Journal of Mathematical Physics

Random positive operator valued measures

Cite as: J. Math. Phys. **61**, 042202 (2020); https://doi.org/10.1063/1.5131028
Submitted: 08 October 2019 . Accepted: 20 March 2020 . Published Online: 17 April 2020

Teiko Heinosaari, Maria Anastasia Jivulescu, and Ion Nechita

Journal of
Mathematical Physics

ARTICLE

scitation.org/journal/jmp

Random positive operator valued measures

Cite as: J. Math. Phys. 61, 042202 (2020); doi: 10.1063/1.5131028
Submitted: 8 October 2019 • Accepted: 20 March 2020 •
Published Online: 17 April 2020

View Online Export Citation CrossMark

Teiko Heinosaari,[1,a)] Maria Anastasia Jivulescu,[2,b)] and Ion Nechita[3,c)]

AFFILIATIONS

[1] Department of Physics and Astronomy, QTF Centre of Excellence, Turku Centre for Quantum Physics, University of Turku, Turku 20014, Finland

[2] Department of Mathematics, Politehnica University of Timişoara, Victoriei Square 2, 300006 Timişoara, Romania

[3] Laboratoire de Physique Théorique, Université de Toulouse, CNRS, UPS, France

[a)] **Electronic mail:** teiko.heinosaari@utu.fi
[b)] **Electronic mail:** maria.jivulescu@upt.ro
[c)] **Author to whom correspondence should be addressed:** nechita@irsamc.ups-tlse.fr

ABSTRACT

We introduce several notions of random positive operator valued measures (POVMs), and we prove that some of them are equivalent. We then study statistical properties of the effect operators for the canonical examples, starting from the limiting eigenvalue distribution. We derive the large system limit for several quantities of interest in quantum information theory, such as the sharpness, the noise content, and the probability range. Finally, we study different compatibility criteria, and we compare them for generic POVMs.

I. INTRODUCTION

In the last few years, significant developments have been reported in quantum information theory as a consequence of applying sophisticated techniques coming from random matrix theory and free probability theory. Indeed, the introduction of suitable models for random quantum states and channels has generated results in various topics, such as quantum entanglement,[1] classical capacity of quantum channels,[2] and additivity question.[3,4] It is of interest to apply such methods to other concepts or open problems from quantum information, such as *positive operator-valued measures* (POVMs).[5]

In this paper, we define *random POVMs* and study thoroughly their properties. Moreover, we ask questions about the (in)-compatibility of two independent random POVMs and find suitable conditions using various criteria from the literature. We actually present several models of randomness for POVMs, and we also study the connections between them. The most natural way to define a random POVM is as the image of diagonal unit rank projections through random unital, completely positive maps coming from Haar isometries. This is the model that we shall consider mostly in this paper:

Definition I.1. Fix an orthonormal basis $\{e_i\}_{i=1}^k$ of $\mathbb{C}^k$ and consider a Haar-distributed random isometry $V : \mathbb{C}^d \to \mathbb{C}^k \otimes \mathbb{C}^n$, for some integers d, k, n with $d \leq kn$. Define the random unital, completely positive map $\Phi(X) = V^*(X \otimes I_n)V$. A Haar-random POVM is the k-tuple $(M_1, \ldots, M_k)$ defined by $M_i := \Phi(|e_i\rangle\langle e_i|)$.

This model for random completely positive maps has been used also in other frameworks; see Ref. 4 (Sec. VI) for a review. Two other models of randomness for POVMs are introduced: one coming from the Lebesgue measure on the compact set of POVMs and the other given by the Wishart-random POVM ensemble. It is relevant to stress that in Theorem V.9, we prove the equivalence of the Wishart-random POVM model to the one coming from the Haar ensemble, while in Corollary V.14, we show that the Lebesgue measure is the special case $n = d$ in the definition above; these facts justify our choice in studying its properties. We would like to mention that random POVMs have been previously considered in the literature: Naimark dilations to a random orthonormal basis of $\mathbb{C}^n \otimes \mathbb{C}^k$ were considered in Ref. 6; in Ref. 7, the authors study Gaussian perturbations of a fixed POVM; normalized unit rank projections on independent, identically distributed (i.i.d.) random vectors were considered in Ref. 8, in a situation where the number of outcomes is larger than the dimension. Finally, in the work[9] (which appeared after the preprint version of our work was made available online), the authors compute several probabilities for the compatibility of independent dichotomic qubit POVMs, parameterized by points on the Bloch sphere.

Using the most general Wishart model, we analyze the spectral distribution of the effect operators, which are elements of the *Jacobi ensemble*.[10] We compute in Proposition VI.1 the moments of the individual effects from a Haar-random POVM using the (graphical) Weingarten calculus. In Proposition VI.2, we re-derive the asymptotic spectral distribution of the random effect as a dilatation of free additive convolution of a Bernoulli measure. These results are of help for deriving auxiliary properties of random POVMs, which involve spectral expressions, such as regularity, the norm-1 property, or the probability range. Furthermore, we study and compare (in)-compatibility criteria for Haar-random POVMs, such as the noise content criterion, the Jordan product criterion, the optimal cloning criterion, the Miyadera–Imai criterion, and the Zhu criterion. Our study shows that, for certifying compatibility for typical random POVMs, it is of interest to check first the Jordan product criterion.

The paper is organized as follows. In Sec. II, we recall basic notions related to POVMs, definitions, relevant examples, and remarks. Section III deals the notion of compatible POVMs and contains a brief presentation of the known incompatibility criteria. In Sec. IV, we review the basic ingredients needed for a good understanding of random matrix theory techniques used in the paper. To this aim, different topics are approached, such as random isometries, Weingarten calculus (also in its graphical incarnation), and some tools from Voiculescu's free probability theory. In Sec. V, we describe in details the models of randomness for POVMs and we state remarks related to their equivalence. We present in Sec. VI the statistical properties of random POVMs, whereas in Sec. VII, we consider incompatibility criteria for them, which are compared in Subsection VII F.

Before we move on, let us introduce some basic notation. We write $[n] := \{1, 2, \ldots, n\}$, and we denote by $\mathcal{S}_n$ the symmetric group acting on $[n]$. For a given permutation $\sigma \in \mathcal{S}_n$, we use the following notations: $\#\sigma$ is the number of cycles of σ and $|\sigma|$ is the length of σ, that is, the minimal number of transpositions that multiply to σ. We denote by $\gamma := (n, \ldots, 3, 2, 1) \in \mathcal{S}_n$ the full cycle permutation (in reverse order). In this paper, the following asymptotic notation is used:

$$x_n \sim y_n \Leftrightarrow \lim_{n\to\infty} \frac{x_n}{y_n} = 1.$$

We denote by $\mathcal{H}_d$ a finite d-dimensional complex Hilbert space and by $\mathcal{L}(\mathcal{H}_d)$ the algebra of linear operators on $\mathcal{H}_d$. Furthermore, we denote by Tr the (un-normalized) trace of matrices.

II. POVMs AND THEIR PROPERTIES

The states of a quantum system are mathematically described as density operators on a complex Hilbert space $\mathcal{H}_d \cong \mathbb{C}^d$, i.e., positive semi-definite operators with unit trace. A measurement is, mathematically speaking, a map that assigns a probability distribution to every state. The probability distribution is interpreted as the distribution of measurement outcomes. The additional requirement is that this kind of map is affine; a convex mixture of states must go into the respective mixture of the probability distributions. It follows that quantum measurements can be identified with positive operator valued measures (POVMs).[5] We will only consider POVMs with a finite number of outcomes, and Hilbert spaces are assumed to be finite dimensional. In this section, we recall some physically motivated properties of POVMs.

A. POVMs

For a POVM A, we denote by Ω_A the set of all outcomes of A, $\Omega_A = \{1, \ldots, k\}$ for some $k \in \mathbb{N}$. A POVM is then a map,

$$A : \Omega_A \to \mathcal{L}(\mathcal{H}), \quad i \mapsto A_i$$

such that $\sum_i A_i = I$ (=the identity operator on $\mathcal{H}_d$) and A_i are positive semi-definite operators, $A_i \geq 0$ for all $i \in \Omega_A$. The operators A_i are called the *effects* of the POVM A.

Example II.1. A POVM T is called trivial if T_i is proportional to the identity operator I for every outcome $i \in \Omega_T$. In this case, there is a probability distribution p on Ω_T such that $T_i = p_i I$.

Example II.2. Let $\{\varphi_i\}_{i=1}^d$ be an orthonormal basis of $\mathcal{H}_d$. We set $A_i = |\varphi_i\rangle\langle\varphi_i|$ for every $i = 1, \ldots, d$, and then A is a POVM. It is called the POVM associated to the orthonormal basis $\{\varphi_i\}_{i=1}^d$.

Both types of POVMs from the previous examples are *commutative*, i.e., $A_iA_j = A_jA_i$ for all $i, j \in \Omega_A$. One can easily construct examples of non-commutative POVMs by mixing two POVMs corresponding to two different orthonormal bases. There are also non-commutative POVMs that are extreme in the set of all POVMs with the same outcome set; we refer to Refs. 11 and 12 for further examples.

B. Operator range and probability range of a POVM

Let us first observe that when we have a measurement device that implements a POVM A, we can obtain not only the numbers $\mathrm{Tr}(\rho A_j)$ but also all sums of these numbers simply by grouping the measurement outcomes differently. For this reason, the following concept is useful when we talk about properties of POVMs.

Definition II.3. For a POVM A and a subset $X \subseteq \Omega_A$, we denote $A_X := \sum_{i \in X} A_i$. The (operator) range of A is the set

$$\mathrm{Ran}(A) := \{A_X : X \subseteq \Omega_A\}.$$

Instead of starting from POVMs, one can consider a measurement as an affine map from the state space to a probability simplex,

$$\Delta_k = \left\{(p_1, \ldots, p_k) : 0 \le p_i \le 1 \text{ for all } 1 \le i \le k, \sum_i p_i = 1\right\}.$$

It is well known that these descriptions are equivalent; any POVM determines such an affine map, and any such affine map determines a unique POVM.[13] From this point of view, it is of equal importance and interest to study both the range of the affine map related to a POVM and its operator range; see also Ref. 14.

Definition II.4. The probability range of a k-outcome POVM A is the convex subset of the probability simplex,

$$\mathrm{ProbRan}(A) := \{(\mathrm{Tr}(\rho A_1), \ldots, \mathrm{Tr}(\rho A_k)) \,:\, \rho \text{ is a density matrix}\} \subseteq \Delta_k.$$

A trivial POVM T, given as $T_i = p_i I$ for a probability distribution t, has a probability range reduced to the single point $p = (p_1, \ldots, p_k)$. On the other side of the spectrum, it is easy to see that a POVM A has full probability range, that is, $\mathrm{ProbRan}(A) = \Delta_k$, if and only if its effects have all unit operator norm, $\|A_i\| = 1$, for all $1 \le i \le k$; see Definition II.8.

We present next two examples of probability ranges. First, let us consider the case of diagonal effects. Let us assume that the POVM effects A_i are diagonal, $A_i = \mathrm{diag}(a_i)$, for some vectors $a_i \in [0,1]^d$ satisfying

$$\forall j \in [d], \qquad \sum_{i=1}^{k} a_i(j) = 1.$$

Considering the vectors $\alpha_j \in [0,1]^k$, for $j \in [d]$, defined by $\alpha_j(i) = a_i(j)$, we have the following result.

Proposition II.5. The probability range of a diagonal POVM A is the polytope $\mathrm{conv}\{\alpha_1, \ldots, \alpha_d\}$.

Proof. First, note that the normalization condition for the POVM A translates to the fact that α_j are probability vectors. Next, for a unit vector $x \in \mathbb{C}^d$, we have

$$[\langle x, A_i x\rangle]_{i=1}^{k} = \left[\sum_{j=1}^{d} |x_j|^2 a_i(j)\right]_{i=1}^{k} = \sum_{j=1}^{d} |x_j|^2 \alpha_j,$$

proving the claim. □

As an example, see the left panel in Fig. 1 where we have depicted the probability range of the following diagonal 3-outcome POVM:

$$\begin{aligned} A_1 &= \mathrm{diag}\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right),\\ A_2 &= \mathrm{diag}\left(\frac{1}{3}, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right),\\ A_3 &= \mathrm{diag}\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}\right). \end{aligned} \tag{1}$$

The following example of a non-trivial probability range is taken from Ref. 15. Consider a 3-outcome qubit POVM, with unit rank effects $A_i = \frac{2}{3}|a_i\rangle\langle a_i|$, where

$$a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} -1/2 \\ \sqrt{3}/2 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} -1/2 \\ -\sqrt{3}/2 \end{bmatrix}. \tag{2}$$

A direct computation shows that the squared distance from a point $(\mathrm{Tr}(\rho A_i))_{i=1}^3$ to the "center" $(1/3, 1/3, 1/3)$ of the probability simplex Δ_3 is less than $[(1-2a)^2 + 4|b|^2]/6$, where ρ is an arbitrary qubit density matrix,

$$\rho = \begin{bmatrix} a & b \\ \bar{b} & 1-a \end{bmatrix}.$$

Using the positivity condition for ρ, i.e., $|b|^2 \leq a(1-a)$, we conclude that the probability range of the POVM A is contained in a circle of radius $1/\sqrt{6}$ around the equiprobability vector $(1/3, 1/3, 1/3)$; doing the computations backward shows that in fact we have equality between the probability range and the aforementioned circle; see the right panel of Fig. 1.

C. Spectral properties of POVMs

This section contains a list of properties of quantum effects and POVMs relevant from the point of view of quantum information theory. We shall introduce them via a list of definitions followed by some simple properties and remarks; in Secs. VI and VII, we shall study these properties for random POVMs. All these properties reduce to some property on the spectrum of the effects. We denote by $\mathrm{spec}(E)$ the spectrum of an operator E.

1. Sharpness and regularity

An effect E is called *sharp* if it is a projection (i.e., $E^2 = E$), and otherwise *unsharp*. Hence, being sharp is equivalent to $\mathrm{spec}(E) \subseteq \{0, 1\}$. A POVM A is called sharp if A_i is sharp for every $i \in \Omega_A$; otherwise A is called unsharp. As a measure of unsharpness, we use the following.

Definition II.6. The unsharpness of an effect E is

$$\sigma(E) := 4 \parallel E - E^2 \parallel. \tag{3}$$

The unsharpness of a POVM A is

$$\sigma(A) := \max_i \sigma(A_i). \tag{4}$$

For quantum effects, we have $0 \leq \sigma(E) \leq 1$, with $\sigma(E) = 0$ iff E is a sharp and $\sigma(E) = 1$ iff $\frac{1}{2} \in \mathrm{spec}(E)$. For POVMs, it also holds that $0 \leq \sigma(A) \leq 1$.

One may ask if there is a qualitative property between sharpness and unsharpness. This kind of property is regularity.[16,17]

Definition II.7. An effect E is called regular if neither $E \leq \frac{1}{2}I$ nor $\frac{1}{2}I \leq E$. A POVM A is called regular if all effects, except 0 *and I, in* $\mathrm{Ran}(A)$ *are regular.*

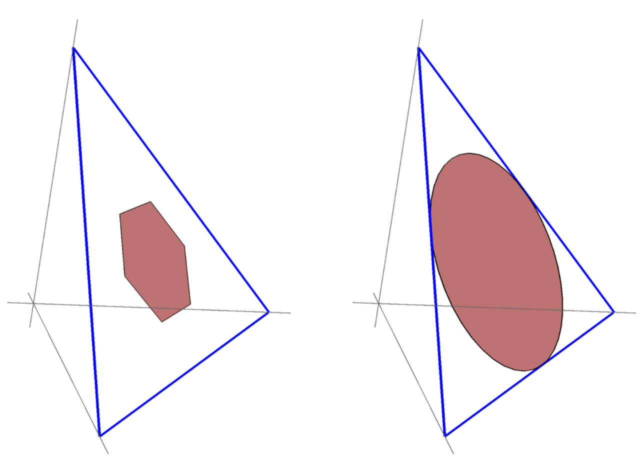

FIG. 1. Examples for the probability range of two POVMs with three outcomes. On the left, the diagonal POVM from (1). On the right, the example from (2); the probability range is a disk around the equiprobability vector $(1/3, 1/3, 1/3)$. The axes are in gray, the probability simplex Δ_3 is the blue triangle, and the probability range is the red convex set.

