Almost no experiments have classical Kirkwood-Dirac representations

Christopher Langrenez[1], Wilfred Salmon[2,3], Stephan De Bièvre[1], Jonathan J. Thio[4], Christopher K. Long[2,4], and David R. M. Arvidsson-Shukur[2]

[1]Univ. Lille, CNRS, Inria, UMR 8524, Laboratoire Paul Painlevé, F-59000 Lille, France

[2]Hitachi Cambridge Lab., J.J. Thomson Avenue, Cambridge CB3 0HE, UK

[3]DAMTP, Centre for Mathematical Sciences, Univ. of Cambridge, CB30WA, UK

[4]Cavendish Lab., Department of Physics, Univ. of Cambridge, Cambridge CB3 0HE, UK

A central problem in quantum information is determining quantum-classical boundaries. In the quasiprobability framework, a state is called classical if it is represented by a quasiprobability distribution that is positive, and thus a probability distribution. In recent years, the Kirkwood-Dirac (KD) distributions have gained much interest due to their numerous applications in modern quantum-information research. A particular advantage of the KD distributions is that they can be defined with respect to arbitrary observables. Here, we show that if two d-dimensional observables are picked at random, the set of classical (positive) states of the resulting KD distribution is a minimal polytope of dimension $2(d-1)$ with $2d$ explicitly known vertices. This implies minimality of the sets of KD-real observables, of KD-positive measurement elements and of KD-positivity-preserving unitaries. We show how these results have implications on robust observations of nonclassical phenomena, on classical simulations of quantum circuits, and on foundations of quantum theory.

Introduction:—Quasiprobability distributions are useful tools to study and characterize quantum experiments. They are objects similar to probability distributions but can assume negative or nonreal values—hallmarks of quantum phenomena. The Wigner function, which represents quantum states with distributions over the eigenvalues of conjugate observables, has played a pivotal role in the development of quan-

Christopher Langrenez: christopher.langrenez@univ-lille.fr

tum optics [1, 2]. More generally, the Wigner function, and other quasiprobability distributions, are practical tools for investigating the classical-quantum boundary in various quantum-information-processing tasks [3–10]. Such advantages can be obtained only if the quasiprobability distributions characterizing the states, observables and processes involved are not all positive [11–13]. This highlights the importance of identifying the (convex) set of positive states, meaning those states that have positive quasiprobability distributions.

For example, the only pure Wigner-positive states are the pure Gaussian states [14, 15]. However, a similar characterization of the set of mixed Wigner-positive states is currently unknown and the subject of significant research [15–18]. In particular, there exist Wigner positive states that are not convex combinations (mixtures) of gaussian states [15, 17, 19, 20].

Recently, the Kirkwood-Dirac (KD) distributions [21–23] have come to the foreground as versatile mathematical tools to study discrete-variable quantum-information processes. Consider the sets of projectors $\mathcal{A} = \{|a_i\rangle\langle a_i|\}$ and $\mathcal{B} = \{|b_j\rangle\langle b_j|\}$ associated with two bases $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$ in a d-dimensional Hilbert space $\mathcal{H}$. Typically, these bases are eigenbases of observables A and B determined by the physical context of interest. Unlike the observables of the Wigner function, A and B need not be conjugate. The KD distribution associates to each quantum state ρ a quasiprobability distribution $Q(\rho)$ with entries given by

$$Q_{ij}(\rho) = \langle b_j|a_i\rangle\langle a_i|\rho|b_j\rangle \in \mathbb{D}, \qquad (1)$$

where $\mathbb{D}$ is the closed complex unit disk. In anal-

ogy with the definition of Wigner-positive states, if $Q_{ij}(\rho) \in [0,1]$, for all i, j, such that it is a proper probability distribution, we call ρ KD-positive, or classical. If $Q(\rho)$ takes negative or nonreal entries, we call ρ KD nonpositive, or non-classical.

The KD distributions' negativity or nonreality underlie quantum advantages in diverse fields [23, 24]. Examples include metrology [8, 10, 25], weak values [26–31], thermodynamics [5, 9, 24, 32, 33] and information scrambling [6, 32, 34–36]. Moreover, negative or nonreal KD distributions can signal contextuality (a strict notion of nonclassicality) [9, 23, 37–39] and violations of Leggett-Garg inequalities (temporal Bell inequalities) [40–43].

In light of its diverse applications, a full characterization of KD-positive states is warranted. However, such a characterization is, as for the Wigner-positive states, nontrivial. The difficulty is exacerbated by the KD distribution's dependence on the two arbitrary bases in its definition. A partial characterization of the KD-positive pure states is obtained in Refs. [44–49], where KD positivity is linked to quantum uncertainty. Moreover, initial results on the KD positivity of mixed states were obtained in specific instances [50–54], which highlight the intricacies of the structure of the set of KD-positive states: This set, which we shall denote by $\mathcal{E}_{\text{KD+}}$, is not necessarily a polytope, and if it is, its extremal points may not be pure states, making them hard to identify. (See Fig. 1.)

In this Article, we fully characterize the set of KD-positive states of almost all KD distributions. For almost all choices of the bases, we show that all KD-positive states are mixtures of the bases' states: They are of the simplest and minimal form. (See Theorem 1.) Our result further implies that the set of observables, unitary transformations, and generalized measurements (defined below) with classical KD representations are also almost always minimal. Moreover, the only unitaries that are KD-positivity preserving are (trivial) global phases. Finally, we discuss applications and implications of our results in relation to witness-construction for nonclassical phenomena, classical simulations of quantum circuits, and foundations of quantum theory.

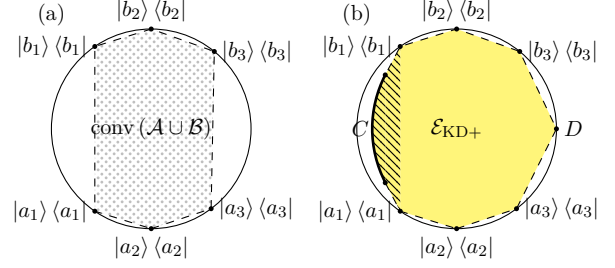*The KD representation of quantum theory:—* The KD distributions, as do other quasiproba-

Figure 1: Schematic representation of possible geometries of $\mathcal{E}_{\text{KD+}}$, when $d = 3$. The disc represents all states, with the boundary representing the pure states. (a) The simplest situation (Theorem 1): $\mathcal{E}_{\text{KD+}} = \text{conv}(\mathcal{A} \cup \mathcal{B})$. (b) A more complex situation: $\text{conv}(\mathcal{A} \cup \mathcal{B}) \subsetneq \text{conv}(\mathcal{E}_{\text{KD+}}^{\text{pure}}) \subsetneq \mathcal{E}_{\text{KD+}}$. The point D represents a pure state in $\mathcal{E}_{\text{KD+}}$. The arc C represents a set of mixed extremal states in $\mathcal{E}_{\text{KD+}}$. The hatched part represents the so-called *exotic* states.

bility distributions, provide useful alternatives to the standard Schrödinger picture of quantum mechanics. If $m_{\mathcal{A},\mathcal{B}} := \min |\langle a_i|b_j\rangle| > 0$ (as we assume henceforth), the KD distribution is informationally complete: Knowledge of $Q(\rho)$ uniquely determines ρ. The marginals of $Q(\rho)$ reproduce the Born rule:

$$\sum_j Q_{ij}(\rho) = \langle a_i|\rho|a_i\rangle, \quad \sum_i Q_{ij}(\rho) = \langle b_j|\rho|b_j\rangle$$
$$\sum_{i,j} Q_{ij}(\rho) = 1. \tag{2}$$

Therefore, $Q(\rho)$ behaves like a joint probability distribution for the observables A and B, except that it may take complex values.

To extend the KD formalism to a full representation of quantum mechanics, we introduce the KD symbol $W(F)$ of an operator F:

$$W_{ij}(F) = \frac{\langle a_i|F|b_j\rangle}{\langle a_i|b_j\rangle} = \frac{Q_{ij}(F)}{|\langle a_i|b_j\rangle|^2}. \tag{3}$$

Here, we extended Eq.(1) by linearity to all operators F on $\mathcal{H}$. Then, the expectation value $\text{Tr}\left(F^\dagger \rho\right)$ of any operator F with respect to any state ρ is [13, 23, 50]

$$\text{Tr}\left(F^\dagger \rho\right) = \sum_{i,j} W_{ij}^*(F)Q_{ij}(\rho). \tag{4}$$

Mathematically, each matrix element $W_{ij}(F)$ is a *weak value* of F [26]. We call an operator KD real if its KD symbol is real. The real vector space of KD-real observables is denoted by V_{KDr}: $F \in V_{\text{KDr}}$ provided that $F = F^\dagger$ and that $W_{ij}(F) \in \mathbb{R}$ for all i, j. The set V_{KDr} of KD-real observables is

in general equally difficult to characterize as the set $\mathcal{E}_{\text{KD}+}$ of KD-positive states.

A generalized measurement (also known as a positive-operator-valued measure) is described by a set $\{M_k\}$, the elements of which are positive semi-definite operators that satisfy $\sum_k M_k = \mathbb{I}$. We denote by $M_{\text{KD}+}$ the space of measurement elements M with $W_{ij}(M) \geq 0$.

Together, the KD distribution Q and the KD symbol W provide a quasiprobability representation of quantum mechanics. Their restrictions to $\mathcal{E}_{\text{KD}+}$, respectively $M_{\text{KD}+}$, provide a proper probability representation of a (classically describable) "fragment" of quantum mechanics. The states of this fragment belong to $\mathcal{E}_{\text{KD}+}$ and its observables are modeled with generalized measurements with elements in $M_{\text{KD}+}$.

