Comonotonicity and Choquet integrals of Hermitian operators and their applications

A. Vourdas
Department of Computing,
University of Bradford,
Bradford BD7 1DP, United Kingdom
a.vourdas@bradford.ac.uk

In a quantum system with d-dimensional Hilbert space, the Q-function of a Hermitian positive semidefinite operator θ, is defined in terms of the d^2 coherent states in this system. The Choquet integral $\mathcal{C}_Q(\theta)$ of the Q-function of θ, is introduced using a ranking of the values of the Q-function, and Möbius transforms which remove the overlaps between coherent states. It is a figure of merit of the quantum properties of Hermitian operators, and it provides upper and lower bounds to various physical quantities in terms of the Q-function. Comonotonicity is an important concept in the formalism, which is used to formalize the vague concept of physically similar operators. Comonotonic operators are shown to be bounded, with respect to an order based on Choquet integrals. Applications of the formalism to the study of the ground state of a physical system, are discussed. Bounds for partition functions, are also derived.

I. INTRODUCTION

There are many quantities which describe quantum properties of quantum systems. The various entropic quantities (von Neumann entropy, Wehrl entropy [1], etc) are examples of this. In this paper we introduce Choquet integrals as indicators of the quantum properties of Hermitian operators. Choquet integrals are used in problems with probabilities, where the various alternatives are not independent, but they overlap with each other.

We consider a quantum system $\Sigma(d)$ with variables in $\mathbb{Z}(d)$ (the integers modulo d), described with the d-dimensional Hilbert space $H(d)$ [2, 3]. Let Ω be the set of d^2 coherent states, associated with the Heisenberg-Weyl group of displacements in this discrete system. The overlapping nature of coherent states is our motivation for the use of Choquet integrals. We map the Q function of a Hermitian positive semidefinite operator θ, into the Choquet integral $\mathcal{C}_Q(\theta)$, which is also a Hermitian positive semidefinite operator. The formalism uses capacities (non-additive probabilities) and Choquet integrals, and we briefly introduce these concepts.

Capacities (non-additive probabilities): The basic property of Kolmogorov probabilities is additivity ($\mu(A \cup B) - \mu(A) - \mu(B) + \mu(A \cap B) = 0$). But in subjects like Artificial Intelligence, Operations Research, Game Theory, Mathematical Economics, etc, nonadditive probabilities have been used extensively (e.g., [4–8]). They are particularly useful in problems where the various alternatives overlap, and they formalize the added value in an aggregation, where the 'whole is greater than the sum of its parts'.

In recent work[9] we have shown that there is a strong link between the non-commutativity of general projectors, and the non-additivity of the corresponding probabilities (Eq.(54) below). This leads naturally to Choquet integrals, which we introduce in this paper in a quantum context, and discuss their use as figures of merit of the quantum properties of Hermitian operators.

Choquet integrals in a classical context: Integration is based on additivity. Integrals with non-additive probabilities require another approach, and this leads to Choquet integration[10], which has been used extensively in Artificial Intelligence[11–15], in Game Theory and its applications in Mathematical Economics[16–19], etc. In a 'weighted average' we have a number of independent alternatives, and we assign a probability to each alternative. The Choquet integral is a 'sophisticated weighted average', for non-additive probabilities related to overlapping alternatives. It replaces the probability distributions used in weighted averages, with the derivative of cumulative functions, and by doing so, it assigns weights to aggregations of alternatives. The weight for an aggregation of alternatives, is in general different from the sum of the weights, of the alternatives it contains.

Consequently, the derivative of cumulative functions is in general different from the probability distributions (they are always equal in the case of additive probabilities).

Choquet integrals in a quantum context: In this paper we map the Q function of a Hermitian positive semidefinite operator θ (defined with respect to the set Ω of coherent states), into the discrete Choquet integral $\mathcal{C}_Q(\theta)$. It is calculated using cumulative projectors, and their discrete derivatives (differences) which are d orthogonal projectors.

An important concept related to Choquet integrals, is comonotonicity of two operators θ, ϕ. There is added value in an aggregation of components with different properties, because the various components play complementary role to each other. In this case the whole is different from the sum of its parts, and the $\mathcal{C}_Q(\theta + \phi)$ is different from $\mathcal{C}_Q(\theta) + \mathcal{C}_Q(\phi)$. But if the components of an aggregation have similar properties, this complementarity and added value are missing, the whole is equal to the sum of its parts, and $\mathcal{C}_Q(\theta + \phi)$ is equal to $\mathcal{C}_Q(\theta) + \mathcal{C}_Q(\phi)$. Comonotonicity defines rigorously the intuitive concept of physically similar operators.

We next compare briefly the Choquet formalism with the spectral formalism of orthogonal projectors, the positive operator valued measures (POVM), and the formalism of frames and wavelets:

- The spectral formalism of eigenvalues and eigenvectors, leads in the case of Hermitian operators to orthogonal projectors, which play a fundamental role in von Neumann's measurement theory.

- The POVM formalism uses projectors related to coherent states (or other non-orthogonal and non-commuting projectors), and it is based on a resolution of the identity. The resolution of the identity is crucial for the calculation of various physical quantities in terms of coherent states.

- The formalism of frames and wavelets, is based on lower and upper bounds to a resolution of the identity, and in this sense it uses approximate resolutions of the identity with bounded error.

- The Choquet formalism uses a 'weak resolution of the identity', that involves not only the non-orthogonal projectors, but also a correction which consists of 'Möbius operators' that eliminate the 'double counting' in the sum of the non-orthogonal projectors (Eq.(61)). The Choquet integral $\mathcal{C}_Q(\theta)$ can be expressed in terms of the Möbius operators (as in proposition IV.1).

Physical applications: The Choquet integral $\mathcal{C}_Q(\theta)$ is a figure of merit of the quantum properties of Hermitian operators. Its physical applications include:

- upper and lower bounds to various physical quantities in terms of the Q-function (proposition IV.3). This includes the derivation of bounds for partition functions (section VIII).

- the study of changes in the ground state of physical systems. Hamiltonians with and without degeneracies are considered, and it is shown that the Choquet integral $\mathcal{C}_Q(\theta)$ detects changes in the ground state of the system (sections VII A, VII B).

- the formalism leads naturally to the concept of comonotonicity. It is shown that comonotonic operators are bounded within certain intervals, with respect to an order based on Choquet integrals, and in this sense they are similar to each other (section V).

A desirable feature of the formalism, is that it is robust in the presence of noise, and yet it is sensitive enough to detect changes in the physical system (e.g., changes in the ground state of the system).

Contents: In section 2, we introduce capacities and Choquet integrals in a classical context. There is much literature on these concepts in other than Physics areas, and here we present briefly the concepts that we are going to bring into Quantum Physics. In section 3, we introduce technical details (cumulative coherent projectors and their discrete derivatives, Möbius operators, etc [20]) which are needed in the calculation of the Choquet integral.

In section 4, we introduce the Choquet integral $\mathcal{C}_Q(\theta)$ of a Hermitian operator θ, and study its properties. In section 5, we discuss the concept of comonotonic operators. In section 6, we introduce an order based on Choquet integrals, and show that comonotonic operators are bounded with respect to this order. This implies

that certain physical quantities are also bounded. In section 7, we apply the formalism to the study of the ground state of a physical system. In section 8, we derive bounds for partition functions. In section 9 we compare and contrast the Choquet formalism, with the spectral formalism of eigenvectors/eigenvalues, the POVM formalism, and the formalism of wavelets (and frames). We conclude in section 10, with a discussion of our results.

II. CAPACITIES AND DISCRETE CHOQUET INTEGRALS IN A CLASSICAL CONTEXT

A. Capacities for overlapping and non-independent alternatives

Kolmogorov probability is a map μ from subsets of a 'set of alternatives' Ω, to $[0,1]$. Its basic property is the additivity relation

$$\delta(A,B) = 0; \quad \delta(A,B) = \mu(A \cup B) - \mu(A) - \mu(B) + \mu(A \cap B); \quad A, B \subseteq \Omega. \tag{1}$$

In the case $A \cap B = \emptyset$ this reduces to

$$A \cap B = \emptyset \;\; \rightarrow \;\; \mu(A \cup B) = \mu(A) + \mu(B) \tag{2}$$

Capacity or nonadditive probability, is a weaker concept which obeys the relations

$$\mu(\emptyset) = 0; \quad \mu(\Omega) = 1$$
$$A \subseteq B \;\; \rightarrow \;\; \mu(A) \leq \mu(B) \tag{3}$$

If the second of these requirements is replaced with the additivity relation of Eq.(2) which is stronger, then the capacity is probability. A prerequisite for the use of probabilities is the assumption that the alternatives in the set Ω are separable from each other, and truly independent. In capacities this assumption is relaxed, the aggregation of some of the alternatives is different from the sum of its parts, and Eq.(2) is not valid.

Capacities have been introduced by Choquet [10], and they have been used extensively in areas like Artificial Intelligence, Operations Research, Game Theory, Mathematical Economics, etc. They describe the added value in an aggregation, where the 'whole is greater than the sum of its parts'. For example, the percentage of votes in a coalition of two political parties, might be greater (or smaller) than the sum of the percentages in the component parties. For capacities the $\delta(A,B)$ can be positive or negative, in which case we say that the capacities are supermodular or submodular.

Remark II.1. In a quantum context the requirement for probabilities, of independent alternatives in Ω, corresponds to the use of an orthonormal basis. In the case of coherent states, the non-validity of the analogue of Eq.(2), is given with the operator in Eq.(49) below, which is non-zero.

B. Ranking and derivatives of cumulative functions in Choquet integrals

The nonadditivity in capacities implies that the concept of integration needs revision. The Choquet integration is appropriate in this case. We consider a function f on the finite set Ω, which takes the real values $f(1), ..., f(N)$. We note that Choquet integrals can also be defined for functions with a continuum of real values, but in this paper we consider the finite case. We relabel this function, using a 'ranking permutation' $i = \sigma(j)$ of the indices, so that

$$f[\sigma(1)] \leq ... \leq f[\sigma(N)]. \tag{4}$$

The Choquet integral of f with respect to the capacity μ is given by

$$
\begin{aligned}
\mathcal{C}(f;\mu) &= \sum_{i=1}^{N} f[\sigma(i)]\nu_f(i) \\
\nu_f(i) &= \mu(\sigma(i),\sigma(i+1),...,\sigma(N)) - \mu(\sigma(i+1),\sigma(i+2),...,\sigma(N)); \quad i=1,...,N-1 \\
\nu_f(N) &= \mu(\sigma(N)); \quad \sum_{i=1}^{N}\nu_f(i) = 1.
\end{aligned}
\tag{5}
$$

The $\mu(\sigma(1),\sigma(2),...,\sigma(i-1))$ is a cumulative function, and

$$
\mu(\sigma(i),\sigma(i+1),...,\sigma(N)) = 1 - \mu(\sigma(1),\sigma(2),...,\sigma(i-1))
\tag{6}
$$

is a complementary cumulative function. The $\nu_f(i)$ can be viewed as 'discrete derivative' of the cumulative function. For additive capacities (additive probabilities) the derivative of the cumulative function is equal to the probability distribution:

$$
\mu(\sigma(i),\sigma(i+1),...,\sigma(N)) - \mu(\sigma(i+1),\sigma(i+2),...,\sigma(N)) = \mu(\sigma(i)); \quad i=1,...,N-1
\tag{7}
$$

but for non-additive capacities this is not true (in general $\nu_f(i) \neq \mu(\sigma(i))$). The $\nu_f(i)$ depends on $\sigma(i+1),...,\sigma(N)$ and therefore it depends on the ranking in Eq.(4). This is indicated in the notation $\nu_f(i)$ with the index f. For two functions f, g, in general $\sigma_f(j) \neq \sigma_g(j)$, and therefore $\nu_f(i) \neq \nu_g(i)$.

In a weighted average we multiply the values of a function with the corresponding probabilities. In a Choquet integral we replace the probabilities with discrete derivatives (differences) of cumulative functions.

Remark II.2. The $\mathcal{C}$ in the notation, indicates Choquet integral. Both the ranking of the function in Eq.(4), and also the derivatives of cumulative functions, play a crucial role in Choquet integrals. Both of them will be linked to non-commutativity, in a quantum context later.

If μ_1 and μ_2 are capacities, then $\mu = p\mu_1 + (1-p)\mu_2$ where $0 \leq p \leq 1$ is also a capacity, and

$$
\mathcal{C}(f;\mu) = p\mathcal{C}(f;\mu_1) + (1-p)\mathcal{C}(f;\mu_2)
\tag{8}
$$

Example II.3. *If $A \subseteq \Omega$ and μ is a capacity such that*

$$
\begin{aligned}
\mu(B) &= 1 \ \text{if} \ A \subseteq B \subseteq \Omega \\
\mu(B) &= 0 \ \text{otherwise},
\end{aligned}
\tag{9}
$$

then Eq.(5) reduces to

$$
\mathcal{C}(f;\mu) = \min[f(A)],
\tag{10}
$$

where $\min[f(A)]$ is the minimum of all the values of the function f in the subset A. Here only the aggregations of alternatives in sets $B \supseteq A$ make a contribution to the integral, because for any other set C, we have $\mu(C) = 0$. In this case, the integral is equal to the minimum of the values of the function f, in the set A. The contribution of the other values of the function f is zero, because the discrete derivatives of the corresponding cumulative function $\mu(\{\sigma(i),\sigma(i+1),...,\sigma(N)\})$, are zero.

C. Möbius transform: how to avoid double-counting

The Möbius transform is used extensively in Combinatorics, after the work by Rota[21, 22]. It is a generalization of the inclusion-exclusion principle that gives the cardinality of the union of overlaping sets. The Möbius

transform describes the overlaps between sets, and it is used to avoid the 'double-counting'. Rota generalized this to partially ordered structures.

The Möbius transform of capacities leads to the following function $\mathfrak{d}(A)$ (where $A \subseteq \Omega$):

$$\mathfrak{d}(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} \mu(B). \tag{11}$$

where $|A|$, $|B|$ are the cardinalities of these sets. For example, if $A = \{a_1, ..., a_m\}$, then

$$\begin{aligned} \mathfrak{d}(a_i) &= \mu(a_i); \quad \mathfrak{d}(a_i, a_j) = \mu(a_i, a_j) - \mu(a_i) - \mu(a_j) \\ \mathfrak{d}(a_i, a_j, a_k) &= \mu(a_i, a_j, a_k) - \mu(a_i, a_j) - \mu(a_i, a_k) - \mu(a_j, a_k) \\ &\quad + \mu(a_i) + \mu(a_j) + \mu(a_k) \end{aligned} \tag{12}$$

etc.

The inverse Möbius transform is the intuitively nice relation

$$\mu(A) = \sum_{B \subseteq A} \mathfrak{d}(B). \tag{13}$$

For sets with one element only, $\mathfrak{d}(A) = \mu(A)$. If $\Omega = \{a_1, ..., a_n\}$, Eq.(13) with $A = \Omega$ becomes

$$\sum_{i=1}^{n} \mu(a_i) + \sum_{i,j} \mathfrak{d}(a_i, a_j) + ... + \mathfrak{d}(a_1, ..., a_n) = 1. \tag{14}$$

There are $2^n - 1$ terms in this sum (there are 2^n subsets of Ω, but we exclude the empty set). Eq.(14) is important because lack of additivity means that in general

$$\sum_{i=1}^{n} \mu(a_i) \neq 1. \tag{15}$$

In the special case that the capacity $\mu(A)$ is additive (i.e., Eq.(2) holds), $\mathfrak{d}(B)$ is zero if the cardinality of B is greater or equal to 2:

$$|B| \geq 2 \quad \rightarrow \quad \mathfrak{d}(B) = 0. \tag{16}$$

Remark II.4. In a quantum context the analogue of the Möbius transforms $\mathfrak{d}$, are the operators in Eq.(59).