For effects, the definition above is equivalent to the fact that the spectrum spec(E) of E is not contained in $[0, \frac{1}{2}]$ or $[\frac{1}{2}, 1]$. Interestingly, it can be shown[16] that a POVM A is regular if and only if Ran(A) is a Boolean lattice with respect to the operator order $\leq$ and the complementation $E \mapsto I - E$ restricted to Ran(A).

2. Norm-1 property

We recall the following definition.[18]

Definition II.8. A POVM A has the norm-1-property if $\| A_i \| = 1$ *for every i.*

Physically, the norm-1-property means that for each outcome i, there is a state ρ_i such that the outcome i occurs with certainty, i.e., $\mathrm{Tr}[\rho_i A_i] = 1$. It follows that $\mathrm{Tr}[\rho_i A_j] = 0$ for $i \neq j$, thereby each operator A_i has both the eigenvalues 0 and 1. In particular, a POVM with the norm-1-property is regular.

We recall another characterization of the norm-1-property that links to a different physical property. An *instrument* $\mathcal{I}$ is a mapping from an outcome set (here, a finite set) to the set of quantum operations (completely positive maps), satisfying the obvious normalization and additivity properties[5] (Sec. 5.1.2). Instruments encode the transformations of a quantum state following a measurement, so they contain more information than POVMs, which only deal with the probabilities of obtaining different outcomes. For any given POVM A, there are several instruments that describe some state transformation associated with some measurement of A.[5] An instrument $\mathcal{I}$ is called *repeatable* if a subsequent measurement with the same device gives the same outcome, i.e.,

$$\mathrm{Tr}[\mathcal{I}_i(\mathcal{I}_j(\rho))] = \delta_{ij}\mathrm{Tr}[\mathcal{I}_i(\rho)].$$

As shown in Ref. 19 (Sec. III.4.6), a POVM A admits a repeatable instrument if and only if A has a norm-1-property. The POVMs with the norm-1-property have also appeared in relation to a strong notion of additivity for quantum channels; see Ref. 15 (Definition 1 and Theorem 4).

3. Noise content

Trivial POVMs (see Example II.1) can be use to describe measurement noise. Namely, if we start from a POVM A and mix it with a trivial POVM T, then we get a noisy version of A. Reversely, we can investigate how much noise a given POVM has. We recall the following definition.[20]

Definition II.9. The noise content $w(A)$ *of a POVM A is defined as*

$$w(A) := \sup\{0 \leq t \leq 1 \,:\, A = tT + (1-t)B \textit{ for some trivial POVM T and some POVM B with } \Omega_T = \Omega_B = \Omega_A\}.$$

It can be shown[20] that

$$w(A) = \sum_i \lambda_{min}(A_i),$$

where $\lambda_{min}(A_i)$ denotes the minimal eigenvalue of an operator A_i.

Instead of considering all trivial POVMs as noise, it is sometimes of interest to take only uniformly distributed trivial POVM as noise. The *uniform noise content* $w^u(A)$ of a POVM A with k outcomes is defined as

$$w^u(A) := \sup\{0 \leq t \leq 1 : A = t\frac{1}{k}I + (1-t)B \quad \text{for some POVM } B \text{ with } \Omega_B = \Omega_A\}.$$

In this case, we define similarly $w^u(A) = \min_i \lambda_{min}(A_i)$. We note that the uniform noise content behaves very differently than the noise content. For instance, $w(T) = 1$ for all trivial observables, whereas $w^u(T) = 0$ for $T = pI$ such that $p_i = 0$ for some outcome i.

III. INCOMPATIBILITY OF POVMs

Mathematically, incompatibility is an n-place relation in the set of n-tuples of POVMs. In this work, we concentrate only on the binary incompatibility relation. Physically speaking, the incompatibility relation describes the impossibility to measure simultaneously two (or more) POVMs. The simplest physical example of incompatible measurements consists of two different spin component measurements.[21] The realm and applications of incompatibility have been extensively developed in the past years. We refer to Ref. 22 for a more extensive explanation and for further references. In this section, we recall all results on incompatibility that are needed later.

A. Definition and basic properties

Given two POVMs A and B, we say that B is a *post-processing* of A if there exists a column stochastic matrix μ such that

$$B(x) = \sum_{y \in \Omega_A} \mu_{xy} A(y)$$

for all $x \in \Omega_B$. The post-processing relation is a preorder on the set of POVMs and has been introduced in Ref. 23.

Two POVMs A and B are *compatible* if there exists a third POVM C such that A and B are both post-processings of C; otherwise, A and B are *incompatible*. The compatibility relation is clearly reflexive and symmetric but not transitive.[24]

We recall that if A and B are compatible, then they are marginals of a third POVM,[25] called their joint POVM. Namely, let us assume that $A(x) = \sum \mu^A_{xz} C(z)$ and $B(y) = \sum \mu^B_{yz} C(z)$ for some POVM C and column stochastic matrices A and B. We then define a new POVM G as

$$G(x, y) = \sum_z \mu^A_{xz} \mu^B_{yz} C(z).$$

Then,

$$\sum_y G(x, y) = A(x), \quad \sum_x G(x, y) = B(y). \tag{5}$$

We see that if A and B are compatible and G satisfies (5), then $\mathrm{Ran}(A) \cup \mathrm{Ran}(B) \subset \mathrm{Ran}(G)$. However, the existence of a POVM G such that $\mathrm{Ran}(A) \cup \mathrm{Ran}(B) \subset \mathrm{Ran}(G)$ does not guarantee the compatibility of A and B.[26]

B. Criteria for compatibility

In the following, we recall three sufficient conditions for compatibility or, in other words, necessary conditions for incompatibility. We present their proofs for the reader's convenience.

Proposition III.1. Noise content criterion[20]: *if two POVMs A and B satisfy*

$$w(A) + w(B) \geq 1, \tag{6}$$

then they are compatible.

Proof. If (6) holds, then there exist trivial POVMs $S_i = p_i I$ and $T_j = q_j I$ and numbers $s, t \in [0, 1]$ such that $s + t = 1$ and $A = sS + (1 - s)A'$ and $B = tT + (1 - t)B'$ for some POVMs A', B'. We define a map M as $M_{ij} := tq_j A'_i + sp_i B'_j$. Then, M is a joint POVM for A and B. □

Proposition III.2. Jordan product criterion[27]*: if two POVMs A and B are such that*

$$\forall i, j : \qquad A_i \circ B_j := A_i B_j + B_j A_i \geq 0, \tag{7}$$

then A and B are compatible.

Proof. We define $M_{ij} = \frac{1}{2} A_i \circ B_j$. Then, $\sum_j M_{ij} = A_i$ and $\sum_i M_{ij} = B_j$. The requirement $A_i \circ B_j \geq 0$ implies that M is a valid POVM. □

This Jordan product criterion covers as a special case the following well-known implication: if A and B commute, then they are compatible.

Proposition III.3. Optimal cloning criterion[28]: *if two POVMs A and B satisfy*

$$\forall i : \lambda_{min}(A_i) \geq \frac{1}{2(1 + d)} \mathrm{Tr}[A_i], \tag{8}$$

$$\forall j : \lambda_{min}(B_j) \geq \frac{1}{2(1 + d)} \mathrm{Tr}[B_j], \tag{9}$$

then they are compatible.

Proof. We recall that the so-called symmetric universal cloning machine Λ, presented in Ref. 29, is defined as

$$\Lambda(\rho) = s_d\, S(\rho \otimes I) S,$$

where S is the projection from $\mathcal{H}_d^{\otimes 2}$ to the symmetric subspace of $\mathcal{H}_d^{\otimes 2}$ and the normalization coefficient s_d is independent of ρ. The state $\tilde{\rho}$ of each approximate copy is obtained as the corresponding marginal of $\Lambda(\rho)$ and, as it was shown in Ref. 30, it reads

$$\tilde{\rho} = c_d\rho + (1 - c_d)\frac{I}{d},$$

where the number c_d is independent of ρ and given by $c_d = (2 + d)/(2 + 2d)$. We are then performing measurements of POVMs A' and B' on the two copies of ρ obtained through the cloning machine. A measurement of A' on the approximate copy $\tilde{\rho}$ gives the same result as the action of the noisy POVM $c_d A' + (1 - c_d)T_{A'}$ on the initial state ρ, where $T_{A'}$ is the trivial POVM related to the probability distribution $\frac{1}{d}\mathrm{Tr}[A'_i]$. By choosing

$$A'_i = \frac{1}{c_d}\left[A_i - \frac{1 - c_d}{d}\mathrm{Tr}[A_i]I\right], \tag{10}$$

the mixture $c_d A' + (1 - c_d)T_{A'}$ is A. The condition is hence that A' in (10) is a valid POVM, meaning that

$$A_i \geq \frac{1 - c_d}{d}\mathrm{Tr}[A_i]$$

for each outcome i. This is equivalent to (8). □

Remark III.4. The conditions (8) and (9) look similar to Eq. (6) from the noise content compatibility criterion. They are, however, incomparable at a fixed dimension. To see this, let A and B be two qubit POVMs that correspond to two different bases $\{\varphi_1, \varphi_2\}$ and $\{\psi_1, \psi_2\}$. We form two families of noisy versions of A and B. First, we define A' and B' as $A'_1 = \frac{1}{2}A_1$, $A'_2 = \frac{1}{2}A_2 + \frac{1}{2}I$ and $B'_1 = \frac{1}{2}B_1$ and $B'_2 = \frac{1}{2}B_2 + \frac{1}{2}I$. The POVMs A' and B' satisfy the condition (6) but not (8) and *(9). Second, we define $A''_1 = \frac{2}{3}A_1 + \frac{1}{6}I$, $A''_2 = \frac{2}{3}A_2 + \frac{1}{6}I$ and $B''_1 = \frac{2}{3}B_1 + \frac{1}{6}I$ and $B''_2 = \frac{2}{3}B_2 + \frac{1}{6}I$. The POVMs A'' and B'' now satisfy the conditions (8)* and *(9) but not (6).*

C. Miyadera–Imai criterion for incompatibility

In Ref. 31 (Corollary 2), Miyadera and Imai provide a condition satisfied by all pairs of compatible POVMs. We recall it below (see also Ref. 22, Sec. 3.2). We denote by $[.,.]$ the commutator, and $\sigma(.)$ is the sharpness measure from Definition II.6.

Proposition III.5. If two POVMs A and B satisfy

$$4\|[A_i, B_j]\|^2 > \sigma(A_i)\cdot\sigma(B_j),$$

for all $i \in \Omega_A$ and $j \in \Omega_B$, then they are incompatible.

This condition covers as a special case following the well-known implication: if A is sharp, then any POVM B compatible with A commutes with A.

D. Zhu's criterion for incompatibility

In the following, we recall Zhu's criterion[32] for detecting incompatible observables, which stems from the application of the Gill–Massar inequality for Fisher information matrices.[33] The criterion has a constructive approach, which we recall briefly, for the reader's convenience.

Given two POVMs A and B, we define the superoperators $\mathcal{G}_A, \mathcal{G}_B \in \mathcal{M}_d(\mathbb{C}) \otimes \mathcal{M}_d(\mathbb{C})$ as

$$\mathcal{G}_A := \sum_i \frac{|A_i\rangle\langle A_i|}{\mathrm{Tr}[A_i]} \qquad \text{and} \qquad \mathcal{G}_B := \sum_j \frac{|B_j\rangle\langle B_j|}{\mathrm{Tr}[B_j]}, \tag{11}$$

where $|A_i\rangle = \mathrm{vec}(A_i) \in \mathbb{C}^d \otimes \mathbb{C}^d$ denotes the vectorization (or flattening) of the matrix A_i: if $X = \sum_{i,j} x_{ij} e_i e_j^*$ is a matrix, then

$$|X\rangle = \mathrm{vec}(X) = \sum_{ij} x_{ij} e_i \otimes e_j,$$

for some orthonormal basis $\{e_i\}_{i=1}^d$ of $\mathbb{C}^d$. By denoting

$$\tau(\mathcal{G}_A, \mathcal{G}_B) = \min_{H \geq \mathcal{G}_A, H \geq \mathcal{G}_B} \mathrm{Tr}[H], \tag{12}$$

Zhu established the following incompatibility criterion[32] [Eq. (10)].

Proposition III.6. If $\tau(\mathcal{G}_A, \mathcal{G}_B) > d$, then the POVMs A, B are incompatible.

It is clear that the quantity τ from (12) is the value of a semidefinite program.[34] Indeed, we can associate to it the Lagrangian

$$\mathcal{L}(H, x, y) = \mathrm{Tr}[H] + \langle x, \mathcal{G}_A - H\rangle + \langle y, \mathcal{G}_B - H\rangle$$

and define

$$\begin{aligned} g(x,y) &:= \min_H \mathcal{L}(H,x,y) = \min_H \langle H, I - x - y\rangle + \langle x, \mathcal{G}_A\rangle + \langle y, \mathcal{G}_B\rangle \\ &= \begin{cases} \langle x, \mathcal{G}_A\rangle + \langle y, \mathcal{G}_B\rangle, & \text{if } x + y = I \\ -\infty, & \text{otherwise.} \end{cases} \end{aligned} \tag{13}$$

Furthermore, by associating the dual condition to (13), it follows that

$$\max_{x,y\geq 0,\, x+y=I} \langle x, \mathcal{G}_A\rangle + \langle y, \mathcal{G}_B\rangle = \max_{0\leq x\leq I} \langle x, \mathcal{G}_A - \mathcal{G}_B\rangle + \mathrm{Tr}[\mathcal{G}_B]. \tag{14}$$

The optimal value for x for (14) is achieved at $x_{opt} = P_+(\mathcal{G}_A - \mathcal{G}_B)$, the orthogonal projection on the eigenspaces corresponding to non-negative eigenvalues of $\mathcal{G}_A - \mathcal{G}_B$; one notices the similarity between this SDP and the one for optimal discrimination of quantum states.[13,35] We conclude that

$$\tau(\mathcal{G}_A, \mathcal{G}_B) = \mathrm{Tr}[(\mathcal{G}_A - \mathcal{G}_B)_+] + \mathrm{Tr}[\mathcal{G}_B] = \frac{1}{2}[\mathrm{Tr}[\mathcal{G}_A] + \mathrm{Tr}[\mathcal{G}_B] + \|\mathcal{G}_A - \mathcal{G}_B\|_1]. \tag{15}$$

IV. INTERLUDE: RANDOM MATRIX THEORY AND FREE PROBABILITY

This section aims to recall basic definitions and concepts necessary for a facile understanding of the current work, rendering it self-contained. The theory of Haar-distributed random unitary operators and isometries is reviewed to be connected to the theory of random quantum channels and random POVMs. In addition, overviews on the (graphical) Weingarten calculus and free probability are given. In each case, the main concepts are presented, and the theorems which shall be used later are stated without proofs; references are given for the reader interested in further exploring these topics.