Main result:—We now explore further the set of KD-positive states and KD-real observables. Given two orthonormal bases $\mathcal{A}$ and $\mathcal{B}$, one can readily confirm that

$$\text{conv}(\mathcal{A} \cup \mathcal{B}) \subseteq \mathcal{E}_{\text{KD}+} \text{ and } \text{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B}) \subseteq V_{\text{KDr}}. \tag{5}$$

Here, $\text{conv}(Z)$ denotes all convex combinations of the elements in Z, $\text{span}_{\mathbb{R}}(Z)$ denotes the real vector space generated by Z, and $\cup$ denotes the union of two sets. Our main result is:

Theorem 1. *Let* $\mathcal{A} = \{|a_i\rangle\langle a_i|\}$ *and* $\mathcal{B} = \{|b_j\rangle\langle b_j|\}$ *be randomly chosen. Then, with probability 1, the sets of KD-positive states, of KD-real observables, of KD-positive measurement elements and of KD-positivity-preserving unitaries* T *(in $d > 2$) are all of minimal size:*

$$\mathcal{E}_{\text{KD}+} = \text{conv}(\mathcal{A} \cup \mathcal{B}), \tag{6}$$

$$V_{\text{KDr}} = \text{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B}), \tag{7}$$

$$M_{\text{KD}+} \subset \mathbb{R}^+\text{conv}(\mathcal{A} \cup \mathcal{B}), \tag{8}$$

$$T = \exp(i\alpha)\mathbb{I}, \alpha \in \mathbb{R}. \tag{9}$$

Equation (6) asserts that a state ρ is KD positive if, and only if, there exist $0 \leq \lambda_i, \mu_j \leq 1$ with $\sum_i \lambda_i + \sum_j \mu_j = 1$, such that

$$\rho = \sum_{i,j} \lambda_i |a_i\rangle\langle a_i| + \mu_j |b_j\rangle\langle b_j|. \tag{10}$$

Thus, $\mathcal{E}_{\text{KD}+}$ is a $2(d-1)$-dimensional polytope with $\{|a_i\rangle\langle a_i|\}$ and $\{|b_j\rangle\langle b_j|\}$ as its $2d$ known vertices. $\mathcal{E}_{\text{KD}+}$ cannot be "smaller" as any projector in $\mathcal{A}$ or in $\mathcal{B}$ is KD positive by construction. Moreover, Eq.(7) asserts that, with probability

1, an observable is KD real if, and only if, there exist $\lambda_i, \mu_j \in \mathbb{R}$ such that

$$F = \sum_{i,j} \lambda_i |a_i\rangle\langle a_i| + \mu_j |b_j\rangle\langle b_j|. \tag{11}$$

Similarly, Eq.(8) asserts that any measurement element $M \in M_{\text{KD}+}$ is of the form

$$M = \sum_{i,j} \lambda_i |a_i\rangle\langle a_i| + \mu_j |b_j\rangle\langle b_j|, \tag{12}$$

with $0 \leq \lambda_i, \mu_j \leq 1$. Finally, Eq. (9) asserts that only trivial unitaries T transform any KD-positive state ρ into a KD-positive state $T\rho T^\dagger$.

Before proving Theorem 1, we explain the precise meaning of the expression "randomly chosen". Whereas both sides of Eqs. (6)–(8) depend on the sets of basis projectors $\mathcal{A}$ and $\mathcal{B}$, the relations themselves depend only on their transition matrix U (with elements $U_{ij} = \langle a_i|b_j\rangle$), not on the individual basis elements. In other words, if $\mathcal{A}'$ and $\mathcal{B}'$ are such that $U' = U$, then, if Eqs. (6)–(8) hold for $\mathcal{A}$ and $\mathcal{B}$, it also holds for $\mathcal{A}'$ and $\mathcal{B}'$ [46]. The unitary group $\mathcal{U}(d)$ carries a unique invariant probability measure, called the Haar measure, that we denote by μ_{H} [55]. Theorem 1 asserts that there is a probability-1 subset of transition matrices for which Eqs. (6)–(8) hold.

We now turn to the proof of the theorem. We denote by Ω the set of all unitary matrices U for which $\min |U_{ij}| > 0$; these are the unitary matrices that contain no zeroes. The set Ω is open and $\mu_{\text{H}}(\Omega) = 1$. In other words, Ω is a probability-1 subset of the unitary group. The latter statement follows directly from Proposition 3. It was established in [50] that, when $U \in \Omega$, $\dim(\text{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B})) = 2d - 1$ and

$$\mathcal{E}_{\text{KD}+} = \text{conv}(\mathcal{A} \cup \mathcal{B}) \Leftrightarrow V_{\text{KDr}} = \text{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B}). \tag{13}$$

In view of Eq.(5), the theorem thus follows once we show that, with probability 1, $V_{\text{KDr}} \subseteq \text{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B})$ or, equivalently, that $\dim V_{\text{KDr}} = 2d - 1$.

We start by noting that

$$V_{\text{KDr}} = \text{Ker}(\text{Im}W) \cap \mathcal{S}(\mathcal{H}) \tag{14}$$

where $\cap$ denotes the intersection of two sets, $\mathcal{S}(\mathcal{H})$ denotes the real vector space of self-adjoint operators on $\mathcal{H}$, and $\text{Im}Z$ denotes the imaginary part of Z. Indeed, an operator F is KD real whenever $\text{Im}W(F) = 0$, *i.e.* $F \in \text{Ker}(\text{Im}W)$.
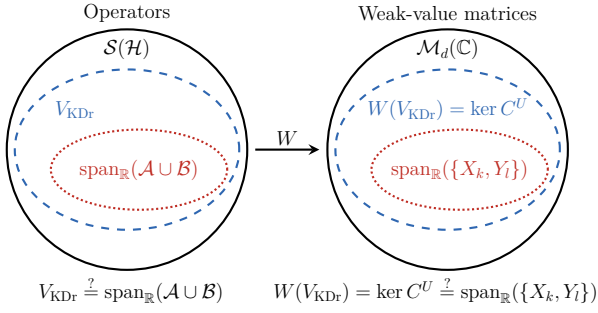
Figure 2: The two circles represent the set of all operators and the corresponding set of KD symbols. Proving the Theorem reduces to showing that, with probability 1, the red (dotted) and the blue (striped) sets coincide. The blue and red sets on the right have simpler mathematical descriptions than they do on the left, facilitating the proof of the Theorem.

As Q is informationally complete (bijective), Eq.(3) implies that the same holds for W. That is, given an arbitrary matrix $w \in \mathcal{M}_d(\mathbb{C})$, there is a unique operator $F \in \mathcal{L}(\mathcal{H})$ such that

$$\langle a_i | F | b_j \rangle = w_{ij} \langle a_i | b_j \rangle. \qquad (15)$$

Hence

$$\dim V_{\mathrm{KDr}} = \dim W(V_{\mathrm{KDr}}). \qquad (16)$$

In what follows, we will show that $\dim W(V_{\mathrm{KDr}}) = 2d - 1$ with probability 1, thereby concluding the proof. We give a schematic overview of our proof strategy in Fig. 2.

We proceed by showing how to characterize the real vector space $W(V_{\mathrm{KDr}})$ as a subspace of $\mathcal{M}_d(\mathbb{R})$, the space of $d \times d$ real matrices. For that purpose, we first identify $W(\mathcal{S}(\mathcal{H}))$. A direct computation, using the definition of W, shows that an operator F is self-adjoint if, and only if, for all i, j,

$$\sum_l (W_{il}(F) - W_{jl}(F)) U_{jl}^* U_{il} = 0. \qquad (17)$$

It follows that $w \in W(V_{\mathrm{KDr}})$ if, and only if, $w \in \mathcal{M}_d(\mathbb{R})$ and, for all i, j,

$$\sum_l (w_{il} - w_{jl}) U_{jl}^* U_{il} = 0. \qquad (18)$$

This observation motivates us to define, for each unitary operator U, the real linear map $C^U : \mathcal{M}_d(\mathbb{R}) \to \mathcal{S}_d(\mathbb{C})$, where

$$C_{jk}^U(w) = i \sum_l (w_{jl} - w_{kl}) U_{kl}^* U_{jl}. \qquad (19)$$

One can therefore conclude that

$$W(V_{\mathrm{KDr}}) = \mathrm{Ker}(C^U). \qquad (20)$$

This implies that the two real vector spaces V_{KDr} and $\mathrm{Ker}(C^U)$ are isomorphic and hence have the same dimension. Additionally, in view of Eq.(5),

$$W(\mathrm{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B})) \subseteq W(V_{\mathrm{KDr}}) = \mathrm{Ker}(C^U). \qquad (21)$$

Hence, $\dim(\mathrm{Ker}(C^U)) \geq 2d - 1$. In other words, given U, the linear algebra question we need to address is: "Is the dimension of $\mathrm{Ker}(C^U)$ equal to $2d - 1$?"

We observe that

$$W(\mathrm{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B})) = \mathrm{span}_{\mathbb{R}}\{X_k, Y_\ell\}, \qquad (22)$$

where $X_k = W(|a_k\rangle\langle a_k|)$ and $Y_k = W(|b_k\rangle\langle b_k|)$ are $d \times d$ matrices given by

$$X_k = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ 1 & \cdots & 1 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \qquad (23)$$

$$Y_k = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}. \qquad (24)$$

Here, X_k is a matrix with ones on its kth row and zeroes elsewhere, and $Y_k = X_k^T$. While both V_{KDr} and the map W depend on the two bases $\mathcal{A}$ and $\mathcal{B}$, the real vector space $W(V_{\mathrm{KDr}}) = \mathrm{Ker}(C^U)$ depends only on their transition matrix U, not on the individual bases. In addition, $W(\mathrm{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B}))$ depends neither on the bases $\mathcal{A}$ and $\mathcal{B}$, nor even on their transition matrix U. This constitutes a considerable simplification that we will exploit to determine the dimension of $\mathrm{Ker}(C_U)$.

For any U, we define

$$\mathcal{T} := \mathrm{span}_{\mathbb{R}}\{X_k, Y_\ell\} \subseteq \mathrm{Ker}(C^U). \qquad (25)$$

$\mathcal{T}$ is well defined even if U does have zeroes since C^U in Eq. (19) is always well defined. Let $\mathcal{N} = \mathcal{T}^{\perp}$ denote the orthogonal complement of $\mathcal{T}$ in $\mathcal{M}_d(\mathbb{R})$ with respect to the Hilbert-Schmidt inner

product on $\mathcal{M}_d(\mathbb{C})$ given by $\langle A, B \rangle = \text{Tr}\left(A^\dagger B\right)$. Then, $\dim \mathcal{N} = d^2 - (2d-1) = (d-1)^2$ and

$$C^U : \mathcal{M}_d(\mathbb{R}) = \mathcal{T} \oplus \mathcal{N} \to \text{Ran}\left(C^U\right) \subseteq \mathcal{S}_d(\mathbb{C}). \tag{26}$$

Here, $\text{Ran}(Z)$ denotes the range of Z. Since $\mathcal{T}$ does not depend on U, the subspace $\mathcal{N} \subsetneq \mathcal{M}_d(\mathbb{R})$ is independent of U as well. We then define the set of "good" unitaries

$$\Gamma = \{U \in \mathcal{U}(d) \mid \text{Ker}(C^U) = \mathcal{T}\}. \tag{27}$$

This terminology is justified by the fact that $\dim \mathcal{T} = 2d - 1$. Thus, the proof of Eqs. (6)–(8) of Theorem 1 is reduced to the proof of the following proposition.

