

# Negativity Bounds for Weyl–Heisenberg Quasiprobability Representations

John B. DeBrota<sup>1</sup> · Christopher A. Fuchs<sup>1</sup>

Received: 1 March 2017 / Accepted: 8 May 2017 / Published online: 17 May 2017  
© Springer Science+Business Media New York 2017

**Abstract** The appearance of negative terms in quasiprobability representations of quantum theory is known to be inevitable, and, due to its equivalence with the onset of contextuality, of central interest in quantum computation and information. Until recently, however, nothing has been known about how much negativity is necessary in a quasiprobability representation. Zhu (Phys Rev Lett 117 (12):120404, 2016) proved that the upper and lower bounds with respect to one type of negativity measure are saturated by quasiprobability representations which are in one-to-one correspondence with the elusive symmetric informationally complete quantum measurements (SICs). We define a family of negativity measures which includes Zhu’s as a special case and consider another member of the family which we call “sum negativity.” We prove a sufficient condition for local maxima in sum negativity and find exact global maxima in dimensions 3 and 4. Notably, we find that Zhu’s result on the SICs does not generally extend to sum negativity, although the analogous result does hold in dimension 4. Finally, the Hoggar lines in dimension 8 make an appearance in a conjecture on sum negativity.

**Keywords** Quasiprobability representations · Symmetric informationally complete measurements · Frame decompositions · Negativity measures

## 1 Introduction

The mathematical machinery of quantum theory has persisted without substantial modification for nearly a century, but we are still waiting for a compelling set of

---

✉ Christopher A. Fuchs  
qbism.fuchs@gmail.com

<sup>1</sup> University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA 02125, USA

physical principles upon which to hang the theory's predictions. It is the hope of some quantum foundations researchers that looking at standard quantum theory re-represented in an appropriate fashion will help suggest these principles [2].

Probabilities are central objects in canonical quantum theory; at the end of a quantum mechanical calculation, we are left with a probability distribution or a simple consequence of one. It is tempting, therefore, to behave as though quantum theory gives us a probability distribution—the probabilities for a set of outcomes. Careful consideration reveals, however, that *given a quantum state*, quantum theory allows us to calculate a probability distribution. But where does the quantum state come from in the first place? The apparatus of quantum theory is unable to say. In practice, an experimenter eventually settles on a quantum state for her preparation procedure after a suite of tests and calibrations, and, ultimately, statistical methods go into the state determination itself. That is, we have probabilities at the beginning and probabilities at the end with the formal apparatus of quantum theory gluing it all together. It is possible that what goes on in between these ends stands alone, but such a circumstance is far from guaranteed. The proper understanding of quantum theory may depend upon conceiving of probability theory in the proper way.

In fact, revisiting and deciding on the proper understanding of probability theory is the starting point of QBism [3–5]. QBists take a strict personalist Bayesian [6–8] stance on probability theory. A probability is a valuation an agent places on his or her degree of belief in a possible outcome, nothing more and nothing less. As a consequence, probabilities are not empirically determinable quantities because they do not independently exist outside of an agent's mind.

As an example, consider repeated flips of a coin. A frequentist conception of probability asserts that the probability of heads for the coin is the long-run ratio of number of heads to number of flips. The Bayesian first points out that to regard the coin flipped at different times as an “equivalent” or “exchangeable” process amounts to a belief the experimenter has about the situation—perhaps nothing is wrong with this belief, but she should be cognizant of its influence on the conclusions of the experiment. Secondly, the Bayesian asks just how many times the experimenter plans to actually flip the coin before she decides that the relative frequency is the probability of heads for that coin. If she flips it a finite number of times, by her own admission, any frequency is technically possible (although she believes some are unlikely). The usual answer to such a question is that she plans to flip it until the deviations in the ratio with further experimentation are small enough so as to be negligible. In other words, in order to *define* the likelihood of an outcome, she asserts that the ratio obtained by a finite series of experiments is *likely* to be close to the “true probability”. The circularity of such an argument should be evident and worrying to anyone espousing the frequentist paradigm. Properly understood, then, probabilities are single-case; no probability is meaningfully right or wrong by any *external* criterion. The knee-jerk reaction to this statement is to exasperatedly throw up one's hands and exclaim that Bayesianism is just the claim that probability theory is useless! Nothing could be further from the truth: although a probability is not subject to objective external validation or invalidation, it commands the same sway over our lives that it would if it were; if an agent wishes to avoid sure loss—to avoid being demonstrably stupid—she must take steps to ensure that she never assigns probabilities which mutually contradict each other.

When a theory tells us what we “should” do or “strive for”, it is a *normative* theory. In other words, the personalist Bayesian view is that probability theory is a normative theory. In this community, compatible probabilities are called *coherent*. Remarkably, nearly all of the standard rules of probability theory are consequences of coherence [7].

One consequence that we will reference soon is the law of total probability (LTP):

$$q(j) = \sum_i p(i)r(j|i). \quad (1)$$

The LTP describes a scenario where two actions are taken, one after another.  $p(i)$  represents the probability we ascribe to getting outcome  $i$  from the first action,  $r(j|i)$  is the probability we ascribe to getting outcome  $j$  from the second action conditioned on outcome  $i$  for the first action, and  $q(j)$  is the probability we ascribe to outcome  $j$  for the second action, not conditioned on anything other than the operational procedure we have laid out. The commitment to coherence alone (and independent from any possible nature of reality) requires that our probabilities assigned at any given moment should hold together in accordance with (1).

As we explain below, it is possible to represent any quantum state as a single probability distribution over the possible outcomes of an appropriately chosen measurement. If we take this fact seriously, a quantum state is conceptually nothing more than a probability distribution. In QBism, all of the personalist Bayesian properties of probability theory carry over to quantum states; that is, quantum states, like probabilities, are valuations of belief for future experiences. However, application of the rules of quantum theory reveal that not all of the probability distributions in the probability simplex correspond to a valid quantum state. Just as in probability theory where an agent strives to be consistent with herself in her probability assignments, an agent should not ascribe a probability distribution she knows to be in conflict with the quantum mechanical formalism. In this way we arrive at an understanding that quantum theory is an empirically-motivated normative addition to probability theory.

If the functional form of the additions to probability calculus are cumbersome, then there may be no reason to adopt it for everyday use—furthermore, it may not shed any light on the “nature of reality.” What would constitute a nice looking addition to probability theory? One possibility would be if the normative rules of quantum theory could be made to mirror those of probability theory in a suggestive way. It turns out that just this sort of situation can be made to occur.

An *informationally complete quantum measurement (IC-POVM)* for a Hilbert space  $\mathcal{H}_d$  is a set of at least  $d^2$  positive semi-definite operators  $E_i$  which span  $\mathcal{L}(\mathcal{H}_d)$ , the vector space of linear operators on  $\mathcal{H}_d$ , and satisfy  $\sum_i E_i = \mathbb{I}$ . When such a measurement consists of exactly  $d^2$  elements, density matrices  $\rho$  and the Born rule probabilities  $p(i) = \text{Tr}(\rho E_i)$  are in bijective correspondence because the  $E_i$  form a basis for  $\mathcal{L}(\mathcal{H}_d)$ . Such minimal IC-POVMs are known to exist in all dimensions [9]. Which one we choose for a representation, however, stands a chance of revealing or obscuring the properties which probability distributions equivalent to quantum states must have. Very often in mathematics and physics, a hard problem becomes easy when we choose the right basis. For example, the Eddington–Finkelstein coordinates

revealed that the event horizon of a non-rotating black hole is not a physical singularity. Is there a best or particularly nice IC-POVM which will reveal hidden properties of quantum theory?

It is not possible for an IC-POVM to be an orthonormal basis [10–12], so perhaps our first hope is ruled out. If the elements of an IC-POVM cannot be orthogonal, can they at least be equiangular? It turns out that they can [13]. That is, we can find a set of  $E_i$  which satisfy  $\text{Tr}(E_i E_j) = c$  for  $i \neq j$ . Can such a POVM consist of only rank-one matrices? Remarkably for such a simply-stated question, it is not generally known. A set of  $d^2$  rank-one matrices  $\Pi_i$  such that

$$\text{Tr}(\Pi_i \Pi_j) = \frac{d\delta_{ij} + 1}{d + 1}, \quad (2)$$

defines an IC-POVM  $E_i = \frac{1}{d}\Pi_i$  called a *Symmetric IC-POVM (SIC)* [14–16]. High-precision numerical SICs have been found in all dimensions 2–151 [17, 18] and in a few sporadic higher dimensions. In many cases, exact SICs have been constructed among these dimensions as well [19]. All indications are that SICs exist in all dimensions, but the proof continues to evade us.

