Upper bounds to the Holevo Cramér-Rao bound for multiparameter quantum metrology

Francesco Albarelli†,‡ and Animesh Datta†,‡

1Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

(Dated: November 26, 2019)

The fundamental attainable scalar bound on the mean square error matrix of an estimator. The scalar Helstrom CRB is defined as the weighted trace of the inverse of the quantum Fisher information matrix (QFIM). However, its evaluation requires solving a Lyapunov equation corresponding to a density matrix and inverting the QFIM, neither of which are known to be possible analytically. Furthermore, the attainability of such a bound is not guaranteed due to the non-commutativity of observables in quantum mechanics. Tighter attainable bounds have therefore been sought and identified since [21–24].

The fundamental attainable scalar bound on the mean square error matrix is the so-called Holevo CRB [23–25]. Evaluating the Holevo CRB requires an optimisation over the set of Hermitian operators which is not known to be possible analytically except in a few non-trivial cases [23, 26–29], leaving it largely unscrutinised. It was recently shown that the evaluation of the Holevo CRB is a convex optimisation problem solvable using a semidefinite program [30].

In this Letter we present two results—one each for scalar and matrix bounds. The former is that the Holevo CRB is never greater that twice the Helstrom CRB. Indeed, we point out a tighter intermediate bound. The latter suggests that a mean square error matrix equaling one-half the QFIM may be attainable asymptotically for arbitrary quantum statistical models.

Our results show that the Holevo CRB cannot provide new information about possible quantum enhancements in scaling in multiparameter estimation that is not already available from the Helstrom CRB. However for judging a given quantum state’s performance in applications such as simultaneous phase and loss estimation in optical interferometry, the Helstrom CRB or even the QFIM is inadequate and the Holevo CRB is necessary [30]. More fundamentally, the Holevo CRB should provide a deeper quantitative understanding of collective quantum measurements which the Helstrom CRB cannot [31]. Since both the Holevo and the Helstrom CRBs are typically evaluated numerically, the supposed ease of computing the latter should not endow it with exaggerated significance, especially since the former can be obtained from a semidefinite program.

Note added. While completing this work we became aware of an independent alternative derivation of inequality (11) by Carollo et al. [32] and of a looser upper bound recently and independently derived by Tsang [33]. We have also found a small gap in the proof of [32], which we fix Appendix B.

II. QUANTUM ESTIMATION THEORY

A quantum statistical model is a set of density operators \( \{ \rho_\theta \} \) labeled by a \( p \)-dimensional vector of real parameters \( \theta \in \Theta \subset \mathbb{R}^p \) that we want to estimate. Given a POVM\(^1\) \( \Pi = \{ \Pi_\omega | \Pi_\omega \geq 0, \sum_{\omega \in \Omega} \Pi_\omega = 1 \} \), where \( \Omega \) is the outcome space. The \( \theta \)-dependent probability distribution is given by the Born rule \( p_\theta(\omega) = \text{Tr}[\rho_\theta \Pi_\omega] \). We consider an unbiased estimator \( \hat{\theta} \) and we quantify the accuracy of the estimate by means of the mean square error matrix (for this class of estimators it coincides with

---

\(^1\) The notation \( A \succeq B \) means that \( A - B \) is a positive semidefinite operator (matrix in finite dimension).
the covariance matrix:
\[
\Sigma_{\mu
u}(\rho_\theta, \Pi, \hat{\theta}) = \sum_{\omega \in \Omega} \rho_\theta(\omega) \left( \hat{\theta}_\mu(\omega) - \theta_\mu \right) \left( \hat{\theta}_\nu(\omega) - \theta_\nu \right).
\]  
(1)

The classical CRB is a lower matrix bound for the mean square error matrix:
\[
\Sigma(\rho_\theta, \Pi, \hat{\theta}) \geq F(\rho_\theta, \Pi)^{-1},
\]  
(2)

that does not depend on the particular estimator; the quantity on the r.h.s. is the classical Fisher information matrix (FIM)
\[
F(\rho_\theta, \Pi)_{\mu\nu} = \sum_{\omega \in \Omega} \rho_\theta(\omega) \frac{\partial \log \rho_\theta(\omega)}{\partial \theta_\mu} \frac{\partial \log \rho_\theta(\omega)}{\partial \theta_\nu}.
\]  
(3)

