Supplementary information : Efficient computation of the Nagaoka–Hayashi bound for multi-parameter estimation with separable measurements

Lorcán O. Conlon,[1,*] Jun Suzuki,[2,†] Ping Koy Lam,[1,3] and Syed M. Assad[1,3,‡]

[1] Centre for Quantum Computation and Communication Technology, Department of Quantum Science, Australian National University, Canberra, ACT 2601, Australia.

[2] Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo 182-8585, Japan

[3] School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639673, Republic of Singapore

Supplementary Note 1 : Summary of Hayashi's results from Ref. [1]

We summarise Hayashi's result [1] which was published in the proceedings of a domestic workshop in the Research Institute for Mathematical Sciences (RIMS) at Kyoto University in Japanese for the reader's convenience. Let $\mathcal{H}_q$ be a finite d-dimensional Hilbert space and consider a set of observables (Hermitian matrices) $X = (X_1, X_2, ..., X_n)^\intercal$ on it. We say that a POVM $\Pi = \{\Pi_m\}$ is a *simultaneous measurement* of the given observables X, if

$$X_j = \sum_m \hat{x}_{jm} \Pi_m \ , \tag{1}$$

holds for all j. In general, a projection measurement does not exist unless the X_j commute with each other, but a POVM Π exists. Given a state S on $\mathcal{H}_q$, we define the expectation value of X_j by

$$x_j := \mathrm{Tr}[S X_j] = \sum_m \hat{x}_{jm} \mathrm{Tr}[S \Pi_m] \ . \tag{2}$$

We define the covariance matrix by

$$\left[\tilde{\mathsf{U}}(\Pi, \hat{x})\right]_{jk} = \sum_m (\hat{x}_{jm} - x_j)(\hat{x}_{km} - x_k) \mathrm{Tr}[S\Pi_m] \tag{3}$$

$$= \sum_m \hat{x}_{jm} \hat{x}_{km} \mathrm{Tr}[S\Pi_m] - x_j x_k \tag{4}$$

$$= [\mathsf{U}(\Pi, \hat{x})]_{jk} - x_j x_k \ . \tag{5}$$

We are interested in minimizing the sum of the diagonal elements of $\tilde{\mathsf{U}}(\Pi, \hat{x})$. As the second term is constant this is equivalent to minimising $\mathsf{Tr}[\mathsf{U}]$. Indeed the second term can be ignored for all practical purposes. We define the precision limit as

$$C = \inf_{\Pi} \left\{ \mathsf{Tr}[\mathsf{U}] - \sum_j x_j^2 \,\middle|\, \Pi : \text{ simultaneous measurement of } X \right\} \ . \tag{6}$$

Note here that C depends on the given state S and the set of observables X. Hayashi derived the following two bounds for C.

Theorem 1 (Hayashi)**.** *The following are lower bounds for C and further that $C \geq C_1 \geq C_2$ holds.*

$$C_1 = \inf_{\mathbb{U}} \left\{ \mathbb{T}\mathrm{r}[\mathbb{U}] - \sum_j x_j^2 \,\middle|\, \mathbb{U}_{jk} = \mathbb{U}_{kj} \text{ Hermitian, } \mathbb{U} \geq \sqrt{\mathbb{S}} X X^\intercal \sqrt{\mathbb{S}} \right\}, \tag{7}$$

$$C_2 = \inf_{\mathsf{U}} \left\{ \mathsf{Tr}[\mathsf{U}] - \sum_j x_j^2 \,\middle|\, \mathsf{U} \text{ Hermitian, } \mathsf{U} \geq \mathrm{Tr}[\sqrt{\mathbb{S}} X X^\intercal \sqrt{\mathbb{S}}] \right\}, \tag{8}$$

---

* lorcan.conlon@anu.edu.au

† junsuzuki@uec.ac.jp

‡ cqtsma@gmail.com

where $\mathbb{S} = 1 \otimes S$ *and* $\mathbb{U}$ *are complex matrices on the extended Hilbert space* $\mathcal{H}_c \otimes \mathcal{H}_q$.