Proposition 2. *The set Γ of unitary matrices U for which $\text{Ker}(C^U) = \mathcal{T}$ is a set of probability 1:*

$$\mu_H(\Gamma) = 1. \tag{28}$$

Proof of Proposition 2. It follows from the rank-nullity theorem that $U \in \Gamma$ if, and only if,

$$C^U|_{\mathcal{N}} : \mathcal{N} \to \mathcal{S}_d(\mathbb{C}) \tag{29}$$

is injective. We now show that this is indeed the case with probability 1.

Our proof relies on the existence of two bases $\mathcal{A}_\star$ and $\mathcal{B}_\star$ for which $U_\star \in \Gamma$. As follows from a simple calculation in the Appendix B, in any dimension d, this property is satisfied by

$$U_\star = \exp(id^{-1}J_d) = I_d + zJ_d. \tag{30}$$

Here, $z = d^{-1}\left(e^i - 1\right)$, I_d is the identity matrix, and J_d is the matrix of ones. We define

$$\mathcal{R}_\star := \text{Ran}\left(C^{U_\star}\right) \subseteq \mathcal{S}_d(\mathbb{C}). \tag{31}$$

Since $U_\star \in \Gamma$, we know that $\dim(\mathcal{R}_\star) = \dim(\mathcal{N}) = (d-1)^2$. Let $\Pi_\star$ be the orthogonal projector onto $\mathcal{R}_\star$:

$$\Pi_\star : \mathcal{S}_d(\mathbb{C}) \to \mathcal{R}_\star \subsetneq \mathcal{S}_d(\mathbb{C}), \quad \Pi_\star^2 = \Pi_\star. \tag{32}$$

For every $U \in \mathcal{U}(d)$ we then define

$$\widetilde{C}^U := \Pi_\star C^U|_{\mathcal{N}} : \mathcal{N} \to \mathcal{R}_\star. \tag{33}$$

If $\widetilde{C}^U$ is a bijection, then $C^U|_{\mathcal{N}}$ is an injection which in turn implies that $U \in \Gamma$. Moreover, by construction, $\widetilde{C}^{U_\star}$ is a bijection. We now show that the unitaries $U \in \mathcal{U}(d)$ for which $\tilde{C}^U$ is a

bijection form a set of probability 1. This will conclude the proof. Consider two bases, one in $\mathcal{N}$ and one in $\mathcal{R}_\star$. Because neither $\mathcal{N}$, nor $\mathcal{R}_\star$ depend on U these bases can be chosen independently of U. We then denote the determinant of $\tilde{C}^U$ with respect to this choice of bases by $\det\left(\tilde{C}^U\right)$. Since the matrix elements of C^U are quadratic polynomials in the real and imaginary parts of the matrix elements of U, it follows that the map

$$c : \mathcal{U}(d) \to \mathbb{R}, \; U \mapsto \det\left(\widetilde{C}^U\right) \tag{34}$$

is also a polynomial in the real and imaginary parts of the entries of U. Since $c(U_\star) \neq 0$, the result now follows from Proposition 3 below. $\square$

Proposition 3. *Let $f : \mathcal{U}(d) \to \mathbb{R}$ be a polynomial in the real and imaginary matrix elements of U. Suppose that there exists at least one unitary matrix $U_\star$ for which $f(U_\star) \neq 0$; then the zero-set of f has vanishing Haar measure.*

We provide a proof in Appendix A (which contains references [56–63] therein). The proof relies on the fact that the zero sets of real analytic functions on $\mathbb{R}^m$ have vanishing Lebesgue measure. Finally, our proof yields as a corollary Eq.(9).

Discussion:—We have shown that the convex set of KD-positive states, $\mathcal{E}_{\text{KD}+}$, is, with probability 1, simple and of minimal size. That is, $\mathcal{E}_{\text{KD}+}$ is almost always the polytope of all mixtures of the states in $\mathcal{A}$ and $\mathcal{B}$, of which the supporting hyperplanes are known [50]. Also the space of KD-real observables, the set of KD-positivity-preserving unitary transformations, and the set of KD-positive measurement elements are almost always minimal. Below, we discuss four applications of our results, with details in Appendices D, E and F.

First, consider the task of preparing a state ρ with the purpose of generating nonclassical phenomena in any of the numerous fields where KD nonpositivity signals quantum advantages. Naturally, one would want ρ to be far from $\mathcal{E}_{\text{KD}+}$, i.e., far from the shaded areas in Fig. 1. In general, it can be a formidable task to characterize the region outside $\mathcal{E}_{\text{KD}+}$. For example, the KD distribution naturally associated to n qubits ($d = 2^n$) has $\mathcal{O}(d^{\log_2(d)})$ KD-positive pure extremal states and additional elusive mixed extremal states. Then $\mathcal{E}_{\text{KD}+}$ is not necessarily a polytope and determining the distance of ρ to this

set is often unfeasible. However, our results show that almost all KD distributions have only $2d$ explicitly known positive extremal states. For such KD distributions, one can lower bound the distance of any given ρ to $\mathcal{E}_{\mathrm{KD}+}$ through the known witnesses associated with the facets of the simple polytope $\mathcal{E}_{\mathrm{KD}+} = \mathrm{conv}\,(\mathcal{A} \cup \mathcal{B})$.

Second, our theorem has applications to the foundational concept of contextuality [64]. Generalizing Bell nonlocality, contextuality describes the nonclassical phenomenon wherein experimentally indistinguishable procedures cannot be represented identically on a hidden-variable level. Prior works have established that experiments measuring the weak values in Eq. (3) are noncontextual (classical) if, and only if, the input state ρ is KD-positive [29, 39, 65]. Applying Theorem 1, we find that the set of input states that yield classical behavior is minimal and well-known for almost all observables A and B.

Third, certain forms of nonclassical phenomena [39] stem from mixed *exotic* states: states that have a positive KD distribution but cannot be decomposed in terms of KD-positive pure states. [See Fig. 1(b).] Our results show that such states can exist only in rare and tailored KD distributions, such as those associated to certain spin-1 [50] or n-qubit systems.

Fourth, Theorem 1 has implications for the prospect of employing the KD distribution for efficient classical simulation of amenable quantum circuits. Quasiprobability distributions can be used to simulate quantum circuits using Monte Carlo sampling [66]. The runtime of such techniques scales polynomially with the accuracy but exponentially with the total KD nonpositivity of the circuit's input state, unitary gates, and measurements. Our results immediately imply that, with probability 1, the only KD-positive circuit is an initialization in $\mathrm{conv}\,(\mathcal{A} \cup \mathcal{B})$ followed immediately by a measurement in $M_{\mathrm{KD}+}$. Consequently, only rare KD distributions can be expected to be of use in standard protocols for efficient quantum-circuit simulation.

To summarize, our results prove that the characterization of nonclassical features connected with KD distributions is almost always vastly simpler than it can be. Our results also show that rare KD distributions must be tailored to unveil certain aspects of quantum theory. Such KD tailoring will benefit from further studies of

the map C^U.

All authors contributed equally. The author ordering was randomized.

References

[1] U. Leonhardt. *Essential Quantum Optics: From Quantum Measurements to Black Holes.* p.67. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511806117.

[2] A. Serafini. *Quantum Continuous Variables: A Primer of Theoretical Methods (1st ed.)* Boca Raton: CRC Press., 2017. DOI: 10.1201/9781315118727.

[3] V. Veitch, C. Ferrie, D. Gross, and J. Emerson. "Negative quasi-probability as a resource for quantum computation". In: *New Journal of Physics* 14.11 (Nov. 2012), p. 113011. DOI: 10.1088/1367-2630/14/11/113011.

[4] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson. "The resource theory of stabilizer quantum computation". In: *New Journal of Physics* 16.1 (Jan. 2014), p. 013009. DOI: 10.1088/1367-2630/16/1/013009.

[5] M. Lostaglio. "Quantum Fluctuation Theorems, Contextuality, and Work Quasiprobabilities". In: *Phys. Rev. Lett.* 120 (4 Jan. 2018), p. 040602. DOI: 10.1103/PhysRevLett.120.040602.

[6] J. R. González Alonso, N. Yunger Halpern, and J. Dressel. "Out-of-Time-Ordered-Correlator Quasiprobabilities Robustly Witness Scrambling". In: *Phys. Rev. Lett.* 122 (4 Feb. 2019), p. 040404. DOI: 10.1103/PhysRevLett.122.040404.

[7] R. Raussendorf, J. Bermejo-Vega, E. Tyhurst, C. Okay, and M. Zurel. "Phase-space-simulation method for quantum computation with magic states on qubits". In: *Phys. Rev. A* 101 (1 Jan. 2020), p. 012350. DOI: 10.1103/PhysRevA.101.012350.

[8] D. R. M. Arvidsson-Shukur, N. Yunger Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd. "Quantum advantage in postselected metrology". In: *Nat. Commun.* 11.1 (2020), p. 3775. DOI: 10.1038/s41467-020-17559-w.

[9] M. Lostaglio. "Certifying Quantum Signatures in Thermodynamics and Metrology via Contextuality of Quantum Linear Response". In: *Phys. Rev. Lett.* 125 (23 Dec. 2020), p. 230603. DOI: 10.1103/PhysRevLett.125.230603.

[10] N. Lupu-Gladstein, Y. B. Yilmaz, D. R. M. Arvidsson-Shukur, A. Brodutch, A. O. Pang, A. M. Steinberg, and N. Yunger Halpern. "Negative quasiprobabilities enhance phase estimation in quantum-optics experiment". In: *Physical Review Letters* 128.22 (2022), p. 220504.

[11] R. W. Spekkens. "Negativity and Contextuality are Equivalent Notions of Nonclassicality". In: *Phys. Rev. Lett.* 101 (2 July 2008), p. 020401. DOI: 10.1103/PhysRevLett.101.020401.

[12] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens. "A structure theorem for generalized-noncontextual ontological models". In: *Quantum* 8 (Mar. 2024), p. 1283. DOI: 10.22331/q-2024-03-14-1283.

[13] D. Schmid, R. D. Baldijão, Y. Yīng, R. Wagner, and J. H. Selby. "Kirkwood-Dirac Representations beyond Quantum States and Their Relation to Noncontextuality". In: *Physical Review A* 110.5 (2024), p. 052206. DOI: 10.1103/PhysRevA.110.052206.