Lemma II.5. *The Choquet integral of Eq.(5) is given in terms of the Möbius transform $\mathfrak{d}$ of the capacities μ, as*

$$\mathcal{C}(f; \mu) = \sum_{A \subseteq \Omega} \mathfrak{d}(A) \min[f(A)]. \tag{17}$$

For all subsets A of Ω, we multiply $\mathfrak{d}(A)$ with the minimum value of the function in this subset, and we add the results.

Proof. Explicit proof is given in [11–15]. Here we only give a hint of the proof, which is based on Eqs.(8),(10). In the special case of the capacity in Eq.(9), we get

$$\begin{aligned} \mathfrak{d}(B) &= 1 \ \text{ if } \ B = A \\ \mathfrak{d}(B) &= 0 \ \text{ otherwise} \end{aligned} \tag{18}$$

and Eq.(17), reduces to Eq.(10). In this sense, Eq.(17), is simply a generalization of Eq.(10). $\qquad\square$

Proposition II.6. *The Choquet integral of Eq.(5) can be written as the sum*

$$\mathcal{C}(f;\mu) = \mathcal{C}_1(f;\mu) + \mathcal{C}_2(f;\mu) + ... + \mathcal{C}_N(f;\mu)$$

$$\mathcal{C}_1(f;\mu) = \sum \mathfrak{d}(a_i)f(a_i) = \sum \mu(a_i)f(a_i)$$

$$\mathcal{C}_2(f;\mu) = \sum \mathfrak{d}(a_i, a_j)\min[f(a_i), f(a_j)]$$

$$\text{..}$$

$$\mathcal{C}_N(f;\mu) = \mathfrak{d}(a_1, ..., a_M)\min[f(a_1), ..., f(a_M)]. \tag{19}$$

In the special case of additive capacities (Kolmogorov probabilities) $\mathcal{C}_2(f;\mu) = ... = \mathcal{C}_N(f;\mu) = 0$, and the Choquet integral is the standard weighted average $\mathcal{C}_1(f;\mu)$.

Proof. We start from Eq.(17), and we group together all terms with $\mathfrak{d}(A)$ where the cardinality of A is equal to k. This gives the $\mathcal{C}_k(f;\mathfrak{d})$. In the special case of additive capacities, all the $\mathfrak{d}(a_{i_1}, ..., a_{i_r})$ with $r \geq 2$ are zero, and consequently all the $\mathcal{C}_k(f;\mu)$ with $k \geq 2$ are equal to zero. $\square$

The Choquet integral is the weighted average given by $\mathcal{C}_1(f;\mu)$, plus the corrections of the other terms which are due to deviations from the additivity of probability. In a quantum context later, these extra corrections are due to higher order Möbius operators and are directly related to the overlaps between the coherent states.

D. Comonotonic functions and the weak additivity property of Choquet integrals

Ranking is important in Choquet integrals. In general two functions have different ranking and

$$\mathcal{C}(f + g;\mu) \neq \mathcal{C}(f;\mu) + \mathcal{C}(g;\mu). \tag{20}$$

We define comonotonic (same ranking) functions[10–15], as follows:

Definition II.7. Two functions $f(i)$ and $g(i)$ on the set Ω, are called comonotonic if the following statements which are equivalent to each other, hold:

(1) The ranking permutation $\sigma(j)$ is the same for the two functions ($\sigma_f(j) = \sigma_g(j)$). Therefore $\nu_f(i) = \nu_g(i)$. The weights $\nu_f(i)$ in Eq.(5) are the same for comonotonic functions, but they are in general different for non-comonotonic functions.

(2) For all i, j

$$[f(i) - f(j)][g(i) - g(j)] \geq 0. \tag{21}$$

Comonotonicity does not obey the transitivity property (i.e., if f_1, f_2 are comonotonic, and f_2, f_3 are also comonotonic, the f_1, f_3 might not be comonotonic).

Proposition II.8.

(1) For comonotonic functions f, g and positive a, b

$$\mathcal{C}(af + bg;\mu) = a\mathcal{C}(f;\mu) + b\mathcal{C}(g;\mu). \tag{22}$$

(2) If μ is an additive capacity (Kolmogorov probability), then for any functions f, g

$$\mathcal{C}(af + bg;\mu) = a\mathcal{C}(f;\mu) + b\mathcal{C}(g;\mu). \tag{23}$$

Proof.

(1) Comonotonic functions have the same weights $\nu_f(i)$ in Eq.(5), and this proves Eq.(22). The a, b are taken to be positive, so that the ranking is preserved.

(2) For additive capacities, all the $\mathfrak{d}(a_{i_1}, ..., a_{i_r})$ with $r \geq 2$ in Eq.(19), are zero. Therefore only the $\mathcal{C}_1(f; \mu)$ is non-zero, and then we can easily prove Eq.(23).

$\square$

For general (non-additive) capacities, additivity of the Choquet integral holds only for comonotonic functions. We refer to this as weak additivity property of Choquet integrals.

A constant function c is comonotonic with any function f and therefore

$$\mathcal{C}(f + c; \mu) = \mathcal{C}(f; \mu) + c. \tag{24}$$

E. Example

Four students A, B, C, D were examined in three modules $1, 2, 3$ and they got the following marks (in the interval $[0, 100]$):

$$
\begin{aligned}
f_A(1) &= 70; & f_A(2) &= 70; & f_A(3) &= 30 \\
f_B(1) &= 90; & f_B(2) &= 50; & f_B(3) &= 80 \\
f_C(1) &= 50; & f_C(2) &= 90; & f_C(3) &= 70 \\
f_D(1) &= 70; & f_D(2) &= 60; & f_D(3) &= 50.
\end{aligned}
\tag{25}
$$

A professor considers them as applicants for a PhD study, taking into account how close the three modules are to the topic of the Ph.D. The assumption of separability and independence of the three modules is too strong (because usually the modules overlap with each other). We adopt the weaker concepts of capacity and Choquet integrals, which allow for an aggregation to be different from the sum of its parts.

In this example, Ω is the set of the three modules $\{1, 2, 3\}$. We will calculate the Choquet integrals using the capacities:

$$
\begin{aligned}
\mu(1) &= 0.3; & \mu(2) &= 0.3; & \mu(3) &= 0.2 \\
\mu(1, 2) &= 1; & \mu(1, 3) &= 0.4; & \mu(2, 3) &= 0.4 \\
\mu(\emptyset) &= 0; & \mu(1, 2, 3) &= 1
\end{aligned}
\tag{26}
$$

It is not a requirement that the $\mu(1) + \mu(2) + \mu(3)$ should be equal to 1 (see Eq.(14)). These capacities reflect the fact that the aggregation of modules 1 and 2 is ideal for the topic of this PhD, and for this reason $\mu(1, 2) > \mu(1) + \mu(2)$ (and in fact $\mu(1, 2) = 1$). The aggregation of modules 1 and 3 is not very important for this Ph.D., and this is reflected in the $\mu(1, 3) < \mu(1) + \mu(3)$. One reason for this may be that there is overlap in the material taught in modules $1, 3$. Similar comment can be made for the aggregation of modules $2, 3$. In the quantum context later, the reason for the non-additivity is the non-zero overlap between coherent states.

For student A we have $f_A(3) \leq f_A(2) \leq f_A(1)$ and therefore

$$
\begin{aligned}
\nu_A(3) &= 1 - \mu(1, 2) = 0; \quad \nu_A(2) = \mu(1, 2) - \mu(1) = 0.7; \quad \nu_A(1) = \mu(1) = 0.3 \\
\mathcal{C}_A(f; \mu) &= f_A(3)\nu_A(3) + f_A(2)\nu_A(2) + f_A(1)\nu_A(1) = 70.
\end{aligned}
\tag{27}
$$

Since $\nu_A(3) = 0$ the lowest mark of this student does not contribute in the calculation. For student B we have $f_B(2) \leq f_B(3) \leq f_B(1)$ and therefore

$$
\begin{aligned}
\nu_B(2) &= 1 - \mu(1, 3) = 0.6; \quad \nu_B(3) = \mu(1, 3) - \mu(1) = 0.1; \quad \nu_B(1) = \mu(1) = 0.3 \\
\mathcal{C}_B(f; \mu) &= f_B(2)\nu_B(2) + f_B(3)\nu_B(3) + f_B(1)\nu_B(1) = 65.
\end{aligned}
\tag{28}
$$

For student C we have $f_C(1) \leq f_C(3) \leq f_C(2)$ and therefore

$$\nu_C(1) = 1 - \mu(2,3) = 0.6; \quad \nu_C(3) = \mu(2,3) - \mu(2) = 0.1; \quad \nu_C(2) = \mu(2) = 0.3$$
$$\mathcal{C}_C(f;\mu) = f_C(1)\nu_C(1) + f_C(3)\nu_C(3) + f_C(2)\nu_C(2) = 64. \tag{29}$$

The marks of the student D are comonotonic to those of the student A. This means that the students A, D have similar academic strengths and weaknesses, with respect to the modules $\{1, 2, 3\}$ (the analogue of this in a quantum context will be physically similar Hermitian operators). Therefore $\nu_D(3) = \nu_A(3) = 0$ and $\nu_D(2) = \nu_A(2) = 0.7$ and $\nu_D(1) = \nu_A(1) = 0.3$. It follows that

$$\mathcal{C}_D(f;\mu) = f_D(3)\nu_D(3) + f_D(2)\nu_D(2) + f_D(1)\nu_D(1) = 63. \tag{30}$$

The Choquet integral is a figure of merit, which orders the students as $A \succ B \succ C \succ D$. Here $A \succ B$ means that A is more (or equally) preferable for Ph.D. than B.

We note that the weight of the same subject is different for different students. For example, $\nu_A(2) = 0.5$, $\nu_B(2) = 0.6$, $\nu_C(2) = 0.3$. This is related to the fact that the three modules are not independent. Cumulative rather than separable weights are used in the calculation. The ranking in Eq.(4) plays an important role in determining the values of ν.

The Möbius transform of the capacities in Eq.(26) gives

$$\mathfrak{d}(1,2,3) = 0; \quad \mathfrak{d}(1,2) = 0.4; \quad \mathfrak{d}(1,3) = -0.1; \quad \mathfrak{d}(2,3) = -0.1$$
$$\mathfrak{d}(1) = 0.3; \quad \mathfrak{d}(2) = 0.3; \quad \mathfrak{d}(3) = 0.2 \tag{31}$$

Then using Eq.(17) we find the same results as above. We present explicitly the calculation for one of them. Taking into account that

$$\min\{f_A(1), f_A(2)\} = f_A(2); \quad \min\{f_A(1), f_A(3)\} = f_A(3); \quad \min\{f_A(2), f_A(3)\} = f_A(3) \tag{32}$$

we get

$$\begin{aligned}
\mathcal{C}_A(f;\mathfrak{d}) &= \mathcal{C}_A^{(1)}(f;\mathfrak{d}) + \mathcal{C}_A^{(2)}(f;\mathfrak{d}) + \mathcal{C}_A^{(3)}(f;\mathfrak{d}) = 70 \\
\mathcal{C}_A^{(1)}(f;\mathfrak{d}) &= \mathfrak{d}(1)f_A(1) + \mathfrak{d}(2)f_A(2) + \mathfrak{d}(3)f_A(3) = 48 \\
\mathcal{C}_A^{(2)}(f;\mathfrak{d}) &= \mathfrak{d}(1,2)f_A(2) + \mathfrak{d}(1,3)f_A(3) + \mathfrak{d}(2,3)f_A(3) = 22 \\
\mathcal{C}_A^{(3)}(f;\mathfrak{d}) &= \mathfrak{d}(1,2,3)f_A(3) = 0,
\end{aligned} \tag{33}$$

which is the same result as in Eq.(27).

We note that if we use the 'standard average' we find

$$\mathfrak{M}_A = \frac{170}{3}; \quad \mathfrak{M}_B = \frac{220}{3}; \quad \mathfrak{M}_C = \frac{210}{3}; \quad \mathfrak{M}_D = \frac{180}{3}. \tag{34}$$

and this leads to the ordering $B \sqsupset C \sqsupset D \sqsupset A$, where $\sqsupset$ is the ordering according to the 'standard averaging'.

III. CUMULATIVE PROJECTORS AND MÖBIUS OPERATORS

A. Coherent states

We consider a quantum system $\Sigma(d)$ with variables in $\mathbb{Z}(d)$, and d-dimensional Hilbert space $H(d)$. We also consider the orthonormal basis of 'position states' $|X; n\rangle$, and through the Fourier transform F, the basis of

momentum states $|P;n\rangle$[2, 3]:

$$F = d^{-1/2} \sum_m \omega(mn)|X;n\rangle\langle X:m|; \quad \omega(\alpha) = \exp\left(\frac{i2\pi\alpha}{d}\right)$$

$$|P;n\rangle = F|X;n\rangle; \quad m,n,\alpha \in \mathbb{Z}(d). \tag{35}$$

Displacement operators in the $\mathbb{Z}(d) \times \mathbb{Z}(d)$ phase space, are given by

$$D(\alpha,\beta) = Z^\alpha X^\beta \omega(-2^{-1}\alpha\beta); \quad Z = \sum_m \omega(m)|X;m\rangle\langle X;m|; \quad X = \sum_m |X;m+1\rangle\langle X;m| \tag{36}$$

The $\{D(\alpha,\beta)\omega(\gamma)\}$ form the Heisenberg-Weyl group of displacements in this system. The formalism of finite quantum systems, is slightly different in the cases of odd and even d. The factor 2^{-1} above, is an element of $\mathbb{Z}(d)$, and it exists only for odd d. Below we assume that the dimension d is an odd integer.

Acting with $D(\alpha,\beta)$ on a (normalized) fiducial vector $|\eta\rangle$, we get the d^2 coherent states[23, 24]:

$$|C;\alpha,\beta\rangle = D(\alpha,\beta)|\eta\rangle; \quad |\eta\rangle = \sum_m \eta_m|X;m\rangle; \quad \sum_m |\eta_m|^2 = 1. \tag{37}$$

The X,P,C in the notation are not variables, but they simply indicate position states, momentum states and coherent states. We call Ω the set of the d^2 coherent states:

$$\Omega = \{|C;\alpha,\beta\rangle \mid \alpha,\beta \in \mathbb{Z}(d)\}. \tag{38}$$

The set Ω is invariant under displacement transformations.

Let $\Pi(\alpha,\beta)$ be the projector to the one-dimensional subspace $H(\alpha,\beta)$ that contains the coherent states $|C;\alpha,\beta\rangle$. Then

$$\frac{1}{d}\sum_{\alpha,\beta}\Pi(\alpha,\beta) = \mathbf{1}; \quad \Pi(\alpha,\beta) = |C;\alpha,\beta\rangle\langle C;\alpha,\beta|$$

$$D(\gamma,\delta)\Pi(\alpha,\beta)D^\dagger(\gamma,\delta) = \Pi(\alpha+\gamma,\beta+\delta) \tag{39}$$

The term 'coherent states' refers to these two properties. They are the analogue of the harmonic oscillator coherent states [25–27], in the context of quantum systems with finite-dimensional Hilbert space.

Let $\mathcal{M}_d$ be the set of $d \times d$ Hermitian positive semidefinite matrices, and $\mathcal{N}_d \subset \mathcal{M}_d$ the set of $d \times d$ density matrices. For $\theta \in \mathcal{M}_d$, the Q-function is given by

$$Q(\alpha,\beta \mid \theta) = \frac{1}{d}\mathrm{Tr}[\Pi(\alpha,\beta)\theta]; \quad \sum_{\alpha,\beta} Q(\alpha,\beta \mid \theta) = \mathrm{Tr}\theta, \tag{40}$$

and the P-function by

$$\theta = \sum_{\alpha,\beta} P(\alpha,\beta \mid \theta)\Pi(\alpha,\beta); \quad \sum_{\alpha,\beta} P(\alpha,\beta \mid \theta) = \mathrm{Tr}\theta. \tag{41}$$

If $\theta_{mn} = \langle X;m|\theta|X;n\rangle$ then

$$Q(\alpha,\beta \mid \theta) = \sum \theta_{mn} A(m,n;\alpha,\beta); \quad A(m,n;\alpha,\beta) = \frac{1}{d}\langle C;\alpha,\beta|X;m\rangle\langle X;n|C;\alpha,\beta\rangle \tag{42}$$

The $A(m,n;\alpha,\beta)$ is a $d^2 \times d^2$ matrix, and the fiducial vector should be such that its determinant is non-zero. Then Eq.(42) is a system with d^2 equations and if the $Q(\alpha,\beta \mid \theta)$ are known we can calculate the θ_{mn}, and vice-versa.

