



Wagner, R., Schwartzman-Nowik, Z., Paiva, I. L., Te'eni, A., Ruiz-Molero, A., Soares Barbosa, R., Cohen, E., & Galvão, E. F. (2024). Quantum circuits for measuring weak values, Kirkwood–Dirac quasiprobability distributions, and state spectra. *Quantum Science and Technology*, 9(1), Article 015030. <https://doi.org/10.1088/2058-9565/ad124c>

Peer reviewed version

License (if available):
CC BY

Link to published version (if available):
[10.1088/2058-9565/ad124c](https://doi.org/10.1088/2058-9565/ad124c)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via IOP Publishing at <https://doi.org/10.1088/2058-9565/ad124c>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Quantum circuits for measuring weak values, Kirkwood–Dirac quasiprobability distributions, and state spectra

Rafael Wagner,^{1,2,*} Zohar Schwartzman-Nowik,^{3,4} Ismael L. Paiva,⁵ Amit Te'eni,⁴
Antonio Ruiz-Molero,^{1,6} Rui Soares Barbosa,¹ Eliahu Cohen,⁴ and Ernesto F. Galvão^{1,7,†}

¹*INL – International Iberian Nanotechnology Laboratory, Braga, Portugal*

²*Centro de Física, Universidade do Minho, Braga, Portugal*

³*School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel*

⁴*Faculty of Engineering and the Institute of Nanotechnology and Advanced Materials, Bar Ilan University, Ramat Gan, Israel*

⁵*H. H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom*

⁶*Departamento de Informática, Universidade do Minho, Braga, Portugal*

⁷*Instituto de Física, Universidade Federal Fluminense, Niterói – RJ, Brazil*

(Dated: January 4, 2024)

Weak values and Kirkwood–Dirac (KD) quasiprobability distributions have been independently associated with both foundational issues in quantum theory and advantages in quantum metrology. We propose simple quantum circuits to measure weak values, KD distributions, and spectra of density matrices without the need for post-selection. This is achieved by measuring unitary-invariant, relational properties of quantum states, which are functions of Bargmann invariants, the concept that underpins our unified perspective. Our circuits also enable experimental implementation of various functions of KD distributions, such as out-of-time-ordered correlators (OTOCs) and the quantum Fisher information in post-selected parameter estimation, among others. An upshot is a unified view of nonclassicality in all those tasks. In particular, we discuss how negativity and imaginarity of Bargmann invariants relate to set coherence.

CONTENTS

1. Introduction	2	5.1. Negativity, imaginarity, and set coherence	18
2. Preliminaries	3	5.1.1. Negativity and imaginarity as coherence witnesses	18
2.1. Weak values	3	5.1.2. Nonclassicality in weak values and KD distributions as relational nonclassicality	18
2.2. Kirkwood–Dirac quasiprobability distribution	4	5.1.3. Set coherence beyond negativity and imaginarity	19
2.2.1. Post-selected quantum Fisher information	4	5.2. Can negativity and imaginarity be witnessed from overlaps?	19
2.2.2. Out-of-time-ordered correlators	5	5.2.1. Dimension-specific KD negativity and imaginarity from overlaps in the whole phase space	19
2.3. Bargmann invariants and how to measure them	6	5.2.2. Inferring higher-order Bargmanns from overlaps: the three-state scenario	19
2.4. Nonclassicality of Bargmann invariants	7	5.2.3. Overlaps are, in general, not enough for negativity and imaginarity	20
3. Bargmann invariants as a unifying concept	8	5.2.4. Negativity and overlaps for real amplitude states	21
4. Quantum circuits for measuring weak values and Kirkwood–Dirac quasiprobability distributions	9	6. Discussion and future directions	22
4.1. Kirkwood–Dirac quasiprobability	9	Acknowledgments	23
4.2. Weak values	11	References	23
4.3. Post-selected quantum Fisher information	11	A. Formal comparison between standard weak measurement and circuit protocol	27
4.4. Out-of-time-ordered correlators	13	1. Same number of samples for numerator and denominator	27
4.5. Comparison with other methods for estimating the Kirkwood–Dirac quasiprobability distribution	13	2. Different number of samples for numerator and denominator	28
4.6. Estimating the spectrum of a quantum state	15	B. Sample and measurement complexity of estimating	
5. Minimal experimental conditions for witnessing nonclassicality	17		

* rafael.wagner@inl.int

† ernesto.galvao@inl.int

the spectrum using the cycle test	30
1. Proof of Theorem 6	30
2. Numerical analysis for the Faddeev–LeVerrier algorithm	31
a. Learning the largest eigenvalue	31
C. Proof of Lemma 11	32

1. INTRODUCTION

Two concepts have profoundly impacted quantum foundations, metrology, and thermodynamics: weak values, introduced in the seminal work of Aharonov, Albert, and Vaidman [1], and the quasiprobability distribution introduced by Kirkwood and Dirac [2, 3]. These were recently connected with one another [4], and both can be experimentally measured using weak measurement schemes involving pre- and post-selection of carefully chosen observables [5]. Despite such great impact, investigations of measurement schemes that can be implemented using currently relevant quantum circuit architectures without post-selection and without weak coupling are only now beginning to appear [5, 6].

Another important carrier of nonclassical information about a quantum state ρ , seemingly unrelated to those discussed above, is the spectrum $\text{Spec}(\rho)$. Spectral properties are well known to capture nonclassical features of states such as entanglement [7] and basis-dependent coherence [8]. Various nonclassicality properties can be learned from the spectrum of a quantum state, or from non-linear functions thereof, such as univariate traces of the form $\text{Tr}(\rho^n)$ for some integer $n \geq 1$. In quantum information, knowledge of such univariate traces has recently been used to witness non-stabilizerness (i.e., “magic”) of quantum states [9], or in subroutines for variational quantum eigensolvers [10].

In this article, we describe a unified framework that enables expressing Kirkwood–Dirac (KD) quasiprobability functions, weak values, and univariate traces $\text{Tr}(\rho^n)$ in terms of more fundamental quantities, known as *Bargmann invariants*. These invariants fully characterize the relational properties of a set of quantum states, that is, all the properties that remain invariant under the application of the same unitary to all states in the set. The simplest non-trivial Bargmann invariant is the overlap $\text{Tr}(\rho\sigma)$ between two states ρ and σ , and it is also the easiest to probe experimentally [11].

The starting point for our conceptual unification is the observation that KD distributions, weak values, univariate traces, and many other constructions of interest can be written as functions of Bargmann invariants. From this simple yet powerful observation, we draw foundational and practical implications. On the one hand, the connection with Bargmann invariants provides a unified view for studying nonclassicality of sets of states for all use cases considered in this work. On the other, it opens the door for the use of a family of circuits that measure Bargmann

invariants to estimate *any* function of Bargmann invariants, including all those mentioned above.

A crucial notion of nonclassicality we consider here is that of *set coherence*, a basis-independent notion of coherence proposed in Ref. [12], whereby a set of quantum states is said to be coherent if and only if it is not pairwise commuting. We show how this form of nonclassicality connects with the theory of Bargmann invariants.

KD distributions can exhibit a different kind of nonclassicality by taking on negative or even non-real complex values. Similarly, the weak value of an observable is considered nonclassical when it does not lie within the observable’s spectrum. As it turns out, the nonclassicality of KD distributions and weak values, which underpins their relevance as quantum information resources, is a relational property described by Bargmann invariants. We show that, in general, learning the values of higher-order invariants (beyond overlaps) is *necessary* to assess nonclassicality. We also show how further assumptions can be used to ascertain nonclassicality using overlaps only.

More pragmatically, the nonclassicality of weak values and KD distributions has been linked to quantum advantage in metrology [13] and to quantification of quantum information scrambling [14, 15]. Our framework describing those quantities via Bargmann invariants allows us to use recently proposed cycle test circuits [16, 17] to directly measure weak values and KD distributions. This enables quantum circuit measurements of functions associated with multiple applications; see Fig. 1 for a conceptual scheme describing our main contributions. We compare the performance of these circuits with the usual strategies for performing weak measurements and quantum state tomography. We also show how out-of-time-ordered correlators (OTOCs), used to quantify information scrambling in quantum dynamics, can be measured using the same type of circuit. We show that this is also true for the quantum Fisher information obtained in post-selected parameter estimation.

We recall how these circuits can also estimate the spectrum of a given d -dimensional quantum state ρ via estimation of $\text{Tr}(\rho^n)$ for $n = 1, \dots, d$. This circuit architecture is well known [18], as is the fact that one can learn spectral properties by classical post-processing using the Faddeev–LeVerrier algorithm [19]. Besides our unified perspective that such quantities correspond to specific Bargmann invariants, we provide new sample complexity arguments and numerical simulations suggesting that learning the spectrum via this technique is useful for near-term applications, and is significantly simpler when compared with other efficient techniques.

We expect our contributions to open the path to the quantification of nonclassicality and many other applications, for example in the characterization of coherence, entanglement, and quantum computational advantage.

Outline. This paper is organized as follows. In Sec. 2 we introduce the basic definitions and properties of KD distributions, weak values arising from weak measure-

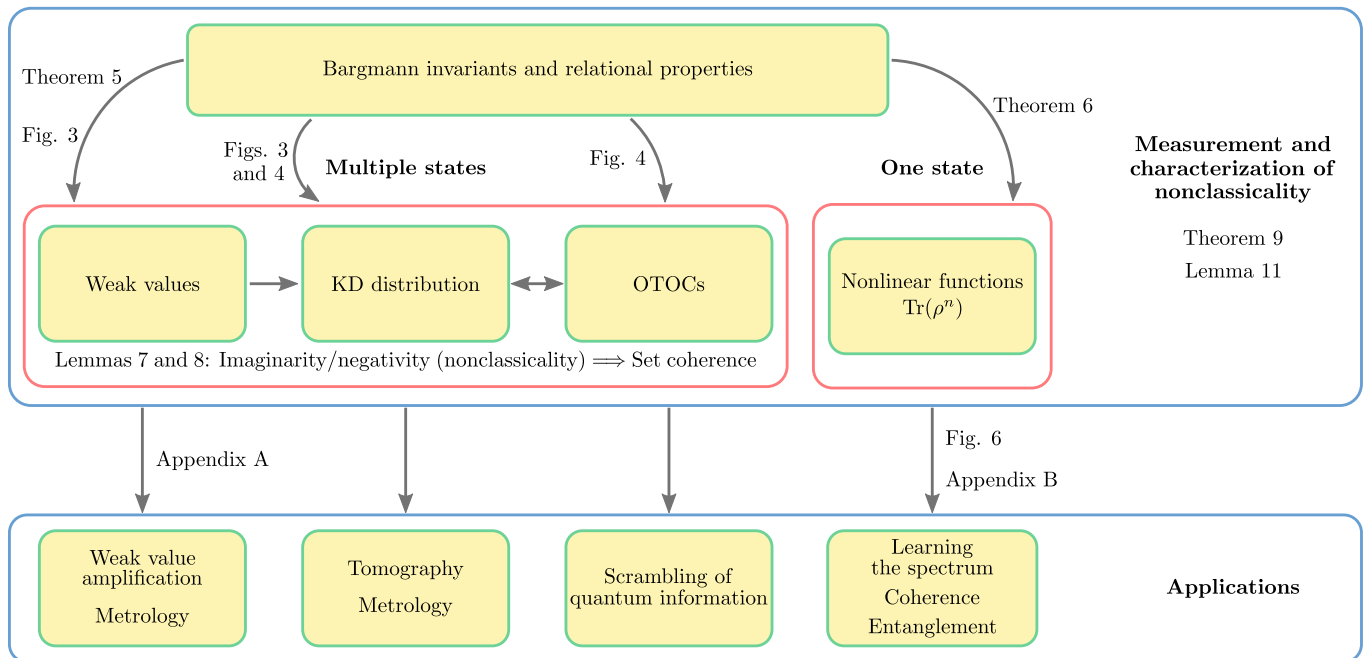


Figure 1. **Conceptual scheme describing the content and main contributions and their inter-relations.** We show how weak values, the Kirkwood–Dirac (KD) quasiprobability distribution, and out-of-time-ordered correlators (OTOCs) can all be expressed as unitary-invariant quantities known as Bargmann invariants in Sec. 3. Circuits to measure them are discussed in Sec. 4. Moreover, results on the nonclassicality of Bargmann invariants, such as those obtained in Sec. 5, enable a better understanding and quantification of quantum advantage, where we show that negativity and imaginarity of Bargmann invariants require set coherence.

ments, extended KD distributions, and Bargmann invariants. In Sec. 3 we then propose Bargmann invariants as the unifying concept connecting all the constructions just described. We proceed to discuss the pragmatic quantum circuit measurement schemes in Sec. 4, and the foundational implications for analyzing nonclassicality in Sec. 5. Specifically, in Sec. 4, we present our main results, describing how one can experimentally access weak values, KD distributions, and state spectra via the measurement of Bargmann invariants. We compare the efficiency and experimental particularities of estimating weak values using the standard weak measurement approach against our proposed quantum circuits. We also compare our scheme with other approaches to estimate the KD distribution. Moreover, we study how the measurement of higher-order Bargmann invariants enables the estimation of the quantum Fisher information obtained in post-selected parameter amplification, OTOCs, and the spectrum of a density matrix. In Sec. 5, we start by formally connecting Bargmann nonclassicality with set coherence, hence connecting KD and weak value nonclassicality with various other notions and results in the literature. We then study how nonclassical properties of invariants help characterize nonclassical quasiprobability distributions. We finish this section by analyzing minimal conditions for characterizing the nonclassicality of KD distributions and related quantities. We conclude in Sec. 6 with an outlook on possible future work.

2. PRELIMINARIES

In this section, we recall the necessary background on weak values, (extended) KD distributions, and Bargmann invariants.

2.1. Weak values

Consider a system prepared (pre-selected) in state $|\psi\rangle$, on which one performs a weak measurement of an observable A , i.e., a measurement associated with a small coupling strength (when compared to the standard deviation of the measuring pointer). Although this measurement has the apparent downside of not generating a significant average shift of the pointer compared to its standard deviation, it has the benefit of causing little disturbance to the system of interest. Finally, let the system be post-selected in the state $|\phi\rangle$. It turns out that for this post-selection, the average shift of the pointer used for the intermediate weak measurement is proportional to

$$A_w := \frac{\langle \phi | A | \psi \rangle}{\langle \phi | \psi \rangle}, \quad (1)$$

which is known as the weak value of A [1]. A_w arises from a first-order approximation of a certain Taylor expansion [20]. This quantity may lie outside the spectrum of the measured operator, either by having a non-zero imaginary

part or by having a real part outside the spectrum of A , in which case the weak value is said to be *anomalous*. In fact, with appropriate choices of pre- and post-selections, anomalous weak values can have arbitrarily large real or imaginary parts. For future reference, we state this as a definition.

Definition 1 (Weak value nonclassicality). Let $A : \mathcal{H} \rightarrow \mathcal{H}$ be a Hermitian operator and $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ two non-orthogonal vectors in a finite-dimensional Hilbert space \mathcal{H} . We say that the weak value $A_w = \langle \phi | A | \psi \rangle / \langle \phi | \psi \rangle$ is *nonclassical* or *anomalous* if $A_w \notin \text{Spec}(A)$, where $\text{Spec}(A)$ denotes the spectrum of A .

Due to the appearance of such anomalous values and the usual way in which these quantities are experimentally obtained, via weak measurements, the quantum nature of weak values has been questioned since their introduction; see Ref. [21] for an overview. More recently, this situation seems to have changed, with the nonclassicality of certain weak values being more well established. Important theoretical results contributing to this understanding state that anomalous weak values obtained from weak measurements constitute proofs of generalized contextuality [22, 23].

Since weak measurements cause little disturbance to the measured system, they also do not extract much information about it. As a result, protocols involving weak measurements typically require large ensembles in order to decrease the variance associated with them; see, however, the recent demonstration in Ref. [24]. In spite of this, weak values have found various practical applications. In particular, through a method known as *weak value amplification*, large weak values are used for realizing extremely sensitive measurements [25–37]. This technique is especially helpful in the presence of technical noise or detector saturation [29, 32].

It is also noteworthy that weak values can be used to estimate wave functions, i.e., a state’s representation in position coordinates $\Psi(x) := \langle x | \psi \rangle$, or specific density matrix terms [38–42]. Observe, however, that the present work focuses exclusively on the finite-dimensional setting.

Considering the spectral decomposition of the observable A , $A = \sum_{a \in \text{Spec}(A)} a |a\rangle\langle a|$, with $\text{Spec}(A)$ denoting the spectrum of A , we see that

$$A_w = \sum_{a \in \text{Spec}(A)} a \frac{\langle \phi | a \rangle \langle a | \psi \rangle}{\langle \phi | \psi \rangle} = \sum_{a \in \text{Spec}(A)} a \frac{\langle \phi | a \rangle \langle a | \psi \rangle \langle \psi | \phi \rangle}{|\langle \phi | \psi \rangle|^2} \quad (2)$$

since, by construction, $\langle \psi | \phi \rangle \neq 0$. Other generalizations of weak values, such as sequential weak values or joint weak values, provide valuable information regarding the behavior of a given state with respect to two—possibly incompatible—observables [43–46]. It so happens that the numerators in the expression for the weak value above, or those of joint weak values, correspond to values of the KD quasiprobability distribution at the phase-space point determined by the pre- and post-selected states [47]. These have their own relevance in quantum information

theory [4]. We now proceed to review some basic facts on the KD quasiprobability distribution.

2.2. Kirkwood–Dirac quasiprobability distribution

Consider a finite discrete phase space $I \times F$, associated with quantum states $\{|i\rangle\}_{i \in I}$ and $\{|f\rangle\}_{f \in F}$ that form two orthonormal bases of a d -dimensional Hilbert space \mathcal{H} . The *KD quasiprobability distribution* for a given state $\rho \in \mathcal{D}(\mathcal{H})$ ¹ is the function $\xi(\rho|\cdot) : I \times F \rightarrow \mathbb{C}$ given by

$$\xi(\rho|i, f) := \langle f | i \rangle \langle i | \rho | f \rangle \quad (3)$$

at each phase-space point $(i, f) \in I \times F$. Under the assumption that $\langle f | i \rangle \neq 0$ for all i, f , the KD distribution provides complete information about the state ρ . This is because, in this case, the values of the KD distribution at phase space points are the decomposition coefficients with respect to the orthonormal basis $\{|i\rangle\langle f| / \langle f | i \rangle\}_{(i, f) \in I \times F}$ of $\mathcal{B}(\mathcal{H})$, the space of bounded operators on \mathcal{H} [48, 49]. The KD distribution has been experimentally measured in Refs. [50–54].

The KD distribution can be extended to be well defined for a larger number of bases, or even general projection-valued measures (PVMs). For a family of PVMs $M_i = \{\Pi_k^i\}_{k \in K_i}$, with $i = 1, \dots, n$, the *extended KD distribution* for a state ρ at a phase space point $(k_1, \dots, k_n) \in K_1 \times \dots \times K_n$ reads

$$\xi(\rho|k_1, \dots, k_n) = \text{Tr}(\Pi_{k_1}^1 \Pi_{k_2}^2 \dots \Pi_{k_n}^n \rho). \quad (4)$$

The following definition formalizes what is understood by nonclassicality in the context of KD distributions.

Definition 2 (KD nonclassicality). Fix $\{|i\rangle\}_{i \in I}$, $\{|f\rangle\}_{f \in F}$ two reference orthonormal bases of a finite-dimensional Hilbert space \mathcal{H} . We say that the KD quasiprobability distribution $\xi(\rho|\cdot)$ associated with a state $\rho \in \mathcal{D}(\mathcal{H})$ is *nonclassical* when there exists a phase-space point $(i, f) \in I \times F$ such that $\xi(\rho|i, f) \notin [0, 1]$.

For extended KD quasiprobability distributions, we similarly define KD nonclassicality to mean taking on values outside the unit interval.

2.2.1. Post-selected quantum Fisher information

The KD distribution and its extended variants are key constructions for witnessing many nonclassical properties of quantum dynamics [4, 5, 55]. Recent developments have shown that the KD distribution is deeply connected to the quantum Fisher information [5]. The latter provides

¹ We denote by $\mathcal{D}(\mathcal{H})$ the set of all quantum states in the Hilbert space \mathcal{H} .

the optimal rate with which one can learn some (set of) parameter(s) θ encoded in quantum states ρ_θ via the Cramér–Rao bound [56, 57]. For good introductory works on Fisher information theory and quantum or classical estimation theory, see, e.g., Refs. [58–60]. It is therefore interesting from a foundational perspective to seek ways of estimating the quantum Fisher information, an important task for the noisy intermediate-scale quantum (NISQ) era of technological capabilities [61–63]. Indeed, this is currently an active research topic [64, 65].

Let us focus on a recently found connection with quantum metrology. In Ref. [13], nonclassicality of the KD distribution was connected with metrological advantage in parameter estimation. In a prepare-and-measure metrological scenario on a d -dimensional Hilbert space \mathcal{H} , an initial pure state $\rho = |\psi\rangle\langle\psi|$ is prepared, information about a parameter θ is encoded through the action of a unitary $U = e^{-i\theta I}$ with $I = \sum_i \lambda_i |i\rangle\langle i|$, evolving to a final state ρ_θ^{ps} , which is successfully post-selected with respect to a projector $F = \sum_f |f\rangle\langle f|$ with probability $p_\theta^{\text{ps}} = \text{Tr}(F\rho_\theta)$. The quantum Fisher information \mathcal{I}^{ps} captures the optimal rate one can learn about the parameter θ in this setup. The following relationship was established in Ref. [13] between the extended KD distribution and the quantum Fisher information of post-selected states:

$$\mathcal{I}^{\text{ps}} = 4 \sum_{i,i',f} \lambda_i \lambda_{i'} \frac{\xi(\rho_\theta|i, i', f)}{p_\theta^{\text{ps}}} - 4 \left| \sum_{i,i',f} \lambda_i \frac{\xi(\rho_\theta|i, i', f)}{p_\theta^{\text{ps}}} \right|^2. \quad (5)$$

The extended KD distribution for this scenario takes the form

$$\xi(\rho_\theta|i, i', f) = \langle i|\rho_\theta|i'\rangle\langle i'|f\rangle\langle f|i\rangle. \quad (6)$$

For a generalization of Eq. (5) to the case of multi-parameter estimation and non-ideal observables F , see Refs. [66, 67]. Importantly, Eq. (5) holds only for pure states $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, and was later generalized in Ref. [68]. Moreover, in Ref. [68], the relationship displayed in Eq. (5) was used to unravel metrological advantages of post-selected parameter amplification for the estimation of a small parameter θ . Also, in Ref. [69], the efficiency of the post-selection advantage was bounded, formally relating the growth of the post-selected version of Fisher information with the factor p_θ^{ps} .

As a remark, the KD distribution also appears as part of the quantum Fisher information without the need for post-selection in the so-called linear response regime, as discussed in Ref. [5].

2.2.2. Out-of-time-ordered correlators

In Ref. [4] it was shown that the extended KD distribution underlies the out-of-time-ordered correlators (OTOCs), commonly used to witness the scrambling of

quantum information. These are given by the expression

$$\text{OTOC}(t) := \text{Tr}(W^\dagger(t)V^\dagger W(t)V\rho). \quad (7)$$

Intuitively, OTOCs witness scrambling of information in the following situation. Consider a many-body system ρ and two observables $W = W(0)$ and V acting over distant regions of the system; the canonical example considers local observables acting on the initial and final spins in a one-dimensional lattice of size $N \gg 1$. The function $\text{OTOC}(t)$ witnesses the noncommutativity of the observable $W(t) = U(t)^\dagger W U(t)$, the unitary evolution of W in the Heisenberg picture by the one-parameter group $U(t)$, with respect to V . It thus signals the delocalization of quantum information. The OTOC can be written as [4]

$$\sum_{v_1, w_2, v_2, w_3} v_1 w_2 v_2^* w_3^* \text{Tr}(\Pi_{w_3}^{W(t)} \Pi_{v_2}^V \Pi_{w_2}^{W(t)} \Pi_{v_1}^V \rho), \quad (8)$$

where v and w range over the eigenvalues of $V = \sum_{v \in \text{Spec}(V)} v \Pi_v^V$ and $W = \sum_{w \in \text{Spec}(W)} w \Pi_w^W$ associated to eigenprojectors Π_v^V and Π_w^W , respectively. We also have that $\Pi_w^{W(t)} = U(t)^\dagger \Pi_w^W U(t)$. The choice of labels made in Eq. (8) comes from the interpretation of Ref. [4], where the OTOC is described by a quasiprobability naturally arising from comparing the forward and backward evolution of the system under a unitary action.