[14] R. L. Hudson. "When is the Wigner quasiprobability density non-negative?" In: *Reports on Mathematical Physics* 6.2 (1974), pp. 249–252.

[15] D. Gross. "Hudson's Theorem for Finite-Dimensional Quantum Systems". In: *Journal of Mathematical Physics* 47.12 (Dec. 2006), p. 122107. DOI: 10.1063/1.2393152.

[16] A. Mandilara, E. Karpov, and N. J. Cerf. "Extending Hudson's theorem to mixed quantum states". In: *Phys. Rev. A* 79 (6 June 2009), p. 062302. DOI: 10.1103/PhysRevA.79.062302.

[17] Z. Van Herstraeten and N. J. Cerf. "Quantum Wigner entropy". In: *Physical Review A* 104.4 (Oct. 13, 2021). Publisher: American Physical Society, p. 042211. DOI: 10.1103/PhysRevA.104.042211.

[18] U. Chabaud, P.-E. Emeriau, and F. Grosshans. "Witnessing Wigner Negativity". In: *Quantum* 5 (2021). Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, p. 471. DOI: 10.22331/q-2021-06-08-471.

[19] M. G. Genoni, M. G. A. Paris, and K. Banaszek. "Measure of the non-Gaussian character of a quantum state". In: *Phys. Rev. A* 76 (4 Oct. 2007), p. 042327. DOI: 10.1103/PhysRevA.76.042327.

[20] A. Hertz and S. De Bièvre. "Decoherence and nonclassicality of photon-added and photon-subtracted multimode Gaussian states". In: *Physical Review A* 107.4 (2023), p. 043713. DOI: 10.1103/PhysRevA.107.043713.

[21] J. G. Kirkwood. "Quantum statistics of almost classical assemblies". In: *Physical Review* 44.1 (1933), p. 31.

[22] P. A. M. Dirac. "On the Analogy Between Classical and Quantum Mechanics". In: *Rev. Mod. Phys.* 17 (2-3 Apr. 1945), pp. 195–199. DOI: 10.1103/RevModPhys.17.195.

[23] D. R. M. Arvidsson-Shukur, W. F. Braasch Jr, S. De Bièvre, J. Dressel, A. N. Jordan, C. Langrenez, M. Lostaglio, J. S. Lundeen, and N. Yunger Halpern. "Properties and Applications of the Kirkwood–Dirac

Distribution". In: *New Journal of Physics* 26.12 (2024), p. 121201. DOI: 10.1088/1367-2630/ada05d.

[24] M. Lostaglio, A. Belenchia, A. Levy, S. Hernández-Gómez, N. Fabbri, and S. Gherardini. "Kirkwood-Dirac quasiprobability approach to the statistics of incompatible observables". In: *Quantum* 7 (2023), p. 1128.

[25] J. H. Jenne and D. R. M. Arvidsson-Shukur. "Unbounded and lossless compression of multiparameter quantum information". In: *Phys. Rev. A* 106 (4 Oct. 2022), p. 042404. DOI: 10.1103/PhysRevA.106.042404.

[26] Y. Aharonov, D. Z. Albert, and L. Vaidman. "How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100". In: *Phys. Rev. Lett.* 60 (14 Apr. 1988), pp. 1351–1354. DOI: 10.1103/PhysRevLett.60.1351.

[27] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. "Colloquium: Understanding quantum weak values: Basics and applications". In: *Rev. Mod. Phys.* 86 (1 Mar. 2014), pp. 307–316. DOI: 10.1103/RevModPhys.86.307.

[28] S. Pang, J. Dressel, and T. A. Brun. "Entanglement-Assisted Weak Value Amplification". In: *Phys. Rev. Lett.* 113 (3 July 2014), p. 030401. DOI: 10.1103/PhysRevLett.113.030401.

[29] M. F. Pusey. "Anomalous Weak Values Are Proofs of Contextuality". In: *Phys. Rev. Lett.* 113 (20 Nov. 2014), p. 200401. DOI: 10.1103/PhysRevLett.113.200401.

[30] J. Dressel. "Weak values as interference phenomena". In: *Phys. Rev. A* 91 (3 Mar. 2015), p. 032116. DOI: 10.1103/PhysRevA.91.032116.

[31] S. Pang and T. A. Brun. "Improving the Precision of Weak Measurements by Postselection Measurement". In: *Phys. Rev. Lett.* 115 (12 Sept. 2015), p. 120401. DOI: 10.1103/PhysRevLett.115.120401.

[32] N. Yunger Halpern. "Jarzynski-like equality for the out-of-time-ordered correlator". In: *Phys. Rev. A* 95 (1 Jan. 2017), p. 012120. DOI: 10.1103/PhysRevA.95.012120.

[33] A. Levy and M. Lostaglio. "Quasiprobability Distribution for Heat Fluctuations in the Quantum Regime". In: *PRX Quantum* 1 (1 Sept. 2020), p. 010309. DOI: 10.1103/PRXQuantum.1.010309.

[34] N. Yunger Halpern, B. Swingle, and J. Dressel. "Quasiprobability behind the out-of-time-ordered correlator". In: *Phys. Rev. A* 97 (4 Apr. 2018), p. 042105. DOI: 10.1103/PhysRevA.97.042105.

[35] R. Mohseninia, J. R. G. Alonso, and J. Dressel. "Optimizing measurement strengths for qubit quasiprobabilities behind out-of-time-ordered correlators". In: *Phys. Rev. A* 100 (6 Dec. 2019), p. 062336. DOI: 10.1103/PhysRevA.100.062336.

[36] S. Majidy, U. Agrawal, S. Gopalakrishnan, A. C. Potter, R. Vasseur, and N. Y. Halpern. "Critical phase and spin sharpening in SU(2)-symmetric monitored quantum circuits". In: *Phys. Rev. B* 108 (5 Aug. 2023), p. 054307. DOI: 10.1103/PhysRevB.108.054307.

[37] H. F. Hofmann. "Quantum paradoxes originating from the nonclassical statistics of physical properties related to each other by half-periodic transformations". In: *Phys. Rev. A* 91 (6 June 2015), p. 062123. DOI: 10.1103/PhysRevA.91.062123.

[38] M. Ji and H. F. Hofmann. "Quantitative relations between different measurement contexts". In: *Quantum* 8 (Feb. 2024), p. 1255. DOI: 10.22331/q-2024-02-14-1255.

[39] J. J. Thio, W. Salmon, C. H. W. Barnes, S. D. Bièvre, and D. R. M. Arvidsson-Shukur. *Contextuality Can be Verified with Noncontextual Experiments.* 2024.

[40] A. N. Jordan, A. N. Korotkov, and M. Büttiker. "Leggett-Garg Inequality with a Kicked Quantum Pump". In: *Phys. Rev. Lett.* 97 (2 July 2006), p. 026805. DOI: 10.1103/PhysRevLett.97.026805.

[41] R. Ruskov, A. N. Korotkov, and A. Mizel. "Signatures of Quantum Behavior in Single-Qubit Weak Measurements". In: *Phys. Rev. Lett.* 96 (20 May 2006), p. 200404. DOI: 10.1103/PhysRevLett.96.200404.

[42] N. S. Williams and A. N. Jordan. "Weak values and the Leggett-Garg inequality in solid-state qubits". In: *Physical review letters* 100.2 (2008), p. 026804.

[43] C. Emary, N. Lambert, and F. Nori. "Corrigendum: Leggett–Garg inequalities (2014 Rep. Prog. Phys. 77 016001)". In: *Reports on Progress in Physics* 77.3 (Feb. 2014), p. 039501. DOI: 10.1088/0034-4885/77/3/039501.

[44] D. R. M. Arvidsson-Shukur, J. C. Drori, and N. Yunger Halpern. "Conditions tighter than noncommutation needed for nonclassicality". In: *Journal of Physics A: Mathematical and Theoretical* 54.28 (June 2021). Publisher: IOP Publishing, p. 284001. DOI: 10.1088/1751-8121/ac0289.

[45] S. De Bièvre. "Complete Incompatibility, Support Uncertainty, and Kirkwood-Dirac Nonclassicality". In: *Physical Review Letters* 127.19 (Nov. 2021), p. 190404. DOI: 10.1103/PhysRevLett.127.190404.

[46] S. De Bièvre. "Relating incompatibility, noncommutativity, uncertainty, and Kirkwood–Dirac nonclassicality". In: *J. Math. Phys.* 64.2 (2023), Paper No. 022202, 31. DOI: 10.1063/5.0110267.

[47] J. Xu. "Classification of incompatibility for two orthonormal bases". In: *Physical Review A: Atomic, Molecular, and Optical Physics* 106.2 (Aug. 2022). Number of pages: 7 Publisher: American Physical Society, p. 022217. DOI: 10.1103/PhysRevA.106.022217.

[48] J. Xu. "Kirkwood-Dirac classical pure states". In: *Physics Letters A* 510 (2024), p. 129529. DOI: https://doi.org/10.1016/j.physleta.2024.129529.

[49] Y.-H. Yang, B.-B. Zhang, X.-L. Wang, S.-J. Geng, and P.-Y. Chen. "Characterizing an Uncertainty Diagram and Kirkwood–Dirac Nonclassicality Based on Discrete Fourier Transform". In: *Entropy* 25.7 (July 2023), p. 1075. DOI: 10.3390/e25071075.

[50] C. Langrenez, D. R. M. Arvidsson-Shukur, and S. De Bièvre. "Characterizing the geometry of the Kirkwood–Dirac-positive states". In: *Journal of Mathematical Physics* 65.7 (July 2024), p. 072201. DOI: 10.1063/5.0164672.

[51] Y.-H. Yang, S. Yao, S.-J. Geng, X.-L. Wang, and P.-Y. Chen. "Geometry of Kirkwood–Dirac classical states: a case study based on discrete Fourier transform". In: *Journal of Physics A: Mathematical and Theoretical* 57.43 (2024). Publisher: IOP Publishing, p. 435303. DOI: 10.1088/1751-8121/ad819a.

[52] J. Xu. *Hermitian Kirkwood-Dirac real operators for discrete Fourier transformations.* Dec. 22, 2024. DOI: 10.48550/arXiv.2412.16945.

[53] S. D. Bièvre, C. Langrenez, and D. Radchenko. *The Kirkwood-Dirac representation associated to the Fourier transform for finite abelian groups: positivity.* 2025.