A. Random isometries and channels

Let us recall here the notion of quantum channel in order to justify the study of random isometries. A *quantum channel* is a linear map $\Psi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$ which is completely positive and trace preserving. Alternatively, using the dual map with respect to the usual scalar product, the map $\Psi^* : \mathcal{M}_k(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ is completely positive and unital. Stinespring's representation theorem [see, e.g., Ref. 5 (Chap. 4) or Ref. 36 (Chap. 2.2)] states that any quantum channel Ψ can be written as

$$\Psi(X) = [\mathrm{id}_k \otimes \mathrm{Tr}_n](VXV^*), \qquad \forall X \in \mathcal{M}_d(\mathbb{C}), \tag{16}$$

where $V : \mathbb{C}^d \to \mathbb{C}^k \otimes \mathbb{C}^n$ is an isometry. In the dual picture, we have the following representation of completely positive, unital maps:

$$\Psi^*(Y) = V^*(Y \otimes I_n)V, \qquad \forall Y \in \mathcal{M}_k(\mathbb{C}). \tag{17}$$

The Stinespring representation works also conversely: any isometry V gives rise to a quantum channel. This fact is used to introduce a random quantum channel, obtained by a random choice of the isometry V in (16) or (17); we explain next what we call a *random isometry*.

The set of all isometries $\{V : \mathbb{C}^d \to \mathbb{C}^D\}$ admits a *unique* left- and right-invariant probability measure, called the Haar measure, which can be obtained from the Haar measure on the unitary group $\mathcal{U}(kn)$[37] (Sec. 4.2) by truncation. More precisely, there is a unique probability measure μ_{Haar} on the set of isometries $\mathbb{C}^d \to \mathbb{C}^D$ which has the property that if $V \sim \mu_{Haar}$, then, for all $U_1 \in \mathcal{U}(d)$ and $U_2 \in \mathcal{U}(D)$, the isometry $U_2 V U_1 \sim \mu_{Haar}$.

Using random isometries, *random quantum channels* were introduced in Ref. 38 by choosing the isometry V appearing in the Stinespring representation from the Haar ensemble. Indeed, for each pair of integers d, k and for all values of the parameter n, the set of all channels $\{\Psi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})\}$ is endowed with the measure induced by the Haar distribution on the set of isometries V by the map (16) that associates to V the channel Ψ. Although there are many other probability distributions on the set of quantum channels, in this paper, we are going to be concerned with the one above.

This model of random quantum channels has been used with great success in the theory of quantum information, starting with the work of Hayden and Winter.[38] Subsequently, several authors[39–41] have studied the application of this model of randomness to the problem of additivity of the minimum output entropy of quantum channels; see Ref. 4 (Sec. 6) for a review.

B. Weingarten formula

In order to compute properties of random quantum channels, one has to integrate over the set of Haar-distributed random isometries, or, equivalently, over the set of Haar-distributed unitary operators. The expectation of products of entries of a random unitary operator has been considered in the physics literature by Weingarten in Ref. 42 for the case of large matrix dimension. The rigorous mathematical analysis at a fixed matrix size is due to Collins[43] and Collins–Śniady,[44] where it was shown, using Schur–Weyl duality, that the moment integrals can be expressed as sums over the symmetric group.

Theorem IV.1. *Let N be a positive integer and $i = (i_1, \ldots, i_n), i' = (i'_1, \ldots, i'_n), j = (j_1, \ldots, j_n), j' = (j'_1, \ldots, j'_n)$ n-tuples of positive integers from $[N] = \{1, 2, \ldots, N\}$. Let $U \in \mathcal{U}(N)$ be an $N \times N$ Haar-distributed unitary random matrix and denote by U_{ij} the (i, j)-th entry of U and $\delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$. Then, we have*

$$\int_{\mathcal{U}(N)} U_{i_1 j_1} \ldots U_{i_n j_n} \bar{U}_{i'_1 j'_1} \ldots \bar{U}_{i'_n j'_n} \mathrm{d}U = \sum_{\alpha, \beta \in \mathcal{S}_n} \delta_{i_1 i'_{\alpha(1)}} \ldots \delta_{i_n i'_{\alpha(n)}} \delta_{j_1 j'_{\beta(1)}} \ldots \delta_{j_n j'_{\beta(n)}} \mathrm{Wg}(N, \alpha^{-1}\beta), \tag{18}$$

where the function Wg *is called the* Weingarten function. *If $n \neq n'$, then*

$$\int_{\mathcal{U}(N)} U_{i_1 j_1} \ldots U_{i_n j_n} \bar{U}_{i'_1 j'_1} \ldots \bar{U}_{i'_{n'} j'_{n'}} \mathrm{d}U = 0. \tag{19}$$

The Weingarten function Wg dates back to Weingarten,[42] but the terminology and the notation were introduced by Collins.[43]

Remark IV.2. For $\alpha \in \mathcal{S}_n$, $n \leq N$ and for $U \in \mathcal{U}(N)$, an $N \times N$ Haar-distributed unitary random matrix, where dU *the normalized Haar measure, we have that*

$$\mathrm{Wg}(N, \alpha) = \int_{\mathcal{U}(N)} U_{11} \ldots U_{nn} \bar{U}_{1\alpha(1)} \ldots \bar{U}_{n\alpha(n)} \mathrm{d}U = \mathbb{E}[U_{11} \ldots U_{nn} \bar{U}_{1\alpha(1)} \ldots \bar{U}_{n\alpha(n)}].$$

In the following, we recall the definition of the Weingarten function, give examples of it, and present some of its properties used in the current paper.

Definition IV.3. The unitary Weingarten function $\mathrm{Wg}(N, \alpha)$, *depending on the dimension parameter N and on the permutation α in the symmetric group $\mathcal{S}_n$, is the inverse of the function $\alpha \mapsto N^{\#\alpha}$ under the following convolution operation for the symmetric group:*

$$\forall \sigma, \pi \in \mathcal{S}_n, \qquad \sum_{\tau \in \mathcal{S}_n} \mathrm{Wg}(N, \sigma^{-1}\tau) N^{\#(\tau^{-1}\pi)} = \delta_{\sigma,\pi}.$$

The Weingarten function has the particularity that it depends only on the cycle structure of the permutation. For example, Wg $(N, [2, 1])$ denotes the value of every permutation in $\mathcal{S}_3$ which decomposition consists of a transposition and a fixed point. It holds that

$$\mathrm{Wg}(N, [2, 1]) = \frac{-1}{(N^2 - 1)(N^2 - 4)}.$$

More details related to the computation of Weingarten functions are given in Ref. 44. The dimension parameter in the notation of Wg can be omitted when there is no confusion ($\mathrm{Wg}(N, \alpha) \equiv \mathrm{Wg}(\alpha)$). To the aim of our paper, it is of interest to present information about the behavior of the Wg function in the large limit of N (when n is kept fixed).

Remark IV.4. The asymptotics of the Weingarten function is given by

$$\mathrm{Wg}(N, \alpha) = N^{-(n+|\sigma|)}(\mathrm{M\ddot{o}b}(\alpha) + \mathcal{O}(N^{-2}),$$

where the Möbius function on the symmetric group is multiplicative with respect to the cycle structure of permutations,

$$\mathrm{M\ddot{o}b}(\alpha) = \prod_{c \text{ cycle of } \alpha} (-1)^{\mathrm{card}(c)-1} \mathrm{Cat}_{\mathrm{card}(c)-1}.$$

Here, Cat_N *is the N-th Catalan number and* $\mathrm{card}(c)$ *denotes the cardinality of the set on which the cycle c acts non-trivially:* $|c| = \mathrm{card}(c) - 1$. *In particular, if* α *is a product of disjoint transpositions, then*

$$\mathrm{M\ddot{o}b}(\alpha) = (-1)^{|\alpha|}.$$

Frequently, we shall use the (justified) notation $\mathrm{M\ddot{o}b}(\alpha^{-1}\beta) := \mathrm{M\ddot{o}b}(\alpha, \beta)$.

C. Graphical calculus for random independent unitary matrices

The integration formula (18) used to evaluate expectation over Haar-distributed unitary random matrices usually involves sums indexed by large sets of indices, which often turns out to be a complicated task to handle. In order to simplify tensors operations, the *graphical Weingarten formalism* was introduced in Ref. 45. It builds up on Penrose's graphical tensor notation[46] where diagrams consisting of boxes, decorations, and wires are used to represent tensors, collection of tensors, their dimensions, and contraction operations on them. In Ref. 45, expectation values of diagrams $\mathcal{D}$ containing random, Haar-distributed unitary matrices U and $\bar{U}$ are computed graphically, using the so-called *removal* procedure. According to (19), if the number of U boxes is different from the number of $\bar{U}$ boxes, then $\mathbb{E}\mathcal{D} = 0$. Otherwise, we shall use a pair of permutation $(\alpha, \beta) \in \mathcal{S}_n^2$ to pair the decorations of the n pairs of boxes $U/\bar{U}$. For each $i = 1, \ldots, n$, wires are used to connect white decorations of the k-th U box with the white decorations of the $\alpha(k)$-th $\bar{U}$ box. By a similar procedure, the black decorations are paired using now the β permutation; see Fig. 2. The next step consists of erasing the $U/\bar{U}$ boxes and denoting by $\mathcal{D}_{\alpha,\beta}$ the resulting diagram. It holds that

$$\mathbb{E}_U(\mathcal{D}) = \sum_{\alpha,\beta} \mathcal{D}_{\alpha,\beta} \mathrm{Wg}(N, \alpha^{-1}\beta). \tag{20}$$

The formula above is just the interpretation of the algebraic expression (18) in the tensor graphical calculus. Explicit examples for the use of (20) are given in Ref. 45.

D. Tools from free probability

This section aims to recall basic statements from free probability needed for a good understanding of the paper. We shall only sketch the concepts and results that shall be used later in the paper; we refer the reader to the monographs[47–49] for details.

A C^* *probability space* is the pair $(\mathcal{A}, \varphi)$, where $\mathcal{A}$ is a unital C^* algebra, with involution $a \mapsto a^*$, endowed with the state φ, i.e., $\varphi : \mathcal{A} \to \mathbb{C}$, φ-positive. The norm satisfies $\|a\| = \lim_{p\to\infty} (\varphi(a^p))^{1/p}$. Given a selfadjoint element a, the distribution of a, denoted by μ_a, is the probability measure on the spectrum of a, given by

$$\int x^p d\mu_a(x) = \varphi(a^p), \quad \forall p \in \mathbb{N}^*.$$

The number $\varphi(a^p), p \in \mathbb{N}^*$ is called the p-th moment of a. The moments of the random variable a are usually identified to the moments of the probability measure μ_a, which are given by $m_p(\mu_a) := \int x^p d\mu_a(x)$. In this paper, we are mostly concerned with the convergence of the

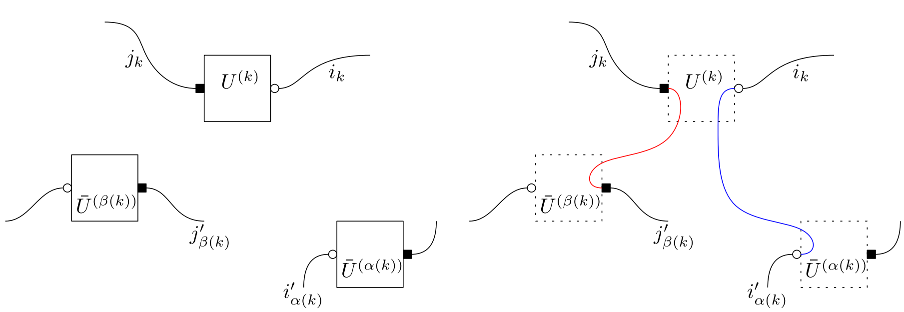

FIG. 2. A part of a tensor diagram before and after the removal procedure. On the right panel, we add new wires to pair the decorations of U with the decorations of $\bar{U}$ boxed according to the permutations α and β; on the left panel. we then delete the boxes corresponding to the Haar-distributed random unitary matrix U.

eigenvalues of random matrices. In C^* probability spaces, one can consider two types of convergence: the convergence in distribution (which is the convergence of all moments if, say, the limit measure has compact support) and the strong convergence (which implies, in particular, the convergence of the extreme eigenvalues of the matrices). It is of interest to recall that the convergence in distribution does not imply strong convergence.

Definition IV.5. Given the C^* probability spaces $(\mathcal{A}, \varphi, \|\cdot\|_\varphi)$ and $(\mathcal{A}^{(N)}, \varphi_N, \|\cdot\|_{\varphi_N})$ with $N \in \mathbb{N}$, where φ and φ_N are faithful traces. For the n-tuple $a = (a_1, \ldots, a_n) \in \mathcal{A}$ and $a^{(N)} = (a_1^{(N)}, \ldots, a_n^{(N)}) \in \mathcal{A}^{(N)}$, we say that

- *$a^{(N)}$ converges in distribution if*

$$\lim_{N\to\infty} \varphi_N[P(a^{(N)}, a^{(N)*})] = \varphi[P(a, a^*)].$$

- *$a^{(N)}$ converges strongly in distribution if, in addition,*

$$\lim_{N\to\infty} \|P(a^{(N)}, a^{(N)*})\|_{\varphi_N} = \|P(a, a^*)\|_\varphi.$$

The theory of free probability is based on new concepts such as *free independence, free cumulants, and free convolution*. In the following, we recall some of them. Given a probability measure μ on the real line with compact support, its free cumulants $k_p(\mu)$ are given by the moment-cumulant formula,[48]

$$m_p(\mu) = \sum_{\pi \in NC(p)} \prod_{b\in\pi} k_{|b|}(\mu). \tag{21}$$

Obviously, the free-cumulants $k_p(\mu)$ contain the same information as the moments of the measure $m_p(\mu)$.

We recall that given two free elements a, b having distributions μ, ν, the distributions of $a + b$ is denoted by $\mu \boxplus \nu$ and it is called the *free additive convolution* of μ and ν; see Ref. 48 (Lecture 12). Given the case of Bernoulli distributions $b_t = (1-t)\delta_0 + t\delta_1$, it holds that [see Ref. 47 (Example 3.6.7) and Ref. 48 (Exercise 14.21)].

Proposition IV.6. For any $T \geq 1$, the free additive power of a Bernoulli distribution is given by

$$b_s^{\boxplus T} = \max(0, 1 - Ts)\delta_0 + \max(0, 1 - T(1-s))\delta_T + \frac{T\sqrt{(\gamma^+(s,T) - x)(x - \gamma^-(s,T))}}{2\pi x(T-x)} \mathbf{1}_{[\gamma^-(s,T),\gamma^+(s,T)]}(x)dx,$$

where $\gamma^\pm = (T-2)s + 1 \pm 2\sqrt{(T-1)s(1-s)}$.

We recall below a lemma related to the push-forward property of the free additive convolution of probability measures. The following notation is used: $f_{\#\mu}$ is the push-forward of a measure μ by a measurable function f; it holds that given a random variable X of distribution μ, then $f(X)$ has distribution $f_{\#\mu}$.

Lemma IV.7. Let μ be a compactly supported probability measure on $\mathbb{R}$ so that, for any $T \geq 1$, the distribution $\mu^{\boxplus T}$ is well-defined. Then, we have, for any $a, b \in \mathbb{R}$,

$$((x \mapsto ax + b)_{\#\mu})^{\boxplus T} = (x \mapsto ax + Tb)_\#(\mu^{\boxplus T}).$$

V. RANDOM POVMs

This section contains one of the main contributions of this work, the definition and the basic properties of *random POVMs*. We focus on one specific model, which we study in detail in Subsection V A; this same model will be used in the rest of the paper to analyze the different quantities and (in-)compatibility criteria introduced in Subsections V B and C. In the second part of this section, we consider an *a priori* different probability distribution over the set of POVMs of a given size with a given number of outcomes, obtained by normalizing independent Wishart random matrices. We then show that, in the range of parameters we are interested in, under some symmetry assumption, this Wishart-like distribution coincides with our main model. Finally, we briefly discuss other possibilities for defining random POVMs in Subsection V C

A. Haar-random POVMs

Our approach to random POVMs comes from the observation that if $\Phi : \mathcal{M}_k(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ is a unital and positive map, then the image of the diagonal projections $\{|e_i\rangle\langle e_i|\}_{i=1}^k$ through Φ form a POVM,

$$M_i := \Phi(|e_i\rangle\langle e_i|), \qquad 1 \le i \le k.$$

Indeed, since the map Φ is positive, the POVM elements M_i are positive semidefinite, and the total probability condition follows from the fact that Φ is unital:

$$\sum_{i=1}^{k} M_i = \Phi\left(\sum_{i=1}^{k} |e_i\rangle\langle e_i|\right) = \Phi(I_k) = I_d.$$

We are going to strengthen the requirements above and consider *random unital, completely positive maps* Φ coming from Haar isometries. Such maps Φ are duals (for the Hilbert–Schmidt scalar product) of *random quantum channels*; see the discussion at the beginning of Sec. IV. As explained, choosing the isometry V appearing in the formula (16) for the Stinespring dilation of a quantum channel [or its dual; see (17)], one induces a random quantum channel. Let us note that a similar model of random POVMs has been introduced in Ref. 6 in relation to the hidden subgroup problem; there, however, the focus was on the distinguishability power of such measurements, and the analytical properties of the random POVMs were not investigated.

