Link to publisher's version: *https://doi.org/10.1016/j.aop.2018.02.010*

Citation: Vourdas A (2018) Independence and totalness of subspaces in phase space methods. Annals of Physics. 391: 83-111.

Independence and totalness of subspaces in phase space methods

A. Vourdas
Department of Computer Science,
University of Bradford,
Bradford BD7 1DP, UK

The concepts of independence and totalness of subspaces are introduced in the context of quasi-probability distributions in phase space, for quantum systems with finite-dimensional Hilbert space. It is shown that due to the non-distributivity of the lattice of subspaces, there are various levels of independence, from pairwise independence up to (full) independence. Pairwise totalness, totalness and other intermediate concepts are also introduced, which roughly express that the subspaces overlap strongly among themselves, and they cover the full Hilbert space. A duality between independence and totalness, that involves orthocomplementation (logical NOT operation), is discussed. Another approach to independence is also studied, using Rota's formalism on independent partitions of the Hilbert space. This is used to define informational independence, which is proved to be equivalent to independence. As an application, the pentagram (used in discussions on contextuality) is analyzed using these concepts.

I. INTRODUCTION

Phase space methods[1–3] play an important role in quantum mechanics. They study various quasi-probability distributions which are the analogues of joint probabilities, for non-commuting variables like the position and momentum. In this paper we consider quantum systems with finite Hilbert space $H(d)$, and study a wide class of such functions $R(i)$ related to projectors in subspaces $H_1, ..., H_n$ of $H(d)$. Special cases are the Q-function, the probability distribution in position space, etc. In this context we study two concepts, independence and totalness.

Linear independence (which for simplicity we call independence) is a very fundamental concept in the theory of vector spaces, and in other areas like quantum mechanics that depend on it. A deeper study of this concept led to the subject of matroids[4–6], which defines independence through some axioms, and then defines the concepts of rank and basis. In a different context, independence has been used within the general framework of the continuous geometries by von Neumann[7, 8]. Another approach by Rota and collaborators [9–15], defines independent partitions of the Hilbert space.

Ideas from all these areas are incorporated in the present formalism. We show that there are various levels of independence for the subspaces $\{H_1, ..., H_n\}$. We use indistinguishably the terms independence and disjointness, but for simplicity in most of the paper we use the term independence only. We also introduce the concept of totalness, which requires strong overlap between the subspaces, and which is dual to the concept of independence. We show that there are various levels of totalness. The existence of various levels of independence and totalness, is intimately related to the non-distributive nature of the lattice of subspaces.

Distributivity is a very fundamental property in classical physics and classical (Boolean) logic. For example, a student studies a compulsory module C, and in addition to that he has to choose one of two optional modules O_1 or O_2. The following statements are equivalent:

- He will study the module C and in addition to that module O_1 or module O_2.
- He will study the modules C and O_1 or he will study the modules C and O_2.

The equivalence looks trivial, because distributivity is deeply embedded in our everyday language and the classsical world, which are formally described with set theory and Boolean algebra. In Quantum Mechanics distributivity does not hold, and we need to develop appropriate language that describes this and plays complementary role to non-commutativity. Concepts which are trivially equivalent in a distributive structure, might become inequivalent in a non-distributive structure. The various levels of independence (or disjointness), and the various levels of totalness of sets $\{H_1, ..., H_n\}$ of subspaces of $H(d)$, are examples of this.

More specifically, in this paper:

- We introduce the concept of n independent subspaces, which is generalization of independent vectors. We show that for $n \geq 3$, independence is stronger concept than pairwise independence. This is related to the non-distributive nature of the lattice of subspaces. We also introduce various intermediate concepts of independence, and the degree of independence.

- We introduce total sets of n subspaces, which are extensions of total sets of vectors. Totalness means not just covering of the full Hilbert space $H(d)$, but also strong overlap between the subspaces. We show that for $n \geq 3$ totalness is stronger than pairwise totalness. This is related to the fact that the lattice of subspaces is non-distributive. We also introduce various intermediate concepts of totalness, and the degree of totalness.

- There is a duality between a set of independent subspaces, and the total set of the orthocomplements of these subspaces. Orthocomplementation (logical NOT operation) transforms independence into totalness. These ideas are interpreted in terms of measurements with projectors to these subspaces.

- A measurement with a projector to a particular subspace, might give the same result for two different states. For a given measurement, we partition the Hilbert space into sets (blocks) of states, such that this measurement gives the same result for all states in each block (when the outcome is 'yes'). We then introduce the concept of informationally independent measurements, and show that it is equivalent to independence. This part, links the present work with Rota's formalism on independent partitions[9–15].

- Using these concepts we discuss the pentagram, which is used in contextuality[16–24]. The pentagram within a non-contextual hidden variable theory, uses marginals of joint probability distributions. They are based on the law of total probability, which in turn depends on distributivity. In non-distributive structures joint probabilities and their marginals are problematic (e.g., joint probabilities of non-commuting variables).

- Within the full lattice of subspaces which is non-distributive, there are sublattices which are distributive (e.g., a sublattice generated by commuting subspaces). In these 'islands', independence is equivalent to pairwise independence, totalness is equivalent to pairwise totalness, joint probabilities and their marginals are well defined, etc.

Overall, the development of such concepts provides a complementary approach to non-commutativity. Quantum theory is usually described through non-commutativity, and in this paper it is described through non-distributivity.

In section II we introduce within set theory, the concept of disjointness or independence and also the dual concept of totalness. In set theory distributivity holds, and there is a single concept of independence and a single concept of totalness. These two concepts define partitions, which are used in the law of total probability, and in defining marginals from probability distributions.

In section III we present briefly the lattice of subspaces. We also use projectors to the subspaces $\{H_i\}$ to define a generalized phase space function $R(i)$.

In section IV we introduce various levels of independence, and define the degree of independence. In section V we introduce various levels of totalness, and define the degree of totalness.

In section VI we use independent partitions, to define the concept of informationally independent subspaces and measurements. We also show that informational independence is equivalent to independence. Weaker concepts of independence (like pairwise independence), are not informationally independent.

As an application of these ideas, we discuss in section VII the pentagram, which is used in discussions on contextuality. We conclude in section VIII with a discussion of our results.

II. DISJOINTNESS AND TOTALNESS IN SET THEORY

We consider the set of all subsets of a finite set Ω (the powerset 2^Ω). In it we define the conjunction (logical AND), disjunction (logical OR), and negation (logical NOT), as the intersection, union and complement:

$$A_1 \wedge A_2 = A_1 \cap A_2; \qquad A_1 \vee A_2 = A_1 \cup A_2; \quad \neg A = \Omega \setminus A. \tag{1}$$

The powerset 2^Ω with these operations is a Boolean algebra. The corresponding partial order $\prec$ is 'subset'. The smallest element is the empty set $\mathcal{O} = \emptyset$, and the largest element is $\mathcal{I} = \Omega$.

Definition II.1.

(1) The subsets $A_1, ..., A_n$ of Ω are independent or disjoint, if

$$(A_1 \vee ... \vee A_{i-1} \vee A_{i+1} \vee ... \vee A_n) \wedge A_i = \emptyset. \tag{2}$$

for all $i = 1, ..., n$.

(2) The subsets $A_1, ..., A_n$ of Ω are pairwise independent or pairwise disjoint, if $A_i \wedge A_j = \emptyset$ for all i, j.

(3) The subsets $A_1, ..., A_n$ of Ω are weakly independent or weakly disjoint, if

$$A_1 \wedge ... \wedge A_n = \emptyset. \tag{3}$$

Definition II.2.

(1) The subsets $A_1, ..., A_n$ of Ω form a total set, if

$$(A_1 \wedge ... \wedge A_{i-1} \wedge A_{i+1} \wedge ... \wedge A_n) \vee A_i = \Omega. \tag{4}$$

for all $i = 1, ..., n$:

(2) The subsets $A_1, ..., A_n$ of Ω form a pairwise total set, if $A_i \vee A_j = \Omega$ for all i, j.

(3) The subsets $A_1, ..., A_n$ of Ω form a weakly total set of subsets, if

$$A_1 \vee ... \vee A_n = \Omega. \tag{5}$$

The subsets $A_1, ..., A_n$ form a partition, if they are disjoint (independent) and they also form a weakly total set.

Proposition II.3. *In set theory:*

(1) Independence is equivalent to pairwise independence. Independence is stronger concept than weak independence (they are equivalent for $n = 2$).

(2) Totalness is equivalent to pairwise totalness. Totalness is stronger concept than weak totalness (they are equivalent for $n = 2$).

Proof. (1) Using the distributivity property of set theory, we rewrite Eq.(2) as

$$(A_1 \wedge A_i) \vee ... \vee (A_{i-1} \wedge A_i) \vee (A_{i+1} \wedge A_i) \vee ... \vee (A_n \wedge A_i) = \emptyset. \tag{6}$$

This shows that $A_i \wedge A_j = \emptyset$ and therefore independence is equivalent to pairwise independence.

(2) Using the distributivity property of set theory, we rewrite Eq.(4) as

$$(A_1 \vee A_i) \wedge ... \wedge (A_{i-1} \vee A_i) \wedge (A_{i+1} \vee A_i) \wedge ... \wedge (A_n \vee A_i) = \Omega. \tag{7}$$

This shows that $A_i \vee A_j = \Omega$ and therefore totalness is equivalent to pairwise totalness.

□

Proposition II.4. *The $A_1, ..., A_n$ are a total set of subsets, if and only if the $\neg A_1, ..., \neg A_n$ are independent subsets of Ω.*

Proof. Using de Morgan's rule, the negation of $A_i \wedge A_j = \emptyset$, gives $\neg A_i \vee \neg A_j = \Omega$. □

A. Marginal distributions: distributivity and the law of total probability

The marginals of joint probability distributions are based on the law of total probability in Kolmogorov's probability theory.

Proposition II.5. *Let Ω be a set of alternatives, $B_1, ..., B_n$ a partition of the set Ω and $A \subseteq \Omega$. The law of total probability states that*

$$p(A) = \sum_i p(A \cap B_i). \tag{8}$$

Proof. Using the distributivity property of set theory we get

$$A = A \cap \Omega = A \cap (B_1 \cup ... \cup B_n) = (A \cap B_1) \cup ... \cup (A \cap B_n). \tag{9}$$

Since $(A \cap B_i) \cap (A \cap B_j) = \emptyset$, we use the additivity property of Kolmogorov probabilities

$$S_1 \cap S_2 = \emptyset \;\rightarrow\; p(S_1 \cup S_2) = p(S_1) + p(S_2), \tag{10}$$

we prove Eq.(8). □

Eq.(8) can be used to define marginals of probability distributions. The ingredients for the law of total probability, are partitions and distributivity[25]. Partitions are based on the concepts of disjointness (independence) and also weak totalness. We will show below that in non-distributive structures, there are various levels of disjointness (independence) and various levels of totalness. Consequently the relationship between a joint probability distribution and its marginals becomes problematic.

III. QUANTUM SYSTEMS WITH VARIABLES IN $\mathbb{Z}(d)$

We consider a quantum system $\Sigma(d)$ with variables in $\mathbb{Z}(d)$ (the integers modulo d), with states in a d-dimensional Hilbert space $H(d)$[26, 27]. We also consider an orthonormal basis of 'position states' which we denote as $|X;\alpha\rangle$, where the $a \in \mathbb{Z}(d)$, and the X in the notation indicates position states. We also consider another orthonormal basis of 'momentum states' which we denote as $|P;\beta\rangle$, where the P in the notation indicates momentum states. They are related to the position states through a finite Fourier transform:

$$|P;\beta\rangle = \frac{1}{\sqrt{d}}\sum_{\alpha}\omega(\alpha\beta)|X;\alpha\rangle;\quad \omega(\alpha)=\exp\left(i\frac{2\pi\alpha}{d}\right);\quad \alpha,\beta\in\mathbb{Z}(d). \tag{11}$$

Displacement operators in the $\mathbb{Z}(d)\times\mathbb{Z}(d)$ phase space of this system, are defined as

$$D(\alpha,\beta)=Z^{\alpha}X^{\beta}\omega(-2^{-1}\alpha\beta);\quad Z=\sum_m \omega(m)|X;m\rangle\langle X;m|;\quad X=\sum_m |X;m+1\rangle\langle X;m| \tag{12}$$

The factor 2^{-1} above, is an element of $\mathbb{Z}(d)$, and it exists only for odd d. The formalism of finite quantum systems, is slightly different in the cases of odd and even d. Below, in the formulas that use the displacement operators, we assume that the dimension d is an odd integer.

Acting with $D(\alpha,\beta)$ on a 'generic' and normalized fiducial vector $|f\rangle$

$$|f\rangle=\sum f_m|X;m\rangle;\quad \sum|f_m|^2=1, \tag{13}$$

we get the following d^2 states which we call coherent states[26, 27]

$$|C;\alpha,\beta\rangle=D(\alpha,\beta)|f\rangle=\sum A_m(\alpha,\beta)|X;m\rangle;\quad A_m(\alpha,\beta)=\omega(am-2^{-1}\alpha\beta)f_{m-\beta};\quad \alpha,\beta\in\mathbb{Z}(d). \tag{14}$$

The C in the notation indicates coherent states. We can write the $A_m(\alpha,\beta)$ as a $d\times d^2$ matrix, with indices m and the pair (α,β) written as one index. Then the requirement of a generic fiducial vector is that the rank of this matrix is d. In this case any d of the d^2 coherent states are linearly independent.

