67, 1965.
[2] R. Renner, *Security of Quantum Key Distribution*, PhD thesis, ETH Zurich, 2005, arXiv:quant-ph/0512258.
[3] R. Renner, (2007), arXiv:quant-ph/0703069.
[4] R. König and R. Renner, J. Math. Phys. **46**, 122108 (2005).
[5] M. Christandl, R. König, G. Mitchison, and R. Renner, (2006), arXiv:quant-ph/0602130.
[6] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim, (2006), arXiv:quant-ph/0608195.
[7] G. Chiribella and G. M. D'Ariano, (2006), arXiv:quant-ph/0608007.
[8] C. A. Fuchs and R. Schack, (2004), arXiv:quant-ph/0404156.
[9] R. L. Hudson, Found. Phys. **11**, 805 (1981).
[10] M. Fannes, H. Spohn, and A. Verbeure, J. Math. Phys. **21**, 355 (1980).
[11] G. A. Raggio and R. F. Werner, Helv. Phys. Acta **62**, 980 (1989).
[12] T. A. Brun, C. M. Caves, and R. Schack, Phys. Rev. A **63**, 042309 (2001).
[13] A. C. Doherty, P. A. Parillo, and F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004).
[14] K. M. R. Audenaert, Proceedings of MTNS2004 (2004), arXiv:quant-ph/0402076.
[15] B. M. Terhal, A. C. Doherty, and D. Schwab, Phys. Rev. Lett **90**, 157903 (2003).
[16] J. Bae and A. Acin, (2006), arXiv:quant-ph/0603078.
[17] C. D'Cruz, T. J. Osborne, and R. Schack, (2006), arXiv:quant-ph/0606139.
[18] G. Mitchison, (2007), arXiv:quant-ph/0701064.
[19] A. Perelomov, *Generalized Coherent States and Their Applications* (Springer, 1986).
[20] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, 1987).
[21] R. Carter, G. Segal, and I. MacDonald, *Lectures on Lie Groups and Lie Algebras*, London Mathematical Society Student Texts Vol. 32, 1 ed. (cup, 1995).
[22] W. Fulton and J. Harris, *Representation Theory : A First Course* (Springer, 1999).
[23] R. Goodman and N. R. Wallach, *Representations and Invariants of the Classical Groups* (Cambridge University Press, 2003).
[24] A. W. Knapp, *Lie groups beyond an introduction* (Birkhäuser, 2005).
[25] S. A. Gaal, *Linear Analysis and Representation Theory* (Springer, 1973).
[26] A. Winter, IEEE Transactions on Information Theory **45** (1999).
[27] T. Ogawa and H. Nagaoka, (2002), arXiv:quant-ph/0208139.
[28] M. Fannes and C. Vandenplas, (2006), arXiv:quant-ph/0605216.