OPEN ACCESS
IOP Publishing

Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **54** (2021) 435302 (16pp)

https://doi.org/10.1088/1751-8121/ac2717

Higher-order uncertainty bounds for mixed states

Alex J Belfield[1] **and Dorje C Brody**[1,2,*]

[1] Department of Mathematics, University of Surrey, Guildford GU2 7XH, United Kingdom
[2] St Petersburg National Research University of Information Technologies, Mechanics and Optics, St Petersburg 197101, Russia

E-mail: d.brody@surrey.ac.uk

Received 3 December 2020, revised 20 August 2021
Accepted for publication 15 September 2021
Published 4 October 2021

Abstract

Uncertainty lower bounds for parameter estimations associated with a unitary family of mixed-state density matrices are obtained by embedding the space of density matrices in the Hilbert space of square-root density matrices. In the Hilbert-space setup the measure of uncertainty is given by the skew information of the second kind, while the uncertainty lower bound is given by the Wigner–Yanase skew information associated with the conjugate observable. Higher-order corrections to the uncertainty lower bound are determined by higher-order quantum skew moments; expressions for these moments are worked out in closed form.

Keywords: uncertainty bound, skew information, quantum state estimation, mixed state

1. Introduction

The Heisenberg uncertainty relation can be interpreted in a variety of ways [1, 2], but perhaps operationally the most direct and mathematically the most straightforward way of understanding the relation is in the context of quantum state estimation. Take, for instance, the energy–time uncertainty relation $\Delta T \Delta H \geqslant \hbar/2$. The statistical interpretation of this relation is as follows. At time zero we prepare the system in a state, say, a pure state $|\psi_0\rangle$, and let the system evolve under the influence of the Hamiltonian $\hat{H}$. At a later point we wish to estimate how much time has elapsed since the initial preparation of the system. This amounts to estimating the parameter t in the one-parameter family of states $|\psi_t\rangle = \mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}|\psi_0\rangle$ in Hilbert

[*] Author to whom any correspondence should be addressed.

space. The estimation process is facilitated by use of an estimator $\hat{T}$, which is a maximally symmetric operator satisfying the properties that $\langle\psi_t|\hat{T}|\psi_t\rangle = t$ for all t and that $\mathrm{i}[\hat{H},\hat{T}] = \hbar$. The existence of such an operator $\hat{T}$, albeit not self-adjoint, for each given $\hat{H}$, is by now well understood [3].

An analogous setup can be envisaged in the case of the position-momentum uncertainty relation $\Delta Q\Delta P \geqslant \hbar/2$. In this case the parametric family of states for the system, say, a particle, is given in the coordinate-space representation by $|\psi(q)\rangle = \mathrm{e}^{-\mathrm{i}\hat{P}q/\hbar}|\psi(0)\rangle$, with a given wave function $|\psi(0)\rangle$ at the origin of the parameter space. Assuming, without loss, that $\langle\psi(0)|\hat{Q}|\psi(0)\rangle = 0$, then because the mean position $\langle\psi(q)|\hat{Q}|\psi(q)\rangle = q$ is given by the parameter q, the parameter estimation for the state $|\psi(q)\rangle$ amounts to estimating the position of the particle.

In either case, to interpret the meaning of the uncertainty relation, we need to consider estimation errors. Specifically, in the context of parameter estimation in classical statistics, the quadratic error bound associated with an estimate is of great interest. In the classical context of parametric statistics one has a model, characterised by a parametric family of density functions $p(x|\theta)$ that is postulated to describe the data generated by sampling the value of a random variable X. The value of the parameter θ, however, is unknown, and will have to be estimated from the given data set. If $\Theta(x)$ were an unbiased estimator for the parameter θ so that $\int \Theta(x)p(x|\theta)\mathrm{d}x = \theta$, then by substituting the sampled value of X in $\Theta(x)$ we get an estimate for θ.

To arrive at the lower bound for the quadratic estimation error (the variance), Rao [4] considered the embedding of the density function in Hilbert space via the square-root map $\xi_\theta(x) = \sqrt{p(x|\theta)}$. Utilising the Dirac notation for (real) Hilbert space operations by writing, for instance, $|\xi_\theta\rangle = \xi_\theta(x)$ and $\langle\xi_\theta|\hat{\Theta}|\xi_\theta\rangle = \int \Theta(x)\xi_\theta(x)^2\,\mathrm{d}x = \theta$, we have $\langle\xi_\theta|(\hat{\Theta}-\theta)|\xi_\theta\rangle = 0$. Differentiating this with respect to θ, and using the relations $\langle\dot{\xi}_\theta|\xi_\theta\rangle = \langle\xi_\theta|\dot{\xi}_\theta\rangle = 0$ and $\langle\dot{\xi}_\theta|\hat{\Theta}|\xi_\theta\rangle = \langle\xi_\theta|\hat{\Theta}|\dot{\xi}_\theta\rangle$, we deduce that $\langle\dot{\xi}_\theta|(\hat{\Theta}-\theta)|\xi_\theta\rangle = \frac{1}{2}$. Then from the Schwarz inequality $\langle\dot{\xi}_\theta|\eta\rangle^2 \leqslant \langle\dot{\xi}_\theta|\dot{\xi}_\theta\rangle\langle\eta|\eta\rangle$ we deduce by setting $|\eta\rangle = (\hat{\Theta}-\theta)|\xi_\theta\rangle$ the Cramér–Rao inequality

$$\Delta\Theta^2 \geqslant \frac{1}{4\langle\dot{\xi}_\theta|\dot{\xi}_\theta\rangle}, \tag{1}$$

where we wrote $\Delta\Theta^2 = \langle\xi_\theta|(\hat{\Theta}-\theta)^2|\xi_\theta\rangle$ for the variance of the estimate. The quantity $4\langle\dot{\xi}|\dot{\xi}\rangle$ appearing on the right side is the Fisher information [5] (in the case of a multi-parameter family of densities the variance is replaced by the covariance matrix, and the Fisher information is replaced by the Fisher information matrix). The Cramér–Rao inequality (1) shows that the more information one can extract from sampling for the value of the unknown parameter θ, the smaller the estimation error bound is.

There is a second, geometric interpretation of the Fisher information, namely, that it gives the speed at which the state $|\xi_\theta\rangle$ changes in Hilbert space, as the parameter θ is varied. Thus, if the state $|\xi_\theta\rangle$ is sensitive to the parameter θ then the estimation error can be made small, whereas if the state hardly changes when the parameter is varied, then the estimation error will be large. The advantage of Rao's Hilbert space formulation of classical statistics is that it is readily applicable, *mutatis mutandis*, to the problem of quantum state estimation. In particular, in the context of the above-posed state estimation problem, a short calculation, making use of the Schrödinger–Kibble dynamical equation $|\dot{\xi}_t\rangle = -\mathrm{i}\hbar^{-1}(\hat{H}-\langle\hat{H}\rangle)|\xi_t\rangle$ that removes the dynamical phase, shows that $\langle\dot{\xi}_t|\dot{\xi}_t\rangle = \Delta H^2/\hbar^2$. It follows that the Cramér–Rao inequality reduces to the Heisenberg uncertainty relations [6].