Wehrl entropy for the Q-function of density matrices: For $\theta \in \mathcal{M}_d$, we define the $\widetilde{\theta} = \theta/\mathrm{Tr}\theta \in \mathcal{N}_d$, which can be viewed as a density matrix. Its Wehrl entropy[1] is given by

$$E(\widetilde{\theta}) = -\sum_{\alpha,\beta} Q(\alpha, \beta \mid \widetilde{\theta}) \log Q(\alpha, \beta \mid \widetilde{\theta}); \qquad \sum_{\alpha,\beta} Q(\alpha, \beta \mid \widetilde{\theta}) = 1. \tag{43}$$

Its maximum value is $d \log d$. Under any permutation $(\gamma, \delta) = \sigma(\alpha, \beta)$ of the indices of the Q-function, the Wehrl entropy $E(\widetilde{\theta})$, does not change:

$$E(\widetilde{\theta}) = -\sum Q(\alpha, \beta \mid \widetilde{\theta}) \log Q(\alpha, \beta \mid \widetilde{\theta}) = -\sum Q[\sigma(\alpha, \beta) \mid \widetilde{\theta}] \log Q[\sigma(\alpha, \beta) \mid \widetilde{\theta}] \tag{44}$$

Therefore the Wehrl entropy does not tell us, for which coherent states we get high (or low) value of the Q-function. The Wehrl entropy shows whether the Q-function is uniform or concentrated in a few coherent states, but in the latter case it does not show where it is concentrated. Depending on the application, this might be a desirable or undesirable property of the Wehrl entropy. This is also seen by the fact that under displacement transformations, the Wehrl entropy does not change:

$$E\left[D(\alpha, \beta)\widetilde{\theta}D^{\dagger}(\alpha, \beta)\right] = E(\widetilde{\theta}). \tag{45}$$

We stress that the $Q(\alpha, \beta \mid \widetilde{\theta})$ are not probabilities, because the coherent states overlap with each other (d^2 coherent states in a d-dimensional space). Related to this, is that the distribution $Q(\alpha, \beta \mid \widetilde{\theta})$ can not be very narrow, and consequently the Wehrl entropy $E(\widetilde{\theta})$ is greater than a certain value (which in the harmonic oscillator case is equal to one[28]). The motivation for introducing Choquet integrals later, is to quantify and elucidate the effects of these overlaps between the coherent states.

B. Two-dimensional cumulative projectors

We consider the two-dimensional space $H(\alpha_1, \beta_1; \alpha_2, \beta_2)$ that contains all superpositions $\kappa|C; \alpha_1, \beta_1\rangle + \lambda|C; \alpha_2, \beta_2\rangle$:

$$H(\alpha_1, \beta_1; \alpha_2, \beta_2) = \mathrm{span}[H(\alpha_1, \beta_1) \cup H(\alpha_2, \beta_2)] \tag{46}$$

In the language of lattices[29–31] this is the disjunction of the one dimensional spaces $H(\alpha_1, \beta_1)$ and $H(\alpha_2, \beta_2)$. We note that the conjuction of these spaces $H(\alpha_1, \beta_1) \cap H(\alpha_2, \beta_2)$ contains only the zero vector.

We denote the projector to the space $H(\alpha_1, \beta_1; \alpha_2, \beta_2)$ as $\Pi(\alpha_1, \beta_1; \alpha_2, \beta_2)$ or if there is no danger of confusion simply as $\Pi(1, 2)$. The $\Pi(1, 2)$ can be calculated with the Gram-Schmidt orthogonalization method, where we take the component of $|C; \alpha_2, \beta_2\rangle$ which is perpendicular to $|C; \alpha_1, \beta_1\rangle$, and we normalize it into a vector with length 1. We express this in terms of projectors as

$$\begin{aligned}
\Pi(1, 2) &= \Pi(1) + \varpi(2|1) \\
\varpi(2|1) &= \frac{\Pi^{\perp}(1)\Pi(2)\Pi^{\perp}(1)}{\mathrm{Tr}[\Pi^{\perp}(1)\Pi(2)]} \\
\Pi^{\perp}(1) &= \mathbf{1} - \Pi(1).
\end{aligned} \tag{47}$$

We call the $\Pi(1, 2)$ cumulative projectors because they project into two-dimensional spaces, and therefore the corresponding probabilities take a range (two) values. The $\varpi(2|1) = \Pi(1, 2) - \Pi(1)$ can be viewed as a discrete derivative (difference) of the cumulative projectors. In additive (Kolmogorov) probabilities, the derivative of the cumulative distributions are the probability distributions. This is not true for capacities (non-additive probabilities), precisely because additivity does not hold. Here this is the fact that the $\varpi(2|1)$ is different from

$\Pi(2)$. From a physical point of view, a measurement with the projector $\Pi^\perp(1)$ (which projects to the orthogonal complement of $H(\alpha_1, \beta_1)$), on the coherent state $|C; \alpha_2, \beta_2\rangle$ (which is described with the density matrix $\Pi(2)$), will collapse it into the $\varpi(2|1)$ with probability $\text{Tr}[\Pi^\perp(1)\Pi(2)]$.

Lemma III.1.

$$D(\gamma, \delta)\varpi(\alpha_2, \beta_2|\alpha_1, \beta_1)D^\dagger(\alpha, \beta) = \varpi(\alpha_2 + \gamma, \beta_2 + \delta|\alpha_1 + \gamma, \beta_1 + \delta) \tag{48}$$

Proof. We multiply both sides of the second of Eq.(47) by $D(\gamma, \delta)$ on the left and $D^\dagger(\gamma, \delta)$ on the right, taking into account Eq.(39). $\square$

In analogy with Eq.(1) we consider the following operator:

$$\mathfrak{D}(1, 2) = \Pi(1, 2) - \Pi(1) - \Pi(2); \quad \text{Tr}[\mathfrak{D}(1, 2)] = 0. \tag{49}$$

A projector to the space $H(\alpha_1, \beta_1) \cap H(\alpha_2, \beta_2)$ should also be added to the right hand side, but as we explained earlier it is zero. The trace of this operator with a density matrix ρ converts the projectors into probabilities, and in this sense the $\mathfrak{D}(1, 2)$ is analogous to $\delta(A, B)$ in Eq.(1). Unlike $\delta(A, B)$, the $\mathfrak{D}(1, 2)$ is in general non-zero, and quantifies deviations from the additivity of probability due to the overlapping nature of coherent states. The resolution of the identity in terms of coherent states, shows that in the corresponding sum these overlaps cancel each other. The following proposition shows that something similar happens with the $\mathfrak{D}(1, 2)$ operators:

Proposition III.2. *For fixed α_i, β_i:*

$$\sum_{\kappa, \lambda} \mathfrak{D}(\alpha_1 + \kappa, \beta_1 + \lambda; \alpha_2 + \kappa, \beta_2 + \lambda) = 0. \tag{50}$$

Proof. Using the resolution of the identity for coherent states, it has been proved (Eq.(119) in ref[3]) that for any operator χ

$$\frac{1}{d}\sum_{\kappa, \lambda} D(\kappa, \lambda)\chi[D(\kappa, \lambda)]^\dagger = \mathbf{1}\text{Tr}\chi. \tag{51}$$

We use this with $\chi = \Pi(\alpha_1, \beta_1; \alpha_2, \beta_2)$, in conjuction with the relation

$$D(\kappa, \lambda)\Pi(\alpha_1, \beta_1; \alpha_2, \beta_2)[D(\kappa, \lambda)]^\dagger = \Pi(\alpha_1 + \kappa, \beta_1 + \lambda; \alpha_2 + \kappa, \beta_2 + \lambda), \tag{52}$$

and we prove that

$$\frac{1}{2d}\sum_{\kappa, \lambda} \Pi(\alpha_1 + \kappa, \beta_1 + \lambda; \alpha_2 + \kappa, \beta_2 + \lambda) = \mathbf{1}. \tag{53}$$

This together with the resolution of the identity for $\Pi(\alpha_1, \beta_1)$ and $\Pi(\alpha_2, \beta_2)$ proves the proposition. $\square$

Remark III.3. The $\mathfrak{D}(1, 2)$ are a special case of more general operators $\mathfrak{D}(H_1, H_2)$ associated with subspaces H_1 and H_2 of $H(d)$, which we have studied in [9]. We have proved there that the commutator of the projectors to these subspaces $[\Pi(H_1), \Pi(H_2)]$ is related to $\mathfrak{D}(H_1, H_2)$, through the relation:

$$[\Pi(H_1), \Pi(H_2)] = \mathfrak{D}(H_1, H_2)[\Pi(H_1) - \Pi(H_2)]. \tag{54}$$

This relation links non-commutativity with non-additive probabilities. For non-commuting projectors, the $\text{Tr}[\rho\mathfrak{D}(H_1, H_2)]$ (where ρ is a density matrix) is non-zero, and we cannot interpret the corresponding probabilities as additive (Kolmogorov) probabilities. In [9], we interpreted quantum probabilities as non-additive (Dempster-Shafer) probabilities, for which the $\delta(A, B)$ of Eq.(1) is in general non-zero.

C.   Multi-dimensional cumulative projectors

We order the coherent states in an arbitrary way and we label them as $|C;\alpha_1,\beta_1\rangle,...,|C;\alpha_{d^2},\beta_{d^2}\rangle$. The formalism in this section depends on this ordering. In the Choquet integrals, the Q-function of an operator θ will define the ordering, as discussed in the next section.

We introduce inductively the space $H(\alpha_i,\beta_i;...;\alpha_{d^2},\beta_{d^2})$ that contains all superpositions of the $d^2-(i-1)$ coherent states $|C;\alpha_i,\beta_i\rangle,...,|C;\alpha_{d^2},\beta_{d^2}\rangle$. We start from d^2 and use 'reverse order' because this is consistent with the ascending ordering in Eq.(4) (and Eq.(66) later), which is standard practice in the Choquet integrals literature. As we go from the space $H(\alpha_{i+1},\beta_{i+1};...;\alpha_{d^2},\beta_{d^2})$ to the space $H(\alpha_i,\beta_i;...;\alpha_{d^2},\beta_{d^2})$, there are two cases:

- The coherent state $|C;\alpha_i,\beta_i\rangle$ is not a linear combination of the coherent states $|C;\alpha_{i+1},\beta_{i+1}\rangle,...,|C;\alpha_{d^2},\beta_{d^2}\rangle$. The projector to the space $H(\alpha_i,\beta_i;...;\alpha_{d^2},\beta_{d^2})$ is

$$\Pi(i,...,d^2)=\Pi(i+1,...,d^2)+\varpi(i|i+1,...,d^2);\qquad i=1,...,d^2$$
$$\varpi(i|i+1,...,d^2)=\frac{\Pi^\perp(i+1,...,d^2)\Pi(i)\Pi^\perp(i+1,...,d^2)}{\mathrm{Tr}[\Pi^\perp(i+1,...,d^2)\Pi(i)]}$$
$$\Pi^\perp(i,...,d^2)=\mathbf{1}-\Pi(i,...,d^2)$$
$$\Pi(i+1,...,d^2)\varpi(i|i+1,...,d^2)=0. \tag{55}$$

  The denominator in this case is different than zero, and the dimension of the space $H(\alpha_i,\beta_i;...;\alpha_{d^2},\beta_{d^2})$ is equal to the dimension of the space $H(\alpha_{i+1},\beta_{i+1};...;\alpha_{d^2},\beta_{d^2})$ plus one. The Gram-Schmidt orthogonalization method is used here. The algorithm can also be implemented with the QR factorization of matrices [32], and is available in computer libraries (eg, in MATLAB). From a physical point of view, a measurement with the projector $\Pi^\perp(i+1,...,d^2)$ (which projects to the orthogonal complement of $H(\alpha_{i+1},\beta_{i+1};...;\alpha_{d^2},\beta_{d^2})$), on the coherent state $|C;\alpha_i,\beta_i\rangle$ (which is described with the density matrix $\Pi(i)$), will collapse it into the $\varpi(i|i+1,...,d^2)$ with probability $\mathrm{Tr}[\Pi^\perp(i+1,...,d^2)\Pi(i)]$.

- The coherent state $|C;\alpha_i,\beta_i\rangle$ is a linear combination of the coherent states $|C;\alpha_{i+1},\beta_{i+1}\rangle,...,|C;\alpha_{d^2},\beta_{d^2}\rangle$. In this case $\varpi(i|i+1,...,d^2)=0$ and the dimension of the space $H(\alpha_i,\beta_i;...;\alpha_{d^2},\beta_{d^2})$ is equal to the dimension of the space $H(\alpha_{i+1},\beta_{i+1};...;\alpha_{d^2},\beta_{d^2})$.

There are d^2 projectors $\varpi_\theta(i|i+1,...,d^2)$ in the d-dimensional space $H(d)$ (with $\varpi(d^2)=\Pi(d^2)$). d^2-d of these projectors are equal to zero, and the rest form an orthogonal and complete set of projectors in $H(d)$:

$$\Pi(i,...,d^2)=\varpi(i|i+1,...,d^2)+...+\varpi(d^2-1|d^2)+\varpi(d^2)$$
$$\sum_{i=1}^{d^2}\varpi(i|i+1,...,d^2)=\mathbf{1}. \tag{56}$$

Relations similar to those in lemma III.1 can also be proved for the projectors $\varpi(i|i+1,...,d^2)$.

Coherent states with a generic fiducial vector: A fiducial vector is called 'generic', if any d of the corresponding coherent states are linearly independent. In this case any set of d or more coherent states is a total set in $H(d)$, i.e., there is not vector which is orthogonal to all these coherent states. Then for any set $A=\{i_1,...,i_d\}$ with d indices

$$\Pi(i_1,...,i_d)=\mathbf{1};\qquad \varpi(j|i_1,...,i_d)=0;\qquad j\in\{1,...,d^2\}-A$$
$$\varpi(i_d|i_{d-1},...,i_1)+...+\varpi(i_2|i_1)+\varpi(i_1)=\mathbf{1}. \tag{57}$$

Apart from position and momentum states, 'most' of the other vectors can be used as generic fiducial vectors.

For simplicity all our examples later, are in the 3-dimensional space $H(3)$, and use coherent states $D(\alpha, \beta)|\eta\rangle$ with respect to the generic fiducial vector

$$|\eta\rangle = \frac{1}{\sqrt{14}} \left(|X; 0\rangle + 2|X; 1\rangle + 3|X; 2\rangle \right). \tag{58}$$

D. Möbius operators

The Möbius transform of Eqs(11),(13), in the present context provides a systematic method for the expression of the $\mathfrak{D}$-operators in terms of the cumulative Π-projectors. If $A = \{(\alpha_1, \beta_1), (\alpha_2, \beta_2), ...\}$ (where $\alpha_i, \beta_i \in \mathbb{Z}(d)$) is a set of pairs of indices, we use the shorthand notation $\mathfrak{D}(A)$ for $\mathfrak{D}(1, 2, ...)$, and $\Pi(A)$ for $\Pi(1, 2, ...)$. $\mathfrak{D}(B)$ is related to the various projectors through the Möbius transform [21, 22]

$$\mathfrak{D}(B) = \sum_{A \subseteq B} (-1)^{|A| - |B|} \Pi(A). \tag{59}$$

For sets with only one pair $\mathfrak{D}(A) = \Pi(A)$. A simple example of this, with two coherent states, is in Eq.(49). We refer to $\mathfrak{D}(B)$ as the Möbius operators. The trace of these operators with a density matrix, lead to probabilistic relations which quantify deviations from the additivity of probability. The inverse Möbius transform is

$$\Pi(A) = \sum_{B \subseteq A} \mathfrak{D}(B). \tag{60}$$

In Eq.(60) we put $A = \mathbb{Z}(d) \times \mathbb{Z}(d)$ (the set of all (α_i, β_i)), and we get

$$\sum_{i=1}^{d^2} \Pi(i) + \sum_{i,j} \mathfrak{D}(i, j) + ... + \mathfrak{D}(1, ..., d^2) = \mathbf{1}. \tag{61}$$

This can be viewed as a kind of weak resolution of the identity, where the 'Möbius operators' eliminate the 'double counting' in the sum of the non-orthogonal projectors. The term 'weak'is used to indicate that in addition to the projectors, the Möbius operators are needed.