If a SIC exists in dimension  $d$ , it is possible to rewrite the Born rule in a uniquely simple form analogous to the LTP, an equation called the *urgleichung* [4, 5], German for “primal equation” in [5]:

$$q(j) = \sum_i \left[ (d + 1)p(i) - \frac{1}{d} \right] r(j|i), \quad (3)$$

where  $q(j)$  is the probability for obtaining outcome  $j$  of a general quantum measurement,  $p(i)$  is the probability an agent ascribes to obtaining outcome  $i$  in the imagined scenario where a SIC measurement is performed on the system instead, and  $r(j|i)$  is the probability for obtaining the equivalent outcome  $j$  conditional on obtaining outcome  $i$  that the agent ascribes in the imagined scenario (see [3] for a detailed exposition). It is essential to recognize the operational difference between the *urgleichung* and the LTP: the *urgleichung* describes a scenario where the first measurement is not *actually* made—just imagined. If we actually planned to implement the first measurement, our probabilities must hold together according to the familiar LTP. We take Eq. (3) very seriously. In fact, it motivated the most recent development in QBism—a reconstruction of quantum theory featuring a generalization of the *urgleichung* as the key assumption [20]. See [21, 22] for critical review and discussion of the *urgleichung* in QBism as well as comparison to other contexts.

Probability theory itself has no tether to physical reality—rather it is a tool that anyone, *anywhere*, can use to manage their expectations for further experiences. Those expectations will certainly be influenced by deeply-held convictions that the agent has about the nature of reality around them, but the way those probabilities must hang together if the agent is to be coherent is unaffected. Quantum mechanics, on the other hand, although evidently a normative theory like probability theory, *is* tethered to physical reality. A reformulation of quantum theory which brings to the forefront this normative structure allows us to examine the threads of this tether without confusing

the subjective and the objective. This way, we may hopefully more readily determine the aspects of reality which forced quantum theory to be the way that it is. The LTP is a direct consequence of coherence in one’s probability assignments. Is the *Urgleichung*, which presents as an almost trivial modification of the LTP, on the right track for expressing the conditions for a kind of quantum coherence that an agent should strive for by virtue of being in our universe? We would like to accumulate as much evidence as possible that it is. Often one can gather more evidence for a sentiment simply by looking where it’s least expected. In this case, Huangjun Zhu recently demonstrated additional evidence for this line of reasoning by instead departing from probability theory [1].

If we formally relax the positivity condition for minimal IC-POVMs (keeping the fact that they sum to the identity), we are dealing with the larger space of Hermitian operator bases. Denoting such an operator  $F_i$ , if we also keep the form of the Born rule for a quantum state  $\rho$ , we obtain a set of real numbers  $p(i) = \text{Tr}(\rho F_i)$ , some of which may be negative, such that  $\sum_i p(i) = 1$ . This set of numbers is referred to as a *quasiprobability vector* (we will always denote quasiprobability vectors with  $\mathfrak{p}$  script) and a quasiprobability vector obtained in this way from a quantum state is called a *quasiprobability representation* of the state.

Quantum opticians have benefited from the ease of plotting quasiprobability distributions over phase space [23], so there is some utility in their use, but what is a quasiprobability? If a probability is a valuation of belief, what meaning can we attach to a quasiprobability? There does not appear to be a simple meaning—some attempts at attaching operational substance to quasiprobabilities have been made, for example, see [24–28], but if these solutions get us no closer to understanding *why* quantum theory is the normative probability calculus an agent of our universe should use to productively navigate, then they amount to duct tape over a structural weakness. However unsatisfying a quasiprobability may be on principled grounds, it turns out that permitting them for the time being gets us something desirable in return: the Born rule obtains an even closer functional analogy to the LTP.

Consider a Hermitian operator basis  $\{F_i\}$  such that  $\sum_i F_i = \mathbb{I}$  and a dual basis<sup>1</sup>  $\{Q_j\}$  constrained to satisfy  $\text{Tr}(Q_j) = 1$ . From a state  $\rho$  and an arbitrary POVM  $\{G_j\}$  we form the quasiprobabilities  $p(i) = \text{Tr}(\rho F_i)$  and the conditional quasiprobabilities  $\tau(j|i) = \text{Tr}(Q_j G_j)$ . Using the identity

$$\sum_j \text{Tr}(B F_j) \text{Tr}(Q_j C) = \text{Tr}(BC), \tag{4}$$

we can rewrite the Born rule as

$$q(j) = \text{Tr}(\rho G_j) = \sum_i p(i) \tau(j|i). \tag{5}$$

Like the *Urgleichung*, (5) looks very much like the LTP. In fact, (5) is functionally *identical* to the LTP. The difference is that this equation is written in terms of quasiprob-

<sup>1</sup> A dual basis is one for which the bases considered together are biorthogonal,  $\text{Tr} F_i Q_j = \delta_{ij}$ .

abilities instead of probabilities. Negativity must pop up somewhere, for it is known that negativity *must*<sup>2</sup> appear in quasiprobability representations of quantum theory [11, 29].

Another advantage when we are not burdened with positivity is that we may choose the  $F_i$  to form an orthogonal basis for operator space. If a basis is orthogonal, it is proportional to its dual basis and called *self-dual*. In this case, the sum constraint on the basis automatically fixes  $\text{Tr}(Q_j) = 1$  and the constant of proportionality  $F_j = \frac{1}{d} Q_j$ . Here we depart from Zhu's terminology and refer to a quasiprobability representation obtained from this sort of self-dual basis as a *Q-rep*. Q-reps account for most of the quasiprobability representations considered in the literature [30–33]. Importantly, however, Q-reps do not account for all finite dimensional quasiprobability representations; any nonorthogonal basis provides an example outside of this set. We can identify a Q-rep with the dual basis,  $\{Q_i\}$ , which defines it. We will reserve  $q$  for quasiprobability representations of states with respect to a Q-rep  $\{Q_j\}$  (when there is no index, we are referring to the full vector).

Mathematically speaking, a probability is a type of quasiprobability; that is, the kind without any negative entries. If we believe that the *urgleichung* differing from the LTP captures some aspect of the essential difference between quantum and classical, then some of this essential difference is also contained in the difference between the LTP and (5), that is, in the appearance of negativity in quasiprobability representations of quantum theory. Indeed, Spekkens showed that the presence of negative elements in quasiprobability representations of quantum theory and the impossibility of noncontextual hidden variable models are equivalent notions of nonclassicality [10]. Therefore, insofar as we think contextuality is an important ingredient in the quantum–classical distinction, we should be interested in negativity as well. Investigating the negativity in Q-reps also seems to be a promising approach to identifying exactly what advantages quantum computation affords us over classical computation, for example, Veitch et al. showed that negativity is a resource for quantum computation [34] and Howard et al. recently showed that contextuality enables universal quantum computation via ‘magic state’ distillation [35]. Additionally, and, as we will see, of particular note for this paper, Pashayan et al. have shown that a value related to the sum of the negative entries in a quasiprobability representation may be thought of as a measure that bounds the efficiency of a classical estimation of probabilities [36]. In light of these facts, Zhu's recent result is especially exciting. In his paper, Zhu developed a natural measure of negativity for Q-reps and established strict upper and lower bounds for this measure in one-to-one correspondence with SICs in each dimension. SICs are related to the bounds of this measure of negativity and the appearance of negativity in quasiprobability representations of quantum theory seems to contain some hints toward what “quantum” really means. Are there more hints to be found? Motivated to answer this question, we investigate how robust Zhu's result is to modifications in the negativity measure.

<sup>2</sup> That is, it is impossible to represent quantum theory in a way which eliminates the appearance of negativity in *both*  $p$  and  $\tau$  in (5) for all quantum states and POVMs. It is possible, in general, to eliminate the negativity appearing in one or the other.