[54] C. Langrenez, S. De Bièvre, and D. R. M. Arvidsson-Shukur. *Convex roofs witnessing Kirkwood-Dirac nonpositivity.* July 5, 2024. DOI: 10.48550/arXiv.2407.04558.

[55] B. Simon. *Representations of finite and compact groups.* 10. American Mathematical Soc., 1996.

[56] B. C. Hall. *Lie Groups, Lie Algebras, and Representations.* Vol. 222. Graduate Texts in Mathematics. New York, NY: Springer, 2003. DOI: 10.1007/978-0-387-21554-9.

[57] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations.* Vol. 102. Graduate Texts in Mathematics. New York, NY: Springer, 1984. DOI: 10.1007/978-1-4612-1126-6.

[58] M. Spivak. *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus.* Avalon Publishing, 1965. 164 pp.

[59] S. G. Krantz and H. R. Parks. *The Implicit Function Theorem: History, Theory, and Applications.* New York, NY: Springer, 2013. DOI: 10.1007/978-1-4614-5981-1.

[60] B. S. Mityagin. "The Zero Set of a Real Analytic Function". In: *Mathematical Notes* 107.3 (2020), pp. 529–530. DOI: 10.1134/S0001434620030189.

[61] W. Rudin. *Real and complex analysis, 3rd ed.* USA: McGraw-Hill, Inc., 1987.

[62] E. M. Alfsen. "A Simplified Constructive Proof of the Existence and Uniqueness of Haar Measure." In: *Mathematica Scandinavica* 12 (Dec. 1963), pp. 106–116. DOI: 10.7146/math.scand.a-10675.

[63] R. Abraham and J. E. Marsden. *Foundations Of Mechanics*. Basic Books, 1978. 854 pp.

[64] R. W. Spekkens. "Contextuality for preparations, transformations, and unsharp measurements". In: *Physical Review A* 71.5 (May 2005). DOI: 10.1103/physreva.71.052108.

[65] R. Kunjwal, M. Lostaglio, and M. F. Pusey. "Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts". In: *Phys. Rev. A* 100 (4 Oct. 2019), p. 042116. DOI: 10.1103/PhysRevA.100.042116.

[66] H. Pashayan, J. J. Wallman, and S. D. Bartlett. "Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities". In: *Phys. Rev. Lett.* 115 (7 Aug. 2015), p. 070501. DOI: 10.1103/PhysRevLett.115.070501.

A   On the zeroes of a non-zero polynomial function on $\mathcal{U}(d)$

For convenience, we recall here Proposition 3. We expect this result to be well known, but we didn't find a proof in the literature.

Proposition 3. *Let* $f : \mathcal{U}(d) \to \mathbb{R}$ *be a polynomial in the real and imaginary matrix elements of* U. *Suppose that there exists at least one unitary matrix* $U_\star$ *for which* $f(U_\star) \neq 0$; *then the zero-set of* f *has vanishing Haar measure.*

We first recall some known properties of the unitary group $\mathcal{U}(d)$. We then show that the zero set of an analytic function on a connected analytic manifold has zero measure, implies Proposition 3.

Note that throughout this Appendix, we call a function analytic if it is a real analytic function.

First, we need some basic facts on the unitary group $\mathcal{U}(d) = \left\{ U \in \mathcal{M}_d(\mathbb{C}), U^\dagger U = I_d \right\}$ that we collect here. We view a matrix $\mathcal{U} \in U(d)$ as a tuple of d columns:

$$U = (U_1, \ldots, U_d). \tag{35}$$

The U_j satisfy the orthonormality relations

$$U_j^\dagger U_k = \delta_{jk}. \tag{36}$$

We write

$$U = R + iS, \quad \text{and} \quad U_j = R_j + iS_j, \tag{37}$$

for real and imaginary parts, and thus identify U with an element of $\mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R}) \simeq \mathbb{R}^{2d^2}$.

The orthonormality relations of equation (36) become, for all j, k:

$$R_j^T R_k + S_k^T S_j = \delta_{jk}, \quad R_j^T S_k - S_j^T R_k = 0. \tag{38}$$

Proposition 4. *$\mathcal{U}(d)$ is compact and connected. Moreover, it is a d^2 dimensional analytic submanifold of $\mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R}) \simeq \mathbb{R}^{2d^2}$.*

The proof that $\mathcal{U}(d)$ is compact and connected is standard, see for example [56]. We did not find a proof of the analytic part of the Proposition in the literature, although it is certainly well known. We provide an argument below. Let us first recall that a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is analytic if it has, for each $x \in \mathbb{R}^n$, a convergent Taylor series expansion (in some neighbourhood of x). A submanifold of $\mathbb{R}^n$ is said to be a k-dimensional analytic submanifold if it can be locally parametrized with analytic coordinates [57].

Theorem 5 states that a level surface of an analytic function f is an analytic submanifold, provided f has no critical points on the level surface. Theorem 5 is the adaptation to the analytic setting of a well-known result in the C^∞ setting, based on the analytic implicit function theorem; we omit the details (see Theorem 5.3 in [58] and Theorem 6.3.2 in [59]).

Theorem 5. *Let* $(n, m) \in \mathbb{N}^2$ *such that* $m < n$. *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$ *be an analytic function and* $k \in \mathbb{R}^m$. *We denote by* $\Sigma_k = \{x \in \mathbb{R}^n \mid f(x) = k\}$. *Suppose that for all* $x \in \Sigma_k$, $Df(x)$ *has rank* m *(i.e. is of maximal rank). Then,* Σ_k *is an analytic submanifold of* $\mathbb{R}^n$ *with dimension* $n - m$.

We can now prove Proposition 4.

Proof of Proposition 4. To prove that $\mathcal{U}(d)$ is an analytic manifold of dimension d^2, we consider the map $f : \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R}) \to \mathcal{S}_d(\mathbb{R}) \times \mathcal{A}_d(\mathbb{R})$, with the notations used in Eq.(38):

$$f(R,S) = \left(R^T R + S^T S, R^T S - S^T R\right). \quad (39)$$

Here, $\mathcal{S}_d(\mathbb{R})$ is the set of real symmetric d by d matrices and $\mathcal{A}_d(\mathbb{R})$ the set of real antisymmetric d by d matrices. Clearly, f is analytic since it is a polynomial in the matrix elements of R and S. Consequently, using the notation in Theorem 5, one sees that $\mathcal{U}(d) = \Sigma_k$ with $k = (I_d, 0_{d,d})$ where $0_{d,d}$ is the zero matrix of dimension $d \times d$ and f is an analytic function.

Now, for a fixed unitary $U = R_U + iS_U$, we can define $g_U : \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R}) \to \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R})$

$$g_U(R,S) = (R_U R - S_U S, R_U S + S_U R). \quad (40)$$

With this definition, one can check that $f \circ g_U = f$ for all $U \in \mathcal{U}(d)$, meaning that f is invariant under the left action of unitary matrices. This implies that it is enough to prove that the derivative of f at $(R,S) = (I_d, 0_{d,d})$, $Df(I_d, 0_{d,d})$, is of full rank. We find, for all $(H,K) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R})$,

$$Df(I_d, 0_{d,d})(H,K) = (H + H^T, K - K^T).$$

As any matrix A in $\mathcal{S}_d(\mathbb{R})$ can be written as $\frac{1}{2}A + \frac{1}{2}A^T$ and any matrix $B \in \mathcal{A}_d(\mathbb{R})$ can be written as $\frac{1}{2}B - \frac{1}{2}B^T$, this proves that $Df(I_d, 0_{d,d})$ is surjective in $\mathcal{S}_d(\mathbb{R}) \times \mathcal{A}_d(R)$, i.e. f is of full rank.

As $\mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R}) \simeq \mathbb{R}^{2d^2}$ and $\mathcal{S}_d(\mathbb{R}) \times \mathcal{A}_d(\mathbb{R}) \simeq \mathbb{R}^{d^2}$, we can apply Theorem 5 to conclude that $\mathcal{U}(d)$ is an analytic submanifold of $\mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_d(\mathbb{R})$ of dimension d^2. $\square$

We now turn to the results we need on the zeroes of analytic functions on compact, connected and analytic manifolds.

Theorem 6. *Let Ω be an open connected subset of $\mathbb{R}^n$ and $f : \Omega \to \mathbb{C}$ be an analytic function. Let $\mathcal{Z}(f) = \{x \in \Omega \mid f(x) = 0\}$ be the set of zeroes of f. Then either $\mathcal{Z}(f)$ has zero Lebesgue measure or f vanishes identically on Ω, meaning that $\mathcal{Z}(f) = \Omega$.*

A proof of this theorem can be found in [60]. To generalise this theorem, we define a set of zero measure on a manifold as follows. A subset N of a manifold M of dimension m is said to have measure zero if on each local chart $\phi : V \to \Omega$, $\phi(V \cap N)$ is a subset of $\mathbb{R}^m$ with vanishing Lebesgue measure. This definition leads to the following theorem on connected and analytic manifolds.

Theorem 7. *Let M be a connected and analytic manifold of dimension m. Let $f : M \to \mathbb{C}$ be an analytic function and $\mathcal{Z}(f) = \{x \in M \mid f(x) = 0\}$ be the set of zeroes of f. Then, either $\mathcal{Z}(f)$ has zero measure or f vanishes identically on M, meaning that $\mathcal{Z}(f) = M$.*

Proof. As M is a manifold, it is second-countable and thus M is covered by a countable set of charts $\{V_n\}_{n \in \mathbb{N}}$, i.e. $\cup_{n \in \mathbb{N}} V_n = M$. Let $\phi_n : V_n \to \Omega_n \subseteq \mathbb{R}^m$ denote the coordinate functions on the charts V_n. Note that, if we fix $n \in \mathbb{N}$, $f \circ \phi_n^{-1} : \Omega_n \to \mathbb{C}$ is an analytic function on an open set of $\mathbb{R}^m$ (because M is analytic). Consequently, by applying Theorem 6, either $\mathcal{Z}(f \circ \phi_n^{-1})$ has a zero Lebesgue measure or $\mathcal{Z}(f \circ \phi_n^{-1}) = \Omega_n$, meaning that f vanishes on V_n.