The coarse-graining hides the possible degeneracy in the spectral decomposition of the observables. In fact, if we take such degeneracy into account, we have that each observable is described by

$$\Pi_v^V = \sum_{\lambda_v} |v, \lambda_v\rangle\langle v, \lambda_v|, \quad (9)$$

where the eigenspace associated with eigenvalue v is described by the complete set of vectors $|v, \lambda_v\rangle$ with λ_v ranging over the degeneracy parameters. We find that the OTOC can be described in a more fine-grained way by

$$\text{OTOC}(t) = \sum_{\substack{(v_1, \lambda_{v_1}), (v_2, \lambda_{v_2}), \\ (w_2, \lambda_{w_2}), (w_3, \lambda_{w_3})}} v_1 w_2 v_2^* w_3^* \tilde{A}_\rho, \quad (10)$$

where $\tilde{A}_\rho = \tilde{A}_\rho(v_1, \lambda_{v_1}; w_2, \lambda_{w_2}; v_2, \lambda_{v_2}; w_3, \lambda_{w_3})$ is the quasiprobability behind the OTOC [4]. This is an extended KD distribution given at each phase-space point by

$$\tilde{A}_\rho = \langle w_3, \lambda_{w_3} | U | v_2, \lambda_{v_2} \rangle \langle v_2, \lambda_{v_2} | U^\dagger | w_2, \lambda_{w_2} \rangle \langle w_2, \lambda_{w_2} | U | v_1, \lambda_{v_1} \rangle \langle v_1, \lambda_{v_1} | \rho U^\dagger | w_3, \lambda_{w_3} \rangle, \quad (11)$$

where each $|v_l, \lambda_{v_l}\rangle\langle v_l, \lambda_{v_l}|$ or $|w_l, \lambda_{w_l}\rangle\langle w_l, \lambda_{w_l}|$ corresponds to the state of the system after a measurement (related to observable V or W) has been performed on it and an outcome (v_l, λ_{v_l}) or (w_l, λ_{w_l}) has been obtained.

As we will see shortly, the various fundamental concepts discussed so far—namely, the KD distribution in Eq. (3),

its extended version in Eq. (4), the quasiprobability behind the OTOC in Eq. (11), the post-selected quantum Fisher information in Eq. (5), and the expression for weak values in Eq. (2)—are all written in terms of Bargmann invariants, which we review next.

2.3. Bargmann invariants and how to measure them

General multivariate traces of quantum states can be directly measured in an efficient manner using the cycle test scheme [16] or constant-depth circuit variations thereof [17]. They are also known as n -th order Bargmann invariants [70]: given an n -tuple of states (ρ_1, \dots, ρ_n) , its Bargmann invariant is

$$\Delta_n(\rho_1, \dots, \rho_n) = \text{Tr}(\rho_1 \dots \rho_n). \quad (12)$$

These quantities are invariant under the simultaneous conjugation $\rho_i \mapsto Z \rho_i Z^{-1}$ by any invertible matrix Z .²

In general, we are interested in studying a set of states $\{\rho_i\}_i$ through all the Bargmann invariants among them: the quantities $\Delta_n(\rho_{i_1}, \dots, \rho_{i_n})$ for tuples of (not necessarily distinct) labels (i_1, \dots, i_n) .

Bargmann invariants play an important role in linear optics [71], characterizing multiphoton interference. In the context of invariant theory, multivariate traces of the form $\text{Tr}(X_1 \dots X_n)$ have been studied in depth since they completely generate the ring of invariants over matrix tuples [72]. In Ref. [16], it was pointed out that this feature can be used to completely characterize the equivalence classes of unitarily equivalent tuples of states, i.e., tuples $(\rho_i)_{i=1}^m$ that are mapped to one another by the action of a unitary map U simultaneously on every state in the tuple: $\rho_i \mapsto U \rho_i U^\dagger$. We refer to properties of sets of states that are invariant under such global unitary symmetry as *relational properties*. This is the most fundamental property of Bargmann invariants: every relational property of a set (or even a tuple) of states must be expressible as a function of the Bargmann invariants among those states.³ As we will see later, in many non-trivial situations, one requires non-linear functions of invariants, as is the case for weak values. Moreover, recognizing that a given function on a set of states can be written solely in terms of Bargmann invariants reveals that such a function captures relational information (see Def. 4 below). We refer to Ref. [72, Sec. 13.9.2], and references therein, for

more information on polynomials over sets of matrices that are unitarily invariant.

If we let C_n be the unitary representation of the cyclic permutation of n elements

$$(1\ 2\ 3\ \dots\ n) = (1\ 2)(2\ 3)\dots(n-1\ n) \quad (13)$$

that acts as

$$(a_1, a_2, \dots, a_{n-1}, a_n) \xrightarrow{C_n} (a_n, a_1, a_2, \dots, a_{n-1}), \quad (14)$$

it is possible to write any Bargmann invariant as

$$\Delta_n(\rho_1, \dots, \rho_n) = \text{Tr}(C_n \rho_1 \otimes \dots \otimes \rho_n). \quad (15)$$

Because C_n is unitary, we can use a quantum circuit known as the Hadamard test, shown in Fig. 2(a), to estimate the Bargmann invariant from Eq. (15). This procedure was detailed in Ref. [16]. To estimate a generic n -order Bargmann invariant $\Delta_n(\rho_1, \dots, \rho_n)$ we input, in parallel, an auxiliary qubit in the state $|0\rangle$ and the quantum states ρ_1, \dots, ρ_n . We apply a Hadamard gate to put the auxiliary qubit in superposition, followed by a controlled unitary operation of the unitary C_n . Depending on whether we want to estimate the real ($s = 0$) or imaginary ($s = 1$) part of the invariant, we apply the gate $P^s = \text{diag}(1, i^s)$ followed by another Hadamard gate. We conclude the protocol by measuring the auxiliary qubit in the Z basis.

There are different circuit constructions of C_n [16, 17], the simplest of which consists of the applying successive SWAP operators, thus translating into circuit form the decomposition of the cyclic permutation into transpositions from Eq. (13):

$$C_n = \text{SWAP}_{1,2} \circ \text{SWAP}_{2,3} \circ \dots \circ \text{SWAP}_{n-1,n}, \quad (16)$$

as depicted in Fig 2(b). The circuit must be run multiple times to estimate the invariant from the relative frequency estimation of a final computational basis measurement of the auxiliary qubit. For the specifics of implementing a qudit SWAP operation, see Ref. [73].

In Sec. 5, we consider how this Hadamard test circuit can be used to estimate weak values, KD distributions, and state spectra.

Among the most commonly considered higher-order Bargmann invariants in the literature are the *univariate traces* of a single quantum state, i.e., quantities of the form $\Delta_n(\rho, \dots, \rho) = \text{Tr}(\rho^n)$. It is known that learning sufficiently many of these can be used to estimate the spectrum $\text{Spec}(\rho)$ of a given mixed state ρ [74], an experimental task relevant for, e.g., quantifying coherence [8] or entanglement [7, 17, 75–78] using resource theory monotonies.

Procedures for measuring the quantities $\{\text{Tr}(\rho^n)\}_{n=1}^d$ have been proposed using visibility-based quantum algorithms [74, 79–81], via similar implementations of the cycle operator [17, 82], and using random measurements [83].

More generally, one may be interested in characterizing all relational (i.e., unitary-invariant) properties of a set of

² Throughout the text, when dealing with pure states ψ , we adopt the terser notation $\Delta_n(\dots, \psi, \dots)$ for $\Delta_n(\dots, |\psi\rangle\langle\psi|, \dots)$.

³ To avert potential misunderstanding, we reserve the letter n to indicate the order of Bargmann invariants. When studying the relational invariant properties of a tuple (or set) of states, we use the letter m to denote its size if its finiteness is relevant, as in $(\rho_i)_{i=1}^m$. Such properties are characterized by Bargmann invariants $\Delta_n(\rho_{i_1}, \dots, \rho_{i_n})$ where $i_1, \dots, i_n \in \{1, \dots, m\}$ are n labels.

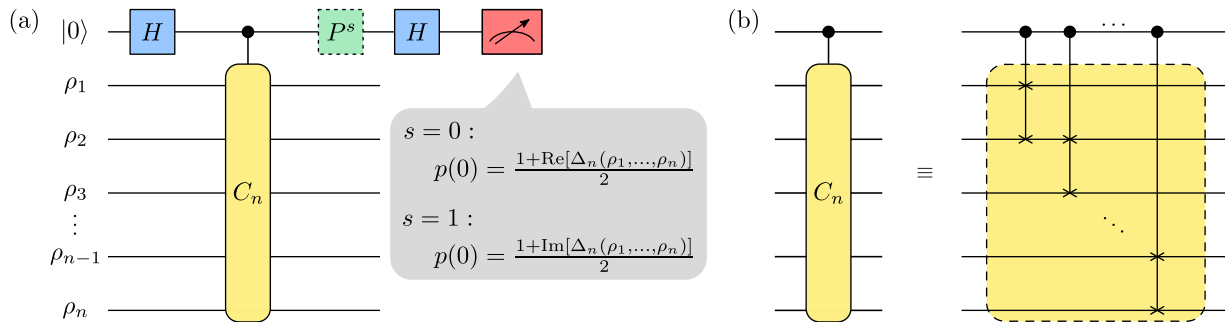


Figure 2. **Hadamard test circuit to measure Bargmann invariants.** (a) This Hadamard test circuit can be used to estimate the real and imaginary part of the n th order Bargmann invariant $\Delta_n(\rho_1, \dots, \rho_n) = \text{Tr}(\rho_1 \dots \rho_n)$. The controlled-unitary (in yellow) is a controlled cycle permutation, and the P^s gate equals identity for $s = 0$, and $\text{diag}(1, i)$ for $s = 1$. (b) Explicit decomposition of the controlled-cycle gate in terms of 3-system controlled-swap gates, also known in the qubit case as Fredkin gates.

m states. For pure states, there are known constructions of complete sets of invariants using $O(m^2)$ invariants of at most order m [16, 84]. Less is known for mixed states, but the minimum number of invariants in a complete set is known to depend also on the system dimension [16].

2.4. Nonclassicality of Bargmann invariants

A particularly important relational property of a set of states, which can be completely characterized by Bargmann invariants, is pairwise commutativity. Given any set of states $\{\rho_i\}_i$, this is equivalent to the existence of some unitary U such that $U\rho_i U^\dagger$ is diagonal for all i . In Ref. [12] the term *set coherence* was used to describe the property of a set of states for which no such unitary exists. For future reference, we write this notion as a definition.

Definition 3 (Set coherence [12]). Let $\{\rho_i\}_i$ be a set of quantum states in a finite-dimensional Hilbert space \mathcal{H} . This set is said to be *set incoherent* if all the states in it pairwise commute, or equivalently, if all these states are simultaneously diagonalizable, i.e. are represented by diagonal density matrices with respect to the same reference basis. Otherwise, we say that the set $\{\rho_i\}_i$ is *set coherent*.

Refs. [85–87] studied this property by analyzing the simplest of Bargmann invariants, two-state overlaps $\Delta_2(\rho_i, \rho_j) = \text{Tr}(\rho_i \rho_j)$.⁴

It is sensible to discuss the different *realizations* of an invariant, or of a tuple of invariants. We say that a value for Δ_n is realisable if there exists a tuple of

quantum states (ρ_1, \dots, ρ_n) , for some Hilbert space \mathcal{H} , such that $\Delta_n(\rho_1, \dots, \rho_n)$ has that value. Typically, there are infinitely many different realizations of a given value; for example, every tuple of the form (ψ, ψ) for a pure state $|\psi\rangle$ realises $\Delta_2(\psi, \psi) = 1$.

More generally, we consider tuples of Bargmann invariants, for example, the three overlaps among three states

$$(\Delta_2(\rho_1, \rho_2), \Delta_2(\rho_1, \rho_3), \Delta_2(\rho_2, \rho_3)) \in [0, 1]^3. \quad (17)$$

If only two-state overlaps appear, we also refer to the tuple as an overlap tuple. It is possible to obtain linear constraints bounding the tuples of Bargmann invariants realizable with sets of incoherent states, as shown in Refs. [85, 86]. We briefly sketch the main ideas of this approach. First, it is shown that some 0/1-valued assignments to the overlaps in a tuple are not realizable by set-incoherent states; in our running example from Eq. (17), this is the case, for instance, for the triple of values $(1, 0, 1)$. We then characterize the set of realizable, coherence-free tuples as the convex hull of all 0/1-valued assignments for the overlap tuple that are realizable by set-incoherent states [86].

As the convex hull of a finite number of points, this forms a polytope, which can also be characterized via a finite system of facet-defining linear inequalities. For bounding triples of overlaps among three set-incoherent states ρ_1, ρ_2 , and ρ_3 , the simplest such inequality is

$$\Delta_2(\rho_1, \rho_2) + \Delta_2(\rho_1, \rho_3) - \Delta_2(\rho_2, \rho_3) \leq 1. \quad (18)$$

This discussion will be relevant in Sec. 5 when we relate nonclassicality of quasiprobability distributions at some phase-space points to violations of these inequalities.

Ref. [85] studied the polytope \mathcal{C}_3^\bullet of overlap triples realizable by three set-incoherent states. This polytope is fully characterized by the inequality of Eq. (18), its index relabellings, plus the trivial inequalities $0 \leq \Delta_2 \leq 1$. We may also consider general quantum realizations, yielding

⁴ A note on terminology: in the literature, the term overlap is sometimes used to denote the inner product between two pure states, rather than its modulus squared, as we do here. Moreover, for pure states, the overlap becomes equivalent to the fidelity between two states.

the set

$$\mathcal{Q}_3^\bullet = \left\{ \begin{pmatrix} \Delta_2(\rho_1, \rho_2) \\ \Delta_2(\rho_1, \rho_3) \\ \Delta_2(\rho_2, \rho_3) \end{pmatrix} \in [0, 1]^3 : (\rho_i)_{i=1}^3 \in \mathcal{D}(\mathcal{H})^3 \right\}. \quad (19)$$

This set was shown numerically to be convex in Ref. [85]. It is sometimes referred to as the body of quantum correlations, and parts of it correspond to the ellipsope of Ref. [88]. In Refs. [85, 88] it was shown that points in \mathcal{Q}_3^\bullet satisfy the inequality

$$\Delta_2(\rho_1, \rho_2) + \Delta_2(\rho_1, \rho_3) + \Delta_2(\rho_2, \rho_3) - 2\sqrt{\Delta_2(\rho_1, \rho_2)\Delta_2(\rho_1, \rho_3)\Delta_2(\rho_2, \rho_3)} \leq 1. \quad (20)$$

It holds that $\mathcal{C}_3^\bullet \subset \mathcal{Q}_3^\bullet$, and any point in $\mathcal{Q}_3^\bullet \setminus \mathcal{C}_3^\bullet$ serves as a witness of set coherence for the triplets realizing the tuple of invariants.

For higher-order invariants, there are also constraints on the values realizable using only set-incoherent states. It was argued in Ref. [16] that negativity or imaginarity of Bargmann invariants witness this form of basis-independent coherence for a set of states, namely set coherence. In Sec. 5 we revisit this idea.

Set coherence in particular (see Def. 3), and noncommutativity of observables in general, is a unitary-invariant property, i.e., it is independent of a choice of reference basis. As such, deciding whether a set of observables pairwise commute can be framed solely in terms of Bargmann invariants, as discussed in Ref. [16]. This, in turn, demands the characterization of sufficiently many Bargmann invariants, forming a so-called *complete set of invariants*, which solves the problem entirely: knowing the values of all Bargmann invariants in that set suffices for deciding the property of interest. For the case of noncommutativity, the complete sets of Bargmann invariants and their associated bounds are not known in general.

A significantly simpler task is to select a single invariant, or an incomplete family of invariants, or functions thereof, and use their values to merely *witness* nonclassicality in the form of set coherence. Each of these witnesses signals a different manifestation of noncommutativity through the possible values attained by them. Because different witnesses usually capture significantly different phenomena, they each deserve individual analysis. Still, as we discuss in this work, all such nonclassicality witnesses can be framed within the following common terminology.

Definition 4 (Relational nonclassicality). Consider sets $\{\rho_i\}_i$ of states in a finite-dimensional Hilbert space \mathcal{H} . A property of such sets is said to be *relational* when it is captured by a function $f(\{\rho_i\}_i)$ which is *unitarily invariant* in that $f(\{\rho_i\}_i) = f(\{U\rho_i U^\dagger\}_i)$ for all unitaries U . We term *relational nonclassicality* the realization of values of such a relational property f that are not realizable with set-incoherent states.

Clearly, Bargmann invariants provide a specific case of function f in the above definition. Since this is a

particularly relevant case for the analysis in this work, we may refer to the appearance of nonclassical values as *Bargmann nonclassicality*.

Next, we start presenting our results, building on the following three aspects of Bargmann invariants. First, we recognize functions of Bargmann invariants that correspond to different concepts we have reviewed, e.g., OTOCs, weak values, and KD distributions. Following that, we point out that cycle test circuits can be used to estimate these quantities without the need for post-selection. Then, we show that any nonclassical property of the functions of invariants we consider can be linked to the nonclassicality of the underlying invariants themselves.

3. BARGMANN INVARIANTS AS A UNIFYING CONCEPT

In this section, we argue that Bargmann invariants serve as a unifying concept for the discussion of nonclassicality in KD distributions and weak values. We show how both are functions of third-order invariants. The same holds for extended KD distributions. A practical by-product of this unified description is the realization that cycle test circuits can be used to measure weak values, KD distributions, and the spectrum of quantum states, as we discuss in Sec. 4.

Starting with the definition of weak value in Eq. (2), we recognize that

$$\langle A \rangle_w = \sum_{a \in \text{Spec}(A)} a \frac{\Delta_3(\phi, a, \psi)}{\Delta_2(\phi, \psi)}. \quad (21)$$

Thus, we regard weak values as functions of second- and third-order invariants, encoding the relational information between the basis associated with observable A , and the two reference states $|\phi\rangle$ and $|\psi\rangle$. While weak values can also be viewed as averages of a conditional KD distribution [5], this requires the introduction of a second complete basis of reference into the weak value setup. Therefore, Eq. (21) expresses the weak value in terms of an economical function of only the strictly necessary unitary invariants, dispensing with the irrelevant, arbitrary choice of a second reference basis.

As weak values are functions of Bargmann invariants, they can be estimated with no need for post-selection, up to any desired accuracy, using the cycle test circuits of [16]. This shows that weak values need not be defined with respect to weak measurement schemes, as they were historically introduced. Eq. (21) provides a natural description of the weak value as an average of the eigenvalues of observable A , weighted by a quasiprobability distribution $\Delta_3(\phi, a, \psi)/\Delta_2(\phi, \psi)$ in terms of states $|\phi\rangle, |\psi\rangle$. Eq. (21) also shows it is not strictly necessary to view the weak value as merely the first term in an infinite series describing the shift of a classical pointer in a weak measurement scheme [20]. Together with our analysis presented in Sec. 4, this strengthens the case for

viewing weak values as interesting constructions of their own, as argued, e.g., in Ref. [6].

Similarly, the KD quasiprobability distribution can be also shown to be a relational property of the states involved, and thus a function of Bargmann invariants. The definition of the KD distribution at a single phase-space point, given by Eq. (3), shows explicitly that $\xi(\rho|i, f) = \Delta_3(i, \rho, f)$, i.e., that its value is a third-order Bargmann invariant. The KD distribution encodes the relational information about the quantum state ρ we want to describe, together with two complete Hilbert space bases $\{|i\rangle\}_{i \in I}$ and $\{|f\rangle\}_{f \in F}$.

Consequently, the question of characterizing the non-classicality of KD distributions (see Def. 2) boils down to characterizing the nonclassicality of individual third-order invariants, which we address in Sec. 5. As pointed out there, this novel perspective may prove useful in the future for connecting KD-nonclassicality with contextual-ity, an important open problem in quantum foundations. From a practical point of view, this simple understanding allows us to propose in Sec. 4 variations of cycle test circuits to measure KD distributions with no need for post-selection, in contrast to previous proposals based on weak measurements.

The observation above carries over to extended KD distributions relative to n bases, which are expressible as $(n+1)$ th order Bargmann invariants. More generally still, one can express an extended KD distribution of the form given by Eq. (4), which is relative to arbitrary PVMs (not necessarily composed only of rank-one projectors, i.e., bases). That additionally involves a factor that accounts for multiplicity. Setting $\rho_k^i := \frac{\Pi_k^i}{\text{Tr}(\Pi_k^i)}$, one has⁵

$$\xi(\rho|k_1, \dots, k_n) = \Delta_{n+1}(\rho, \rho_{k_1}^1, \dots, \rho_{k_n}^n) \prod_{i=1}^n \text{Tr}(\Pi_{k_i}^i). \quad (22)$$

This expression shows that at each phase-space point (k_1, \dots, k_n) the extended KD distribution describes relational properties between the probed state ρ together with all the states $\{\rho_{k_i}^i\}_i$. Any nonclassical property of the quasiprobability distribution must be due to the nonclassical properties of such higher-order Bargmann invariants.

A number of constructions that are derived from KD distributions are, as a result, also functions of Bargmann invariants. The extended KD distribution used in Eq. (5) is simply the collection of all invariants of the form

$$\xi(\rho_\theta|i, i', f) = \Delta_4(i, \rho_\theta, i', f). \quad (23)$$

This leads to the representation of the post-selected quantum Fisher information \mathcal{I}^{ps} purely in terms of Bargmann

invariants:

$$\mathcal{I}^{\text{ps}} = 4 \sum_{i, i', f} \lambda_i \lambda_{i'} \frac{\Delta_4(i, \rho_\theta, i', f)}{\sum_f \Delta_2(\rho_\theta, f)} - 4 \left| \sum_{i, i', f} \lambda_i \frac{\Delta_4(i, \rho_\theta, i', f)}{\sum_f \Delta_2(\rho_\theta, f)} \right|^2. \quad (24)$$

Similarly, the quasiprobability distribution behind the OTOC is an extended KD distribution:

$$\tilde{A}_\rho = \Delta_5(\lambda_{w_3}, \lambda_{v_2}^{(t)}, \lambda_{w_2}, \lambda_{v_1}^{(t)}, \rho(t)),$$

where we use the simplified notation $|\lambda_{w_l}\rangle := |w_l, \lambda_{w_l}\rangle$, $|\lambda_{v_l}^{(t)}\rangle := U(t)|v_l, \lambda_{v_l}\rangle$, and $\rho(t) := U(t)\rho U(t)^\dagger$.

Finally, it is important to remember the reasoning behind invariant theory: every relational property between states can be expressed in terms of a set of Bargmann invariants. This holds true for the spectrum of a quantum state, with the peculiarity that here we must consider invariants describing a (multi)set with n copies of a single quantum state. Estimating quantities of the form

$$\Delta_n(\rho, \dots, \rho) = \text{Tr}(\rho^n) \quad (25)$$

for sufficiently large n , one can learn several interesting properties of a single state, in particular its spectrum, as we recall later in Sec. 4.6.

There are two upshots from the unified perspective in terms of Bargmann invariants established in the present section:

- First, the fact that KD distributions, OTOCs, and weak values are functions of Bargmann invariants allows for a unified discussion of nonclassicality in terms of the values taken by them, in the sense of Def. 4. We use this insight in Sec. 5, where we establish connections between the resource theory of set coherence (discussed above, see Def. 3) and known results on the nonclassicality of quasiprobability distributions.
- The fact that all the constructions presented can be written as functions of Bargmann invariants means that cycle test circuits can be employed to estimate all of them, as we discuss in the next section.