In Sect. 2 we motivate and define a general negativity measure for quasiprobability representations which includes Zhu’s measure and the measure of principle interest in this paper, called sum negativity, as special cases and state Zhu’s theorem that the bounds on his measure of negativity are achieved by Q-reps in one-to-one correspondence with SICs. In Sect. 3 we address the sum negativity for the SIC Q-reps in the first few dimensions. In Sect. 4 we argue that Weyl–Heisenberg (WH) covariant Q-reps are a natural subset to consider while looking for counterexamples to Zhu’s theorem and establish the explicit conditions for a WH covariant Q-rep in dimension 3. Section 5 contains the main results: we explicitly demonstrate that Zhu’s theorem does not generally extend to sum negativity in either bound, we prove a general sufficiency theorem for a Q-rep being a local maximum for sum negativity, use this theorem to prove the exact upper bound for sum negativity among Q-reps in dimension 3, and state a conjecture regarding the lower bound among WH Q-reps. In Sect. 6 we again apply our theorem to prove that, although not generally the case, one of the SIC Q-reps achieves the exact upper bound for sum negativity in dimension 4. We also briefly discuss the sum negativity for one of the Hoggar SIC Q-reps in dimension 8. In Sect. 7 we discuss further questions and directions.

## 2 Negativity and Sum Negativity

Setting aside Q-rep vectors for a moment, we start by proposing a family of negativity measures for general quasiprobability vectors. Qualitatively speaking, we want a measure of the “amount” of negativity that appears in a vector with entries which sum to 1. A few candidates immediately stand out as especially natural measures: perhaps a measure proportional to the sum of the negative elements or to the most negative element appearing in the quasiprobability vector—indeed, as we will see, the latter choice is taken by Zhu [1]. We want a family of measures of negativity to meaningfully capture the deviation or “distance” from quasiprobabilities which have no negative elements. So we are faced with the task of measuring something we might understand as a distance for elements in a finite dimensional vector space (constrained by the normalization condition, of course). From this vantage point, the  $L^p$ -norms offer a very compelling family of generalized distance measures which we might utilize. Recall the definition of the  $L^p$ -norm of a vector  $\mathbf{x}$ ,

$$\|\mathbf{x}\|_p := \left( \sum_{i=1}^n |x(i)|^p \right)^{1/p}, \tag{6}$$

and the limiting expression for  $p \rightarrow \infty$ ,

$$\|\mathbf{x}\|_\infty = \max \{ |x_1|, |x_2|, \dots, |x_n| \}. \tag{7}$$

We might hope that the  $L^p$ -norm of a quasiprobability vector will itself be a useful quantity which is immediately related to the negativity present in the vector. If  $p = 1$ , for example, we can see that if there are any negative elements in the quasiprobability

vector  $\mathbf{p}$  then  $\|\mathbf{p}\|_1 > 1$ , so perhaps we could associate the amount by which the  $L^p$ -norm is larger than 1 with the negativity of  $\mathbf{p}$ . However,  $\|\mathbf{p}\|_p \leq \|\mathbf{p}\|_q$  when  $p > q$  and, in fact,  $\|\mathbf{p}\|_p$  need not be greater than 1 even when negativity is present. So what we really want is to measure the deviation of only the negative part of the quasiprobability vector from the zero vector, disregarding all positive entries entirely. With this in mind, define the *negative part* of a quasiprobability vector  $\mathbf{p}$  to be the vector

$$\mathbf{p}^{(-)}(j) := \frac{|\mathbf{p}(j)| - \mathbf{p}(j)}{2}, \quad (8)$$

which replaces the positive elements of  $\mathbf{p}$  with zero and the negative elements with their absolute value. This definition makes it easy to isolate properties of the negative elements of a quasiprobability vector. Now that we have done away with the positive entries in our quasiprobability vector, we define the  $N^p$  *negativity* of a quasiprobability vector  $\mathbf{p}$  to be the  $L^p$ -norm of the negative part of a quasiprobability vector:

$$N^p(\mathbf{p}) := \left\| \mathbf{p}^{(-)} \right\|_p. \quad (9)$$

We will refer to the special cases  $N^1$  and  $N^\infty$ , which we see are equivalent to the two natural candidates proposed above, as the *sum negativity*<sup>3</sup> and the *ceiling negativity* respectively.

Zhu defines the negativity of a quantum state  $\rho$  with respect to a Q-rep  $\{Q_j\}$  to be  $d$  times the magnitude of most negative element appearing in the quasiprobability representation  $q(j) = \text{Tr}(\rho F_j)$ . In our framework, this corresponds to  $d$  times  $N^\infty(q)$ . He then defines the negativity of a Q-rep itself to be the maximum of this value over all of quantum state space. In our framework, this corresponds to  $d$  times  $\max_\rho N^\infty(q)$ . Thus, in the general case, we define the  $N^p$  *negativity*<sup>4</sup> of a quantum state  $\rho$  with respect to a Q-rep  $\{Q_j\}$  to be

$$N^p(\rho, \{Q_j\}) := N^p(q), \quad (10)$$

and the  $N^p$  *negativity of a Q-rep*  $\{Q_j\}$  to be

$$N^p(\{Q_j\}) := \max_\rho N^p(\rho, \{Q_j\}). \quad (11)$$

Since  $L^p$ -norms are convex and nondecreasing for  $p \geq 1$  on the positive reals [38] and the negative part map (8) is a convex function, the composition (10) is also convex [39]. This means the maximum occurs on the boundary of the domain so we may take the maximum in (11) to be over pure states in these cases.

<sup>3</sup> Veitch et al. use the term *sum negativity* specifically for the sum of the negative elements of the discrete Wigner function for a quantum state [37] whereas we will be considering the equivalent notion with respect to any Q-rep.

<sup>4</sup> We have chosen to omit multiplication by  $d$  in our negativity definitions so that the negativity can more immediately be associated with the negative values in a quasiprobability vector. As the dual basis is calculationally easier to work with, a downside of our convention is that factors of  $1/d$  crop up more frequently.



As it will be useful later, note that the  $L^2$  norm of a Q-rep vector corresponding to a pure state always equals  $\sqrt{1/d}$ . To see this, let  $B = C = \rho$  and  $\text{Tr}(\rho^2) = 1$  in (4). This is another reason why the  $L^p$  norms are not a good choice for a family of negativity measures.

Zhu notes that the ceiling negativity of a Q-rep  $\{Q_j\}$  takes the simple form

$$N^\infty(\{Q_j\}) = \frac{1}{d} \left| \min_j \lambda_{\min}(Q_j) \right|, \tag{12}$$

where  $\lambda_{\min}(Q_j)$  is the minimal eigenvalue of  $Q_j$ .

If a SIC, denoted  $\{\Pi_j\}$ , exists in dimension  $d$ , we may construct two Q-reps  $\{Q_j^+\}$  and  $\{Q_j^-\}$  called the *SIC Q-reps*:

$$Q_j^\pm = \mp(\sqrt{d+1})\Pi_j + \frac{1}{d}(1 \pm \sqrt{d+1})\mathbb{I}, \tag{13}$$

which have ceiling negativities

$$N^+ \equiv N^\infty(\{Q_j^+\}) = \frac{(d-1)\sqrt{d+1}-1}{d^2}, \quad N^- \equiv N^\infty(\{Q_j^-\}) = \frac{\sqrt{d+1}-1}{d^2}. \tag{14}$$

We introduce these Q-reps because of the following theorem.

**Theorem 1** (Zhu) *Every Q-rep  $\{Q_j\}$  in dimension  $d$  satisfies  $N^- \leq N^\infty(\{Q_j\}) \leq N^+$ . The lower bound is saturated if and only if  $\{Q_j\}$  has the form  $\{Q_j^-\}$  where  $\Pi_j$  is a SIC. If  $\{Q_j\}$  is group covariant, then the upper bound is saturated if and only if  $\{Q_j\}$  has the form  $\{Q_j^+\}$ .*

Zhu’s theorem identifies SICs as centrally important to the study of Q-reps and more broadly for quantum theory because the Q-reps which achieve both bounds on ceiling negativity in any dimension are related to SICs by a simple affine transformation. Is the ceiling negativity unique in this way? If so, it would be interesting to understand why. If not, where does it fail?

We will address this question for the sum negativity. Like ceiling negativity, there is a more manageable expression for the sum negativity of a Q-rep. The following argument is due to Appleby and Zhu.

**Lemma 1** *An equivalent form of the sum negativity of a Q-rep  $\{Q_j\}$  is*

$$N^1(\{Q_j\}) = -\frac{1}{d} \min \left\{ \lambda_{\{1\}}, \lambda_{\{2\}}, \dots, \lambda_{\{2^{d^2-1}\}} \right\}, \tag{15}$$

where  $\lambda_{\{i\}}$  is the minimal eigenvalue of the  $i$ th partial sum matrix of the  $\{Q_j\}$  matrices.