Definition V.1. Fix an orthonormal basis $\{e_i\}_{i=1}^{k}$ of $\mathbb{C}^k$ and consider a Haar-distributed random isometry $V : \mathbb{C}^d \to \mathbb{C}^k \otimes \mathbb{C}^n$, for some triple of integers (d, k, n) with $d \le kn$. Define the unital, completely positive map

$$\begin{aligned} \Phi : \mathcal{M}_k(\mathbb{C}) &\to \mathcal{M}_d(\mathbb{C}) \\ X &\mapsto V^*(X \otimes I_n)V. \end{aligned}$$

A Haar-random POVM of parameters $(d, k; n)$ is the k-tuple $(M_1, \ldots, M_k)$ defined by

$$\mathcal{M}_d(\mathbb{C}) \ni M_i := \Phi(|e_i\rangle\langle e_i|).$$

Remark V.2. In Definition V.1, the parameter n can be any integer satisfying $n \ge d/k$. However, the distribution of the random POVM M makes sense for all real values $n \in [d/k, \infty)$; see Remark V.10.

Remark V.3. The POVM elements M_i can be written as

$$M_i = V_i^* V_i, \qquad 1 \le i \le k,$$

where V_i are the $n \times d$ blocks of V, $V = \sum_{i=1}^{k} |e_i\rangle \otimes V_i$.

Remark V.4. A decomposition $M_i = V_i^* V_i$ defines an instrument of the POVM M. Namely, the map $\rho \mapsto V_i \rho V_i^*$ has the properties that

$$\mathrm{Tr}\left[V_i \rho V_i^*\right] = \mathrm{Tr}[\rho M_i]$$

and $\rho \mapsto \sum_i V_i \rho V_i^*$ is a quantum channel. We remark that not all instruments of M are of this form; in general, an instrument $\mathcal{I}$ of M is given as

$$\mathcal{I}_i(\rho) = \sum_{j \in X_i} V_{ij} \rho V_{ij}^*,$$

with

$$\sum_{j \in X_i} V_{ij}^* V_{ij} = M_i,$$

where $X_1, \ldots, X_k$ form a partition of some index set $\{1, \ldots, n\}$ into disjoint subsets.

Remark V.5. The same approach for constructing random POVMs is used in the function randomPOVM *of the QETLAB library[50], with the particular choice $n = d/k$. Our MATLAB routine[51] is more general and allows for arbitrary integer values of n (satisfying $n \ge d/k$).*

We gather in the next proposition some basic facts about Haar-random POVMs.

Proposition V.6. Let $M = (M_1, \ldots, M_k)$ be a Haar-random POVM of parameters $(d, k; n)$. The random k-tuple M is a permutation invariant; for any permutation $\sigma \in \mathcal{S}_k$, the random variables

$$(M_1, \ldots, M_k) \quad \text{and} \quad (M_{\sigma(1)}, \ldots, M_{\sigma(k)})$$

have the same distribution. In particular, the random matrices $\{M_i\}_{i=1}^k$ *are identically distributed. Moreover, with probability one, the rank of a POVM element* M_i *is* $\min(d, n)$.

Proof. The first assertion follows from the invariance of the Haar distribution of the random isometry V from Definition V.1; for any permutation $\sigma \in \mathcal{S}_k$, the isometries V and $(P_\sigma \otimes I_n)V$ have the same distribution (here, P_σ is the permutation matrix corresponding to σ).

The second assertion follows from the fact that the rank of any sub-matrix of a Haar-distributed random unitary matrix is the minimum of its dimensions (i.e., the maximum rank allowed). Indeed, if a sub-matrix had smaller rank, one could find a polynomial in the matrix entries which would vanish; it is a classical result in algebraic geometry (see, e.g., Ref. 52, Lemma 4.3 and the references within) that such a polynomial either vanishes on the whole unitary group or it vanishes on a set of measure zero. By constructing an explicit example, one can see that the former situation cannot happen, and the proof is complete. □

Remark V.7. In Sec. V, we will vary the parameter n to interpolate between POVMs having elements with small rank ($n \ll d$) *and POVMs with invertible elements* ($n \geq d$) *allowing us to test the strength of various necessary (respectively, sufficient) conditions for compatibility found in the literature. One of the reasons we prefer this model of randomness for POVMs is the existence (at fixed* d, k) *of this 1-parameter family of probability measures. A similar framework was developed for the study of random quantum states; see Refs. 53 or 4 (Sec. IV.A.2).*

B. Wishart-random POVMs

We consider in this section another model of randomness than the one stemming from (duals of) random quantum channels. The starting point here is the Wishart ensemble of random matrix theory[54] [see also Ref. 55 (Chap. 3) for a textbook introduction]. Recall that a Wishart random matrix with parameters (d, s) is given by $W = G^* G$, where $G \in \mathcal{M}_{s\times d}(\mathbb{C})$ is a *Ginibre random matrix*, that is, a matrix with i.i.d. complex standard Gaussian entries. Wishart matrices are, by construction, positive semidefinite, so one needs to apply a normalization procedure in order to construct a POVM. A similar model has been considered in Ref. 8, where instead of normalizing independent Wishart matrices, the authors consider independent rank one projection (this imposes that the number of outcomes should be larger than the dimension).

We summarize the construction in the following definition.

Definition V.8. A Wishart-random POVM of parameters $(d, k; s_1, \ldots, s_k)$ *is a k-tuple of matrices* $(M_1, \ldots, M_k)$, *where*

$$M_i = S^{-1/2} W_i S^{-1/2}$$

with $S = \sum_{i=1}^k W_i$ *and* $\{W_i\}_{i=1}^k$ *is a family of independent Wishart matrices of respective parameters* (d, s_i) *for* $1 \leq i \leq k$.

The Wishart-POVM ensemble might be useful in practice in the presence of an *a priori* requiring different distribution for the POVM elements. Note that, in the case when $s_1 = \cdots = s_k$, the distribution of the POVM elements $\{M_i\}$ is a permutation invariant. In fact, in the case where the common value of the parameters is an integer, the distribution of a Wishart-random POVM is exactly the distribution from Definition V.1.

Theorem V.9. *The distribution of a Wishart-random POVM of parameters* $(d, k; n, n, \ldots, n)$ *is equal to the distribution of a Haar-random POVM of parameters* $(d, k; n)$.

Proof. Consider a Wishart-random POVM obtained from independent complex Gaussian matrices $G_1, \ldots, G_k \in \mathcal{M}_{n\times d}(\mathbb{C})$. Stack the G_i matrices on top of each other to form

$$G := \sum_{i=1}^{k} |i\rangle \otimes G_i \in \mathcal{M}_{kn\times d}(\mathbb{C}).$$

The matrix G is again a Gaussian matrix since its entries are independent and follow a standard complex Gaussian distribution. Hence, its polar decomposition $G = VP$ can be chosen in such way that

1. the positive part is $P = (G^* G)^{1/2} \geq 0$, with $P \in \mathcal{M}_d(\mathbb{C})$ and
2. the angular part $V : \mathbb{C}^d \to \mathbb{C}^{kn}$ is Haar distributed.

The latter condition follows from the unitary invariance of the Gaussian ensemble (note that $d \leq kn$). We have

$$W_i = G_i^* G_i = G^* (I_n \otimes |i\rangle\langle i|) G,$$

and thus, $S = \sum_{i=1}^k W_i = G^*G = P^2$. It follows that $S^{-1/2} = P^{-1}$ (where one might need to use the pseudo-inverse), and thus,

$$M_i = S^{-1/2} W_i S^{-1/2} = P^{-1} G^* (I_n \otimes |i\rangle\langle i|) G P^{-1} = V^* (I_n \otimes |i\rangle\langle i|) V = V_i^* V_i,$$

where we have decomposed

$$V = \sum_{i=1}^k |i\rangle \otimes V_i.$$

Since V was chosen to be a Haar isometry, the conclusion follows. □

Remark V.10. Since Wishart matrices with parameters (d, s) can be defined not only for integer s but also for all real $s \geq d$, one can consider (Haar or Wishart)-random POVMs of parameters $(d, k; n)$ for any integers d, k and

$$n \in \left\{ \left\lceil \frac{d}{k} \right\rceil, \left\lceil \frac{d}{k} \right\rceil + 1 \ldots, d-1 \right\} \cup [d, \infty).$$

Remark V.11. In practice, it is computationally cheaper to sample random Haar POVMs using Wishart matrices, than using Haar-distributed random isometries. However, from an analytical perspective, it is often more enlightening to use Definition V.1 of random POVMs.

Remark V.12. Let us also point out that the distribution of a single effect of a Wishart-random POVM is given by the Jacobi (or double Wishart) distribution from random matrix theory. Indeed. if M is a Wishart-random POVM of parameters $(d, k; s_1, \ldots, s_k)$, then the random matrix M_i has the same distribution as $(A+B)^{-1/2} A (A+B)^{-1/2}$, where A has a Wishart distribution of parameters (d, s_i) and B is another Wishart matrix, independent from A, with parameters $(d, \check{s}_i)$, with $\check{s}_i = \sum_{j \neq i} s_j$.

One can explicitly compute the density of a Wishart-random POVM with respect to the Lebesgue measure on k-tuples of Hermitian matrices using the matrix Dirac delta function;[56,57] we defer the proof to the Appendix.

Theorem V.13. *The distribution of a Wishart-random POVM of parameters $(d, k; s_1, s_2, \ldots, s_k)$ has the following density at a point $m = (m_1, \ldots, m_k)$:*

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{dLeb}}(m_1, \ldots, m_k) = C_{d,k,s_1,\ldots,s_k} \mathbf{1}_{\sum_j m_j = I_d} \prod_i \mathbf{1}_{m_i \geq 0} \det(m_i)^{s_i - d}, \tag{22}$$

where $C_{d,k,s_1,\ldots,s_k}$ is a normalization constant.

Corollary V.14. In the particular case where $s_1 = \cdots = s_k = d$, the density above is flat, so one recovers the Lebesgue measure on the set of POVMs.

C. Other distributions

A third model of random POVMs comes from the notion of *random bases*. Consider, for fixed d, a random basis $\{e_1, \ldots, e_d\}$ of $\mathbb{C}^d$, which can be obtained from the columns of a Haar-distributed, random unitary matrix U. For a mixing parameter $t \in [0,1]$, define the effect operators $M_i = t|e_i\rangle\langle e_i| + (1-t)I/d$, for all $i \in [d]$. This procedure defines a random d-outcome POVM in $\mathcal{M}_d(\mathbb{C})$, depending on the parameter t. For $t = 1$, we obtain a random von Neumann measurement on the vectors e_i, while for $t = 0$, we get the trivial uniform POVM $(I/d, \ldots, I/d)$. Note that for this model of random POVMs, the number of outcomes is equal to the dimension of the effects.

A fourth model is provided by the *Lebesgue measure*. By Corollary V.14, this measure is a special case of the parametric families we consider; we can obtain it either as a Haar-random POVM of parameters $(d, k; d)$ or as a Wishart-random POVM of parameters $(d, k; d, \ldots, d)$.

Finally, let us mention that random perturbations by Gaussian noise of a fixed given POVM have been considered in Ref. 7 in a numerical algorithm used to find the optimal POVM for some particular state-estimation problem.

VI. STATISTICAL PROPERTIES OF RANDOM POVMs

We consider in this section the statistical properties of the effects $M_1, \ldots, M_k$ of a random POVM M, sampled from the ensemble introduced in Sec. V A, Definition V.1. We shall be interested in the asymptotic spectrum of the individual effects M_i. These effect operators are elements of the *Jacobi ensemble*, introduced by Wachter[10] and studied thoroughly in the random matrix theory literature [Ref. 58 (Sec. 13.2), Refs. 59 and 55 (Theorem 4.10), and Ref, 60 (Sec. 3.6)]. We use the graphical Weingarten calculus from Ref. 45 to obtain moment formulas in a simple, combinatorial way and then use free probability to re-derive the limiting spectral distribution.

A. Exact moments of random effects

In the following proposition, we aim to compute explicitly the moments of a POVM element M_i from the Haar-POVM ensemble; note that since the distribution of the random POVM M is permutationally invariant, the value of i is irrelevant, so we shall set $i = 1$. We shall use the graphical Weingarten calculus introduced in Sec. IV C.

Proposition VI.1. For any integer dimensions parameters n, d, the moments of the random matrix $M_1 \in \mathcal{M}_{nd}(\mathbb{C})$ are given by

$$\forall p \geq 1, \qquad \mathbb{E}\,\mathrm{Tr}\, M_1^p = \sum_{\alpha,\beta\in\mathcal{S}_p} n^{\#\alpha} d^{\#(\beta\gamma^{-1})}\mathrm{Wg}(kn, \alpha^{-1}\beta). \tag{23}$$

Proof. Without loss of generality, we can replace the random isometry V in the definition of a random POVM by a random Haar-distributed unitary matrix $U \in \mathcal{U}_{kn}$. We aim to compute, for $\forall p \geq 1$, the moment $\mathbb{E}\,\mathrm{Tr}\, M_1^p$; using indices, this reads

$$\mathbb{E}\,\mathrm{Tr}\, M_1^p = \mathbb{E}\sum_{i_1,\ldots,i_p=1}^{d} M_1(i_1,i_2)M_1(i_2,i_3)\cdots M_1(i_p,i_1).$$

In the graphical notation, we aim to compute the expectation of the diagram $\mathcal{D}$ in Fig. 3. We use the formula (20) to compute the expectation value with respect to the random unitary matrix U. We use the removal algorithm, which assumes the rules recalled below:

- replace U^* boxes by $\bar{U}$, as the removal procedure is requiring to pair decorations of the same color; the resulting diagram is presented in Fig. 4,
- round decorations correspond to $\mathbb{C}^n$, whereas the square ones correspond to $\mathbb{C}^d$. Diamond shaped decorations correspond to $\mathbb{C}^k$, but they are not important in what follows since their contribution will be trivial,
- we aim to wire p groups of $(U, \bar{U})$,
- using formula (20), the expectation of the diagram is a weighted sum (with Weingarten weights) of diagrams $\mathcal{D}_{\alpha,\beta}$, obtained after the removal of U and $\bar{U}$, and
- the loops in the diagram are of two types: the ones connecting round decorations(each having a value of n) and the others are connecting square decorations (each having a value of d).