For quantum state estimation, starting from the Heisenberg relation there are two directions in which the analysis can be extended. The first concerns the investigation to obtain sharper bounds for the variance. For sure for some states the quadratic error is small (e.g. for a coherent state the position variance is equal to the inverse of the Fisher information), but for other states this is not the case. The higher-order corrections to the Heisenberg relation, leading to sharper uncertainty relations, were first obtained in [6], where by 'order' we mean the degree of dispersion of the conjugate operator (but not in terms of, say, powers of Planck's constant so that contributions from higher-order terms need not be small). Thus, for instance, one can deduce that

$$\Delta Q^2 \Delta P^2 \geqslant \frac{\hbar^2}{4}\left(1 + \frac{(\Delta P^4 - 3(\Delta P^2)^2)^2}{\Delta P^6 \Delta P^2 - (\Delta P^4)^2}\right) \tag{2}$$

and so on [7], where we have written ΔP^n to mean the nth central moment of $\hat{P}$. In fact, it is possible to work out an infinitely many such higher-order corrections explicitly [8].

The second direction concerns the parameter estimation when the state of the system is described by a mixed-state density matrix. A mixed quantum state represents a probabilistic mixture of pure states. Thus, the problem of parameter estimation becomes more difficult, because on top of the intrinsic quantum-mechanical uncertainty represented by the pure states, there is an added, essentially classical, uncertainty regarding which pure state might be the 'correct' state to describe the system. As a consequence, the Fisher information—the information that can be extracted from sampling, or measurements—is necessarily reduced from the variance of the conjugate variable, in the context of parameter estimation associated with a unitary curve in the space of mixed states. Specifically, it reduces to another information measure introduced by Wigner and Yanase [9], as demonstrated in [10–12].

There is in fact a long history of research in developing a theoretical framework for investigating the quantum state estimation problem. In particular, starting with the work of Helstrom [14], the Fisher information associated with a parametric family of density matrices is often investigated by use of the technique of a symmetric logarithmic derivative. This is quite natural inasmuch as in classical statistics the Fisher information is commonly represented in terms of the log-likelihood function $\log p(x|\theta)$ (except in the work of Rao where it is represented in terms of the square-root density function $\sqrt{p(x|\theta)}$). Indeed, we have the identity

$$\int p(x|\theta)\left(\frac{\mathrm{d}\,\log\, p(x|\theta)}{\mathrm{d}\theta}\right)^2 \mathrm{d}x = 4\int \left(\frac{\mathrm{d}\sqrt{p(x|\theta)}}{\mathrm{d}\theta}\right)^2 \mathrm{d}x. \tag{3}$$

Utilising the technique of symmetric logarithmic derivatives, Braunstein and Caves [15] maximised a representation for the Fisher information over all quantum states to show that for pure states the Cramér–Rao inequality reduces to the Heisenberg uncertainty relation.

Although the use of symmetric logarithmic derivatives in the context of quantum state estimation has been very popular in the literature since the work of [15], there are several limitations associated with this approach when attempting to obtain sharper uncertainty relations for mixed states. For the same token, the relation to the geometry of the underlying Hilbert space, as elegantly exploited by Rao in the case of classical statistics, becomes obscure. The point is that in classical parametric statistics one often works in the space of log-density functions, as opposed to the space of densities directly; the latter is just the subspace of the first Lebesgue class functions $\mathcal{L}^1(\mathbb{R})$. Thus, for instance, Efron's measure for information loss in higher-order asymptotic inference is given by the curvature of the parametric curve $l_\theta(x) = \log p(x|\theta)$ in the space of log-likelihood functions [16]. In contrast, Rao considered the Hilbert-space embedding of the parametric family of density functions, for which the more familiar machineries on the space $\mathcal{L}^2(\mathbb{R})$ of second Lebesgue class functions in mathematical analysis can be applied.

In particular, the geometric meaning of the Fisher information then becomes immediately apparent [4].

In a similar vein, when working with density matrices in quantum theory, one can either consider the log-density matrix or the square-root density matrix; but it is the latter that lends itself with the geometry of the associated Hilbert space. The idea is that by considering a Hermitian square root $\hat{\xi}$ of a density matrix $\hat{\rho} = \hat{\xi}^2$ (any real root would suffice), we can embed the space of density matrices in a (real) Hilbert space of 'pure' states $\hat{\xi}$ equipped with the Hilbert–Schmidt (or just the trace) inner product. In contrast, the log-density matrix is not an element of Hilbert space. This line of thinking was exploited in [12] to derive a modified form of uncertainty relations that arise naturally in the context of the Hilbert space of square-root density matrices.

With these preliminaries, the purpose of the present paper is to derive higher-order corrections to the uncertainty lower bounds obtained in [10–12]. We shall follow, in particular, the approach developed in [8], but instead of pure states considered therein we will be working with square roots of mixed state density matrices. The analysis presented in what follows, however, are not mere generalisations of the work in [8] from pure to mixed states in that, perhaps surprisingly, in the limit of pure states our results do not reproduce those of [8]. Thus our results give rise also to a genuinely new set of higher-order corrections to the uncertainty lower bounds for pure states as well. More generally, for mixed states, we shall encounter modified forms of higher-order central moments that arise naturally. We shall refer to these as the higher-order 'quantum skew moments' in analogy with the Wigner–Yanase skew information, which is a modified form of the second central moment (variance). We shall determine the form of these skew moments, and demonstrate how corrections to the uncertainty relation, to an arbitrary high order, can be derived explicitly in a recursive manner, expressed in terms of the various skew moments.

2. Mixed state uncertainties and the skew information

Let us begin by briefly reviewing the approach proposed in [12] for benefits of readers less acquainted with the material. This will also be useful in setting the notations. Starting from a one-parameter family of states $\hat{\rho}_t = \exp(-\mathrm{i}\hat{H}t)\hat{\rho}_0 \exp(\mathrm{i}\hat{H}t)$ associated with a prescribed initial state $\hat{\rho}_0$ we let $\hat{\xi}_t$ be an arbitrary Hermitian square-root of $\hat{\rho}_t$ so that $\hat{\xi}_t^2 = \hat{\rho}_t$. The objective in state estimation here is to determine the elapse time t since the initial preparation of the state. We let $\hat{T}$ be an unbiased estimator for the time parameter t so that $\mathrm{tr}(\hat{T}\hat{\xi}_t^2) = t$.

To proceed it will be useful to introduce the notation for the mean-adjusted operator by writing $\tilde{T} = \hat{T} - t\mathbb{1}$ so that the variance of $\hat{T}$ is given by $\Delta T^2 = \mathrm{tr}(\tilde{T}^2\hat{\xi}_t^2)$. Differentiating the normalisation condition $\mathrm{tr}(\hat{\xi}_t^2) = 1$ in t and writing $\hat{\xi}_t' = \partial_t\hat{\xi}_t$, we obtain $\mathrm{tr}(\hat{\xi}_t'\hat{\xi}_t) = 0$. Differentiating $\mathrm{tr}(\tilde{T}\xi_t^2) = 0$ in t and using $\tilde{T}' = -\mathbb{1}$ we find $\mathrm{tr}[(\tilde{T}\hat{\xi}_t + \hat{\xi}_t\tilde{T})\hat{\xi}_t'] = 1$. Applying the matrix Schwarz inequality, we deduce a form of the quantum Cramér–Rao inequality

$$\Delta T^2 + \delta T^2 \geqslant \frac{1}{2\,\mathrm{tr}(\hat{\xi}_t'\hat{\xi}_t')}, \tag{4}$$

where $\delta T^2 = \mathrm{tr}(\tilde{T}\hat{\xi}_t\tilde{T}\hat{\xi}_t) - t^2$. As for the Fisher information, by use of the evolution equation $\hat{\xi}_t' = -\mathrm{i}(\hat{H}\hat{\xi}_t - \hat{\xi}_t\hat{H})$, where we shall be working in units $\hbar = 1$, we find

$$\mathrm{tr}(\hat{\xi}_t'\hat{\xi}_t') = 2\left[\mathrm{tr}(\hat{H}^2\hat{\xi}_t^2) - \mathrm{tr}(\hat{H}\hat{\xi}_t\hat{H}\hat{\xi}_t)\right], \tag{5}$$

which is twice the skew information introduced by Wigner and Yanase [9].