This inverse Möbius transform involves the d^2 projectors $\Pi(i)$, and all the Möbius $\mathfrak{D}$-operators, whose role is to remove the overlaps between the $\Pi(i)$ so there is no double-counting. Eq.(61) is the quantum analogue of Eq.(14). From Eqs(39), (61) it follows that

$$\sum_{i,j} \mathfrak{D}(i, j) + ... + \mathfrak{D}(1, ..., d^2) = (1 - d)\mathbf{1}. \tag{62}$$

The amount of double counting in the sum $\sum \Pi(i) = d\mathbf{1}$ is $(d-1)\mathbf{1}$, and it is cancelled by the above sum of Möbius $\mathfrak{D}$-operators.

In the case of coherent states with a generic fiducial vector, we insert in Eq.(60) any set with d pairs of indices, $A = \{i_1, ..., i_d\}$, and we get the following inverse Möbius transform that involves only d of the d^2 coherent states, and the corresponding Möbius operators.

$$\sum_{j=1}^{d} \Pi(i_j) + \sum_{i_j, i_k} \mathfrak{D}(i_j, i_k) + ... + \mathfrak{D}(i_1, ..., i_d) = \mathbf{1}. \tag{63}$$

Remark III.4. The trace of the projectors $\Pi(A)$ times a density matrix, gives capacities. In this sense, the projectors $\Pi(A)$ are the quantum analogue of the capacities μ in the classical case. The $\Pi(A \cup B) \neq \Pi(A) + \Pi(B)$ corresponds to the non-additivity of capacities. The operators $\mathfrak{D}$ are the quantum analogue of the $\mathfrak{d}$ in the classical case.

Example III.5. *In the three-dimensional space $H(3)$ we consider coherent states with a generic fiducial vector. For any triplet of indices i, j, k (from $1, ..., 9$) we consider the Möbius operators:*

$$\mathfrak{D}(i) = \Pi(i); \quad \mathfrak{D}(i, j) = \Pi(i, j) - \Pi(i) - \Pi(j)$$
$$\mathfrak{D}(i, j, k) = \mathbf{1} - \Pi(i, j) - \Pi(i, k) - \Pi(j, k) + \Pi(i) + \Pi(j) + \Pi(k) \tag{64}$$

If A is a set with three of the indices $1, ..., 9$, then

$$\sum_i \Pi(i) + \sum_{i,j} \mathfrak{D}(i, j) + \mathfrak{D}(i, j, k) = \mathbf{1}; \quad i, j, k \in A. \tag{65}$$

This involves 3 (from the total of 9) coherent states, and the corresponding Möbius operators.

IV.   THE DISCRETE CHOQUET INTEGRAL FOR THE Q-FUNCTION

The formalism below is presented with the Q-function of operators $\theta \in \mathcal{M}_d$, but it can also be used with the P-function, for operators θ with non-negative P-function. The formalism can be extended to the more general case where Q and P take all real values (i.e., all Hermitian operators), but we do not discuss this in the present paper.

We relabel the $Q(\alpha, \beta \mid \theta)$ as $Q(i \mid \theta)$ $(i = 1, ..., d^2)$ so that

$$0 \leq Q(1 \mid \theta) \leq Q(2 \mid \theta) \leq ... \leq Q(d^2 \mid \theta). \tag{66}$$

We use here a ranking permutation

$$i = \sigma(\alpha, \beta \mid \theta), \tag{67}$$

of the d^2 indices $(\alpha, \beta) \in \mathbb{Z}(d) \times \mathbb{Z}(d)$ which depends on the operator θ. Accordingly, we relabel the subspaces $H(\alpha, \beta)$ as $H_\theta(i)$, and the projectors $\Pi(\alpha, \beta)$ as $\Pi_\theta(i)$. The index θ indicates that the labelling depends on θ (on the reordering in Eq.(67)). We note here that for large d, the ordering of the $Q(\alpha, \beta | \theta)$ can be a practically difficult problem, but there are computer programmes which do this (e.g., in MATLAB).

In analogy to the classical case in Eq.(5), we introduce the Choquet integral $\mathcal{C}_Q(\theta)$ of the Q-function of θ, as

$$\mathcal{C}_Q(\theta) = \sum_{i=1}^{d^2} dQ(i \mid \theta) \varpi_\theta(i | i + 1, ..., d^2), \tag{68}$$

where $\varpi_\theta(i | i + 1, ..., d^2)$ are the discrete derivatives (differences) of the cumulative projectors

$$\varpi_\theta(i | i + 1, ..., d^2) = \Pi_\theta(i; i + 1; ...; d^2) - \Pi_\theta(i + 1; i + 2; ...; d^2) \tag{69}$$

These projectors are the same as Eq.(55), but here the labelling depends on the ranking of the Q-function of θ. Since $d^2 - d$ of the projectors $\varpi_\theta(i | i + 1, ..., d^2)$ are zero, it follows that only d of the d^2 values of $Q(i \mid \theta)$, contribute to the $\mathcal{C}_Q(\theta)$. The $\mathcal{C}_Q(\theta)$ is a Hermitian positive semidefinite operator with eigenprojectors the $\varpi_\theta(i | i + 1, ..., d^2)$ (the ones which are non-zero), and eigenvalues the corresponding $dQ(i \mid \theta)$. There is a finite number of sets of projectors $\{\varpi_\theta(i | i + 1, ..., d^2)\}$ (given in Eqs(101),(102) below), and therefore the set of all $\mathcal{C}_Q(\theta)$ is a subset $\mathcal{C}_d$ of $\mathcal{M}_d$. The Choquet integral is a map from $\mathcal{M}_d$ to $\mathcal{C}_d$.

In the case of coherent states with generic fiducial vectors, only the d highest values of $Q(i \mid \theta)$ enter in Eq.(68). In the rest of the paper we consider generic fiducial vectors, and

$$\mathcal{C}_Q(\theta) = \sum_{i=d^2-d+1}^{d^2} dQ(i \mid \theta) \varpi_\theta(i | i + 1, ..., d^2) \tag{70}$$

We refer to the $|C; \alpha_i, \beta_i\rangle$, $\Pi(\alpha_i, \beta_i)$, $Q(i \mid \theta)$ with $i = d^2 - d + 1, ..., d$, which enter in Eq.(70), as dominant coherent states, dominant projectors and dominant values of the Q-function, for the operator θ. We also refer to the $|C; \alpha_i, \beta_i\rangle$, $\Pi(\alpha_i, \beta_i)$, $Q(i \mid \theta)$ in the 'tail' $i = 1, ..., d^2 - d + 1$, as inferior coherent states, inferior projectors and inferior values of the Q-function, for the operator θ.

If two of the dominant values of the Q-function, $Q(i \mid \theta)$ and $Q(i + 1 \mid \theta)$, are equal to each other, there are two different orderings of the corresponding coherent states that can be used in Eq.(70), and they both lead to the same result. Indeed, the contribution of these two terms to $\mathcal{C}_Q(\theta)$, is

$$dQ(i \mid \theta)\varpi_\theta(i|i + 1, ..., d^2) + dQ(i + 1 \mid \theta)\varpi_\theta(i + 1|i + 2, ..., d^2)$$
$$= dQ(i \mid \theta)[\Pi_\theta(i; i + 1; ...; d^2) - \Pi_\theta(i + 2; ...; d^2)] \tag{71}$$

The $\Pi_\theta(i; i + 1; ...; d^2)$ does not change if we swap the two coherent states ranked with i and $i + 1$. In this case there is a degeneracy in the eigenvalues of $\mathcal{C}_Q(\theta)$. $Q(i \mid \theta)$ and $Q(i + 1 \mid \theta)$ are two eigenvalues equal to each other, and $\Pi_\theta(i; i + 1; ...; d^2) - \Pi_\theta(i + 2; ...; d^2)$ is the corresponding eigenprojector to a two dimensional space.

It is easily seen that $\mathcal{C}_Q(a\theta) = a\mathcal{C}_Q(\theta)$ for $a \geq 0$. In general $\mathcal{C}_Q(\theta) + \mathcal{C}_Q(\phi) \neq \mathcal{C}_Q(\theta + \phi)$. The question under what conditions we have additivity, leads naturally to the concept of comonotonicity, which we discussed in a classsical context earlier, and which is discussed in a quantum context later.

Proposition IV.1. $\mathcal{C}_Q(\theta)$ can be written as

$$\mathcal{C}_Q(\theta) = \mathcal{C}_{Q,1}(\theta) + \mathcal{C}_{Q,2}(\theta) + ... + \mathcal{C}_{Q,d^2}(\theta)$$
$$\mathcal{C}_{Q,1}(\theta) = d \sum \Pi(i)Q(i| \theta)$$
$$\mathcal{C}_{Q,2}(\theta) = d \sum \mathfrak{D}(i, j) \min\{Q(i| \theta), Q(j| \theta)\}$$
$$\mathcal{C}_{Q,3}(\theta) = d \sum \mathfrak{D}(i, j, k) \min\{Q(i| \theta), Q(j| \theta), Q(k| \theta)\}$$
$$\text{..}$$
$$\mathcal{C}_{Q,d^2}(\theta) = d\mathfrak{D}(1, ..., d^2)Q(1| \theta) \tag{72}$$

Proof. We start from Eq.(68), and we group together all terms that involve the operators $\mathfrak{D}$ with k Hilbert spaces. This gives the $\mathcal{C}_{Q,k}(\theta)$. $\qquad\square$

Choquet integrals, are designed for cases where the various alternatives are not independent, but they overlap with each other. The Möbius transforms studied in section 124, quantify these overlaps. The term $d\mathfrak{D}(i, j) \min\{Q(i| \theta), Q(j| \theta)\}$ is a 'correction' related to the overlap between two coherent states i, j. The term $d\mathfrak{D}(i, j, k) \min\{Q(i| \theta), Q(j| \theta), Q(k| \theta)\}$ is a 'correction' related to the overlap between three coherent states i, j, k, etc. Adding all of them together, we remove the double-counting due to overlaps between the coherent states.

The following proposition gives the $\mathcal{C}_Q(\theta)$ in some special cases.

Proposition IV.2.

(1) If the d dominant values of $Q(i \mid \theta)$ are equal to each other, then

$$\mathcal{C}_Q(\theta) = d \max\{Q(\alpha, \beta|\theta)\}\mathbf{1}. \tag{73}$$

(2) Let $\theta = \sum_m \lambda_m |X; m\rangle\langle X; m|$ with $\lambda_m \geq 0$ (so that it is positive semidefinite operator). Then

$$\mathcal{C}_Q(\theta) = d \max\{Q(\alpha, \beta|\theta)\}\mathbf{1}. \tag{74}$$

Similar result holds for $\theta = \sum_m \lambda_m |P; m\rangle\langle P; m|$.

(3)

$$\mathcal{C}_Q(\mathbf{1}) = \mathbf{1}; \quad \mathcal{C}_Q(\theta + \lambda\mathbf{1}) = \mathcal{C}_Q(\theta) + \lambda\mathbf{1}. \tag{75}$$

Proof.

(1) This follows immediately from Eq.(70) and the fact that the $\varpi_\theta(i|i+1,...,d^2)$ are an orthogonal and complete set of projectors.

(2) For $\theta = \sum_m \lambda_m |X;m\rangle\langle X;m|$, the $D(\alpha,\beta)\theta[D(\alpha,\beta)]^\dagger$ do not depend on α, and consequently the d dominant values of $Q(\alpha,\beta|\theta)$ are equal to each other. From this follows Eq.(74).

(3) This follows from the fact that $Q(\alpha,\beta \mid \mathbf{1}) = \frac{1}{d}$.

$\square$

One of the applications of the Choquet integral is that it provides bounds for various physical quantities. The following proposition provides bounds to $\mathrm{Tr}(\theta)$, $\mathrm{Tr}(\rho\theta)$ (where ρ is a density matrix), and $\mathrm{Tr}(\theta\phi)$, in terms of $\mathrm{Tr}[\mathcal{C}_Q(\theta)]$, $\mathrm{Tr}[\mathcal{C}_Q(\phi)]$. It also shows that $\mathrm{Tr}[\mathcal{C}_Q(\theta)]$ is a convex function. We note that the calculation of $\mathrm{Tr}[\mathcal{C}_Q(\theta)]$ only requires the calculation of the Q-function and its ranking in Eq.(66). It does not require the calculation of the projectors $\varpi_\theta(i|i+1,...,d^2)$. Indeed, from Eq.(70) it follows that

$$\mathrm{Tr}[\mathcal{C}_Q(\theta)] = \sum_{i=d^2-d+1}^{d^2} dQ(i \mid \theta). \tag{76}$$

Proposition IV.3. *Let $\theta, \phi \in \mathcal{M}_d$.*

(1) For $\theta \neq 0$

$$\frac{1}{d}\mathrm{Tr}[\mathcal{C}_Q(\theta)] < \mathrm{Tr}(\theta) \leq \mathrm{Tr}[\mathcal{C}_Q(\theta)]. \tag{77}$$

For $\theta = \mathbf{1}$ the right hand side inequality becomes equality. The left hand side is a strict inequality.

(2) For any density matrix ρ

$$\mathrm{Tr}(\rho\theta) \leq \mathrm{Tr}[\mathcal{C}_Q(\theta)]. \tag{78}$$

(3)

$$\mathrm{Tr}(\theta\phi) \leq \mathrm{Tr}[\mathcal{C}_Q(\theta)]\mathrm{Tr}[\mathcal{C}_Q(\phi)]. \tag{79}$$

(4) $\mathrm{Tr}[\mathcal{C}_Q(\theta)]$ is a convex function:

$$\mathrm{Tr}[\mathcal{C}_Q(a\theta + (1-a)\phi)] \leq a\mathrm{Tr}[\mathcal{C}_Q(\theta)] + (1-a)\mathrm{Tr}[\mathcal{C}_Q(\phi)]; \quad 0 \leq a \leq 1. \tag{80}$$

Proof.

(1)

$$\mathrm{Tr}[\mathcal{C}_Q(\theta)] - \mathrm{Tr}(\theta) = (d-1)\sum_{i=d^2-d+1}^{d} Q(i|\theta) - \sum_{i=1}^{d^2-d} Q(i|\theta) \tag{81}$$

There are $d^2 - d$ terms in both of these sums and any term in the first sum is greater or equal to any term in the second sum. This proves that $\text{Tr}(\theta) \leq \text{Tr}[\mathcal{C}_Q(\theta)]$. Also

$$d\text{Tr}(\theta) - \text{Tr}[\mathcal{C}_Q(\theta)] = d\sum_{i=1}^{d^2-d} Q(i|\theta) \tag{82}$$

For $\theta \neq 0$ this is always a positive number. Indeed, θ is a positive semidefinite operator and $\theta = \sum \theta_\nu \mathfrak{P}_\nu$ where $\theta_\nu \geq 0$ are its eigenvalues, and $\mathfrak{P}_\nu$ its eigenprojectors. In this case

$$Q(i|\theta) = \sum_\nu \theta_\nu Q(i|\mathfrak{P}_\nu). \tag{83}$$

We have explained earlier that for generic fiducial vectors, d or more coherent states form a total set of vectors in $H(d)$. Therefore for every ν, there exists at least one i for which the $Q(i|\mathfrak{P}_\nu)$ is positive, and then the left hand side of Eq.(82) is positive.