4. QUANTUM CIRCUITS FOR MEASURING WEAK VALUES AND KIRKWOOD–DIRAC QUASIPROBABILITY DISTRIBUTIONS

4.1. Kirkwood–Dirac quasiprobability

Fig. 3(a) presents a circuit for measuring the value of the KD quasiprobability associated with pure state $|\psi\rangle$ at a given phase-space point (i, f) : $\xi(\psi|i, f) = \langle f|i\rangle \langle i|\psi\rangle \langle \psi|f\rangle = \Delta_3(i, \psi, f)$. The circuit, based on a cycle test [16], can be straightforwardly generalized to measure any value $\text{Tr}(\Pi_{k_1}^1 \dots \Pi_{k_n}^n \rho)$ of extended KD distributions; cf. Secs. 4.3 and 4.4. While we here focus on the case of pure states, the same architecture is capable of characterizing d -dimensional mixed input states.

⁵ This expression was suggested to us by one of the anonymous referees.

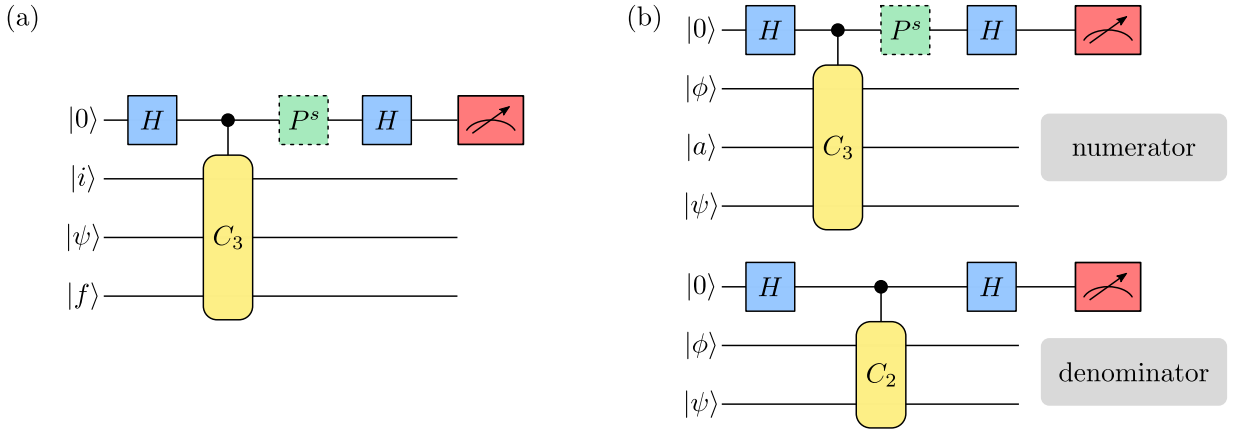


Figure 3. **Quantum circuits for measuring the third-order Bargmann invariants characterizing (a) Kirkwood–Dirac quasiprobability distribution and (b) weak values.** P^s corresponds either to applying a phase-gate $P = \text{diag}(1, i)$ in case we want to estimate the imaginary part of the invariant ($s = 1$), or an identity $P^0 = \mathbb{1} = \text{diag}(1, 1)$ in case we want to estimate the real part ($s = 0$). Circuit (a) assumes that one can prepare basis states from the two bases $\{|i\rangle\}_i, \{|f\rangle\}_f$ defining the KD distribution, as well as the state $|\psi\rangle$ to be characterized. Two variations of the circuit estimate the real and imaginary parts of the value of the KD distribution at a chosen phase-space point. For the circuit in (b), we require input states that are eigenvectors $\{|a\rangle\}_a$ of an observable A , as well as the pre- and post-selection states used in the standard weak measurement protocol. On top of that, anomalous weak values result from precisely selecting the overlaps $|\langle\phi|\psi\rangle|^2$, which can be directly measured using the two-state SWAP test.

The circuit shown in Fig. 3(a) has four wires or systems: the first is an auxiliary qubit, while the remaining are d -dimensional systems described by the Hilbert space \mathcal{H} under study. Given a fixed choice of i and f , the circuit initializes by preparing the product state $|0\rangle \otimes |i\rangle \otimes |\psi\rangle \otimes |f\rangle$. The specific order is relevant as the value of the Bargmann invariant changes for different choices of orders, due to noncommutativity of the rank-one projectors involved. The circuit then proceeds to put the auxiliary qubit in a superposition and perform a cascade of controlled-SWAP operations between the remaining states, as mentioned in Fig. 2.

Assuming that one can prepare the states $|\psi\rangle$, $|i\rangle$, and $|f\rangle$, the circuit in Fig. 3(a) measures the (real and imaginary parts of) the third-order Bargmann invariant defining $\xi(\psi|i, f)$, to precision ε , with high probability, using $O(1/\varepsilon^2)$ samples of the triplet of states [17], and therefore constant order of samples of the state $|\psi\rangle$. Consequently, the protocol can estimate the entire KD distribution of a state in any finite dimension d to precision ε with high probability using a total of $\tilde{O}(d^2/\varepsilon^2)$ samples, where the tilde hides $\log(d)$ terms.

Assuming that $\langle i|f\rangle \neq 0$, learning the full KD distribution of a state can be used to perform full tomography [48]. Tomography via KD distribution is neither better nor worse in terms of sample and measurement complexity than textbook quantum state tomography using Pauli measurements. The latter requires $O(d^4/\varepsilon^2)$ samples [89, 90] to achieve a precision ε relative to the trace distance with high probability. Now, if we learn the value of the KD distribution of a state ρ at every phase-space point, we can completely reconstruct the

state by

$$\rho = \sum_{(i,f) \in I \times F} \frac{|i\rangle\langle f|}{\langle i|f\rangle} \xi(\rho|i, f) \quad (26)$$

(in fact, this equation holds for any bounded operator, not just density matrices [4]). We count the number of samples and measurements required to perform tomography in this fashion, where a measurement consists of one run of the cycle test circuit to estimate some $\xi(\rho|i, f)$. In order to estimate the value at each phase-space point up to precision ε with probability $1 - \delta$,⁶ we need $O(\ln(2/\delta)/\varepsilon^2)$ samples, a bound provided by the Hoeffding concentration inequality [91, 92]. Assuming that we want to estimate at every phase-space point with the same precision ε and probability $1 - \delta$, the number of samples is of the order of $\tilde{O}(d^2/\varepsilon^2)$, where we hide the dependence on δ , as it is commonly assumed to be fixed.

To compare this scaling in samples with standard tomography, we must quantify the success of the procedure in terms of the distance induced by the 1-norm. Write $\hat{\xi}$ for the estimated KD distribution from the cycle test and $\hat{\rho}$ for the corresponding estimate of the state ρ , obtained via Eq. (26). Assuming that the values of KD are estimated with precision ε_1 at every phase-space point, i.e., that $|\xi(\rho|i, f) - \hat{\xi}(\rho|i, f)| \leq \varepsilon_1$ for all $(i, f) \in I \times F$,

⁶ Whenever we say that our learning task is successful with high probability, we mean that we are considering $1 - \delta$ with $\delta > 0$ a fixed small number.

we then have

$$\begin{aligned} \|\rho - \hat{\rho}\|_1 &= \left\| \sum_{i,f} \frac{|i\rangle\langle f|}{\langle i|f\rangle} \left(\xi(\rho|i, f) - \hat{\xi}(\rho|i, f) \right) \right\|_1 \\ &\leq \varepsilon_1 \left\| \sum_{i,f} \frac{|i\rangle\langle f|}{\langle i|f\rangle} \right\|_1 = \varepsilon_1 \|\mathbb{1}_{d \times d}\|_1 = \varepsilon_1 d, \end{aligned} \quad (27)$$

using that the rank-one operators $|i\rangle\langle f|/\langle i|f\rangle$ form an orthonormal basis of $\mathcal{B}(\mathcal{H})$. Therefore, in total, one needs $\tilde{O}(d^4/\varepsilon^2)$ samples to perform full KD state tomography ε close in the trace distance.

However, from the above calculation, it is clear that performing full tomography is *not necessary* for achieving arbitrarily good precision in learning the KD distribution of a given state ρ . Therefore, in scenarios where one is interested in learning solely the KD distribution $\{\xi(\rho|i, f)\}_{(i,f) \in I \times F}$, the required number of $\tilde{O}(d^2/\varepsilon^2)$ samples constitutes a polynomial advantage in sample complexity with respect to performing Pauli-based full tomography, which requires $O(d^4/\varepsilon^2)$ samples, or even efficient incoherent full tomography [93], which requires $O(d^3/\varepsilon^2)$ samples.

4.2. Weak values

Fig. 3(b) presents a circuit for estimating the weak value A_w of an observable A with respect to pre- and post-selection states ψ and ϕ . It works similarly to the cycle test described in Sec. 4.1. The main difference in the estimation of the weak value stems from the fact that it requires running two different circuits. The first, used to estimate the numerator of the expression for the weak value in Eq. (21), works just like the one shown in Fig. 3(a), while the second corresponds to a simple SWAP-test for estimating the denominator of Eq. (21). Such a process estimates a single summand $\Delta_3(\phi, a, \psi)/\Delta_2(\phi, \psi)$. To obtain the entire weak value, the procedure must be repeated for each eigenvector, and the estimated quantities then summed after being weighted by their corresponding eigenvalue.

The whole procedure assumes that we are able to not only prepare the selected states $|\phi\rangle$ and $|\psi\rangle$, but also each eigenvector $|a\rangle$ of A . This could be feasible given the mathematical description of the spectral decomposition of the operator A . But even in a more operational setup, it could be achieved, for example, if assuming the ability to perform a (strong) measurement of A on the totally mixed state.

The nonclassicality of weak values, witnessed by anomalous negative or nonreal values, can also be directly estimated using the above circuit and analyzed through the lens of Bargmann invariant nonclassicality. This result dissociates once more weak values from weak measurements, similarly to the discussions in Refs. [6, 94, 95].

The lack of post-selection might suggest that the sample complexity for measuring weak values using the cycle test improves over the standard weak measurement protocol. The following theorem shows that this is not the case. In fact, if the overlap between the pre- and post-selected states is small, our bounds show that the cycle test performs worse.

Theorem 5. *For estimating the weak value A_w with precision ε and high probability, one needs $N^{(weak)} = O(|A_w|^2/\varepsilon^2 \Delta_2(\phi, \psi))$ samples when using the standard weak measurement scheme, with $\Delta_2(\phi, \psi) = |\langle \phi|\psi\rangle|^2$ being the probability of successful post-selection. Moreover, one needs $N^{(cycle-test)} = O(|A_w|^2/\varepsilon^2 \Delta_2(\phi, \psi)^2)$ samples when using the procedure based on cycle tests.*

We formally prove this result in Appendix A. Intuitively, the difference in the number of required samples comes from the fact that the estimate using cycle tests is constructed from the estimation of two quantities—a second- and a third-order invariants—while the standard weak measurement scheme directly estimates the weak value from the pointer’s position. Since to observe anomalous weak values in most relevant situations—in particular in weak-value amplification schemes—the overlap $\Delta_2(\phi, \psi)$ needs to be close to zero,⁷ this difference in sample complexity makes the protocol suboptimal. Therefore, one needs to look for specific cases of interest for which the cycle test may be more relevant. A clear example is provided by experimental situations in which one utilizes purely *imaginary* weak values [96–98]. In these cases, simply witnessing imaginarity boils down to observing imaginary third-order invariants related to the numerator of the weak value A_w , a task that (with high probability and to precision ε) requires only $O(d/\varepsilon^2)$ samples.

4.3. Post-selected quantum Fisher information

We have seen that KD distributions can be extended to involve a larger number of bases (or PVMs) as described by Eq. (4), and that such constructions have appeared in the literature expressing important quantities behind OTOCs or quantum advantage in post-selected metrology, for example. Their description effectively requires Bargmann invariants of order higher than 3; see Eq. (22). In Fig. 4, we describe circuits specifically targeted to measuring such quantities.

The circuit in Fig. 4(a) measures the fourth-order Bargmann invariant in the expression for the post-selected

⁷ However, note that weak values with a small imaginary part do not require small overlaps between the pre- and post-selection states, and are also anomalous weak values. A simple example is the weak value of σ_x for a system pre-selected in $|z-\rangle$ and post-selected in $|y+\rangle$, which is $(\sigma_x)_w = -i$. A similar argument can be made when the real part of a weak value lies slightly outside the interval determined by the spectrum of the operator.

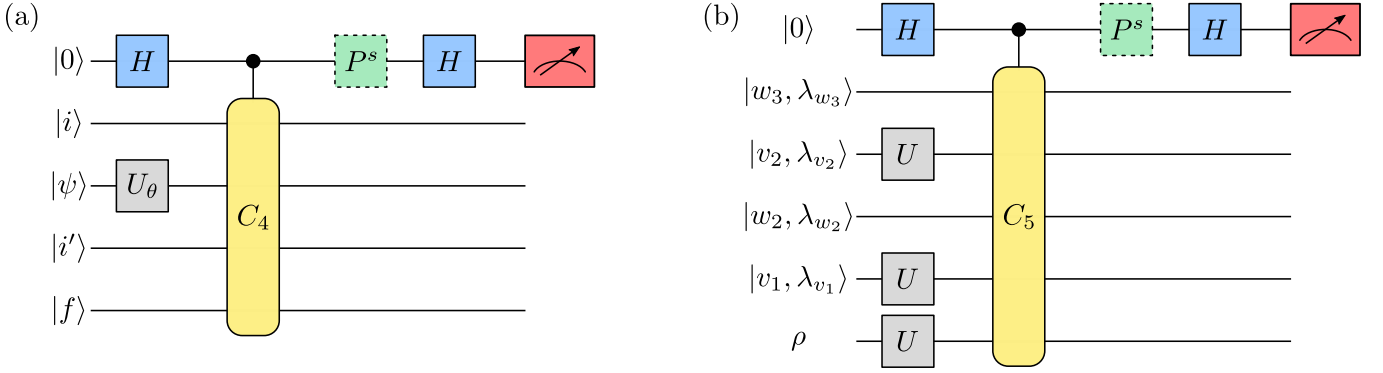


Figure 4. **Bargmann invariant circuits for measuring the extended Kirkwood–Dirac quasiprobability distributions behind (a) the post-selected quantum Fisher information and (b) the out-of-time-ordered correlator.** P^s is as described in Fig. 3. C_4 and C_5 correspond to the circuit implementation of the unitary discussed in Sec. 2.3. Part (a) shows how one can measure the post-selected quantum Fisher information using the cycle test and applying a unitary operator $U_\theta = e^{-i\theta I}$ over the state $|\psi\rangle$ considered in the protocol. It is assumed to be possible to prepare, in parallel, states $|f\rangle$ associated with the post-selection and the eigenstates related to the operator I . Part (b) shows the protocol for estimating the out-of-time-ordered correlator from the related Bargmann invariants. A possibly scrambling unitary U is applied to ρ and to the projectors related to the eigenspaces of v and w defined by Π_v^V and Π_w^W , respectively. It is possible to avoid an effective backwards-in-time evolution to estimate the OTOC, for any pair of observables V and W .

quantum Fisher information; see Eq. (24). This quantity is of interest, for example, in proof-of-principle experiments such as that performed in Ref. [68]. We discuss in detail how to make such a measurement using this simple quantum circuit.

Recall the metrological scenario for parameter estimation considered in Sec. 2.2.1, and the notation therein. The circuit in Fig. 4(a) consists of one (auxiliary) qubit and four systems described by the d -dimensional Hilbert space of interest. It is initialized by preparing the product state

$$|0\rangle \otimes |i\rangle \otimes |\psi\rangle \otimes |i'\rangle \otimes |f\rangle, \quad (28)$$

where $|i\rangle$, $|i'\rangle$ are elements of an eigenbasis of (the Hamiltonian) I and $|f\rangle$ an element of a basis for the range of the (post-selecting) projector F . The order in which those states are considered matters, being related to the order in the KD distribution presented in Eq. (6). The circuit then proceeds by applying a local unitary $U_\theta = e^{-i\theta I}$ on the state $|\psi\rangle$, which encodes information about the value of θ .

A cycle test is then performed over the four states, similarly to the procedures discussed in Sec. 4.1 or Sec. 4.2. The auxiliary qubit is put in a superposition and three controlled SWAP operations are applied: first between $|i\rangle$ and $|\psi_\theta\rangle$, then between $|\psi_\theta\rangle$ and $|i'\rangle$, and finally between $|i'\rangle$ and $|f\rangle$. As before, to estimate the real part of the invariant from Eq. (6) one simply measures the auxiliary qubit in the X basis, while to estimate the imaginary part of the invariant one first applies a phase gate $P = \text{diag}(1, i)$ and then measures the auxiliary qubit in the X basis.

The description above corresponds to a single run of the protocol, which estimates a specific value of the extended KD distribution of $|\psi_\theta\rangle$ for a given phase-space point. To estimate the entire function \mathcal{I}^{ps} , one must generate

statistics for each preparation of the form in Eq. (28) for all the labels i , i' , and f . The basis $\{|i\rangle\}_i$ has d elements, corresponding to the Hilbert space dimension, while $\{|f\rangle\}_f$ has fewer labels than d .

A note on the measurement and sample complexity related to estimating the post-selected quantum Fisher information in particular, and the extended KD distribution more generally, is necessary. Estimating \mathcal{I}^{ps} by this method requires $\tilde{O}(d^3/\varepsilon^2)$ samples and measurements, as all the information from the entire extended KD distribution is relevant. This does not represent an information complexity gain to experimentally probe this quantity as there are other approaches that require fewer samples. For instance, as we will recall in Sec. 4.6, some techniques to estimate the spectrum (which are also used to perform full tomography) solve the problem using $O(d^2/\varepsilon^2)$ samples. Yet, due to the intricate aspects of performing ideal protocols such as these, the circuit from Fig. 4(a) has the potential advantage of providing a simple and fixed structure for the measurement to be performed.

It is worth mentioning that, as in the experiment performed in Ref. [68], one may be interested in estimating a single phase-space point of the extended KD distribution, in order to monitor the presence of negative values. In such a case, provided we have access to some triplet of states $\{|i\rangle, |i'\rangle, |f\rangle\}$ with $i \neq i'$, estimating a single phase-space point has a simple scaling in terms of complexity relative to the one described in Sec. 4.1. To estimate any extended KD phase-space point value $\xi(\rho|\Pi_{k_1}, \dots, \Pi_{k_n})$, one needs $\tilde{O}(1/\varepsilon^2)$ samples, while a number of samples of order $\tilde{O}((n+1)d^2/\varepsilon^2)$ is required to perform full tomography of all the states and projectors involved in the calculation.

As discussed in Ref. [13, Supp. Note 2], imaginary values of the post-selected quantum Fisher information

do not contribute to metrological advantages since they cannot increase the second term of \mathcal{I}^{Ps} as described by Eq. (5). Hence, if one is only interested in witnessing the presence of the resource, namely negativity, one can use the circuit version of Fig. 4(a) with $s = 0$ (i.e., $P^s = 1$) to estimate the real part of the invariant.

To conclude this section, we remark that our circuits can be applied, after straightforward adaptations, to estimate more general definitions of the post-selected quantum Fisher information *matrix*, relevant for multi-parameter estimation, e.g., from Ref. [67, Eq. (37)].

4.4. Out-of-time-ordered correlators

The circuit in Fig. 4(b) applies the cycle test to measure OTOCs. From Eq. (8), we may use the fact that U is unitary and that the equivalence between the Heisenberg and Schrödinger pictures to rewrite the Bargmann invariant defining the OTOC as

$$\begin{aligned} & \text{Tr}(\Pi_{w_3}^{W(t)} \Pi_{v_2}^V \Pi_{w_2}^{W(t)} \Pi_{v_1}^V \rho) \\ &= \text{Tr}(U^\dagger \Pi_{w_3}^W U \Pi_{v_2}^V U^\dagger \Pi_{w_2}^W U \Pi_{v_1}^V \rho) \\ &= \text{Tr}(\Pi_{w_3}^W U \Pi_{v_2}^V U^\dagger \Pi_{w_2}^W U \Pi_{v_1}^V U^\dagger U \rho U^\dagger) \\ &= \text{Tr}(\Pi_{w_3}^W \mathcal{U}(\Pi_{v_2}^V) \Pi_{w_2}^W \mathcal{U}(\Pi_{v_1}^V) \mathcal{U}(\rho)), \end{aligned} \quad (29)$$

where $\mathcal{U}(\cdot) := U(\cdot)U^\dagger$ is the dynamical evolution in the Schrödinger picture. This allows us to have the cycle test that evaluates the OTOC without the need for applying the reversed evolution \mathcal{U}^\dagger . As compared with the protocols for estimating OTOCs reported in Ref. [4], not only does the cycle test avoid an effective backwards-in-time evolution, it also does away with the need for the auxiliary qubit to remain coherent during the scrambling dynamics. The protocol estimates the extended KD distribution behind the OTOC in either its coarse-grained version or the fine-grained description of Eq. (11), which carries nonclassical information that brings further relevant tools to the analysis of the scrambling dynamics [14, 15].

These observations place the cycle test as an interesting new paradigm for estimating OTOCs, in comparison with all protocols reported in Ref. [4, Table I, pg. 12]. Two important drawbacks of this protocol are: the fact that it requires quantum information processing of possibly many degrees of freedom in parallel, and the assumption that one is capable of preparing the states associated with the eigenprojectors Π_w^W and Π_v^V . Note that the latter assumption can be reduced to being able to measure the observables W and V (on any state)

A similar research direction was pursued in Ref. [99], where the authors analyzed when indirect metrological inference can be replaced by direct measurement processes, focusing on the Hadamard test. Our approach allows us to make claims that are similar in nature. For instance, Bargmann invariants can *also* be used to directly measure OTOCs.

4.5. Comparison with other methods for estimating the Kirkwood–Dirac quasiprobability distribution

We have seen how to use cycle test circuits to estimate the KD distribution in Sec. 4.1, weak values in Sec. 4.2, and also two specific examples of interest in Secs. 4.3 and 4.4. We proceed to situate our measurement scheme among those reviewed in Refs. [4, 5]. Table I compares our proposal with several protocols found in the literature. The first column presents the references that introduced the measurement scheme reviewed, which in some cases were not directly linked to estimating the KD distribution but were later shown to serve for this purpose in Ref. [5]. The second column shows the output of the protocol after gathering statistics, which in most cases is later used to infer the KD distribution. The third and fourth columns show the number of measurements and samples of the state $\rho \in \mathcal{D}(\mathcal{H})$ *per trial* of the experiment, where we consider weakly coupling to a pointer as another measurement. The fifth column indicates whether weak measurements are involved in the scheme, meaning that weak couplings are necessary. Finally, the last column indicates the required number of auxiliary quantum systems.

As these protocols are quite different in nature, it is challenging to make fair comparisons. In what follows, we detail the specifics of each protocol and explain the reasoning for the information in Table I. We choose to present the number of measurements and samples per trial as most proposals lack a formal complexity analysis (the exception is the protocol presented in Ref. [105]). Making a complexity analysis for each protocol would be beyond the scope of this work, and could potentially hide relevant constant factors in the number of measurements or samples needed.

Table I analyzes the resources for estimating just the real part $\text{Re}[\xi(\rho|i, f)]$ with i, f fixed. Estimating the imaginary part constitutes a different experiment in all cases except for the interferometric protocol from Ref. [104], which estimates inner products; in this case, our analysis considers the estimation of the value $\xi(\rho|i, f)$. In the other cases, to also take into account the estimation of $\text{Im}[\xi(\rho|i, f)]$ one simply has to multiply all measurements and samples by 2. Whenever there is more than one output (e.g., second row), parallel estimations need to take place to reconstruct the KD distribution phase-space point $\xi(\rho|i, f)$, increasing the number of measurements, samples, and auxiliary systems per trial. We now proceed to discuss each protocol in more detail.