In the quantum domain one can introduce matrix bounds that depend only on the quantum statistical model and not on the particular POVM. There is no unique way to do so, but the most common choice is the QFIM introduced by Helstrom:
\[
J^S_{\mu\nu} = \frac{1}{2} \text{Tr} [\rho_\theta (L^S_\mu L^S_\nu + L^S_\nu L^S_\mu)],
\]  
(4)

written in terms of the symmetric logarithmic derivatives (SLDs), defined by the equation \(2 \frac{\partial \rho_\theta}{\partial \theta_\mu} = L^S_\mu \rho_\theta + \rho_\theta L^S_\mu\). A well-known alternative is the matrix \(J^{R}_{\mu\nu} = \text{Tr}[\rho_\theta L^R_\mu L^R_\nu]\), obtained from right logarithmic derivatives (RLDs) that satisfy the equation \(\frac{\partial \rho_\theta}{\partial \theta_\mu} = \rho_\theta L^R_\mu\). These matrices give matrix upper bounds to the classical FIM as follows \(F(\rho_\theta, \Pi) \leq J^{S,R}(\mu, \Pi)\), the corresponding matrix Helstrom CRB is
\[
\Sigma(\rho_\theta, \Pi, \hat{\theta}) \geq (J^S)^{-1}.
\]  
(5)

In multiparameter estimation it is necessary to consider scalar figures of merit and an usual choice is the trace of the covariance matrix, weighted by some positive \(p \times p\) matrix \(W \succ 0\), for which we have the following inequalities\(^2\)
\[
\text{tr}[W \Sigma] \geq \text{tr}[WF^{-1}] \geq C^H(W) \geq \max[C^S(W), C^R(W)]
\]  
(6)

where \(C^S(W) = \text{tr}[W (J^S)^{-1}]\) is the scalar Helstrom CRB and \(C^R(W) = \text{tr}[W \text{Re}(J^R)^{-1}] + \|\sqrt{W} \text{Im}(J^R)^{-1} \sqrt{W}\|_1\) is the scalar RLD-CRB.\(^4\), while \(C^H(W)\) is the Holevo CRB, defined as
\[
C^H(W) = \min_{X \in X_\theta} \left\{ \text{tr}[W \text{Re} Z(X)] + \|\sqrt{W} \text{Im} Z(X) \sqrt{W}\|_1 \right\},
\]  
(7)

where \(Z[X]_{\mu\nu} = \text{Tr}[\rho_\theta X_\mu X_\nu]\) and \(X\) is a vector of \(p\) Hermitian operators belonging to the set
\[
X_\theta = \left\{ X_i \in L(H) | \text{Tr}[\frac{\partial \rho_\theta}{\partial \theta_\mu} X_\nu] = \delta_{\mu\nu}, \text{Tr}[\rho_\theta X_\mu] = 0 \right\}.
\]  
(8)

We also introduce the incompatibility matrix \(D\), also known as the mean Uhlmann curvature, defined as
\[
D_{\mu\nu} = \frac{1}{2i} \text{Tr}[\rho_\theta (L^S_\mu L^S_\nu - L^S_\nu L^S_\mu)].
\]  
(9)

The Holevo CRB and the scalar Helstrom CRB coincide if and only if all the elements of this matrix are zero [36]. The Holevo CRB can also be equal to the RLD bound when the quantum statistical model satisfies a condition called D-invariance [26, 29].

III. UPPER BOUNDS TO THE HOLEVO CRB

Our first main result is
\[
C^H(W) \leq C^S(W) + \|\sqrt{W} (J^S)^{-1} D (J^S)^{-1} \sqrt{W}\|_1 \leq 2C^S(W).
\]  
(10)

(11)

The first upper bound (10) can be obtained by restricting the optimization (7) to the real span of the SLDs. The quantity on the r.h.s. of (10) is attainable and equal to \(C^H(W)\) when the model is D-invariant [26]: when this is not the case, the minimization is carried out on a smaller set and therefore the value of the minimum is greater or equal to the Holevo CRB. This inequality was already presented in [26] and in Appendix A 1 we provide an alternative derivation.