In consequence, the diagram $\mathcal{D}_{\alpha,\beta}$ consists of a collection of loops that correspond to different vector spaces, as follows:

- $\#\alpha$ loops of dimension n, corresponding to the round-shaped white labels. These decorations are actually connected to the identity permutation (in the original diagram) and the graphical expansion connects them by α. The resulting number of loops is $\#\alpha = \#(\alpha \cdot \mathrm{id}^{-1})$
- $\#(\beta\gamma^{-1})$ loops of dimension d, corresponding to square-shaped black labels. The square decorations are initially connected with the permutation

$$\gamma := (p\,p-1\,\cdots\,3\,2\,1) \in \mathcal{S}_p$$

that allows links of the form $l \to l-1$ and the graphical expansion connects them with the permutation β. The total number of loops is $\#(\beta\gamma^{-1})$.

Putting together the contributions above, weighted by the Weingarten factors, we obtain the claimed formula. □

Let us consider now the simplest cases of the formula in the result above, $p = 1$ and $p = 2$, respectively.

At $p = 1$, there is only one term in the sum, and we obtain

$$\mathbb{E}\,\mathrm{Tr}\, M_1 = dn\mathrm{Wg}(kn, (1)) = \frac{dn}{kn} = \frac{d}{k}. \tag{24}$$

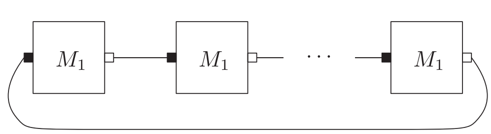

FIG. 3. The diagram corresponding to the moment $\mathbb{E}\,\mathrm{Tr}\, M_1^p$. There are p copies of the box M_1. The square labels attached to the boxes correspond to the space $\mathbb{C}^d$.

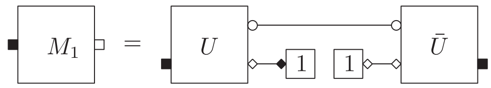

FIG. 4. The diagram for the random matrix M_1. We just write 1 for the basis element $e_1 \in \mathbb{C}^k$.

This result was to be expected since we know that

$$d = \operatorname{Tr} I_d = \sum_{i=1}^{k} \mathbb{E} \operatorname{Tr} M_i = k \mathbb{E} \operatorname{Tr} M_1.$$

For $p = 2$, the result is already non-trivial. We have that $\mathbb{E} \operatorname{Tr} M_1^2$ is a sum of four terms, corresponding to $\alpha, \beta \in \{\mathrm{id}, (12)\}$; the corresponding diagrams $\mathcal{D}_{\alpha,\beta}$ are depicted in Fig. 5. The terms are as follows: the wiring $\alpha = \beta = \mathrm{id}$ gives a contribution of $n^2\ dWg(kn, \mathrm{id})$. When $\alpha = \mathrm{id}$ and β is the transposition (12), we get the term $n^2 d^2 \mathrm{Wg}(kn, (12))$; but, if $\alpha = (12)$ and $\beta = \mathrm{id}$, the contribution to the sum is $nd\mathrm{Wg}(kn, (12))$. The final situation, corresponding to $\alpha = \beta = (12)$, yields the term $nd^2 \mathrm{Wg}(nk, \mathrm{id})$. In conclusion, the total sum reads

$$\mathbb{E} \operatorname{Tr} M_1^2 = (n^2 d + nd^2)\mathrm{Wg}(kn, \mathrm{id}) + (n^2 d^2 + nd)\mathrm{Wg}(kn, (1\,2)).$$

Using the corresponding values for the Weingarten functions

$$\mathrm{Wg}(kn, \mathrm{id}) = \frac{1}{(kn)^2 - 1}, \qquad \mathrm{Wg}(kn, (1\,2)) = \frac{-1}{kn((kn)^2 - 1)},$$

it follows that

$$\mathbb{E} \operatorname{Tr} M_1^2 = (n^2 d + nd^2)\frac{1}{(kn)^2 - 1} + (n^2 d^2 + nd)\frac{-1}{kn((kn)^2 - 1)} = \frac{d(kn^2 + dn(k-1) - 1)}{k((kn)^2 - 1)}. \tag{25}$$

A similar computation gives the covariance between two different random POVM elements M_1 and M_2. The diagram for $\operatorname{Tr}[M_1 M_2]$ consists of the product of two copies of the diagram in Fig. 4, with the "1" replaced by a "2" in the second copy. This fact imposes the constraint $\beta = \mathrm{id}$ in the Weingarten sum: terms with $\beta = (12)$ are zero because of the scalar product $\langle e_1, e_2 \rangle$. We have, thus,

$$\mathbb{E}\operatorname{Tr}[M_1 M_2] = n^2 d\mathrm{Wg}(kn, \mathrm{id}) + n^2 d^2 \mathrm{Wg}(kn, (1, 2)) = \frac{nd(kn - d)}{k((kn)^2 - 1)}. \tag{26}$$

B. The asymptotical spectral distribution of random POVM effects

With the help of free probability theory, we give here a simple derivation of the formula for the distribution of a POVM element M_i from the Haar-POVM ensemble, in the large d limit; for different approaches; see, e.g., Ref. 55 (Theorem 4.10). To be more precise, we shall consider the following asymptotical regime:

- k, the number of outcomes of the POVM, is fixed,
- d, the dimension of the POVM effects, grows to infinity,

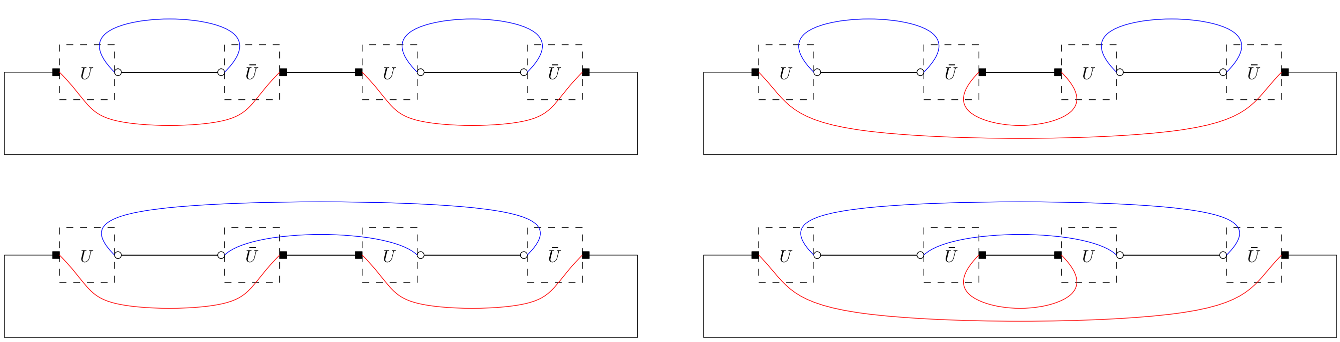

FIG. 5. The four diagrams appearing in the graphical expansion of $\mathbb{E} \operatorname{Tr} M_1^2$. From top to bottom and left to right, the diagrams correspond to $(\alpha, \beta) = (\mathrm{id}, \mathrm{id})$, $(\mathrm{id}, (12))$, $((12), \mathrm{id})$, $((12), (1(2))$. The permutation α is drawn on top, in blue, while β is drawn downward, in red. The value of each diagram is given by the number of loops in blue/red.

- n, the parameter appearing in the definition of Haar-random POVMs, grows to infinity, in such a way that

$$\lim_{d\to\infty} \frac{d}{kn} = t,$$

where $t \in (0,1]$ is a constant.

Proposition VI.2. In the asymptotical regime where k is fixed and $d, n \to \infty$ in such a way that $d \sim tkn$ for some constant $t \in (0,1]$, the distribution of a POVM element M_i from the Haar-POVM ensemble of parameters $(d,k;n)$ converges in moments toward the probability measure,

$$\begin{aligned} D_t\Big[b_{k^{-1}}^{\boxplus t^{-1}}\Big] &= \max(0, 1 - t^{-1}k^{-1})\delta_0 + \max(0, 1 - t^{-1} + t^{-1}k^{-1})\delta_1 \\ &\quad + \frac{\sqrt{(x-\varphi_-)(\varphi_+ - x)}}{2\pi t x(1-x)}\mathbf{1}_{[\varphi_-,\varphi_+]}(x)\mathrm{d}x, \end{aligned} \tag{27}$$

where

$$\varphi_\pm = t + k^{-1} - 2tk^{-1} \pm 2\sqrt{t(1-t)k^{-1}(1-k^{-1})}. \tag{28}$$

Above, $D_\cdot$ is the dilation operator (if X has distribution μ, then aX has distribution $D_a\mu$), b is the Bernoulli distribution ($b_p = (1-p)\delta_0 + p\delta_1$), and $\boxplus$ is the free additive convolution operation defined in Sec. IV D.

Moreover, the convergence also holds strongly, in the sense of Ref. 61. *In particular, the extremal eigenvalues of M_i converge almost surely to the edges of the support of the measure from* (27).

Proof. The result follows from the large d, n limit of the formula (23). We shall study the terms which contribute asymptotically and then we shall identify the limiting probability distribution with the help of its free cumulants.

To this end, we recall that the Weingarten function $\mathrm{Wg}(nk, \alpha^{-1}\beta)$ may be approximated to second order by $(nk)^{-p-|\alpha^{-1}\beta|}\mathrm{Möb}(\alpha^{-1}\beta)$, for permutations $\alpha, \beta \in \mathcal{S}_p$. Consequently, the moments behave as

$$d^{-1}\mathbb{E}\,\mathrm{Tr}\,M_i^p \sim (ntk)^{-1}\sum_{\alpha,\beta\in\mathcal{S}_p} n^{\#\alpha}(ntk)^{\#(\beta\gamma^{-1})}(nk)^{-p-|\alpha^{-1}\beta|}\mathrm{Möb}(\alpha,\beta).$$

Above, the only non-vanishing terms, as $d, n \to \infty$, are the ones containing the largest power of n. A straightforward analysis shows that

$$\text{power of } n = -1 + \#\alpha + \#(\beta\gamma^{-1}) - p - |\alpha^{-1}\beta| = p - 1 - (|\alpha| + |\alpha^{-1}\beta| + |\beta^{-1}\gamma|) \le 0,$$

where we have used the relation $|\alpha| = p - \#\alpha$ and the triangle inequality

$$|\alpha| + |\alpha^{-1}\beta| + |\beta^{-1}\gamma| \ge |\gamma| = p - 1.$$

The above inequality is saturated if and only if the both α and β lay on the *geodesic* between the identity permutation and the full cycle γ; we write $\mathrm{id} - \alpha - \beta - \gamma$. Here, the notion of geodesic is in relation to the following distance function on the symmetric group $\mathcal{S}_p$:

$$\mathrm{dist}(\sigma,\pi) := |\sigma^{-1}\pi|.$$

We say that a permutaion χ lies on the geodesic between σ and π if χ saturates the triangle inequality,

$$\mathrm{dist}(\sigma,\chi) + \mathrm{dist}(\chi,\pi) \ge \mathrm{dist}(\sigma,\pi).$$

Hence, we obtain the asymptotic moments

$$\lim_{n\to\infty} d^{-1}\mathbb{E}\,\mathrm{Tr}\,M_i^p = \sum_{id-\alpha-\beta-\gamma} t^{-1+\#(\beta^{-1}\gamma)}k^{-1-|\beta^{-1}\gamma|-|\alpha^{-1}\beta|}\mathrm{Möb}(\alpha,\beta).$$

Using the fact that, for geodesic permutations α, β, $-1 - |\alpha^{-1}\beta| - |\beta^{-1}\gamma| = -p + |\alpha| = -\#\alpha$, the equation above may be rewritten as

$$\sum_{id-\alpha-\beta-\gamma} t^{-1+\#(\beta^{-1}\gamma)}k^{-\#(\alpha)}\mathrm{Möb}(\alpha,\beta) = \sum_{id-\alpha-\beta-\gamma} t^{p-\#(\beta)}k^{-\#(\alpha)}\mathrm{Möb}(\alpha,\beta). \tag{29}$$

We now fix $\beta \in \mathcal{S}_p$ and use the moment-cumulant formula[48] in free probability to write

$$\sum_{\mathrm{id}-\alpha-\beta} k^{-\#(\alpha)}\mathrm{Möb}(\alpha,\beta) = \sum_{id-\alpha-\beta} m_p(b_{k^{-1}})\mathrm{Möb}(\alpha,\beta) = \mathcal{K}_\beta(b_{k^{-1}}),$$

where $b_{k^{-1}}$ is the Bernoulli distribution $b_{k^{-1}} = (1-k^{-1})\delta_0 + k^{-1}\delta_1$.

Therefore, Eq. (29) becomes

$$\sum_{id-\beta-\gamma} t^{p-\#(\beta)}\mathcal{K}_\beta(b_{k^{-1}}) = t^p \sum_{id-\beta-\gamma} t^{-\#(\beta)}\mathcal{K}_\beta(b_{k^{-1}}) = t^p m_p(b_{k^{-1}}^{\boxplus t^{-1}}) = m_p\Big(D_t\big[b_{k^{-1}}^{\boxplus t^{-1}}\big]\Big),$$

proving the first claim. In the following, we aim to express the distribution $D_t[b_{k^{-1}}^{\boxplus t^{-1}}]$ in the form presented in the statement of the theorem. Indeed, using Proposition IV.6, we get that

$$D_t\big[b_{k^{-1}}^{\boxplus t^{-1}}\big] = \{x \mapsto tx\}_\#(b_{k^{-1}}^{\boxplus t^{-1}}) = \max(0, 1 - t^{-1}k^{-1})\delta_0 + \max(0, 1 - t^{-1}(1 - k^{-1}))\delta_1 + \frac{1/t\sqrt{(\gamma^+ - \frac{x}{t})(\frac{x}{t} - \gamma^-)}}{2\pi\frac{x}{t}(\frac{1}{t} - \frac{x}{t})}\mathbf{1}_{[\gamma^-,\gamma^+]}(\frac{x}{t})\frac{\mathrm{d}x}{t},$$

where $\gamma^\pm(1/k, 1/t) = (\frac{1}{t} - 2)\frac{1}{k} + 1 \pm 2\sqrt{(\frac{1}{t} - 1)\frac{1}{k}(1 - \frac{1}{k})}$. By denoting $t\gamma^\pm(1/k, 1/t) = \varphi^\pm(1/k, t)$, we obtain the result announced in (27).

The strong convergence follows from the strong asymptotic freeness results of Collins and Male[61] (Theorem 1.4) applied to the Haar-distributed random unitary matrices U_n and a sequence of deterministic projections. □

Remark VI.3. For $t = 1$, the measure in the theorem is the Bernoulli measure $b_{k^{-1}}$.

Remark VI.4. Since the probability distribution (27) can have Dirac masses at 0 or 1 (never at both end points), its support may be non-convex. This happens whenever one or the other Dirac mass is present, that is, when $t < 1/k$ (Dirac mass at 0) or when $t > 1 - 1/k$ (Dirac mass at 1).

Remark VI.5. In light of the results from Ref. 62, the distribution above is equal to the free multiplicative convolution of two Bernoulli distributions of parameters $1/k$ and t, respectively. We do not discuss this equivalent point of view here.

We now present some immediate consequences of the theorem above. These results are about quantities of interest in quantum information theory, such as regularity or the norm-1 property; we refer the reader to Sec. II C for the definitions.