(2) For Hermitian positive semidefinite operators A, B, it is known[33] that $\text{Tr}(AB) \leq \text{Tr}(A)\text{Tr}(B)$. We use this in conjuction with Eq.(77), and we get:

$$\text{Tr}(\rho\theta) \leq \text{Tr}(\rho)\text{Tr}(\theta) = \text{Tr}(\theta) \leq \text{Tr}[\mathcal{C}_Q(\theta)]. \tag{84}$$

(3) We use the formula $\text{Tr}(AB) \leq \text{Tr}(A)\text{Tr}(B)$ in conjuction with Eq.(77), and we get:

$$\text{Tr}(\theta\phi) \leq \text{Tr}(\theta)\text{Tr}(\phi) \leq \text{Tr}[\mathcal{C}_Q(\theta)]\text{Tr}[\mathcal{C}_Q(\phi)]. \tag{85}$$

(4) We first prove that

$$\text{Tr}[\mathcal{C}_Q(\theta + \phi)] \leq \text{Tr}[\mathcal{C}_Q(\theta)] + \text{Tr}[\mathcal{C}_Q(\phi)]. \tag{86}$$

We start from the relation

$$Q(d^2|\theta + \phi) = Q(i_1|\theta) + Q(j_1|\phi). \tag{87}$$

Here the coherent state labelled with d^2 in the ordering of $Q(\alpha, \beta|\theta + \phi)$, is labelled with i_1 in the ordering of $Q(\alpha, \beta|\theta)$, and with j_1 in the ordering of $Q(\alpha, \beta|\phi)$. Similarly

$$Q(d^2 - 1|\theta + \phi) = Q(i_2|\theta) + Q(j_2|\phi), \tag{88}$$

etc. Adding these equations we get

$$\text{Tr}[\mathcal{C}_Q(\theta + \phi)] = d\sum_{\ell=1}^{d}[Q(i_\ell|\theta) + Q(j_\ell|\phi)]. \tag{89}$$

The indices $i_1, ..., i_d$ are different from each other, and take values in the set $A \cup B$ where $A = \{1, ..., d^2 - d\}$ and $B = \{d^2 - d + 1, ..., d^2\}$. If the $Q(i_\ell|\theta)$ has index $i_\ell \in A$, we replace it with another $Q(i'_\ell|\theta)$ with index i'_ℓ in $B - \{i_1, ..., i_d\}$ (in a way that at the end all indices are diferent from each other). This increases the sum, and therefore

$$d\sum_{\ell=1}^{d} Q(i_\ell|\theta) \leq \text{Tr}[\mathcal{C}_Q(\theta)]. \tag{90}$$

We do the same with the $Q(j_\ell|\phi)$ and we prove Eq.(86). From this follows easily Eq.(80).

□

We define the 'dominance ratio'

$$r(\theta) = \frac{\mathrm{Tr}[\mathcal{C}_Q(\theta)]}{d\mathrm{Tr}(\theta)} = \frac{\sum\limits_{i=d^2-d+1}^{d^2} Q(i|\theta)}{\sum\limits_{i=1}^{d^2} Q(i|\theta)}; \quad \frac{1}{d} \leq r(\theta) < 1. \tag{91}$$

It gives the percentage of the sum of the dominant values of the Q-function, with respect to the sum of all values of the Q-function. For any $\lambda > 0$, the operators θ and $\lambda\theta$ have the same dominance ratio. In examples later, we present values of this quantity.

Proposition IV.4. *Displacement transformations on θ, imply displacement transformations on $\mathcal{C}_Q(\theta)$:*

$$\mathcal{C}_Q\left[D(\alpha,\beta)\theta D^\dagger(\alpha,\beta)\right] = D(\alpha,\beta)\mathcal{C}_Q(\theta)D^\dagger(\alpha,\beta) \tag{92}$$

Proof. We first use the definition of Eq.(40) in conjuction with Eq.(39) to prove that

$$Q\left[\gamma,\delta|D(\alpha,\beta)\theta D^\dagger(\alpha,\beta)\right] = Q(\gamma-\alpha,\delta-\beta|\theta) \tag{93}$$

Then we use Eq.(70) in conjuction with Eq.(48) (which are generalized for all $\varpi_\theta(i|i+1,...,d^2)$), and prove Eq.(92). □

The Choquet integral $\mathcal{C}_Q(\theta)$ is based on a ranking formalism and it depends strongly on the dominant coherent states that give a high value of the Q-function. Under displacement transformations, the $\mathcal{C}_Q(\theta)$ transforms as in Eq.(92). In contrast to this, the Wehrl entropy does not change (Eq.(45)). The Wehrl entropy shows whether the Q-function is uniform or concentrated in a few coherent states, but in the latter case it does not show where it is concentrated. Furthermore, $\mathrm{Tr}[\mathcal{C}_Q(\theta)]$ is a convex function while entropy is a concave function, i.e., mixing of two density matrices θ, ϕ into $a\theta + (1-a)\phi$, decreases the $\mathrm{Tr}(\mathcal{C}_Q)$ and increases the entropy. Therefore the Choquet integral contains complementary information to the Wehrl entropy.

A.   Robustness of the formalism in the presence of noise:

The Choquet formalism is robust in the presence of noise. This is because the formalism is based on the ranking in Eq.(66). Noise affects all values of the Q-function in approximately equal way, and it is unlikely that it will change the ranking drastically.

We present a numerical example which shows this. In the 3-dimensional space $H(3)$, we consider a Hermitian operator θ, and add noise in its elements as follows:

$$\theta = \begin{pmatrix} 8+r_1 & 1+r_2+ir_3 & -5+r_4+ir_5 \\ 1+r_2-ir_3 & 4+r_6 & 2+r_7+ir_8 \\ -5+r_4-ir_5 & 2+r_7-ir_8 & 7+r_9 \end{pmatrix} \tag{94}$$

$r_1,...,r_9$ are uniformly distributed random numbers in the region $(-1,1)$. We have calculated the eigenvalues e_1, e_2, e_3 (where $e_1 \leq e_2 \leq e_3$), the corresponding eigenvectors $|v_1\rangle, |v_2\rangle, |v_3\rangle$, and the function $Q(\alpha,\beta)$, of this operator. In table I, we present results for the case without noise (first row), and for five cases with noise. The three dominant values of $Q(\alpha,\beta)$, the eigenvalues, and the dominance ratio $r(\theta)$ (Eq.(91)) are shown. For the eigenvectors we show their overlaps $\tau_i = |\langle u_i|v_i\rangle|^2$ with their counterparts $|u_i\rangle$ in the noiseless case.

It is seen that the dominant coherent states and the corresponding dominant values of $Q(\alpha,\beta|\theta)$ change only slightly. The lowest eigenvalue is sensitive to noise. Overall, the Choquet formalism is robust in the presence of noise.

V. COMONOTONIC OPERATORS

We generalize the concept of comonotonic functions discussed in section II D, into operators in $\mathcal{M}_d$. Comonotonic operators is one way of making precise the intuitive concept of physically similar operators. This is analogous to students with comonotonic marks in section II E, which have similar academic strengths and weaknesses. The Choquet integral of the sum of comonotonic operators, is equal to the sum of the Choquet integrals of the operators. This is used in the next section (corollary VI.3), to derive bounds for the trace of Choquet integrals, which physically are related to mild changes in the physical system.

Definition V.1. Two operators $\theta, \phi \in \mathcal{M}_d$ are comonotonic, if the following statements, which are equivalent to each other, hold:

(1) The ranking permutation of Eq.(66) is the same for both operators: $\sigma(\alpha, \beta \mid \theta) = \sigma(\alpha, \beta \mid \phi)$. The θ, ϕ have the same dominant projectors, and the corresponding $\mathcal{C}_Q(\theta), \mathcal{C}_Q(\phi)$ have the same eigenprojectors and commute:

$$\varpi_\theta(i|i+1,...,d^2) = \varpi_\phi(i|i+1,...,d^2); \qquad [\mathcal{C}_Q(\theta), \mathcal{C}_Q(\phi)] = 0. \tag{95}$$

(2) for the d dominant values of $Q(\alpha, \beta| \theta)$ and $Q(\alpha, \beta|\phi)$

$$[Q(\alpha, \beta| \theta) - Q(\gamma, \delta| \theta)][Q(\alpha, \beta| \phi) - Q(\gamma, \delta| \phi)] \geq 0 \tag{96}$$

It is easily seen that:

- The $\mathbf{1}$ is comonotonic to any other operator.

- For $\lambda \geq 0$, the θ and $\lambda\theta$ are comonotonic.

- If θ, ϕ are comonotonic and $\lambda, \mu \geq 0$, then the $\theta, \phi, \lambda\theta + \mu\phi$ are pairwise comonotonic.

- If θ, ψ are comonotonic, and ϕ, ψ are comonotonic, then the $\theta + \phi, \psi$ are comonotonic.

- If $\mathcal{C}_Q(\theta) = \mathcal{C}_Q(\phi)$, then the operators θ, ϕ are comonotonic.

Proposition V.2. *If θ, ϕ are comonotonic operators, then*

$$\mathcal{C}_Q(a\theta + b\phi) = a\mathcal{C}_Q(\theta) + b\mathcal{C}_Q(\phi); \qquad a, b \geq 0. \tag{97}$$

Proof. For comonotonic operators θ, ϕ, the eigenprojectors $\{\varpi_\theta(i|i+1,...,d^2)\}$ of $\mathcal{C}_Q(\theta)$ are the same with eigenprojectors $\{\varpi_\phi(i|i+1,...,d^2)\}$ of $\mathcal{C}_Q(\phi)$ (Eq.(95)). Then Eq.(97) follows easily. $\qquad\square$

Additivity holds only for comonotonic operators, and we refer to this as the weak additivity property of Choquet integrals.

Proposition V.3. *If θ, ϕ are comonotonic operators, then the $D(\alpha, \beta)\theta D^\dagger(\alpha, \beta)$ and $D(\alpha, \beta)\phi D^\dagger(\alpha, \beta)$ are also comonotonic operators.*

Proof. Since θ, ϕ are Q-comonotonic

$$[Q(\gamma, \delta \mid \theta) - Q(\epsilon, \zeta \mid \theta)][Q(\gamma, \delta \mid \phi) - Q(\epsilon, \zeta \mid \phi)] \geq 0. \tag{98}$$

We insert $\gamma = \gamma' - \alpha$, $\delta = \delta' - \beta$, $\epsilon = \epsilon' - \alpha$, $\zeta = \zeta' - \beta$ and we get

$$[Q(\gamma' - \alpha, \delta' - \beta \mid \theta) - Q(\epsilon' - \alpha, \zeta' - \beta \mid \theta)][Q(\gamma' - \alpha, \delta' - \beta \mid \phi) - Q(\epsilon' - \alpha, \zeta' - \beta \mid \phi)] \geq 0. \tag{99}$$

Taking into account Eq.(93) we rewrite this as

$$[Q(\gamma', \delta' \mid D(\alpha, \beta)\theta D^\dagger(\alpha, \beta)) - Q(\epsilon', \zeta' \mid D(\alpha, \beta)\theta D^\dagger(\alpha, \beta))]$$
$$\times [Q(\gamma', \delta' \mid D(\alpha, \beta)\phi D^\dagger(\alpha, \beta)) - Q(\epsilon', \zeta' \mid D(\alpha, \beta)\phi D^\dagger(\alpha, \beta))] \geq 0. \tag{100}$$

and this proves that the $D(\alpha, \beta)\theta D^\dagger(\alpha, \beta)$ and $D(\alpha, \beta)\phi D^\dagger(\alpha, \beta)$ are comonotonic operators. $\qquad\square$

A. Equivalence classes of comonotonic operators in $\mathcal{M}'_d$

Comonotonicity is not transitive in the set $\mathcal{M}_d$. For example, $\mathbf{1}$ is comonotonic to every operator and yet there are operators which are not comonotonic. We define a subset of $\mathcal{M}_d$ where comonotonicity is transitive.

Definition V.4.

(1) $\mathcal{M}'_d$ is a subset of $\mathcal{M}_d$ which contains operators θ for which the d dominant values of $Q(i \mid \theta)$ are different from each other. $\mathcal{N}'_d \subset \mathcal{M}'_d$ is the set of such operators with trace equal to one.

(2) Through the Choquet integral map, $\mathcal{M}'_d$ is mapped into $\mathcal{C}'_d \subset \mathcal{C}_d$ which contains Choquet integrals with eigenvalues $Q(i \mid \theta)$ which are different from each other.

Comonotonicity is transitive in $\mathcal{M}'_d$. In this case we have a strict inequality in Eq.(96). This is analogous to commutativity which is not transitive in general, but it is transitive if we restrict ourselves to matrices with eigenvalues which are different from each other.

In $\mathcal{M}'_d$ (and $\mathcal{N}'_d$) comonotonicity is an equivalence relation, which partitions $\mathcal{M}'_d$ (and $\mathcal{N}'_d$) into equivalence classes, which we denote as $\mathcal{M}'_d(\nu)$ (and $\mathcal{N}'_d(\nu)$). We denote comonotonic operators in these classes with $\theta_1 \sim \theta_2$. It is easily seen that if $\theta_1 \sim \theta_2$ then $a\theta_1 + b\theta_2 \sim \theta_1 \sim \theta_2$, where $a, b \geq 0$.

Through the Choquet integral map, $\mathcal{C}'_d$ is also partitioned into equivalence classes, which we denote as $\mathcal{C}'_d(\nu)$. There is an ordered set of d coherent states associated with each equivalence class $\mathcal{M}'_d(\nu)$. The number of such classes is

$$n_d = d^2(d^2 - 1)...(d^2 - d + 1) = \frac{d^2!}{(d^2 - d)!} \tag{101}$$

We prove this by taking one coherent state from the set of d^2 coherent states, and then another coherent state from the remaining set of $d^2 - 1$ coherent states (which we use together with the first coherent state for the two-dimensional cumulative projectors in section III B), etc.

So there is a finite number of sets of projectors in the formalism:

$$\mathcal{S}_\nu = \{\varpi_\nu(i|i + 1, ..., d^2)] \, |i = d^2 - d + 1, ..., d^2\}; \quad \nu = 1, ..., n_d. \tag{102}$$

All Choquet integrals in the same equivalence class $\mathcal{C}'_d(\nu)$ commute with each other, and have the same eigenprojectors: Through the Choquet integral map, the property comonotonicity in $\mathcal{M}'_d$, becomes commutativity in $\mathcal{C}'_d$.

In the case that two of the dominant values of the Q-function are equal to each other (i.e., for operators $\theta \in \mathcal{M}_d - \mathcal{M}'_d$), the corresponding sums of projectors

$$\varpi_\nu(i|i + 1, ..., d^2) + \varpi_\nu(i + 1|i + 2, ..., d^2) = \Pi_\theta(i; i + 1; ...; d^2) - \Pi_\theta(i + 2; ...; d^2) \tag{103}$$

enter into $\mathcal{C}_Q(\theta)$ as explained in Eq.(71).

B. Comonotonicity intervals of operators $\theta(\lambda)$ and crossings of the Q-function

In many cases the operator θ is a function of a real parameter λ. Examples are:

- A Hamiltonian $\theta(\lambda) = \theta_1 + \lambda\theta_2$, where θ_1 is the free part, θ_2 the interaction part, and λ the coupling constant (an example is given in section V C below).

- $\theta(\lambda) = |g(\lambda)\rangle\langle g(\lambda)|$ where $|g(\lambda)\rangle$ is the ground state of a system described by a Hamiltonian $\mathfrak{H}(\lambda)$ with coupling constant λ (two examples are given in section VII, with and without degeneracies in the eigenvalues).