The protocol first introduced in Ref. [48] was recently investigated in Ref. [5] and implemented in Ref. [54]. It demands the parallel implementation of three different schemes, each of which uses one sample of ρ per trial: a two-point measurement (TPM) [106] to learn p_{if}^{TPM} , another similar scheme with a non-selective measure-

Protocol	Output	# Meas.	# Samples of $\rho \in \mathcal{D}(\mathcal{H})$	Weak Meas.	# Aux. Systems
1) Fig. (3)	$\xi(\rho i, f)$	1	1	No	1 qubit + 2 systems \mathcal{H}
2) Ref. [48]	$p_{if}^{\text{TPM}}, p_f^{\text{END}}, p_{if}^{\text{WTPM}}$	5	3	No	0
3) Ref. [100]	$\chi(u, v)$	1	1	No	1 environment
4) Ref. [101]	$p_{\pm, if}$	3	1	No	1 system \mathcal{H}
5) Ref. [102]	Pointer shift	3	1	Yes	0
6) Ref. [103]	Pointer shift	3	1	Yes	0
7) Ref. [104]	$\langle i \rho f\rangle, \langle f i\rangle$	6	1	No	2 qubits
8) Ref. [104]	$\mathcal{P}_{\text{weak}}$	3	1	Yes	0
9) Ref. [105]	$\xi(\rho i, f)$	1	1	No	3 qubits

Table I. **Comparison between our approach and other methods for estimating the Kirkwood–Dirac quasiprobability distribution.** We compare previously proposed protocols (rows 2–9) with the one studied here (row 1). The first column shows the general output of the protocol; in the main text, we explain how one can recover the KD distribution from such outputs. Other columns indicate the number of measurements, samples, and auxiliary systems necessary for each single trial of the protocol, and whether it requires weakly coupling to a pointer, i.e., weak measurements. These single trials are solely for the estimation of the real part of the KD-value $\xi(\rho|i, f)$ for fixed ρ , i , and f . All protocols are also capable of estimating the imaginary part with similar schemes. The protocol in row 6 estimates inner products, hence both real and imaginary parts, which makes a fair comparison more subtle. *Abbreviations:* Meas.: measurement, Aux.: auxiliary.

ment⁸ to estimate p_{if}^{WTPM} , and a final experiment with a single measurement to estimate p_f^{END} . The real part of the KD distribution is then recovered by computing $\text{Re}[\xi(\rho|i, f)] = p_{if}^{\text{TPM}} + \frac{1}{2}(p_f^{\text{END}} - p_{if}^{\text{WTPM}})$, while the imaginary part can also be estimated with a similar setup. It is unclear how this protocol scales in terms of sample and measurement complexity. Our protocol demands fewer measurements and samples at the cost of quantum memory for the states $|i\rangle$ and $|f\rangle$.

The protocol from Ref. [100] estimates the characteristic function $\chi(u, v) = \sum_{if} \xi(\rho|i, f) e^{i\lambda_f v + i\lambda_i u}$ of the KD distribution with an interferometric scheme. It needs to couple the system of interest to an auxiliary system (called an environment) for the estimation of the real and imaginary parts of $\chi(u, v)$. Importantly, to recover the KD distribution, one needs to post-process the final result by applying an inverse Fourier transform over χ . This might impact on the scalability, in case the classical processing demands more precision in the estimation, as does the classical post-processing for estimating the spectrum, which we will see in Sec. 4.6.

While the protocols discussed above were studied with a focus on quantum foundations and quantum thermodynamics, the one from Ref. [101] provides a general way of estimating $\text{Tr}(\rho AB)$ for two observables A and B . It demands an auxiliary system described by the d -dimensional Hilbert space \mathcal{H} prepared in a maximally

mixed state. It then has one joint projective measurement over the target and auxiliary systems $M_{\pm} := \{\Pi^+, \Pi^-\}$ with $\Pi^{\pm} := (\mathbb{1} \pm \text{SWAP})/2$, together with local projective measurements in each system, $\{|i\rangle\langle i|\}_i$ and $\{|f\rangle\langle f|\}_f$. The probability of obtaining \pm from the first measurement M_{\pm} followed by i and f from the local measurements is denoted by $p_{\pm, if}$. One then recovers the real part of $\xi(\rho|i, f)$ by

$$\text{Re}[\xi(\rho|i, f)] = \frac{d+1}{2} p_{+, if} - \frac{d-1}{2} p_{-, if}. \quad (30)$$

Note that this needs more measurements per trial than our protocol, but, due to the probabilistic nature of quantum measurements, one cannot target specific phase-space points unless outcomes from the local measurements are post-selected.

The protocol from Ref. [102] is based on joint weak measurements together with the estimation of higher-order weak values, which implies that the effective output to be measured is the statistics of an average of a pair of pointer shifts. It is important to point out that Ref. [102] does not directly link their findings to estimating the KD distribution, while we (and Ref. [5]) argue that, with a direct re-reading of the protocol, one can connect the average shift of the pointer to KD values. The number of measurements in total is two weak measurements involving interactions with two different pointers in sequence, followed by final projective measurements. Since the protocol is based on standard weak measurement techniques, there is the need for post-selection.

While the protocol just mentioned is capable of estimating the KD distribution using joint weak measurements (requiring at least two systems) a similar approach using sequential weak measurements from Ref. [103], and

⁸ This procedure was termed weak TPM in Ref. [5]. While no weak measurement is needed, the fact that a non-selective (unsharp) measurement is used resembles the statistics that can be obtained from a weak measurement.

experimentally implemented in Refs. [38, 107], is also capable of estimating the KD distribution using a weak measurement scheme.

The interference scheme from Ref. [104, Appendix B] consists of estimating each inner-product $\langle f|i\rangle$ and $\langle i|\rho|f\rangle$ separately. The general way proposed to estimate $\langle i|\rho|f\rangle$ is using quantum state tomography and for estimating $\langle f|i\rangle$ the interferometer needs to estimate the real and imaginary parts separately by performing two parallel experiments: a SWAP test and another more involved one. The overall protocol becomes fairly demanding in terms of measurements and samples, especially due to the need to reconstruct each inner product and the estimation of $\langle i|\rho|f\rangle$. The same reference proposes a better scheme to estimate the extended KD distribution relevant for OTOC measurements that remains valid for general extended KD distributions (see Table I). In each trial of this weak-measurement-based scheme, for a general state ρ , one performs one weak measurement and two strong measurements in the pre/post-processing parts.⁹ For more details, see Ref. [104, Appendix A, Eq. (A5)] and Ref. [4].

Finally, we comment on the more recent protocol introduced in Ref. [105]. The ideas present there greatly resemble our scheme since something similar to a Hadamard test is performed (and a careful analysis of sample complexity is presented). The main difference is that instead of using auxiliary systems to prepare states $|i\rangle$ and $|f\rangle$ entering a generic cycle test, the method uses auxiliary qubits and performs block-encoding unitaries U_i and U_f that act on the joint system composed of $|\psi\rangle$ and the auxiliary qubits. As far as we know, this protocol is the only one that requires a similar number of samples and measurements to our approach without the need for major classical post-processing. The main difference between the two is that we avoid the need for different block-encoding unitaries U_i and U_f for each phase-space instance, at the cost of having access to a quantum memory with prepared states $|i\rangle$ and $|f\rangle$. Sidestepping such encoding issues clarifies the relational properties of the states involved, and may simplify the classical pre-processing for larger phase spaces, favoring our protocol. Meanwhile, for near-term devices, the lack of a large quantum memory may be more crucial, favoring the protocol introduced in Ref. [105].

In contrast to the other methods listed in Table I, our protocol works even when no classical description of the states being compared is available. In other words, the cycle tests we use can receive as inputs unknown states provided by third parties, which may suit applications involving, for example, state comparisons in networks. In

that, our protocol is akin to the programmable quantum gate arrays, first proposed in Ref. [108], where a fixed circuit has different functionalities depending on the input program states.

We sum up with a more detailed comparison of our proposal with those in Table I. Our protocol requires fewer measurements or samples than those from Refs. [48, 101, 102, 104], does not require post-selection or weak measurements as that from Ref. [102], does not require classical post-processing of data such as that from Ref. [100] (which applies the inverse Fourier transform), and it does not require a different Hadamard test, nor complete information on the states $|i\rangle$ and $|f\rangle$, for each KD phase-space point as that from Ref. [105]. Similarly to the protocol from Ref. [105], our protocol allows the estimation of the distribution at a single phase-space point.

4.6. Estimating the spectrum of a quantum state

So far, all quantities considered have depicted some relational aspect among different states in a set. Interestingly, cycle test circuits can also be used to estimate quantities associated with a single quantum state. In this section, we start by discussing quantum circuits that can estimate univariate trace polynomials $\text{Tr}(\rho^n)$, reminiscent of known approaches [80]. We carefully address the numerics of estimating the spectrum using these quantities due to the post-processing of classical data required by the Faddeev–LeVerrier algorithm. Finally, we provide a careful analysis in terms of the sample complexity needed to perform the task, together with a comparison with the best protocol available, known as the empirical Young diagram algorithm. We revisit the task of learning the spectrum, known from the early 2000s [74], in light of the results from Ref. [109] that provided optimal bounds for sample complexity of learning the spectrum, arguing that with our circuits it is possible to achieve the same efficiency with a simplified construction.

In Fig. 6, we describe the cycle test circuit that estimates $\text{Tr}(\rho^n)$ for any given $n \in \mathbb{N}$. It uses a single auxiliary qubit and n controlled-SWAP gates over n copies of the state ρ . This proposal is not optimal in the number of gates, however. In Refs. [16, 17], families of circuits were proposed with either logarithmic or constant depth. Yet, we observe that such advantage in depth is gained at the cost of an increase in space, as the proposals require the preparation of auxiliary multi-party Greenberger–Horne–Zeilinger (GHZ) states [110].

The Faddeev–LeVerrier algorithm, which uses the Newton identities [19, Chapter 3], allows us to compute the characteristic polynomial of the density matrix ρ from the quantities $\{\text{Tr}(\rho^n)\}_{n=1}^d$. The eigenvalues are then obtained by finding the roots of the characteristic equation. We provide code for this computation, as well as supporting information, in the repository [111], where we present these ideas with an experimentally-friendly approach.

⁹ In fact, this proposal requires three weak measurements since they are essentially estimating a higher-order invariant appearing in the OTOC. We use fewer weak measurements since we consider only the usual KD distribution. While this makes it more difficult to compare with our approach, this protocol is still an interesting measurement scheme to be considered for estimating the KD distribution.

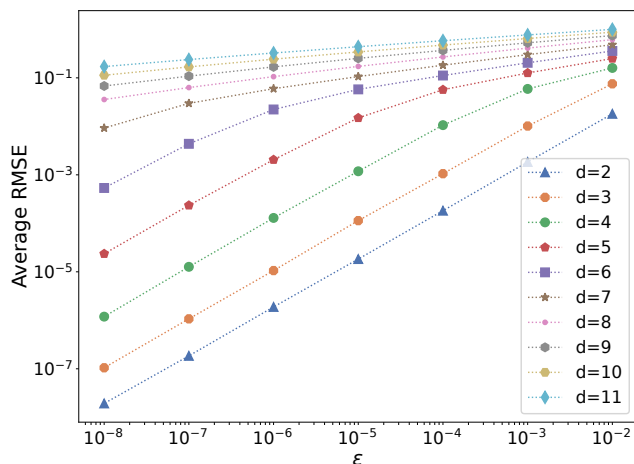


Figure 5. **Average root-mean-squared error (RMSE) of the estimate for the real part of the eigenvalues of random mixed states, under Gaussian noise ϵ .** We start from a data set of exact values of $\Delta_n := \text{Tr}(\rho^n)$, $\{\widehat{\Delta}_n\}_{n=2}^d$, introduce Gaussian noise with standard deviation ϵ , and plot the average root-mean-squared error (RMSE) of the estimated eigenvalues under noise. The spectrum reconstruction uses the Faddeev–LeVerrier algorithm based on Newton’s identities, and the average RMSE is over 5000 samples, each used to generate 1000 noisy samples. We use the Ginibre random ensemble of mixed states of rank larger than one, employing the algorithm introduced in Ref. [112].

This algorithm has some limitations, associated with finding the roots of the characteristic polynomial. As the dimension of the system increases, the coefficients of lower-degree monomials decrease, making them prone to numerical inaccuracies. We find that, for dimensions greater than 9, the predicted roots have an imaginary part larger than 10^{-8} . This imaginary part appears because of numerical inaccuracies, and it is independent of statistical or experimental noise. Therefore, it does not affect the error in the real part of the predicted eigenvalues, and we may simply discard the imaginary part of the results. Appendix B presents a thorough study of this behavior.

In real experiments, noisy data will distort the predictions of the algorithm. In Fig. 5, we study the aggregated root-mean-squared error (RMSE) of the real part of the eigenvalues of random mixed density matrices considering different amounts of noise in the experimental data. For a given set of traces computed from a random state, we can simulate the results from experiments by adding Gaussian noise with standard deviation ϵ . This noisy set of traces represents the average results from experiments accurate up to an error ϵ . The RMSE is computed by comparing the predictions of the algorithm using the noisy traces with the noiseless ones. We repeat the procedure and average with 1×10^3 noisy versions for each state, for 5×10^4 random states using the Ginibre ensemble [112]. These numbers are not related to the experimental samples required to estimate the traces, but to give enough

statistics to see the response of the algorithm to noisy data. Due to the estimation of the entire spectrum, it is expected that higher dimensions incur a high RMSE even for small ϵ . Typical error ranges in experimental realizations, e.g., $\epsilon \sim 10^{-3}$, would allow us to estimate the entire spectrum for dimensions 3 or 4 with good accuracy. A lower error ϵ in the experimental data could be attainable by increasing the number of samples. We remark that we only take statistical errors into account, as more details about the specific experimental implementation would be needed for further error analysis.

A simpler task involves learning the largest eigenvalue of ρ with high precision. This task should give better results as the higher-degree coefficients of the characteristic polynomial have a greater size than those of lower degrees, and are thus less affected by noise. We study the scaling of the error in the estimation of the largest eigenvalue in Appendix B. As expected, higher dimensional systems can be studied. For instance, for an experimental error of $\epsilon = 10^{-4}$, the average RMSE is around 5×10^{-3} for dimension 6. However, measuring the complete set of traces for estimating the largest eigenvalue is not an optimal procedure. While the coefficients of higher-degree monomials of the polynomial require traces of lower orders in powers of the density matrix, coefficients of lower degrees require traces of higher orders. As these traces require a more complex circuit, they will also be noisier. A sensible alternative for computing the largest eigenvalue is to truncate the polynomial to a certain lower degree k [78], that is, to compute the higher-degree coefficients of the polynomial using traces of lower powers of ρ , $\{\text{Tr}(\rho^2), \dots, \text{Tr}(\rho^{d-k})\}$. In this approximation, the traces that are more difficult to measure are not taken into account, as their contribution is smaller.

Oddly enough, estimating the spectrum of ρ from $\{\text{Tr}(\rho^n)\}_{n=1}^d$ for $d > 2$ has not, to the best of our knowledge, received any attention from experimental investigations yet, even though one can show that the required number of samples is optimal.

Theorem 6. *The number of samples needed to estimate $\{\text{Tr}(\rho^n)\}_{n=1}^d$ up to precision ϵ in all quantities with high probability is $N = \tilde{O}(d^2/\epsilon^2)$ with the cycle test. The number of measurements needed over auxiliary qubits is $\tilde{O}(d/\epsilon^2)$.*

We prove this result in Appendix B. The protocol of Ref. [17] improves this sample complexity by logarithmic factors, obtaining that order $O(d^2/\epsilon^2)$ samples are needed with high probability. One can learn the spectrum of a given state ρ using an optimal number of samples and measurements using the cycle test circuits described in this work or those described in Ref. [17]. The best bounds on the number of samples needed to estimate the spectrum $\text{Spec}(\rho)$ for d -dimensional mixed states ρ were first rigorously obtained in Refs. [113, 114] with the use of the empirical Young diagram (EYD) algorithm, later connected with complexity-theoretic arguments [109, 115], and formally used to show that the same sample complex-

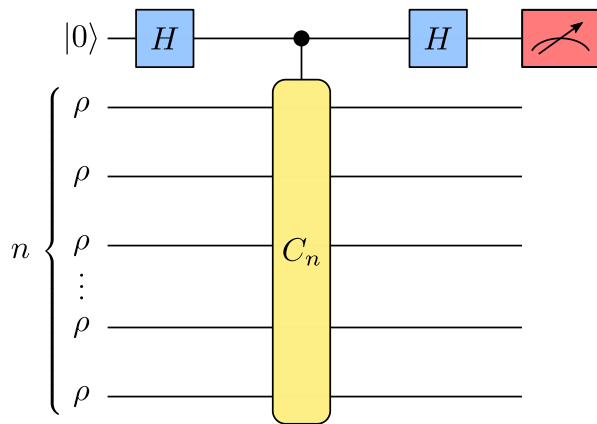


Figure 6. **Circuit for measuring $\text{Tr}(\rho^n)$.** The input consists of n copies of a d -dimensional state ρ . Measurements of $\text{Tr}(\rho^n)$ for $n = 1, 2, \dots, d$ gives us the spectrum of ρ , via the Newton identities.

ity is optimal for full state tomography [93, 116–118].

We provide a brief review of the EYD algorithm. Recall that by Schur–Weyl duality, the n -tensor product of \mathbb{C}^d decomposes as a direct sum of irreducible representations of $U(d)$:

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda} (V_{\lambda})^{\otimes m_{\lambda}} \quad (31)$$

where the direct sum goes over all partitions λ of n as a sum of no more than d positive integers. The direct sum contains m_{λ} copies of the irreducible representation V_{λ} , where m_{λ} is the number of standard Young tableaux corresponding to the partition λ . The irreducible representations V_{λ} (also known as *Weyl modules* in this context) can be constructed from $(\mathbb{C}^d)^{\otimes n}$ as the images of certain projections. Essentially, for each standard Young tableau, the corresponding projection acts on $(\mathbb{C}^d)^{\otimes n}$ by symmetrizing the tensor indices $1, \dots, n$ that correspond to each row of the tableau, and anti-symmetrizing the indices corresponding to each column. Now, the EYD algorithm is performed by fixing a large value of n , and then measuring $\rho^{\otimes n}$ with respect to the projections $\{\Pi_{\lambda}\}$, where Π_{λ} projects onto the subspace indexed by λ in Eq. (31). The measurement outcome λ is a set of positive integers that sum to n . After dividing each one by n , they comprise an estimate for the nonzero eigenvalues of ρ .

Some important differences between the two methods for learning the spectrum—the EYD algorithm and the protocol using the cycle test—deserve to be mentioned. To implement the EYD algorithm, one needs to measure $\rho^{\otimes n}$ in the highly entangled Schur basis (an orthonormal basis comprised of eigenvectors of the projections Π_{λ}), and to obtain the spectrum with precision ε , one needs $n_{\text{EYD}} = O(d^2/\varepsilon^2)$. In terms of sample complexity, this is optimal: one requires $N = n_{\text{EYD}}$ copies of ρ , which is the optimal sample complexity—and the same as in our protocol, as stated before.

At first glance, the EYD algorithm may seem to be simpler, as it requires implementing only a single quantum circuit. Indeed, one should choose a large enough value of n_{EYD} , and then implement the Schur transform (mapping the Schur basis to the computational one) and measure the final state only for that fixed n_{EYD} . In contrast, our approach requires estimating $\text{Tr}(\rho^n)$ for $n = 1, \dots, d$, where every increment of n requires a slightly larger circuit. However, this “partition” of the algorithm into d quantum circuits holds a huge practical advantage. As is evident from the previous paragraph, the EYD algorithm requires a quantum computer that can reliably sustain a highly entangled state in a Hilbert space of dimension $d^{n_{\text{EYD}}}$. This dimension grows larger for smaller desired precision ε . In contrast, in our approach, the largest circuit acts on dimension d^d , which still grows large for higher values of d but does not depend on ε . Moreover, the gate complexity of the Schur transform is polynomial in both d and n_{EYD} [119], while the complexity of the circuit for measuring $\text{Tr}(\rho^n)$ only depends on n .¹⁰ Thus, the gate complexity of our approach does not depend on ε and is generally much smaller when compared with EYD.

It should be noted that the authors of Ref. [120] overcome the issue of gate complexity by considering a specific experimental setup where the symmetry of the interaction Hamiltonian is utilized to implement the Schur transform efficiently.¹¹

Moreover, other methods have been proposed for estimating $\text{Tr}(\rho^n)$. One such test, known as Brun’s multicopy method [121], uses the fact that any polynomial $f(\rho)$ on the matrix elements of ρ , with respect to some fixed basis, can be expressed as the expectation of some operator acting over $\rho^{\otimes m}$ for some natural number m . Observables acting over such tensor products are sometimes called *multicopy observables* and have been proposed to access nonclassicality of quantum states in optics [122]. For a formal comparison between our scheme and Brun’s, see Ref. [121, Sec. V]. It is also worth mentioning that the cycle test was used in earlier works [79], focusing on the estimation of $\text{Tr}(\rho^n)$ only, without considering more general Bargmann invariants. Another recently introduced methodology for general estimation of Bargmann invariants is presented in Ref. [17].

5. MINIMAL EXPERIMENTAL CONDITIONS FOR WITNESSING NONCLASSICALITY

Most of the constructions discussed in the last section capture the community’s interest because they present

¹⁰ Or even better, achieving $\log(n)$, as shown in Ref. [16], or constant depth, as shown in Ref. [17].

¹¹ While acknowledging that some of the experimental aspects needed to implement the EYD algorithm “seem like a daunting task in practice” [120, pg. 3].

features that are puzzling to be described classically—notably, values that are negative or have nonzero imaginary part. Given that the protocols discussed in this work are capable of estimating these quantities, we are now interested in studying minimal experimental setups for which it is possible to witness nonclassicality in various guises.

We first discuss how the negativity or imaginarity of (third- and higher-order) Bargmann invariants, and in particular that of weak values or KD distributions, can be understood as witnessing quantum (set) coherence.

However, following Ref. [86] (see Sec. 2.4), set coherence can also be witnessed from two-state overlaps. So, afterwards, we consider to what extent these two ways of witnessing coherence are interrelated. In particular, we consider the question of whether anything can be inferred about the negativity or imaginarity of third-order invariants from knowledge of pairwise overlaps alone. In the simplest scenario, with three states, it turns out that, in general, overlaps alone are not enough to decide about negativity or imaginarity of third-order invariants, but if one restricts to real amplitude states, then overlap nonclassicality and negativity of the third-order invariant are mutually exclusive (complementary) witnesses of coherence.

5.1. Negativity, imaginarity, and set coherence

5.1.1. Negativity and imaginarity as coherence witnesses

As reviewed in Sec. 2.4, the assumption of set incoherence for a set $\{\rho_i\}_i$ of quantum states, described in Def. 3, imposes restrictions on tuples of Bargmann invariants among states drawn from that set. This was extensively studied in Refs. [85, 86] for the case of overlaps, and briefly mentioned in an appendix of Ref. [16] for more general Bargmann invariants. One such restriction is that all the invariants must necessarily be real and nonnegative (in fact, must take values in the interval $[0, 1]$), as we now see.

Consider a set of states $\{\rho_i\}_i$ in a finite-dimensional Hilbert space \mathcal{H} , and suppose that it is set incoherent. This means that its elements are simultaneously diagonalizable, i.e., there is a basis with respect to which all those states ρ_i are represented by a diagonal matrix, say $\text{diag}(p_i^{(1)}, \dots, p_i^{(d)})$, where $d = \dim(\mathcal{H})$. Any Bargmann invariant drawn from this set then takes the form

$$\Delta_n(\rho_{i_1}, \dots, \rho_{i_n}) = \sum_{k=1}^d p_{i_1}^{(k)} \cdots p_{i_n}^{(k)}. \quad (32)$$

This value can be understood as the probability that, upon measuring the n states $\rho_{i_1}, \dots, \rho_{i_n}$ with an observable corresponding to the basis that diagonalizes all the states in $\{\rho_i\}_i$, they all return the same outcome.

Nonreal or negative values are blatantly at odds with this interpretation, and thus immediately rule out

a set-incoherent realization. More generally, the specific form in Eq.(32) constrains the possible values of Bargmann invariants—and of tuples of Bargmann invariants—admitting a set-incoherent realization.