Proposition VI.6. In the asymptotical regime where k is fixed and $d, n \to \infty$ in such a way that $d \sim tkn$ for some constant $t \in (0, 1]$, the first two limiting moments of the random effects M_i read

$$\lim_{n\to\infty} \frac{1}{d}\mathbb{E}\mathrm{Tr}[M_i] = \frac{1}{k},$$
$$\lim_{n\to\infty} \frac{1}{d}\mathbb{E}\mathrm{Tr}\big[M_i^2\big] = \frac{tk + 1 - t}{k^2},$$

while the asymptotic covariance of two different effects ($i \neq j$) behaves like

$$\lim_{n\to\infty} \frac{1}{d}\mathbb{E}\mathrm{Tr}[M_iM_j] = \frac{1 - t}{k^2}.$$

Proof. The first two formulas follow either from Proposition VI.2 for $p = 1, 2$ or from taking the limit in (24) and (25). The covariance formula follows from Eq. (26). □

Proposition VI.7. In the asymptotical regime where k is fixed and $d, n \to \infty$ in such a way that $d \sim tkn$ for some constant $t \in (0, 1]$, a random POVM M is regular (see Definition II.7) iff

$$t \in \left(\frac{1}{2} - \frac{2\sqrt{k-1}}{k}, \frac{1}{2} + \frac{2\sqrt{k-1}}{k}\right).$$

Proof. The condition from the statement is equivalent to asking that 1/2 is not an element of the support of the limiting spectral distribution of the random effects (27). □

Proposition VI.8. In the asymptotical regime where k is fixed and $d, n \to \infty$ in such a way that $d \sim tkn$ for some constant $t \in (0, 1]$, a random POVM M has the norm-1 property (see Definition II.8) iff $t > 1 - 1/k$.

Proof. This follows from (27), by asking that the weight of the Dirac mass δ_1 is positive. □

We display Monte Carlo simulations of a Haar-random POVM element, together with the theoretical curve from the theorem above in Fig. 6. Different statistical properties of these POVM elements will be analyzed in Sec. VII.

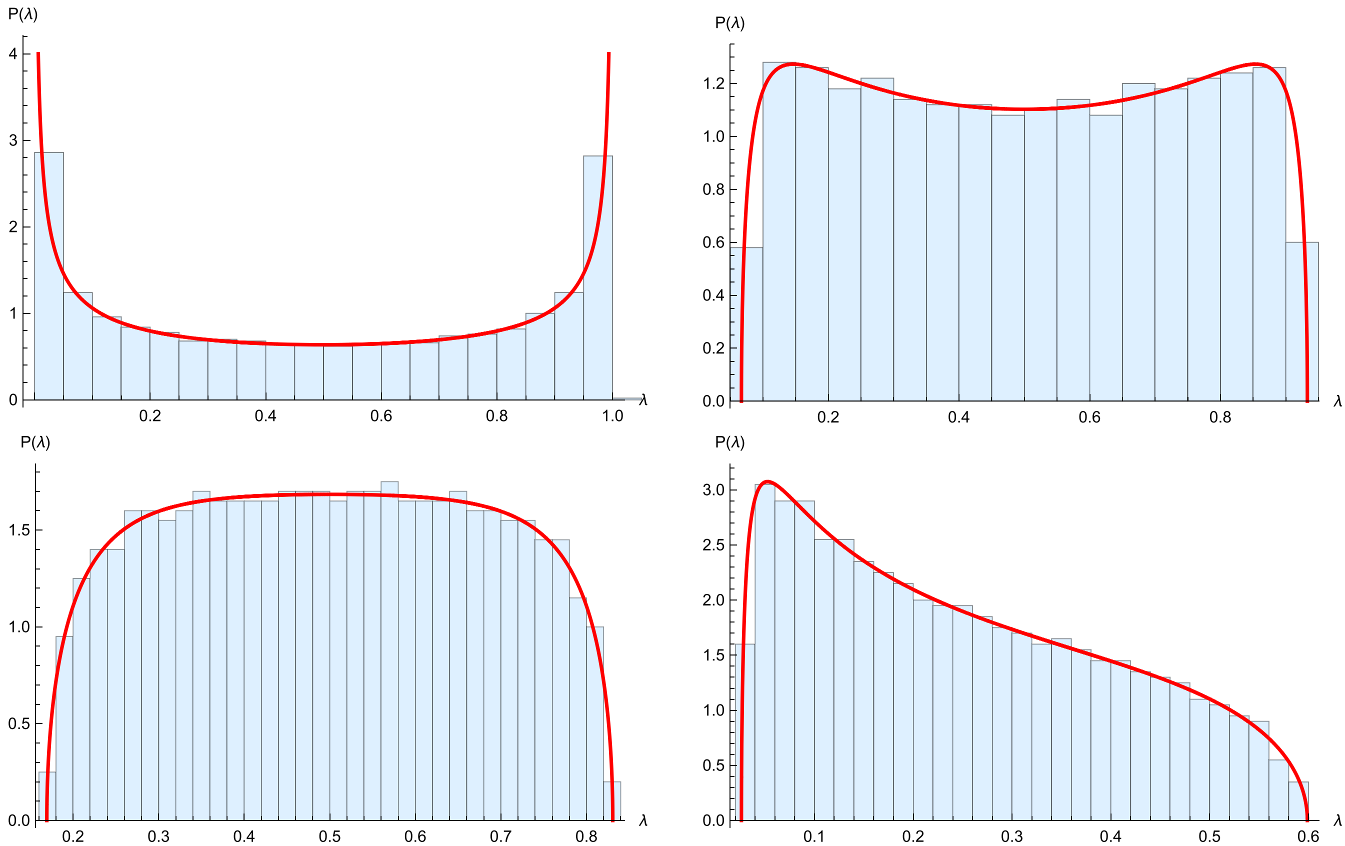

FIG. 6. Monte Carlo simulations vs theoretical curves for the eigenvalues of Haar-POVM elements with the following $(d, k; n)$ triples: top-left (1000, 2; 1000), top-right (1000, 2; 2000), bottom-left (1000, 2; 4000), and bottom-right (1000, 4; 2000). Since the first three examples are dichotomic POVMs, the plots are symmetric with respect to $x = 1/2$. The histogram from each plot corresponds to the eigenvalues of a single sample.

C. The probability range of random POVMs

We now discuss the probability range of random POVMs. Since there is a close connection between the probability range and the output set of unital, completely positive maps, we shall use the results from Ref. 63 in the latter setting to obtain a characterization of the asymptotic probability range in the large dimension limit. Before we do this, let us provide a heuristic argumentation for Theorem VI.9. Consider a random quantum channel

$$\Psi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C}), \quad \Psi(\rho) = [\mathrm{id}_k \otimes \mathrm{Tr}_n](V\rho V^*),$$

where $V : \mathbb{C}^d \to \mathbb{C}^k \otimes \mathbb{C}^n$ is a Haar-distributed random isometry. We know from Sec. V that a random POVM has effects $M_i = \Psi^*(|i\rangle\langle i|)$, where Ψ^* is the Hilbert–Schmidt adjoint of Ψ. Using this duality, we have

$$[\mathrm{Tr}(\rho M_i)]_{i=1}^k = [\mathrm{Tr}(\rho\Psi^*(|i\rangle\langle i|))]_{i=1}^k = [\mathrm{Tr}(\Psi(\rho)|i\rangle\langle i|)]_{i=1}^k = \mathrm{diag}\,\Psi(\rho). \tag{30}$$

First, note that, given an arbitrary fixed pure quantum state $|\psi\rangle \in \mathbb{C}^d$, the distribution of the (random) probability vector

$$(\langle\psi|M_1|\psi\rangle, \dots, \langle\psi|M_k|\psi\rangle) \in \Delta_k$$

is the *Dirichlet distribution* of parameter n,

$$\mathrm{Dir}_k^{(n)}(p_1, \dots, p_k) \sim p_1^{n-1} p_2^{n-1} \cdots p_k^{n-1}.$$

Indeed, this follow from (30) and the fact that the diagonal of a random density matrix from the induced ensemble of parameters (k, n) is $\mathrm{Dir}_k^{(n)}$; see Ref. 64 (Sec. VIII).

Moreover, the probability range of a random POVM is related to the diagonals of the output set of a random quantum channel. In order to state and prove the main theorem, let us recall the definition of the (t)-*norm* from Ref. 63. To any vector $x \in \mathbb{R}^k$ associate a self-adjoint element in the non-commutative probability space $(\mathbb{C}^k, \mathrm{tr})$, where we denote by $\mathrm{tr} := \frac{1}{k}\mathrm{Tr}[\cdot]$ the normalized trace. Consider also the projection p of trace $t \in (0,1)$ living in the non-commutative probability space $(\mathbb{C}^2, \mathrm{tr})$. We define the (t)-norm of x as

$$\|x\|_{(t)} := \|pxp\|,$$

where the elements in the right-hand side live in the *free product* of the two non-commutative probability spaces mentioned above. Moreover, let us define the set

$$K_{k,t} := \{\lambda \in \Delta_k \,:\, \forall a \in \Delta_k,\ \langle \lambda, a\rangle \leq \|a\|_{(t)}\}.$$

Theorem VI.9. *Consider a sequence $(M^{(n)})_n$ of k-valued random POVMs, with effects $M_i^{(n)} \in \mathcal{M}_{d_n}(\mathbb{C})$. The effect dimensions scale as $d_n \sim tkn$, for some constant $t \in (0,1)$. Almost surely, the probability ranges of the random POVMs $M^{(n)}$ converge to the deterministic convex set $K_{k,t}$, in the following sense:*

$$K_{k,t}^{\circ} \subseteq \liminf_{n\to\infty} \mathrm{ProbRan}(M^{(n)}) \subseteq \limsup_{n\to\infty} \mathrm{ProbRan}(M^{(n)}) \subseteq K_{k,t}.$$

Proof. The result for the output sets of the random quantum channels $\Psi^{(n)}$ is Ref. 65 (Theorem 6.2), which in turn builds on Ref. 63 (Theorem 5.4). Restricting to diagonals obviously preserves the upper bound, by the Schur–Horn theorem: for any Hermitian matrix A, $\mathrm{diag}(A) \prec \mathrm{spec}(A)$, where $\prec$ denotes the majorization relation; see Ref. 66 (Exercise II.1.12). For the lower bound, note that Ref. 65 (Theorem 6.2) is stated at the level of matrices: any self-adjoint matrix having its spectrum in the interior of $K_{k,t}$ will eventually be in the output set of $\Psi^{(n)}$; in particular, this holds for diagonal matrices. □

In general, computing the (t)-norm of vectors in $\mathbb{R}^k$ requires solving polynomial equations of high degree. The only analytical result in a closed form is the value of the (t)-norm for bi-valued vectors. First, note that for non-negative reals $0 \leq a \leq b$, we have

$$\|(a, a, \ldots, a, b, b, \ldots, b\|_{(t)} = a + \|(0, 0, \ldots, 0, b-a, b-a, \ldots, b-a\|_{(t)}.$$

Then, it follows from Ref. 63 (Proposition 3.6) that

$$\|(\underbrace{1, 1, \ldots, 1}_{j\text{ times}}, \underbrace{0, 0, \ldots, 0}_{k-j\text{ times}})\|_{(t)} = \begin{cases} t + u - 2tu + 2\sqrt{tu(1-t)(1-u)} & \text{if } t + u < 1 \\ 1 & \text{if } t + u \geq 1, \end{cases}$$

where $u = j/k \in [0,1]$. We have, thus, a complete picture of the asymptotic probability range for $k = 2$: $K_{2,t} = \{(p, 1-p) : |p - 1/2| \leq x_t\}$, with

$$x_t = \begin{cases} \sqrt{t(1-t)}, & \text{if } t \leq 1/2 \\ 1/2, & \text{if } t > 1/2. \end{cases}$$

VII. (IN-)COMPATIBILITY CRITERIA FOR RANDOM POVMs

Having developed in Secs. V and VI the theory of random POVMs, we turn in this section to the question of compatibility of generic POVMs. The fundamental question here is the following:

Given two independent random POVMs, what is the probability that they are compatible?

More precisely, the two random POVMs are chosen independently from the Haar ensembles with parameters $(d_i, k_i; n_i)$, respectively $(i = 1, 2)$; we assume obviously that $d_1 = d_2$. Since compatibility of random POVMs can be formulated as a semidefinite program, the considerations in this section could also be seen as giving bounds for the existence of solutions of random SDPs (see, e.g., Ref. 67).

As it is often the case in random matrix theory, we shall focus on the asymptotic regime where the Hilbert space size $d = d_1 = d_2$ grows to infinity. We shall keep the number of effects $k_{1,2}$ in the POVMs constant, and the respective parameters $n_{1,2}$ will follow linear scalings with respect to d; this is precisely the asymptotical regime studied in Proposition VI.2.

Sections VII A, VII B, and VII C deal with *compatibility criteria* that is sufficient conditions for compatibility. Sections VII D and VII E are focused on *incompatibility criteria*, i.e., necessary conditions for compatibility; it turns out that the two such criteria we discuss are not informative in the asymptotical regime we investigate. Finally, we compare the noise content and the Jordan product criteria in Sec. VII F.

A. The noise content criterion

We analyze in this section the *noise content criterion*, stated in Proposition III.1, when applied to Haar-random POVMs.

We know from Proposition VI.2 that, for a Haar-random POVM A with parameters $(d, k; n)$, in the asymptotic regime where k is fixed and $d, n \to \infty$ in such a way that $d \sim skn$ for some constant $s \in (0, 1]$, the smallest eigenvalue of some POVM element A_i converges, almost surely, to the constant φ_- from (28),

$$\varphi_-(s,k) = \begin{cases} s + \dfrac{1-2s}{k} - \dfrac{2}{k}\sqrt{s(1-s)(k-1)}, & \text{if } s < \dfrac{1}{k} \\ 0, & \text{if } s \geq \dfrac{1}{k}. \end{cases}$$

The formula above allows us to obtain the limiting noise content of random POVMs; for a sequence of random POVMs of parameters $(d, k; n_d)$, where n_d is a sequence of integers with the property that $d \sim skn_d$ (as $d \to \infty$), the noise content $w(M) = \sum_{i=1}^k \lambda_{\min}(M_i)$ converges, almost surely as $d \to \infty$ and k, s fixed, to the quantity $k\varphi_-(s, k)$.

Using this result, we obtain the following compatibility criterion for Haar-random POVMs.

Theorem VII.1. *Let $(A^{(d)})$ and $(B^{(d)})$ be two sequences of random POVMs of respective parameters $(d, k; n_d)$ and $(d, l; m_d)$, where n_d and m_d are two integer sequences growing to infinity in such a way that $d \sim skn_d$ and $d \sim tlm_d$ for two constants $s, t \in (0, 1]$. If*

$$k\varphi_-(s,k) + l\varphi_-(t,l) > 1, \tag{31}$$

then, almost surely as $d \to \infty$, the Haar-random POVMs $A^{(d)}$ and $B^{(d)}$ are asymptotically compatible.

Proof. From Proposition VI.2, we know that for individual POVM operators $A_i^{(d)}$ (respectively, $B_j^{(d)}$), the minimum eigenvalue converges, almost surely as $d \to \infty$, to the corresponding value φ_- (here, we need the strong convergence flavor of the theorem). Taking the intersection of $k + l$ almost sure events, we obtain the simultaneous almost sure convergence of the sum of minimum eigenvalues to the left-hand-side of (31). The conclusion follows from a standard countable approximation argument. □

Remark VII.2. Note that in Theorem VII.1, we do not need to make any assumptions on the joint distribution of the random POVMs A and B (such as independence). This is due to the fact that the minimum eigenvalue compatibility criterion we are using only depends on individual spectral characteristics of the two POVMs.

Corollary VII.3. In the case $k = l \geq 2$ and $s = t$ (identically distributed Haar-random POVMs), the condition from (31) simplifies to

$$s < \frac{1}{6k - 4 + 4\sqrt{(k-1)(2k-1)}}.$$

Corollary VII.4. In the case $k = l = 2$ and s, t arbitrary (dichotomic POVMs), the condition from (31) simplifies to

$$t < \frac{1}{2} - \sqrt{\sqrt{s(1-s)} - s(1-s)} \quad \text{and} \quad s < \frac{1}{2}.$$

The condition $s < \frac{1}{2}$ appears because one needs to take the first branch of the definition of the function φ_- in order to satisfy (31). The inequalities of Corollaries VII.3 and VII.4 are depicted in Fig. 8.