For a pure state satisfying $\hat{\xi}_t^2 = \hat{\xi}_t$, the skew information is maximised and (5) reduces to twice the energy variance $2\Delta H^2$, whereas $\delta T^2 = 0$ for pure states. Thus for pure state we recover the standard uncertainty relation. For mixed states, however, we obtain a modified form of uncertainty relation

$$\Delta T^2 + \delta T^2 \geqslant \frac{1}{4(\Delta H^2 - \delta H^2)}. \tag{6}$$

Note that

$$\Delta H^2 - \delta H^2 = \mathrm{tr}(\hat{H}^2\hat{\rho}) - \mathrm{tr}(\hat{H}\sqrt{\hat{\rho}}\hat{H}\sqrt{\hat{\rho}}) \tag{7}$$

appearing on the right side of (6) is the Wigner–Yanase skew information associated with $\hat{H}$, which is a positive quantity that is strictly smaller than the variance ΔH^2 for mixed states (because $\Delta H^2 \geqslant \delta H^2 \geqslant 0$ for any state [13]) and is equal to the variance for pure states (because $\delta H^2 = 0$ for pure states). The skew information can be worked out explicitly for any given initial state $\hat{\rho}_0$ and the Hamiltonian $\hat{H}$ of the system. As an elementary example, let the initial state of, say, a two-level system be diagonal with elements p and $1 - p$, whereas the Hamiltonian is given by $\hat{H} = \omega\hat{\sigma}_x$ for some real parameter ω. Then a short calculation shows that

$$\Delta H^2 - \delta H^2 = \omega^2 \left(1 - 2\sqrt{p(1-p)}\right), \tag{8}$$

which explicitly shows that the skew information is a monotonically increasing function of the purity of the state.

The quantity appearing on the left side of (6),

$$\Delta T^2 + \delta T^2 = \mathrm{tr}(\hat{T}^2\hat{\rho}) + \mathrm{tr}(\hat{T}\sqrt{\hat{\rho}}\hat{T}\sqrt{\hat{\rho}}) - 2\left(\mathrm{tr}(\hat{T}\hat{\rho})\right)^2, \tag{9}$$

is strictly greater than the variance ΔT^2 for mixed states and is equal to the variance of $\hat{T}$ for pure states. We shall refer to (9) as the 'skew information of the second kind'.

What the generalised uncertainty relation (6) suggests is that for mixed states, perhaps the more appropriate measure of uncertainty error in parameter estimation is the skew information of the second kind, rather than the variance, whereas the sensitivity of the state in parameter variation, as measured by the Fisher information, is given by the Wigner–Yanase skew information, again rather than the variance.

Let us elaborate on this point so as to clarify the interpretation of (6). For this purpose, suppose that the state $\hat{\rho}$ is a stationary state satisfying $[\hat{\rho}, \hat{H}] = 0$. Then we have $\delta H^2 = \Delta H^2$, and hence the uncertainty lower bound for time estimation on the right side of (6) diverges. This indeed has to be the case because if the state is stationary, then no information about the time parameter can be obtained from any experiment. For pure states, one arrives at the same conclusion from the variance-based Heisenberg relation because $\Delta H^2 = 0$ for stationary states. However, for a mixed stationary state we necessarily have $\Delta H^2 > 0$, even though $\Delta T^2 = \infty$. It follows that the variance-based measures are not appropriate for fully characterising estimation errors and their bounds when it comes to mixed states. In contrast, inequality (6) fully captures the notion of an error bound in all circumstances. Further, because the Wigner–Yanase skew information in the present context is the Fisher information in the sense of Rao, the skew information of the second kind $\Delta T^2 + \delta T^2$, which is attainably bounded by the inverse of the Fisher information, has to be interpreted as the correct measure of estimation error for mixed states.

3. Geometric derivation of the quantum Cramér–Rao inequality

The idea that we shall exploit in order to obtain higher-order corrections to the uncertainty relation (6) is as follows. We consider the Hilbert space $\mathcal{H}$ of square-root density matrices equipped with the trace inner product. Note that $\mathcal{H}$ is not the Hilbert space of states that describes the system; rather, it is a larger Hilbert space in which the density matrices are embedded via the square-root map. Thus, on $\mathcal{H}$, the square-root of a mixed state $\hat{\rho}$ is interpreted like a 'pure' state vector, where the inner product of two such vectors $\hat{\xi}$ and $\hat{\eta}$ is defined by $\mathrm{tr}(\hat{\xi}\hat{\eta})$. The advantage of working on $\mathcal{H}$ is that vectorial operations that have been used effectively in [6–8] for statistical analysis of pure states can be extended into the domain of mixed states, albeit the calculations do become more elaborate.

To begin, we shall consider level surfaces $t(\hat{\xi}) = c$ in $\mathcal{H}$ associated with constant expectation values of the operator $\hat{T}$:

$$t(\hat{\xi}) = \frac{\mathrm{tr}(\hat{T}\hat{\xi}^2)}{\mathrm{tr}(\hat{\xi}^2)}. \tag{10}$$

Note that elements of $\mathcal{H}$ satisfy the condition $\mathrm{tr}(\hat{\xi}^2) < \infty$, and the normalisation condition $\mathrm{tr}(\hat{\xi}^2) = 1$ is always imposed after performing calculations. In this way, projectively meaningful results can be obtained. Next, we consider the vector in $\mathcal{H}$ that is normal to the level surface $t(\hat{\xi}) = c$ at $\hat{\xi}$. A calculation shows that

$$\hat{\nabla}t = \left.\frac{\partial t}{\partial \hat{\xi}}\right|_{\mathrm{tr}(\hat{\xi}^2)=1} = \hat{\xi}\hat{T} + \hat{T}\xi - 2\,\mathrm{tr}(\hat{T}\hat{\xi}^2)\,\hat{\xi}, \tag{11}$$

whose squared magnitude gives

$$|\hat{\nabla}t|^2 = 2\left(\mathrm{tr}(\hat{T}^2\hat{\xi}^2) + \mathrm{tr}(\hat{T}\hat{\xi}\hat{T}\hat{\xi}) - 2\Big(\mathrm{tr}(\hat{T}\hat{\xi}^2)\Big)^2\right). \tag{12}$$

Observe from (9) that this is precisely twice the left side of (6) when $\hat{\xi}$ is the solution to the dynamical equation $\hat{\xi}'_t = -\mathrm{i}(\hat{H}\hat{\xi}_t - \hat{\xi}_t\hat{H})$. It follows that the uncertainties associated with estimating the parameter t is given exactly by (rather than bounded by) half the squared magnitude of the normal vector $\hat{\nabla}t$ in $\mathcal{H}$.