*Proof* For a quasiprobability representation  $q$  of a state  $\rho$  with respect to Q-rep  $\{Q_j\}$ , if  $S := \{i | \text{Tr}(\rho Q_i) < 0\}$ , then

$$N^1(q) = \left| \frac{1}{d} \sum_{i \in S} \text{Tr}(\rho Q_i) \right| = \left| \frac{1}{d} \text{Tr} \left( \rho \sum_{i \in S} Q_i \right) \right|. \tag{16}$$

As in (12), the minimal value of the expression  $\text{Tr}(\rho F)$  over quantum state space is the minimal eigenvalue of  $F$  and the state  $\rho$  which minimizes the expression is the corresponding eigenvector. Thus, for a fixed subset  $\mathcal{G}$  of the  $\{Q_j\}$  matrices, the state which minimizes  $\text{Tr}(\rho \sum_{i \in \mathcal{G}} Q_i)$  is the minimal eigenvector of the matrix  $\sum_{i \in \mathcal{G}} Q_i$ . Via the definition above, for each  $\rho$  there is a subset  $\mathcal{S}$ ; in particular, there is a subset  $\mathcal{S}'$  for a state whose quasiprobability representation has the sum negativity value  $N^1(\{Q_j\})$  and, furthermore, the magnitude of the minimal eigenvalue of the matrix  $\sum_{i \in \mathcal{S}'} Q_i$  is equal to  $d$  times  $N^1(\{Q_j\})$ . Thus determining the sum negativity is equivalent to looking for the minimal eigenvalue over all partial sum matrices of  $\{Q_j\}$ . There are  $2^{d^2} - 1$  entries to minimize over because we ignore the partial sum corresponding to the empty subset.  $\square$

Does Theorem 1 extend to sum negativity? In Sect. 5, we will demonstrate that it generally does not with explicit counterexamples in the first nontrivial dimension,  $d = 3$ .

### 3 Sum Negativities of $\{Q_j^+\}$ and $\{Q_j^-\}$

Sum negativity is notably harder to work with than ceiling negativity, both analytically and numerically. As such, using the method described in Lemma 1, analytic results for the sum negativity of  $\{Q_j^+\}$  and  $\{Q_j^-\}$  have only so far been obtained for dimensions 2–4. Numerically exact results have also been obtained for  $d = 5$ .

For  $d = 2$ , sum negativity and ceiling negativity are equivalent measures because a  $d = 2$  Q-rep vector cannot contain more than one negative element. This property can be proven easily with Eq. (4) and the fact quasiprobabilities are normalized. In addition, the ceiling negativities for  $\{Q^+\}$  and  $\{Q^-\}$  are equivalent. Thus,

$$N^1(\{Q_j^\pm\}) = N^\infty(\{Q_j^\pm\}) = \frac{\sqrt{3} - 1}{4}. \quad (17)$$

In fact, all  $N^p$  negativities for the SIC Q-reps are equal in dimension 2. This is a reflection of the fact that all Q-reps are equivalent to the Wootters discrete Wigner function in this dimension [1].

There is a continuous one-parameter family of SICs in dimension 3 [40]. For any of them, we may construct the SIC Q-reps. In Zhu's paper we easily see that ceiling negativity is insensitive to the SIC chosen—all that matters is that the defining property of a SIC is satisfied. This turns out to also be true for sum negativity.<sup>5</sup> What's more,

<sup>5</sup> We won't explore it further here, but the situation is more interesting. Although the sum negativity is insensitive to the value of the parameter  $t$  which defines the inequivalent SIC Q-reps, the quantum states whose quasiprobability representations achieve these sum negativity values *do* depend on the parameter. It turns out that the sum negativity for  $\{Q_j^-\}$  constructed from the Hesse SIC ( $t = 0$  in [40]) is achieved by a complete set of mutually unbiased bases [41], that is, 12 [=  $d(d + 1)$ ] vectors which form four orthogonal bases such that any vector from one basis has an equal overlap with any vector from another basis. For all the other SICs in dimension 3, the states which achieve the sum negativity of  $\{Q^-\}$  form a single basis instead. The complete set of mutually unbiased bases also turns out to be the set of states which minimize the Shannon entropy in the Hesse SIC representation [42, 43].

the sum negativities of  $\{Q_j^+\}$  and  $\{Q_j^-\}$  are equal:

$$N^1(\{Q_j^\pm\}) = \frac{1}{3}. \tag{18}$$

Recall that Theorem 1 establishes that the upper and lower bounds for ceiling negativity over all Q-reps are achieved by the SIC Q-reps. We see now that in dimension 3, the sum negativities of  $\{Q_j^+\}$  and  $\{Q_j^-\}$  are equal. This tells us that if any  $d = 3$  Q-rep has a sum negativity other than  $1/3$ , the analog of Theorem 1 does not hold for sum negativity in dimension 3.

For  $d = 4$ ,

$$N^1(\{Q_j^+\}) = \frac{1}{2}, \tag{19}$$

and

$$N^1(\{Q_j^-\}) = -\frac{1}{16} \left( 5 + \sqrt{5} - 2\sqrt{2(1+\sqrt{5})} - 2\sqrt{23 - 2\sqrt{5} + 2\sqrt{-22 + 10\sqrt{5}}} \right) \approx 0.420967. \tag{20}$$

Value (19) is surprisingly nice. We will comment briefly on this in Sect. 6. Value (20) is shocking. However, in light of recent results relating the SIC problem to algebraic number theory ([44], see the contributions of Appleby et al [45] and Bengtsson [46] to this volume for a review), it is worth mentioning a few possibly relevant facts about this number and how it arose.

The sum negativity for the  $d = 4$  SIC Q-rep is (1/4 times) the minimal eigenvalue of certain 7-element partial sums of the SIC Q-rep matrices. The characteristic polynomial which has this eigenvalue as a root is:

$$x^4 - 7x^3 + \frac{21}{2}x^2 + \left( \frac{129}{8} - \frac{35\sqrt{5}}{8} + \sqrt{2(31 + 17\sqrt{5})} \right)x - \frac{1293}{32} + \frac{293\sqrt{5}}{32} - \sqrt{5(22 + 29\sqrt{5})} \tag{21}$$

The coefficients in Eq. (21) are members of the field  $\mathbb{Q}(\sqrt{5} + \sqrt{362 + 313\sqrt{5}})$ . The factor multiplying  $-1/16$  in (20) is an algebraic integer, but not an algebraic unit. Finally, the minimal polynomial for (20) is degree 8.

For  $d = 5$ ,  $N^1(\{Q^+\}) \approx 0.584277$  and  $N^1(\{Q^-\}) \approx 0.501957$ . These answers are numerically correct, but do not lend themselves readily to conversion to exact values.

### 4 Weyl–Heisenberg Q-reps in $d = 3$

We say that a set of vectors is *group covariant* if it is the orbit of some group action on an initial vector, which we call the *fiducial*. All known SICs are group covariant

and all but one of those are covariant with respect to the *WH group*. In dimension  $d$ , let  $\omega_d = e^{2\pi i/d}$  be a  $d$ th root of unity, and define the shift and phase operators

$$X|j\rangle = |j+1\rangle, \quad Z|j\rangle = \omega_d^j|j\rangle, \quad (22)$$

where the shift is modulo  $d$ . Products of powers of  $X$  and  $Z$  and powers of  $\omega_d$  define the WH group. The order of the WH group in dimension  $d$  is  $d^3$ , but for the purposes of constructing a measurement operator or Q-rep, we can neglect the phase factors. In other words, for some fiducial matrix  $Q_0$ , the other matrices in the orbit would take the form

$$Q_{ij} = X^i Z^j Q_0 (X^i Z^j)^\dagger. \quad (23)$$

The only known exception to WH covariance for SICs is the Hoggar SIC in dimension 8, but even this outlier is group covariant with respect to the tensor product of three  $d = 2$  WH groups. What about Q-reps? By their construction, the non-Hoggar SIC Q-reps are WH covariant and these Q-reps achieve the bounds for ceiling negativity in all dimensions (provided a SIC exists in that dimension). Thus, if we were to consider any subset of the full space of Q-reps for computational study, the set of WH covariant Q-reps (WH Q-reps) is likely the best starting point. In any case, due to the ubiquity of the WH group in quantum information theory, the bounds of sum negativity within WH Q-reps may be of independent interest.