Two observations are in order. Firstly, this is an upper bound more informative than \(2C^S\), yet obtainable only from the SLDs. Secondly, this upper bound can be a loose restriction on the difference \(C^H - C^S\), which can be small and cannot be estimated from the SLDs alone without evaluating \(C^H\). The problem of noisy 3D magnetometry with multi-qubit systems offers such an illustration [30].

The inequality (11) follows from
\[
\|\sqrt{W} (J^S)^{-1} D (J^S)^{-1} \sqrt{W}\|_1 \leq C^S(W),
\]  
(12)

which is obtained by applying a lemma from Holevo [24] (originally due to Belavkin and Grishanin [37]), as we show explicitly in Appendix A 2.

We stress that these inequalities hold for arbitrary quantum systems, irregardless of their dimension, since the derivation only relies on algebraic manipulations of the quantity that enters in the evaluation of the Holevo CRB\(^5\).

\(^2\) In particular, we will always restrict to regular quantum statistical models for which the parametrization is sufficiently smooth, the QFIM is non-singular \(J^S \succ 0\) and the rank of \(\rho_\theta\) is fixed, to avoid discontinuities and related problems with CRBs [34, 35].

\(^3\) We suppress the dependencies of \(\Sigma\) and \(F\) for brevity.

\(^4\) The trace norm is defined as \(\|A\|_1 = \text{tr} \left[ \sqrt{A^* A} \right]\).

\(^5\) This is the reason why traces in Hilbert space are denoted as \(\text{Tr}\), while traces of matrices are denoted as \(\text{tr}\): when the Hilbert space is infinite dimensional they are different operations.
The only assumption is to have a regular, non-singular \((J^S > 0)\) quantum statistical model. Furthermore, these bounds hold for any \(W \geq 0\), including rank-deficient ones which are relevant when dealing with nuisance parameters [11, 17, 38, 39].

Quantum tomography of pure states is an estimation problem for which our bounds are saturated, that is, \(C^S = C^H = 2C^S\). Li et al. [40] have shown that for the estimation of all the \(2d - 2\) parameters of a pure state there exists a POVM having a FIM proportional to the QFIM and also saturating the Gill-Massar inequality [41]. In particular these POVs satisfy \(F = \frac{1}{2}J^S\), regardless of \(d\). For pure states, a POVM that saturates the Gill-Massar inequality also attains the Holevo CRB [42] and implies our claim that \(C^H = 2C^S\) for this model.

**IV. ATTAINING ONE-HALF OF THE QFIM IN GAUSSIAN SHIFT MODELS**

A Gaussian state \(\rho^G_\theta\) is the thermal state (ground state if pure) of a Hamiltonian quadratic in the canonical operators \(r = (x_1, p_1, \ldots, x_k, p_k)\) for \(k\) bosonic modes, satisfying \([r_i, r_j] = i\Omega_{ij}\) with \(\Omega = i \oplus_{i=1}^k \sigma_y\). The parameter-dependent first moments \(\bar{r}_\theta\) of the Gaussian state are defined as \((\bar{r}_\theta)_i = \text{Tr}[\rho_G r_i]\), while the (parameter-independent) covariance matrix (CM) is \(\sigma_{ij} = \text{Tr}[\rho_G \{r_i - (\bar{r}_\theta)_i\}, (r_j - (\bar{r}_\theta)_j)]\), where \(\{A, B\} = AB + BA\). A generic Gaussian measurement is defined by a physical CM \(\sigma_m \geq i\Omega\) and a vector of \(k\) outcomes \(r_{\text{out}}\), physically it is implemented by noisy general-dyne detection. The corresponding probability distribution reads [43, 44]

\[
p(r_{\text{out}}|\theta) = \frac{\exp\left[-(r_{\text{out}} - \bar{r}_\theta)^T(\sigma + \sigma_m)^{-1}(r_{\text{out}} - \bar{r}_\theta)\right]}{\pi^k \sqrt{\det(\sigma + \sigma_m)}}
\]

and the associated classical FIM is

\[
F(\bar{r}_\theta, \sigma, \sigma_m) = 2(\partial_\theta \bar{r}_\theta)^T (\sigma + \sigma_m)^{-1} (\partial_\theta \bar{r}_\theta).
\]

where we have introduced the \(2k \times p\) Jacobian matrix with elements \((\partial_\theta \bar{r}_\theta)_{ij} = \partial(\bar{r}_\theta)_i/\partial \theta_j\). The QFIM for a Gaussian shift model has exactly the same form as the FIM for a classical Gaussian distribution [43, 45]

\[
J^S(\bar{r}_\theta, \sigma) = 2(\partial_\theta \bar{r}_\theta)^T \sigma^{-1} (\partial_\theta \bar{r}_\theta).
\]