With this observation we are in the position to estimate the squared length of the vector $\hat{\nabla}t$. To this end we use the elementary fact that the squared norm of a vector is equal to the sum of the squared norms of its orthogonal components with respect to any choice of an orthonormal frame. To fix a frame for $\mathcal{H}$ we shall be using the vector $\hat{\xi}_t$ and its higher-order derivatives in t, and then orthogonalise them using the standard Gram–Schmidt scheme. The first step therefore is to identify the derivatives of the state. From $\hat{\xi}_t = \mathrm{e}^{-\mathrm{i}\hat{H}t}\,\hat{\xi}_0\,\mathrm{e}^{\mathrm{i}\hat{H}t}$ we have $\hat{\xi}'_t = \partial_t\hat{\xi}_t = \mathrm{i}(\hat{\xi}_t\hat{H} - \hat{H}\hat{\xi}_t)$. Another differentiation then gives

$$\hat{\xi}''_t = -\hat{H}^2\hat{\xi}_t + 2\hat{H}\hat{\xi}_t\hat{H} - \hat{\xi}_t\hat{H}^2, \tag{13}$$

and so on. More generally, a short calculation shows that

$$\hat{\xi}^{(n)}_t = (-\mathrm{i})^n \sum_{k=0}^{n} (-1)^k \binom{n}{k} \hat{H}^{n-k}\,\hat{\xi}_t\,\hat{H}^k. \tag{14}$$

The idea is to use these derivatives to form the proper velocity, acceleration, and higher-order analogues of them, which in turn form the basis for $\mathcal{H}$. Letting $\{\hat{\Psi}_n\}$ denote the resulting orthogonal vectors, we thus find

$$\Delta T^2 + \delta T^2 = \frac{1}{2}\sum_{n=0}^{\infty} \frac{\left[\mathrm{tr}\left((\hat{\xi}_t\hat{T} + \hat{T}\hat{\xi}_t - 2t\hat{\xi})\hat{\Psi}_n\right)\right]^2}{\mathrm{tr}(\hat{\Psi}_n\hat{\Psi}_n)}. \tag{15}$$

The next step is to determine the set of vectors $\{\hat{\Psi}_n\}$. We let $\hat{\Psi}_0 = \hat{\xi}_t$. We then see at once that $\mathrm{tr}\left((\hat{\xi}_t\hat{T} + \hat{T}\hat{\xi}_t - 2t\hat{\xi})\hat{\Psi}_0\right) = 0$. In other words, $\hat{\Psi}_0 = \hat{\xi}_t$ is orthogonal to $\hat{\nabla}t$ so that the $n = 0$ term in the sum (15) makes no contribution. For $n = 1$ we let $\hat{\Psi}_1$ be the component of the first derivative $\hat{\xi}_t'$ orthogonal to $\hat{\xi}_t$. However, because $\mathrm{tr}(\hat{\xi}_t'\hat{\xi}_t) = 0$ we have $\hat{\Psi}_1 = \hat{\xi}_t'$ for the proper velocity vector. To identify the $n = 1$ contribution we need to calculate $\mathrm{tr}\left((\hat{\xi}_t\hat{T} + \hat{T}\hat{\xi}_t)\hat{\xi}_t'\right)$. In fact we have already deduced this trace (see just above equation (4)), but let us calculate this explicitly. Using the dynamical equation $\hat{\xi}_t' = -\mathrm{i}(\hat{H}\hat{\xi}_t - \hat{\xi}_t\hat{H})$ and the cyclic property of the trace operation we get $\mathrm{tr}\left((\hat{\xi}_t\hat{T} + \hat{T}\hat{\xi}_t)\hat{\xi}_t'\right) = -\mathrm{tr}\left(\mathrm{i}(\hat{\xi}_t^2\hat{T}\hat{H} - \hat{\xi}_t^2\hat{H}\hat{T})\right)$, so from the commutation relation $\mathrm{i}[\hat{H}, \hat{T}] = 1$ we see that the dependence on the estimator $\hat{T}$ drops out. Together with (5) we thus deduce (6) at once because the truncation of the sum at $n = 1$ is necessarily smaller than or equal to the entire sum.

This geometric derivation of (6) can be visualised as follows. In Hilbert space $\mathcal{H}$ we have a curve $\hat{\xi}_u$, where $u \geqslant 0$. The arc length $s(t)$ of the curve is defined by

$$s(t) = \int_0^t |\hat{\xi}_u'|\,\mathrm{d}u = \sqrt{2(\Delta H^2 - \delta H^2)}\,t, \tag{16}$$

which evidently is proportional to t because the skew information of the energy is constant along the curve of unitary motion. Thus, the objective of time estimation is in effect to estimate the arc length of the curve. We now slice $\mathcal{H}$ by 'space-like' hypersurfaces of constant expectations values of $\hat{T}$ along the curve. If the normal vector to the hypersurface at $\hat{\xi}_t$ is parallel to the tangent vector of the curve, then the estimation error is minimised and we have a minimum uncertainty state. Indeed, the unit normal vector to the surface at $\hat{\xi}_t$ is given by

$$\hat{\boldsymbol{n}} = \frac{\hat{\xi}_t\hat{T} + \hat{T}\hat{\xi}_t - 2t\hat{\xi}}{\sqrt{2(\Delta T^2 + \delta T^2)}}, \tag{17}$$

whereas the unit tangent vector of the curve at $\hat{\xi}_t$ is given by

$$\hat{\boldsymbol{e}}_1 = \frac{-\mathrm{i}(\hat{H}\hat{\xi}_t - \hat{\xi}_t\hat{H})}{\sqrt{2(\Delta H^2 - \delta H^2)}}. \tag{18}$$

The angular separation of these vectors, defined by $\cos\theta = \hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{e}}_1$, is thus given by

$$\theta = \cos^{-1}\left(\frac{1}{2\sqrt{(\Delta T^2 + \delta T^2)(\Delta H^2 - \delta H^2)}}\right). \tag{19}$$

For a minimum uncertainty state we have $\theta = 0$, but in a generic case we have $\theta \neq 0$. The angle θ thus can be viewed as the universal measure for the uncertainty associated with quantum state estimation in the case of a one-parameter family of quantum states.

4. Quantum skew moments

To work out the higher-order uncertainty bounds we shall encounter inner products of the form $\mathrm{tr}(\hat{\xi}_t^{(m)}\hat{\xi}_t^{(n)})$, where $n+m$ is even. Before we proceed to examine the bounds more closely, let us analyse this inner product first because it leads to the notion of *higher-order quantum skew central-moments*, first envisaged in [12] but has not been worked out previously. Specifically, in our analysis it is the even-order skew moments that determine the higher-order corrections, in a way analogous to the bound given in (2). We shall define the $(n+m)$th order skew central moment of the Hamiltonian by $S_{n+m} = \mathrm{tr}(\hat{\xi}_t^{(n)}\hat{\xi}_t^{(m)})$ if $n+m = 4k+2$; whereas $S_{n+m} = -\mathrm{tr}(\hat{\xi}_t^{(n)}\hat{\xi}_t^{(m)})$ if $n+m = 4k$ for all $k = 0, 1, 2, \ldots$. The intuition of the appearance of the minus sign here when $n+m$ is a multiple of four is as follows. For a central moment we demand that the highest-order moment $\mathrm{tr}(\hat{H}^{n+m}\hat{\xi}^2)$ appearing inside to have the positive sign, but when the states are differentiated in the inner product $\mathrm{tr}(\hat{\xi}_t^{(n)}\hat{\xi}_t^{(m)})$ we get a factor of $-\mathrm{i}^{n+m}$ in front of the highest-order moment term. Hence we insert an extra minus sign to define the skew moments when $-\mathrm{i}^{n+m} = -1$, or equivalently, when $n+m = 4k$. With these preliminaries we have the following:

Lemma. *The higher-order quantum skew central-moments are given by*

$$S_{4L+2} = 2\sum_{k=0}^{2L}(-1)^k \binom{4L+2}{k} \mathrm{tr}(\hat{H}^{4L+2-k}\hat{\xi}_t\hat{H}^k\hat{\xi}_t) - \binom{4L+2}{2L+1} \mathrm{tr}(\hat{H}^{2L+1}\hat{\xi}_t\hat{H}^{2L+1}\hat{\xi}_t) \tag{20}$$

and

$$S_{4L} = 2\sum_{k=0}^{2L-1}(-1)^k \binom{4L}{k} \mathrm{tr}(\hat{H}^{4L-k}\hat{\xi}_t\hat{H}^k\hat{\xi}_t) + \binom{4L}{2L} \mathrm{tr}(\hat{H}^{2L}\hat{\xi}_t\hat{H}^{2L}\hat{\xi}_t) \tag{21}$$

for all $L > 0$, *whereas for* $L = 0$ *we have* $S_0 = 1$.