B. The Jordan product criterion

In this section, we focus on the compatibility criterion given by the *Jordan product*; see Proposition III.2. To apply this criterion to Haar-random POVMs A and B, one has to compute the minimum eigenvalue of the Jordan product $A_i \circ B_j$ of two (independent) random matrices having limiting eigenvalue distributions such as in Proposition VI.2. The computation of the distribution of the anti-commutator of two random matrices is an important problem in the general theory of random matrices, which has received some attention in the last years, especially in the framework of free probability.[68,69] A nice description of the anti-commutator of a pair of free random variables remains elusive in the most general case, despite some partial results [e.g., for even random variables, see Ref. 68 (Proposition 1.10)] and some implicit characterizations [see Ref. 69 (Theorem 2.2)].

In the absence of an analytical description of the smallest eigenvalue of the Jordan product of two random POVM elements, we rely here on the following general lower bound. For a positive definite matrix X, we denote

$$R(X) := \frac{\lambda_{\max}(X)}{\lambda_{\min}(X)} \in [1, \infty).$$

Lemma VII.5 (Refs. 70–72). *Let $X, Y \in \mathcal{M}_d(\mathbb{C})$ be two positive definite matrices. If any of the two equivalent conditions below holds*

- $(\sqrt{R(X)} - 1)(\sqrt{R(Y)} - 1) < 2$ and
- $(R(X) - 1)^2(R(Y) - 1)^2 < 16R(X)R(Y)$,

then $Z = X \circ Y = XY + YX$ is a positive definite.

The next result uses the previous lemma to provide a sufficient criterion for the asymptotic compatibility of Haar-random POVMs. We omit the proof since it is very similar to the Proof of Theorem VII.1. We need the following notation ($k \ge 2$ and $0 < s \le 1$):

$$R(k,s) := \begin{cases} \dfrac{s + \frac{1-2s}{k} + \frac{2}{k}\sqrt{s(1-s)(k-1)}}{s + \frac{1-2s}{k} - \frac{2}{k}\sqrt{s(1-s)(k-1)}}, & \text{if } s < \dfrac{1}{k} \\ +\infty, & \text{if } s \ge \dfrac{1}{k}. \end{cases}$$

Theorem VII.6. *Let $(A^{(d)})$ and $(B^{(d)})$ be two sequences of random POVMs of respective parameters $(d, k; n_d)$ and $(d, l; m_d)$, where n_d and m_d are two integer sequences growing to infinity in such a way that $d \sim skn_d$ and $d \sim tlm_d$ for two constants $s, t \in (0, 1]$. If*

$$\left(\sqrt{R(k,s)} - 1\right)\left(\sqrt{R(l,t)} - 1\right) < 2, \tag{32}$$

then, almost surely as $d \to \infty$, the Haar-random POVMs $A^{(d)}$ and $B^{(d)}$ are asymptotically compatible.

Remark VII.7. As for Theorem VII.1, we do not need to make any assumptions on the joint distribution of the random POVMs A and B. Although the Jordan product compatibility criterion depends jointly and in a non-trivial manner on the POVM elements of both A and B, the inequality from Lemma VII.5 separates these contributions, allowing for the very general bound (32).

Corollary VII.8. In the case $k = l \ge 2$ and $s = t$ (identically distributed Haar-random POVMs), the condition from (32) simplifies to $R(k,s) < 3 + 2\sqrt{2}$, which, after some algebra, yields

$$s < \frac{k(3 - 2\sqrt{2}) + 2(\sqrt{2} - 1)}{k^2 + 4k - 4} < \frac{1}{k}.$$

Corollary VII.9. In the case $k = l = 2$ and s, t arbitrary (dichotomic POVMs), the condition from (32) simplifies to

$$\sqrt{s(1-s)} + \sqrt{t(1-t)} < \frac{1}{4} \iff t < \frac{1}{2} - \sqrt{s(1-s)} \quad \text{and} \quad s < \frac{1}{2}.$$

The inequalities of Corollaries VII.8 and VII.9 are depicted in Fig. 8.

C. The optimal cloning map criterion

We briefly discuss here the optimal cloning compatibility criterion presented in Proposition III.3 for Haar-random POVMs. The relevant quantities here are the minimal eigenvalues of the effects (which were discussed at length in Proposition VI.2 and used in Theorem VII.1) and the traces of the effects. Regarding the latter quantities, we know from Proposition VI.2 that, almost surely as $d \to \infty$,

$$\forall i\ :\ \lim \frac{\mathrm{Tr}\left[A_i^{(d)}\right]}{d} = \frac{1}{k}$$

for a sequence of random Haar-POVMs $A^{(d)}$ with parameters $(d, k; n_d)$ in the scaling $d \sim skn_d$. It follows that the asymptotical version of Eq. (8) reads

$$\varphi_-(s,k) > \frac{1}{2k}.$$

Assuming also the corresponding condition for a second sequence of random Haar-POVMs $B^{(d)}$ with parameters $(d, l; m_d)$, we recover by summing them [Eq. (31)], showing that, asymptotically, for Haar-random POVMs, the optimal cloning criterion is weaker that the noise content criterion. Note, however, that this is not the case at fixed dimension d as it was pointed out in Remark III.4.

D. Unsharpness and the Miyadera–Imai criterion

Our goal in this section is to analyze under which conditions independent random POVMs are certified incompatible by the Miyadera–Imai criterion recalled in Sec. III C.

Since the sharpness measure from Definition II.6 plays an important role in the Miyadera–Imai criterion, let us study it in the case of the random POVMs defined in Sec. V.

Proposition VII.10. Let $(A^{(d)})$ *be a sequence of random POVMs of parameters* $(d, k; n_d)$*, where* n_d *is an integer sequence growing to infinity in such a way that* $d \sim skn_d$ *for a constant* $s \in (0, 1]$*. Then, almost surely, for all* $i = 1, \ldots, k$*,*

$$\lim_{d\to\infty} \sigma(M_i^{(d)}) = \sigma(k,s) := \begin{cases} 4\varphi_+(s,k^{-1})(1-\varphi_+(s,k^{-1})), & \text{if } s \in [0, s_0) \\ 1, & \text{if } s \in [s_0, 1-s_0] \\ 4\varphi_-(s,k^{-1})(1-\varphi_-(s,k^{-1})), & \text{if } s \in (1-s_0, 1) \\ 0, & \text{if } s = 1, \end{cases} \tag{33}$$

where $\sigma(.)$ *is the sharpness measure from (3),* $\varphi_\pm$ *are the constants defined in (28), and*

$$s_0 = \frac{1}{2} - \frac{\sqrt{k-1}}{k} \in [0, 1/2).$$

Proof. The result follows from a simple analysis of the support of the measure (27). □

Remark VII.11. It is easy to see that the limiting value $\sigma(k, s)$ *is symmetric with respect to* $s = 1/2$*:* $\sigma(k, s) = \sigma(k, 1 - s)$*. Moreover, for all* k *and* $s \in (0, 1)$*,* $\sigma(k, s) \geq k^{-1} - k^{-2}$*. At* $s = 1$*, the random POVM elements* $M_1, \ldots, M_k$ *are random projections summing to the identity, hence, the unsharpness is null.*

Remark VII.12. For $k = 2$*, we have* $s_0 = 0$ *and thus, for all* $s \in [0, 1]$*,* $\sigma(2, s) = 1$*. This is because* $1/2$ *is, asymptotically, almost surely an element of the spectrum of both effects of a binary random POVM.*

In Fig. 7, we plot the limiting value of the unsharpness $\sigma(A_i)$ as a function of s, for fixed k.

Regarding now the application of the Miyadera–Imai criterion, since this is only a necessary condition for compatibility of POVMs, the only scenario in which it can be used is if

$$4\|[A_i, B_j]\|^2 > \sigma(A_i)\sigma(B_j), \tag{34}$$

in which case the POVMs are guaranteed to be incompatible. Above, A_i and B_j are quantum effects belonging to two POVMs A and B. In the case of random POVMs A and B, the difficulty lies in computing the left-hand side. For example, in the most natural setting, when $A^{(d)}$ and $B^{(d)}$ are sequences of independent, identically distributed Haar-random POVMs as in Proposition VII.10, one needs to compute the limiting eigenvalue distribution of the random matrix $A_iB_j - B_jA_i$ in a strong sense (in order to also obtain the convergence of the operator norm). As it was argued in Sec. VII B, the computations of the limiting distributions of commutators and anti-commutators is a highly

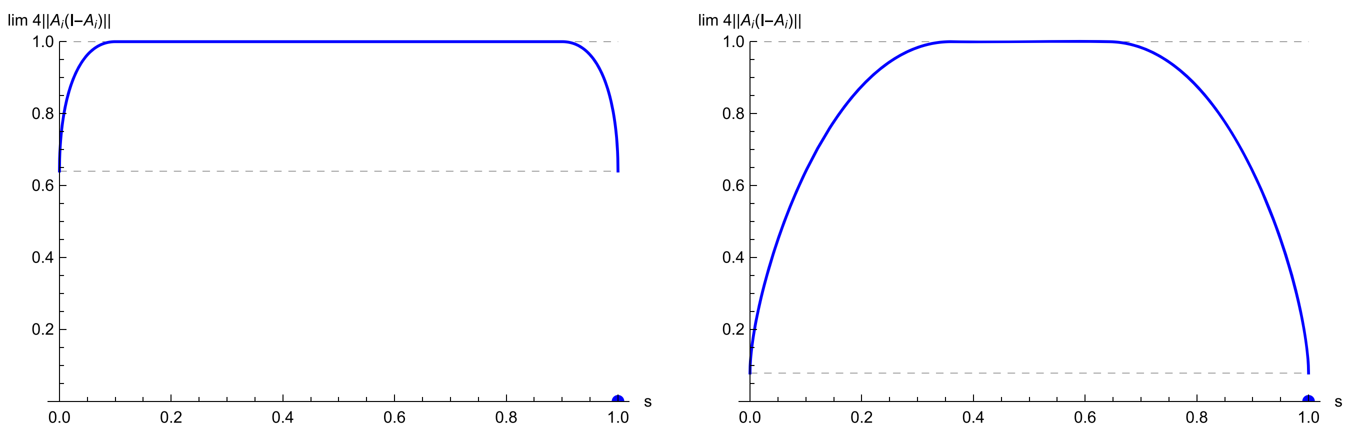

FIG. 7. The limiting value of the unsharpness of a random POVM as a function of s, for $k = 5$ (left) and $k = 50$ (right). The dotted horizontal lines correspond to the minimal ($4(k^{-1} - k^{-2})$) and maximal (1) values of the unsharpness, corresponding to a fixed value of k.

TABLE I. The average value of $4\|[A_i, B_j]\|^2$ for 10 pairs of independent quantum effects A_i, B_j from the Haar-random POVM ensemble of parameters $(d = \lfloor skn \rfloor, k; n = 1000)$ for $s = 0.1, 0.3, 0.5, 0.7, 0.9$ and $k = 2, 3, 5$. In all these cases, the right-hand side of (34) is equal to 1, asymptotically, so the application of the Miyadera–Imai criterion is inconclusive.

k (s)	0.1	0.3	0.5	0.7	0.9
2	0.025 522	0.199 76	0.472 05	0.756 35	0.963
3	0.020 689	0.170 55	0.426 74	0.722 26	0.957 12
5	0.011 771	0.112 01	0.317 46	0.603 86	0.914 34

non-trivial question in free probability, so we lack a precise answer in our setting. More recent theoretical results based on the theory of operator-valued free probability[73] might be the right framework to tackle such questions; we leave this question open. Numerical simulations[74] seem to suggest, however, that the Miyadera–Imai criterion does not allow to conclude that Haar-random POVMs are incompatible; see Table I.

E. The Zhu criterion

The Zhu criterion from Proposition III.6 is unfortunately uninformative in the asymptotical regime we are interested in. Indeed, using the triangle inequality, we upper bound the expression of τ from (15) by

$$\frac{1}{2}\left[\mathrm{Tr}[\mathcal{G}_A] + \mathrm{Tr}[\mathcal{G}_B] + \|\mathcal{G}_A\|_1 + \|\mathcal{G}_B\|_1\right] = \mathrm{Tr}[\mathcal{G}_A] + \mathrm{Tr}[\mathcal{G}_B].$$

Since $\mathrm{Tr}[\mathcal{G}_A] = \sum_{i=1}^{k} \mathrm{Tr}\left[A_i^2\right]/\mathrm{Tr}[A_i] \leq k$, to obtain a violation of the inequality from Proposition III.6, one needs $k + l > d$, where k, respectively, l is the number of outcomes of the POVMs A, respectivel, B. In the regime we are interested in (fixed number of outcomes, large dimension), this inequality cannot hold, so the Zhu criterion is not applicable in our setting.

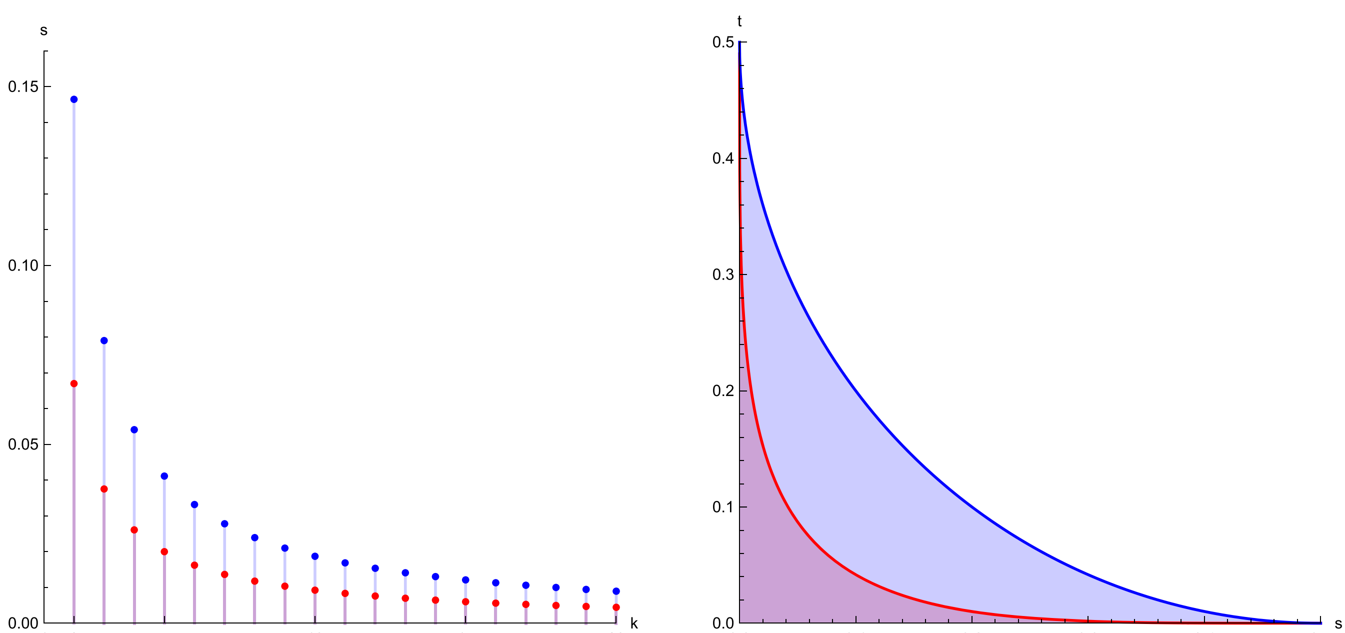

FIG. 8. Efficiency of the noise content and Jordan product criteria. Left: the range of values (k, s) for which one can infer the compatibility of two identically distributed Haar-random POVMs (see Corollaries VII.3 and VII.8). Right: the range of values (s, t) for which one can infer the compatibility of two dichotomic Haar-random POVMs (see Corollaries VII.4 and VII.9). The top curves (in blue) correspond to the Jordan product criterion, while the bottom ones (in red) to the noise content criterion.