Let us remark that Bargmann nonclassicality, as in Def. 4, is not equivalent to set coherence or incompatibility, from Def. 3. Bargmann nonclassicality is associated with functions of Bargmann invariants attaining values inaccessible by set-incoherent realizations. Each departure from such a set incoherent explanation signals a specific form of nonclassicality. A few examples that we have already encountered are (i) $\text{Re}[\Delta_n(\rho_1, \dots, \rho_n)] < 0$ (negativity), (ii) $\text{Im}[\Delta_n(\rho_1, \dots, \rho_n)] \neq 0$ (imaginarity), and (iii) violations of overlap inequalities as in Eq. (18) (interpretable as coherence from two-state overlaps, as per Ref. [86]), each of which expresses an inequivalent form of Bargmann nonclassicality. Imaginarity, in particular, received some attention in Ref. [16].

5.1.2. Nonclassicality in weak values and KD distributions as relational nonclassicality

Negativity and imaginarity have also been studied as markers of nonclassicality for weak values or KD distributions. Given our rendering of these quantities in terms of Bargmann invariants, the observations above admit a reading in those terms. The following two lemmas are an instantiation of the observations presented in Sec. 5.1.1 and make explicit this trivial—albeit important—connection between set coherence and negativity or imaginarity of weak values or KD distributions.

Lemma 7. *In a finite-dimensional Hilbert space \mathcal{H} , let $|\phi\rangle, |\psi\rangle$ two non-orthogonal vectors, $A : \mathcal{H} \rightarrow \mathcal{H}$ a Hermitian operator, and $a \in \text{Spec}(A)$ an eigenvalue of A with corresponding eigenstate $|a\rangle$. Then, negativity or imaginarity of the weak value P_w^a of the eigenprojector $P^a = |a\rangle\langle a|$, relative to pre- and post-selected states ψ and ϕ , witness the set coherence of $\{|a\rangle, |\psi\rangle, |\phi\rangle\}$.*

Lemma 8. *Let $\{|i\rangle\}_{i \in I}, \{|f\rangle\}_{f \in F}$ be two (reference) orthonormal bases of a finite-dimensional Hilbert space \mathcal{H} , and $\rho \in \mathcal{D}(\mathcal{H})$ be a state. Then, negativity or imaginarity of $\xi(\rho|i, f)$, the value of the KD distribution associated with ρ at a specific phase-space point $(i, f) \in I \times F$, witness the set coherence of $\{\rho, |i\rangle, |f\rangle\}$. An analogous statement also holds for any extended KD distribution.*

Note that for weak values, given that we assume pure pre- and post-selections, the measurement of $\Delta_2(|\psi\rangle, |\phi\rangle)$ alone guarantees coherence. Recently, Ref. [123] considered in more depth the implications of Lemma 7, showing that *any* anomalous weak values require coherence, in the more general case where pre- and post-selection states can be mixed.

5.1.3. Set coherence beyond negativity and imaginarity

The analysis of Bargmann invariants for diagonal states sheds some light on the reason why negativity and imaginarity require set coherence or noncommutativity, as noticed in Ref. [124], while the converse is not the case.

Indeed, as already discussed, these conditions on Bargmann invariants are only two among *many* that such quantities should satisfy in order to corroborate the assumptions connected to the classical probabilistic interpretation provided by Eq. (32). It is, therefore, to be expected that negativity or imaginarity *alone* are incapable of completely characterizing set coherence. To illustrate this point, consider the following three states:

$$\begin{aligned} |\psi_1\rangle &:= |0\rangle, \\ |\psi_2\rangle &:= \frac{|0\rangle + \sqrt{3}|1\rangle}{2}, \\ |\psi_3\rangle &:= \frac{\sqrt{3}|0\rangle + |1\rangle}{2}. \end{aligned} \quad (33)$$

These states satisfy $\Delta_3(\psi_1, \psi_2, \psi_3) = 3/8 > 0$ and similarly for the other orientation. As such, all Bargmann invariants of third (and therefore of *any*) order among states drawn from the set $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ are real and positive. However, this set of three states is set coherent. This can be seen from the fact that they violate the coherence-witnessing inequalities from Refs. [11, 85], reviewed in Section 2.4. Concretely, they violate the inequality

$$-\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_3) + \Delta_2(\psi_2, \psi_3) \leq 1, \quad (34)$$

whose left-hand side, for these states, evaluates to

$$-\frac{1}{4} + \frac{3}{4} + \frac{3}{4} = \frac{5}{4} > 1. \quad (35)$$

Later on, we revisit these two-state overlap inequalities in a different context; see Eq. (44). For proof that these inequalities are witnesses of coherence, we refer the reader to Refs. [11, 85] (see also the brief summary in Sec. 2.4).

5.2. Can negativity and imaginarity be witnessed from overlaps?

We have seen that the negativity or imaginarity of (any) Bargmann invariant witness set coherence. In other situations, knowledge of two-state overlaps is enough to witness set coherence; see Refs. [85, 86].

A natural question is to understand the extent to which such witnesses capture distinct aspects of nonclassicality. Specifically, we ask whether—or to what extent—it is possible to infer negativity and imaginarity of higher-order Bargmann invariants from the knowledge of overlaps (i.e., second-order Bargmann invariants) alone.

5.2.1. Dimension-specific KD negativity and imaginarity from overlaps in the whole phase space

We first review a known result of this kind. Refs. [124–128] identified conditions that imply the presence of *some* negativity or imaginarity in the KD distribution of a pure state $|\psi\rangle$, using only knowledge of the overlaps between $|\psi\rangle$ and the states in each of the reference bases. In fact, the condition only requires *possibilistic* information about such overlaps, i.e., whether they are equal to zero (orthogonality) or not.

Concretely, let $\{|i\rangle\}_{i \in I}$ and $|f\rangle_{f \in F}$ be two orthonormal bases of a Hilbert space of dimension d , to be used as reference for the KD distribution, and make the usual assumption that $\langle i|f\rangle \neq 0$ for all i, f . Given a pure state $|\psi\rangle$, define

$$n_I(\psi) := |\{i \in I : \langle i|\psi\rangle \neq 0\}| \quad (36)$$

to be the number of elements of the basis $\{|i\rangle\}_{i \in I}$ not orthogonal to $|\psi\rangle$, or equivalently, with nonzero overlap $\Delta_2(i, \psi)$; and similarly define n_F with respect to the basis $|f\rangle_{f \in F}$.

Then, one obtains negativity or imaginarity at *some* phase-space point $\xi(\psi|i, f)$ if

$$n_I(\psi) + n_F(\psi) > d + 1. \quad (37)$$

If, moreover, the two reference bases are mutually unbiased, the KD distribution $\xi(\psi|\cdot)$ is classical, i.e., it has real and non-negative values at every phase-space point, if and only if [128]

$$n_I(\psi)n_F(\psi) = d. \quad (38)$$

Note that, in both cases, the condition depends on the dimension of the physical system. Moreover, in order to apply those results, one needs (some) information about all the overlaps $\Delta_2(i, f)$, $\Delta_2(i, \psi)$, $\Delta_2(\psi, f)$, namely whether they are zero or not.

We aim to address a related question, still about witnessing negativity or imaginarity from overlaps alone, but adopting a somewhat different perspective. First, we focus on one phase-space point at a time, i.e., on each triplet $|\psi\rangle, |i\rangle, |f\rangle$ with fixed i, f . Second, we are interested in dimension-independent witnesses, which avoid assuming a specific Hilbert state dimension.

5.2.2. Inferring higher-order Bargmanns from overlaps: the three-state scenario

The preceding discussion leads us to consider the minimal scenario where the question of witnessing negativity or imaginarity from two-state overlaps is meaningful. Suppose that there are three (unknown) pure states $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$, and that we are able to obtain knowledge of the three pairwise overlaps $\Delta_2(\psi_1, \psi_2)$, $\Delta_2(\psi_2, \psi_3)$, and $\Delta_2(\psi_1, \psi_3)$. We ask the following question: is such information enough to infer about the negativity or imaginarity of the third-order Bargmann invariant $\Delta_3(\psi_1, \psi_2, \psi_3)$?

In the remainder of this Sec. 5, we obtain and review some impossibility results and some partial results in this direction.

We set some notation to facilitate precisely stating our results about the three-state overlap scenario. Recall from Eq. (19) the definition of the set $\mathcal{Q}_3^\bullet \subseteq [0, 1]^3$ of overlap triples realizable by quantum states, and from the preceding paragraph in Sec. 2.4 that of the subset $\mathcal{C}_3^\bullet \subseteq \mathcal{Q}_3^\bullet$ of those overlap triples realizable by set-incoherent states. We are interested in distinguishing various cases for the (not directly available) value of the third-order invariant among the states. We write \mathcal{R}_3^\bullet (resp. \mathcal{I}_3^\bullet , \mathcal{P}_3^\bullet , \mathcal{N}_3^\bullet , \mathcal{Z}_3^\bullet) for the set of overlap triples realizable by three states ρ_1, ρ_2, ρ_3 whose third-order invariant $\Delta_3(\rho_1, \rho_2, \rho_3)$ is real (resp. nonreal, positive, negative, zero); for example,

$$\mathcal{R}_3^\bullet = \left\{ \left(\begin{array}{c} \Delta_2(\rho_1, \rho_2) \\ \Delta_2(\rho_1, \rho_3) \\ \Delta_2(\rho_2, \rho_3) \end{array} \right) \in [0, 1]^3 : \begin{array}{l} \rho_1, \rho_2, \rho_3 \in \mathcal{D}(\mathcal{H}), \\ \text{Im}[\Delta_3(\rho_1, \rho_2, \rho_3)] \neq 0 \end{array} \right\}. \quad (39)$$

Many of our results concern realizations with *pure* states. We write \mathcal{Q}_3° for the overlap triples realizable by three *pure* states,

$$\mathcal{Q}_3^\circ = \left\{ \left(\begin{array}{c} \Delta_2(\psi_1, \psi_2) \\ \Delta_2(\psi_1, \psi_3) \\ \Delta_2(\psi_2, \psi_3) \end{array} \right) \in [0, 1]^3 : |\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \in \mathcal{H} \right\}. \quad (40)$$

Similarly, \mathcal{C}_3° , \mathcal{R}_3° , \mathcal{I}_3° , \mathcal{P}_3° , \mathcal{N}_3° , \mathcal{Z}_3° denote the sets of overlap triples admitting a realization with pure states that moreover satisfies the condition corresponding to each letter as above. Finally, we use the superscript $(\cdot)^{(d)}$ to restrict the condition to realizations by states in d -dimensional Hilbert space. Trivially, from reasoning at the level of realizations, $\mathcal{X}_3^{\circ(d)} \subseteq \mathcal{X}_3^{\bullet(d)}$, $\mathcal{X}_3^{\circ(d)} \subseteq \mathcal{X}_3^{\bullet(d+1)}$, and $\mathcal{X}_3^\bullet = \bigcup_d \mathcal{X}_3^{\bullet(d)}$, for all $\mathcal{X} \in \{\mathcal{Q}, \mathcal{C}, \mathcal{R}, \mathcal{I}, \mathcal{P}, \mathcal{N}, \mathcal{Z}\}$ and $\bullet \in \{\circ, \bullet\}$.

By the same token, it is clear that $\mathcal{P}_3^\bullet \cup \mathcal{N}_3^\bullet \cap \mathcal{Z}_3^\bullet = \mathcal{R}_3^\bullet$ and $\mathcal{R}_3^\bullet \cup \mathcal{I}_3^\bullet = \mathcal{Q}_3^\bullet$, but observe that, crucially, these unions are not *a priori* disjoint. This is because an element of these sets does not carry all the information about a realization, but only the values of the overlaps among the three realizing states; a particular overlap triple typically admits more than one realization. Such limited information might not even be sufficient to distinguish real and imaginary values (or positive and negative values) for the third-order invariant of the realizing states. In other words, the same tuple of overlaps could be realizable by different sets of three states, with different values for the third-order invariant, including negative and positive, or with and without imaginary part.¹² Our first result below (Theorem 9) establishes that this indeed happens.

Before stating it, we make a useful remark regarding imaginarity. A set of states can be simultaneously represented with only real amplitudes (i.e., with respect to the same basis) if and only if all its Bargmann invariants are real. Consequently, in the case of pure states, since a set of three states is fully characterized by Bargmanns up to order 3, the set \mathcal{R}_3° is equivalently described as the set of overlap values realizable by a choice of three pure states with *real amplitudes*¹³.

5.2.3. *Overlaps are, in general, not enough for negativity and imaginarity*

We now show that overlaps are, in general, not sufficient to decide negativity or imaginarity of third-order invariants, even in the case of pure states.

Theorem 9. *Let $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$ be three (unknown) pure states in a finite-dimensional Hilbert space \mathcal{H} . Suppose that one is given the values of the three pairwise overlaps $\Delta_2(\psi_1, \psi_2)$, $\Delta_2(\psi_2, \psi_3)$, and $\Delta_2(\psi_1, \psi_3)$.*

1. *In general, it is not possible to discriminate between real and strictly complex third-order Bargmann invariant, i.e., to decide whether $\text{Im}[\Delta_3(\psi_1, \psi_2, \psi_3)] = 0$ or $\text{Im}[\Delta_3(\psi_1, \psi_2, \psi_3)] \neq 0$.*
2. *Further assuming that $\Delta_3(\psi_1, \psi_2, \psi_3)$ is real, it is in general not possible to discriminate between positive and negative third-order Bargmann invariants, i.e., to decide whether $\Delta_3(\psi_1, \psi_2, \psi_3) > 0$ or $\Delta_3(\psi_1, \psi_2, \psi_3) < 0$.*

In summary, $\mathcal{R}_3^\circ \cap \mathcal{I}_3^\circ \neq \emptyset$ and $\mathcal{P}_3^\circ \cap \mathcal{N}_3^\circ \neq \emptyset$.

Proof. We prove statement 1 by constructing a simple counterexample. The idea is to find realizations with equal overlaps but different imaginary parts of the third-order invariant.

Considering the triplet of states $(|0\rangle, |+\rangle, |i+\rangle)$, where $|i+\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$, we have that all pairwise overlaps are equal to $1/2$ while their third-order invariants are $\Delta_3 = (1 \pm i)/4$ (with the sign depending on the order), hence $\text{Im}[\Delta_3] \neq 0$.

We now give a different triplet of states with all pairwise overlaps equal to $1/2$, but whose states can all be described with only real amplitudes. The simplest such example would be the triplet consisting of three maximally mixed qubits, i.e., $(\mathbb{1}/2, \mathbb{1}/2, \mathbb{1}/2)$. This would suffice to establish the result for mixed states, or even to show $\mathcal{I}_3^\circ \cap \mathcal{R}_3^\bullet \neq \emptyset$ (even more specifically, $\mathcal{I}_3^{\circ(2)} \cap \mathcal{R}_3^{\bullet(2)} \neq \emptyset$). For the stated result, however, we need an example with pure states.

¹² The case \mathcal{Z}_3° when the third-order Bargmann invariant $\Delta_3(\psi_1, \psi_2, \psi_3)$ is zero is easier to separate because it implies one of the pairwise two-order invariants also vanish.

¹³ For mixed states ρ_1, ρ_2, ρ_3 , it is conceivable that $\Delta_3(\rho_1\rho_2\rho_3)$ is real but imaginarity still appears in higher-order invariants, such as $\Delta_4(\rho_1, \rho_2, \rho_3, \rho_2)$. As far as we know, it is an open question whether this can happen. For now, only one of the inclusions holds: the set of overlaps realizable by three states with real density matrices is contained in \mathcal{R}_3^\bullet .

The following states (in a Hilbert space of dimension ≥ 3) have real amplitudes and pairwise overlaps $1/2$:

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}}, \frac{|1\rangle + |2\rangle}{\sqrt{2}}, \frac{|0\rangle + |2\rangle}{\sqrt{2}}. \quad (41)$$

We have thus established that $(1/2, 1/2, 1/2) \in \mathcal{I}_3^{\circ(2)} \cap \mathcal{R}_3^{\circ(3)}$.

We now proceed to prove statement 2. We do so, again, by constructing a counterexample where two different realizations return the same two-state overlaps but third-order invariants that are real and of opposite sign. For the three states

$$|0\rangle, \frac{|0\rangle + \sqrt{3}|1\rangle}{2}, \frac{|0\rangle - \sqrt{3}|1\rangle}{2}, \quad (42)$$

all three overlaps equal $1/4$, while $\Delta_3 = -1/8 < 0$. On the other hand, for the three states

$$|0\rangle, \frac{|0\rangle + \sqrt{3}|1\rangle}{2}, \frac{3|0\rangle + \sqrt{3}|1\rangle + \sqrt{24}|2\rangle}{6}, \quad (43)$$

all overlaps, once more, equal $1/4$, while $\Delta_3 = +1/8 > 0$. This shows that $(1/4, 1/4, 1/4) \in \mathcal{N}_3^{\circ(2)} \cap \mathcal{P}_3^{\circ(3)}$. \square

The theorem above shows that the knowledge of two-state overlaps for a given triple of states, in general, provides insufficient information about their third-order invariant to decide its imaginarity or negativity. The following theorem, pieced together from results in Ref. [85], is a result in a similar direction, for imaginarity.

Theorem 10. *A triple of overlaps is realizable by three quantum states if and only if it is a convex combination of triples of overlaps realizable by pure states with real amplitudes. In other words, \mathcal{Q}_3^{\bullet} is the convex hull of \mathcal{R}_3° .*

Proof. It was shown in Ref. [85] that all the overlap triples that are extremal points of the set \mathcal{Q}_3^{\bullet} are realizable using real-amplitude pure states. For completeness, we proceed to sketch the proof of this result.

First, note that a Bargmann invariant for mixed states can be written as a convex combination of Bargmann invariants of pure states, using a convex decomposition of each mixed state into pure states. Fixing a decomposition for each mixed state in a set $\{\rho_i\}_i$, all Bargmann invariants drawn from the set can be written as the *same* convex mixture of Bargmann invariants among pure states. In particular, this applies to overlap triples, hence \mathcal{Q}_3^{\bullet} is contained in the convex hull of \mathcal{Q}_3° . It thus suffices to look at realizations by pure states to characterize the border of \mathcal{Q}_3^{\bullet} .

Now, any three pure states can be unitarily sent to a concrete representation of the form

$$\begin{aligned} |\psi_1\rangle &= |0\rangle, \\ |\psi_2\rangle &= \cos(\beta)|0\rangle + \sin(\beta)|1\rangle, \\ |\psi_3\rangle &= \cos(\gamma)|0\rangle + e^{i\phi}\sin(\gamma)\sin(\alpha)|1\rangle + \sin(\gamma)\cos(\alpha)|2\rangle, \end{aligned}$$

with $\alpha, \beta, \gamma \in [0, \pi/2)$ and $\phi \in [0, 2\pi)$. This fixes $\Delta_2(\psi_1, \psi_2) = \cos^2(\beta)$ and $\Delta_2(\psi_1, \psi_3) = \cos^2(\gamma)$ to be

real. The existence of imaginary values in the representation of the states may therefore only affect the third overlap $\Delta_2(\psi_2, \psi_3)$. To conclude the argument, the maximal values achievable by $\Delta_2(\psi_2, \psi_3)$ for any fixed values of β, γ can be found as the extreme points of the multivariate function $\Delta_2(\psi_2(\alpha, \phi), \psi_3(\alpha, \phi))$. Ref. [85, Appendix A] performs these calculations explicitly, showing this happens either if $\sin(\alpha) = 0$ or if $\sin(\phi) = 0$, both conditions that imply the three states have real amplitudes. We conclude that the overlap triples in the boundary of \mathcal{Q}_3° are realizable by pure states with real amplitudes only. Therefore, \mathcal{Q}_3^{\bullet} is in the convex hull of \mathcal{R}_3° .

For the other direction, it suffices to establish that \mathcal{Q}_3^{\bullet} is convex. In fact, \mathcal{Q}_3° (or even \mathcal{R}_3°) is. This was observed numerically in Ref. [85]. Moreover, its bounds were analytically found to correspond to the boundaries of an *elliptope*, discussed in detail in Ref. [88].¹⁴ More formally, we establish convexity by showing that if a triple of overlaps $(\Delta_{12}, \Delta_{13}, \Delta_{23})$ is pure-state realizable, then so is any triple in the line segment connecting it to $(0, 0, 0)$, i.e. any triple of the form $(\lambda\Delta_{12}, \lambda\Delta_{13}, \lambda\Delta_{23})$ for $\lambda \in [0, 1]$. Starting from $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$ the pure states in a Hilbert space \mathcal{H} realizing $(\Delta_{12}, \Delta_{13}, \Delta_{23})$, consider the larger Hilbert space $\mathcal{H} \oplus \mathbb{C}^3$ with $\{|e_i\rangle\}_{i=1}^3$ a basis for the additional summand, and take $|\psi'_i\rangle := \sqrt{\lambda}\lambda|\psi_i\rangle \oplus \sqrt{1-\lambda}|e_i\rangle$. Then, for $i \neq j$, $\langle\psi'_i|\psi'_j\rangle = \sqrt{\lambda}\langle\psi_i|\psi_j\rangle + (1-\sqrt{\lambda})\langle e_i|e_j\rangle = \sqrt{\lambda}\langle\psi_i|\psi_j\rangle$, and so $\Delta_2(\psi'_i, \psi'_j) = \lambda\Delta_2(\psi_i, \psi_j)$. Therefore, $\{|\psi'_i\rangle\}_{i=1}^3$ realizes $(\lambda\Delta_{12}, \lambda\Delta_{13}, \lambda\Delta_{23})$, as wanted. \square

In other words, this theorem shows that the realizable overlap triples can be described as convex combinations of overlap triples realizable by pure states that are real with respect to some reference basis. If Theorem 9 showed that an overlap triple does not *always* give enough information to decide imaginarity of the third-order Bargmann, Theorem 10 seems to indicate that it can *never*—or *hardly*—provide useful information in this regard.

These results highlight that, to test imaginarity using overlaps, one needs to make further assumptions on the states, such as on purity or the Hilbert space dimension. Indeed, a careful look at the proof of Theorem 9 suggests there is scope for such examples if one restricts to *pure* single-qubit states, $\mathcal{Q}_3^{\circ(2)}$.

It turns out that, for instance, three *pure* single-qubit states that have pairwise overlaps of $1/2$ must have an imaginary third-order invariant.

5.2.4. Negativity and overlaps for real amplitude states

In contrast to imaginarity, for negativity, we are able to obtain some positive results: restricting to realizations

¹⁴ There is a dual relationship between the scenarios investigated in Ref. [88] and those in Ref. [85]. For more information about these aspects, we refer the reader to the discussion in Ref. [86].

with real-amplitude states (i.e., to \mathcal{R}_3°), it is possible to draw conclusions about the negativity of third-order Bargmann invariants given only their pairwise overlaps. In fact, it turns out that negativity and overlap nonclassicality, as witnessed by the overlap inequalities such as Eq. 18 studied in Refs. [85–87], witness mutually exclusive forms of set coherence.

We first consider the two-dimensional case, for which we are able to obtain a full discriminating result, partitioning $\mathcal{R}_3^\circ(2)$. Given pure states $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$, the following inequalities are witnesses of set coherence: the first three are the overlap inequalities defining the polytope \mathcal{C}_3° ; the last is nonnegativity of the third-order Bargmann invariant.

$$\begin{aligned} +\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_2, \psi_3) - \Delta_2(\psi_1, \psi_3) &\leq 1, \\ +\Delta_2(\psi_1, \psi_2) - \Delta_2(\psi_2, \psi_3) + \Delta_2(\psi_1, \psi_3) &\leq 1, \\ -\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_2, \psi_3) + \Delta_2(\psi_1, \psi_3) &\leq 1, \\ \Delta_3(\psi_1, \psi_2, \psi_3) &\geq 0. \end{aligned} \quad (44)$$

If the ψ_i are one-qubit pure states with real amplitudes with respect to some basis, it turns out that no two of these inequalities can be violated. For any such choice of three states, there exists a unitary that maps them to the rebit subspace. This implies that we can parametrize the states in the following way;

$$\begin{aligned} |\psi_1\rangle &= |0\rangle, \\ |\psi_2\rangle &= \cos(\theta) |0\rangle + \sin(\theta) |1\rangle, \\ |\psi_3\rangle &= \cos(\phi) |0\rangle + \sin(\phi) |1\rangle. \end{aligned} \quad (45)$$

with $\theta, \phi \in [0, \pi]$ being sufficient to capture all possible values. This yields the following values for the Bargmann invariants:

$$\begin{aligned} \Delta_2(\psi_1, \psi_2) &= \cos^2(\theta), \\ \Delta_2(\psi_1, \psi_3) &= \cos^2(\phi), \\ \Delta_2(\psi_2, \psi_3) &= \cos^2(\theta - \phi), \\ \Delta_3(\psi_1, \psi_2, \psi_3) &= \cos(\theta - \psi) \cos(\theta) \cos(\phi). \end{aligned} \quad (46)$$

The domains in which each of the inequalities from Eq. (44) is violated are plotted in Fig. 7, in terms of the parameters θ and ϕ . There, we can observe that each value of θ and ϕ defining a triplet of states, except when those states are set-incoherent corresponding to borders of the colored regions, violates one and only one inequality.