The coherent states obey the resolution of the identity:

$$\frac{1}{d}\sum_{\alpha,\beta}|C;\alpha,\beta\rangle\langle C;\alpha,\beta|=\mathbf{1}. \tag{15}$$

We will use the notation $H(X;\alpha)$, $H(P;\alpha)$, $H(C;\alpha,\beta)$ for the one-dimensional subspaces of $H(d)$ that contain the states $|X;\alpha\rangle$, $|P;\alpha\rangle$, $|C;\alpha,\beta\rangle$, correspondingly. We will also use the notation $\Pi[H(X;\alpha)]$, $\Pi[H(P;\alpha)]$, $\Pi[H(C;\alpha,\beta)]$ for the projectors to these subspaces:

$$\begin{aligned}
&\Pi[H(X;\alpha)]=|X;\alpha\rangle\langle X;\alpha|;\quad \sum_\alpha \Pi[H(X;\alpha)]=\mathbf{1}\\
&\Pi[H(P;\alpha)]=|P;\alpha\rangle\langle P;\alpha|;\quad \sum_\alpha \Pi[H(P;\alpha)]=\mathbf{1};\\
&\Pi[H(C;\alpha,\beta)]=|C;\alpha,\beta\rangle\langle C;\alpha,\beta|;\quad \frac{1}{d}\sum_{\alpha,\beta}\Pi[H(C;\alpha,\beta)]=\mathbf{1}.
\end{aligned} \tag{16}$$

A. The lattice $\mathcal{L}(d)$ of subspaces

The Birkhoff-von Neumann lattice of the closed subspaces of the Hilbert space, with the operations of conjunction, disjunction and complementation, has been studied extensively in the literature [28–33].

We consider the finite-dimensional Hilbert space $H(d)$, describing the system $\Sigma(d)$. In the set of subspaces of $H(d)$, we define the conjunction (logical AND) and disjunction (logical OR) [34–38]:

$$H_1\wedge H_2=H_1\cap H_2;\qquad H_1\vee H_2=\mathrm{span}(H_1\cup H_2). \tag{17}$$

We stress that the logical OR is not just the union, but it contains superpositions of states in the two spaces. This will lead later to the distinction between pairwise independence and independence.

The set of subspaces of $H(d)$ with these operations is a lattice, which we denote as $\mathcal{L}(d)$. The corresponding partial order $\prec$ is 'subspace'. The smallest element is $\mathcal{O} = H(0)$ (the zero-dimensional subspace that contains only the zero vector), and the largest element is $\mathcal{I} = H(d)$.

The lattice $\mathcal{L}(d)$ is not distributive. $\mathcal{L}(d)$ is a modular orthocomplemented lattice. Modularity is a weak version of distributivity, and is related to independence. Birkhoff discussed the link between matroids (which introduce independence in an abstract way) and modular lattices[34].

Modularity states that

$$H_1 \prec H_3 \;\rightarrow\; H_1 \vee (H_2 \wedge H_3) = (H_1 \vee H_2) \wedge H_3. \tag{18}$$

Equivalent to this is the following relation which is valid for any H_1, H_2, H_3:

$$H_1 \wedge (H_2 \vee H_3) = H_1 \wedge [H_3 \vee (H_2 \wedge (H_1 \vee H_3))]. \tag{19}$$

Each subspace has an infinite number of complements. The orthocomplement of H_1 is unique, and is another subspace which we denote as $H_1^\perp$, with the properties

$$\begin{aligned} &H_1 \wedge H_1^\perp = \mathcal{O}; \quad H_1 \vee H_1^\perp = \mathcal{I} = H(d); \quad (H_1^\perp)^\perp = H_1 \\ &(H_1 \wedge H_2)^\perp = H_1^\perp \vee H_2^\perp; \quad (H_1 \vee H_2)^\perp = H_1^\perp \wedge H_2^\perp \\ &\dim(H_1) + \dim(H_1^\perp) = d. \end{aligned} \tag{20}$$

Orthocomplementation is related to logical NOT, in the description of quantum measurements. We will use the notation $\Pi(H_1)$ for the projector to the subspace H_1. Then

$$\Pi^\perp(H_1) = \Pi(H_1^\perp) = \mathbf{1} - \Pi(H_1). \tag{21}$$

$H_1^\perp$ is the null space of $\Pi(H_1)$, and H_1 is the null space of $\Pi(H_1^\perp)$.

A measurement with $\Pi(H_1)$ on a state $|s\rangle$, will give:

- 'yes' with probability $p = \langle s|\Pi(H_1)|s\rangle$, in which case the state will collapse into $\frac{1}{\sqrt{p}}\Pi(H_1)|s\rangle$
- 'no' with probability $1 - p = \langle s|\Pi(H_1^\perp)|s\rangle$, in which case the state will collapse into $\frac{1}{\sqrt{1-p}}\Pi(H_1^\perp)|s\rangle$.

An important property of modular lattices [34], is that

$$\dim(H_1 \vee H_2) + \dim(H_1 \wedge H_2) = \dim(H_1) + \dim(H_2). \tag{22}$$

Definition III.1. H_1 commutes with H_2 (we denote this as $H_1 \mathcal{C} H_2$) if

$$H_1 = (H_1 \wedge H_2) \vee (H_1 \wedge H_2^\perp) \tag{23}$$

It can be proved that $H_1 \mathcal{C} H_2$ if and only if $[\Pi(H_1), \Pi(H_2)] = 0$. Commutativity of subspaces is equivalent to commutativity of the projectors to these subspaces.

It is easily seen that:

- $H_1 \prec H_2$ implies that $H_1 \mathcal{C} H_2$. Therefore H_1 commutes with $H_1 \vee H_2$ and $H_1 \wedge H_2$ (for any H_2).
- Since $\mathcal{L}(d)$ is an modular lattice, if $H_1 \mathcal{C} H_2$, then $H_2 \mathcal{C} H_1$ and also $H_1 \mathcal{C} H_2^\perp$, $H_1^\perp \mathcal{C} H_2$.
- Every subspace commutes with $H(d)$ and $\mathcal{O}$.
- If $H_1 \mathcal{C} H_2$ and $H_2 \mathcal{C} H_3$, then the H_1, H_3 might not commute (transitivity does not hold).

Within the lattice $\mathcal{L}(d)$ which is non-distributive, there are sublattices which are distributive. For example, any sublattice of $\mathcal{L}(d)$ generated by commuting subspaces, is distributive. In these 'islands' results similar to classical physics do hold. For example, the law of the total probability holds, joint probability distributions and their marginals are well defined, etc.

B. Quasi-probability distributions

Below we consider a set $\{H_1, ..., H_n\}$ of $n \geq 2$ proper subspaces of $H(d)$ (which might not have the same dimension).

Notation III.2.

$$\begin{aligned}\mathfrak{H}_i &= \bigvee_{j \neq i} H_j = H_1 \vee ... \vee H_{i-1} \vee H_{i+1} \vee ... \vee H_n \\ \mathfrak{H}_i^{\perp} &= \bigwedge_{j \neq i} H_j^{\perp} = H_1^{\perp} \wedge ... \wedge H_{i-1}^{\perp} \wedge H_{i+1}^{\perp} \wedge ... \wedge H_n^{\perp}\end{aligned} \tag{24}$$

Also

$$\begin{aligned}\mathfrak{h}_i &= \bigwedge_{j \neq i} H_j = H_1 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n \\ \mathfrak{h}_i^{\perp} &= \bigvee_{j \neq i} H_j^{\perp} = H_1^{\perp} \vee ... \vee H_{i-1}^{\perp} \vee H_{i+1}^{\perp} \vee ... \vee H_n^{\perp}\end{aligned} \tag{25}$$

Let ρ be a density matrix, and

$$R(i) = \mathrm{Tr}[\rho \Pi(H_i)] \geq 0; \quad i = 1, ..., n \tag{26}$$

For a given i, $R(i)$ is the probability that the measurement $\Pi(H_i)$ will give the outcome 'yes'. However the set $\{R(i)|i = 1, ..., n\}$ is not in general a probability distribution, but it can be viewed as a quasi-probability distribution. This is related to the lack of independence between the subspaces H_i, and we study this in depth taking into account the non-distributivity of the quantum structure. The index i might be a k-tuple $(\alpha_1, ..., \alpha_k)$ which takes a finite number of values.

The concepts of independence and totalness underpin this formalism. Due to the non-distributivity of the quantum structure, both of these concepts are more complex than in set theory discussed earlier in section II.

Later, in the study of independence and totalness, we will use two more quasi-probability distributions:

$$\widetilde{R}(i) = \mathrm{Tr}[\rho \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)]; \quad \widehat{R}(i) = \mathrm{Tr}[\rho \Pi(\mathfrak{h}_i \vee H_i)]. \tag{27}$$

Since $\mathfrak{H}_i^{\perp} \wedge H_i \prec H_i \prec \mathfrak{h}_i \vee H_i$, it follows that

$$0 \leq \widetilde{R}(i) \leq R(i) \leq \widehat{R}(i). \tag{28}$$

The

$$\widetilde{R}(i) = \mathrm{Tr}[\rho \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)] = \mathrm{Tr}[\rho \Pi(H_1^{\perp} \wedge ... \wedge H_{i-1}^{\perp} \wedge H_i \wedge H_{i+1}^{\perp} \wedge ... \wedge H_n^{\perp})], \tag{29}$$

involves the part of the space H_i which overlaps with all $H_j^{\perp}$, and therefore it does not overlap with any of the H_j, for $j \neq i$. The $\Pi(\mathfrak{H}_i^{\perp} \wedge H_i)$ commutes with all $\Pi(H_j)$:

$$[\Pi(H_j), \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)] = 0 \tag{30}$$

The $\widetilde{R}(i)$ is the probability that a measurement $\Pi(\mathfrak{H}_i^{\perp} \wedge H_i)$ on a system with density matrix ρ, will give 'yes'. In this case the state belongs to H_i and it also belongs to all $H_j^{\perp}$, with $j \neq i$. Therefore a simultaneous measurement with $\Pi(H_j)$ will give 'no', if $j \neq i$.

The

$$\widehat{R}(i) = \mathrm{Tr}[\rho\Pi(\mathfrak{h}_i \vee H_i)] = \mathrm{Tr}\{\rho\Pi[(H_1 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n) \vee H_i]\}, \tag{31}$$

involves the disjunction of H_i with the overlap of all H_j (with $j \neq i$). The $\widehat{R}(i)$ is the probability that a measurement $\Pi(\mathfrak{h}_i \vee H_i)$ on a system with density matrix ρ, will give 'yes'. In this case the state collapses to a superposition of a state in H_i and another state which belongs to all H_j with $j \neq i$.

Example III.3. *For the d subspaces $H(X;\alpha)$ we get*

$$R(\alpha) = \langle X;\alpha|\rho|X;\alpha\rangle; \quad \sum_\alpha R(\alpha) = 1; \quad \alpha \in \mathbb{Z}(d). \tag{32}$$

This is the probability distribution in the position space.

Example III.4. *For the $2d$ subspaces*

$$\begin{aligned} H_i &= H(X;i); \quad i = 0, ..., d-1 \\ H_i &= H(P;i-d); \quad i = d, ..., 2d-1 \end{aligned} \tag{33}$$

we get

$$\begin{aligned} &R(i) = \langle X;i|\rho|X;i\rangle; \quad \text{if } i = 0, ..., d-1 \\ &R(i) = \langle P;i-d|\rho|P;i-d\rangle; \quad \text{if } i = d, ..., 2d-1 \\ &\frac{1}{2}\sum_i R(i) = 1. \end{aligned} \tag{34}$$

This distribution consists of both probabilities in position space and probabilities in momentum space. Although such a distribution is not used in the literature, it is interesting to apply the concepts of this paper, to it.

Example III.5. *For the d^2 subspaces $H(C;\alpha,\beta)$ we get*

$$R(\alpha,\beta) = \langle C;\alpha,\beta|\rho|C;\alpha,\beta\rangle; \quad \frac{1}{d}\sum_{\alpha,\beta} R(\alpha,\beta) = 1; \quad \alpha,\beta \in \mathbb{Z}(d), \tag{35}$$

Here the index i is the pair $(\alpha,\beta) \in \mathbb{Z}(d) \times \mathbb{Z}(d)$. $R(\alpha,\beta)$ is the Q-function in the $\mathbb{Z}(d) \times \mathbb{Z}(d)$ phase space.

IV. LEVELS OF INDEPENDENCE

A. Independence

Proposition IV.1. *The subspaces $H_1, ..., H_n$ of $H(d)$ are independent, if one of the following statements, which are equivalent to each other, holds:*

(1) For all $i = 1, ..., n$,

$$\mathfrak{H}_i \wedge H_i = \mathcal{O}. \tag{36}$$

A state cannot belong to both H_i AND to $\mathfrak{H}_i$ (which contains superpositions of states in all H_j with $j \neq i$).