- The $\theta(\lambda) = \exp(-\lambda\mathfrak{H})$ where $\mathfrak{H}$ is a Hamiltonian and λ the inverse temperature. The trace of this operator is the partition function, and bounds for it are given in section VII below.

If $\theta(\lambda)$ is a continuous function of λ, the $Q[\alpha, \beta|\theta(\lambda)]$ are also continuous functions of λ. Consequently, there are intervals of the parameter λ, where the ranking of the d highest values of $Q[\alpha, \beta|\theta(\lambda)]$ remains unchanged. We call them comonotonicity intervals. By definition, if λ_1, λ_2 belong to the same comonotonicity interval, then the $\theta(\lambda_1)$, $\theta(\lambda_2)$ are comonotonic. But it is not necessary that all operators in a comonotonicity interval belong in the same equivalence class of $\mathcal{M}'_d$ (there are pairs of comonotonic operators in $\mathcal{M}_d - \mathcal{M}'_d$, but transitivity might not hold).

There might be values λ_i where we have crossings of the d highest values of the Q-function:

$$Q[\alpha_1, \beta_1|\theta(\lambda_i)] = Q[\alpha_2, \beta_2|\theta(\lambda_i)] \tag{104}$$

We call them crossings of the Q-function. At these points a change of the ranking occurs, and the matrix $\mathcal{C}_Q[\theta(\lambda)]$ has a discontinuity. The $\text{Tr}\{\mathcal{C}_Q[\theta(\lambda)]\}$ is continuous at these points (as sum of continuous functions), but its derivative with respect to λ, might have discontinuities.

If $\{\lambda_i\}$ is the set of the crossings of the Q-function, the λ-axis is partitioned to many intervals $(\lambda_i, \lambda_{i+1})$, and within each interval all the $\theta(\lambda)$ are pairwise comonotonic operators.

Remark V.5. A phenomenon analogous to 'avoided crossing' of the energy levels, might occur. A small external perturbation can invalidate the equality in Eq.(104). For example, a small amount of noise added into $\theta(\lambda)$ will make it $\theta(\lambda) + \Delta\theta$ where $\Delta\theta$ is an infinitesimal matrix (which we assume to be Hermitian). Then

$$Q[\alpha_i, \beta_i|\theta(\lambda) + \Delta\theta] = Q[\alpha_i, \beta_i|\theta(\lambda)] + Q[\alpha_i, \beta_i|\Delta\theta] \tag{105}$$

and in general $Q[\alpha_1, \beta_1|\Delta\theta] \neq Q[\alpha_2, \beta_2|\Delta\theta]$. In this case the curve $Q[\alpha_1, \beta_1|\theta(\lambda) + \Delta\theta]$ on the left of the crossing point λ_i will join the curve $Q[\alpha_2, \beta_2|\theta(\lambda) + \Delta\theta]$ on the right of the crossing point (and the curve $Q[\alpha_2, \beta_2|\theta(\lambda) + \Delta\theta]$ on the left of the crossing point will join the curve $Q[\alpha_1, \beta_1|\theta(\lambda) + \Delta\theta]$ on the right of the crossing point). Therefore λ_i will not be a crossing point of the Q-function, and the left and right comonotonicity intervals, will join to become one comonotonicity interval. We call this 'avoided crossings of the Q-function'. It is a phenomenon which should be studied in its own right. In the examples below, we assume the absence of such perturbations, and the absence of the 'avoided crossings of the Q-function'.

C. Example

We consider the operator (in the position basis):

$$\theta(\lambda) = \theta_1 + \lambda\theta_2; \quad \theta_1 = \begin{pmatrix} 6 & 0 & i \\ 0 & 12 & 0 \\ -i & 0 & 15 \end{pmatrix}; \quad \theta_2 = \begin{pmatrix} 7 & 3 & 6 \\ 3 & 7 & 0 \\ 6 & 0 & 7 \end{pmatrix}. \tag{106}$$

The values of $Q(\alpha, \beta \,|\theta)$ are:

$$Q(0,0|\theta) = 4.5 + 3.476\lambda; \quad Q(0,1|\theta) = 3 + 4.476\lambda; \quad Q(0,2|\theta) = 3.5 + 3.762\lambda$$
$$Q(1,0|\theta) = 4.623 + 1.761\lambda; \quad Q(1,1|\theta) = 3.247 + 1.261\lambda; \quad Q(1,2|\theta) = 3.582 + 1.619\lambda$$
$$Q(2,0|\theta) = 4.376 + 1.762\lambda; \quad Q(2,1|\theta) = 2.752 + 1.261\lambda; \quad Q(2,2|\theta) = 3.417 + 1.619\lambda \tag{107}$$

The three dominant values, which in our notation are $Q(9|\theta), Q(8|\theta), Q(7|\theta)$, depend on the value of λ. We consider the interval $0 \leq \lambda \leq 0.7$, and ordering of the $Q(\alpha, \beta \,|\theta)$ shows that it consists of five comonotonicity intervals. In table II we show these comonotonicity intervals, the corresponding three dominant values of $Q(\alpha, \beta)$, and the dominance ratio $r[\theta(\lambda)]$. The $r[\theta(\lambda)]$ is a continuous function of λ, but its derivative with respect to λ has discontinuiuties at $\lambda = 0.44$ and $\lambda = 0.6$.

We present the Choquet integral, for the first two intervals. In the comonotonicity interval $(0, 0.06)$:

$$\varpi_\theta(9) = \Pi(1,0); \quad \varpi_\theta(8) = \Pi(0,0;1,0) - \Pi(1,0); \quad \varpi_\theta(7) = \mathbf{1} - \Pi(0,0;1,0) \tag{108}$$

Therefore

$$\mathcal{C}_Q[\theta(\lambda)] = 3[Q(7|\theta)\varpi_\theta(7) + Q(8|\theta)\varpi_\theta(8) + Q(9|\theta)\varpi_\theta(9)] = A_1 + \lambda B_1$$

$$A_1 = \begin{pmatrix} 13.25 & 0.04 - 0.16i & 0.01 + 0.14i \\ 0.04 + 0.16i & 13.53 & -0.05 - 0.17i \\ 0.01 - 0.14i & -0.05 + 0.17i & 13.70 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} 6.28 & 1.31 - i & 1.18 + 0.13i \\ 1.31 + i & 8.01 & 1.41 + 1.36i \\ 1.18 - 0.13i & 1.41 - 1.36i & 6.7 \end{pmatrix}; \quad [A_1, B_1] = 0.. \tag{109}$$

In the comonotonicity interval $(0.06, 0.44)$,

$$\varpi_\theta(9) = \Pi(0,0); \quad \varpi_\theta(8) = \Pi(1,0;0,0) - \Pi(0,0); \quad \varpi_\theta(7) = \mathbf{1} - \Pi(1,0;0,0) \tag{110}$$

Therefore

$$\mathcal{C}_Q[\theta(\lambda)] = 3[Q(7|\theta)\varpi_\theta(7) + Q(8|\theta)\varpi_\theta(8) + Q(9|\theta)\varpi_\theta(9)] = A_2 + \lambda B_2$$

$$A_2 = \begin{pmatrix} 13.29 & 0.08 - 0.23i & 0.01 + 0.15i \\ 0.08 + 0.23i & 13.62 & -0.11 - 0.07i \\ 0.01 - 0.15i & -0.11 + 0.07i & 13.57 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 5.65 & 0.73 & 1.10 \\ 0.73 & 6.75 & 2.20 \\ 1.10 & 2.20 & 8.59 \end{pmatrix}; \quad [A_2, B_2] = 0. \tag{111}$$

Within each of the comonotonic intervals the $\theta(\lambda)$ are comonotonic operators, and the $\mathcal{C}_Q[\theta(\lambda)]$ commute with each other and have the same eigenprojectors. At the crossing points of the Q-function, the $\mathcal{C}_Q[\theta(\lambda)]$ has a discontinuity.

VI. BOUNDS FOR COMONOTONIC OPERATORS

We have seen earlier (proposition IV.3) that the trace of the Choquet integral is a bound for physical quantities like $\mathrm{Tr}(\theta)$, $\mathrm{Tr}(\rho\theta)$, $\mathrm{Tr}(\phi\theta)$, etc. This is our physical motivation for using it in this section, to define an order among the Hermitian operators.

An order is useful if it has certain properties, and a natural property is that addition should preserve the order (the analogue of $a \geq b$ implies that $a + c \geq b + c$ in real numbers). We show that this property is valid in the case of comonotonic operators. This already shows that in some sense comonotonic operators are physically similar. More importantly, a whole family of comonotonic operators are bounded, with respect to this order (corollary VI.3 below). This means that the trace of the Choquet integral is bounded, and therefore the other physical quantities to which this is a bound, are also bounded.

We introduce the 'greater trace of the Choquet integral' preorder, as follows:

Definition VI.1. $\theta_1 \succ \theta_2$ if $\mathrm{Tr}[\mathcal{C}_Q(\theta_1)] \geq \mathrm{Tr}[\mathcal{C}_Q(\theta_2)]$.

$\succ$ is transitive, but the antisymmetry property does not hold ($\theta_1 \succ \theta_2$ and $\theta_1 \prec \theta_2$ implies that $\mathrm{Tr}[\mathcal{C}_Q(\theta_1)] = \mathrm{Tr}[\mathcal{C}_Q(\theta_2)]$, but it does not follow that $\theta_1 = \theta_2$). Therefore $\succ$ is a preorder, rather than a partial order. It is a total preorder because for any θ_1, θ_2, either $\theta_1 \prec \theta_2$ or $\theta_1 \succ \theta_2$.

The following proposition shows that for comonotonic operators, addition preserves the $\succ$ preorder:

Proposition VI.2.

(1) If θ_1, θ_3 are comonotonic, and θ_2, θ_3 are comonotonic, then

$$\theta_1 \succ \theta_2 \quad \rightarrow \quad \theta_1 + \theta_3 \succ \theta_2 + \theta_3. \tag{112}$$

(2) For comonotonic θ_1, θ_2

$$\theta_1 \succ \theta_2 \quad \rightarrow \quad \theta_1 \succ a\theta_1 + (1-a)\theta_2 \succ \theta_2; \quad 0 \le a \le 1. \tag{113}$$

Proof.

(1) We have

$$\theta_1 \succ \theta_2 \quad \rightarrow \quad \mathrm{Tr}[\mathcal{C}_Q(\theta_1)] \ge \mathrm{Tr}[\mathcal{C}_Q(\theta_2)] \rightarrow \quad \mathrm{Tr}[\mathcal{C}_Q(\theta_1) + \mathcal{C}_Q(\theta_3)] \ge \mathrm{Tr}[\mathcal{C}_Q(\theta_2) + \mathcal{C}_Q(\theta_3)] \tag{114}$$

Using the additivity of the Choquet integral for comonotonic operators, we rewrite this as

$$\mathrm{Tr}[\mathcal{C}_Q(\theta_1 + \theta_3)] \ge \mathrm{Tr}[\mathcal{C}_Q(\theta_2 + \theta_3)], \tag{115}$$

and this proves the proposition.

(2) $\theta_1 \succ \theta_2$, implies that $(1-a)\theta_1 \succ (1-a)\theta_2$. We add $a\theta_1$ on both sides, and using Eq.(112), and we get $\theta_1 \succ a\theta_1 + (1-a)\theta_2$. In a similar way we prove that $a\theta_1 + (1-a)\theta_2 \succ \theta_2$.

$\square$

Corollary VI.3. *Let $\theta(\lambda)$ be an operator which is a linear function of λ, within a comonotonicity interval I. For $\lambda_1 < \lambda_2$ where $\lambda_1, \lambda_2 \in I$, we assume that $\theta(\lambda_1) \succ \theta(\lambda_2)$. Then at any point $\lambda \in [\lambda_1, \lambda_2]$*

$$\theta(\lambda_1) \succ \theta(\lambda) \succ \theta(\lambda_2) \tag{116}$$

Proof. For $\lambda \in [\lambda_1, \lambda_2]$

$$\theta(\lambda) = a\theta(\lambda_1) + (1-a)\theta(\lambda_2); \quad a = \frac{\lambda_2 - \lambda}{\lambda_2 - \lambda_1}, \tag{117}$$

and use of Eq.(113) proves the statement. $\square$

This result shows that for comonotonic operators, the $\mathrm{Tr}\{\mathcal{C}_Q[\theta(\lambda)]\}$ is bounded by $\mathrm{Tr}\{\mathcal{C}_Q[\theta(\lambda_1)]\}$ and $\mathrm{Tr}\{\mathcal{C}_Q[\theta(\lambda_2)]\}$. Therefore physical quantities to which the $\mathrm{Tr}\{\mathcal{C}_Q[\theta(\lambda)]\}$ is a bound, are also bounded.

The corollary assumes that $\theta(\lambda)$ is a linear function of λ. If $\theta(\lambda)$ is a non-linear function of λ, a comonotonicity interval can be divided into many small subintervals, and within each of them $\theta(\lambda)$ is approximately a linear function of λ, and the corollary can be used. An example of this is discussed in section VII A below.

Remark VI.4. We rewrite the result in Eq.(116), as

$$f(\lambda_1) \ge f(\lambda) \ge f(\lambda_2); \quad f(\lambda) = \mathrm{Tr}\{\mathcal{C}_Q[\theta(\lambda)]\}. \tag{118}$$

Since $\theta(\lambda)$ is a continuous function of λ, the $f(\lambda)$ is also a continuous function of λ. The intermediate value theorem for continuous functions states that for any $f_0 \in [f(\lambda_2), f(\lambda_1)]$, there exists $\lambda_0 \in [\lambda_1, \lambda_2]$ such that $f(\lambda_0) = f_0$. We point out that this is weaker than our result in Eq.(116), for comonotonic operators.

VII.   APPLICATIONS TO THE STUDY OF THE GROUND STATE OF PHYSICAL SYSTEMS

The study of the ground state of a large physical system as a function of the coupling constant, is important for phase transitions. We study a toy model which shows how our formalism can be used for the study of the ground state of a physical system. For practical reasons, we consider a small system described with the 3-dimensional space $H(3)$, and study two cases of Hamiltonians with and without degeneracies in their eigenvalues.

A.   The ground state of a physical system with Hamiltonian without degeneracies

We consider the following Hamiltonian which is a non-linear function of the coupling constant λ:

$$\mathfrak{H}(\lambda) = \begin{pmatrix} 7 & 3i\lambda + \lambda^2 & 6i\lambda + 2\lambda^2 \\ -3i\lambda + \lambda^2 & 9 & 5\lambda + 4\lambda^2 \\ -6i\lambda + 2\lambda^2 & 5\lambda + 4\lambda^2 & 11 \end{pmatrix} \tag{119}$$

We have calculated numerically the eigenstate $|g(\lambda)\rangle$ which corresponds to the lowest eigenvalue of $\mathfrak{H}(\lambda)$, and then calculated the $Q[\alpha, \beta|\mathfrak{P}(\lambda)]]$ where $\mathfrak{P}(\lambda) = |g(\lambda)\rangle\langle g(\lambda)|$. The three dominant values of $Q[\alpha, \beta|\mathfrak{P}(\lambda)]$ are shown in table III for $\lambda = 0.1, 0.2, ..., 1$. It is seen that the comonotonicity intervals are $[0, \lambda_1]$, $[\lambda_1, \lambda_2]$, $[\lambda_2, \lambda_3]$, and $[\lambda_3, 1]$, where $\lambda_1 \approx 0.3$, $\lambda_2 \approx 0.4$ and $\lambda_3 \approx 0.7$. The dominance ratio $r[\mathfrak{P}(\lambda)]$ (Eq.(91)), is also shown.