We thus conclude that $\mathcal{R}_3^\circ(2)$ partitions into two *disjoint* subsets: $\mathcal{P}_3^\circ(2)$ and $\mathcal{C}_3^\circ(2)$, with the latter being the disjoint union of $\mathcal{N}_3^\circ(2)$ and $\mathcal{Z}_3^\circ(2)$ (the only triples with a set-incoherent realization, where at least one of the pairwise overlaps vanishes).

Turning our attention to arbitrary dimension (i.e., to \mathcal{R}_3°), this observation leads us to search for similar results, converging on the lemma below. Even in this more general case, violation of overlap inequalities and negativity of Δ_3 are mutually exclusive, i.e., $\mathcal{N}_3^\circ \subseteq \mathcal{C}_3^\circ$.

Lemma 11. *Let $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$ be pure states that are real with respect to some basis, in a finite-dimensional*

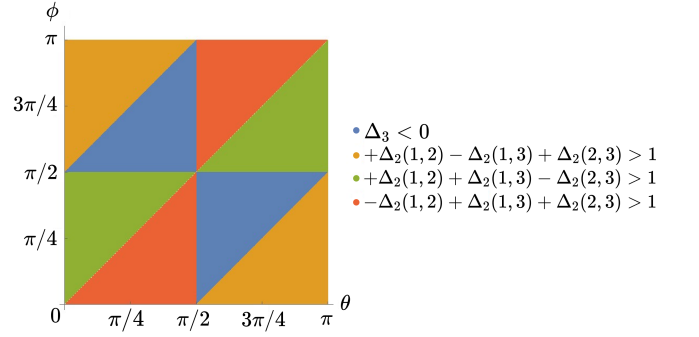


Figure 7. **Negativity of third-order invariants from measurements of overlaps only.** For real-amplitude one-qubit states, we show a complementarity relation between negativity of third-order invariants and overlap inequalities. Here we use the notation $\Delta_2(i, j) = |\langle \psi_i | \psi_j \rangle|^2$ and $\Delta_3 = \langle \psi_1 | \psi_2 \rangle \langle \psi_2 | \psi_3 \rangle \langle \psi_3 | \psi_1 \rangle$.

Hilbert space \mathcal{H} of dimension d . Then, violation of any of the inequalities

$$\begin{aligned} \Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_3) - \Delta_2(\psi_2, \psi_3) &\leq 1, \\ \Delta_2(\psi_1, \psi_2) - \Delta_2(\psi_1, \psi_3) + \Delta_2(\psi_2, \psi_3) &\leq 1, \\ -\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_3) + \Delta_2(\psi_2, \psi_3) &\leq 1. \end{aligned} \quad (47)$$

implies that

$$\Delta_3(\psi_1, \psi_2, \psi_3) > 0. \quad (48)$$

In other words, $(\mathcal{R}_3^\circ \setminus \mathcal{C}_3^\circ) \subseteq \mathcal{P}_3^\circ$ and thus $\mathcal{N}_3^\circ \subseteq \mathcal{C}_3^\circ$.

If, moreover, $d = 2$, the converse also holds: if all overlap inequalities are strictly satisfied (i.e., strictly less than 1), we obtain $\Delta_3 < 0$.

We prove this lemma in Appendix C. The part of the lemma regarding dimension $d = 2$ was already shown by our previous argument and is depicted in Fig. 7. Overlaps provide a simple test for when the third-order invariants are positive, in which case it might simplify the search for negativity.

6. DISCUSSION AND FUTURE DIRECTIONS

In this work, we have presented two distinct approaches for the study of central concepts in quantum mechanics, namely, weak values, KD distributions, and state spectra.

In this work, we have explored two complementary approaches to a better understanding of various selected constructions useful in quantum information processing, including weak values, KD distributions, and state spectra. The first approach is more practical and application-focused, where we show simple quantum circuits that can be used to estimate various constructions of fundamental interest by measuring multivariate traces known as Bargmann invariants. Specifically, we have described circuits to experimentally estimate the spectrum of any

mixed quantum state, weak values, KD quasiprobability distribution, post-selected quantum Fisher information, and OTOCs. The second approach is more foundational and investigates conditions for the nonclassicality of these functions, in particular, negativity and imaginarity, relating it to the presence of a recently introduced quantum resource termed set coherence, which acknowledges coherence as a basis-independent, relational property of a set of states.

These results bring proof-of-principle tests of the nonclassicality of quantum theory—a vital resource for information processing—closer to current experimental capabilities. Indirect measurements of unitary invariants up to fourth order have been made in different platforms [71, 129, 130]. We believe that the unified view of nonclassicality we discussed here may boost the interest in a systematic exploration of nonclassicality witnessed and quantified by unitary invariants. Moreover, the connection between coherence theory and all the above constructions motivates not only the development of a formal resource theory based on Bargmann invariants but also the experimental investigation of coherence for higher-dimensional systems through the lens of resource-theoretic monotones.

The relationship between Bargmann invariants and KD distributions could bring quantitative understanding to other problems of interest. For instance, nonclassicality of Bargmann invariants understood as witness of coherence may clarify the connection between the nonclassicality of OTOCs and scrambling of information [14] for quantum computation. This new view could lead to more robust and rigorous ways of benchmarking the scrambling of information against decoherence that goes beyond numerical evidence [131]. Another interesting point to notice is that the tools described in Ref. [16] and further developed here clarify the understanding of when it is indeed *crucial* to experimentally estimate higher-order invariants in extended KD distributions. It would be interesting to characterize the minimum cardinality and/or order of sets of Bargmann invariants that are necessary for quantifying nonclassicality in particular applications.

We believe that our circuits and the unified picture in terms of Bargmann invariants will motivate further progress in analyzing and quantifying nonclassicality and its relation with quantum information advantage for different tasks. As an example, in Ref. [87], the authors prove a quantum advantage for the task of quantum interrogation relying on a function of overlaps. Our work suggests that higher-order invariants may provide a more complete characterization of advantages for general tasks.

Conceptually, we expect future work to consolidate—or reject—the view that violations of different constraints over Bargmann invariants, in the form of inequalities,

are a helpful guide to identifying different nonclassical resources provided by quantum mechanics. We view some of our results as preliminary steps in this direction. A complete answer to these questions, together with a formal framework for addressing the possible consequences for quantum information science, has yet to be investigated.

As a final technical remark, even though some of the results regarding the use of the univariate traces $\text{Tr}(\rho^n)$ for $n = 2, \dots, d$ have previously appeared in the literature, we stress that our novel sample complexity analysis and numerical experiments have shown concrete benefits in learning the spectrum using the higher-order Bargmann invariants. Interesting future work in this direction could involve experimentally probing $\text{Tr}(\rho^n)$ for small $n = 2, 3, 4$ and using the obtained information to estimate the spectrum of density matrices. These estimates can serve as subroutines in variational quantum eigensolvers aimed at preparing Gibbs states, such as those in Ref. [10].

Given the foundational importance of KD distributions and weak values, we expect that our unified framework based on measurements of Bargmann invariants will provide both theoretical insight and practical recipes for experimental implementations of the many applications we analyze here.

ACKNOWLEDGMENTS

We would like to thank Stephan De Bièvre, Mafalda Francisco Ramôa da Costa Alves, Francesco Buscemi, and Agung Budiyo for providing useful comments on an early version of this work. We would also like to thank an anonymous referee for pointing out the simple yet elegant Eq. (22), which provides a link between the extended KD distribution and higher-order Bargmann invariants.

We thank the International Iberian Nanotechnology Laboratory (INL) and the Bar-Ilan Institute of Nanotechnology and Advanced Materials (BINA) for their support, which was instrumental in kick-starting this collaboration. RW and ARM acknowledge support from FCT—Fundação para a Ciência e a Tecnologia (Portugal) through PhD Grants SFRH/BD/151199/2021 and SFRH/BD/151453/2021, respectively. RSB and EFG also acknowledge support from FCT, CEECINST/00062/2018. ILP acknowledges support from the ERC Advanced Grant FLQuant. The work of the INL-based authors (RW, ARM, RSB, EFG) was supported by the Digital Horizon Europe project FoQaCiA, GA no. 101070558. EC was supported by the Israeli Innovation Authority under Project 73795 and the Eureka project, by Elta Systems Ltd., by the Pazy Foundation, by the Israeli Ministry of Science and Technology, and by the Quantum Science and Technology Program of the Israeli Council of Higher Education.

[1] 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, [Phys. Rev.](#)

- Lett.* **60**, 1351 (1988).
- [2] J. G. Kirkwood, Quantum statistics of almost classical assemblies, *Phys. Rev.* **44**, 31 (1933).
 - [3] P. A. M. Dirac, On the analogy between classical and quantum mechanics, *Rev. Mod. Phys.* **17**, 195 (1945).
 - [4] N. Y. Halpern, B. Swingle, and J. Dressel, Quasiprobability behind the out-of-time-ordered correlator, *Phys. Rev. A* **97**, 042105 (2018).
 - [5] M. Lostaglio, A. Belenchia, A. Levy, S. Hernández-Gómez, N. Fabbri, and S. Gherardini, Kirkwood-Dirac quasiprobability approach to the statistics of incompatible observables, *Quantum* **7**, 1128 (2023).
 - [6] E. Cohen and E. Pollak, Determination of weak values of quantum operators using only strong measurements, *Phys. Rev. A* **98**, 042112 (2018).
 - [7] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
 - [8] A. Streltsov, G. Adesso, and M. B. Plenio, Colloquium: Quantum coherence as a resource, *Rev. Mod. Phys.* **89**, 041003 (2017).
 - [9] E. Tirrito, P. S. Tarabunga, G. Lami, T. Chanda, L. Leone, S. F. E. Oliviero, M. Dalmonte, M. Collura, and A. Hama, [Quantifying non-stabilizerness through entanglement spectrum flatness](#), arXiv:2304.01175 [quant-ph] (2023).
 - [10] Y. Wang, G. Li, and X. Wang, Variational quantum Gibbs state preparation with a truncated Taylor series, *Phys. Rev. Appl.* **16**, 054035 (2021).
 - [11] T. Giordani, C. Esposito, F. Hoch, G. Carvacho, D. J. Brod, E. F. Galvão, N. Spagnolo, and F. Sciarrino, Witnesses of coherence and dimension from multiphoton indistinguishability tests, *Phys. Rev. Res.* **3**, 023031 (2021).
 - [12] S. Designolle, R. Uola, K. Luoma, and N. Brunner, Set coherence: Basis-independent quantification of quantum coherence, *Phys. Rev. Lett.* **126**, 220404 (2021).
 - [13] D. R. M. Arvidsson-Shukur, N. Y. Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, Quantum advantage in postselected metrology, *Nat. Comm.* **11**, 3775 (2020).
 - [14] J. R. González Alonso, N. Yunger Halpern, and J. Dressel, Out-of-time-ordered-correlator quasiprobabilities robustly witness scrambling, *Phys. Rev. Lett.* **122**, 040404 (2019).
 - [15] J. R. G. Alonso, N. Shammah, S. Ahmed, F. Nori, and J. Dressel, [Diagnosing quantum chaos with out-of-time-ordered-correlator quasiprobability in the kicked-top model](#), arXiv:2201.08175 [quant-ph] (2022).
 - [16] M. Oszmaniec, D. J. Brod, and E. F. Galvão, [Measuring relational information between quantum states, and applications](#), arXiv:2109.10006 [quant-ph] (2021).
 - [17] Y. Quek, M. M. Wilde, and E. Kaur, [Multivariate trace estimation in constant quantum depth](#), arXiv:2206.15405 [quant-ph] (2022).
 - [18] J. Yirka and Y. Subaşı, Qubit-efficient entanglement spectroscopy using qubit resets, *Quantum* **5**, 535 (2021).
 - [19] J.-P. Escofier, *Galois Theory* (Springer-Verlag, New York, NY, 2001).
 - [20] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Colloquium: Understanding quantum weak values: Basics and applications, *Rev. Mod. Phys.* **86**, 307 (2014).
 - [21] L. Vaidman, Weak value controversy, *Philos. Trans. R. Soc. A* **375**, 20160395 (2017).
 - [22] M. F. Pusey, Anomalous weak values are proofs of contextuality, *Phys. Rev. Lett.* **113**, 200401 (2014).
 - [23] R. Kunjwal, M. Lostaglio, and M. F. Pusey, Anomalous weak values and contextuality: robustness, tightness, and imaginary parts, *Phys. Rev. A* **100**, 042116 (2019).
 - [24] E. Rebuffello, F. Piacentini, A. Avella, M. A. d. Souza, M. Gramegna, J. Dziewior, E. Cohen, L. Vaidman, I. P. Degiovanni, and M. Genovese, Anomalous weak values via a single photon detection, *Light Sci. Appl.* **10**, 106 (2021).
 - [25] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Ultrasensitive beam deflection measurement via interferometric weak value amplification, *Phys. Rev. Lett.* **102**, 173601 (2009).
 - [26] M. D. Turner, C. A. Hagedorn, S. Schlamminger, and J. H. Gundlach, Picoradian deflection measurement with an interferometric quasi-autocollimator using weak value amplification, *Opt. Lett.* **36**, 1479 (2011).
 - [27] Y. Susa, Y. Shikano, and A. Hosoya, Optimal probe wave function of weak-value amplification, *Phys. Rev. A* **85**, 052110 (2012).
 - [28] J. Dressel, K. Lyons, A. N. Jordan, T. M. Graham, and P. G. Kwiat, Strengthening weak-value amplification with recycled photons, *Phys. Rev. A* **88**, 023821 (2013).
 - [29] A. N. Jordan, J. Martínez-Rincón, and J. C. Howell, Technical advantages for weak-value amplification: when less is more, *Phys. Rev. X* **4**, 011031 (2014).
 - [30] S. Pang, J. Dressel, and T. A. Brun, Entanglement-assisted weak value amplification, *Phys. Rev. Lett.* **113**, 030401 (2014).
 - [31] G. B. Alves, B. Escher, R. L. de Matos Filho, N. Zagury, and L. Davidovich, Weak-value amplification as an optimal metrological protocol, *Phys. Rev. A* **91**, 062107 (2015).
 - [32] J. Harris, R. W. Boyd, and J. S. Lundeen, Weak value amplification can outperform conventional measurement in the presence of detector saturation, *Phys. Rev. Lett.* **118**, 070802 (2017).
 - [33] M. Pfender, P. Wang, H. Sumiya, S. Onoda, W. Yang, D. B. R. Dasari, P. Neumann, X.-Y. Pan, J. Isoya, R.-B. Liu, *et al.*, High-resolution spectroscopy of single nuclear spins via sequential weak measurements, *Nat. Commun.* **10**, 594 (2019).
 - [34] K. S. Cujia, J. M. Boss, K. Herb, J. Zopes, and C. L. Degen, Tracking the precession of single nuclear spins by weak measurements, *Nature* **571**, 230 (2019).
 - [35] S.-Z. Fang, H.-T. Tan, G.-X. Li, and Q.-L. Wu, Weak value amplification for angular velocity measurements, *Appl. Opt.* **60**, 4335 (2021).
 - [36] J.-H. Huang, X.-Y. Duan, and X.-Y. Hu, Amplification of rotation velocity using weak measurements in Sagnac's interferometer, *Eur. Phys. J. D* **75**, 114 (2021).
 - [37] I. L. Paiva, R. Lenny, and E. Cohen, Geometric phases and the Sagnac effect: Foundational aspects and sensing applications, *Adv. Quantum Tech.* **5**, 2100121 (2022).
 - [38] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen, Direct Measurement of the Density Matrix of a Quantum System, *Phys. Rev. Lett.* **117**, 120401 (2016).
 - [39] J. S. Lundeen and C. Bamber, Procedure for Direct Measurement of General Quantum States Using Weak Measurement, *Phys. Rev. Lett.* **108**, 070402 (2012).
 - [40] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, Direct measurement of the quantum

- wavefunction, *Nature* **474**, 188 (2011).
- [41] Z. Shi, M. Mirhosseini, J. Margiewicz, M. Malik, F. Rivera, Z. Zhu, and R. W. Boyd, Scan-free direct measurement of an extremely high-dimensional photonic state, *Optica* **2**, 388 (2015).
- [42] M. Malik, M. Mirhosseini, M. P. J. Lavery, J. Leach, M. J. Padgett, and R. W. Boyd, Direct measurement of a 27-dimensional orbital-angular-momentum state vector, *Nat. Commun.* **5**, 3115 (2014).
- [43] J. Lundeen and K. Resch, Practical measurement of joint weak values and their connection to the annihilation operator, *Phys. Lett. A* **334**, 337 (2005).
- [44] J. S. Lundeen and A. M. Steinberg, Experimental joint weak measurement on a photon pair as a probe of hardy's paradox, *Phys. Rev. Lett.* **102**, 020404 (2009).
- [45] A. C. Martinez-Becerril, G. Bussi eres, D. Curic, L. Giner, R. A. Abrahao, and J. S. Lundeen, Theory and experiment for resource-efficient joint weak-measurement, *Quantum* **5**, 599 (2021).
- [46] Y. Kim, Y.-S. Kim, S.-Y. Lee, S.-W. Han, S. Moon, Y.-H. Kim, and Y.-W. Cho, Direct quantum process tomography via measuring sequential weak values of incompatible observables, *Nat. Commun.* **9**, 192 (2018).
- [47] L. M. Johansen and A. Luis, Nonclassicality in weak measurements, *Phys. Rev. A* **70**, 052115 (2004).
- [48] L. M. Johansen, Quantum theory of successive projective measurements, *Phys. Rev. A* **76**, 012119 (2007).
- [49] L. M. Johansen and P. A. Mello, Quantum mechanics of successive measurements with arbitrary meter coupling, *Phys. Lett. A* **372**, 5760 (2008).
- [50] C. Bamber and J. S. Lundeen, Observing Dirac's Classical Phase Space Analog to the Quantum State, *Phys. Rev. Lett.* **112**, 070405 (2014).
- [51] G. S. Thekkadath, R. Y. Saaltink, L. Giner, and J. S. Lundeen, Determining Complementary Properties with Quantum Clones, *Phys. Rev. Lett.* **119**, 050405 (2017).
- [52] H. F. Hofmann, How Weak Values Emerge in Joint Measurements on Cloned Quantum Systems, *Phys. Rev. Lett.* **109**, 020408 (2012).
- [53] J. Z. Salvail, M. Agnew, A. S. Johnson, E. Bolduc, J. Leach, and R. W. Boyd, Full characterization of polarization states of light via direct measurement, *Nat. Photonics* **7**, 316 (2013).
- [54] S. Hern andez-G omez, S. Gherardini, A. Belenchia, M. Lostaglio, A. Levy, and N. Fabbri, *Projective measurements can probe non-classical work extraction and time-correlations*, arXiv:2207.12960 [quant-ph] (2023).
- [55] A. Budiyo and H. K. Dipojono, Quantifying quantum coherence via Kirkwood-Dirac quasiprobability, *Phys. Rev. A* **107**, 022408 (2023).
- [56] H. Cram er, *Mathematical methods of statistics* (Princeton University Press, Princeton, NJ, 1946).
- [57] C. R. Rao, Information and the accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.* **37**, 81 (1945).
- [58] A. Ly, M. Marsman, J. Verhagen, R. Grasman, and E.-J. Wagenmakers, A tutorial on Fisher information, *J. Math. Psychol.* **80**, 40 (2017).
- [59] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, Quantum Fisher information matrix and multiparameter estimation, *J. Phys. A* **53**, 023001 (2019).
- [60] M. G. A. Paris, Quantum estimation for quantum technology, *Int. J. Quantum Inf.* **7**, 125 (2009).
- [61] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [62] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
- [63] K. Bharti, Fisher information: A crucial tool for NISQ research, *Quantum Views* **5**, 61 (2021).
- [64] J. J. Meyer, Fisher information in noisy intermediate-scale quantum applications, *Quantum* **5**, 539 (2021).
- [65] X. Zhang, X.-M. Lu, J. Liu, W. Ding, and X. Wang, Direct measurement of quantum Fisher information, *Phys. Rev. A* **107**, 012414 (2023).
- [66] F. Salvati, W. Salmon, C. H. W. Barnes, and D. R. M. Arvidsson-Shukur, Compression of metrological quantum information in the presence of noise (2023), arXiv:2307.08648 [quant-ph].
- [67] J. H. Jenne and D. R. M. Arvidsson-Shukur, Unbounded and lossless compression of multiparameter quantum information, *Phys. Rev. A* **106**, 042404 (2022).
- [68] N. Lupu-Gladstein, Y. B. Yilmaz, D. R. Arvidsson-Shukur, A. Brodutch, A. O. Pang, A. M. Steinberg, and N. Y. Halpern, Negative quasiprobabilities enhance phase estimation in quantum-optics experiment, *Phys. Rev. Lett.* **128**, 220504 (2022).
- [69] S. Das, S. Modak, and M. N. Bera, Saturating quantum advantages in postselected metrology with the positive Kirkwood-Dirac distribution, *Phys. Rev. A* **107**, 042413 (2023).
- [70] V. Bargmann, Note on Wigner's theorem on symmetry operations, *J. Math. Phys.* **5**, 862 (1964).
- [71] A. E. Jones, S. Kumar, S. D'Aurelio, M. Bayerbach, A. J. Menssen, and S. Barz, *Distinguishability and mixedness in quantum interference*, arXiv:2201.04655 [quant-ph] (2022).
- [72] A. Wigderson, Mathematics and computation, in *Mathematics and Computation* (Princeton University Press, Princeton, NJ, 2019).
- [73] J. C. Garcia-Escartin and P. Chamorro-Posada, A SWAP gate for qudits, *Quantum Inf. Process.* **12**, 3625 (2013).
- [74] P. Horodecki, From limits of quantum operations to multicopy entanglement witnesses and state-spectrum estimation, *Phys. Rev. A* **68**, 052101 (2003).
- [75] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying entanglement, *Phys. Rev. Lett.* **78**, 2275 (1997).
- [76] F. A. Bovino, G. Castagnoli, A. Ekert, P. Horodecki, C. M. Alves, and A. V. Sergienko, Direct Measurement of Nonlinear Properties of Bipartite Quantum States, *Phys. Rev. Lett.* **95**, 240407 (2005).
- [77] A. J. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, Measuring entanglement growth in quench dynamics of bosons in an optical lattice, *Phys. Rev. Lett.* **109**, 020505 (2012).
- [78] Y. Subaşı, L. Cincio, and P. J. Coles, Entanglement spectroscopy with a depth-two quantum circuit, *J. Phys. A* **52**, 044001 (2019).
- [79] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Direct estimations of linear and nonlinear functionals of a quantum state, *Phys. Rev. Lett.* **88**, 217901 (2002).
- [80] C. M. Alves, P. Horodecki, D. K. L. Oi, L. C. Kwek, and A. K. Ekert, Direct estimation of functionals of density operators by local operations and classical communica-