(2) For all $i = 1, ..., n$,

$$[H_1 \vee ... \vee H_{i-1}] \wedge H_i = \mathcal{O}. \quad (37)$$

(3) Any n vectors $|v_1\rangle \in H_1, ..., |v_n\rangle \in H_n$ (one vector from each of the subspaces H_i), are independent:

$$\lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle = 0 \quad \rightarrow \quad \lambda_1 = ... = \lambda_n = 0. \quad (38)$$

Proof. (1) We prove that the first two statements are equivalent. Our proof is related to the one in [8]. The fact that the first statement implies the second one, is trivial. We next prove that if

$$A \wedge H_i = (A \vee H_i) \wedge H_{i+1} = \mathcal{O}; \quad A = H_1 \vee ... \vee H_{i-1} \quad (39)$$

then

$$(A \vee H_{i+1}) \wedge H_i = \mathcal{O}. \quad (40)$$

We use the identity of Eq.(19) with

$$H_1 \; \rightarrow \; H_i; \quad H_2 \; \rightarrow \; H_{i+1}; \quad H_3 \; \rightarrow \; A. \quad (41)$$

and we get

$$H_i \wedge (H_{i+1} \vee A) = H_i \wedge [A \vee (H_{i+1} \wedge (H_i \vee A))]. \quad (42)$$

The assumptions in Eq.(39) show that the right hand side is $\mathcal{O}$, and therefore the left hand side is $\mathcal{O}$. We next use the identity of Eq.(19) with

$$H_1 \; \rightarrow \; H_i; \quad H_2 \; \rightarrow \; H_{i+2}; \quad H_3 \; \rightarrow \; B = A \vee H_{i+1}. \quad (43)$$

and we get

$$H_i \wedge (H_{i+2} \vee B) = H_i \wedge [B \vee (H_{i+2} \wedge (H_i \vee B))]. \quad (44)$$

Using the extra assumption $(H_1 \vee ... \vee H_{i+1}) \wedge H_{i+2} = \mathcal{O}$, we prove that the right hand side is $\mathcal{O}$, and therefore the left hand side is $\mathcal{O}$. We continue in this way, and we prove that the second statement implies the first one.

(2) We prove that the first and third statements, are equivalent. We assume that Eq.(36) holds, and prove that if $\lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle = 0$, then $\lambda_1 = ... = \lambda_n = 0$. Indeed,

$$\lambda_1|v_1\rangle + ... + \lambda_{i-1}|v_{i-1}\rangle + \lambda_{i+1}|v_{i+1}\rangle + ... + \lambda_n|v_n\rangle = -\lambda_i|v_i\rangle \quad (45)$$

From this follows that $\lambda_i = 0$ because the left hand side belongs to $\mathfrak{H}_i$, the right hand side to H_i and $\mathfrak{H}_i \wedge H_i = \mathcal{O}$. Conversely, if $\lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle = 0$ implies that $\lambda_1 = ... = \lambda_n = 0$, then Eq.(36) holds, because if $\mathfrak{H}_i \wedge H_i \neq \mathcal{O}$ then we have solution to Eq.(52) with $\lambda_i \neq 0$.

□

Proposition IV.2.

(1) If the set $\{H_1, ..., H_n\}$ contains independent subspaces, then the subspaces in any subset (with cardinality at least 2) are also independent.

(2) If the subspaces $H_1, ..., H_n$ of $H(d)$ are independent, then

$$\dim(H_1) + ... + \dim(H_n) = \dim(H_1 \vee ... \vee H_n) \leq d. \tag{46}$$

Proof.

(1) If Eq.(36) holds for the set $\{H_1, ..., H_n\}$, then analogous equation holds for any subset of it.

(2) The proof is based on Eq.(22). From Eq.(36) with $i = 1$, it follows that

$$\dim(H_1) + \dim(H_2 \vee ... \vee H_n) = \dim(H_1 \vee ... \vee H_n). \tag{47}$$

The $H_2, ..., H_n$ are independent, and in the same way we prove that

$$\dim(H_2) + \dim(H_3 \vee ... \vee H_n) = \dim(H_2 \vee ... \vee H_n). \tag{48}$$

These two equations give

$$\dim(H_1) + \dim(H_2) + \dim(H_3 \vee ... \vee H_n) = \dim(H_1 \vee ... \vee H_n). \tag{49}$$

We continue in the same way and we prove the proposition.

□

B. Pairwise independence

Definition IV.3. The subspaces $H_1, ..., H_n$ are pairwise independent, if $H_i \wedge H_j = \mathcal{O}$ for all i, j.

Proposition IV.4.

(1) The non-distributivity of the lattice $\mathcal{L}(d)$, implies that independence is stronger concept than pairwise independence.

(2) For subspaces within a distributive sublattice of $\mathcal{L}(d)$, independence is equivalent to pairwise independence. An example, is when the $H_1, ..., H_n$ commute with each other.

Proof.

(1) In every lattice[34–38]

$$\begin{aligned}&(H_1 \wedge H_i) \vee ... \vee (H_{i-1} \wedge H_i) \vee (H_{i+1} \wedge H_i) \vee ... \vee (H_n \wedge H_i)\\ &\prec [H_1 \vee ... \vee H_{i-1} \vee H_{i+1} \vee ... \vee H_n] \wedge H_i; \quad i = 1, ..., n.\end{aligned} \tag{50}$$

Independence implies that the right hand side is $\mathcal{O}$, and then the left hand side is $\mathcal{O}$. This leads to $H_i \wedge H_j = \mathcal{O}$ for all i, j, i.e., pairwise independence. Therefore independence implies independence of every pair of subspaces.

The converse is not true. Pairwise independence implies that the left hand side is $\mathcal{O}$, but this does not imply that the right hand side is $\mathcal{O}$. Therefore pairwise independence does not imply independence.

(2) In distributive lattices Eq.(50) becomes equality. Therefore within a distributive sublattice of $\mathcal{L}(d)$, independence is equivalent to pairwise independence.

□

C. Degree of independence

We have seen that pairwise independence is weaker concept than independence. Between these two concepts, we introduce intermediate concepts which we quantify with the degree of independence.

Proposition IV.5. *Let* $\{H_1, ..., H_n\}$ *be* $n \geq 3$ *pairwise independent subspaces, and* $\mathfrak{H}_i$ *the subspaces in Eq.(24). If* $|v_i\rangle \in \mathfrak{H}_i^{\perp} \wedge H_i$*, then*

$$\lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle = 0 \quad \rightarrow \quad \lambda_1 = ... = \lambda_n = 0. \tag{51}$$

Proof. We assume that $\lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle = 0$. Then

$$\lambda_1|v_1\rangle + ... + \lambda_{i-1}|v_{i-1}\rangle + \lambda_{i+1}|v_{i+1}\rangle + ... + \lambda_n|v_n\rangle = -\lambda_i|v_i\rangle \tag{52}$$

From this follows that $\lambda_i = 0$ because the left hand side belongs to $\mathfrak{H}_i$, the right hand side belongs to $\mathfrak{H}_i^{\perp} \wedge H_i$, and

$$\mathfrak{H}_i \wedge (\mathfrak{H}_i^{\perp} \wedge H_i) = \mathcal{O}. \tag{53}$$

This completes the proof. □

It is seen that the independence relation in Eq.(38) is valid here only for vectors in the subspace $\mathfrak{H}_i^{\perp} \wedge H_i$ of H_i. This is the motivation for introducing in Eq.(27), the quasi-probability distribution $\widetilde{R}(i)$. The degree of independence compares the subspaces $\mathfrak{H}_i^{\perp} \wedge H_i$ and H_i or equivalently the $\widetilde{R}(i)$ with $R(i)$.

Definition IV.6. Let ρ be a density matrix. The matrix for the degree of independence $\mathcal{A}$, and the degree of independence $\eta(\rho)$, are given by

$$\mathcal{A} = \frac{1}{n}\sum_i [\Pi(H_i) - \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)]; \quad \eta(\rho) = \frac{1}{n}\sum_i [R(i) - \widetilde{R}(i)] = \mathrm{Tr}(\rho\mathcal{A}). \tag{54}$$

Each $\Pi(H_i) - \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)$ is a projector. As a sum of projectors, $\mathcal{A}$ is a $d \times d$ positive semidefinite matrix. The various $\Pi(H_i) - \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)$ do not commute, and the corresponding $\mathrm{Tr}\{\rho[\Pi(H_i) - \Pi(\mathfrak{H}_i^{\perp} \wedge H_i)]\}$ can be measured using different ensembles described by the same density matrix ρ.

There are two extreme cases and many intermediate cases:

- If $\mathfrak{H}_i^{\perp} \wedge H_i = H_i$ for all i, then $\mathcal{A} = 0$ and $\eta(\rho) = 0$. Therefore the $\{H_1, ..., H_n\}$ are independent. In this case proposition IV.5 reduces to proposition IV.1, and the independence relation in Eq.(51) holds for all $|v_i\rangle \in H_i$. This is the strongest form of independence.

- If $\mathfrak{H}_i^{\perp} \wedge H_i = \mathcal{O}$ for all i,

$$\mathcal{A} = \frac{1}{n}\sum_i \Pi(H_i); \quad \eta(\rho) = \frac{1}{n}\sum R(i) \tag{55}$$

  and the $\{H_1, ..., H_n\}$ are pairwise independent. The independence relation in Eq.(51) does not hold. This is the weakest form of independence.

- Between these two extreme cases, the $\{H_1, ..., H_n\}$ are partially independent. For a given ρ, $\eta(\rho)$ takes values in the interval

$$0 \leq \eta(\rho) \leq \frac{1}{n}\sum R(i) \tag{56}$$

  In this case the independence relation in Eq.(51) does not hold for all vectors $|v_i\rangle \in H_i$ (it only holds when $|v_i\rangle$ is in the subspace $\mathfrak{H}_i^{\perp} \wedge H_i$ of H_i).

TABLE I: Various levels of independence (or disjointness) and various levels of totalness, for the subspaces $\{H_1, ..., H_n\}$ of $H(d)$.

pairwise independence: $H_i \wedge H_j = \mathcal{O}$	pairwise totalness: $H_i \vee H_j = H(d)$
independence: $H_i \wedge \mathfrak{H}_i = \mathcal{O}$	totalness: $H_i \vee \mathfrak{h}_i = H(d)$
weak independence: $H_1 \wedge ... \wedge H_n = \mathcal{O}$	weak totalness: $H_1 \vee ... \vee H_n = H(d)$
matrix for degree of independence: $\mathcal{A} = \frac{1}{n}\sum[\Pi(H_i) - \Pi(\mathfrak{H}_i^\perp \wedge H_i)]$	matrix for the degree of totalness: $\mathcal{T} = \frac{1}{n}\sum[\Pi(\mathfrak{h}_i \vee H_i) - \Pi(H_i)]$
degree of independence: $\eta(\rho) = \mathrm{Tr}(\rho\mathcal{A}) = \frac{1}{n}\sum[R(i) - \widetilde{R}(i)]$	degree of totalness: $\epsilon(\rho) = \mathrm{Tr}(\rho\mathcal{T}) = \frac{1}{n}\sum[\widehat{R}(i) - R(i)]$

Proposition IV.4 shows that the non-equivalence of independence and pairwise independence (which leads to intermediate concepts) is related to the non-distributivity of the lattice $\mathcal{L}(d)$. In distributive sublattices of $\mathcal{L}(d)$, independence is equivalent to pairwise independence.

A summary of the various levels of independence is shown in table I.

Example IV.7. *This is related to example III.3. We consider the set $\{H(X;\alpha)\}$ with d subspaces of $H(d)$, where $\alpha \in \mathbb{Z}(d)$. These subspaces are pairwise independent. Also*

$$\mathfrak{H}(X;\alpha) = \bigvee_{\beta\neq\alpha} H(X;\beta); \quad [\mathfrak{H}(X;\alpha)]^\perp = H(X;\alpha). \tag{57}$$

and

$$\mathcal{A} = 0; \quad \eta(\rho) = 0. \tag{58}$$

Therefore these subspaces are independent. This is the highest level of independence. In this example $\widetilde{R}(\alpha) = R(\alpha)$ where $R(\alpha)$ has been given in Eq.(32).

Example IV.8. *This is related to example III.4. We consider the set $\{H(X;\alpha), H(P;\beta)\}$ with $2d$ subspaces of $H(d)$, labelled as in Eq.(33). These subspaces are pairwise independent ($H_i \wedge H_j = \mathcal{O}$).*

In this case

$$\begin{aligned} i = 0, ..., (d-1) \;\rightarrow\; \mathfrak{H}_i &= \left(\bigvee_{\beta\neq\alpha} H(X;\beta)\right) \vee \left(\bigvee_{\gamma} H(P;\gamma)\right) = H(d); \quad \mathfrak{H}_i^\perp = \mathcal{O} \\ i = d, ..., (2d-1) \;\rightarrow\; \mathfrak{H}_i &= \left(\bigvee_{\beta} H(X;\beta)\right) \vee \left(\bigvee_{\gamma\neq\alpha} H(P;\gamma)\right) = H(d); \quad \mathfrak{H}_i^\perp = \mathcal{O}. \end{aligned} \tag{59}$$

and

$$\mathcal{A} = \frac{1}{2d}\sum \Pi(H_i) = \frac{1}{d}\mathbf{1}; \quad \eta(\rho) = \frac{1}{d}. \tag{60}$$

Therefore these subspaces are pairwise independent. This is the lowest level of independence. In this example, $\widetilde{R}(i) = 0$.