We define two sets: I is the set of $n \in \mathbb{N}$ for which $\mathcal{Z}(f \circ \phi_n^{-1})$ has zero Lebesgue measure and J is the set of $n \in \mathbb{N}$ for which f vanishes on V_n. Let $W = \cup_{i \in I} V_i$ and $Z = \cup_{j \in J} V_j$. Note that W, Z are the union of open sets, so are open. Moreover, as the sets V_n cover M, we find $M = W \cup Z$. Additionally, suppose there exists some $p \in W \cap Z$, so that $p \in V_i \cap V_j$ for some $i \in I$ and $j \in J$. Note that, by definition of J, $f \circ \phi_i^{-1}$ vanishes on $\phi_i(V_i \cap V_j) \subseteq \Omega_i$. However, $\phi_i(V_j \cap V_i)$ is open in Ω_i and non-empty. Thus, it has positive measure in Ω_i, so that $f \circ \phi_i^{-1}$ vanishes on a positive measure subset of Ω_i. This contradicts the definition of I, and we deduce $W \cap Z = \emptyset$.

Therefore, as M is connected, either $W = M$ or $Z = M$. If $Z = M$, then f vanishes identically on M. If $M = W$, then for any $n \in \mathbb{N}$, $\phi_n(\mathcal{Z}(f) \cap V_n) = \mathcal{Z}(f \circ \phi_n^{-1})$ has zero Lebesgue measure. Then, for any chart $\phi : V \to \Omega \subseteq \mathbb{R}^m$,

$$\phi(\mathcal{Z}(f) \cap V) = \bigcup_{\substack{n \in \mathbb{N} \\ V_n \cap V \neq \emptyset}} \phi(\mathcal{Z}(f) \cap V \cap V_n)$$

and thus,

$$\phi(\mathcal{Z}(f) \cap V) = \bigcup_{\substack{n \in \mathbb{N} \\ V_n \cap V \neq \emptyset}} (\phi \circ \phi_n^{-1})[\mathcal{Z}(f \circ \phi_n^{-1}) \cap \phi_n(V \cap V_n)].$$

Moreover, for $n \in \mathbb{N}$, $(\phi \circ \phi_n^{-1})$ is analytic, so differentiable, and thus maps measure zero sets to

measure zero sets (see [61] Lemma 7.25). Since the countable union of measure zero sets is itself measure zero, we deduce that $\phi(\mathcal{Z}(f) \cap V)$ is measure zero. Since this is true for any chart (V, ϕ), we conclude that $\mathcal{Z}(f)$ is measure zero in M.

$\square$

Applying Theorem 7 to the special case of $\mathcal{U}(d)$, we obtain Proposition 3.

Proof. Since $\mathcal{U}(d)$ is a compact group, it has a unique measure that is left invariant, called the Haar measure [62]. Moreover, $\mathcal{U}(d)$ is a compact Lie group [57]. From Proposition 4.1.14 in [63], the Haar measure on $\mathcal{U}(d)$ is a volume form, meaning that the Haar measure is absolutely continuous with respect to the Lebesgue measure on each local coordinate patch. Thus, it suffices to show that $\mathcal{Z}(f) := \{U \in \mathcal{U}(d) : f(U) = 0\}$ has zero measure in the sense of Theorem 7.

Note that the matrix entries U_{ij}, U_{ij}^* are analytic functions on $\mathcal{U}(d)$ (by the implicit function theorem, Theorem 6.3.2 in [59]). Since a polynomial of analytic functions is also analytic, we deduce that $f : \mathcal{U}(d) \to \mathbb{R}$ is analytic. Finally, by assumption $f \neq 0$ and we apply Theorem 7. $\square$

As a final remark, we explain how our results imply that the simple structure of $\mathcal{E}_{\mathrm{KD+}}$ is robust under small perturbations of the bases. Recall that Ω is defined as the set of all unitary matrices U that contain no zeroes. Then, with Γ as defined in Eq.(27), the set $\Omega \cap \Gamma$ is open. Consequently, if U has no zeroes and is a "good" unitary ($U \in \Gamma$), then all sufficiently small perturbations of U are also good. Thus, the simple structure of $\mathcal{E}_{\mathrm{KD+}}$ is robust under small perturbations of the bases. Conversely, since $\Omega \cap \Gamma$ has full measure, it is also dense and thus if two bases $\mathcal{A}$ and $\mathcal{B}$ are such that $\mathrm{conv}(\mathcal{A} \cup \mathcal{B}) \subsetneq \mathcal{E}_{\mathrm{KD+}}$, then there exist nearby bases $\mathcal{A}'$ and $\mathcal{B}'$ for which $\mathrm{conv}(\mathcal{A}' \cup \mathcal{B}') = \mathcal{E}_{\mathrm{KD+}}$. In other words, there exists a small perturbation of the bases $\mathcal{A}$ and $\mathcal{B}$ that do have a simple associated $\mathcal{E}_{\mathrm{KD+}}$.

B Proof that $U_\star$ satisfy $\mathrm{Ker}(C^{U_\star}) = \mathrm{span}_{\mathbb{R}} \{X_k, Y_\ell \mid k, \ell \in [\![1, d]\!]\}$.

We recall that $U_\star$ is defined in Eq.(30). First, as $\mathcal{N}$ is defined as $(\mathrm{span}_{\mathbb{R}} \{X_k, Y_\ell \mid k, \ell \in [\![1, d]\!]\})^\perp$,

a matrix $N \in \mathcal{N}$ satisfies the following equations:

$$\forall i \in [\![1, d]\!], \ 0 = \langle X_i, N \rangle = \sum_{j=1}^{d} N_{ij}$$

$$\forall j \in [\![1, d]\!], \ 0 = \langle Y_j, N \rangle = \sum_{i=1}^{d} N_{ij}, \quad (41)$$

where we recall that orthogonality is with respect to the Hilbert-Schmidt inner product $\langle A, B \rangle = \mathrm{Tr}(A^\dagger B)$. This means that a matrix N is in $\mathcal{N}$ if and only if the sum over every column and over every row is zero. Next, we consider $U_\star = \exp(id^{-1}J_d) = I_d + zJ_d$ where $z = d^{-1}(e^i - 1)$, I_d is the identity matrix, and J_d is the matrix of ones. Note that $\mathrm{Re}(z) \neq 0$ and $\mathrm{Im}(z) \neq 0$. We want to prove that $C^{U_\star}|_{\mathcal{N}}$ is injective. Suppose that there exists $N \in \mathcal{N}$ with $C^{U_\star}(N) = 0$. Then, for all j, k:

$$-iC_{jk}^{U_\star}(N) = \sum_{l=1}^{d} (N_{jl} - N_{kl})(U_\star)_{kl}^*(U_\star)_{jl}$$

$$= |z|^2 \sum_{l=1}^{d} (N_{jl} - N_{kl}) + (N_{jj} - N_{kj})z^*$$

$$+ (N_{jk} - N_{kk})z \quad (42)$$

$$= (N_{jj} - N_{kj})z^* + (N_{jk} - N_{kk})z, (43)$$

where the first term from the second line is zero as the sum over every column of N is zero by Eq.(41). Since $C^{U_\star}(N) = 0$, the real and imaginary parts of Eq.(42) are zero, yielding that, for all $(j, k) \in [\![1, d]\!]^2$:

$$(N_{jj} - N_{kj})\mathrm{Re}(z) = (N_{kk} - N_{jk})\mathrm{Re}(z); \quad (44)$$

$$(N_{kj} - N_{jj})\mathrm{Im}(z) = (N_{kk} - N_{jk})\mathrm{Im}(z). \quad (45)$$

As $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ are not zero, adding Eq.(44) and (45) yields that, for all $(j, k) \in [\![1, d]\!]^2$:

$$N_{kk} = N_{jk}. \quad (46)$$

Finally, for a fixed $k \in [\![1, d]\!]$, Eq.(41) implies that:

$$0 = \sum_{j=1}^{d} N_{jk} = \sum_{j=1}^{d} N_{kk} = dN_{kk}. \quad (47)$$

Thus, for all $k \in [\![1, d]\!]$, $N_{kk} = 0$. Eq.(46) then implies that $N = 0$. Consequently, as $\mathrm{Ker}(C^{U_\star}|_{\mathcal{N}}) = \{0\}$, $C^{U_\star}|_{\mathcal{N}}$ is injective, concluding our proof. $\square$

C   On the absence of KD-positivity preserving unitaries.

Let $(|a_i\rangle)_{i\in[\![1,d]\!]}$ and $(|b_j\rangle)_{j\in[\![1,d]\!]}$ be two orthonormal bases of a Hilbert space of dimension d. We denote by U the transition unitary matrix between the two bases and by $\mathcal{E}_{\mathrm{KD}+}$ the set of KD positive states associated to the unitary matrix U. We say that $T \in \mathcal{U}(d)$ is a KD-positivity preserving unitary if and only if $T\mathcal{E}_{\mathrm{KD}+}T^{\dagger} = \mathcal{E}_{\mathrm{KD}+}$. The goal of this section is to prove the following proposition, regarding KD-preserving unitaries.

Proposition 9. *In any dimension $d \geqslant 3$, there exists a set $\mathcal{V}(d)$ of full Haar measure among the set of unitaries $\mathcal{U}(d)$ for which the only KD-positivity preserving unitaries T are given by the family $\left(T(\theta) = e^{i\theta}\mathrm{id}_{\mathcal{H}}\right)_{\theta\in\mathbb{R}}$.*

We now suppose that $U \in \mathcal{W}(d)$, where $\mathcal{W}(d) = \Omega \cap \Gamma$. Here Ω is the set of all unitaries that do not have zeroes, and Γ is defined in Eq.(30). We thus have that $m_{\mathcal{A},\mathcal{B}} > 0$ and $\mathcal{E}_{\mathrm{KD}+} = \mathrm{conv}(\mathcal{A} \cup \mathcal{B})$. This implies that $\mathcal{E}_{\mathrm{KD}+}^{\mathrm{pure}} = \mathcal{A} \cup \mathcal{B}$. As T preserves the extreme points of $\mathcal{E}_{\mathrm{KD}+}$, T realises a permutation of the set of $2d$ elements $\mathcal{A} \cup \mathcal{B}$. The following Lemma shows that only specific permutations are allowed.