- tion, *Phys. Rev. A* **68**, 032306 (2003).
- [81] P. Horodecki and A. Ekert, Method for direct detection of quantum entanglement, *Phys. Rev. Lett.* **89**, 127902 (2002).
- [82] T. Tanaka, Y. Ota, M. Kanazawa, G. Kimura, H. Nakazato, and F. Nori, Determining eigenvalues of a density matrix with minimal information in a single experimental setting, *Phys. Rev. A* **89**, 012117 (2014).
- [83] S. J. van Enk and C. W. J. Beenakker, Measuring $\text{Tr}\rho^n$ on single copies of ρ using random measurements, *Phys. Rev. Lett.* **108**, 110503 (2012).
- [84] T.-Y. Chien and S. Waldron, A characterization of projective unitary equivalence of finite frames and applications, *SIAM J. Discrete Math.* **30**, 976 (2016).
- [85] E. F. Galvão and D. J. Brod, Quantum and classical bounds for two-state overlaps, *Phys. Rev. A* **101**, 062110 (2020).
- [86] R. Wagner, R. S. Barbosa, and E. F. Galvão, *Inequalities witnessing coherence, nonlocality, and contextuality*, arXiv:2209.02670 [quant-ph] (2022).
- [87] R. Wagner, A. Camillini, and E. F. Galvão, *Coherence and contextuality in a Mach-Zehnder interferometer*, arXiv:2210.05624 [quant-ph] (2022).
- [88] T. P. Le, C. Meroni, B. Sturmfels, R. F. Werner, and T. Ziegler, Quantum correlations in the minimal scenario, *Quantum* **7**, 947 (2023).
- [89] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, UK, 2000).
- [90] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators, *New J. Phys.* **14**, 095022 (2012).
- [91] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Am. Stat. Assoc.* **58**, 13 (1963).
- [92] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms* (Cambridge University Press, New York, NY, 2014).
- [93] S. Chen, B. Huang, J. Li, A. Liu, and M. Sellke, *When does adaptivity help for quantum state learning?*, arXiv:2206.05265 [quant-ph] (2022).
- [94] L. M. Johansen, Reconstructing weak values without weak measurements, *Phys. Lett. A* **366**, 374 (2007).
- [95] G. Vallone and D. Dequal, Strong measurements give a better direct measurement of the quantum wave function, *Phys. Rev. Lett.* **116**, 040502 (2016).
- [96] N. Brunner and C. Simon, Measuring small longitudinal phase shifts: weak measurements or standard interferometry?, *Phys. Rev. Lett.* **105**, 010405 (2010).
- [97] C.-F. Li, X.-Y. Xu, J.-S. Tang, J.-S. Xu, and G.-C. Guo, Ultrasensitive phase estimation with white light, *Phys. Rev. A* **83**, 044102 (2011).
- [98] J. Dressel and A. N. Jordan, Significance of the imaginary part of the weak value, *Phys. Rev. A* **85**, 012107 (2012).
- [99] K. Mitarai and K. Fujii, Methodology for replacing indirect measurements with direct measurements, *Phys. Rev. Res.* **1**, 013006 (2019).
- [100] L. Mazzola, G. De Chiara, and M. Paternostro, Measuring the characteristic function of the work distribution, *Phys. Rev. Lett.* **110**, 230602 (2013).
- [101] F. Buscemi, M. Dall’Arno, M. Ozawa, and V. Vedral, *Direct observation of any two-point quantum correlation function*, arXiv:1312.4240 [quant-ph] (2013).
- [102] K. J. Resch and A. M. Steinberg, Extracting joint weak values with local, single-particle measurements, *Phys. Rev. Lett.* **92**, 130402 (2004).
- [103] G. Mitchison, R. Jozsa, and S. Popescu, Sequential weak measurement, *Phys. Rev. A* **76**, 062105 (2007).
- [104] N. Yunger Halpern, Jarzynski-like equality for the out-of-time-ordered correlator, *Phys. Rev. A* **95**, 012120 (2017).
- [105] P. Rall, Quantum algorithms for estimating physical quantities using block encodings, *Phys. Rev. A* **102**, 022408 (2020).
- [106] M. Esposito, U. Harbola, and S. Mukamel, Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems, *Rev. Mod. Phys.* **81**, 1665 (2009).
- [107] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa, and M. Genovese, Measuring incompatible observables by exploiting sequential weak values, *Phys. Rev. Lett.* **117**, 170402 (2016).
- [108] M. A. Nielsen and I. L. Chuang, Programmable quantum gate arrays, *Phys. Rev. Lett.* **79**, 321 (1997).
- [109] R. O’Donnell and J. Wright, Quantum spectrum testing, in *Proceedings of the forty-seventh annual ACM symposium on Theory of computing* (ACM, 2015) pp. 529–538.
- [110] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going beyond Bell’s theorem, in *Bell’s Theorem, Quantum Theory and Conceptions of the Universe* (Kluwer Academic Publishers, Amsterdam, NL, 1989) pp. 69–72.
- [111] Link for the Newton’s identity algorithm that recovers the spectrum from the polynomials:, <https://github.com/Rumoa/example-f1> (2022).
- [112] W. Bruzda, V. Cappellini, H.-J. Sommers, and K. Życzkowski, Random quantum operations, *Phys. Lett. A* **373**, 320 (2009).
- [113] M. Keyl and R. F. Werner, Estimating the spectrum of a density operator, *Phys. Rev. A* **64**, 052311 (2001).
- [114] R. Alicki, S. Rudnicki, and S. Sadowski, Symmetry properties of product states for the system of N n -level atoms, *J. Math. Phys.* **29**, 1158 (1988).
- [115] J. Wright, *How to learn a quantum state*, Ph.D. thesis, Carnegie Mellon University (2016).
- [116] R. O’Donnell and J. Wright, Efficient quantum tomography, in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing* (ACM, 2016) pp. 899–912.
- [117] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, Sample-optimal tomography of quantum states, in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing* (ACM, 2016) pp. 913–925.
- [118] H. Yuen, An improved sample complexity lower bound for (fidelity) quantum state tomography, *Quantum* **7**, 890 (2023).
- [119] D. Bacon, I. L. Chuang, and A. W. Harrow, Efficient quantum circuits for Schur and Clebsch-Gordan transforms, *Phys. Rev. Lett.* **97**, 170502 (2006).
- [120] M. E. Beverland, J. Haah, G. Alagic, G. K. Campbell, A. M. Rey, and A. V. Gorshkov, Spectrum estimation of density operators with Alkaline-Earth atoms, *Phys. Rev. Lett.* **120**, 025301 (2018).
- [121] T. A. Brun, Measuring polynomial functions of states, *Quantum Inf. Comput.* **4**, 401 (2004).
- [122] M. Arnhem, C. Griffet, and N. J. Cerf, Multicopy observ-

ables for the detection of optically nonclassical states, *Phys. Rev. A* **106**, 043705 (2022).

- [123] R. Wagner and E. F. Galvão, Simple proof that anomalous weak values require coherence, *Phys. Rev. A* **108**, L040202 (2023).
- [124] D. R. M. Arvidsson-Shukur, J. C. Drori, and N. Y. Halpern, Conditions tighter than noncommutation needed for nonclassicality, *J. Phys. A* **54**, 284001 (2021).
- [125] S. De Bièvre, Complete incompatibility, support uncertainty, and Kirkwood-Dirac nonclassicality, *Phys. Rev. Lett.* **127**, 190404 (2021).
- [126] S. D. Bièvre, Relating incompatibility, noncommutativity, uncertainty, and Kirkwood-Dirac nonclassicality, *J. Math. Phys.* **64**, 022202 (2023).
- [127] J. Xu, Classification of incompatibility for two orthonormal bases, *Phys. Rev. A* **106**, 022217 (2022).
- [128] J. Xu, *Kirkwood-Dirac classical pure states*, arXiv:2210.02876 [quant-ph] (2022).
- [129] T. Bitter and D. Dubbers, Manifestation of Berry's topological phase in neutron spin rotation, *Phys. Rev. Lett.* **59**, 251 (1987).
- [130] D. Suter, K. T. Mueller, and A. Pines, Study of the Aharonov-Anandan quantum phase by NMR interferometry, *Phys. Rev. Lett.* **60**, 1218 (1988).
- [131] J. Harris, B. Yan, and N. A. Sinitsyn, Benchmarking information scrambling, *Phys. Rev. Lett.* **129**, 050602 (2022).
- [132] A. Hariri, D. Curic, L. Giner, and J. S. Lundeen, Experimental simultaneous readout of the real and imaginary parts of the weak value, *Phys. Rev. A* **100**, 032119 (2019).

Appendix A: Formal comparison between standard weak measurement and circuit protocol

In this section, we estimate the sample complexity of finding weak values using the Bargmann invariants scheme and compare it with the sample complexity of measuring weak values using the standard protocol, known as weak measurement. Operationally, the two procedures are drastically different, which makes the comparison in terms of purely complexity arguments difficult. Therefore, for simplicity of the argument, we consider observables $A = |a\rangle\langle a|$. Recall that weak values can be written as

$$A_w = \frac{\Delta_3(\phi, a, \psi)}{\Delta_2(\phi, \psi)}. \quad (\text{A1})$$

1. Same number of samples for numerator and denominator

In order to compute A_w with the cycle test, we must compute the quantities in the numerator and denominator separately: using the SWAP-test for one and the cycle test associated with the operator C_3 for the other. We can focus solely on the complexity of estimating the real part A_w . The same analysis can be repeated for the imaginary part. Then, since $\Delta_2(\phi, \psi)$ is real, we must determine the sample complexity related to estimating

the quantity $\text{Re}[A_w] = \text{Re}[\Delta_3(\phi, a, \psi)]/\Delta_2(\phi, \psi)$, i.e., we can focus on the real part of $\Delta_3(\phi, a, \psi)$ for the estimation of the numerator of A_w .

The results provided by runs of such tests can be described by random variables X_i taking values in $\{\pm 1\}$, with i denoting each run. Denoting by $p_+ = (1 + \text{Re}[\Delta_3(\phi, a, \psi)])/2$ and $p_- = 1 - p_+$ the probabilities of X_i being $+1$ and -1 , respectively, we write X_i 's expectation value as $\mathbb{E}[X_i] = +p_+ - p_- = \text{Re}[\Delta_3(\phi, a, \psi)]$.

Assuming we have $N^{(3)}$ runs associated with X_i , where superscript (3) refers to the fact that we are estimating third-order invariant, we consider the total random variable $X := \sum_{i=1}^{N^{(3)}} X_i$. By linearity of \mathbb{E} , $\mathbb{E}[X] = N^{(3)}\text{Re}[\Delta_3(\phi, a, \psi)]$. We, then, apply Hoeffding's inequality, a standard method in learning theory and statistical analysis [91, 92]. According to this bound, if a_i and b_i are real constants such that $a_i \leq X_i \leq b_i$ for every $i \in \{1, \dots, N\}$ and $X = \sum_i X_i$, we have that, for any $t > 0$,

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq t] \leq 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^{N^{(3)}} (b_i - a_i)^2}\right). \quad (\text{A2})$$

In our case, $a_i = -1$ and $b_i = +1$ for every i , which allows us to rewrite the above expression as

$$\mathbb{P}\left[|X - N^{(3)}\text{Re}[\Delta_3(\phi, a, \psi)]| \geq t\right] \leq 2e^{-2t^2/4N^{(3)}}, \quad (\text{A3})$$

which implies that

$$\mathbb{P}\left[\left|\frac{X}{N^{(3)}} - \text{Re}[\Delta_3(\phi, a, \psi)]\right| \geq \frac{t}{N^{(3)}}\right] \leq 2e^{-t^2/2N^{(3)}}. \quad (\text{A4})$$

Defining $\delta^{(3)} := 2e^{-t^2/2N^{(3)}}$, we get that $\delta^{(3)}/2 = e^{-t^2/2N^{(3)}} \Rightarrow t = \sqrt{2N^{(3)} \ln(2/\delta^{(3)})}$. Hence, we can ensure that, with probability greater than $1 - \delta^{(3)}$, the estimator $X/N^{(3)}$ gives the true value of the real third-order invariant within an error of $\sqrt{2 \ln(2/\delta^{(3)})}/N^{(3)}$, i.e.,

$$\mathbb{P}\left[\left|\frac{X}{N^{(3)}} - \text{Re}[\Delta_3(\phi, a, \psi)]\right| \geq \sqrt{\frac{2 \ln(2/\delta^{(3)})}{N^{(3)}}}\right] \leq \delta^{(3)}. \quad (\text{A5})$$

Similarly, we can introduce random variables Y_i for $N^{(2)}$ runs of the SWAP test, each with expectation $\mathbb{E}[Y_i] = \Delta_2(\phi, \psi)$, and the respective counterpart of other relevant parameters to write

$$\mathbb{P}\left[\left|\frac{Y}{N^{(2)}} - \Delta_2(\phi, \psi)\right| \geq \sqrt{\frac{2 \ln(2/\delta^{(2)})}{N^{(2)}}}\right] \leq \delta^{(2)}, \quad (\text{A6})$$

where we have used the fact that $\Delta_2(a, b) \equiv \text{Re}[\Delta_2(a, b)]$ for any pair of states $|a\rangle, |b\rangle \in \mathcal{H}$.

In our protocol, we have two different procedures: one in which we use $N^{(3)}$ samples to estimate $\text{Re}[\Delta_3(\phi, a, \psi)]$ and another in which we use $N^{(2)}$ samples to estimate

$\Delta_2(\phi, \psi)$. Defining $\varepsilon^{(n)} := \sqrt{2 \ln(2/\delta^{(2)})/N^{(2)}}$, we can write

$$\Delta_2(\phi, \psi) - \varepsilon^{(2)} \leq \frac{Y}{N^{(2)}} \leq \Delta_2(\phi, \psi) + \varepsilon^{(2)} \quad (\text{A7})$$

or simply

$$\left| \frac{Y}{N^{(2)}} \right| \geq |\Delta_2(\phi, \psi) - \varepsilon^{(2)}|. \quad (\text{A8})$$

At first, let us assume that we have equal choices for the precision with a high probability of learning the related second- and third-order invariants, i.e., $\varepsilon^{(2)} = \varepsilon^{(3)} \equiv \varepsilon$ and $\delta^{(2)} = \delta^{(3)} \equiv \delta$. This implies that $N^{(2)} = N^{(3)} \equiv N$. Then, with a probability greater than $1 - 2\delta$, we have

$$\begin{aligned} \left| \frac{X}{Y} - \frac{\text{Re}[\Delta_3(\phi, a, \psi)]}{\Delta_2(\phi, \psi)} \right| &= \\ &= \left| \frac{\frac{X}{N}\Delta_2 - \frac{Y}{N}\text{Re}[\Delta_3]}{\frac{Y}{N}\Delta_2} \right| \\ &= \left| \frac{\frac{X}{N}\Delta_2 - \Delta_2\text{Re}[\Delta_3] + \Delta_2\text{Re}[\Delta_3] - \frac{Y}{N}\text{Re}[\Delta_3]}{\frac{Y}{N}\Delta_2} \right| \\ &\leq \frac{\Delta_2|\frac{X}{N} - \text{Re}[\Delta_3]| + |\text{Re}[\Delta_3]||\frac{Y}{N} - \Delta_2|}{\Delta_2\frac{Y}{N}} \\ &\leq \frac{\Delta_2\varepsilon + |\text{Re}[\Delta_3]|\varepsilon}{\Delta_2|\Delta_2 - \varepsilon|} = \frac{1 + |\text{Re}[A_w]|}{\left| \frac{\Delta_2}{\varepsilon} - 1 \right|}. \end{aligned} \quad (\text{A9})$$

In the last inequality, we have used that $|Y/N| \geq |\Delta_2 - \varepsilon|$, as seen in Eq. (A8). In the above expression, we have also simplified the notation using $\Delta_2 \equiv \Delta_2(\phi, \psi)$ and $\Delta_3 \equiv \Delta_3(\phi, a, \psi)$. As a result, we conclude that

$$\mathbb{P} \left[\left| \frac{X}{Y} - \text{Re}[A_w] \right| \geq \frac{1 + |\text{Re}[A_w]|}{\left| \frac{\Delta_2}{\varepsilon} - 1 \right|} \right] \leq 2\delta, \quad (\text{A10})$$

i.e., the resulting accuracy ε_T of estimating the real part of the weak value is

$$\varepsilon_T = \frac{1 + |\text{Re}[A_w]|}{\left| \frac{\Delta_2}{\varepsilon} - 1 \right|}. \quad (\text{A11})$$

Since $\varepsilon = \sqrt{\ln(2/\delta)/N}$, we see that the number of samples needed scales as

$$N = O \left(\frac{\ln(2/\delta)|\text{Re}[A_w]|^2}{\Delta_2^2 \varepsilon_T^2} \right). \quad (\text{A12})$$

This allows us to compare the sample complexity of measuring weak values using the cycle test with the sample complexity associated with standard weak measurement schemes¹⁵.

In the standard weak measurement, one has a successful post-selection given by $N_s = |\langle \phi | \psi \rangle|^2 N$ with N the total number of samples used. For the successful incidents, the measurement outcomes are assumed to be Gaussian distributed with variance satisfying $\sigma_i^2 \gg \gamma^2 |A_w|^2$.

Then, for every $i \in \{1, \dots, N_s\}$, $\mathbb{E}[X_i] = \gamma \text{Re}[A_w]$. Also, $\text{Var}[X] = N_s \sigma_i^2$. Given that $\mathbb{E}[X/N_s \gamma] = \text{Re}[A_w]$ and because the distribution of successful incidents is Gaussian distributed, we have that the probability within a certain region is guaranteed by the error function

$$\mathbb{P} \left[\left| \frac{X}{\gamma N_s} - \text{Re}[A_w] \right| \geq \varepsilon \right] \leq 1 - \text{erf} \left(\frac{\varepsilon}{\sqrt{2 \text{Var}(X/\gamma N_s)}} \right). \quad (\text{A13})$$

To simplify the notation, we use that $\text{Var}(X/\gamma N_s) = \text{Var}(X)/\gamma^2 N_s^2 = \sigma_i^2/\gamma^2 N_s$ and define

$$\delta := \text{erfc} \left(\sqrt{\frac{\varepsilon^2 \gamma^2 \Delta_2 N}{2 \sigma_i^2}} \right), \quad (\text{A14})$$

where $\text{erfc}(x) = 1 - \text{erf}(x)$ is the complement of the error function. Then, the necessary total number of samples N to estimate $\text{Re}[A_w]$ using $X/\gamma N_s$ is

$$N = 2(\text{erfc}^{-1}(\delta))^2 \frac{\sigma_i^2}{\varepsilon^2 \gamma^2 \Delta_2} \gg 2(\text{erfc}^{-1}(\delta))^2 \frac{|A_w|^2}{\varepsilon^2 \Delta_2}, \quad (\text{A15})$$

which gives an order of samples better than the case of the cycle test when $\Delta_2 \rightarrow 0$ since, in this case,

$$N^{(\text{weak})} = O \left(\text{erfc}^{-1}(\delta)^2 \frac{|A_w|^2}{\varepsilon^2 \Delta_2} \right). \quad (\text{A16})$$

This implies that there might be no complex theoretic advantage in using the cycle test protocol to implement weak value amplification, for which $\Delta_2(\phi, \psi)$ must be extremely small. However, this does not exclude the possibility that, for specific setups, there might be benefits in using the cycle test—e.g., in cases where the constant C related to the weak coupling is exponentially big. As pointed out in the main text, for simply witnessing *non-classical values*, the cycle test *does* provide a sample complexity advantage since one does not need to estimate $\Delta_2(\phi, \psi)$. In particular, this feature might be interesting in specific setups of quantum sensing using the imaginary part of weak values.

2. Different number of samples for numerator and denominator

Given that we expect $\Delta_2(\phi, \psi) \ll 1$ in various applications of weak values, e.g., weak value amplification, it is relevant to consider the use of more samples to estimate the overlap such that $N^{(2)} > N^{(3)}$. We can, therefore, analyze this particular situation to observe if (and how) the sample complexity of estimating the weak value changes.

If we do not assume that $N^{(2)} = N^{(3)}$, we have

$$\begin{aligned}
& \left| \frac{X}{Y} - \frac{N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)]}{N^{(2)} \Delta_2(\phi, \psi)} \right| = \left| \frac{N^{(2)} \Delta_2(\phi, \psi) X - N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)] Y}{Y N^{(2)} \Delta_2(\phi, \psi)} \right| = \\
& = \left| \frac{N^{(2)} \Delta_2(\phi, \psi) X - N^{(2)} N^{(3)} \Delta_2(\phi, \psi) \operatorname{Re}[\Delta_3(\phi, a, \psi)] + N^{(2)} N^{(3)} \Delta_2(\phi, \psi) \operatorname{Re}[\Delta_3(\phi, a, \psi)] - N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)] Y}{Y N^{(2)} \Delta_2(\phi, \psi)} \right| \\
& \leq \frac{|N^{(2)} \Delta_2(\phi, \psi) X - N^{(2)} N^{(3)} \Delta_2(\phi, \psi) \operatorname{Re}[\Delta_3(\phi, a, \psi)]| + |N^{(2)} N^{(3)} \Delta_2(\phi, \psi) \operatorname{Re}[\Delta_3(\phi, a, \psi)] - N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)] Y|}{Y N^{(2)} \Delta_2(\phi, \psi)} \\
& = \frac{N^{(2)} \Delta_2(\phi, \psi) |X - N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)]| + N^{(3)} |\operatorname{Re}[\Delta_3(\phi, a, \psi)]| |N^{(2)} \Delta_2(\phi, \psi) - Y|}{Y N^{(2)} \Delta_2(\phi, \psi)}. \tag{A17}
\end{aligned}$$

Hoeffding's inequality leads to

$$\mathbb{P}[|X - N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)]| \geq \sqrt{2N^{(3)} \ln(2/\delta^{(3)})}] \leq \delta^{(3)} \tag{A18}$$

and

$$\mathbb{P}[|Y - N^{(2)} \Delta_2(\phi, \psi)| \geq \sqrt{2N^{(2)} \ln(2/\delta^{(2)})}] \leq \delta^{(2)}, \tag{A19}$$

implying that, with probability $\geq 1 - \delta^{(2)}$,

$$\frac{1}{Y} \leq \frac{1}{N^{(2)} \Delta_2(\phi, \psi) - \sqrt{2N^{(2)} \ln(2/\delta^{(2)})}} \approx \frac{1}{N^{(2)} \Delta_2(\phi, \psi)} \tag{A20}$$

for a sufficiently large $N^{(2)}$. In turn, this leads to

$$\begin{aligned}
\left| \frac{X}{Y} - \frac{N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)]}{N^{(2)} \Delta_2(\phi, \psi)} \right| & \leq \frac{N^{(2)} \Delta_2(\phi, \psi) \sqrt{2N^{(3)} \ln(2/\delta^{(3)})} + N^{(3)} |\operatorname{Re}[\Delta_3(\phi, a, \psi)]| \sqrt{2N^{(2)} \ln(2/\delta^{(2)})}}{N^{(2)} N^{(3)} \Delta_2(\phi, \psi)^2} \\
\left| \frac{X}{Y} - \frac{N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)]}{N^{(2)} \Delta_2(\phi, \psi)} \right| & \leq \frac{\sqrt{N^{(3)}}}{N^{(2)} \Delta_2(\phi, \psi)} \sqrt{2 \ln(2/\delta^{(3)})} + \frac{N^{(3)}}{(N^{(2)})^{3/2} \Delta_2(\phi, \psi)^2} |\operatorname{Re}[\Delta_3(\phi, a, \psi)]| \sqrt{2 \ln(2/\delta^{(2)})}, \tag{A21}
\end{aligned}$$

from which we conclude that

$$\left| \frac{\frac{X}{N^{(3)}} - \frac{\operatorname{Re}[\Delta_3(\phi, a, \psi)]}{\Delta_2(\phi, \psi)}}{\frac{Y}{N^{(2)}}} \right| = \frac{N^{(2)}}{N^{(3)}} \left| \frac{X}{Y} - \frac{N^{(3)} \operatorname{Re}[\Delta_3(\phi, a, \psi)]}{N^{(2)} \Delta_2(\phi, \psi)} \right| \leq \underbrace{\frac{\sqrt{2 \ln(2/\delta^{(3)})}}{\sqrt{N^{(3)}} \Delta_2(\phi, \psi)} + \frac{|\operatorname{Re}[\Delta_3(\phi, a, \psi)]| \sqrt{2 \ln(2/\delta^{(2)})}}{\sqrt{N^{(2)}} \Delta_2(\phi, \psi)^2}}_{=:\varepsilon}. \tag{A22}$$

Defining $\delta := \delta^{(2)} + \delta^{(3)}$, we write

$$\mathbb{P} \left[\left| \frac{\frac{X}{N^{(3)}} - \frac{\operatorname{Re}[\Delta_3(\phi, a, \psi)]}{\Delta_2(\phi, \psi)}}{\frac{Y}{N^{(2)}}} \right| \geq \varepsilon \right] \leq \delta. \tag{A23}$$

Since we have already analyzed the case in which $N^{(2)} = N^{(3)}$, we now focus on the case in which they differ. Let us capture this difference with a parameter c , letting $N^{(2)} \sim N^{(3)} / \Delta_2(\phi, \psi)^c$. Since our main interest lies in scenarios $\Delta_2(\phi, \psi)$ is assumed to be small, we want to study the situation $N^{(2)} \gg N^{(3)}$ to see if we have some improvement in the sample complexity. Then, we assume c is positive and $\Delta_2(\phi, \psi)^2 \ll 1$. As a result, the total number of samples satisfies $N := N^{(2)} + N^{(3)} \simeq N^{(3)} / \Delta_2(\phi, \psi)^c$. This implies that

$$\varepsilon = \frac{\sqrt{2 \ln(2/\delta^{(3)})}}{\sqrt{N^{(3)}} \Delta_2(\phi, \psi)} + \frac{|\operatorname{Re}[\Delta_3(\phi, a, \psi)]| \sqrt{2 \ln(2/\delta^{(2)})}}{\sqrt{N^{(2)}} \Delta_2(\phi, \psi)^2} = \frac{\sqrt{2(\ln(4/\delta))}}{\sqrt{N}} \left(\frac{1}{\Delta_2(\phi, \psi)^{1+c/2}} + \frac{|\operatorname{Re}[\Delta_3(\phi, a, \psi)]|}{\Delta_2(\phi, \psi)^2} \right), \tag{A24}$$

¹⁵ It is noteworthy that, differently from the protocol we consider next, there are weak measurement schemes able to estimate

both the real and imaginary parts of the weak value simultaneously [132].

from which we can extract the dependency in the number of samples with precision ε in learning the real part of the weak value:

$$N \simeq \frac{2 \ln(4/\delta)}{\varepsilon^2} \left(\frac{1}{\Delta_2(\phi, \psi)^{1+c/2}} + \frac{|\operatorname{Re}[\Delta_3(\phi, a, \psi)]|}{\Delta_2(\phi, \psi)^2} \right)^2. \quad (\text{A25})$$

Finally, let us study the number of samples by considering the possible values of the parameter c :

1. Case $c = 2$: We have $1 + c/2 = 2$. Then, we are left with an order of number of samples equal to the one when we had $N^{(2)} = N^{(3)}$ since $N = \frac{2 \ln(4/\delta)}{\varepsilon^2 \Delta_2(\phi, \psi)^4} (1 + |\operatorname{Re}[\Delta_3(\phi, a, \psi)]|)^2$. Note that, in the results presented in the previous subsection, the number of samples refers to the estimation of the real part of the weak value while, here, it refers to the estimation of the third-order invariant. This is the reason for the difference seen in the order of the overlap.
2. Case $0 < c < 2$: Assuming that $|\operatorname{Re}[\Delta_3(\phi, a, \psi)]| = O(1)$ and that it does not scale with $\Delta_2(\phi, \psi)$, it follows that $|\operatorname{Re}[\Delta_3(\phi, a, \psi)]|/\Delta_2(\phi, \psi) \gg 1/\Delta_2(\phi, \psi)^{1+c/2}$ for small $\Delta_2(\phi, \psi)$. After substituting these approximations into the equation for N , we are left with the same sample complexity as before.
3. Case $c > 2$: Direct calculation leads to $N \simeq \frac{2 \ln(4/\delta)}{\varepsilon^2 \Delta_2(\phi, \psi)^{2+c}}$, which is worse in sample complexity than the previous results, and worsens the larger the value the parameter c takes.