Considering the measurement with the same CM as the state, that is \(\sigma + \sigma_m = 2\sigma\), we get to our second main result

\[
F(\bar{r}_\theta, \sigma, \sigma) = \frac{1}{2}J^S(\bar{r}_\theta, \sigma).
\]

Thus, the classical scalar CRB is indeed twice the scalar Helstrom CRB \(\text{tr}[W F(\bar{r}_\theta, \sigma, \sigma)] = 2C^S\). This is an alternative proof that \(C^H \leq 2C^S\) for Gaussian shift models (for which the Holevo CRB itself is attainable with single-copy measurements). This halving of the available information manifests the additional noise of measuring complementary observables simultaneously as suggested by Arthurs and Kelly [46–48].

**V. DISCUSSION**

The theory of quantum local asymptotic normality (QLAN) [49–55] maps an asymptotically large number of identical copies of a finite-dimensional quantum statistical model to a Gaussian shift model\(^6\). Since for such models there always exists a POVM with a FIM equal to one-half of the QFIM as shown by (16), we have \(C^H \leq 2C^S\), consistently with (11).

This strongly suggests that for any quantum statistical model there will be a sequence of collective POVs \(\Pi_n\) and classical estimators \(\tilde{\theta}_n\) such that

\[
\lim_{n \to \infty} n\Sigma(\rho^G_{\tilde{\theta}_n}, \Pi_n, \tilde{\theta}_n) = 2(J^S)^{-1}
\]

However, due to the technical assumptions behind QLAN, our suggestion does not constitute a rigorous proof, for which one would need to show explicitly the sequence of POVs and estimators. We stress that, on the contrary, the derivation of the scalar inequality \(C^H \leq 2C^S\) is based purely on the 

\[
\text{evaluation}
\]

of the Holevo CRB, which is a well-defined “single letter” calculation, regardless of asymptotics.

**ACKNOWLEDGMENTS**

We thankfully acknowledge M. Tsang for a fruitful discussion that led to the start of this work. We thank M. G. Genoni both for several fruitful discussions and for making us aware of Ref. [32]. We thank D. Branford, J. F. Friel and J. Suzuki for useful discussions. This work has been supported by the UK EPSRC (EP/K04057X/2) and the UK National Quantum Technologies Programme (EP/M01326X/1, EP/M013243/1).

**Appendix A: Algebraic derivation of the upper bound**

1. **First inequality**

Here we explicitly show how to derive the inequality (10), in doing so we exploit an alternative expression of the Holevo CRB due to Suzuki [29], given in (A12).

\(^6\) This has been used to show the asymptotic attainability of the Holevo CRB, although the precise technical assumptions differ in the various formulations.
We start by considering the Gram matrix of the SLDs with respect to the so-called RLD inner product, i.e. defined by the matrix elements

$$Z(L^S)_{\mu\nu} = \text{Tr}[\rho_0 L^S_\mu L^S_\nu], \quad (A1)$$

where $L^S = (L^S_1, L^S_2, \ldots, L^S_p)$ is the vector of SLD operators. Being a Gram matrix, it is positive semi-definite:

$$Z(L^S) = J^S + iD \succeq 0, \quad (A2)$$

where we have used the fact that its real and imaginary part are respectively the QFIM $J^S = \text{Re} Z(L^S)$, cf. Eq. (4), and the incompatibility matrix $D = \text{Im} Z(L^S)$, cf. Eq. (9). As usual, we assume a well defined quantum statistical model ($J^S > 0$) such that $J^S$ is invertible.