Example IV.9. *This is related to example III.5. We consider the set* $\{H(C;\alpha,\beta)\}$ *with* d^2 *subspaces of* $H(d)$*. These subspaces are pairwise independent. We have explained earlier that any* d *of the* d^2 *coherent states are linearly independent, and therefore*

$$\mathfrak{H}(C;\alpha_0,\beta_0)=\bigvee_{\alpha\neq\alpha_0,\beta\neq\beta_0}H(C;\alpha,\beta)=H(d);\quad [\mathfrak{H}(C;\alpha_0,\beta_0)]^{\perp}=\mathcal{O} \tag{61}$$

and

$$\mathcal{A}=\frac{1}{d^2}\sum\Pi[H(C;\alpha,\beta)]=\frac{1}{d}\mathbf{1};\quad \eta(\rho)=\frac{1}{d}. \tag{62}$$

Therefore these subspaces are pairwise independent. This is the lowest level of independence. In this example, $\widetilde{R}(\alpha,\beta)=0$*.*

Example IV.10. *In* $H(6)$ *we consider the following two-dimensional subspaces:*

$$H_1=\left\{\begin{pmatrix}a\\b\\0\\0\\0\\0\end{pmatrix}\right\};\quad H_2=\left\{\begin{pmatrix}0\\0\\a\\b\\0\\0\end{pmatrix}\right\};\quad H_3=\left\{\begin{pmatrix}0\\a\\0\\0\\a\\b\end{pmatrix}\right\}. \tag{63}$$

Here we give a generic vector within these subspaces, which depends on two variables because the subspaces are two-dimensional. Then we calculate the subspaces $\mathfrak{H}_i$ *(Eq.(24)):*

$$\mathfrak{H}_1=H_2\vee H_3=\left\{\begin{pmatrix}0\\a\\b\\c\\a\\d\end{pmatrix}\right\};\quad \mathfrak{H}_2=H_1\vee H_3=\left\{\begin{pmatrix}a\\b\\0\\0\\c\\d\end{pmatrix}\right\};\quad \mathfrak{H}_3=H_1\vee H_2=\left\{\begin{pmatrix}a\\b\\c\\d\\0\\0\end{pmatrix}\right\}. \tag{64}$$

They are four-dimensional subspaces, and the vectors depend on four variables.

We also consider their orthocomplements which are the subspaces:

$$\mathfrak{H}_1^{\perp}=\left\{\begin{pmatrix}a\\b\\0\\0\\-b\\0\end{pmatrix}\right\};\quad \mathfrak{H}_2^{\perp}=\left\{\begin{pmatrix}0\\0\\a\\b\\0\\0\end{pmatrix}\right\};\quad \mathfrak{H}_3^{\perp}=\left\{\begin{pmatrix}0\\0\\0\\0\\a\\b\end{pmatrix}\right\}. \tag{65}$$

They are two-dimensional subspaces, and the vectors depend on two variables. We then calculate the spaces that *are used in proposition IV.5.*

$$\mathfrak{H}_1^{\perp} \wedge H_1 = \left\{ \begin{pmatrix} a \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}; \quad \mathfrak{H}_2^{\perp} \wedge H_2 = \left\{ \begin{pmatrix} 0 \\ 0 \\ a \\ b \\ 0 \\ 0 \end{pmatrix} \right\}; \quad \mathfrak{H}_3^{\perp} \wedge H_3 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ a \end{pmatrix} \right\}. \tag{66}$$

The corresponding projectors are calculated as follows. Let $a_1, ..., a_k$ be k independent vectors, and M the $d \times k$ matrix $(a_1, ..., a_k)$ which has as columns these vectors. The projector to the space spanned by these k vectors is

$$\Pi = M(M^{\dagger}M)^{-1}M^{\dagger}. \tag{67}$$

Using this we calculated the matrix for the degree of independence:

$$\mathcal{A} = \frac{1}{6} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{68}$$

We also calculated the distributions $R(i)$, $\widetilde{R}(i)$ and the degree of independence. We consider an orthonormal basis $|k\rangle$ where $k \in \mathbb{Z}(6)$, and the density matrix

$$\rho = |s\rangle\langle s|; \quad |s\rangle = \frac{1}{\sqrt{7}}(|0\rangle + |1\rangle + 2|3\rangle + |5\rangle). \tag{69}$$

We found that

$$\begin{aligned} &R(1) = 0.285; \quad R(2) = 0.571; \quad R(3) = 0.214 \\ &\widetilde{R}(1) = 0.142; \quad \widetilde{R}(2) = 0.571; \quad \widetilde{R}(3) = 0.142 \end{aligned} \tag{70}$$

Therefore

- *Measurement with $\Pi(H_1)$ will give 'yes' with probability $R(1) = 0.285$.*
- *Measurement with $\Pi(H_1 \wedge \mathfrak{H}_1^{\perp}) = \Pi(H_1 \wedge H_2^{\perp} \wedge H_3^{\perp})$ will give 'yes' with probability $\widetilde{R}(1) = 0.142$. In this case a simultaneous measurement with $\Pi(H_2)$ (which commutes with $\Pi(H_1 \wedge H_2^{\perp} \wedge H_3^{\perp})$) will give 'no'.*

The result 'yes' in the first measurement, means that the system collapses to a state that belongs to H_1. The result 'yes' in the second measurement, means that the system collapses to a state that belongs to H_1 and also to $H_2^{\perp}$ and also $H_3^{\perp}$ (therefore it does not belong to H_2 and it does not belong to H_3). Analogous comments can be made for the other $R(i)$ and $\widetilde{R}(i)$.

The degree of independence is $\eta(\rho) = 0.071$. Therefore we have an intermediate level of independence.

In this example, the independence relation in Eq.(51) is valid for:

$$|v_1\rangle \in \mathfrak{H}_1^{\perp} \wedge H_1 \prec H_1; \quad |v_2\rangle \in \mathfrak{H}_2^{\perp} \wedge H_2 = H_2; \quad |v_3\rangle \in \mathfrak{H}_3^{\perp} \wedge H_3 \prec H_3. \tag{71}$$

It is seen that there are vectors in $|v_1\rangle \in H_1$ and $|v_3\rangle \in H_3$ for which the implication in Eq.(51) is not valid.

D. The partial preorder of the various levels of independence

Definition IV.11. In $H(d)$ (with fixed d), we consider various sets of subspaces $S_1 = \{H_1, ..., H_n\}$, $S_2 = \{H'_1, ..., H'_m\}$, etc, with matrices for the degree of independence $\mathcal{A}_1$, $\mathcal{A}_2$, etc. The set of subspaces S_1 is more independent than S_2 (we denote this as $S_1 \sqsupset S_2$), if $\mathcal{A}_1 - \mathcal{A}_2$ is a negative semidefinite matrix (denoted as $\mathcal{A}_1 - \mathcal{A}_2 \leq 0$). In this case $\eta_1(\rho) \leq \eta_2(\rho)$ for all density matrices ρ.

Proposition IV.12. *$\sqsupset$ is a partial preorder.*

Proof. We consider the following properties:

- Reflexivity: $S_1 \sqsupset S_1$. This holds, because $\mathcal{A}_1 - \mathcal{A}_1 = 0$ is a negative semidefinite matrix.
- Transitivity: if $S_1 \sqsupset S_2$ and $S_2 \sqsupset S_3$ then $S_1 \sqsupset S_3$. This holds because if $\mathcal{A}_1 - \mathcal{A}_2$ and $\mathcal{A}_2 - \mathcal{A}_3$ are negative semidefinite matrices, then $\mathcal{A}_1 - \mathcal{A}_3$ is a negative semidefinite matrix.
- Antisymmetry: if $S_1 \sqsupset S_2$ and $S_2 \sqsupset S_1$ then $S_1 = S_2$. This does not hold. If $\mathcal{A}_1 - \mathcal{A}_2$ and $\mathcal{A}_2 - \mathcal{A}_1$ are negative semidefinite matrices, then $\mathcal{A}_1 = \mathcal{A}_2$, but this does not imply $S_1 = S_2$.

Since the first two properties hold, but not the last one, the $\sqsupset$ is a partial preorder. □

E. Weakly independent subspaces

We introduce the concept of weakly independent subspaces, which is dual through orthocomplementation, to a weakly total set of subspaces introduced later.

Definition IV.13. The subspaces $H_1, ..., H_n$ of $H(d)$ are weakly independent, if $H_1 \wedge ... \wedge H_n = \mathcal{O}$.

Proposition IV.14. *An independent set of subspaces, is also weakly independent set of subspaces. The converse is true when $n = 2$, but it is not true when $n \geq 3$.*

Proof. The relation

$$[H_1 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n] \prec [H_1 \vee ... \vee H_{i-1} \vee H_{i+1} \vee ... \vee H_n]. \tag{72}$$

proves that

$$H_1 \wedge ... \wedge H_n \prec [H_1 \vee ... \vee H_{i-1} \vee H_{i+1} \vee ... \vee H_n] \wedge H_i. \tag{73}$$

For independent subspaces, the right hand side is zero, and therefore the left hand side is zero. In Eq.(72), we have equality when $n = 2$, and inequality when $n \geq 3$. This means that the converse is true when $n = 2$, but it is not true when $n \geq 3$. □

V. LEVELS OF TOTALNESS

In this section we introduce the concept of totalness, which is dual (through orthocomplementation) to independence.

A. Total sets of subspaces

Proposition V.1. *The subspaces $H_1, ..., H_n$ of $H(d)$ are a total set, if one of the following statements, which are equivalent to each other, holds:*

(1) for all $i = 1, ..., n$:

$$\mathfrak{h}_i \vee H_i = H(d), \tag{74}$$

where $\mathfrak{h}_i$ has been defined in Eq.(25). Every vector $|v\rangle \in H(d)$, can be written (not uniquely) as a superposition of a vector in H_i, and a vector which is common in all other subspaces H_j with $j \neq i$:

$$|v\rangle = \lambda_i |a_i\rangle + \mu_i |b_i\rangle; \quad |a_i\rangle \in H_i; \quad |b_i\rangle \in \mathfrak{h}_i. \tag{75}$$

(2) For all $i = 1, ..., n$,

$$[H_1 \wedge ... \wedge H_{i-1}] \vee H_i = H(d). \tag{76}$$

(3) for all $i = 1, ..., n$, there is no vector in $H(d)$, which is perpendicular to both $\mathfrak{h}_i$ and H_i . In other words, if

$$\langle u|v\rangle = 0; \quad |v\rangle = \lambda_i |a_i\rangle + \mu_i |b_i\rangle; \quad , \tag{77}$$

for all $|a_i\rangle \in H_i$, and all $|b_i\rangle \in \mathfrak{h}_i$, then $|u\rangle$ is the zero vector.

Proof. (1) Orthocomplementation of Eqs(36), (37), proves the equivalence of the first two parts.

(2) We assume that the first part of the proposition holds, and prove that the third part also holds. Taking the orthocomplement of both sides in Eq.(74), we get $\mathfrak{h}_i^\perp \wedge H_i^\perp = \mathcal{O}$. This shows that there is no vector perpendicular to both $\mathfrak{h}_i$ and H_i .

Conversely, if the third part of the proposition holds, then $\mathfrak{h}_i^\perp \wedge H_i^\perp = \mathcal{O}$, and by taking the orthocomplement we get Eq.(74).

□

Remark V.2. Eq.(74) shows that in a total set of subspaces, there is a strong overlap between the subspaces:

$$\mathfrak{h}_i = H_1 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n \neq \mathcal{O}; \quad i = 1, ..., n. \tag{78}$$

Any $d - 1$ of the subspaces have vectors in common.

Proposition V.3. *The $H_1, ..., H_n$ are a total set of subspaces, if and only if the $H_1^\perp, ..., H_n^\perp$ are independent subspaces of $H(d)$.*

Proof. Independence is defined in Eq.(36), which we express in terms of the $H_1^\perp, ..., H_n^\perp$ as:

$$\mathfrak{h}_i^\perp \wedge H_i^\perp = \mathcal{O}. \tag{79}$$

From this follows that

$$\mathfrak{h}_i \vee H_i = H(d). \tag{80}$$

Therefore, according to the definition V.1, the $H_1, ..., H_n$ are a total set of subspaces. This argument also holds in the opposite direction, and proves that the converse is true. □

It is seen that orthocomplementation (the logical NOT operation) converts independence into totalness. The following proposition is dual to proposition IV.2.

Proposition V.4.

(1) If the $\{H_1, ..., H_n\}$ is a total set of subspaces, then any subset (with cardinality at least 2) is also a total set of subspaces.

(2) If $\{H_1, ..., H_n\}$ is a total set, then

$$\dim(H_1) + ... + \dim(H_n) = \dim(H_1 \wedge ... \wedge H_n) + (n-1)d \tag{81}$$

Proof.

(1) If the $S_1 = \{H_1, ..., H_n\}$ is a total set of subspaces, then the $S_2 = \{H_1^\perp, ..., H_n^\perp\}$ is a set of independent subspaces. According to proposition IV.2, any subset of S_2 is a set of independent subspaces, and consequently the corresponding subset of S_1, that contains the orthocomplements of these subspaces, is a total set of subspaces.

For an alternative direct proof we consider, as an example, the subset $\{H_2, ..., H_n\}$ of $\{H_1, ..., H_n\}$. If Eq.(74) holds, the fact that

$$\mathfrak{h}_i \prec H_2 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n, \tag{82}$$

implies that

$$H(d) = \mathfrak{h}_i \vee H_i \prec (H_2 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n) \vee H_i. \tag{83}$$

Therefore $(H_2 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n) \vee H_i = H(d)$. This proves that this particular subset is also a total set of subspaces. The proof for any other subset is analogous to this.