A Q-rep is associated with an orthogonal basis of operators  $\{Q_j\}$  with norm 3, that is,

$$\text{Tr}(Q_i Q_j) = 3\delta_{ij}. \quad (24)$$

Therefore, the general conditions for a WH Q-rep may be obtained by requiring that the elements of the WH orbit of a general unit-trace Hermitian matrix

$$\begin{bmatrix} z & y & x \\ y^* & w & v \\ x^* & v^* & 1 - z - w \end{bmatrix}, \quad (25)$$

satisfy Eq. (24). Imposing this condition results in a number of equations which can be algebraically simplified to the following three:

$$\begin{aligned} z^2 + zw + w^2 &= z + w, \\ |y|^2 + |x|^2 + |v|^2 &= 1, \\ xy + y^*v + v^*x^* &= 0. \end{aligned} \quad (26)$$

In terms of real variables, an arbitrary unit-trace Hermitian matrix

$$\begin{bmatrix} a & b+ic & d+ie \\ b-ic & f & g+ih \\ d-ie & g-ih & 1-a-f \end{bmatrix}, \quad (27)$$

is a WH Q-rep fiducial if

$$\begin{aligned}
 a^2 + af + f^2 &= a + f, \\
 b^2 + c^2 + d^2 + e^2 + g^2 + h^2 &= 1, \\
 dg + bd + bg + ch &= ec + eh, \\
 cd + be + bh &= cg + dh + eg.
 \end{aligned}
 \tag{28}$$

From (26) or (28) we can see that  $d = 3$  WH Q-rep fiducials define a four dimensional subspace of the eight dimensional space of  $3 \times 3$  unit-trace Hermitian matrices. The main diagonal elements are independent of the off-diagonal elements and satisfy the equation of an ellipse [the first equation in (26) or (28)]. The magnitude of the off-diagonal elements in (25) lie on the unit 2-sphere, but their exact values only lie at points where the expression  $xy + y^*v + v^*x^*$  vanishes.

The following two matrices are explicit examples of  $d = 3$  Non-SIC WH Q-rep fiducials:

$$Q^{\min} = \begin{bmatrix} 0 & -\frac{1}{3} + \frac{i}{3} & \frac{2}{3} - \frac{i}{3} \\ -\frac{1}{3} - \frac{i}{3} & 1 & -\frac{1}{3} + \frac{i}{3} \\ \frac{2}{3} + \frac{i}{3} & -\frac{1}{3} - \frac{i}{3} & 0 \end{bmatrix},
 \tag{29}$$

$$Q^{\max} = \begin{bmatrix} 0 & \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) + \frac{i}{4} & \left(\frac{2}{3} + \frac{\sqrt{7}}{12}\right) - \frac{i}{4} \\ \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) - \frac{i}{4} & 1 & \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) + \frac{i}{4} \\ \left(\frac{2}{3} + \frac{\sqrt{7}}{12}\right) + \frac{i}{4} & \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) - \frac{i}{4} & 0 \end{bmatrix}.
 \tag{30}$$

In the following section we will see why they are presented with the designations “max” and “min”.

### 5 Sum Negativity Bounds in $d = 3$

The WH Q-reps generated by (29) and (30) will be denoted  $\{Q_j^{\min}\}$  and  $\{Q_j^{\max}\}$ , respectively. Their sum negativities are

$$N^1(\{Q_j^{\min}\}) = \frac{1}{3} \left(2 \cos \frac{\pi}{9} - 1\right) \approx 0.293128,
 \tag{31}$$

and

$$N^1(\{Q_j^{\max}\}) = \frac{2}{9} (\sqrt{7} - 1) \approx 0.365723.
 \tag{32}$$

Recall from Sect. 3 that the sum negativities of both SIC Q-reps is  $1/3$ , so  $\{Q_j^{\min}\}$  and  $\{Q_j^{\max}\}$  are explicit counterexamples to Zhu’s theorem for sum negativity. It turns out that (32) is a strict upper bound on the sum negativity, not only of WH Q-reps, but of all Q-reps in dimension 3. We start with two lemmas.

**Lemma 2** *For quasiprobability vectors with  $d^2$  elements lying in the sphere of radius  $\sqrt{1/d}$ , the stationary points for sum negativity are:*

- (1) *Those vectors whose entries consist only of two distinct values.*  
 (2) *Those vectors whose entries consist only of three distinct values including zero.*

*Proof* Quasiprobability vectors in the sphere of radius  $\sqrt{1/d}$  satisfy the following constraints:

$$\sum_{j=1}^{d^2} p(j) = 1 \quad \text{and} \quad \sum_{j=1}^{d^2} p(j)^2 = \frac{1}{d}. \quad (33)$$

The definition of the sum negativity of a quasiprobability vector (34) gives us

$$N^1(\mathbf{p}) = \left\| \mathbf{p}^{(-)} \right\|_1 = \sum_{j=1}^{d^2} \frac{|p(j)| - p(j)}{2} = \frac{1}{2} \sum_{j=1}^{d^2} |p(j)| - \frac{1}{2}. \quad (34)$$

From this it is clear that the stationary points of the sum negativity are exactly the stationary points of the sum of the absolute values. Absolute values are often difficult to deal with in optimization problems, but it turns out that the function  $|x|$  may be approximated efficiently by  $\sqrt{x^2 + c}$  where  $c$  is taken to zero after any differentiation [47]. Thus, we want to extremize the function

$$\sum_{j=1}^{d^2} \sqrt{p(j)^2 + c}, \quad (35)$$

subject to constraints (33) in the small  $c$  limit. To do this we construct a Lagrangian

$$\mathcal{L}(\mathbf{p}, \lambda, \mu) = \sum_{j=1}^{d^2} \left( \sqrt{p(j)^2 + c} \right) - \lambda \left( \sum_{j=1}^{d^2} p(j) - 1 \right) - \mu \left( \sum_{j=1}^{d^2} p(j)^2 - \frac{1}{d} \right), \quad (36)$$

where  $\lambda$  and  $\mu$  are Lagrange multipliers. Varying this Lagrangian, we see that the stationary points must satisfy

$$\frac{p(j)}{\sqrt{p(j)^2 + c}} - \lambda - 2\mu p(j) = 0, \quad (37)$$

for all  $j$ . This expression has four solutions, two of which become zero in the  $c = 0$  limit:

$$p(j) = \frac{\pm 1 - \lambda}{2\mu} \quad \text{or} \quad p(j) = 0. \quad (38)$$

Consider the case where  $p(j) \neq 0$  for all  $j$ . Now, in order for the constraints (33) to hold, some number  $n$  of the entries are  $\frac{-1-\lambda}{2\mu}$  and the other  $d^2 - n$  are  $\frac{1-\lambda}{2\mu}$ . That is,

$$\begin{aligned} n \left( \frac{-1-\lambda}{2\mu} \right) + (d^2 - n) \left( \frac{1-\lambda}{2\mu} \right) &= 1 \quad \text{and} \\ n \left( \frac{-1-\lambda}{2\mu} \right)^2 + (d^2 - n) \left( \frac{1-\lambda}{2\mu} \right)^2 &= \frac{1}{d}. \end{aligned} \quad (39)$$

It is easy to verify that  $0 < n < d^2$ . In this case the quasiprobability vector  $\mathbf{p}$  consists of two distinct values. If  $p(j) = 0$  for  $m$  of the indices then  $d^2 - n - m$  of the entries are  $\frac{1-\lambda}{2\mu}$  and the appropriately modified form of (39) holds in which case the quasiprobability vector  $\mathbf{p}$  consists of three distinct values including zero.  $\square$

A stationary point can be a local maximum, a local minimum, or a saddle point. Sufficient conditions for maxima and minima in Lagrangian systems with equality constraints are known. We will need the following tool (which can be found in Chap. 2 of [48]):

**Theorem 2** (Sufficient Conditions for Constrained Maxima) *Consider a constrained maximization problem for a twice-differentiable function of  $n$  variables  $y(\mathbf{x})$  with  $m$  twice-differentiable equality constraints  $f_i(\mathbf{x}) = 0, i = 1, \dots, m$ . The Lagrangian is*

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = y(\mathbf{x}) - \sum_{i=1}^m \lambda_i f_i(\mathbf{x}), \tag{40}$$

where  $\lambda_i$  are Lagrange multipliers. If there exist vectors  $\mathbf{x}^*$  and  $\boldsymbol{\lambda}^*$  such that  $\partial_i \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0, i = 1, \dots, n$  and  $f_i(\mathbf{x}^*) = 0, i = 1, \dots, m$  and if

$$(-1)^s \begin{vmatrix} \partial_{11}\mathcal{L} & \cdots & \partial_{1s}\mathcal{L} & \partial_1 f_1 & \cdots & \partial_1 f_m \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_{1s}\mathcal{L} & \cdots & \partial_{ss}\mathcal{L} & \partial_s f_1 & \cdots & \partial_s f_m \\ \partial_1 f_1 & \cdots & \partial_s f_1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_s f_m & 0 & \cdots & 0 \end{vmatrix} > 0, \tag{41}$$

for  $s = m+1, \dots, n$  ( $\partial_{ij}\mathcal{L}$  indicates a second partial derivative of  $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$  with respect to  $x_i$  and  $x_j$  evaluated at  $\mathbf{x}^*$  and  $\boldsymbol{\lambda}^*$  and  $\partial_i f_j$  indicates the first partial derivative of constraint function  $f_j(\mathbf{x})$  with respect to  $x_i$  evaluated at  $\mathbf{x}^*$ ), then  $y(\mathbf{x})$  has a strict local maximum at  $\mathbf{x}^*$ .