Following Suzuki [29] we can explicitly implement the linear constraints in the optimization appearing in (7) for evaluating the Holevo CRB:

$$X_\mu = \sum_{\nu} (J^S)_{\mu\nu}^{-1} L^S_\nu + K_\mu = L_\mu + K_\mu \quad (A3)$$

where the operators in the vector $K = (K_1, \ldots, K_p)$ are perpendicular to the SLDs with respect to the SLD inner product

$$\text{Tr}[\rho_0 \{ L^S_\mu, K_\nu \}] = 0 \quad \forall \mu, \nu \quad (A4)$$

and also perpendicular to the identity

$$\frac{1}{2} \text{Tr}[\rho_0 \{ 1, K_\mu \}] = \text{Tr}[\rho_0 K_\mu] = 0 \quad \forall \mu. \quad (A5)$$

In (A3) we have also introduced the dual SLD operators $L_\mu = \sum_j (J^S)^{-1}_{\mu j} L^S_j$, which are linear combinations of SLDs with real coefficients, so by linearity they also satisfy $\text{Tr}[\{ K_\mu, L_\nu \} \rho_0] = 0 \quad \forall \mu, \nu$.

It is easy to see that with this choice the constraint are satisfied:

$$\text{Tr}[X_\mu \partial_\mu \rho] = \frac{1}{2} \text{Tr}[\rho_0 \{ X_\mu, L_\nu \}] = \sum_\gamma (J^S)^{-1}_{\mu\gamma} \frac{1}{2} \text{Tr}[\rho_0 \{ L^S_\gamma, L^S_\nu \}] + \frac{1}{2} \text{Tr}[\rho_0 \{ K_\mu, L^S_\nu \}] = \sum_\gamma (J^S)^{-1}_{\mu\gamma} J^S_{\gamma\nu} + 0 = \delta_{\mu\nu}, \quad (A6)$$

where we used the definition of the SLD operators and the definition of the QFIM (4), furthermore we also have

$$\text{Tr}[\rho_0 X_\mu] = 0 \quad \forall \mu, \quad (A7)$$

since $\text{Tr}[\rho_0 L^S_\mu] = 0 \forall \mu$.

We can decompose the Gram matrix of the operators $X$ into the quadratic contributions in $L$ and $K$ plus the cross terms$^9$

$$Z(X) = Z(L) + Z(K) + Y(L, K) + Y(K, L), \quad (A8)$$

where the two matrices containing the scalar products between elements in the two sets are defined as $Y(K, L)_{\mu\nu} = \text{Tr}[\rho K_\mu L_\nu]$ and $Y(L, K)_{\mu\nu} = \text{Tr}[\rho L_\mu K_\nu]$. They are both purely imaginary $\text{Re} Y(L, K) = \text{Re} Y(K, L) = 0$, because of the orthogonality with respect to the SLD scalar product, and they satisfy the relationship

$$\text{Im} Y(L, K) = -\text{Im} Y(K, L)^T, \quad (A9)$$

so that the imaginary part of their sum $\tilde{Y}(L, K) = Y(L, K) + Y(K, L)$ is a skew-symmetric matrix:

$$\text{Im} \tilde{Y}(L, K) = \text{Im} Y(L, K) + \text{Im} Y(K, L) = -\text{Im} Y(K, L)^T - \text{Im} Y(L, K)^T \quad (A10)$$

$$= -\text{Im} \tilde{Y}(L, K)^T.$$

With a little algebra the function to minimize in (7) becomes

$$\text{tr}[W \text{Re} Z(X)] + \left\| \sqrt{W} \text{Im} Z(X) \sqrt{W} \right\|_1 = \text{tr}[W (J^S)^{-1}] + \text{tr}[W \text{Re} Z(K)]$$

$$\text{tr}[W \text{Re} Z(K)] + \left\| \sqrt{W} (J^S)^{-1} D (J^S)^{-1} + \text{Im} \tilde{Y} (L, K) + \text{Im} Z(K) \right\| \sqrt{W} \right\|_1.$$