Proof. To verify these, let us first consider the case $m+n = 4L+2$. Then we have $n = 4L+2-m$ and hence from (14) we find

$$\mathrm{tr}(\hat{\xi}_t^{(n)}\hat{\xi}_t^{(m)}) = -\sum_{k=0}^{4L+2-m}\sum_{l=0}^{m}(-1)^{k+l}\binom{4L+2-m}{k} \times \binom{m}{l} \mathrm{tr}\left(\hat{H}^{4L+2+l-k-m}\hat{\xi}_t\hat{H}^{m+k-l}\hat{\xi}_t\right). \tag{22}$$

Now let $\alpha = m+k-l$. Then α ranges from 0 to $4L+2$ (specifically, $\alpha = 0$ when $k = 0$ and $l = m$; whereas $\alpha = 4L+2$ when $l = 0$ and $k = 4L+2$). The trace term in (22) can then be written in the form $\mathrm{tr}(\hat{H}^{4L+2-\alpha}\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t)$. For $\alpha = 0$, corresponding to $(k, l) = (0, m)$, the product of the binomial coefficients in (22) gives 1. For $\alpha = 1$, corresponding to $(k, l) = (0, m-1)$ or $(k, l) = (1, m)$, the product of the binomial coefficients give m or $4L+2-m$, both with the same sign, so they add up to give $4L+2 = (4L+2)!/1!(4L+2-1)!$. For $\alpha = 2$, corresponding to $(k, l) = (0, m-2)$, $(k, l) = (1, m-1)$, or $(k, l) = (2, m)$, the product of the binomial coefficients give $m(m-1)$, $m(4L+2-m)$, or $(4L+2-m)(4L+2-m-1)/2$, again all with the same sign, so they add up to give $(4L+2)!/2!(4L+2-2)!$. Continuing along this line we observe that the double sum in (22) can be replaced with a single sum over α, along with the binomial coefficient $(4L+2)!/\alpha!(4L+2-\alpha!)$ and with alternating signs:

$$\mathrm{tr}(\hat{\xi}_t^{(n)}\hat{\xi}_t^{(m)}) = (-1)^{m+1}\sum_{\alpha=0}^{4L+2}(-1)^{\alpha}\begin{pmatrix}4L+2\\ \alpha\end{pmatrix}\mathrm{tr}\left(\hat{H}^{4L+2-\alpha}\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t\right). \tag{23}$$

Notice that the sum contains an odd number of terms. Apart from the middle term in the sum for which $\alpha = 2L+1$, each term appears twice on account of the symmetry of binomial coefficients, so we can truncate the sum in (23) at $\alpha = 2L$, double the result, and add the term for $\alpha = 2L+1$; this gives (20). An essentially identical line of argument then gives (21) when $m+n = 4L$. □

We note that the even-order skew moment S_{2n} defined here does not reduce to the nth central moment in the pure-state limit (except for $n = 1$). To see this it suffices to consider S_4. Then from (21) we have

$$S_4 = 2\left[\mathrm{tr}(\hat{H}^4\hat{\xi}_t^2) - 4\,\mathrm{tr}(\hat{H}^3\hat{\xi}_t\hat{H}\hat{\xi}_t) + 3\,\mathrm{tr}(\hat{H}^2\hat{\xi}_t\hat{H}^2\hat{\xi}_t)\right], \tag{24}$$

and hence, for a pure state for which $\hat{\xi} = \hat{\xi}^2$ we obtain

$$S_4 = 2\left[\langle\hat{H}^4\rangle - 4\langle\hat{H}^3\rangle\langle\hat{H}\rangle + 3\langle\hat{H}^2\rangle^2\right]. \tag{25}$$

Evidently, this is distinct from the fourth central moment of the Hamiltonian

$$\mu_4 = \langle\hat{H}^4\rangle - 4\langle\hat{H}^3\rangle\langle\hat{H}\rangle + 6\langle\hat{H}^2\rangle\langle\hat{H}\rangle^2 - 3\langle\hat{H}\rangle^4. \tag{26}$$

The fact that the skew moments do not reduce to central moments for pure states will have an implication when we compare the pure-state limit of the results obtained here to those obtained in [8].

5. Higher-order variance bounds

The decomposition (15) has one apparent disadvantage in that while the left side represents the magnitude of quadratic error resulting from using the estimator $\hat{T}$, the right side in general is also dependent on $\hat{T}$. For a statistically meaningful error bound we do not wish it to have a dependence on the choice of the estimator. On the other hand, the dependence on $\hat{T}$ dropped out in the contribution from the $n = 1$ term, owing to the commutation relation of $\hat{T}$ and $\hat{H}$. In general, we shall find that the dependence on $\hat{T}$ drops out from all the odd-order terms due to the commutation relation, whereas the even-order terms give rise to the anticommutator of $\hat{T}$ and $\hat{H}$, which we do not know how to evaluate. Thus the approach we have taken here gives rise to a generalisation of the so-called Robertson–Schrödinger uncertainty relation, where the uncertainty lower bound has terms proportional to the commutator, and terms proportional to the anticommutator. That is, the $n = 1$ term gives the Heisenberg-type relation; the $n = 1$ and $n = 2$ terms combined give the Robertson–Schrödinger-type relation; and further additional terms give higher-order corrections to these relations. (An analogous structure emerges in the analysis of [8] for pure states.) In what follows, we shall be focussed on the odd-order terms that we can evaluate explicitly.

The general Gram–Schmidt scheme for generating the orthogonal frame is given as follows:

$$\begin{aligned}
\hat{\Psi}_0 &= \hat{\xi}_t \\
\hat{\Psi}_1 &= \hat{\xi}_t' - \frac{\mathrm{tr}(\hat{\xi}'\hat{\Psi}_0)}{\mathrm{tr}(\hat{\Psi}_0\hat{\Psi}_0)}\hat{\Psi}_0 \\
\hat{\Psi}_2 &= \hat{\xi}_t'' - \frac{\mathrm{tr}(\hat{\xi}_t''\hat{\Psi}_1)}{\mathrm{tr}(\hat{\Psi}_1\hat{\Psi}_1)}\hat{\Psi}_1 - \frac{\mathrm{tr}(\hat{\xi}_t''\hat{\Psi}_0)}{\mathrm{tr}(\hat{\Psi}_0\hat{\Psi}_0)}\hat{\Psi}_0 \\
\hat{\Psi}_3 &= \hat{\xi}_t''' - \frac{\mathrm{tr}(\hat{\xi}_t'''\hat{\Psi}_2)}{\mathrm{tr}(\hat{\Psi}_2\hat{\Psi}_2)}\hat{\Psi}_2 - \frac{\mathrm{tr}(\hat{\xi}_t'''\hat{\Psi}_1)}{\mathrm{tr}(\hat{\Psi}_1\hat{\Psi}_1)}\hat{\Psi}_1 - \frac{\mathrm{tr}(\hat{\xi}_t'''\hat{\Psi}_0)}{\mathrm{tr}(\hat{\Psi}_0\hat{\Psi}_0)}\hat{\Psi}_0 \\
&\vdots
\end{aligned} \tag{27}$$