(2) This follows from Eq.(46) and the fact that $\dim(H_i) + \dim(H_i^\perp) = d$.

□

B. Pairwise total subspaces

Definition V.5. The subspaces $H_1, ..., H_n$ are pairwise total, if $H_i \vee H_j = H(d)$ for all i, j.

The following proposition is dual to proposition IV.4.

Proposition V.6.

(1) The non-distributivity of the lattice $\mathcal{L}(d)$, implies that totalness is stronger concept than pairwise totalness.

(2) For subspaces in a distributive sublattice of $\mathcal{L}(d)$, totalness is equivalent to pairwise totalness.

Proof. (1) In every lattice

$$(H_1 \vee H_i) \wedge ... \wedge (H_{i-1} \vee H_i) \wedge (H_{i+1} \vee H_i) \wedge ... \wedge (H_n \vee H_i) \succ \mathfrak{h}_i \vee H_i; \quad i = 1, ..., n. \tag{84}$$

Totalness implies that the right hand side is $H(d)$, and then the left hand side is $H(d)$. This leads to $H_i \vee H_j = H(d)$ for all i, j, i.e., pairwise totalness. Therefore totalness implies pairwise totalness.

The converse is not true. Pairwise totalness implies that the left hand side is $H(d)$, but this does not imply that the right hand side is $H(d)$. Therefore pairwise totalness does not imply totalness.

(2) Within a distributive sublattice of $\mathcal{L}(d)$, Eq.(84) is equality, and totalness is equivalent to pairwise totalness.

□

Proposition V.7. *The $H_1, ..., H_n$ are a pairwise total set of subspaces, if and only if the $H_1^\perp, ..., H_n^\perp$ are pairwise independent subspaces of $H(d)$.*

Proof. Pairwise independence for $H_1^\perp, ..., H_n^\perp$ is defined as

$$H_i^\perp \wedge H_j^\perp = \mathcal{O}. \tag{85}$$

Orthocomplementation of this gives

$$H_i \vee H_j = H(d). \tag{86}$$

Therefore the $H_1, ..., H_n$ are a pairwise total set of subspaces. This argument also holds in the opposite direction, and proves that the converse is true. □

C. Degree of totalness

Pairwise totalness is weaker concept than totalness, and there are intermediate concepts between the two. As we go from pairwise totalness to totalness, the overlap between the subspaces increases.

Proposition V.8. *Let $\{H_1, ..., H_n\}$ be subspaces which are pairwise total, and $\mathfrak{h}_i$ the subspaces in Eq.(25). Every vector $|v_i\rangle \in \mathfrak{h}_i \vee H_i$, can be written (not uniquely) as a sum*

$$|v_i\rangle = \lambda_i |a_i\rangle + \mu_i |b_i\rangle; \quad |a_i\rangle \in H_i; \quad |b_i\rangle \in \mathfrak{h}_i. \tag{87}$$

Proof. This follows from the definition of the disjunction $\mathfrak{h}_i \vee H_i$, which is that a vector in this space can be written as in Eq.(87). □

It is seen that the totalness relation in Eq.(75) is valid here only for vectors in the subspace $\mathfrak{h}_i \vee H_i$ of $H(d)$. This is the motivation for introducing in Eq.(27) the quasi-probability distribution $\widehat{R}(i)$. The degree of totalness compares the subspaces $\mathfrak{h}_i \vee H_i$ and H_i or equivalently $\widehat{R}(i)$ and $R(i)$.

Definition V.9. The matrix for the degree of totalness $\mathcal{T}$, and the degree of totalness $\epsilon(\rho)$, are given by

$$\mathcal{T} = \frac{1}{n} \sum_i [\Pi(\mathfrak{h}_i \vee H_i) - \Pi(H_i)]; \quad \epsilon(\rho) = \mathrm{Tr}(\rho \mathcal{T}) = \frac{1}{n} \sum [\widehat{R}(i) - R(i)] \tag{88}$$

Each $\Pi(\mathfrak{h}_i \vee H_i) - \Pi(H_i)$ is a projector. As a sum of projectors, $\mathcal{T}$ is a $d \times d$ positive semidefinite matrix. There are two extreme cases and many intermediate cases:

- If $\mathfrak{h}_i \vee H_i = H(d)$ for all i, the $\{H_1, ..., H_n\}$ are by definition a total set. In this case

$$\mathcal{T} = \mathbf{1} - \frac{1}{n} \sum_i \Pi(H_i); \quad \epsilon(\rho) = \mathrm{Tr}(\rho \mathcal{T}) = 1 - \frac{1}{n} \sum_i R(i). \tag{89}$$

  Eq.(75) holds for all vectors in $H(d)$, and proposition V.8 reduces to proposition V.1. This is the strongest form of totalness.

- If $\mathfrak{h}_i \vee H_i = H_i$ for all i, Eq.(75) does not hold. In this case the $\{H_1, ..., H_n\}$ are a pairwise total set, $\mathcal{T} = 0$ and $\epsilon(\rho) = 0$. This is the weakest form of totalness.
- Between these two extreme cases, the $\{H_1, ..., H_n\}$ are a partially total set, in the sense that Eq.(75) does not hold for all vectors in $H(d)$. In this case the degree of totalness takes values in the interval

$$0 \leq \epsilon(\rho) \leq 1 - \frac{1}{n}\sum_i R(i). \tag{90}$$

A summary of the various levels of totalness is shown in table I.

Example V.10. *We consider the set $\{H(X;\alpha)\}$ with d subspaces of $H(d)$, where $\alpha \in \mathbb{Z}(d)$, and we get*

$$\mathfrak{h}(X;\alpha) = \bigwedge_{\beta\neq\alpha} H(X;\beta) = \mathcal{O}. \tag{91}$$

Therefore $\mathfrak{h}(X;\alpha) \vee H(X;\alpha) = H(X;\alpha)$ and

$$\widehat{R}(\alpha) = \langle X;\alpha|\rho|X;\alpha\rangle. \tag{92}$$

Taking into account the results in examples III.3, IV.7, we see that in this case $\widetilde{R}(\alpha) = R(\alpha) = \widehat{R}(\alpha)$. Therefore $\mathcal{T} = 0$, $\epsilon(\rho) = 0$, and the $\{H(X;\alpha)\}$ is a pairwise total set of subspaces.

A different problem is to study the totalness of the orthocomplements $\{[H(X;\alpha)]^\perp\}$. We have seen in example IV.7, that the $\{H(X;\alpha)\}$ are a set of independent subspaces, and this implies that their orthocomplements $\{[H(X;\alpha)]^\perp\}$ are a total set of subspaces. In order to verify this directly, we show that

$$\mathfrak{h}_{\rm ortho}(X;\alpha) = \bigwedge_{\beta\neq\alpha}[H(X;\beta)]^\perp = \left[\bigvee_{\beta\neq\alpha}[H(X;\beta)]\right]^\perp = H(X;\alpha). \tag{93}$$

Consequently

$$[H(X;\alpha)]^\perp \vee \mathfrak{h}_{\rm ortho}(X;\alpha) = H(X;\alpha)^\perp \vee H(X;\alpha) = H(d). \tag{94}$$

Therefore

$$\mathcal{T}_{\rm ortho} = \mathbf{1} - \frac{1}{d}\sum \Pi\{H(X;\alpha)^\perp\} = \frac{1}{d}\sum \Pi[H(X;\alpha)] = \frac{1}{d}\mathbf{1};\quad \epsilon_{\rm ortho}(\rho) = \frac{1}{d}. \tag{95}$$

This confirms that the $\{[H(X;\alpha)]^\perp\}$ is a total set of subspaces.

Example V.11. *We consider the set $\{H(X;\alpha), H(P;\beta)\}$ with $2d$ subspaces of $H(d)$, labelled as in Eq.(33). In this case*

$$i = 0, ..., (d-1)\ \rightarrow\ \mathfrak{h}_i = \left(\bigwedge_{\beta\neq i} H(X;\beta)\right) \wedge \left(\bigwedge_{\gamma} H(P;\gamma)\right) = \mathcal{O}$$
$$i = d, ..., (2d-1)\ \rightarrow\ \mathfrak{h}_i = \left(\bigwedge_{\beta} H(X;\beta)\right) \wedge \left(\bigwedge_{\gamma\neq i} H(P;\gamma)\right) = \mathcal{O}, \tag{96}$$

and $\mathfrak{h}_i \vee H_i = H_i$. Therefore

$$\mathcal{T} = 0;\quad \epsilon(\rho) = 0. \tag{97}$$

It is seen that the $\{H(X;\alpha), H(P;\beta)\}$ is a pairwise total set of subspaces. Taking into account the results in examples III.4, IV.8, we see that in this case $\widetilde{R}(i) = 0$ and $R(i) = \widehat{R}(i)$.

Example V.12. *We consider the set* $\{H(C;\alpha,\beta)\}$ *that contains* d^2 *subspaces of* $H(d)$*, and we get*

$$\mathfrak{h}(C;\alpha_0,\beta_0) = \bigwedge_{\alpha\neq\alpha_0,\beta\neq\beta_0} H(C;\alpha,\beta) = \mathcal{O}, \tag{98}$$

and $\mathfrak{h}(C;\alpha_0,\beta_0) \vee H(C;\alpha,\beta) = H(C;\alpha,\beta)$*. Therefore*

$$\mathcal{T} = 0; \quad \epsilon(\rho) = 0. \tag{99}$$

It is seen that the $\{H(C;\alpha,\beta)\}$ *is a pairwise total set of subspaces. Taking into account the results in examples III.5, IV.9, we see that in this case* $\widetilde{R}(\alpha,\beta) = 0$ *and* $R(\alpha,\beta) = \widehat{R}(\alpha,\beta)$*.*

Example V.13. *In* $H(6)$ *we consider the following two-dimensional subspaces:*

$$H_1 = \left\{\begin{pmatrix} a\\ b\\ 0\\ 0\\ 0\\ 0\end{pmatrix}\right\}; \quad H_2 = \left\{\begin{pmatrix} 0\\ 0\\ a\\ 0\\ 0\\ b\end{pmatrix}\right\}; \quad H_3 = \left\{\begin{pmatrix} 0\\ a\\ 0\\ 0\\ 0\\ b\end{pmatrix}\right\}. \tag{100}$$

In this case

$$\mathfrak{h}_1 = H_2 \wedge H_3 = \left\{\begin{pmatrix} 0\\ 0\\ 0\\ 0\\ 0\\ a\end{pmatrix}\right\}; \quad \mathfrak{h}_2 = H_1 \wedge H_3 = \left\{\begin{pmatrix} 0\\ a\\ 0\\ 0\\ 0\\ 0\end{pmatrix}\right\}; \quad \mathfrak{h}_3 = H_1 \wedge H_2 = \mathcal{O}. \tag{101}$$

They are one-dimensional subspaces. Then

$$\mathfrak{h}_1 \vee H_1 = \left\{\begin{pmatrix} a\\ b\\ 0\\ 0\\ 0\\ c\end{pmatrix}\right\}; \quad \mathfrak{h}_2 \vee H_2 = \left\{\begin{pmatrix} 0\\ a\\ b\\ 0\\ 0\\ c\end{pmatrix}\right\}; \quad \mathfrak{h}_3 \vee H_3 = H_3. \tag{102}$$

We used Eq.(67) to calculate the projectors and we found that

$$\mathcal{T} = \frac{1}{3}\begin{pmatrix} 0&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&1\end{pmatrix}. \tag{103}$$

TABLE II: Some sets of subspaces $\{H_1, ..., H_n\}$ of $H(d)$, and the corresponding $R(i)$, $\widetilde{R}(i)$, $\widehat{R}(i)$.

$\{H(X;\alpha)\}$	$\{H(X;\alpha), H(P;\beta)\}$	$\{H(C;\alpha,\beta)\}$
$R(\alpha) = \langle X;\alpha\|\rho\|X;\alpha\rangle$ $\sum R(\alpha) = 1$	$R(i) = \langle X;i\|\rho\|X;i\rangle$ $R(i) = \langle P;i-d\|\rho\|P;i-d\rangle$ $\frac{1}{2}\sum R(i) = 1$	$R(\alpha,\beta) = \langle C;\alpha,\beta\|\rho\|C;\alpha,\beta\rangle$ $\frac{1}{d}\sum R(\alpha,\beta) = 1$
$\widetilde{R}(\alpha) = R(\alpha)$	$\widetilde{R}(i) = 0$	$\widetilde{R}(\alpha,\beta) = 0$
$\widehat{R}(\alpha) = R(\alpha)$	$\widehat{R}(i) = R(i)$	$\widehat{R}(\alpha,\beta) = R(\alpha,\beta)$

We also calculated the $R(i)$, $\widehat{R}(i)$ and the degree of totalness for the density matrix

$$\rho = |s\rangle\langle s|; \quad |s\rangle = \frac{1}{\sqrt{15}}(|0\rangle + |1\rangle + 2|4\rangle + 3|5\rangle). \tag{104}$$

We found that

$$\begin{aligned} &R(1) = 0.133; \quad R(2) = 0.600; \quad R(3) = 0.666 \\ &\widehat{R}(1) = 0.733; \quad \widehat{R}(2) = 0.666; \quad \widehat{R}(3) = 0.666 \end{aligned} \tag{105}$$

Therefore

- *Measurement with $\Pi(H_1)$ will give 'yes' with probability $R(1) = 0.133$.*
- *Measurement with $\Pi(H_1 \vee \mathfrak{h}_1^\perp) = \Pi[H_1 \vee (H_2 \wedge H_3)]$ will give 'yes' with probability $\widehat{R}(1) = 0.733$.*

The result 'yes' in the first measurement means that the system collapses to a state that belongs to H_1. The result 'yes' in the second measurement means that the system collapses to a superposition of a state in H_1 and another state which belongs to both H_2 and H_3. Analogous comments can be made for the other $R(i)$ and $\widehat{R}(i)$.