With this substitution the Holevo CRB is found by minimizing over the operators $K$:

$$C^H = C^S + \max_K \left\{ \text{tr}[W \text{Re} Z(K)] \right\}$$

$$\left\| \sqrt{W} (J^S)^{-1} D (J^S)^{-1} + \text{Im} \tilde{Y} (L, K) + \text{Im} Z(K) \sqrt{W} \right\|_1 \right\}, \quad (A12)$$

without further constraints other than orthogonality to the SLDs (A4) and to the identity (A5). Using the triangular inequality of the trace norm we can upper bound the function to be minimized

$$\text{tr}[W \text{Re} Z(K)] + \left\| \sqrt{W} (J^S)^{-1} D (J^S)^{-1} + \text{Im} \tilde{Y} (L, K) + \text{Im} Z(K) \sqrt{W} \right\|_1$$

$$\leq \text{tr}[W \text{Re} Z(K)] + \left\| \sqrt{W} (J^S)^{-1} D (J^S)^{-1} \sqrt{W} \right\|_1 + \left\| \sqrt{W} \text{Im} \tilde{Y} (L, K) \sqrt{W} \right\|_1 + \left\| \sqrt{W} \text{Im} Z(K) \sqrt{W} \right\|_1.$$
Therefore
\[
C^H \leq C^S + \left\| \sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W} \right\|_1 + \min_K \left\{ \text{tr}[W\text{Re}Z(K)] + \left\| \sqrt{W}\text{Im}\tilde{Y}(\mathcal{L},K)\sqrt{W} \right\|_1 \right\} + \left\| \sqrt{W}\text{Im}Z(K)\sqrt{W} \right\|_1,
\]
but since the function to minimize on the r.h.s. is the sum of three positive terms, the minimum is 0 and it is achieved when the \(K_\mu = 0\) \(\forall \mu\). Therefore we have found that
\[
C^H \leq C^S + \left\| \sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W} \right\|_1. \tag{A15}
\]

2. Second inequality

Here we show how to derive inequality (12). This is just an application of a lemma by Holevo [24] (Lemma 6.6.1, p. 244), originally due to Belavkin and Grishanin [37]. We reproduce here a step by step derivation for the reader’s convenience; see also a similar derivation in the appendix of Ref. [13]. Notice also that the same lemma was applied by Nagaoka [25] to obtain the Holevo CRB in the form (7), which is an alternative formulation of Holevo’s original result [23, 24].

From (A2) we have that \(J^S \geq iD\), which is equivalent to \(J^S \geq -iD\), since the eigenvalues of an Hermitian matrix are the same as those of its transpose. From these inequalities we derive
\[
\sqrt{W}(J^S)^{-1}\sqrt{W} \geq \pm i\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}, \tag{A16}
\]
by multiplying from left and right by \((J^S)^{-1}\) and then by \(\sqrt{W}\). The matrix \(\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}\) is still skew-symmetric (and thus it is Hermitian if multiplied by \(\pm i\)). Now from the inequality (A16) it follows that
\[
v^\dagger\left[\sqrt{W}(J^S)^{-1}\sqrt{W}\right]v \geq \pm iv^\dagger\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}v, \tag{A17}
\]
for any arbitrary complex valued vector \(v\). We remark that the quantity \(v^\dagger\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}v\) must be purely imaginary due to the skew-symmetry of the matrix. Therefore \(v^\dagger\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}v = \pm iv^\dagger\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}v\), depending on the sign.

Putting these inequalities together we get
\[
v^\dagger\left[\sqrt{W}(J^S)^{-1}\sqrt{W}\right]v \geq \left|v^\dagger\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}v\right|, \tag{A18}
\]
In particular, we can sum all the scalar inequalities obtained by choosing the eigenvectors of the matrix \(\sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W}\) to obtain
\[
\text{Tr}\left[\sqrt{W}(J^S)^{-1}\sqrt{W}\right] = \text{Tr}\left[W(J^S)^{-1}\right] = C^S \geq \left\| \sqrt{W}(J^S)^{-1}D(J^S)^{-1}\sqrt{W} \right\|_1, \tag{A19}
\]
where we have exploited the fact that for a normal matrix the trace norm is the sum of the absolute values of the eigenvalues (and a skew-symmetric matrix is normal).

Appendix B: Added note on a related work

While completing this work we became aware of the recently published work of Carollo et. al [32], where the authors also derive the upper bound (11) using a different intermediate bound. While the final result is correct, the paper [32] contains a small mistake that makes the proof incomplete for generic weight matrices \(W\); here we show how to fix this.