Lemma 10. *Suppose that $(|a_i\rangle)_{i\in[\![1,d]\!]}$ and $(|b_j\rangle)_{j\in[\![1,d]\!]}$ be two orthonormal bases such that $m_{\mathcal{A},\mathcal{B}} > 0$ and $\mathcal{E}_{\mathrm{KD}+} = \mathrm{conv}(\mathcal{A} \cup \mathcal{B})$. If T is KD-positivity preserving, then one of the two alternatives occurs:*

1. *There exist permutations (σ, τ) and phases $(\eta_i)_{i\in[\![1,d]\!]}, (\theta_j)_{j\in[\![1,d]\!]} \in \mathbb{R}^{2d}$ such that*

$$\forall i \in [\![1,d]\!], T|a_i\rangle = e^{i\eta_i}|a_{\sigma(i)}\rangle,$$
$$\forall j \in [\![1,d]\!], T|b_j\rangle = e^{i\theta_j}|b_{\tau(j)}\rangle; \quad (48)$$

2. *There exist permutations (σ, τ) and phases $(\eta_i)_{i\in[\![1,d]\!]}, (\theta_j)_{j\in[\![1,d]\!]} \in \mathbb{R}^{2d}$ such that*

$$\forall i \in [\![1,d]\!], T|a_i\rangle = e^{i\eta_i}|b_{\sigma(i)}\rangle,$$
$$\forall j \in [\![1,d]\!], T|b_j\rangle = e^{i\theta_j}|a_{\tau(j)}\rangle. \quad (49)$$

Proof. Suppose that we are not in case *1*. By reordering the bases $(|a_i\rangle)_{i\in[\![1,d]\!]}$ and $(|b_j\rangle)_{j\in[\![1,d]\!]}$, we can suppose that $T|a_1\rangle = e^{i\eta_1}|b_1\rangle$, where $\eta_1 \in$

$\mathbb{R}$. Then, suppose that, for some $\ell \in [\![1,d]\!]$ and some $\eta_2 \in \mathbb{R}$, $T|a_2\rangle = e^{i\eta_2}|a_l\rangle$. This implies that

$$0 = \langle a_1|a_2\rangle = \langle a_2|T^{\dagger}T|a_1\rangle = e^{i(\eta_1-\eta_2)}\langle a_\ell|b_1\rangle. \quad (50)$$

Since $m_{\mathcal{A},\mathcal{B}} > 0$, we obtain a contradiction. Moreover, as T is unitary and preserves the set $\mathcal{A} \cup \mathcal{B}$, there exist $j \in [\![2,d]\!]$ and $\eta_2 \in \mathbb{R}$ such that $T|a_2\rangle = e^{i\eta_2}|b_j\rangle$. Repeating this argument, we obtain that there exist a permutation σ and phases $(\eta_i)_{i\in[\![1,d]\!]}$ such that

$$\forall i \in [\![1,d]\!], \ T|a_i\rangle = e^{i\eta_i}|b_{\sigma(i)}\rangle. \quad (51)$$

Now, we turn our attention to $(T|b_j\rangle)_{j\in[\![1,d]\!]}$. Suppose that there exists $j \in [\![1,d]\!]$ and $\theta_1 \in \mathbb{R}$ such that $T|b_1\rangle = e^{i\theta_1}|b_j\rangle$. Then, there exists $i \in [\![1,d]\!]$ such that $\sigma(i) \neq j$. Thus,

$$\langle a_i|b_1\rangle = \langle a_i|T^{\dagger}T|b_1\rangle = e^{i(\theta_1-\eta_i)}\langle b_{\sigma(i)}|b_j\rangle = 0, \quad (52)$$

which contradicts $m_{\mathcal{A},\mathcal{B}} > 0$. Thus, $T\mathcal{B}T^{\dagger} = \mathcal{A}$. As above, there exist a permutation τ and phases $(\theta_j)_{j\in[\![1,d]\!]} \in \mathbb{R}^d$ such that

$$\forall j \in [\![1,d]\!], \ T|b_j\rangle = e^{i\theta_j}|a_{\tau(j)}\rangle. \quad (53)$$

This proves that we are in case *2*. $\qquad\square$

Proof. Let $U \in \mathcal{W}(d)$. Suppose that there exists a KD-positivity preserving unitary T. By Lemma 10, we have to consider two cases. In case *1*, there exist two permutations (σ, τ) and $2d$ phases $(\eta_i)_{i\in[\![1,d]\!]}, (\theta_j)_{j\in[\![1,d]\!]}$ such that

$$\forall i \in [\![1,d]\!], T|a_i\rangle = e^{i\eta_i}|a_{\sigma(i)}\rangle$$
$$\forall j \in [\![1,d]\!], T|b_j\rangle = e^{i\theta_j}|b_{\tau(j)}\rangle. \quad (54)$$

Using Eq.(54), we obtain that for all $(j, k) \in [\![1,d]\!]^2$:

$$e^{i\eta_k}\langle a_k|b_j\rangle = e^{i\theta_j}\langle a_{\sigma(k)}|b_{\tau(j)}\rangle. \quad (55)$$

Suppose that $(\sigma, \tau) \neq (\mathrm{id}, \mathrm{id})$. Without loss of generality, we can suppose that $\sigma \neq \mathrm{id}$. Thus, there exists $p \in [\![1,d]\!]$ such that $\sigma(p) \neq p$. Relations (55) then imply that the set of unitary matrices $U \in \mathcal{W}(d)$ for which the unitary T is KD-positivity preserving is included in $\mathscr{Z}(f_{\sigma,\tau})$ where

$$f_{\sigma,\tau} : U \in \mathcal{U}(d) \rightarrow |U_{p1}|^2 - \left|U_{\sigma(p)\tau(1)}\right|^2 \in \mathbb{R}. \quad (56)$$

We recall that $\mathscr{Z}(f_{\sigma,\tau})$ denotes the set of zeroes of the function $f_{\sigma,\tau}$. Note that $f_{\sigma,\tau}$ is a polynomial

in the real and imaginary parts of the elements of the matrix.

We now construct a unitary matrix $U_{\star,1}$ such that $f_{\sigma,\tau}(U_{\star,1}) \neq 0$. We define two orthonormal vectors of $\mathbb{C}^d$ given by:

$$\widetilde{u}_\star^1 = \frac{1}{\sqrt{d(d-1)}} \left(1 \cdots 1 - (d-1) 1 \cdots 1\right)^{\mathrm{T}} \quad (57)$$

and

$$\widetilde{u}_\star^2 = \frac{1}{\sqrt{d}} \left(1 \cdots 1\right)^{\mathrm{T}} \quad (58)$$

where the coefficient $-\sqrt{\frac{d-1}{d}}$ is in place p. By the Gram-Schmidt process, there exists $(\widetilde{u}_\star^j)_{j \in [\![3,d]\!]}$ such that $(\widetilde{u}_\star^j)_{j \in [\![1,d]\!]}$ is an orthonormal basis of $\mathbb{C}^d$. We now define $U_{\star,1}$ to be

$$\begin{cases} \left(\widetilde{u}_\star^1 \; \widetilde{u}_\star^{\tau(1)} \; \widetilde{u}_\star^3 \; \cdots \; \widetilde{u}_\star^{\tau(1)-1} \; \widetilde{u}_\star^2 \; \widetilde{u}_\star^{\tau(1)+1} \; \cdots \; \widetilde{u}_\star^d\right) & \text{if } \tau(1) \neq 1 \\ \left(\widetilde{u}_\star^1 \; \cdots \; \widetilde{u}_\star^d\right) & \text{if } \tau(1) = 1 \end{cases}$$
$$(59)$$

We then compute

$$f_{\sigma,\tau}(U_{\star,1}) = \begin{cases} \frac{d-2}{d} \neq 0 & \text{if } \tau(1) \neq 1 \\ \frac{2-d}{d(d-1)} \neq 0 & \text{if } \tau(1) = 1 \end{cases} . \quad (60)$$

Consequently, $\mathcal{Z}(f_{\sigma,\tau})$ is, by Proposition 2.4, a set of zero Haar measure. To conclude case *1*, suppose that $(\sigma,\tau) = (\mathrm{id},\mathrm{id})$. The relations (55) imply that

$$\forall (j,k) \in [\![1,d]\!]^2, e^{i\eta_j} = e^{i\theta_k}. \quad (61)$$

We thus get that, according to the definition of T given by equations (54), $\theta_j = \eta_1 \; [2\pi]$ and $\eta_j = \eta_1 \; [2\pi]$ for all $j \in [\![1,d]\!]$. Then, $T = e^{i\eta_1}\mathrm{id}_\mathcal{H}$.

In case *2*, there exist two permutations $(\sigma,\tau) \in \Sigma(d)$ and $2d$ phases $(\eta_i)_{i \in [\![1,d]\!]}$ and $(\theta_j)_{j \in [\![1,d]\!]}$ such that

$$\forall i \in [\![1,d]\!], T|a_i\rangle = e^{i\eta_i}|b_{\sigma(i)}\rangle,$$
$$\forall j \in [\![1,d]\!], T|b_j\rangle = e^{i\theta_j}|a_{\tau(j)}\rangle. \quad (62)$$

We obtain similarly that, for all $(j,k) \in [\![1,d]\!]$:

$$e^{i\eta_k}\langle a_k|b_j\rangle = e^{i\theta_j}\langle b_{\sigma(k)}|a_{\tau(j)}\rangle. \quad (63)$$

Note that there exists $p \in [\![1,d]\!]$ such that $\sigma(p) \neq 1$. Thus, relations (63) imply that the set of unitary matrices $U \in \mathcal{W}(d)$ for which the unitary T is KD-positivity preserving is included in $\mathcal{Z}(g_{\sigma,\tau})$ where

$$g_{\sigma,\tau} : U \in \mathcal{U}(d) \to |U_{p1}|^2 - \left|U_{\tau(1)\sigma(p)}\right|^2 \in \mathbb{R}. \quad (64)$$

Note that $g_{\sigma,\tau}$ is a polynomial in the real and imaginary parts of the elements of the matrix.

We now construct a unitary matrix $U_{\star,2}$ such that $g_{\sigma,\tau}(U_{\star,2}) \neq 0$. We define $U_{\star,2}$ to be

$$U_{\star,2} = \left(\widetilde{u}_\star^1 \; \widetilde{u}_\star^{\sigma(p)} \; \widetilde{u}_\star^3 \; \cdots \; \widetilde{u}_\star^{\sigma(p)-1} \; \widetilde{u}_\star^2 \; \widetilde{u}_\star^{\sigma(p)+1} \; \cdots \; \widetilde{u}_\star^d\right).$$
$$(65)$$

We then compute

$$g_{\sigma,\tau}(U_{\star,2}) = \frac{d-2}{d} \neq 0. \quad (66)$$

Consequently, $\mathcal{Z}(g_{\sigma,\tau})$ is, by Proposition 2.4, a set of zero Haar measure.