A consequence of this is the following:

**Lemma 3** *The stationary points for sum negativity of quasiprobability vectors with  $d^2$  elements lying in the sphere of radius  $\sqrt{1/d}$  are all local maxima or global minima.*

*Proof* From Lemma 2 we know that the stationary points for sum negativity of quasiprobabilities lying on the sphere of radius  $\sqrt{1/d}$  are those consisting of exactly two distinct elements or those consisting of three if one of them is zero. If there are two distinct elements, they can either both be nonnegative or of opposite signs. If they are both nonnegative, the sum negativity is zero which is the global minimum sum negativity value. If there are three distinct elements and they are all nonnegative, then it is also a global minimum for sum negativity. We will now show that when the nonzero elements of a stationary point are of opposite signs, it is a local maximum.

Consider first the stationary points with two distinct elements. In terms of the Lagrange multipliers from the proof of Lemma 2, the stationary quasiprobability vectors are comprised of opposite signed values when  $\mu \neq 0$  and  $-1 < \lambda < 1$ . Without loss of generality we choose  $\frac{-1-\lambda}{2\mu}$  to be the negative value so that we may enumerate the number of negative entries with index  $n$  as in (39). This amounts to the further restriction  $\mu > 0$ . Note that by substituting  $\frac{-1-\lambda}{2\mu} = a$  and  $\frac{1-\lambda}{2\mu} = b$  in (39), we can solve for the positive and negative entries in the quasiprobability vector in terms of the dimension  $d$  and number of negative entries  $1 \leq n < d(d - 1)$ :

$$a = \frac{1}{d^2} + \frac{1}{d^2} \sqrt{\frac{n(d-1)}{d^2-n}} - \sqrt{\frac{d-1}{n(d^2-n)}} < 0, \tag{42}$$

$$b = \frac{1}{d^2} + \frac{1}{d^2} \sqrt{\frac{n(d-1)}{d^2-n}} > 0. \tag{43}$$

Sum negativity is invariant to the ordering of the entries in a quasiprobability vector so we need only prove that one ordering is a local maximum for each  $n$ . Without loss of generality, we fix  $p(1) = a$  and demand that  $p(2)$  and  $p(3)$  are not both also equal to  $a$ . Consider again the Lagrangian (40). Taking the relevant derivatives and  $c$  to zero, we may construct the bordered Hessian matrix from (41) for our problem. Therefore, if  $\mathbf{p}$  is a stationary point and if

$$(-1)^s \begin{vmatrix} -2\mu & \cdots & 0 & 1 & 2p(1) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & -2\mu & 1 & 2p(s) \\ 1 & \cdots & 1 & 0 & 0 \\ 2p(1) & \cdots & 2p(s) & 0 & 0 \end{vmatrix} > 0, \tag{44}$$

for  $s = 3, \dots, d^2$ , then  $\mathbf{p}$  is a local maximum in sum negativity.

Recall the following identity for matrix blocks  $A, B, C$ , and  $D$  which holds when  $A$  is invertible:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \left| D - CA^{-1}B \right|. \tag{45}$$

Using (45), the determinant of the bounded Hessian matrix in (44) is

$$(-1)^s (2\mu)^{s-1} \left( 4s \sum_{i=1}^s p(i)^2 - 4 \sum_{i,j=1}^s p(i)p(j) \right). \tag{46}$$

With the additional  $(-1)^s$  term in Eq. (44), we see that the inequality is satisfied if

$$s \sum_{i=1}^s p(i)^2 - \sum_{i,j=1}^s p(i)p(j) > 0, \tag{47}$$



for all  $s$ . Let  $l$  denote the number of negative elements in the truncated quasiprobability vectors appearing in (44). Note that  $1 \leq l < s$  by our earlier assumption about the first three elements. Then

$$\begin{aligned}
 & s \sum_{i=1}^s p(i)^2 - \sum_{i,j=1}^s p(i)p(j) \\
 &= s \left( la^2 + (s-l)b^2 \right) - (la + (s-l)b)^2 = (a-b)^2(s-l)l > 0. \quad (48)
 \end{aligned}$$

Thus the stationary points consisting of two distinct values, one negative and one positive, are local maxima.

The values of  $a$  and  $b$  in the case with three distinct values are more complicated, but may still be obtained. Otherwise the proof in this case carries through in the same way as above with a slight modification in the last step. Let  $m$  denote the number of elements equal to zero in the truncated quasiprobability vectors appearing in (44). Note that  $l + m \leq s$ . Then it can be verified that

$$s \sum_{i=1}^s p(i)^2 - \sum_{i,j=1}^s p(i)p(j) = s \left( la^2 + (s-l-m)b^2 \right) - (la + (s-l-m)b)^2 > 0, \quad (49)$$

which completes the proof. □

We may now return to the question of extremality among Q-reps. We say that a pure state  $\rho$  achieves the sum negativity if it is an eigenvector with eigenvalue magnitude equal to  $d$  times  $N^1(\{Q_j\})$  of one of the partial sum matrices of the Q-rep  $\{Q_j\}$ .

**Theorem 3** *If the quasiprobability representation of a state which achieves the sum negativity of a Q-rep  $\{Q_j\}$  consists of two distinct elements or three including zero, then  $N^1(\{Q_j\})$  is a local maximum among all Q-reps.*

*Proof* Recall that the quasiprobability representation of a pure state with respect to a Q-rep lies in the sphere of radius  $\sqrt{1/d}$ . Also recall that the convexity of the sum negativity function implies that any state which achieves the sum negativity for a Q-rep  $\{Q_j\}$  is a pure state. Therefore, if, as we vary  $\{Q_j\}$  over the space of Q-reps, the quasiprobability representation of a pure state which achieves the sum negativity of  $\{Q_j\}$  consists of two distinct values or three distinct values including zero for some Q-rep  $\{Q'_j\}$ , then by Lemmas 2 and 3, there is a local maximum or global minimum of the sum negativity function at  $\{Q'_j\}$ . As we know the appearance of negativity is inevitable, the global minimum possibility is avoided. □

**Theorem 4** *The exact upper bound for sum negativity among Q-reps in dimension 3 is  $\frac{2}{9}(\sqrt{7} - 1)$ .*

*Proof* When  $d = 3$ , explicit calculation reveals that the sum negativity of a quasiprobability vector lying in the sphere of radius  $\sqrt{1/d}$  constructed with  $n$  values equal to  $a < 0$ ,  $m$  values equal to 0, and  $9 - n - m$  values equal to  $b > 0$  is maximized when  $n = 2$  and  $m = 0$ . Those values are, from (42) and (43),

$$a = \frac{1}{9}(1 - \sqrt{7}) \quad \text{and} \quad b = \frac{1}{63}(7 + 2\sqrt{7}). \quad (50)$$

Lemmas 2 and 3 and the fact that our domain has no boundary [lying on the sphere of radius  $\sqrt{1/d}$  and quasiprobability normalization together define a  $(d^2 - 2)$ -sphere] imply that the sum negativity of this quasiprobability vector is the global maximum value over this domain. The quasiprobability representations of the states which achieve (32) with respect to the Q-rep  $\{Q_j^{\max}\}$  consist of these values, and so, by Theorem 3,  $\frac{2}{9}(\sqrt{7} - 1)$  is the maximum value for sum negativity over all Q-reps in dimension 3.  $\square$

It is important to note that while Theorem 4 shows that (32) is the strict upper bound for sum negativity among all Q-reps, it does not imply that  $\{Q_j^{\max}\}$  is the unique Q-rep which achieves this bound. Numerical searching suggests that, for the fiducial main diagonal  $\{0, 1, 0\}$ , the WH orbit of (30) is the unique Q-rep which achieves the sum negativity (32), but that there is at least one other fiducial main diagonal which achieves this bound, namely the main diagonal corresponding to the major axis vertices of the ellipse defined by the first equation in (26):  $\left\{\frac{1}{3}, \frac{1}{3} + \frac{1}{\sqrt{3}}, \frac{1}{3} - \frac{1}{\sqrt{3}}\right\}$ . Unfortunately, for this fiducial, we were unable to convert the numerical result to exact values.