The derivation of [32] starts from a wrong expression of the second term appearing in (7): \(\|W\text{Im}Z(X)\|_1 \) instead of \(\|\sqrt{W}\text{Im}Z(X)\sqrt{W}\|_1^{10}\). However, the derived upper bound is still valid, since for a skew-symmetric \(A\) we have the following inequality
\[
\left\| \sqrt{WA\sqrt{W}} \right\|_1 \leq \|WA\|_1. \tag{B1}
\]
This comes from the fact that the matrices \(M_1 = WA\) and \(M_2 = \sqrt{WA\sqrt{W}}\) are isospectral, but only \(M_2\) is normal (being skew-symmetric). For a normal matrix the singular values are equal to the absolute values of the eigenvalues. However, for a non-normal matrix such as \(M_1\) the ordered vector containing the absolute values of the eigenvalues is majorized by the ordered vector of singular values [59]; in turn, this fact implies (B1).

[1] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Phys. Rev. Lett. 111, 070403 (2013).

[2] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Koltzhammer, M. S. Kim, A. Datta, M. Barbieri, and I. A. Walmsley, Nat. Commun. 5, 3532 (2014).

[3] T. Baumgratz and A. Datta, Phys. Rev. Lett. 116, 030801 (2016).

[4] M. Szczykulska, T. Baumgratz, and A. Datta, Adv. Phys. X 1, 621 (2016).

[5] This misunderstanding appeared also in some previous works [57, 58], but since the matrix \(W = 1\) is then considered the results are valid (it is enough that \(W\) commutes with \(A\)).
[5] C. N. Gagatsos, D. Branford, and A. Datta, Phys. Rev. A 94, 042342 (2016).
[6] A. Chrostowski, R. Demkowicz-Dobrzański, M. Jarzyna, and K. Banaszek, Int. J. Quantum Inf. 15, 1740005 (2017).
[7] L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, Phys. Rev. Lett. 119, 130504 (2017).
[8] E. Roccia, I. Gianani, L. Mancino, M. Sbroscia, F. Somma, M. G. Genoni, and M. Barbieri, Quantum Sci. Technol. 3, 01LT01 (2018).
[9] E. Roccia, V. Cimini, M. Sbroscia, I. Gianani, L. Ruggiero, L. Mancino, M. G. Genoni, M. A. Ricci, and M. Barbieri, Optica 5, 1171 (2018).
[10] M. Gessner, L. Pezzè, and A. Smerzi, Phys. Rev. Lett. 121, 130503 (2018).
[11] J. Yang, S. Pang, Y. Zhou, and A. N. Jordan, Phys. Rev. A 100, 032104 (2019).
[12] M. G. Genoni and T. Tufarelli, J. Phys. A Math. Theor. 52, 434002 (2019).
[13] W. Gorecki, S. Zhou, L. Jiang, and R. Demkowicz-Dobrzański, arXiv:1901.00896 (2019).
[14] J. Rubio and J. Dunningham, arXiv:1906.04123 (2019).
[15] E. Bisketzi, D. Branford, and A. Datta, New J. Phys. in press, (2019).
[16] J. Liu, H. Yuan, X.-m. Lu, and X. Wang, arXiv:1907.08037 (2019).
[17] M. Tsang, IEEE Trans. Inf. Theory 19, 740 (1973).
[18] V. P. Belavkin, Theor. Math. Phys. 26, 213 (1976).
[19] A. S. Holevo, in Proc. Third Japan — USSR symp. Probab. Theory, Lecture Notes in Mathematics, Vol. 550, edited by G. Maruyama and J. V. Prokhorov (Springer, Berlin, Heidelberg, 1976).
[20] A. S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory, 2nd ed. (Edizioni della Normale, Pisa, 2011).
[21] H. Nagaoka, IEICE Tech. Rep. IT 89-42, 9 (1989).
[22] J. Suzuki, J. Math. Phys. 57, 042201 (2016).
[23] M. Bradshaw, S. M. Assad, and P. K. Lam, Phys. Lett. A 381, 2598 (2017).
[24] M. Bradshaw, P. K. Lam, and S. M. Assad, Phys. Rev. A 97, 012106 (2018).
[25] J. Suzuki, Entropy 21, 703 (2019).