We have already observed, however, that the inner product $\mathrm{tr}(\hat{\xi}'\hat{\Psi}_0)$ appearing in $\hat{\Psi}_1$ vanishes so that $\hat{\Psi}_1 = \hat{\xi}_t'$. In fact, there are various other cancellations that simplify the expressions for the $\{\hat{\Psi}_n\}$. To this end let us first explore the inner product $\mathrm{tr}(\hat{\xi}_t^{(m)}\hat{\xi}_t^{(n)})$. This inner product vanishes when $n+m$ is an odd number. To see this, it suffices to consider the inner product $\mathrm{tr}(\hat{\xi}_t^{(2m)}\hat{\xi}_t^{(2n+1)})$ because the sum of two integers is odd only if one is even and the other is odd. Then from (14) we obtain

$$\begin{aligned}
\mathrm{tr}\left(\hat{\xi}_t^{(2m)}\hat{\xi}_t^{(2n+1)}\right) = -\mathrm{i}(-1)^{m+n}\sum_{k=0}^{2m}\sum_{l=0}^{2n+1}(-1)^{k+l}\binom{2m}{k}\binom{2n+1}{l} \\
\times\,\mathrm{tr}\left(\hat{\xi}\hat{H}^{2n+1+k-l}\hat{\xi}\hat{H}^{2m-k+l}\right).
\end{aligned} \tag{28}$$

The vanishing of the double sum for all m, n then follows from the symmetry of the binomial coefficients. Specifically, on inspection of the summand:

$$(-1)^{k+l}\frac{\mathrm{tr}\left(\hat{\xi}\hat{H}^{2n+1+k-l}\hat{\xi}\hat{H}^{2m-k+l}\right)}{k!\,(2m-k)!\,l!\,(2n+1-l)!},$$

where we have taken the overall multiplicable constant $(2m)!(2n+1)!$ outside of the sum, we observe the following. For each $(k,l) = (p,q)$ there is another term in the sum with $(k,l) = (2m-p, 2n+1-q)$, which has an identical expression but with an opposite sign, when $p \neq m$. Thus they cancel. When $k = p = m$, there are even terms in the sum and for each term with $l = q$ there is another identical term with $l = 2n+1-q$ having the opposite sign, so they also cancel. This shows that all the terms on the right side of (28) cancel, from which we deduce that $\mathrm{tr}(\hat{\xi}_t^{(2m)}\hat{\xi}_t^{(2n+1)}) = 0$. The elements of the odd-order orthogonal frame that we are interested (the even terms give rise to anticommutator of $\hat{T}$ and $\hat{H}$) are therefore given by the set

$$\begin{aligned}
\hat{\Psi}_1 &= \hat{\xi}_t^{(1)} \\
\hat{\Psi}_3 &= \hat{\xi}_t^{(3)} - \frac{\mathrm{tr}(\hat{\xi}_t^{(3)}\hat{\Psi}_1)}{\mathrm{tr}(\hat{\Psi}_1\hat{\Psi}_1)}\hat{\Psi}_1 \\
\hat{\Psi}_5 &= \hat{\xi}_t^{(5)} - \frac{\mathrm{tr}(\hat{\xi}_t^{(5)}\hat{\Psi}_3)}{\mathrm{tr}(\hat{\Psi}_3\hat{\Psi}_3)}\hat{\Psi}_3 - \frac{\mathrm{tr}(\hat{\xi}_t^{(5)}\hat{\Psi}_1)}{\mathrm{tr}(\hat{\Psi}_1\hat{\Psi}_1)}\hat{\Psi}_1
\end{aligned} \tag{29}$$

$$\hat{\Psi}_7 = \hat{\xi}_t^{(7)} - \frac{\mathrm{tr}(\hat{\xi}_t^{(7)}\hat{\Psi}_5)}{\mathrm{tr}(\hat{\Psi}_5\hat{\Psi}_5)}\hat{\Psi}_5 - \frac{\mathrm{tr}(\hat{\xi}_t^{(7)}\hat{\Psi}_3)}{\mathrm{tr}(\hat{\Psi}_3\hat{\Psi}_3)}\hat{\Psi}_3 - \frac{\mathrm{tr}(\hat{\xi}_t^{(7)}\hat{\Psi}_1)}{\mathrm{tr}(\hat{\Psi}_1\hat{\Psi}_1)}\hat{\Psi}_1$$

$$\vdots$$

Our next step is to work out the norms $N_n = \mathrm{tr}(\hat{\Psi}_n\hat{\Psi}_n)$ appearing in (29). For $n = 1$ this is given in (5) by twice the Wigner–Yanase skew information of the energy. Let us examine the case $n = 3$. From the expression for $\hat{\Psi}_3$ in (29) and the orthogonality of the vectors $\{\hat{\Psi}_n\}$ we see at once that

$$\mathrm{tr}(\hat{\Psi}_3\hat{\Psi}_3) = \mathrm{tr}(\hat{\xi}_t^{(3)}\hat{\xi}_t^{(3)}) - \frac{[\mathrm{tr}(\hat{\xi}_t^{(3)}\hat{\xi}_t^{(1)})]^2}{\mathrm{tr}(\hat{\Psi}_1\hat{\Psi}_1)}. \tag{30}$$

More generally, on a closer inspection of (29) we deduce that the norm N_n can be constructed recursively from the norms of the lower-order vectors along with the inner products of the form $\mathrm{tr}(\hat{\xi}_t^{(m)}\hat{\xi}_t^{(n)})$ where $n + m$ is even. These inner products, however, are given by the higher-order skew moments.

With these observations at hand, let us now work out the expressions for the norms $N_n = \mathrm{tr}(\hat{\Psi}_n\hat{\Psi}_n)$, $n = 1, 3, 5, \ldots$, in terms of the skew moments. In fact, a short calculation at once reveals the recursive structure of these norms, for, we have

$$N_1 = S_2, \qquad N_3 = \frac{\begin{vmatrix} S_6 & S_4 \\ S_4 & S_2 \end{vmatrix}}{S_2}, \qquad N_5 = \frac{\begin{vmatrix} S_{10} & S_8 & S_6 \\ S_8 & S_6 & S_4 \\ S_6 & S_4 & S_2 \end{vmatrix}}{\begin{vmatrix} S_6 & S_4 \\ S_4 & S_2 \end{vmatrix}}, \tag{31}$$

and so on. The norms are therefore given by the ratios of the determinants of the matrices of even-order central moments. Let us denote these determinants by D_{2n}:

$$D_{2n} = \begin{vmatrix} S_{2n} & S_{2n-2} & \ldots & S_{n+1} \\ S_{2n-2} & S_{2n-4} & \ldots & S_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n+1} & S_{n-1} & \ldots & S_2 \end{vmatrix} \tag{32}$$

for $n = 1, 3, 5, \ldots$. Then with the convention that $D_{-2} = 1$, we have

$$N_n = \frac{D_{2n}}{D_{2n-4}} \tag{33}$$

for these odd-order norms. The overall structure of the normalisation $\{N_n\}$ thus is identical to the one obtained in [8] for pure states, except that the norms obtained here are different from the ones obtained in [8] because the skew moment S_n defined here does not reduce to the nth central moment μ_n in the pure-state limit, except for $n = 2$, as indicated above.