The degree of totalness is $\epsilon(\rho) = 0.222$. In this example we have an intermediate level of totalness.

The results for some of the above examples are summarized in table II.

Remark V.14. Propositions IV.4, V.6 show that the following are related:

- The lattice $\mathcal{L}(d)$ is not distributive.
- Independence is stronger concept than pairwise independence.
- Totalness is stronger concept than pairwise totalness.

D. Weakly total sets of subspaces

We introduce the concept of weak totalness which is dual to the weak independence in section IV E. Weak totalness is the same as totalness when $n = 2$, and weaker than totalness when $n \geq 3$.

Proposition V.15. *The $H_1, ..., H_n$ are a weakly total set of subspaces of $H(d)$, if one of the following statements, which are equivalent to each other, holds:*

(1)

$$H_1 \vee ... \vee H_n = H(d). \tag{106}$$

Then every vector $|v\rangle \in H(d)$ is a superposition of vectors in H_i. Therefore it can be written (not uniquely) as a sum

$$|v\rangle = \lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle; \quad |v_i\rangle \in H_i \tag{107}$$

(2) There is no vector in $H(d)$, which is perpendicular to all subspaces $H_1, ..., H_n$. In other words, if

$$\langle a|v\rangle = 0; \quad |v\rangle = \lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle \tag{108}$$

for all $|v_i\rangle \in H_i$ and all $\lambda_i \in \mathbb{C}$, then $|a\rangle$ is the zero vector.

Proof. We assume that the first part of the proposition holds, and prove that the second part also holds. Taking the orthocomplement of both sides in Eq.(106), we get $H_1^\perp \wedge ... \wedge H_n^\perp = \mathcal{O}$. This shows that there is no vector perpendicular to all subspaces $H_1, ..., H_n$, and Eq.(108).

Conversely, if the second part of the proposition holds, then $H_1^\perp \wedge ... \wedge H_n^\perp = \mathcal{O}$. The orthocomplement of this proves Eq.(106), and proves the first part of the proposition.

□

Proposition V.16. *If $H_1, ..., H_n$ is a weakly total set of independent subspaces of $H(d)$, then the expansion in Eq.(107) is unique.*

Proof. The $H_1, ..., H_n$ are a total set, and therefore there are expansions of an arbitrary vector $|v\rangle$ in $H(d)$, as

$$|v\rangle = \lambda_1|v_1\rangle + ... + \lambda_n|v_n\rangle = \mu_1|v_1\rangle + ... + \mu_n|v_n\rangle; \quad |v_i\rangle \in H_i. \tag{109}$$

This implies that

$$(\lambda_1 - \mu_1)|v_1\rangle + ... + (\lambda_n - \mu_n)|v_n\rangle = 0. \tag{110}$$

Since the $H_1, ..., H_n$ are independent subspaces, it follows that $\lambda_i = \mu_i$, and this proves the uniqueness of the expansion. □

Proposition V.17. *A total set of subspaces, is also a weakly total set of subspaces. The converse is true when $n = 2$, but it is not true when $n \geq 3$.*

Proof. By definition of a total set of subspaces $\mathfrak{h}_i \vee H_i = H(d)$. Also the obvious relation $\mathfrak{h}_i \prec \mathfrak{H}_i$ gives

$$H(d) = \mathfrak{h}_i \vee H_i \prec \mathfrak{H}_i \vee H_i = (H_1 \vee ... \vee H_n). \tag{111}$$

Therefore $H_1 \vee ... \vee H_n = H(d)$. For $n = 2$, we get $\mathfrak{h}_i = \mathfrak{H}_i$. This means that the converse is true when $n = 2$, but it is not true when $n \geq 3$. □

It is seen that when $n \geq 3$, totalness is a stronger concept than weak totalness. For $n = 2$, they are the same.

Proposition V.18. *The $H_1, ..., H_n$ are a total set of weakly independent subspaces of $H(d)$, if and only if the $H_1^\perp, ..., H_n^\perp$ are a weakly total set of independent subspaces of $H(d)$. In this case the expansion in Eq.(75) is unique.*

Proof. (1) We have already shown that the $H_1, ..., H_n$ are a total set of subspaces, if and only if the $H_1^\perp, ..., H_n^\perp$ are independent. If in addition to that the $H_1, ..., H_n$ are weakly independent, $H_1 \wedge ... \wedge H_n = \mathcal{O}$ and $H_1^\perp \vee ... \vee H_n^\perp = H(d)$, which implies that $H_1^\perp, ..., H_n^\perp$ are a weakly total set of subspaces of $H(d)$. The converse is also true.

(2) Let

$$|v\rangle = \lambda_i |a_i\rangle + \mu_i |b_i\rangle = \lambda_i' |a_i\rangle + \mu_i' |b_i\rangle; \quad |a_i\rangle \in H_i \tag{112}$$

be two expansions of a vector $|v\rangle$, where $|b_i\rangle \in (H_1 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n)$. Then

$$(\lambda_i - \lambda_i')|a_i\rangle = (\mu_i' - \mu_i)|b_i\rangle; \quad |a_i\rangle \in H_i; \quad |b_i\rangle \in (H_1 \wedge ... \wedge H_{i-1} \wedge H_{i+1} \wedge ... \wedge H_n). \tag{113}$$

It follows that $\lambda_i - \lambda_i' = \mu_i' - \mu_i = 0$, and therefore the expansion is unique.

□

It is seen that with orthocomplementation, independence and weak independence, become totalness, and weak totalness, correspondingly.

E. Orthogonalization

In addition to the expansion in Eq.(107) which involves non-orthogonal components, we can have an orthogonal expansion as discussed below.

Proposition V.19. *Let $H_1, ..., H_n$ be a weakly total set of independent subspaces of $H(d)$. We introduce the following spaces and the corresponding projectors:*

$$\begin{aligned}
&\mathcal{H}_1 = H_1; \quad \mathfrak{P}_1 = \Pi(H_1) \\
&\mathcal{H}_2 = (H_1 \vee H_2) \wedge H_1^\perp; \quad \mathfrak{P}_2 = \Pi(H_1 \vee H_2) - \Pi(H_1) \\
&...... \\
&\mathcal{H}_i = [(H_1 \vee ... \vee H_{i-1}) \vee H_i] \wedge (H_1 \vee ... \vee H_{i-1})^\perp; \quad \mathfrak{P}_i = \Pi(H_1 \vee ... \vee H_i) - \Pi(H_1 \vee ... \vee H_{i-1}) \\
&...... \\
&\mathcal{H}_n = (H_1 \vee ... \vee H_{n-1})^\perp; \quad \mathfrak{P}_n = \mathbf{1} - \Pi(H_1 \vee ... \vee H_{n-1})
\end{aligned} \tag{114}$$

Then:

$$\mathcal{H}_1 \vee ... \vee \mathcal{H}_n = H(d); \quad \mathfrak{P}_1 + ... + \mathfrak{P}_n = \mathbf{1}; \quad \mathfrak{P}_i \mathfrak{P}_j = \mathfrak{P}_i \delta(i,j). \tag{115}$$

Proof. We prove that $\mathcal{H}_1 \vee ... \vee \mathcal{H}_n = H(d)$ using the modularity property in Eq.(18). For example

$$\mathcal{H}_1 \vee \mathcal{H}_2 = H_1 \vee [H_1^\perp \wedge (H_1 \vee H_2)] = (H_1 \vee H_1^\perp) \wedge (H_1 \vee H_2) = H(d) \wedge (H_1 \vee H_2) = H_1 \vee H_2. \tag{116}$$

We continue in the same way and we prove that

$$\mathcal{H}_1 \vee ... \vee \mathcal{H}_n = H_1 \vee ... \vee H_n = H(d). \tag{117}$$

The fact that $\mathcal{H}_1 \vee \mathcal{H}_2 = H_1 \vee H_2$ and $\mathcal{H}_1 \wedge \mathcal{H}_2 = \mathcal{O}$ shows that $\mathfrak{P}_2 = \Pi(H_1 \vee H_2) - \Pi(H_1)$. In analogous way we prove the expressions given above, for the rest of the projectors.

Direct multiplication proves that they are orthogonal projectors. For example

$$\mathfrak{P}_1 \mathfrak{P}_2 = \Pi(H_1)[\Pi(H_1 \vee H_2) - \Pi(H_1)] = \Pi(H_1) - \Pi(H_1) = 0, \tag{118}$$

and

$$\begin{aligned}\mathfrak{P}_2\mathfrak{P}_2 &= [\Pi(H_1 \vee H_2) - \Pi(H_1)][\Pi(H_1 \vee H_2) - \Pi(H_1)] \\ &= \Pi(H_1 \vee H_2) - \Pi(H_1) - \Pi(H_1) + \Pi(H_1) = \mathfrak{P}_2.\end{aligned} \tag{119}$$

This completes the proof. □

Using Eq.(115) we can express an arbitrary state in terms of orthogonal components:

$$|v\rangle = \sum \mathfrak{P}_i|v\rangle = \sum \lambda_i|v_i\rangle; \quad |\lambda_i|^2 = \langle v|\mathfrak{P}_i|v\rangle; \quad |v_i\rangle \in \mathcal{H}_i. \tag{120}$$

If we change the order of the subspaces, we get different projectors and a different expansion.

VI. INFORMATIONALLY INDEPENDENT SUBSPACES AND MEASUREMENTS

It has been pointed out in a pure mathematics context [9–15], that the lattices describing finite quantum systems (and also the normal subgroups of a group), are a special case of modular orthocomplemented lattices, with extra stronger properties. They are lattices of commuting equivalence relations (also called linear lattices by Rota and collaborators[11–15]). The lattices of commuting equivalence relations are modular, but the converse is not true in general.

Equivalence relations are intimately related to partitions of the Hilbert space $H(d)$, and it is the language of partitions that we use below. Two partitions are independent, if knowledge of the block of the first partition to which an element belongs, provides no information about the block of the second partition to which this element belongs.

Based on the concept of independent partitions, we introduce in this section informationally independent subspaces, in a physical context. Physically, each subspace H_1 leads naturally to a partition $\varpi(H_1)$ of the Hilbert space $H(d)$, into blocks which are sets but not subspaces. Measurement with the projector $\Pi(H_1)$ gives the same result for all states in each block of the partition $\varpi(H_1)$ (when the outcome is 'yes').

We show that informational independence is equivalent to independence. Weaker concepts of independence, are not informationally independent.

A. Partitions of the Hilbert space $H(d)$ and their role in quantum measurements

Definition VI.1. If H_1 is a subspace of $H(d)$, $\varpi(H_1)$ is the partition of $H(d)$ into 'blocks' $|v\rangle + H_1$ that contain vectors $|v\rangle \in H_1^\perp$ modulo vectors $|a\rangle \in H_1$:

$$|v\rangle + H_1 = \{|v\rangle + |a\rangle \mid |a\rangle \in H_1\}; \quad |v\rangle \in H_1^\perp. \tag{121}$$

The blocks are sets, but they are **not** subspaces. The partition $\varpi[H(d)] = \varpi(\mathcal{I})$ has only one block, the $H(d)$. In the partition $\varpi(\mathcal{O})$, each block contains one vector $|v\rangle$ only.

Although it is not essential, it is convinient to consider below normalized vectors:

$$\langle v|v\rangle + \langle a|a\rangle = 1; \quad \langle v|a\rangle = 0. \tag{122}$$

Remark VI.2.

- Each normalized vector $|s\rangle$ in $H(d)$, can be written uniquely as

$$|s\rangle = \Pi(H_1^\perp)|s\rangle + \Pi(H_1)|s\rangle. \tag{123}$$

Therefore $|s\rangle$ belongs to exactly one block within the partition $\varpi(H_1)$ (the block $\Pi(H_1^\perp)|s\rangle + H_1$). $|s\rangle$ also belongs to exactly one block within the partition $\varpi(H_1^\perp)$ (the block $\Pi(H_1)|s\rangle + H_1^\perp$).

- With a measurement $\Pi(H_1^\perp)$:
  - The block $|v\rangle + H_1$ in the partition $\varpi(H_1)$, contains states $|s\rangle$ which when the outcome is 'yes', collapse into the same state

$$\frac{1}{\sqrt{p}}|v\rangle \in H_1^\perp; \quad p = \langle v|v\rangle = \langle s|\Pi(H_1^\perp)|s\rangle \tag{124}$$

    with the same probability p. The measurement $\Pi(H_1^\perp)$, cannot distinguish the states in the block $|v\rangle + H_1$, when the outcome is 'yes'.
  - The block $|a\rangle + H_1^\perp$ in the partition $\varpi(H_1^\perp)$, contains states $|s\rangle$ which when the outcome is 'no', collapse into the same state

$$\frac{1}{\sqrt{1-p}}|a\rangle \in H_1; \quad 1-p = \langle a|a\rangle = \langle s|\mathbf{1} - \Pi(H_1^\perp)|s\rangle \tag{125}$$

    with the same probability $1-p$. The measurement $\Pi(H_1^\perp)$, cannot distinguish the states in the block $|a\rangle + H_1^\perp$, when the outcome is 'no'.