The $\mathfrak{H}(\lambda), \mathfrak{P}(\lambda)$ are non-linear functions of λ. However if we divide each comonotonicity interval into small subintervals, we can assume that $\mathfrak{P}(\lambda)$ is approximately linear within each subinterval, and use corollary VI.3. For example we consider the comonotonicity interval $(0.5, 0.7)$ where $\mathfrak{P}(0.5) \prec \mathfrak{P}(0.7)$ because $\text{Tr}\{\mathcal{C}_Q[\mathfrak{P}(0.5)]\} = r[\mathfrak{P}(0.5)] = 0.638$ and $\text{Tr}\{\mathcal{C}_Q[\mathfrak{P}(0.7)]\} = r[\mathfrak{P}(0.7)] = 0.710$ (table III). We divide it into the subintervals $(0.5, 0.6)$ and $(0.6, 0.7)$, and then

$$\begin{aligned} 0.5 \leq \lambda \leq 0.6 &\rightarrow \mathfrak{P}(0.5) \prec \mathfrak{P}(\lambda) \prec \mathfrak{P}(0.6) \\ 0.6 \leq \lambda \leq 0.7 &\rightarrow \mathfrak{P}(0.6) \prec \mathfrak{P}(\lambda) \prec \mathfrak{P}(0.7) \end{aligned} \tag{120}$$

Similarly in the comonotonicity interval $(0.8, 1)$ we have $\mathfrak{P}(0.8) \succ \mathfrak{P}(1)$. We divide it into the subintervals $(0.8, 0.9)$ and $(0.9, 1)$, and then

$$\begin{aligned} 0.8 \leq \lambda \leq 0.9 &\rightarrow \mathfrak{P}(0.8) \succ \mathfrak{P}(\lambda) \succ \mathfrak{P}(0.9) \\ 0.9 \leq \lambda \leq 1 &\rightarrow \mathfrak{P}(0.9) \succ \mathfrak{P}(\lambda) \succ \mathfrak{P}(1). \end{aligned} \tag{121}$$

The Wehrl entropy $E[\mathfrak{P}(\lambda)]$ (Eq.(43)) that involves all 9 values of $Q[\alpha, \beta|\mathfrak{P}(\lambda)]$, is also shown (we used natural logarithms and the result is in nats). The Wehrl entropy has been used in the literature as an indicator of phase transitions (e.g[34]). The $E[\mathfrak{P}(\lambda)]$ has local maxima and minima at the values $\lambda_1' \approx 0.3$ and $\lambda_2' \approx 0.6$, which agrees roughly with the values where $\mathcal{C}_Q[\mathfrak{P}(\lambda)]$ is discontinuous.

The overlap $|\langle g(0)|g(\lambda)\rangle|^2$ of the ground state $|g(\lambda)\rangle$ when the coupling constant is equal to λ, with the ground state $|g(0)\rangle$ when the coupling constant is equal to 0, is a measure of how much the ground state changes. It is given in table III, and it is seen that the biggest change occurs in the region of $\lambda \sim (0.4, 0.7)$

Therefore different quantities confirm that at the crossing points of the Q-function, stronger physical changes occur into the system.

B.   The ground state of a physical system with Hamiltonian with degeneracies

We consider the Hamiltonian:

$$\mathfrak{H}(\lambda) = \begin{pmatrix} 1.500 & 1.414 + \lambda & 1.732 \\ 1.414 + \lambda^* & 2.500 & 2.449 \\ 1.732 & 2.449 & 3.500 \end{pmatrix} \tag{122}$$

The eigenvalues e_1, e_2, e_3 of this Hamiltonian for $\lambda = -0.01, 0, 0.01, 0.01i, -0.01i$ are given in table IV. For $\lambda = 0$ we have a degeneracy, and the two lowest eigenvalues are equal to each other.

In the cases $\lambda = -0.01, 0.01, -0.01i, 0.01i$ that there is no degeneracy, we have calculated numerically the eigenstate $|g(\lambda)\rangle$ which corresponds to the lowest eigenvalue of $\mathfrak{H}(\lambda)$, and then calculated the $Q[\alpha, \beta|\mathfrak{P}(\lambda)]]$ where $\mathfrak{P}(\lambda) = |g(\lambda)\rangle\langle g(\lambda)|$. We also calculated the Wehrl entropy $E[\mathfrak{P}(\lambda)]$. For $\lambda = 0$, the $\mathfrak{P}(0)$ is the projector to the two-dimensional eigenspace corresponding to the two lowest eigenvalues. In this case, we calculated the $Q[\alpha, \beta|\frac{1}{2}\mathfrak{P}(\lambda)]$ (so that the sum of all the Q-values is 1), and it is these values that we used to calculate the Wehrl entropy.

The three dominant values of $Q[\alpha, \beta|\mathfrak{P}(\lambda)]$ are shown in table IV. Although λ changes by a small amount, and the eigenvalues also change by a small amount, the dominant coherent states change drastically as we go from $\lambda = -0.01$ to $\lambda = 0$ (where we get degeneracy), and then to $\lambda = 0.01$ Similar comment can be made for going from $\lambda = -0.01i$ to $\lambda = 0$, and then to $\lambda = 0.01i$. If we compare the cases $\lambda = 0.01$ and $\lambda = 0.01i$, there is also a change in the dominant coherent states. This is because the eigenvector corresponding to the lowest eigenvalue in the $\lambda = 0.01$ case, is very different from the eigenvector corresponding to the lowest eigenvalue in the $\lambda = 0.01i$ case (we have found that $|\langle g(0.01i)|g(0.01)\rangle|^2 = 0.5$).

The results show that the method is sensitive enough to detect changes in the ground state in the case of degeneracies. A change in the ground state, changes some values of the Q-function more than others, and this changes the ranking of the Q-function, and for this reason it is easily detected by the Choquet formalism. In contrast to this, we have seen in section IV A, that random noise affects all values of the Q-function in approximately equal way, the ranking remains the same, and for this reason the formalism is robust in the presence of noise.

VIII.   BOUNDS FOR PARTITION FUNCTIONS

Inequalities between quantities that involve matrices (e.g.,[35]) have many applications in Physics (e.g., [36, 37]), and also in other subjects like Control Theory in Electrical Engineering, Operational Research, etc. In this general context, this paper uses Choquet integrals in conjuction with total sets of vectors (like coherent states). In this section we derive bounds for the partition function, which together with proposition IV.3, show the use of the formalism for bounds of physical quantities.

If θ is a Hamiltonian and λ the inverse temperature, then $\mathrm{Tr}\exp(-\lambda\theta)$ is a partition function. Below we derive upper and lower bounds for the partition function. We also show that $\mathcal{C}_Q[\exp(-\lambda\theta)] \geq \exp[-\lambda\mathcal{C}_Q(\theta)]$ (the $\theta_1 \geq \theta_2$ denotes the fact that $\theta_1 - \theta_2$ is a positive semidefinite Hermitian operator).

Proposition VIII.1.

(1)

$$\mathrm{Tr}\mathcal{C}_Q[\exp(-\lambda\theta)] \geq \mathrm{Tr}\exp(-\lambda\theta) \geq A$$

$$A = \max\left(\frac{1}{d}\mathrm{Tr}\mathcal{C}_Q[\exp(-\lambda\theta)], \frac{1}{d}\sum_{\alpha,\beta}\exp[-d\lambda Q(\alpha, \beta|\theta)]\right). \tag{123}$$

(2) If θ and $\exp(-\lambda\theta)$ (where $\lambda \geq 0$) are comonotonic operators, then

$$\mathcal{C}_Q[\exp(-\lambda\theta)] \geq \exp[-\lambda\mathcal{C}_Q(\theta)]. \tag{124}$$

Proof.

(1) The left part of the inequality folows immediately from the inequality in Eq.(77). For the right part of the inequality, we have

$$\mathrm{Tr}\exp(-\lambda\theta) = \sum_{\alpha,\beta}Q[\alpha, \beta|\exp(-\lambda\theta)]. \tag{125}$$

The Bogoliubov inequality states that for any state ψ and Hermitian operator ϕ

$$\langle\psi|\exp(\phi)|\psi\rangle \geq \exp[\langle\psi|\phi|\psi\rangle].\tag{126}$$

Consequently

$$dQ[\alpha,\beta|\exp(-\lambda\theta)] \geq \exp[-d\lambda Q(\alpha,\beta|\theta)].\tag{127}$$

Therefore

$$\sum_{\alpha,\beta} Q[\alpha,\beta|\exp(-\lambda\theta)] \geq \frac{1}{d}\sum_{\alpha,\beta}\exp[-d\lambda Q(\alpha,\beta|\theta)].\tag{128}$$

From Eqs(125),(128), it follows that [38]

$$\mathrm{Tr}\exp(-\lambda\theta) \geq \frac{1}{d}\sum_{\alpha,\beta}\exp[-d\lambda Q(\alpha,\beta|\theta)].\tag{129}$$

But we also have

$$\mathrm{Tr}\exp(-\lambda\theta) > \frac{1}{d}\mathrm{Tr}\mathcal{C}_Q[\exp(-\lambda\theta)],\tag{130}$$

from Eq.(77). This completes the proof.

(2)

$$\mathcal{C}_Q[\exp(-\lambda\theta)] = \sum_{i=d^2-d+1}^{d^2} dQ[\exp(-\lambda\theta)]\varpi_{\exp(-\lambda\theta)}(i|i+1,...,d^2).\tag{131}$$

Since the operators θ and $\exp(-\lambda\theta)$ are comonotonic

$$\varpi_{\exp(-\lambda\theta)}(i|i+1,...,d^2) = \varpi_\theta(i|i+1,...,d^2).\tag{132}$$

Using Eq.(127) which is based on the Bogoliubov inequality, we get

$$\sum_{i=d^2-d+1}^{d^2} dQ[\exp(-\lambda\theta)]\varpi_{\exp(-\lambda\theta)}(i|i+1,...,d^2) \geq \sum_{i=d^2-d+1}^{d^2}\exp[-d\lambda Q(\theta)]\varpi_\theta(i|i+1,...,d^2)$$
$$= \exp\left\{\sum_{i=d^2-d+1}^{d^2}[-d\lambda Q(\theta)\varpi_\theta(i|i+1,...,d^2)]\right\} = \exp[-\lambda\mathcal{C}_Q(\theta)]\tag{133}$$

This completes the proof.

$\square$

There are two lower bounds in Eq.(123), which involve the Q-function of $\exp(-\lambda\theta)$ and the Q-function of θ. We give two examples which show that sometimes the first is better lower bound, while other times the second is better lower bound. The first example is

$$\theta = \begin{pmatrix} 8 & 1+i & -5 \\ 1-i & 4 & 2 \\ -5 & 2 & 7 \end{pmatrix}\tag{134}$$

For $\lambda = 1$, we get

$$\text{Tr} \exp(-\lambda\theta) = 0.440; \quad \frac{1}{d}\text{Tr}\mathcal{C}_Q[\exp(-\lambda\theta)] = 0.073; \quad \frac{1}{d}\sum_{\alpha,\beta}\exp[-d\lambda Q(\alpha,\beta|\theta)] = 0.013. \tag{135}$$

Here the $\frac{1}{d}\text{Tr}\mathcal{C}_Q[\exp(-\lambda\theta)]$ is a better lower bound.

The second example is $\theta = \mathbf{1}$, in which case

$$\text{Tr} \exp(-\lambda\theta) = d\exp(-\lambda); \quad \frac{1}{d}\text{Tr}\mathcal{C}_Q[\exp(-\lambda\theta)] = \exp(-\lambda); \quad \frac{1}{d}\sum_{\alpha,\beta}\exp[-d\lambda Q(\alpha,\beta|\theta)] = d\exp(-\lambda). \tag{136}$$

Here the $\frac{1}{d}\sum_{\alpha,\beta}\exp[-d\lambda Q(\alpha,\beta|\theta)]$ is a better lower bound.

IX. SPECTRAL FORMALISM, POVM, WAVELETS AND THE CHOQUET FORMALISM

In this section we compare and contrast the Choquet formalism with the spectral formalism of eigenvalues and eigenvectors, the POVM formalism, and the formalism of frames and wavelets. Let θ be a Hermitian operator.

- **Spectral formalism of eigenvalues and eigenvectors:**
  - It uses the complete set of the d eigenvectors of θ, which are orthogonal to each other. This set is not fixed, but depends on θ.
  - $\theta = \sum e_i\mathfrak{P}_i$, where $\mathfrak{P}_i$ are the eigenprojectors and $e_i = \text{Tr}(\theta\mathfrak{P}_i)$ the eigenvalues of θ.
  - If two operators commute, they have the same eigenprojectors $\mathfrak{P}_i$.

- **POVM formalism:**
  - It uses the set Ω of d^2 coherent states. The central feature is the resolution of the identity in Eq.(39), which is used in expressing various physical quantities in terms of coherent states.
  - $\theta = \sum P(\alpha,\beta|\theta)\Pi(\alpha,\beta)$ in terms of the projectors $\Pi(\alpha,\beta)$ and the P-function $P(\alpha,\beta|\theta)$, as explained in Eq.(41).

- **Frames and wavelets:**
  - A frame is a family of states $|v_i\rangle$, such that for all (normalized) states $|f\rangle$ in the Hilbert space

$$A \leq \sum_i |\langle v_i|f\rangle|^2 \leq B$$

$$|f\rangle = \sum_i (S^{-1}|v_i\rangle)\langle v_i|f\rangle; \quad S = \sum_i |v_i\rangle\langle v_i|. \tag{137}$$

  A, B are constants called lower and upper bound.
  - The philosophy here that if we do not know an exact resolution of the identity, we should try to find lower and upper bounds for it. In this sense, the formalism uses an approximate resolution of the identity, with bounded error.

- **Choquet formalism:**

- It uses the set Ω of d^2 coherent states, but it does not use their resolution of the identity of Eq.(39). The formalism introduces its own 'weak resolution of the identity' of Eq.(61), that involves the non-orthogonal projectors and also the Möbius operators that eliminate the double counting. $Q(\alpha, \beta | \theta) = \frac{1}{d} \text{Tr}[\theta \Pi(\alpha, \beta)]$ is the Q-function of a Hermitian operator θ. Based on the ranking in Eq.(66), the coherent states, projectors and $Q(\alpha, \beta | \theta)$ are divided into two groups 'dominant' and 'inferior', which depend on θ.

- $\mathcal{C}_Q(\theta) = \sum Q(i|\theta) \varpi_\theta(i; i+1; ...; d^2)$. The projectors $\varpi_\theta(i|i+1, ..., d^2)$ are discrete derivatives (differences) of the cumulative projectors $\Pi_\theta(i; i+1; ...; d^2)$. The $\varpi_\theta(i|i+1, ..., d^2)$ form an orthogonal set of d projectors, and they are different from the projectors $\Pi(i)$, associated to coherent states. The $Q(i|\theta)$ and $\varpi_\theta(i; i+1; ...; d^2)$ are eigenvalues and eigenprojectors of $\mathcal{C}_Q(\theta)$. The $\mathcal{C}_Q(\theta)$ is a figure of merit for θ, and is in general different from θ. The $\text{Tr}[\mathcal{C}_Q(\theta)]$ is an upper bound for various physical quantities as shown in proposition IV.3 and in section VIII.

- If two operators are comonotonic, they have the same $\varpi_\theta(i|i+1, ..., d^2)$ projectors. Comonotonic operators have the same dominant coherent states and projectors, and their Choquet integrals commute. Comonotonicity formalizes the vague concept of physically similar operators.

- The frames and wavelets formalism, uses approximate resolutions of the identity with bounded error. The Choquet formalism corrects this error with the Möbius operators, and uses the 'weak resolution of the identity' of Eq.(61).

In this paper we used the Choquet formalism with coherent states, but as we explained the formalism introduces its own 'weak resolution of the identity' of Eq.(61), and it does not use the resolution of the identity in Eq.(39). Therefore the formalism can be used with total sets of states, for which we do not know explicitly a resolution of the identity (a set of states is called total, if there is no state in the Hilbert space which is orthogonal to all states in the set). The Choquet formalism introduces a 'weak resolution of the identity', that involves the Möbius operators in addition to the projectors. It is robust in the presence of noise, and it can be used as bound for various physical quantities, in the study of the ground state of physical systems, etc.