The structural similarity of the determinant in (32) to the one obtained in [8] involving central moments, however, implies that the geometric interpretation discussed in [8] is applicable in the present case. Namely, regarding the state $\hat{\xi}_t$ as a curve in Hilbert space, D_{2n} for each n can be expressed as the trace norm of the tensor obtained by a totally skew-symmetric product of the vectors $\hat{\xi}_t^{(1)}, \hat{\xi}_t^{(3)}, \ldots, \hat{\xi}_t^{(n-2)}, \hat{\xi}_t^{(n)}$. It follows that the norms $\{N_n\}$ can be interpreted as representing higher-order torsions of the curve in Hilbert space.

We recall that our objective is to determine the right side of (15) over odd n. Having identified the denominators in (15), let us turn our attention to the numerators. On account of the fact that $\hat{\xi}_t$ and $\hat{\Psi}_n$ are orthogonal, the term containing the parameter t drops out and the numerator reduces to the square of $\mathrm{tr}\left((\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t)\hat{\Psi}_n\right)$; but $\hat{\Psi}_n$ in turn is expressed as a linear combination of odd-order derivatives of $\hat{\xi}$. Before we proceed further, let us first establish the following:

Lemma. *The inner product* $\mathrm{tr}\left((\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t)\hat{\Psi}_n\right)$ *is independent of* $\hat{T}$ *for n odd, whereas for n even it is determined by the anticommutators of* $\hat{T}$ *and* $\hat{H}^p$ *for a range of p.*

Proof. Let us first consider the case where n is odd. Then $\hat{\Psi}_n$ is expressed as a linear combination of odd-order derivatives of the state $\hat{\xi}_t$, so let us examine more explicitly an expression of the form

$$\mathrm{tr}\left(\hat{\xi}_t^{(2m+1)}(\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t)\right)=(-1)^{n+1}\mathrm{i}\sum_{k=0}^{2m+1}(-1)^k\binom{2m+1}{k}\mathrm{tr}\left(\hat{H}^{2m+1-k}\hat{\xi}_t\hat{H}^k(\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t)\right), \tag{34}$$

where we have made use of (14). On a closer inspection we find that for $k=\alpha$ the right side is

$$(-1)^{m+1}\mathrm{i}(-1)^{\alpha}\frac{(2m+1)!}{\alpha!(2n+1-\alpha)!}\,\mathrm{tr}\left(\hat{T}\hat{H}^{2m+1-\alpha}\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t+\hat{H}^{\alpha}\hat{T}\hat{\xi}_t\hat{H}^{2m+1-\alpha}\hat{\xi}_t\right),$$

whereas for $k=2m+1-\alpha$ it gives

$$-(-1)^{m+1}\mathrm{i}(-1)^{\alpha}\frac{(2m+1)!}{\alpha!(2n+1-\alpha)!}\mathrm{tr}\left(\hat{H}^{2m+1-\alpha}\hat{T}\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t+\hat{T}\hat{H}^{\alpha}\hat{\xi}_t\hat{H}^{2m+1-\alpha}\hat{\xi}_t\right).$$

Adding these together we find

$$(-1)^{m+1-\alpha}\frac{(2m+1)!}{\alpha!(2n+1-\alpha)!}\left[\mathrm{tr}\left(\mathrm{i}[\hat{T},\hat{H}^{2m+1-\alpha}]\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t\right)+\mathrm{tr}\left(\mathrm{i}[\hat{H}^{\alpha},\hat{T}]\hat{\xi}_t\hat{H}^{2m+1-\alpha}\hat{\xi}_t\right)\right].$$

We therefore see that the dependence on the estimator $\hat{T}$ drops out on account of the commutation relation $\mathrm{i}[\hat{H}^p,\hat{T}]=p\hat{H}^{p-1}$, and we obtain

$$\mathrm{tr}\left(\hat{\xi}_t^{(2m+1)}(\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t)\right)=\sum_{\alpha=0}^{m}(-1)^{m+1-\alpha}\frac{(2m+1)!}{\alpha!(2m+1-\alpha)!}\times\left[\alpha\,\mathrm{tr}(\hat{H}^{\alpha-1}\hat{\xi}_t\hat{H}^{2m+1-\alpha}\hat{\xi}_t)-(2m+1-\alpha)\mathrm{tr}(\hat{H}^{2m-\alpha}\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t)\right]. \tag{35}$$

This establishes the claim that the dependence of the bound on the estimator $\hat{T}$ drops out from the odd-order terms on account of the commutation relation. When n is even, however, this argument shows that the two corresponding terms for $k=\alpha$ and $k=2m-\alpha$ add up to give anticommutators of the form $\{\hat{H}^p,\hat{T}\}$; while for $k=m$ the two terms resulting from $\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t$ also gives an anticommutator of $\hat{T}$ and $\hat{H}^m$. □

On a closer examination we see that the expression in the right side of (35) can in fact be simplified further. To see this we note that the first sum in (35) can be expressed alternatively in the form

$$\sum_{\alpha=0}^{m-1}(-1)^{m-\alpha}\frac{(2m+1)!}{\alpha!(2m-\alpha)!}\,\mathrm{tr}(\hat{H}^{\alpha}\hat{\xi}_t\hat{H}^{2m-\alpha}\hat{\xi}_t)$$

by shifting the summation variable $\alpha \to \alpha + 1$. On the other, the second sum in (35) can be expressed alternatively in the form

$$\sum_{\alpha=0}^{m-1}(-1)^{m-\alpha}\frac{(2m+1)!}{\alpha!(2m-\alpha)!}\,\mathrm{tr}(\hat{H}^{2m-\alpha}\hat{\xi}_t\hat{H}^{\alpha}\hat{\xi}_t)-\frac{(2m+1)!}{(m!)^2}\,\mathrm{tr}(\hat{H}^{m}\hat{\xi}_t\hat{H}^{m}\hat{\xi}_t).$$

Therefore, the two sums combine, and if we compare the result thus obtained with the expressions in (20) and (21) we deduce that

$$\mathrm{tr}\left(\hat{\xi}_t^{(2m+1)}(\hat{\xi}_t\hat{T}+\hat{T}\hat{\xi}_t)\right)=(-1)^{m+2}\,(2m+1)\,S_{2m}. \tag{36}$$

With these results at hand we are now in the position to examine the numerator terms in (15). Our approach will be to deduce a recursion relation for the numerator $U_n := \mathrm{tr}((\hat{\xi}\hat{T}+\hat{T}\hat{\xi})\hat{\Psi}_n)$, because $\hat{\Psi}_n$ is expressed in terms of a linear combination of $\hat{\xi}_t^{(n)}$ and $\hat{\Psi}_k$ with $k = n-2$, $n-4, \ldots$. To deduce a recursion relation we need to work out the coefficients

$$F_{n,k} := \frac{\mathrm{tr}(\hat{\xi}_t^{(n)}\hat{\Psi}_k)}{\mathrm{tr}(\hat{\Psi}_k\hat{\Psi}_k)}, \quad (k=1,3,5,\ldots,n-2) \tag{37}$$

of $\hat{\Psi}_k$ in $\hat{\Psi}_n$. In terms of these coefficients we therefore have

$$\hat{\Psi}_n=\hat{\xi}_t^{(n)}-\sum_{k=1,3,5,\ldots}^{n-2}F_{n,k}\hat{\Psi}_k. \tag{38}$$

A calculation analogous to the ones outlined above then shows that

$$F_{n,k}=\frac{(-1)^{\frac{1}{2}(n+k)-1}}{D_{2k}}\begin{vmatrix} S_{n+k} & S_{n+k-2} & \ldots & S_{n+1}\\ S_{2k-2} & S_{2k-4} & \ldots & S_{k-1}\\ \vdots & \vdots & \ddots & \vdots\\ S_{k+1} & S_{k-1} & \ldots & S_2\end{vmatrix}. \tag{39}$$

Therefore, it follows that the recursion relation satisfied by the numerator terms is given by

$$U_n=(-1)^{\frac{1}{2}(n-1)}\,nS_{n-1}-\sum_{k=1,3,5,\ldots}^{n-2}F_{n,k}U_k, \tag{40}$$

along with the initial condition $U_1 = 1$.