F. Comparing the different compatibility criteria

We compare in this section the different compatibility criteria described previously. As it was noted, the optimal cloning criterion is asymptotically weaker than the noise content criterion, so we do not discuss it here. The two relevant criteria are the Jordan product criterion and the noise content criterion. We compare them in Fig. 8, and we note that the Jordan product criterion performs systematically better. This result is surprising since we have used several inequalities in our analysis from Sec. VII B in order to be able to apply the criterion to random matrices.

We conclude that, in the presence of typical random POVMs, one has interest in checking first the Jordan product criterion in order to certify compatibility.

We would like also to point out that, at this time, we do not know of any incompatibility criteria that would give any insightful information in the asymptotic regime studied in this paper (fixed number of outcomes, large matrix dimension). It would be interesting to develop such criteria, adapted to noisy POVMs.

ACKNOWLEDGMENTS

I.N's research was supported by ANR Project Nos. StoQ ANR-14-CE25-0003-01 and NEXT ANR-10-LABX-0037-NEXT. I.N. acknowledges the hospitality of the Universitatea Politehnica Timişoara, and Turku. T.H. acknowledges financial support from the Academy of Finland via the Centre of Excellence Program (Project No. 312058 as well as Project No. 287750).

APPENDIX: DENSITY OF WISHART-RANDOM POVMs

We prove in this appendix Theorem V.13 with the help of the matrix Dirac delta functions.[56,57] Let us recall that a Wishart-random POVM M of parameters $(d, k; s_1, \ldots, s_k)$ is obtained by normalizing k independent Wishart matrices $W_1, \ldots, W_k$ of respective parameters (d, s_i),

$$M_i = S^{-1/2} W_i S^{-1/2}, \qquad \text{where } S = \sum_{j=1}^{k} W_j.$$

We also recall that a Wishart random matrix of parameters (d, s) has density

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{dLeb}}(w) = C_{d,s} \mathbf{1}_{w \geq 0} \exp(-\operatorname{Tr} w)(\det w)^{s-d}.$$

Proof of Theorem V.13. In the course of the proof, we shall not keep track of constants although this could be done with the help of the Weyl integration formula[75] (Proposition 4.1.3) and the Selberg integral[76] [Eq. (17.6.5)]. We have

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}}{\mathrm{dLeb}}(m_1, \ldots, m_k) &\sim \int \prod_{i=1}^{k} \mathrm{d}W_i \mathbf{1}_{W_i \geq 0} \exp(-\operatorname{Tr} W_i)(\det W_i)^{s_i - d} \\ &\qquad \cdot \delta\left[m_i - \left(\sum_j W_j\right)^{-1/2} W_i \left(\sum_j W_j\right)^{-1/2}\right] \\ &= \int \mathrm{d}S \left(\prod_{i=1}^{k} \mathrm{d}W_i\right) \delta\left(S - \sum_j W_j\right) \prod_{i=1}^{k} \mathbf{1}_{W_i \geq 0} \exp(-\operatorname{Tr} W_i)(\det W_i)^{s_i - d} \delta(m_i - S^{-1/2} W_i S^{-1/2}).\end{aligned}$$

We shall now make a change of variables $W_i = S^{1/2} m_i^{1/2} Y_i m_i^{1/2} S^{1/2}$, where S and m_i are treated like constants and Y_i are the new variables. Computing the Jacobian of this transformation [see also Ref. 77 (Proposition 3.7)], we have

$$\mathrm{d}W_i = (\det S)^d (\det m_i)^d \mathrm{d}Y_i.$$

Factorizing the expressions appearing in the delta functions and using[57] Proposition 3.3, we get

$$\begin{aligned}\delta\left[m_i - \left(\sum_j W_j\right)^{-1/2} W_i \left(\sum_j W_j\right)^{-1/2}\right] &= (\det m_i)^{-d} \delta(I_d - Y_i) \\ \delta\left(S - \sum_j W_j\right) &= (\det S)^{-d} \delta\left(I_d - \sum_j m_j^{1/2} Y_j m_j^{1/2}\right).\end{aligned}$$

Plugging everything into the expression for the density, we obtain

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{dLeb}}(m_1,\ldots,m_k) \sim \int \mathrm{d}S\left(\prod_{i=1}^k \mathrm{d}Y_i\right)\mathbf{1}_{S\geq 0}\,\exp(-\operatorname{Tr}S)(\det S)^{(k-1)d}\delta\left(I_d - \sum_j m_j^{1/2}Y_j m_j^{1/2}\right)$$
$$\cdot\prod_{i=1}^k \mathbf{1}_{Y_i\geq 0}(\det S)^{s_i-d}(\det m_i)^{s_i-d}(\det Y_i)^{s_i-d}\delta(I_d - Y_i)$$
$$= \delta\left(I_d - \sum_j m_j\right)\prod_{i=1}^k(\det m_i)^{s_i-d}\int \mathrm{d}S\mathbf{1}_{S\geq 0}\,\exp(-\operatorname{Tr}S)(\det S)^{\sum_j s_j - d}$$
$$\sim \delta\left(I_d - \sum_j m_j\right)\prod_{i=1}^k(\det m_i)^{s_i-d},$$

which is formula (22), finishing the proof. □

REFERENCES

[1] G. Aubrun, S. J. Szarek, and D. Ye, "Entanglement thresholds for random induced states," Commun. Pure Appl. Math. **67**, 129–171 (2014).
[2] M. Fukuda and I. Nechita, "On the minimum output entropy of random orthogonal quantum channels," IEEE Trans. Inform. Theory **64**, 1374–1384 (2018).
[3] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," Nat. Phys. **5**, 255–257 (2009).
[4] B. Collins and I. Nechita, "Random matrix techniques in quantum information theory," J. Math. Phys. **57**, 015215 (2016).
[5] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement* (Cambridge University Press, 2012).
[6] J. Radhakrishnan, M. Rötteler, and P. Sen, "Random measurement bases, quantum state distinction and applications to the hidden subgroup problem," Algorithmica **55**, 490–516 (2009).
[7] D. Petz and L. Ruppert, "Optimal quantum-state tomography with known parameters," J. Phys. A: Math. Theor. **45**, 085306 (2012).
[8] G. Aubrun and C. Lancien, "Zonoids and sparsification of quantum measurements," Positivity **20**, 1–23 (2016).
[9] L. Zhang, H. Xiang, X. Li-Jost, and S.-M. Fei, "Incompatibility probability of random quantum measurements," Phys. Rev. E **100**, 062139 (2019).
[10] K. W. Wachter, "The limiting empirical measure of multiple discriminant ratios," Ann. Stat. **8**, 937–957 (1980).
[11] G. M. D'Ariano, P. L. Presti, and P. Perinotti, "Classical randomness in quantum measurements," J. Phys. A: Math. Gen. **38**, 5979 (2005).
[12] E. Haapasalo, T. Heinosaari, and J.-P. Pellonpää, "Quantum measurements on finite dimensional systems: Relabeling and mixing," Quantum Inform. Process. **11**, 1751–1763 (2012).
[13] A. S. Holevo, "Statistical decision theory for quantum systems," J. Multivar. Anal. **3**, 337–394 (1973).
[14] P. Busch, G. Cassinelli, and P. J. Lahti, "Probability structures for quantum state spaces," Rev. Math. Phys. **07**, 1105–1121 (1995).
[15] M. Fukuda, I. Nechita, and M. M. Wolf, "Quantum channels with polytopic images and image additivity," IEEE Trans. Inform. Theory **61**, 1851–1859 (2015).
[16] A. Dvurečenskij and S. Pulmannová, "Difference posets, effects, and quantum measurements," Int. J. Theor. Phys. **33**, 819–850 (1994).
[17] P. Lahti and S. Pulmannová, "Coexistent observables and effects in quantum mechanics," Rep. Math. Phys. **39**, 339–351 (1997).
[18] T. Heinonen, P. Lahti, J.-P. Pellonpää, S. Pulmannova, and K. Ylinen, "The norm-1-property of a quantum observable," J. Math. Phys. **44**, 1998–2008 (2003).
[19] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer, 1996).
[20] S. N. Filippov, T. Heinosaari, and L. Leppäjärvi, "Necessary condition for incompatibility of observables in general probabilistic theories," Phys. Rev. A **95**, 032127 (2017).
[21] P. Busch, "Unsharp reality and joint measurements for spin observables," Phys. Rev. D **33**, 2253–2261 (1986).
[22] T. Heinosaari, T. Miyadera, and M. Ziman, "An invitation to quantum incompatibility," J. Phys. A: Math. Theor. **49**, 123001 (2016).
[23] H. Martens and W. M. de Muynck, "Nonideal quantum measurements," Found. Phys. **20**, 255–281 (1990).
[24] T. Heinosaari, D. Reitzner, and P. Stano, "Notes on joint measurability of quantum observables," Found. Phys. **38**, 1133–1147 (2008).
[25] S. T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo, "Commutative POVMs and fuzzy observables," Found. Phys. **39**, 593–612 (2009).
[26] D. Reeb, D. Reitzner, and M. M. Wolf, "Coexistence does not imply joint measurability," J. Phys. A: Math. Theor. **46**, 462002 (2013).
[27] T. Heinosaari, "A simple sufficient condition for the coexistence of quantum effects," J. Phys. A: Math. Theor. **46**, 152002 (2013).
[28] T. Heinosaari, J. Schultz, A. Toigo, and M. Ziman, "Maximally incompatible quantum observables," Phys. Lett. A **378**, 1695–1699 (2014).
[29] M. Keyl and R. F. Werner, "Optimal cloning of pure states, testing single clones," J. Math. Phys. **40**, 3283–3299 (1999).
[30] R. F. Werner, "Optimal cloning of pure states," Phys. Rev. A **58**, 1827 (1998).
[31] T. Miyadera and H. Imai, "Heisenberg's uncertainty principle for simultaneous measurement of positive-operator-valued measures," Phys. Rev. A **78**, 052119 (2008).
[32] H. Zhu, "Information complementarity: A new paradigm for decoding quantum incompatibility," Sci. Rep. **5**, 14317 (2015).
[33] R. D. Gill and S. Massar, "State estimation for large ensembles," Phys. Rev. A **61**, 042312 (2000).
[34] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
[35] C. W. Helstrom, "Quantum detection and estimation theory," J. Stat. Phys. **1**, 231–252 (1969).
[36] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
[37] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables, and Entropy*, Mathematical Surveys and Monographs Vol. 77 (American Mathematical Society, 2000).
[38] P. Hayden and A. Winter, "Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$," Commun. Math. Phys. **284**, 263–280 (2008).
[39] M. Fukuda and C. King, "Entanglement of random subspaces via the hastings bound," J. Math. Phys. **51**, 042201 (2010).
[40] G. Aubrun, S. Szarek, and E. Werner, "Hastings' additivity counterexample via Dvoretzky's theorem," Commun. Math. Phys. **305**, 85–97 (2011).
[41] S. T. Belinschi, B. Collins, and I. Nechita, "Almost one bit violation for the additivity of the minimum output entropy," Commun. Math. Phys. **341**, 885–909 (2016).

[42] D. Weingarten, “Asymptotic behavior of group integrals in the limit of infinite rank,” J. Math. Phys. **19**, 999–1001 (1978).
[43] B. Collins, “Moments and cumulants of polynomial random variables on unitarygroups, the itzykson-zuber integral, and free probability,” Int. Math. Res. Not. **2003**, 953–982.
[44] B. Collins and P. Śniady, “Integration with respect to the haar measure on unitary, orthogonal and symplectic group,” Commun. Math. Phys. **264**, 773–795 (2006).
[45] B. Collins and I. Nechita, “Random quantum channels I: Graphical calculus and the Bell state phenomenon,” Commun. Math. Phys. **297**, 345–370 (2010).
[46] E. Penrose “Applications of negative dimensional tensors,” in *Combinatorial Mathematics and its Applications*, edited by D. Welsh (Academic Press, New York, 1971), pp. 221–244.
[47] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables* (American Mathematical Soc., 1992), Vol. 1.
[48] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability* (Cambridge University Press, 2006), Vol. 13.
[49] J. A. Mingo and R. Speicher, *Free Probability and Random Matrices* (Springer, 2017), Vol. 35.
[50] N. Johnston, QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9, 2016, http://qetlab.com.
[51] Code for the MATLAB routine RandomHaarPOVM.m (as well as for other numerical functions used in this work) can be found in the supplementary material of the arXiv version of this paper.
[52] I. Nechita and C. Pellegrini, “Random repeated quantum interactions and random invariant states,” Probab. Theory Relat. Fields **152**, 299–320 (2012).
[53] K. Zyczkowski and H.-J. Sommers, “Induced measures in the space of mixed quantum states,” J. Phys. A: Math. Gen. **34**, 7111 (2001).
[54] J. Wishart, “The generalised product moment distribution in samples from a normal multivariate population,” Biometrika **20A**, 32–52 (1928).
[55] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices* (Springer, 2010), Vol. 20.
[56] R. F. Hoskins, *Delta Functions: Introduction to Generalised Functions* (Elsevier, 2009).
[57] L. Zhang, “Dirac delta function of matrix argument,” arXiv:1607.02871 (2016).
[58] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (Wiley, New York, 2003).
[59] I. M. Johnstone, “Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy–Widom limits and rates of convergence,” Ann. Stat. **36**, 2638 (2008).
[60] P. J. Forrester, *Log-Gases and Random Matrices (LMS-34)* (Princeton University Press, 2010).
[61] B. Collins and C. Male, “The strong asymptotic freeness of Haar and deterministic matrices,” Ann. Sci. l’Ecole Norm. Supér. **47**, 147–163 (2014).
[62] B. Collins, “Product of random projections, Jacobi ensembles and universality problems arising from free probability,” Probab. Theory Relat. Fields **133**, 315–344 (2005).
[63] S. Belinschi, B. Collins, and I. Nechita, “Eigenvectors and eigenvalues in a random subspace of a tensor product,” Invent. Math. **190**, 647–697 (2012).
[64] Z. Puchała, Ł. Pawela, and K. Życzkowski, “Distinguishability of generic quantum states,” Phys. Rev. A **93**, 062112 (2016).
[65] B. Collins, M. Fukuda, and I. Nechita, “On the convergence of output sets of quantum channels,” J. Oper. Theory **73**, 333–360 (2015).
[66] R. Bhatia, *Matrix Analysis* (Springer Science & Business Media, 1997), Vol. 169.
[67] D. Amelunxen and P. Bürgisser, “Intrinsic volumes of symmetric cones and applications in convex programming,” Math. Program. **149**, 105–130 (2015).
[68] A. Nica and R. Speicher, “Commutators of free random variables,” Duke Math. J. **92**, 553–592 (1998).
[69] V. Vasilchuk, “On the asymptotic distribution of the commutator and anticommutator of random matrices,” J. Math. Phys. **44**, 1882–1908 (2003).
[70] W. G. Strang, “Eigenvalues of Jordan products,” Am. Math. Mon. **69**, 37–40 (1962).
[71] D. W. Nicholson, “Eigenvalue bounds for $AB + BA$, with A, B positive definite matrices,” Linear Algebra Appl. **24**, 173–184 (1979).
[72] N. Alikakos and P. W. Bates, “Estimates for the eigenvalues of the Jordan product of hermitian matrices,” Linear Algebra Appl. **57**, 41–56 (1984).
[73] S. T. Belinschi, T. Mai, and R. Speicher, “Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem,” J. Reine Angew. Math. (Crelles J.) **2017**, 21–53.
[74] Numerical routines for this paper can be found in the supplementary material of the arXiv version.eprint
[75] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices* (Cambridge University Press, 2010).
[76] M. L. Mehta, *Random Matrices* (Academic Press, 2004), Vol. 142.
[77] L. Zhang, “Volumes of orthogonal groups and unitary groups,” arXiv:1509.00537v5 (2015).