- Vectors in the two blocks $|v\rangle + H_1$ and $\lambda|v\rangle + H_1$ (with $|\lambda| \le 1$) with a measurement $\Pi(H_1^\perp)$ that gives the outcome 'yes', will collapse into the same state given in Eq.(124), with different probabilities p and $|\lambda|^2 p$, correspondingly.

Proposition VI.3. *The partition $\varpi(H_1)$, with the following operation defining superpositions between its blocks*

$$\lambda_1(|s_1\rangle + H_1) + \lambda_2(|s_2\rangle + H_1) = (\lambda_1|s_1\rangle + \lambda_2|s_2\rangle) + H_1; \quad |s_1\rangle, |s_2\rangle \in H_1^\perp, \tag{126}$$

is a Hilbert space isomorphic to $H_1^\perp$. The zero vector in $\varpi(H_1)$, is the block H_1.

Proof. The notation that we introduced above shows that there is a bijective map between the partition $\varpi(H_1)$ and the Hilbert space $H_1^\perp$. It is also easily seen that the sum of blocks in $\varpi(H_1)$ defined in Eq.(126) corresponds to the sum of vectors in $H_1^\perp$. □

Corollary VI.4. *In the set of partitions*

$$\{\varpi(H_1) \mid H_1 \prec H(d)\} \tag{127}$$

we define the following operations:

(1) Disjunction

$$\varpi(H_1) \vee \varpi(H_2) = \varpi(H_1 \wedge H_2) \tag{128}$$

The blocks in $\varpi(H_1) \vee \varpi(H_2)$ are

$$|v\rangle + H_1 \wedge H_2; \quad |v\rangle \in (H_1 \wedge H_2)^\perp = H_1^\perp \vee H_2^\perp. \tag{129}$$

Special cases are:

$$\varpi(H_1) \vee \varpi(\mathcal{O}) = \varpi(\mathcal{O}); \quad \varpi(H_1) \vee \varpi(\mathcal{I}) = \varpi(H_1); \quad \varpi(H_1) \vee \varpi(H_1^\perp) = \varpi(\mathcal{O}). \tag{130}$$

(2) Conjunction

$$\varpi(H_1) \wedge \varpi(H_2) = \varpi(H_1 \vee H_2) \tag{131}$$

The blocks in $\varpi(H_1) \wedge \varpi(H_2)$ *are*

$$|v\rangle + H_1 \vee H_2; \quad |v\rangle \in (H_1 \vee H_2)^\perp = H_1^\perp \wedge H_2^\perp. \tag{132}$$

Special cases are:

$$\varpi(H_1) \wedge \varpi(\mathcal{O}) = \varpi(H_1); \quad \varpi(H_1) \wedge \varpi(\mathcal{I}) = \varpi(\mathcal{I}); \quad \varpi(H_1) \wedge \varpi(H_1^\perp) = \varpi(\mathcal{I}). \tag{133}$$

(3) Orthocomplement

$$[\varpi(H_1)]^\perp = \varpi(H_1^\perp) \tag{134}$$

The blocks in $[\varpi(H_1)]^\perp$ *are* $|v\rangle + H_1^\perp$ *where* $|v\rangle \in H_1$. *Special cases are:*

$$[\varpi(\mathcal{O})]^\perp = \varpi(\mathcal{I}); \quad [\varpi(\mathcal{I})]^\perp = \varpi(\mathcal{O}). \tag{135}$$

(4) The partial order is 'refinement'. $\varpi(H_1) \prec \varpi(H_2)$ *if* $H_1 \succ H_2$, *in which case every block in* $\varpi(H_2)$ *is contained in some block in* $\varpi(H_1)$. *For every* $\varpi(H_1)$, *we get* $\varpi(\mathcal{I}) \prec \varpi(H_1) \prec \varpi(\mathcal{O})$.

Then the set of partitions is a lattice dually isomorphic to $\mathcal{L}(d)$ *(i.e., the conjunction and disjunction exchange roles).*

Proof. This follows from the fact that $\varpi(H_1^\perp)$ is isomorphic to H_1, and Eqs.(20). □

Definition VI.5. The $\varpi(H_1), ..., \varpi(H_n)$ are independent, a total set or a strongly total set, if the $H_1^\perp, ..., H_n^\perp$ are independent, a total set or a strongly total set, correspondingly.

B. Two informationally independent subspaces and measurements

Two partitions $\varpi(H_1)$, $\varpi(H_2)$ are informationally independent, if knowledge that a vector belongs to a particular block of the partition $\varpi(H_1)$, gives no information about the block of the partition $\varpi(H_2)$ which contains this vector. This is the motivation for the definition below.

Definition VI.6. Two distinct partitions $\varpi(H_1)$, $\varpi(H_2)$ are informationally independent, if the intersection of any block of $\varpi(H_1)$, with any block of $\varpi(H_2)$, is non-empty:

$$(|v_1\rangle + H_1) \cap (|v_2\rangle + H_2) \neq \emptyset; \quad |v_1\rangle \in H_1^\perp; \quad |v_2\rangle \in H_2^\perp. \tag{136}$$

If for some blocks $(|v_1\rangle + H_1) \cap (|v_2\rangle + H_2) = \emptyset$, then knowledge that a vector belongs to the block $|v_1\rangle + H_1$ of the partition $\varpi(H_1)$, implies that this vector does not belong to the block $|v_2\rangle + H_2$ of the partition $\varpi(H_2)$. In this case the partitions $\varpi(H_1)$, $\varpi(H_2)$, are not informationally independent.

Partitions are isomorphic to subspaces, and the concept of informationally independent partitions leads to the following definition of informationally independent subspaces.

Definition VI.7. Two distinct subspaces H_1, H_2 are informationally independent, if for any pair of vectors $|v_1\rangle \in H_1$ and $|v_2\rangle \in H_2$, the intersection of the block $|v_1\rangle + H_1^\perp$ with the block $|v_2\rangle + H_2^\perp$, is non-empty:

$$(|v_1\rangle + H_1^\perp) \cap (|v_2\rangle + H_2^\perp) \neq \emptyset; \quad |v_1\rangle \in H_1; \quad |v_2\rangle \in H_2. \tag{137}$$

The motivation for this definition in terms of quantum measurements, is as follows. We asssume that for some blocks $(|v_1\rangle + H_1^\perp) \cap (|v_2\rangle + H_2^\perp) = \emptyset$, in which case if a state $|s\rangle$ belongs to the block $|v_1\rangle + H_1^\perp$, it cannot belong to the block $|v_2\rangle + H_2^\perp$. Then the measurement $\Pi(H_1)$ on $|s\rangle$ will collapse it into $\mathcal{N}_1|v_1\rangle$, if the outcome is 'yes'. But the measurement $\Pi(H_2)$ on $|s\rangle$, cannot collapse it into $\mathcal{N}_2|v_2\rangle$. Therefore knowledge of the outcome of the $\Pi(H_1)$ measurement, gives information about the outcome of the $\Pi(H_2)$ measurement. In this case, the H_1, H_2 are not informationally independent.

The following lemma will be used below to show that informational independence is the same concept as independence.

Lemma VI.8. *The following statements are equivalent:*

- *The subspaces H_1, H_2 are informationally independent*
- *The H_1, H_2 are independent.*
- *The $H_1^\perp, H_2^\perp$ are a total set of subspaces.*

Proof. We will prove the equivalence of the first two statements. The equivalence of the last two statements has been proved earlier.

We assume that H_1, H_2 are independent, i.e., that $H_1 \wedge H_2 = \mathcal{O}$. For any vectors $|v_1\rangle \in H_1$ and $|v_2\rangle \in H_2$, we have to show that there exists vectors $|a_1\rangle \in H_1^\perp$ and $|a_2\rangle \in H_2^\perp$, such that

$$|v_1\rangle + |a_1\rangle = |v_2\rangle + |a_2\rangle. \tag{138}$$

Because in this case Eq.(137) holds. In other words we have to show that there exists solution to the equation:

$$|a_2\rangle - |a_1\rangle = |v_1\rangle - |v_2\rangle; \quad |v_1\rangle - |v_2\rangle \in (H_1 \vee H_2); \quad |a_2\rangle - |a_1\rangle \in (H_1^\perp \vee H_2^\perp). \tag{139}$$

Here the unknowns are the vectors $|a_2\rangle, |a_1\rangle$. We want to have a solution for all vectors $|v_1\rangle \in H_1$ and $|v_2\rangle \in H_2$, and therefore the $|v_1\rangle - |v_2\rangle$ can be any vector in $H_1 \vee H_2$. In order to have a solution the $H_1 \vee H_2$ should be a subspace of $H_1^\perp \vee H_2^\perp$. If this is not the case then for $|v_1\rangle - |v_2\rangle$ in the set $(H_1 \vee H_2) \setminus (H_1^\perp \vee H_2^\perp)$, there exist no solution.

The fact that $H_1 \wedge H_2 = \mathcal{O}$, implies that $H_1^\perp \vee H_2^\perp = H(d)$. Therefore in this case, $H_1 \vee H_2 \prec H_1^\perp \vee H_2^\perp = H(d)$, and we can then find vectors $|a_1\rangle$, $|a_2\rangle$ which satisfy this equation. This proves that H_1, H_2 are informationally independent.

Conversely, we assume that H_1, H_2 are informationally independent, i.e., that Eq.(139) has a solution for all $|v_1\rangle - |v_2\rangle \in (H_1 \vee H_2)$. Then

$$H_1 \vee H_2 \prec H_1^\perp \vee H_2^\perp = (H_1 \wedge H_2)^\perp. \tag{140}$$

From this follows that

$$H_1 \wedge H_2 \prec H_1 \vee H_2 \prec (H_1 \wedge H_2)^\perp. \tag{141}$$

This is possible only if $H_1 \wedge H_2 = \mathcal{O}$. Therefore the H_1, H_2 are independent. □

C. Several informationally independent subspaces and measurements

Definition VI.9. The subspaces $H_1, ..., H_n$ of $H(d)$ are informationally independent, if for all $i = 1, ..., n$, the pairs of subspaces $\mathfrak{H}_i$ and H_i are informationally independent (according to the definition VI.6).

Proposition VI.10. *The following statements are equivalent:*

- *The subspaces $H_1, ..., H_n$ are informationally independent.*
- *The $H_1, ..., H_n$ are independent.*
- *The $H_1^\perp, ..., H_n^\perp$ are a strongly total set of subspaces.*

Proof. We will prove the equivalence of the first two statements. The equivalence of the last two statements has been proved earlier.

We assume that the $H_1, ..., H_n$ are independent, in which case according to Eq.(36) the $\mathfrak{H}_i$ and H_i are independent subspaces, for all $i = 1, ..., n$. Then lemma VI.10 shows that they are also informationally independent subspaces. Therefore according to the definition VI.9, the $H_1, ..., H_n$ are informationally independent subspaces.

This argument is also valid in the opposite direction, and it proves the converse. □

In view of this result we will use the simpler term independence for both independence and informational independence..

VII. APPLICATION: THE PENTAGRAM IN $H(3)$

An example of a formalism that requires a deeper understanding of the underlying concepts, is the pentagram which has been studied in the context of contextuality [16–24]. In this section we apply our formalism to the pentagram.

A. Background

A context is a set $\mathfrak{C} = \{H_1, ..., H_n\}$, of subspaces which commute pairwise ($H_i \mathcal{C} H_j$ for all i, j) or equivalently the corresponding projectors commute pairwise ($[\Pi(H_i), \Pi(H_j)] = 0$ for all i, j). The sublattice of $\mathcal{L}(d)$ generated by a context $\mathfrak{C}$ is distributive. We consider two contexts

$$\mathfrak{C}_1 = \{H_1, H_2, ..., H_n\}; \quad \mathfrak{C}_2 = \{H_1, h_2, ..., h_m\}. \tag{142}$$

The subspaces in the first context commute pairwise, the subspaces in the second context commute pairwise, but in general the H_i does not commute with h_j. We will call them overlapping contexts because H_1 belongs to both of these contexts.

We consider the pentagram in $H(3)$, which has been studied originally in [20]. To be specific, we consider the following states:

$$|s_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \quad |s_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}; \quad |s_2\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}; \quad |s_3\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}; \quad |s_4\rangle = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}. \tag{143}$$

The indices of these states belong to $\mathbb{Z}(5)$ (integers modulo 5). Any three of these vectors are independent. It is easily seen that

$$\langle s_i | s_{i+1} \rangle = 0; \quad i \in \mathbb{Z}(5). \tag{144}$$

We call H_i the one-dimensional subspace of $H(3)$, which contains the states $a|s_i\rangle$:

$$H_i = \{a|s_i\rangle\}; \quad i = 0, ..., 4. \tag{145}$$

Their orthocomplements are the two-dimensional spaces

$$H_0^\perp = \left\{ \begin{pmatrix} 0 \\ a \\ b \end{pmatrix} \right\};\ \ H_1^\perp = \left\{ \begin{pmatrix} a \\ b \\ -b \end{pmatrix} \right\};\ \ H_2^\perp = \left\{ \begin{pmatrix} a \\ b \\ a+b \end{pmatrix} \right\};\ \ H_3^\perp = \left\{ \begin{pmatrix} 2a \\ 2b \\ -(a+b) \end{pmatrix} \right\};\ \ H_4^\perp = \left\{ \begin{pmatrix} a \\ b \\ 2b \end{pmatrix} \right\} \tag{146}$$

We also consider the projectors $\Pi(H_i)$ to the subspaces H_i:

$$\Pi(H_i) = |s_i\rangle\langle s_i|; \quad \Pi(H_i)\Pi(H_{i+1}) = 0; \quad [\Pi(H_i), \Pi(H_{i+1})] = 0; \quad i \in \mathbb{Z}(5). \tag{147}$$

The terms exclusivity or local orthogonality are used in the literature, for the relation $\Pi(H_i)\Pi(H_{i+1}) = 0$ and its physical implications.