Consequently, the set of matrices U that admit a KD-preserving unitary that is not a global phase change is included in the set

$$\Lambda(d) = \mathcal{W}(d)^{\mathrm{C}} \cup \bigcup_{\substack{(\sigma,\tau) \\ (\sigma,\tau) \neq (\mathrm{id},\mathrm{id})}} \mathcal{Z}(f_{\sigma,\tau}) \cup \bigcup_{(\sigma,\tau)} \mathcal{Z}(g_{\sigma,\tau}).$$
$$(67)$$

As $\mathcal{W}(d)$ is of full Haar measure, we obtain that $\Lambda(d)$ is of zero Haar measure. Defining $\mathcal{V}(d)$ to be the complementary set of $\Lambda(d)$, we conclude the proof. $\qquad \square$

This proposition thus proves that, with probability one in any dimension $d \geqslant 3$, the only KD-positivity preserving unitaries are given by $(e^{i\theta}\mathrm{id}_\mathcal{H})_{\theta \in \mathbb{R}}$ and these unitaries act trivially on $\mathcal{A} \cup \mathcal{B}$.

D KD distribution on a n-qubit system

We have shown that, with probability one, the set of KD-positive states $\mathcal{E}_{\mathrm{KD+}}$ is an explicitly known polytope with $2d$ vertices. In this section we show through an example that in specific cases, the structure of $\mathcal{E}_{\mathrm{KD+}}$ can be much more involved.

We consider for that purpose a KD distribution naturally associated to a n-qubit system. For $n = 1$, we consider the computational basis $\mathbf{A}_1 = (|0\rangle, |1\rangle)$ and the basis $\mathbf{B}_1 = (|+\rangle, |-\rangle)$. The transition matrix between these two bases is the real Hadamard matrix, which is also the DFT matrix. The only KD-positive pure states are in that case the bases states.

If $n \in \mathbb{N}^*$, we consider the KD representation of a system of n qubits given by the following two

tensor product bases :

$$\mathbf{A}_n = \left\{ \bigotimes_{i=1}^{n} |\varepsilon_i\rangle \mid \varepsilon_i \in \{0,1\} \right\}$$

$$\mathbf{B}_n = \left\{ \bigotimes_{i=1}^{n} |\eta_i\rangle \mid \eta_i \in \{-,+\} \right\}. \qquad (68)$$

This KD representation can be viewed as being associated to the group $(\mathbb{Z}_2)^n$ and its Heisenberg group as described in [53]. This allows one to identify the set of KD-positive pure states, which we denote by $\mathcal{E}_{\mathrm{KD},n}^{\mathrm{pure}}$. The case $n = 2$ is explicitly worked out below. One can then determine the cardinality $|\mathcal{E}_{\mathrm{KD},n}^{\mathrm{pure}}|$ of $\mathcal{E}_{\mathrm{KD},n}^{\mathrm{pure}}$:

$$\left| \mathcal{E}_{\mathrm{KD},n}^{\mathrm{pure}} \right| = \sum_{p=0}^{n} 2^n \frac{\prod_{k=1}^{p} (2^n - 2^{k-1})}{\prod_{k=1}^{p} (2^p - 2^{k-1})}. \qquad (69)$$

One can prove the following inequalities :

$$2^{\frac{n^2}{4}} \leqslant \left| \mathcal{E}_{\mathrm{KD},n}^{\mathrm{pure}} \right| \leqslant 2^{(n+1)^2+1}. \qquad (70)$$

This proves that the KD distribution naturally associated to n qubits has $O(d^{\log_2(d)})$ KD-positive pure states, where $d = 2^n$ is the dimension of the underlying Hilbert space. This observation illustrates the complexity of the structure of $\mathrm{conv}\left(\mathcal{E}_{\mathrm{KD},n}^{\mathrm{pure}}\right)$ in comparison with the simplest situation that occurs with probability 1 where $\left| \mathcal{E}_{\mathrm{KD}+}^{\mathrm{pure}} \right| = |\mathcal{A} \cup \mathcal{B}| = 2d$.

When $n = 2$, we list in table 1 the 20 KD-positive pure states. In this case, the underlying Hilbert space is of dimension 4. The number of KD-positive pure states in the simplest situation would be 8, occurring with probability 1.

$|00\rangle$	$|01\rangle$	$|10\rangle$	$|11\rangle$
$\frac{|00\rangle+|01\rangle}{\sqrt{2}}$	$\frac{|00\rangle-|01\rangle}{\sqrt{2}}$	$\frac{|10\rangle+|11\rangle}{\sqrt{2}}$	$\frac{|10\rangle-|11\rangle}{\sqrt{2}}$
$\frac{|00\rangle+|10\rangle}{\sqrt{2}}$	$\frac{|00\rangle-|10\rangle}{\sqrt{2}}$	$\frac{|01\rangle+|11\rangle}{\sqrt{2}}$	$\frac{|01\rangle-|11\rangle}{\sqrt{2}}$
$\frac{|00\rangle+|11\rangle}{\sqrt{2}}$	$\frac{|00\rangle-|11\rangle}{\sqrt{2}}$	$\frac{|01\rangle+|10\rangle}{\sqrt{2}}$	$\frac{|01\rangle-|10\rangle}{\sqrt{2}}$
$|++\rangle$	$|+-\rangle$	$|-+\rangle$	$|--\rangle$

Table 1: List of KD-positive pure states for the KD representation associated with 2 qubits.

We note that, as shown in [53], for the 2 qubit KD-representation, we have that :

$$\mathrm{conv}\left(\mathcal{E}_{\mathrm{KD},2}^{\mathrm{pure}}\right) \subsetneq \mathcal{E}_{\mathrm{KD}+}. \qquad (71)$$

In other words, in that case, the set of KD-positive states may not be a polytope and contains exotic states. We expect this to be true in all dimensions.

E Witnesses of $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$

In this section, we show how the simple structure of $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$ can be used to construct estimates of the distance of a state $\rho_\star$ to $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$. In [50], the authors proved that the facets of $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$ are given by $\mathrm{conv}\left(\mathcal{A}_i \cup \mathcal{B}_j\right)$ for all $(i,j) \in [\![1,d]\!]^2$ where

$$\mathcal{A}_i = \{|a_\ell\rangle \langle a_\ell| \mid \ell \in [\![1,d]\!] \backslash \{i\}\}$$
$$\mathcal{B}_j = \{|b_k\rangle \langle b_k| \mid k \in [\![1,d]\!] \backslash \{j\}\}.$$

For a fixed $(i,j) \in [\![1,d]\!]^2$, the bounding plane associated to the facet is given by

$$|a_p\rangle \langle a_p| + \qquad (72)$$
$$\mathrm{span}_{\mathbb{R}} \{|a_\ell\rangle \langle a_\ell| - |a_p\rangle \langle a_p|, |b_k\rangle \langle b_k| - |a_p\rangle \langle a_p|\}_{\substack{\ell \in [\![1,d]\!]\backslash\{i\} \\ k \in [\![1,d]\!]\backslash\{j\}}}$$

where $p = \min [\![1,d]\!] \backslash \{i\}$. This set is an affine hyperplane inside the set $E = \mathrm{span}_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B}) \cap \{F \in \mathcal{S}_d(\mathbb{C}) \mid \mathrm{Tr}(F) = 1\}$ and a bounding plane of $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$. Thus, there exists $\eta_{ij} \in E$ such that :

$$\forall \rho \in \mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right), \mathrm{Tr}(\eta_{ij}\rho) \leqslant \mathrm{Tr}(\eta_{ij}|a_p\rangle \langle a_p|). \qquad (73)$$

Such witnesses can be explicitly constructed thanks to the simple structure of $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$. As an example, when the transition matrix is the DFT matrix in dimension 3, the witnesses are given by

$$\forall (i,j) \in [\![1,3]\!]^2, \eta_{ij} = \mathbb{I}_3 - |a_i\rangle \langle a_i| - |b_j\rangle \langle b_j|. \qquad (74)$$

Suppose that such a witness η_{ij} is given and that $\rho_\star$ is a quantum state satisfying that $\mathrm{Tr}(\eta_{ij}\rho_\star) > \mathrm{Tr}(\eta_{ij}|a_p\rangle \langle a_p|)$. Then,

$$0 < \frac{\mathrm{Tr}\left(\eta_{ij}\left(\rho_\star - |a_p\rangle \langle a_p|\right)\right)}{||\eta_{ij}||_2} \leqslant \mathrm{d}\left(\rho_\star, \mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)\right) \qquad (75)$$

where $\mathrm{d}\left(\rho_\star, A\right)$ denotes the distance of $\rho_\star$ to the set A.

One can then use Theorem 1 from the paper in the following manner. If the transition matrix is in $\Gamma \cap \Omega$, we know $\mathcal{E}_{\mathrm{KD}+} = \mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right)$, so that we obtain a lower bound on the distance of $\rho_\star$ to $\mathcal{E}_{\mathrm{KD}+}$, given by

$$\mathrm{d}\left(\rho_\star, \mathcal{E}_{\mathrm{KD}+}\right) \geqslant \frac{\mathrm{Tr}(\eta_{ij}\rho_\star) - \mathrm{Tr}(\eta_{ij}|a_p\rangle \langle a_p|)}{||\eta_{ij}||_2} > 0. \qquad (76)$$

Obtaining such a bound whenever $\mathrm{conv}\left(\mathcal{A} \cup \mathcal{B}\right) \subsetneq \mathcal{E}_{\mathrm{KD}+}$ is considerably more

difficult because in that case $\mathcal{E}_{\mathrm{KD}+}$ may not be a polytope, and even when it is, it may be very difficult to determine its bounding planes.

F Simulability

A KD-positive quantum circuit is composed of a KD-positive input state, KD-positive unitary gates and KD-positive measurement elements. KD-positive states and KD-positive measurement elements were defined in the main text. KD-positive gates are defined as follows. Let T be a unitary on $\mathcal{H}$, then there exist $K_{ij:i'j'} \in \mathbb{C}$ so that

$$Q(T\rho T^\dagger)_{ij} = \sum_{i'j'} K_{ij:i'j'} Q_{i'j'}(\rho), \qquad (77)$$

with

$$\sum_{ij} K_{ij:i'j'} = 1, \forall i', j'.$$

A unitary T is said to be KD-positive if all $K_{ij:i'j'} \geqslant 0$. As explained in the Discussion section, such circuits can be efficiently simulated classically.

To see how Theorem 1 impacts the problem of efficient simulability, we first note that, clearly, a KD-positive unitary is KD-positivity preserving. Hence Theorem 1 implies that, with probability one, the only quantum circuits that are KD-positive are those that have as input states (convex mixtures of) the basis states, followed immediately by a KD-positive measurement.