X. DISCUSSION

The Choquet integral is used in problems with probabilities, which involve overlapping (non-independent) alternatives. In this paper, we have used it in a quantum context with the Q-function of Hermitian positive semidefinite operators. The Q-function is defined in terms of coherent states, which overlap with each other, and this motivates the use of this approach. The Choquet integral uses the ranking of the values of the Q-function in Eq.(66), and it is given by Eq.(70).

The formalism uses the Möbius operators $\mathfrak{D}(\alpha_1, \beta_1; \alpha_2, \beta_2)$, $\mathfrak{D}(\alpha_1, \beta_1; \alpha_2, \beta_2; \alpha_3, \beta_3)$, etc, to quantify the overlaps between coherent states. They enter in the Choquet integral as described in proposition IV.1. The Möbius operators are interpreted in the context of non-additive probabilities (capacities), and they are related to commutators as in Eq.(54), which shows that they are non-zero if the projectors do not commute.

A central concept in the formalism, which is novel in Physics, is comonotonicity. It is used to formalize the vague concept of physically similar operators. Comonotonic operators are bounded as in Eq.(116), with respect to the $\prec$ preorder. This means that the values of $\text{Tr}\mathcal{C}_Q(\theta)$ are bounded within a certain interval, and consequently other physical quantities (like $\text{Tr}(\rho\theta)$ with any density matrix ρ) to which $\text{Tr}\mathcal{C}_Q(\theta)$ is a bound, are also bounded.

In terms of applications, the Choquet integral has been used to derive bounds for various physical quantities (proposition IV.3, and section VIII for the partition function). A desirable feature of the formalism, is that it is robust in the presence of noise. The reason is that noise affects in a uniform way all coherent states, and does not change the ranking significantly. At the same time the formalism is sensitive enough to detect changes in the ground state of physical systems, because they affect the ranking. Examples of this have been given in sections VII A, VII B.

From a practical point of view, calculations are easy if they involve only the $\mathrm{Tr}\mathcal{C}_Q(\theta)$. This simply requires the Q-function and its ranking in Eq.(66) (see Eq.(76)). If the full $\mathcal{C}_Q(\theta)$ is required, as for example in Eq.(124), then the calculation of the projectors $\varpi_\theta(i|i+1,...,d^2)$ is needed, and this can be computationally more intensive.

There are many figures of merit in Physics. They are used in bounds for the values of physical quantities. They are also used to derive orders in sets of physical quantities (e.g., various entropic quantities define 'more mixed' or 'more entangled', etc). In this paper we introduced the Choquet integral and the concept of comonotonicity, which are motivated by non-additive probabilities associated with overlapping alternatives, and which we have used to derive bounds to physical quantities, and study the lowest state of physical systems.

We have considered positive semidefinite operators, but the work could be extended to all Hermitian operators. Also we have used the Q-function, but a similar formalism that involves the P-function can also be developed. The work provides a deeper insight to the use of non-orthogonal overcomplete sets of states (like coherent states) for the study of physical problems.

[1]   A. Wehrl, Rev. Mod. Phys. 50, 221 (1978)
[2]   A. Vourdas, Rep. Prog. Phys. 67, 267 (2004)
[3]   A. Vourdas, J. Phys. A40, R285 (2007)
[4]   A.P. Dempster, Annal. Math. Stat. 38, 325 (1967)
[5]   G. Shafer, 'A Mathematical Theory of Evidence', (Princeton Univ. Press, Princeton, 1976)
[6]   J.Y. Halpern, R. Fagin, Artificial Intelligence 54, 275 (1992)
[7]   R. Yager, L. Liu (Ed.), 'Classic works of the Dempster-Shafer theory of belief functions' (Springer, Berlin, 2008)
[8]   C. Camerer, M. Weber, J. Risk and Uncertainty, 5, 325 (1992)
[9]   A. Vourdas, J. Math. Phys. 55, 082107 (2014)
[10] G. Choquet, Ann. Inst. Fourier, 5,131 (1953)
[11] M. Grabisch, C. Labreuche, Ann. Operations Res. 175, 247 (2010)
[12] D. Donneberg, 'Non-additive measure and integral' (Kluwer, Dordrecht, 1984)
[13] D. Schmeidler, Proc. Am. Math. Soc. 97, 255 (1986)
[14] I. Gilboa, D. Schmeidler, Ann. Oper. Research 52, 43 (1994)
[15] T. Murofushi, M. Sugeno, J. Math. Anal. Appl. 159, 532 (1991)
[16] P. Wakker, J. Econ. Theory, 52, 453 (1990)
[17] M.J. Machina, Econometrica, 50, 277 (1982)
[18] M.E. Yaari, Econometrica, 55, 95 (1987)
[19] J. Quiggin, J. Econ. Behav. Organ. 3, 323 (1982)
[20] A. Vourdas, J. Geom. Phys. 101, 38 (2016)
[21] G.C. Rota, Z. Wahrseheinlichkeitstheorie 2, 340 (1964)
[22] M. Barnabei, A. Brini, G.C. Rota, Russian Math. Surveys, 41, 135 (1986)
[23] S. Zhang, A. Vourdas, J. Phys. A37, 8349 (2004)
[24] N. Cotfas, J-P Gazeau, A. Vourdas, J. Phys. A44, 175303 (2011)
[25] J.R. Klauder, B-S Skagerstam (Ed.) 'Coherent states' ((World Sci., Singapore, 1985)
[26] S.T. Ali, J-P Antoine, J-P Gazeau, 'Coherent states, wavelets and their generalizations' (Springer, Berlin, 2000)
[27] A.M. Perelomov, 'Generalized coherent states and their applications' (Springer, Berlin, 1986)
[28] E.H. Lieb, Commun. Math. Phys. 62, 35 (1978)
[29] G. Birkhoff, J. von Neumann, Ann. Math. 37, 823 (1936)
[30] C. Piron, 'Foundations of quantum physics', Benjamin, New York, 1976
[31] J. Jauch, 'Foundations of quantum mechanics', Addison-Wesly, Reading, 1968
[32] G.H. Golup, C.F. van Loan, 'Matrix Computations', (John Hopkins Univ. Press, Maryland, 1989)
[33] X.M. Yang, X.Q. Yang, K.L. Teo, J. Math. Anal. Appl. 263, 327 (2001)
[34] E. Romera, R. del Real, M. Calixto, Phys. Rev. A85, 053831 (2012)
[35] M. Marcus, H. Minc, 'A survey of matrix theory and matrix inequalities', (Dover, New York, 1992)
[36] E. Carlen, Contemp. Math., 529, 73 (2009)
[37] M.B. Ruskai, J. Math. Phys., 43, 4358 (2002)
[38] E.H. Lieb, Commun. Math. Phys. 31, 327 (1973)

TABLE I: The three dominant values of $Q(\alpha, \beta)$, the eigenvalues e_1, e_2, e_3, and the dominance ratio $r(\theta)$ of the operator θ in Eq.(94). In the first rwo $r_i = 0$ (there is no noise). In the other five rows r_i are uniformly distributed random numbers in the interval $(-1, 1)$. The overlaps $\tau_i = |\langle u_i | v_i \rangle|^2$ of the eigenvectors in the noisy cases, with their counterparts in the noiseless case are also shown.

$Q(9|\theta)$	$Q(8|\theta)$	$Q(7|\theta)$	e_1	e_2	e_3	$r(\theta)$	τ_1	τ_2	τ_3
$Q(1,2) = 3.023$	$Q(1,1) = 3.023$	$Q(0,2) = 2.095$	0.942	5.488	12.569	0.428	1	1	1
$Q(1,2) = 3.447$	$Q(1,1) = 2.926$	$Q(0,2) = 2.171$	0.604	4.993	13.235	0.453	0.973	0.979	0.992
$Q(1,2) = 3.173$	$Q(1,1) = 2.897$	$Q(0,0) = 2.398$	1.337	6.743	12.230	0.416	0.987	0.990	0.996
$Q(1,1) = 2.911$	$Q(1,2) = 2.506$	$Q(0,1) = 1.865$	0.809	4.245	11.380	0.443	0.990	0.985	0.980
$Q(1,2) = 3.157$	$Q(1,1) = 2.962$	$Q(0,2) = 2.278$	0.747	4.454	13.111	0.458	0.997	0.988	0.988
$Q(1,2) = 3.316$	$Q(1,1) = 3.180$	$Q(0,1) = 2.436$	0.836	5.774	13.671	0.413	0.967	0.950	0.980

TABLE II: Comonotonicity intervals, the corresponding three dominant values of $Q(\alpha, \beta)$, and the dominance ratio $r[\theta(\lambda)]$, for the operator $\theta(\lambda)$ in Eq.(106).

intervals of λ	$Q[9|\theta(\lambda)]$	$Q[8|\theta(\lambda)]$	$Q[7|\theta(\lambda)]$	$r[\theta(\lambda)]$
$I_1 = (0, 0.06)$	$Q(1,0)$	$Q(0,0)$	$Q(2,0)$	$\frac{13.5 + 7\lambda}{99 + 63\lambda}$
$I_2 = (0.06, 0.44)$	$Q(0,0)$	$Q(1,0)$	$Q(2,0)$	$\frac{13.5 + 7\lambda}{99 + 63\lambda}$
$I_3 = (0.44, 0.56)$	$Q(0,0)$	$Q(1,0)$	$Q(0,2)$	$\frac{12.62 + 9\lambda}{99 + 63\lambda}$
$I_4 = (0.56, 0.6)$	$Q(0,0)$	$Q(0,2)$	$Q(1,0)$	$\frac{12.62 + 9\lambda}{99 + 63\lambda}$
$I_5 = (0.6, 0.7)$	$Q(0,0)$	$Q(0,2)$	$Q(0,1)$	$\frac{11 + 11.7\lambda}{99 + 63\lambda}$

TABLE III: The three dominant values of $Q[\alpha, \beta | \mathfrak{P}(\lambda)]$, and the dominance ratio $r[\mathfrak{P}(\lambda)]$ as a function of λ, for the $\mathfrak{P}(\lambda) = |g(\lambda)\rangle\langle g(\lambda)|$ where $|g(\lambda)\rangle$ is the ground state of the system described with the Hamiltonian $\mathfrak{H}(\lambda)$ in Eq.(119). The Wehrl entropy $E[\mathfrak{P}(\lambda)]$ (in nats) and the $|\langle g(0)|g(\lambda)\rangle|^2$ are also shown. Horizontal lines indicate that we cross from one equivalence class to another

λ	$Q[9|\mathfrak{P}(\lambda)]$	$Q[8|\mathfrak{P}(\lambda)]$	$Q[7|\mathfrak{P}(\lambda)]$	$r[\mathfrak{P}(\lambda)]$	$E[\mathfrak{P}(\lambda)]$	$|\langle g(0)|g(\lambda)\rangle|^2$
0.0	$Q[1,2|\mathfrak{P}(\lambda)]=0.214$	$Q[1,0|\mathfrak{P}(\lambda)]=0.214$	$Q[1,1|\mathfrak{P}(\lambda)]=0.214$	0.642	1.929	1
0.1	$Q[1,2|\mathfrak{P}(\lambda)]=0.228$	$Q[1,0|\mathfrak{P}(\lambda)]=0.209$	$Q[1,1|\mathfrak{P}(\lambda)]=0.191$	0.628	1.948	0.971
0.2	$Q[1,2|\mathfrak{P}(\lambda)]=0.245$	$Q[1,0|\mathfrak{P}(\lambda)]=0.202$	$Q[1,1|\mathfrak{P}(\lambda)]=0.166$	0.613	1.972	0.929
0.3	$Q[1,2|\mathfrak{P}(\lambda)]=0.268$	$Q[1,0|\mathfrak{P}(\lambda)]=0.192$	$Q[1,1|\mathfrak{P}(\lambda)]=0.138$	0.598	1.980	0.885
0.4	$Q[1,2|\mathfrak{P}(\lambda)]=0.297$	$Q[1,0|\mathfrak{P}(\lambda)]=0.176$	$Q[2,2|\mathfrak{P}(\lambda)]=0.120$	0.593	1.959	0.812
0.5	$Q[1,2|\mathfrak{P}(\lambda)]=0.313$	$Q[0,2|\mathfrak{P}(\lambda)]=0.168$	$Q[2,2|\mathfrak{P}(\lambda)]=0.157$	0.638	1.889	0.666
0.6	$Q[1,2|\mathfrak{P}(\lambda)]=0.298$	$Q[0,2|\mathfrak{P}(\lambda)]=0.216$	$Q[2,2|\mathfrak{P}(\lambda)]=0.187$	0.701	1.816	0.477
0.7	$Q[1,2|\mathfrak{P}(\lambda)]=0.268$	$Q[0,2|\mathfrak{P}(\lambda)]=0.242$	$Q[2,2|\mathfrak{P}(\lambda)]=0.200$	0.710	1.825	0.330
0.8	$Q[0,2|\mathfrak{P}(\lambda)]=0.252$	$Q[1,2|\mathfrak{P}(\lambda)]=0.241$	$Q[2,2|\mathfrak{P}(\lambda)]=0.202$	0.695	1.845	0.240
0.9	$Q[0,2|\mathfrak{P}(\lambda)]=0.256$	$Q[1,2|\mathfrak{P}(\lambda)]=0.221$	$Q[2,2|\mathfrak{P}(\lambda)]=0.200$	0.677	1.862	0.185
1.0	$Q[0,2|\mathfrak{P}(\lambda)]=0.257$	$Q[1,2|\mathfrak{P}(\lambda)]=0.207$	$Q[2,2|\mathfrak{P}(\lambda)]=0.198$	0.662	1.873	0.149

TABLE IV: The three eigenvalues e_1, e_2, e_3 of the Hamiltonian in Eq.(122). The three dominant values of $Q[\alpha, \beta | \mathfrak{P}(\lambda)]$ for the $\mathfrak{P}(\lambda) = |g(\lambda)\rangle\langle g(\lambda)|$ where $|g(\lambda)\rangle$ is the ground state of the system, are also shown. In the case $\lambda = 0$ the two lowest eigenvalues are equal to each other, and the $\mathfrak{P}(0)$ is the projector to the corresponding two-dimensional eigenspace. In this case we present the $Q[\alpha, \beta | \frac{1}{2}\mathfrak{P}(0)]$

λ	e_1	e_2	e_3	$Q[9|\mathfrak{P}(\lambda)]$	$Q[8|\mathfrak{P}(\lambda)]$	$Q[7|\mathfrak{P}(\lambda)]$	$E[\mathfrak{P}(\lambda)]$
-0.01	0.494	0.509	6.495	$Q[0,1|\mathfrak{P}(-0.01)]=0.194$	$Q[0,2|\mathfrak{P}(-0.01)]=0.194$	$Q[1,1|\mathfrak{P}(0.01)]=0.163$	1.960
0.00	0.500	0.500	6.500	$Q[1,1|\frac{1}{2}\mathfrak{P}(0)]=0.160$	$Q[1,2|\frac{1}{2}\mathfrak{P}(0)]=0.160$	$Q[2,1|\frac{1}{2}\mathfrak{P}(0)]=0.155$	2.088
0.01	0.490	0.505	6.505	$Q[2,1|\mathfrak{P}(0.01)]=0.218$	$Q[2,2|\mathfrak{P}(0.01)]=0.218$	$Q[1,1|\mathfrak{P}(0.01)]=0.155$	1.890
$-0.01i$	0.492	0.507	6.500	$Q[2,2|\mathfrak{P}(-0.01i)]=0.292$	$Q[1,2|\mathfrak{P}(-0.01i)]=0.291$	$Q[0,2|\mathfrak{P}(-0.01i)]=0.255$	1.627
$0.01i$	0.493	0.507	6.500	$Q[2,1|\mathfrak{P}(0.01i)]=0.292$	$Q[1,1|\mathfrak{P}(0.01i)]=0.291$	$Q[0,1|\mathfrak{P}(0.01i)]=0.255$	1.627