Putting these together, we finally deduce that the generalised uncertainty relation takes the form

$$\Delta\hat{T}^2+\delta\hat{T}^2 \geqslant \frac{1}{2}\sum_{n=1,3,5,\ldots}\frac{U_n^2}{N_n}, \tag{41}$$

where the denominator N_n is given by the ratio of matrix determinant (33) and the numerator U_n is obtained recursively by (40). In this way, we are able to work out as many higher-order corrections as we wish to the generalised uncertainty relation associated with generic mixed-state density matrices. For example, truncating the sum at $n = 3$ we deduce that

$$(\Delta\hat{T}^2 + \delta\hat{T}^2)(\Delta\hat{H}^2 - \delta\hat{H}^2) \geqslant \frac{1}{4}\left(1 + \frac{(S_4 - 3S_2^2)^2}{S_6 S_2 - S_4^2}\right). \tag{42}$$

The structure of the bound here is therefore identical to that of (2) originally obtained in [8], except that in the pure-state limit (42) does not reduce to (2), because while the left side becomes $\Delta\hat{T}^2\Delta\hat{H}^2$, the higher-order skew moments S_4 and S_6 appearing on the right side do not reduce to central moments. Because the method of obtaining the higher-order bounds presented here is different from that used in [8], albeit the concept of the approach is identical, it is not *a priori* possible to determine which approach gives rise to a sharper bound for the product of the variances in the pure-state limit. For instance, if a system in a given state has a gamma-distributed energy with shape parameter α and rate parameter β, then it is shown in [8] that the second-order correction gives

$$\Delta\hat{T}^2\Delta\hat{H}^2 \geqslant \frac{1}{4}\left(1 + \frac{18}{3\alpha^2 + 47\alpha + 42}\right), \tag{43}$$

which is experimentally significant (e.g. for an exponential distribution for which $\alpha = 1$ the first-order correction already gives about 20% sharper bound). On the other, the pure-state limit of (42) for this system is given by

$$\Delta\hat{T}^2\Delta\hat{H}^2 \geqslant \frac{1}{4}\left(1 + \frac{3}{2\alpha^2 + 9\alpha + 7}\right), \tag{44}$$

which is close to (43) but is a strictly weaker bound for all values of $\alpha > 0$ (for $\alpha = 1$ it provides about a 17% improvement). In general, by taking the difference of the two bounds we find that the pure-state limit of (42) is sharper than that obtained in [8] if

$$\begin{aligned} &9m_4(m_2 - m_1^2) + 4m_3m_1(6m_2 - m_1^2) + 12m_2m_1^4 - 10m_3^2 \\ &\quad - 9m_2^2(m_2 + m_1^2) - 4m_1^6 > 0, \end{aligned} \tag{45}$$

where $m_k = \langle\psi|\hat{H}|\psi\rangle$ denotes the kth moment of the energy.

6. Discussion

In summary, we have worked out the higher-order corrections, in the sense of orders of the moments, to the uncertainty lower bound associated with a generic mixed state density matrix using the Hilbert space method. We have shown that an appropriate measure of estimation uncertainty, in the case of mixed quantum states, should be the skew information of the second kind introduced here, as opposed to the conventional variance measure. This follows from the facts (a) that the Fisher information in the sense of Rao does not bound the variance but is a bound for the skew information of the second kind; and (b) that for a mixed stationary state the variance remains positive while the Fisher information vanishes, making a variance-based bound not entirely adequate when dealing with mixed states. The uncertainty lower bounds are expressed in terms of the skew moments of the first kind, which we have worked out explicitly. The notion of these skew moments of both kinds arise naturally in analysing mixed quantum states, and has no analogue when dealing with pure states.

We conclude by remarking that for a given observable $\hat{X}$ and a given state $\hat{\rho}$, the skew informations of both kinds $\Delta X^2 \pm \delta X^2$ can be worked out explicitly. Hence if experimental data for estimating the state of the system are available then they can be used to infer these information measures. However, from the viewpoint of estimation theory, the error measure is typically difficult to determine, whereas its lower bounds are easier to work out. The idea therefore is to determine the magnitude of uncertainty in an estimation problem for a given state, and this can be achieved by use of the inequalities derived here in terms of quantum skew moments. In particular, the lowest such moment corresponds to the Wigner–Yanase skew information; but here we have derived expressions for its higher-order analogues.

Acknowledgments

DCB acknowledges support from the Russian Science Foundation, Grant 20-11-20226. The authors thank anonymous referees for comments and D O Soares-Pinto for drawing our attention to [17] in which the problem addressed here has also been investigated and higher-order corrections to mixed-state uncertainty have been obtained independently using a similar method.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Dorje C Brody https://orcid.org/0000-0002-9242-9554

References

[1] Kobe D H and Aguilera-Navarro V C 1994 Derivation of the energy–time uncertainty relation *Phys. Rev.* A **50** 933–8
[2] Bush P 2007 The time-energy uncertainty relation *Time in Quantum Mechanics* ed G Muga (Berlin: Springer)
[3] Holevo A 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[4] Rao C R 1945 Information and the accuracy attainable in the estimation of statistical parameters *Bull. Calcutta Math. Soc.* **37** 81–91
[5] Fisher R A 1925 Theory of statistical estimation *Math. Proc. Camb. Phil. Soc.* **22** 700–25
[6] Brody D C and Hughston L P 1996 Geometry of quantum statistical inference *Phys. Rev. Lett.* **77** 2851–4
[7] Brody D C and Hughston L P 1998 Statistical geometry in quantum mechanics *Proc. R. Soc.* A **454** 2445–75
[8] Brody D C and Hughston L P 1997 Generalised Heisenberg relations for quantum statistical estimation *Phys. Lett.* A **236** 257–62
[9] Wigner E P and Yanase M M 1963 Information contents of distributions *Proc. Natl Acad. Sci.* **49** 910–8
[10] Luo S 2003 Wigner–Yanase skew information and uncertainty relations *Phys. Rev. Lett.* **91** 180403
[11] Luo S 2005 Heisenberg uncertainty relation for mixed states *Phys. Rev.* A **72** 042110
[12] Brody D C 2011 Information geometry of density matrices and state estimation *J. Phys. A: Math. Theor.* **44** 252002

[13] Luo S and Zhang Q 2004 On skew information *IEEE Trans. Inf. Theor.* **50** 1778–82
[14] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[15] Braunstein S L and Caves C M 1994 Statistical distance and the geometry of quantum states *Phys. Rev. Lett.* **72** 3439–43
[16] Efron B 1975 Defining curvature on statistical model *Ann. Stat.* **3** 1189–242
[17] Magno G F, Grossi C H, Adesso G and Soares-Pinto D O 2020 Information geometric approach to mixed state quantum estimation (arXiv:2012.03080)