The $\mathfrak{C}_{i-1} = \{H_{i-1}, H_i\}$ and $\mathfrak{C}_i = \{H_i, H_{i+1}\}$ are overlapping contexts for all i, and

$$[\Pi(H_{i-1}), \Pi(H_i)] = [\Pi(H_i), \Pi(H_{i+1})] = 0; \quad [\Pi(H_{i-1}), \Pi(H_{i+1})] \neq 0. \tag{148}$$

We perform measurements with the projector $\Pi(H_i)$ on an ensemble of states with density matrix ρ, and get the result 'yes' with probability $p_i = \mathrm{Tr}[\rho\Pi(H_i)]$, and the result 'no' with probability $1 - p_i = \mathrm{Tr}[\rho\Pi(H_i^\perp)]$. We use the notation $1, 0$, for 'yes' and 'no' correspondingly. The projectors $\Pi(H_i), \Pi(H_j)$ do not commute in general, and the corresponding probabilities will be measured using different ensembles of states described by the same density matrix ρ. If $a_i = 1, 0$ is the outcome of the measurement $\Pi(H_i)$, the distributions $p(a_i)$ and $p(a_i, a_{i+1})$ are measurable.

B. A pentagram inequality within a non-contextual distributive hidden variable theory

In a non-contextual hidden variable theory, we assume that there exists a joint probability distribution $p(a_0, a_1, a_2, a_3, a_4)$, for the outcomes of the measurements in the previous subsection, which has as marginals the measurable distributions $p(a_i)$:

$$p(a_i) = \sum_{a_j \neq a_i} p(a_0, a_1, a_2, a_3, a_4). \tag{149}$$

We emphasize in this paper that this uses the law of total probability (proposition II.5), which is based on distributivity within the theory of Kolmogorov probabilities. So the non-contextual hidden variable theory, is assumed to be distributive, and this is consistent with the classical nature of the hidden variable theory.

Since $\Pi(H_i)\Pi(H_{i+1}) = 0$, the commuting measurements $\Pi(H_i), \Pi(H_{i+1})$ cannot both give 1, i.e., the a_i, a_{i+1} cannot both be equal to 1. Consequently, if the $(a_0, a_1, a_2, a_3, a_4)$ has more than two '1', the probability $p(a_0, a_1, a_2, a_3, a_4)$ is zero. This shows that the average number of 'yes' answers, satisfies the inequality

$$\sum_{a_i} (a_0 + a_1 + a_2 + a_3 + a_4) p(a_0, a_1, a_2, a_3, a_4) \leq 2. \tag{150}$$

Using Eq.(149) (which is based on distributivity), we rewrite this in terms of the marginal distributions, as

$$\sum_i a_i p(a_i) \leq 2. \tag{151}$$

In the quantum language, this is

$$\sum_{i=0}^{4} \mathrm{Tr}[\rho\Pi(H_i)] \leq 2, \tag{152}$$

where ρ is the density matrix of the system. It is known [16–24] that quantum mechanics violates this inequality. The left hand side can take the maximum value $\sqrt{5}$, and this proves that there exists no joint probability distribution $p(a_0, a_1, a_2, a_3, a_4)$. Quantum mechanics is a contextual theory where distributivity is replaced by the weaker property of modularity..

C. Degree of independence in the pentagram

We have explained above, that lack of distributivity makes problematic the use of marginal distributions in Eq.(149). We now delve deeper into this, and we stress that in set theory there is a unique concept of disjointness, which leads to a clear concept of partition, and to the law of the total probability. In the lattice of subspaces we have various levels of disjointness (independence), and consequently, the meaning of the joint probability and its marginals in Eq.(149), become problematic. This is another way of expressing the problems associated with joint probabilities that involve non-commuting observables. Only in the case of commuting observables, disjointness and pairwise disjointness are equivalent, the corresponding sublattice is distributive, and joint probabilities and their marginals are well defined.

We next show that in the pentagram we have the lowest level of disjointness (pairwise disjointness). It is easily seen that any pair of the subspaces H_0, H_1, H_2, H_3, H_4 (given in Eq.(145)), are independent. Also $\mathfrak{H}_i = H(3)$ and $\mathfrak{H}_i^\perp = \mathcal{O}$. Therefore $\mathfrak{H}_i^\perp \wedge H_i = \mathcal{O}$ for all i, and

$$\mathcal{A} = \frac{1}{5}\sum_i \Pi(H_i) = \begin{pmatrix} 0.30 & 0.10 & 0 \\ 0.10 & 0.36 & 0.02 \\ 0 & 0.02 & 0.34 \end{pmatrix}; \quad \eta(\rho) = \mathrm{Tr}(\mathcal{A}\rho). \tag{153}$$

Our degree of independence $\eta(\rho)$, is the quantity in the inequality of Eq.(152) divided by 5, for normalization purposes. The H_0, H_1, H_2, H_3, H_4 are pairwise independent, which is the weakest form of independence.

The eigenvalues of the matrix $\mathcal{A}$ are

$$\lambda_1 = 0.225; \quad \lambda_2 = 0.338; \quad \lambda_3 = 0.437. \tag{154}$$

Therefore $\eta(\rho) = \mathrm{Tr}(\mathcal{A}\rho)$ can reach the value 0.437, which violates the inequality in Eq.(152) that has normalized upper bound 2/5. In fact the value 0.437 is only slightly lower than the maximum value $\sqrt{5}/5 = 0.447$ given in the literature.

D. A pentagram inequality within quantum theory

The following inequality gives an upper bound 2.5 for $\sum \mathrm{Tr}[\rho\Pi(H_i)]$. The same upper bound has also been given in [39] through a different argument. Here it is easily proved using the language of lattice theory. This upper bound is of course higher than the maximum value for this quantity which is known to be $\sqrt{5}$.

Proposition VII.1. *We consider the subspaces H_i of the pentagram (an example of which is given in Eqs.(143), (145)). If ρ is the density matrix of the system, then*

$$\sum_{i=0}^{4} \mathrm{Tr}[\rho\Pi(H_i)] \le 2.5 \tag{155}$$

Proof. The fact that $\Pi(H_i)\Pi(H_{i+1}) = 0$, implies that $H_i \prec H_{i+1}^\perp$. Consequently,

$$\mathrm{Tr}[\rho\Pi(H_i)] \le \mathrm{Tr}[\rho\Pi(H_{i+1}^\perp)]. \tag{156}$$

We use these inequalities for all values i, and adding them we prove that

$$\sum_{i=0}^{4} \mathrm{Tr}[\rho\Pi(H_i)] \leq \sum_{i=0}^{4} \mathrm{Tr}[\rho\Pi(H_i^{\perp})]. \tag{157}$$

But we also have

$$\sum_{i=0}^{4} \mathrm{Tr}[\rho\Pi(H_i)] + \sum_{i=0}^{4} \mathrm{Tr}[\rho\Pi(H_i^{\perp})] = 5. \tag{158}$$

From the last two relations, follows the inequality in the proposition. □

VIII. DISCUSSION

An important property in classical physics and classical logic, formalized with Boolean algebra, is distributivity. In quantum physics and quantum logic, formalized with the Birkhoff-von Neumann lattice of subspaces, it is replaced by the weaker property of modularity (in systems with finite-dimensional Hilbert space). This has profound implications, some of which are discussed in this paper. Of course, within the lattice $\mathcal{L}(d)$ there are sublattices which are distributive (e.g., when the subspaces commute), and in those 'islands' results similar to classical physics do hold.

Within the formalism of phase space methods, we have considered the subspaces $H_1, ..., H_n$ of $H(d)$, and the quasi-probability distributions $R(i)$ in Eq.(26). We also introduced the quasi-probability distributions $\widetilde{R}(i)$ and $\widehat{R}(i)$ in Eq.(27), and discussed their physical meaning in terms of measurements.

In this general context, we have introduced the concepts of independence and totalness. We have shown that in quantum theory there are many levels of independence, from pairwise independence up to independence, and they are quantified with the degree of independence that compares the distributions $R(i)$ and $\widetilde{R}(i)$. There are also many levels of totalness, from pairwise totalness up to totalness, and they are quantified with the degree of totalness that compares the distributions $R(i)$ and $\widehat{R}(i)$. The existence of various levels of independence and totalness, is intimately related to the lack of distributivity in quantum theory. In set theory where distributivity holds, there is a single concept of independence and a single concept of totalness.

There is a duality between a set of independent subspaces, and the total set of the orthocomplements of these subspaces. Orthocomplementation (logical NOT operation) transforms independence into totalness.

One application of these ideas, is the law of total probability which is used to define marginals of probability distributions. We have explained that its proof relies on the distributivity property, and its application in non-distributive structures is problematic. This has been used in the pentagram, where a non-contextual distributive hidden variable theory leads to the inequality in Eq.(152), which is violated by quantum mechanics.

The work studies quantum theory from the angle of non-distributivity. It introduces novel concepts like the various levels of independence and the various levels of totalness, which can play a complementary role to non-commutativity, for the description of quantum phenomena.

---

[1] C. Zachos, D. Fairlie, T. Cartright, (Ed.)'Quantum Mechanics in Phase space' (World Scientific, Singapore, 2005)
[2] W. Schleich, 'Quantum Optics in Phase Space' (Wiley, Berlin, 2001)
[3] A. Vourdas, J. Phys. A39, R65 (2006)
[4] H. Whitney, Amer. J. Math., 55, 509 (1935)
[5] S. MacLane, Amer. J. Math., 58, 236 (1936)

[6] J. Oxley, 'Matroid theory' (Oxford Univ. Press, Oxford, 1992)
[7] J. von Neumann, 'Continuous geometry' (Princeton Univ. Press, Princeton, 1960)
[8] I. Halperin, Trans. Am. Math. Soc. 44, 537 (1938)
[9] B. Jonsson, Math. Scand. 1, 13 (1953)
[10] B. Jonsson, Math. Scand. 2, 295 (1954)
[11] G-C Rota, Notices of the Am. Math. Soc. 44, 1440 (1997)
[12] D. Finberg, M. Mainetti, G.C. Rota, in 'Lecture Notes in Pure and Applied Mathematics, 180, 69 (1996)
[13] M. Haiman, Adv. Math. 58, 209 (1985)
[14] C.H. Yan, Discrete Mathematics, 181, 295 (1998)
[15] C.H. Yan, Discrete Mathematics, 183, 285 (1998)
[16] J.S. Bell, Physics 1, 195 (1964)
[17] S. Kochen, E.P.Specker, J. Math. Mech. 17, 59 (1967)
[18] J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, Phys. Rev. Lett., 23, 880 (1969)
[19] N.D. Mermin, Rev. Mod. Phys. 65, 803 (1993)
[20] A.A. Klyachko, M.A. Can, S. Binicioglu, A.S. Shumovsky, Phys. Rev. Lett. 101, 020403 (2008)
[21] H. Bartosik, et al, Phys. Rev. Lett. 103, 40403 (2009)
[22] G. Kirchmair, et al, Nature 460, 494 (2009)
[23] T. Fritz, A.B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, A. Acin, Phys. Rev. A89, 032117
[24] A. Cabello, S. Severini, A. Winter, Phys. Rev. Lett. 112, 040401 (2014)
[25] A. Vourdas, J. Geom. Phys. 101, 38 (2016)
[26] A. Vourdas, Rep. Prog. Phys. 67, 1 (2004)
[27] A. Vourdas, *Finite and profinite quantum systems*, (Springer, Berlin, 2017)
[28] G. Birkhoff, J. von Neumann, Ann. Math. 37, 823 (1936)
[29] C. Piron, 'Foundations of quantum physics', Benjamin, New York, 1976
[30] G.W. Mackey, 'Mathematical foundations of quantum mechanics', Benjamin, New York, 1963
[31] J. Jauch, 'Foundations of quantum mechanics', Addison-Wesley, Reading, 1968
[32] V.S. Varadarajan, 'Geometry of quantum theory' (Springer, Berlin, 1968)
[33] E. Beltrametti, G. Cassinelli, 'The Logic of Quantum Mechanics', Addison-Wesley, Reading, 1981
[34] G. Birkhoff 'Lattice theory' (Amer. Math. Soc., Rhode Island, 1995)
[35] G. Szasz, 'Introduction to lattice theory' (Academic, London, 1963)
[36] G.A.Gratzer, 'General lattice theory' (Springer, Berlin, 2003)
[37] G. Kalmbach, 'Orthomodular lattices', Academic, London, 1983
[38] P. Ptak, S. Pulmannova, 'Orthomodular lattices as quantum logics', Kluwer, Dordrecht, 1991
[39] R. Ramanathan, A. Soeda, P. Kurzynski, D. Kaszlikowski, Phys. Rev. Lett. 109, 050404 (2012)