In summary, these calculations show that the sample complexity does not improve statistically if one estimates the overlaps with a larger, or even much larger, number of samples than the number of samples necessary to estimate the third-order invariant in the measurement of weak values.

Appendix B: Sample and measurement complexity of estimating the spectrum using the cycle test

1. Proof of Theorem 6

To estimate $\{\operatorname{Tr}(\rho^2), \dots, \operatorname{Tr}(\rho^d)\}$, we want to learn with high probability $1 - \delta$ the entire set of estimators within an error ε . Using the union bound,

$$\begin{aligned} & \mathbb{P} \left[\bigcup_{i=2}^d \left\{ \hat{\rho}^i : |\operatorname{Tr}(\hat{\rho}^i) - \operatorname{Tr}(\rho^i)| \geq \varepsilon \right\} \right] \\ & \leq \sum_{i=2}^d \mathbb{P} \left[|\operatorname{Tr}(\hat{\rho}^i) - \operatorname{Tr}(\rho^i)| \geq \varepsilon \right] \\ & = (d-1) \mathbb{P} \left[|\operatorname{Tr}(\hat{\rho}^i) - \operatorname{Tr}(\rho^i)| \geq \varepsilon \right]. \end{aligned} \quad (\text{B1})$$

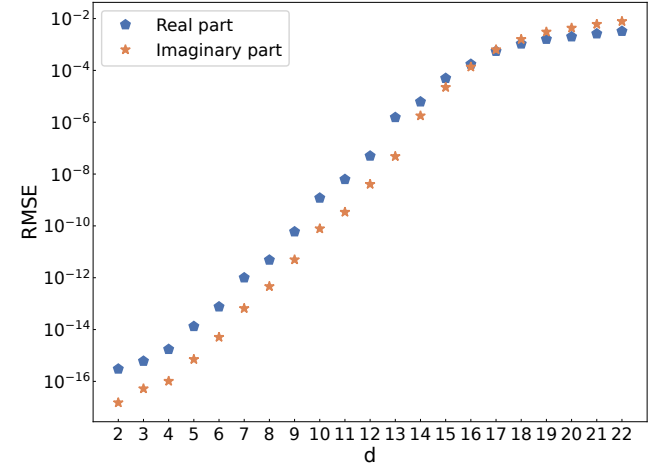


Figure 8. **Aggregated root-mean-squared error (RMSE) spectrum estimation using the Faddeev-LeVerrier Algorithm.** Eigenvalues of mixed states computed from multivariate traces $\{\operatorname{Tr}(\rho^n)\}_{n=1}^d$ are compared with the ones obtained by diagonalization of the density matrix using *numpy*. For each dimension, a data set of random 1000 density matrices ρ is used. We sample the states using the Ginibre ensemble.

We fix $\mathbb{P} \left[|\operatorname{Tr}(\hat{\rho}^i) - \operatorname{Tr}(\rho^i)| \geq \varepsilon \right] = \delta/(d-1)$ such that the total error probability is δ .

From Hoeffding's inequality, for N independent samples $X_i \in [a, b]$ and mean estimator, the number of *statistical* samples needed to estimate each trace is, as already mentioned in Appendix A, $O(\ln(2/\delta)/\varepsilon^2)$. But each statistical sample requires $O(d^2)$ states to compute $\{\operatorname{Tr}(\rho^2), \dots, \operatorname{Tr}(\rho^d)\}$ using the cycle test. The total number of states is, therefore, $Nd^2 = \frac{d^2}{\varepsilon^2} \ln \sqrt{2(d-1)}/\delta$ and, as a consequence, $N = O\left(\frac{d^2}{\varepsilon^2} \ln(d/\delta)\right)$ samples estimate the chosen set of traces to precision ε with probability $1 - \delta$. Since the number of measurements equals the number of statistical samples times d , which accounts for one measurement for each element in the set of traces, we have that the number of measurements is of order $O\left(\frac{d}{\varepsilon^2} \ln(d/\delta)\right)$. This concludes the proof.

In the main text, we reported the sample complexity $\tilde{O}(d^2/\varepsilon^2)$ that hides $\log(d)$ factors and fixed δ to be a reference value to prescribe what is meant by "high probability," usually taken to be equal to $1 - \delta = 2/3$.

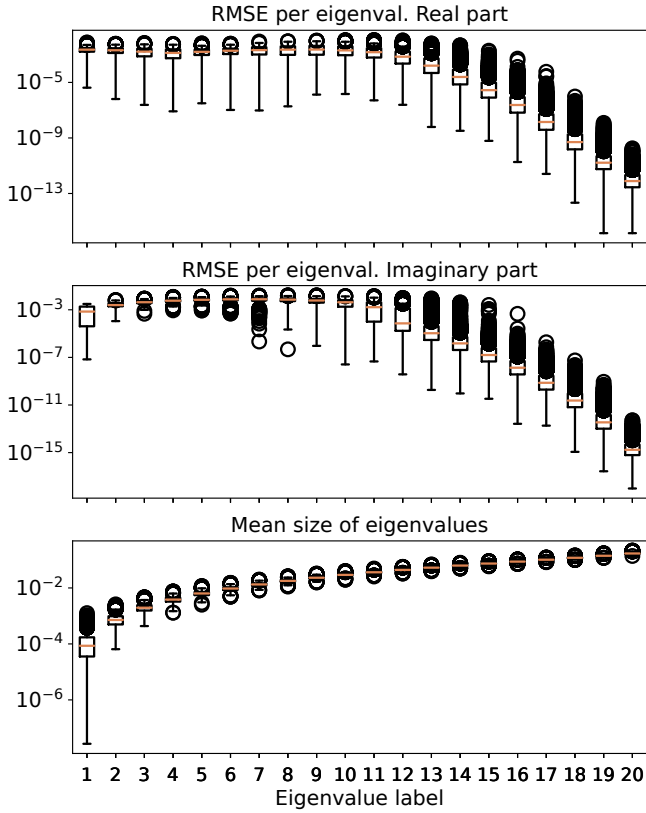


Figure 9. **Mean size of eigenvalues and RMSE per eigenvalue.** For 1000 random mixed density matrices of fixed Hilbert space dimension 20, we study the real and imaginary parts of the spectrum given from the Faddeev–LeVerrier algorithm.

2. Numerical analysis for the Faddeev–LeVerrier algorithm

In this section, we conduct a numerical study of the behavior of the Faddeev–LeVerrier algorithm [19, Chapter 3], which returns the estimates of the eigenvalues of ρ , given $\{\text{Tr}(\rho^i)\}_{i=2}^d$ as an input.

In Fig. 8, using noiseless experimental data, i.e., the traces $\{\text{Tr}(\rho^i)\}_i$, we show the aggregated RMSE of the real and imaginary parts of the spectrum of ρ . We compare the predicted spectrum with the one computed via diagonalization with *numpy*.

Ideally, the spectrum should be real-valued. However, as mentioned in the main text, numerical instabilities generate imaginary parts in the outputs of the algorithm. At a certain dimension, the RMSE of the imaginary part of all eigenvalues *exceeds* the one for the real part. This is expected, as numerical inaccuracies occur when very small roots need to be estimated. As a result, the aggregated RMSE for high dimensions can be orders of magnitude bigger than the actual size of some eigenvalues. In what follows, we study if the appearance of these imaginary values, as well as their size, significantly influences the comparison between the real part of the spectrum and

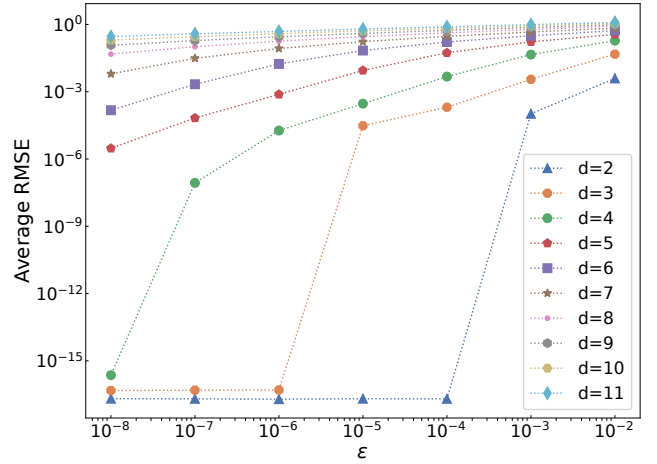


Figure 10. **Average RMSE of the imaginary parts of the spectrum estimation of 5000 random ρ .** This figure represents the imaginary counterpart of Fig. 5. Each matrix generates 1000 noisy samples to compute the RMSE.

the ideal spectrum.

Let us take as a case study the error *per eigenvalue* for a sufficiently large Hilbert space dimension. In Fig. 9, we show the average size of the eigenvalues using 1000 random density matrices of dimension 20. The biggest eigenvalues have sizes of order 10^{-2} while the associated error between the real and imaginary part is smaller than 10^{-13} . On the other hand, the smallest eigenvalues have an error approximately equal to the size of the eigenvalue, which explains the aggregated RMSE from Fig. 8 in higher dimensions.

Although the predicted eigenvalues can contain non-negligible imaginary parts, their presence does not affect the prediction of the real parts. These imaginary values come from the presence of small coefficients in the characteristic polynomial. In Fig. 10, we show the RMSE of the imaginary part in the estimation of the spectrum using noisy data. The sudden increase for a certain dimension is closely related to the amount of noise introduced. For larger dimensions, smaller amounts of noise modify the coefficients enough to produce the appearance of imaginary values. However, by comparing with Fig. 5, the trend of the RMSE of real parts is not affected by these “bumps” in the imaginary parts, allowing us to safely discard them from the predictions.

a. Learning the largest eigenvalue

As smaller eigenvalues produce bigger errors, the Faddeev–LeVerrier algorithm can be used to estimate only the largest eigenvalues. In Fig. 11, we show the RMSE of the real part in the prediction of the largest eigenvalue using the same noisy dataset as the one used for the prediction of the entire spectrum. Since the largest

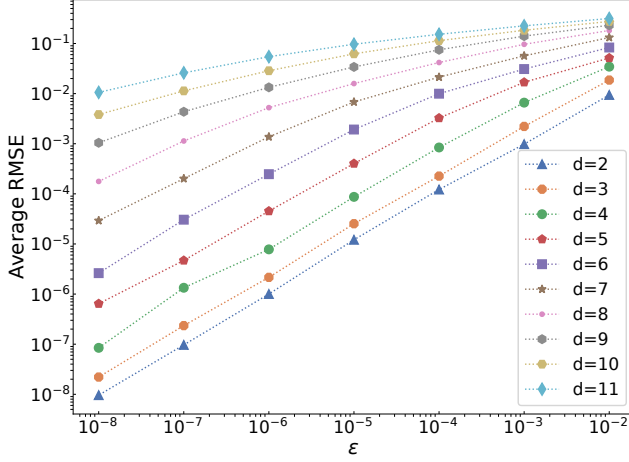


Figure 11. **Average RMSE of the real part of largest eigenvalue for 5000 random density matrices.** Each matrix generates 1000 noisy samples to compute the RMSE. We see that, when focusing on the largest eigenvalue, the accuracy of the estimation improves when compared to the results for the entire spectrum estimation from Fig. 5 in the main text.

eigenvalue depends mostly on the coefficients of the highest order, the size of these makes them less sensitive to the noise and allows the obtaining of a RMSE of 10^{-3} in dimension 6 with an error $\varepsilon = 10^{-4}$ for the traces. However, as stated in the main text, computing the entire set of traces is not recommended in this task as higher powers in the traces of ρ are more difficult to estimate and contribute less to the largest eigenvalue.

Appendix C: Proof of Lemma 11

Consider a general triplet of pure states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ real with respect to some basis and choose the following parametrization for it:

$$\begin{aligned} |\psi_1\rangle &= |0\rangle \\ |\psi_2\rangle &= \cos(\beta)|0\rangle + \sin(\beta)|1\rangle \\ |\psi_3\rangle &= \cos(\gamma)|0\rangle + \sin(\gamma)\sin(\alpha)|1\rangle + \sin(\gamma)\cos(\alpha)|2\rangle, \end{aligned} \quad (\text{C1})$$

where $\beta, \gamma \in [0, \pi]$ and $\alpha \in [0, 2\pi)$. Under this choice, made without loss of generality, we can define the function

$$\begin{aligned} h_1(\alpha, \beta, \gamma) &= -\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_3) + \Delta_2(\psi_2, \psi_3), \\ h_2(\alpha, \beta, \gamma) &= +\Delta_2(\psi_1, \psi_2) - \Delta_2(\psi_1, \psi_3) + \Delta_2(\psi_2, \psi_3), \\ h_3(\alpha, \beta, \gamma) &= +\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_3) - \Delta_2(\psi_2, \psi_3), \end{aligned} \quad (\text{C2})$$

where each overlap depends on the variables α, β, γ . We can, then, study what the constraints $h_i > 1$ represent,

i.e.,

$$\begin{aligned} h_1 > 1 &\Rightarrow \\ 0 < -1 - \cos^2(\beta) + \cos^2(\gamma) + \cos^2(\beta)\cos^2(\gamma) &+ \underbrace{\sin^2(\beta)\sin^2(\gamma)\sin^2(\alpha)}_{\in[0,1]} + \frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha) \\ &\leq -1 - \cos^2(\beta) + \cos^2(\gamma) + \cos^2(\beta)\cos^2(\gamma) + \frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha) \\ &= -2\cos^2(\beta) + 2\cos^2(\beta)\sin^2(\gamma) \\ &\quad + \frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha) \\ &= -2\cos^2(\beta) + 2\Delta_3(\alpha, \beta, \gamma) \leq 2\Delta_3(\alpha, \beta, \gamma), \end{aligned} \quad (\text{C3})$$

which implies that $\Delta_3(\alpha, \beta, \gamma) \equiv \Delta_3(\psi_1, \psi_2, \psi_3) > 0$, where we have used that

$$\begin{aligned} \Delta_3(\alpha, \beta, \gamma) &= \langle \psi_1 | \psi_2 \rangle \langle \psi_2 | \psi_3 \rangle \langle \psi_3 | \psi_1 \rangle \\ &= \cos^2(\beta)\cos^2(\gamma) + \frac{1}{4}\sin(2\beta)\sin(2\gamma)\sin(\alpha). \end{aligned} \quad (\text{C4})$$

The same reasoning applies for $h_2 > 1$. For $h_3 > 1$, the situation is less straightforward since the two functions $h_3(\alpha, \beta, \gamma) - 1$ and $\Delta_3(\alpha, \beta, \gamma)$ are not totally ordered in the domain $[0, 2\pi) \times [0, \pi]^2$. However, we can analyze what $\Delta_3(\alpha, \beta, \gamma) > 0$ corresponds to in terms of the parameters β and γ when α lies in some domains. From the above, it follows that

$$\tan(\beta)\tan(\gamma) > -\frac{1}{\sin(\alpha)} \Rightarrow \Delta_3(\alpha, \beta, \gamma) > 0 \quad (\text{C5})$$

if $\alpha \in [0, \pi]$. The case in which $\alpha \in [\pi, 2\pi]$ can be treated similarly. The function $h_3(\alpha, \beta, \gamma) - 1$ can be written as

$$\begin{aligned} h_3(\alpha, \beta, \gamma) - 1 &= \\ &= -1 + \cos^2(\beta) + \cos^2(\gamma) - \cos^2(\beta)\cos^2(\gamma) \\ &\quad - \sin^2(\beta)\sin^2(\gamma)\sin^2(\alpha) - \frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha) \\ &= -\sin^2(\beta)\sin^2(\gamma) - \sin^2(\beta)\sin^2(\gamma)\sin^2(\alpha) \\ &\quad - \frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha) \\ &= -\sin^2(\beta)\sin^2(\gamma)(1 + \sin^2(\alpha)) \\ &\quad - \frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha). \end{aligned} \quad (\text{C6})$$

The condition that $h_3(\alpha, \beta, \gamma) - 1 > 0$ implies that $\sin^2(\beta)\sin^2(\gamma)(1 + \sin^2(\alpha)) < -\sin(2\beta)\sin(2\gamma)\sin(\alpha)/2$.

Lemma 12. For any $\alpha, \beta, \gamma \in [0, \pi]$, if

$$\sin^2(\beta)\sin^2(\gamma)(1 + \sin^2(\alpha)) < -\frac{1}{2}\sin(2\beta)\sin(2\gamma)\sin(\alpha), \quad (\text{C7})$$

then

$$\tan(\beta)\tan(\gamma) > -\frac{1}{\sin(\alpha)}. \quad (\text{C8})$$

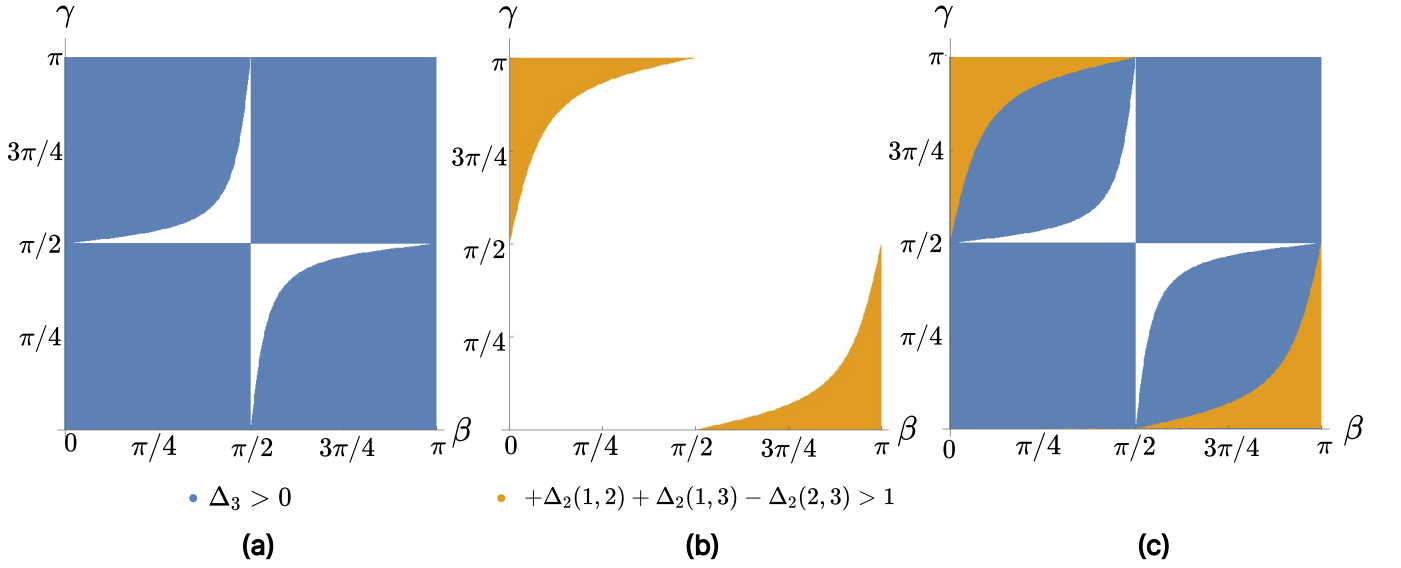


Figure 12. **Witnesses of third-order positivity.** (a) The white regions in the center are the area corresponding to $\Delta_3 \leq 0$. The blue region is given by $\Delta_3(\alpha, \beta, \gamma) > 0$ for $\alpha = 0.11$. (b) The light-brown region overlaps with the blue region completely, and it is given by $h_3(\alpha, \beta, \gamma) - 1 > 0$, showing that this condition implies $\Delta_3(\alpha, \beta, \gamma) > 0$. However, the converse does not necessarily hold. In fact, it is true only for $\alpha \in [0, \pi]$, as can be seen in (c), where we combine both configurations (a) and (b).

Proof. This lemma can be easily seen to hold numerically. In the range of $\alpha, \beta, \gamma \in [0, \pi]$, we see that there is always an overlap between the region for which the first inequality is satisfied simultaneously with the second inequality. The converse is not necessarily true. \square

The above lemma, together with Eq. (C5), implies that $h_3 - 1 > 0 \Rightarrow \Delta_3 > 0$, as we wanted to show. The case

$\alpha \in [\pi, 2\pi]$ is treated similarly, as already mentioned.

Fixing $\alpha = 0.11$, Fig. 12 shows the region of values $(\beta, \gamma) \in [0, \pi]$ for which the third-order invariant positivity can be witnessed with the overlap inequality. This provides a visual way to see that $\Delta_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_3) - \Delta_2(\psi_2, \psi_3) > 1$ implies positivity of the third-order Bargmann invariant, while the converse does not always hold.