What can be said about the lower bound for sum negativity in dimension 3? So far, less is known, but we present the following numerically motivated conjecture:

**Conjecture 1** *The exact lower bound for sum negativity among WH Q-reps in dimension 3 is  $\frac{2}{3}(\cos \frac{\pi}{9} - \frac{1}{2})$ .*

This statement resisted our attempts to prove it in a fashion similar to Theorem 4 because for every Q-rep there exist states with zero negative elements in their quasiprobability representation (for example, the maximally mixed state). The Q-rep vector corresponding to the minimal eigenstate over all partial sums of  $\{Q_j^{\min}\}$  consists of three distinct values which each appear three times:

$$\frac{1}{9}\left(1 - 2\cos \frac{\pi}{9}\right), \quad \frac{1}{9}\left(1 + 2\cos \frac{2\pi}{9}\right), \quad \text{and} \quad \frac{1}{9}\left(1 + 2\cos \frac{\pi}{9} - 2\cos \frac{2\pi}{9}\right). \quad (51)$$

The values in (51) may not have significance as deep as those in the upper bound quasiprobability vector, however, as numerical searching has revealed that (31) is achieved by a WH Q-rep for *every* valid main diagonal. The off-diagonal elements in other cases were too difficult to convert to exact values. The fact that there seems to be a WH Q-rep which achieves (31) for any main diagonal satisfying the first equation in (26) suggests the lower bound among Q-reps may not be saturated by WH Q-reps. Additionally, the extremal properties of general quasiprobability vectors lying on the sphere of radius  $\sqrt{1/d}$  do not come to our aid when we try to find the lower bound because there is no obvious reason to hope that the process of maximizing over quantum state space (in the definition of a Q-rep sum negativity) followed by minimizing over all Q-reps (to find the lower bound for sum negativity) should result in one of the local maxima for general quasiprobability vectors on the sphere of radius  $\sqrt{1/d}$  (recall, of course, that it cannot result in the global minimum of zero negativity

on this sphere because we know that the appearance of negativity is inevitable). Before, we were maximizing over both; in some sense we got lucky that the global maximum sum negativity for a quasiprobability vector on this sphere was achieved by a Q-rep vector.

## 6 Further Observations About SIC Q-reps

In Sect. 3, we noted the appearance of a rational value for the sum negativity of  $\{Q_j^+\}$  in dimension 4. The quasiprobability vectors which achieve the sum negativity of  $\{Q_j^+\}$  are of a special and familiar form; they consist of only the values  $-1/8$  and  $1/8$ . We know from Theorem 3, therefore, that this Q-rep is a local maximum for sum negativity among Q-reps in dimension 4. In fact, following the exact same procedure as in Theorem 4, we see that  $N^1(\{Q_j^+\}) = 1/2$  is the exact upper bound over all Q-reps!

In footnote 5, we mentioned that the states which achieve the sum negativity of  $\{Q_j^-\}$  for the Hesse SIC in dimension 3 form a complete set of mutually unbiased bases. The states which achieve the sum negativity for  $\{Q_j^-\}$  in dimension 4 also form a structure of possible interest. They consist of a set of 16 vectors which have two nontrivial squared overlaps. In the terminology of reference [14], these states form a quantum design of degree 2.

Although dimension 5 was the last in which we were able to explicitly calculate the sum negativity for the SIC Q-reps by exhaustive combinatorial searching, we suspect that we have found the correct sum negativity for  $\{Q_j^-\}$  constructed with the Hoggar SIC in dimension 8. Rather than calculating the eigenvalues of every partial sum matrix (since this is infeasible for  $2^{64}$ ,  $8 \times 8$  matrices), we used a numerical local maximization procedure and around  $10^6$  random pure state seeds. The overall maximum value we found,  $7/8$ , occurred frequently in our data and is significantly larger than all of the smaller local maxima. Of course, we could still be falling short of the global maximum value if it occurs at very hard to access positions. The states whose quasiprobability representations achieve the sum negativity of  $7/8$  consist of 28 copies of value  $-1/32$  and 36 copies of value  $5/96$ . Therefore, by Theorem 3, if these states achieve the actual sum negativity, then  $\{Q_j^-\}$  constructed with the Hoggar SIC is a local maxima among all Q-reps in dimension 8. Surprisingly, these quantum states also minimize the Shannon entropy of their Hoggar SIC representations and thus compose the “twin” Hoggar SIC [43, 49]. This result and the one for  $\{Q_j^-\}$  in dimension 4 in the previous paragraph parallel the one mentioned in footnote 5.

## 7 Discussion

In QBism, quantum states are probability distributions over a set of possible outcomes for an appropriately chosen measurement. With the understanding that quantum theory is an addition to coherence which rational agents should use to help inform their expectations for future experiences in terms of their past ones, QBists hope that the structure of quantum theory, and in particular its boundaries, can be made to suggest

nature's motives. A century of quantum foundational debate should by now have convinced us that these motives will not conform to our prejudices about reality. Although the weirdness of quantum theory has convinced some that we can no longer pretend physics is more than an exercise in instrumentalism, QBists are optimistic that there are ubiquitous and recognizably physical statements about nature yet to be made.

Prior to Zhu's paper, we had focused on characterizing the bounds of quantum theory only from within probability theory. His results reveal this approach was nearsighted. In the initial stages of this project, we had hoped to find further evidence for the centrality of SICs in quantum theory by showing that Zhu's theorem extends to another natural measure of negativity for quasiprobability representations. Indeed, it is interesting to find that the SICs do not generally play the same role in this alternate context. Why do they not? Do they still always play *some* role which is not immediately apparent? Maybe there are other families of Q-reps constructed from SICs or another structure which naturally play the same role for sum negativity.  $\{Q_j^+\}$  in dimension 4 did achieve the upper bound for sum negativity. Does this happen again? Although we cannot calculate it exactly, we have some numerical evidence that the sum negativities for the Hoggar SIC Q-reps differ from the non-Hoggar SIC Q-reps in dimension 8. This suggests that there might be an essential relation between sum negativity and group covariance. Furthermore, the appearance of the complete set of mutually unbiased bases in dimension 3 (mentioned in footnote 5) and the "twin" Hoggar SIC in dimension 8 suggest a deep connection between sum negativity and minimizing Shannon entropy which warrants further exploration.

Due to the connection to contextuality, the bounds on sum negativity and  $N^p$  negativity in general are likely to be of interest to the quantum computation community. A natural further direction for this research is the consideration of negativities other than ceiling and sum negativity. Perhaps the next to consider is the only non-convex integral negativity,  $N^0$ , which tells us the maximum number of negative elements which can appear in a Q-rep vector. On the other hand,  $N^2$  negativity makes use of the most familiar distance function, and, as such, may warrant special attention. In Sect. 4 we established the general conditions for a  $d = 3$  WH Q-rep. This result and any analogous results in higher dimensions<sup>6</sup> may be of independent interest. Likewise, in order to pursue the exact lower bound for sum negativity in dimension 3, we need strategies to construct non-WH Q-reps. Towards this, the general structure and symmetries inherent in Q-reps warrants exploration. It may further be interesting to consider what can be said about quasiprobability representations obtained from non-orthogonal bases or even redundant operator frames in the negativity context.

**Acknowledgements** The authors would like to thank Marcus Appleby, Blake Stacey, and Huangjun Zhu for helpful discussions and suggestions. This research was supported by the Foundational Questions Institute Fund on the Physics of the Observer (Grant FQXi-RFP-1612), a Donor Advised Fund at the Silicon Valley Community Foundation.

<sup>6</sup> And in dimension 8, it may be interesting to look at general  $WH \otimes WH \otimes WH$  covariant Q-reps.

## References

- Zhu, H.: Quasiprobability representations of quantum mechanics with minimal negativity. *Phys. Rev. Lett.* **117** (12), 120404 (2016). [arXiv:1604.06974](https://arxiv.org/abs/1604.06974) [quant-ph]
- Weinberg, S.: The trouble with quantum mechanics. *N. Y. Rev. Books* **64**(1). <http://www.nybooks.com/articles/2017/01/19/trouble-with-quantum-mechanics/> (2017)
- Fuchs, C.A., Stacey, B.C.: QBist Quantum Mechanics: Quantum Theory as a Hero's Handbook. Preprint (2016). [arXiv:1612.07308](https://arxiv.org/abs/1612.07308) [quant-ph]
- Fuchs, C.A.: QBism, the Perimeter of Quantum Bayesianism. Preprint (2010). [arXiv:1003.5209](https://arxiv.org/abs/1003.5209) [quant-ph]
- Fuchs, C.A., Schack, R.: Quantum-Bayesian Coherence. *Rev. Mod. Phys.* **85**, 1693 (2013). [arXiv:0906.2187](https://arxiv.org/abs/0906.2187) [quant-ph]
- de Finetti, B.: *Theory of Probability*. Wiley, Chichester (1990)
- Lad, F.: *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*. Wiley-Interscience, New York (1996)
- Bernardo, J.M., Smith, A.F.M.: *Bayesian Theory*. Wiley, Chichester (1994)
- Caves, C.M., Fuchs, C.A., Schack, R.: Unknown quantum states: the quantum de Finetti representation. *J. Math. Phys.* **43**, 4537 (2002)
- Spekkens, R.W.: Negativity and contextuality are equivalent notions of nonclassicality. *Phys. Rev. Lett.* **101**, 020401 (2008)
- Ferrie, C., Emerson, J.: Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations. *J. Phys. A* **41**, 352001 (2008)
- Appleby, D.M., Dang, H.B., Fuchs, C.A.: Symmetric Informationally-Complete Quantum States as Analogues to Orthonormal Bases and Minimum Uncertainty States. *Entropy* **16**(3), 1484–92 (2014). [arXiv:0707.2071](https://arxiv.org/abs/0707.2071) [quant-ph]
- Appleby, D.M.: Symmetric informationally complete measurements of arbitrary rank. *Opt. Spectrosc.* **103**, 416–428 (2007)
- Zauner, G.: *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*. PhD Thesis, University of Vienna, 1999. <http://www.gerhardzauner.at/qdmye.html>. Published in English translation: Zauner, G.: Quantum designs: foundations of a noncommutative design theory. *Int. J. Quantum Inf.* **9**, 445–508 (2011)
- Caves, C.M.: Symmetric informationally complete POVMs. <http://info.phys.unm.edu/~caves/reports/infopovm.pdf> (2002)
- Renes, J.M., Blume-Kohout, R., Scott, A.J., Caves, C.M.: Symmetric informationally complete quantum measurements. *J. Math. Phys.* **45**, 2171 (2004)
- Fuchs, C.A., Hoang, M.C., Stacey, B.C.: The SIC question: history and state of play. Preprint (2017). [arXiv:1703.07901](https://arxiv.org/abs/1703.07901) [quant-ph]
- Scott, A.J.: SICs: Extending the list of solutions. Preprint (2017). [arXiv:1703.03993](https://arxiv.org/abs/1703.03993) [quant-ph]
- Appleby, M., Chien, T.Y., Flammia, S., Waldron, S.: Constructing exact symmetric informationally complete measurements from numerical solutions. Preprint (2017). [arXiv:1703.05981](https://arxiv.org/abs/1703.05981) [quant-ph]
- Appleby, M., Fuchs, C.A., Stacey, B.C., Zhu, H.: Introducing the Qplex: a novel arena for quantum theory. [arXiv:1612.03234](https://arxiv.org/abs/1612.03234) [quant-ph]
- Khrennikov, A.: External observer reflections on QBism. Preprint (2016). [arXiv:1512.07195](https://arxiv.org/abs/1512.07195) [quant-ph]
- Khrennikov, A.: Towards better understanding QBism. *Found. Sci.* **18**, 1–15 (2017)
- Schleich, W.P.: *Quantum Optics in Phase Space*. Wiley-VCH, Berlin (2001)
- Prugovecki, E.: Simultaneous measurement of several observables. *Found. Phys.* **3**, 3–18 (1973)
- Feynman, R.P.: Negative probability. In: Hiley, B.J., Peat, F.D. (eds.) *Quantum Implications, Essays in Honour of David Bohm*, pp. 235–246. Routledge and Kegan Paul, London (1987)
- Khrennikov, A.: On the physical interpretation of negative probabilities in Prugovecki's empirical theory of measurement. *Can. J. Phys.* **75**, 291–298 (1997)
- Khrennikov, A.: *Interpretations of Probability*, 2nd edn. De Gruyter, Berlin (2009)
- Burgin, M.: Interpretations of negative probabilities. Preprint (2010). [arXiv:1008.1287](https://arxiv.org/abs/1008.1287) [physics.data-an]
- Ferrie, C., Morris, R., Emerson, J.: Necessity of negativity in quantum theory. *Phys. Rev. A* **82**, 044103 (2010)
- Wooters, W.K.: A Wigner-function formulation of finite-state quantum mechanics. *Ann. Phys.* **176**, 1 (1987)

31. Gibbons, K.S., Hoffman, M.J., Wootters, W.K.: Discrete phase space based on finite fields. *Phys. Rev. A* **70**, 062101 (2004)
32. Ferrie, C., Emerson, J.: Framed Hilbert space: hanging the quasi-probability pictures of quantum theory. *N. J. Phys.* **11**, 063040 (2009)
33. Ferrie, C.: Quasi-probability representations of quantum theory with applications to quantum information science. *Rep. Prog. Phys.* **74**, 116001 (2011)
34. Veitch, V., Ferrie, C., Gross, D., Emerson, J.: Negative quasi-probability as a resource for quantum computation. *N. J. Phys.* **14**, 113011 (2012)
35. Howard, M., Wallman, J., Veitch, V., Emerson, J.: Contextuality supplies the ‘magic’ for quantum computation. *Nature* **510**, 351–355 (2014)
36. Pashayan, H., Wallman, J.J., Bartlett, S.D.: Estimating outcome probabilities of quantum circuits using quasiprobabilities. *Phys. Rev. Lett.* **115**, 070501 (2015)
37. Veitch, V., Mousavian, S.A.H., Gottesman, D., Emerson, J.: The resource theory of stabilizer quantum computation. *N. J. Phys.* **16**, 013009 (2014)
38. Bhatia, R.: *Matrix Analysis*. Graduate Texts in Mathematics. Springer, New York (1997)
39. Bertsekas, D.: *Convex Analysis and Optimization*. Athena Scientific, Belmont (2003)
40. Tabia, G.N.M., Appleby, D.M.: Exploring the geometry of qutrit state space using symmetric informationally complete probabilities. *Phys. Rev. A* **88**, 012131 (2013)
41. Wootters, W.K., Fields, B.D.: Optimal state-determination by mutually unbiased measurements. *Ann. Phys.* **191**, 363–381 (1989)
42. Szymusiak, A.: Maximally informative ensembles for SIC-POVMs in dimension 3. *J. Phys. A* **47**, 445301 (2014)
43. Stacey, B.C.: Geometric and information-theoretic properties of the Hoggar lines. Preprint (2016). [arXiv:1609.03075](https://arxiv.org/abs/1609.03075) [quant-ph]
44. Appleby, D.M., Flammia, S., McConnell, G., Yard, J.: Generating ray class fields of real quadratic fields via complex equiangular lines. Preprint (2016). [arXiv:1604.06098](https://arxiv.org/abs/1604.06098) [math.NT]
45. Appleby, M., Flammia, S., McConnell, G., Yard, J.: SICs and algebraic number theory. Preprint (2017). [arXiv:1701.05200](https://arxiv.org/abs/1701.05200) [quant-ph]
46. Bengtsson, I.: The number behind the simplest SIC-POVM. Preprint (2016). [arXiv:1611.09087](https://arxiv.org/abs/1611.09087) [quant-ph]
47. Ramirez, C., Sanchez, R., Kreinovich, V., Argaez, M.:  $\sqrt{x^2 + c}$  is the most computationally efficient smooth approximation to  $|x|$ . *J. Uncertain Syst.* **8**, 205–210 (2014)
48. Pike, R.: *Optimization for Engineering Systems*. Van Nostrand Reinhold, New York (1986)
49. Szymusiak, A., Słomczyński, W.: Informational power of the Hoggar symmetric informationally complete positive operator-valued measure. *Phys. Rev. A* **94**, 012122 (2016)