Hayashi's first bound C_1 is considered as the generalisation of the Nagaoka bound for simultaneous measurement of non-commuting observables [2]. Using the linear programming approach, Hayashi further derived the following alternative forms for C_1 and C_2

$$C_1 = \mathbb{T}\mathrm{r}\Big[\mathrm{Sym}_+\left(\sqrt{\mathbb{S}}XX^\intercal\sqrt{\mathbb{S}}\right)\Big] + \inf_{\mathbb{V}}\left\{\mathbb{T}\mathrm{r}[\mathbb{V}] \,\Big|\, \mathbb{V} \geq 0,\, \mathrm{Sym}_-(\mathbb{V}) = -\mathrm{Sym}_-\left(\sqrt{\mathbb{S}}XX^\intercal\sqrt{\mathbb{S}}\right)\right\} - \sum_j x_j^2, \tag{9}$$

$$C_2 = \mathbb{T}\mathrm{r}\Big[\mathrm{Sym}_+\left(\sqrt{\mathbb{S}}XX^\intercal\sqrt{\mathbb{S}}\right)\Big] + \mathsf{TrAbs}\Big[\mathrm{Tr}\Big[\mathrm{Sym}_-\left(\sqrt{\mathbb{S}}XX^\intercal\sqrt{\mathbb{S}}\right)\Big]\Big] - \sum_j x_j^2\,, \tag{10}$$

where $\mathrm{Sym}_\pm(\mathbb{A}) = \frac{1}{2}(\mathbb{A} \pm \mathbb{A}^\intercal)$ is the symmetrised (anti-symmetrized) matrix of $\mathbb{A}$ on $\mathcal{H}_c \otimes \mathcal{H}_q$ with respect to the classical index.

Finding the fundamental limit C is still an open problem. For two observables, Nagaoka conjectured that the bound C_1 is tight [3]. In other words, $C = C_1$.

Supplementary Note 2 : Nagaoka bound for two parameter estimation

The Nagaoka bound for the two parameter estimation case is [2]

$$c_{\mathrm{N}} = \min_X \left\{\mathrm{Tr}[S_\theta X_1 X_1 + S_\theta X_2 X_2] + \mathrm{TrAbs}\, S_\theta [X_1, X_2]\right\} \tag{11}$$

with X_j Hermitian satisfying (4) in the main text. In this appendix we show that in the two-parameter case, the Nagaoka–Hayashi bound, (5) in the main text, coincides with the original Nagaoka bound. When $n = 2$, the Nagaoka–Hayashi bound is

$$c_{\mathrm{NH}} = \min_{\mathbb{L},\, X}\left\{\mathbb{T}\mathrm{r}[\mathbb{S}_\theta \mathbb{L}] \,\middle|\, \begin{pmatrix} \mathbb{L}_{11} & \mathbb{L}_{12} \\ \mathbb{L}_{12} & \mathbb{L}_{22}\end{pmatrix} \geq \begin{pmatrix} X_1X_1 & X_1X_2 \\ X_2X_1 & X_2X_2\end{pmatrix}\right\}\,, \tag{12}$$

with $\mathbb{L}_{jk}$ Hermitian and X_j Hermitian satisfying (4) in the main text. We can write the condition in (12) as

$$\begin{pmatrix} \mathbb{L}_{11} & \mathbb{L}_{12} \\ \mathbb{L}_{12} & \mathbb{L}_{22}\end{pmatrix} - \begin{pmatrix} X_1X_1 & \frac{1}{2}\{X_1, X_2\} \\ \frac{1}{2}\{X_2, X_1\} & X_2X_2\end{pmatrix} \geq \begin{pmatrix} 0 & \frac{1}{2}[X_1, X_2] \\ \frac{1}{2}[X_2, X_1] & 0\end{pmatrix}\,. \tag{13}$$

Recognising that $[X_1, X_2]/2$ is an antihermitian matrix which we label as $\mathrm{i}H$, we can rewrite the condition as

$$\begin{pmatrix} \mathbb{L}'_{11} & \mathbb{L}'_{12} - \mathrm{i}H \\ \mathbb{L}'_{12} + \mathrm{i}H & \mathbb{L}'_{22}\end{pmatrix} \geq 0\,, \tag{14}$$

where $\mathbb{L}'$ denotes the matrix on the left hand side of (13). In order for this matrix to be positive we require [4]

$$\|\mathbb{L}'_{11} + \mathbb{L}'_{22}\| \geq \|2\mathrm{i}H\|\,, \tag{15}$$

for any unitarily invariant norm. This inequality can be saturated by the choice

$$\mathbb{L}' = \begin{pmatrix} |H| & 0 \\ 0 & |H|\end{pmatrix}\,, \tag{16}$$

where $|H| = \sqrt{H^2}$. The following lemma ensures (14) is satisfied.

Lemma 2 (Bhatia, corollary 1.3.7 [5]). *Let* A *be any matrix. Then the matrix* $\begin{pmatrix} |A| & A^\dagger \\ A & |A|\end{pmatrix}$ *is positive.*

The matrix $\mathbb{L}'$ can be chosen in this way by optimising over the matrix $\mathbb{L}$ so that

$$\begin{pmatrix} \mathbb{L}_{11} & \mathbb{L}_{12} \\ \mathbb{L}_{12} & \mathbb{L}_{22} \end{pmatrix} - \begin{pmatrix} X_1X_1 & \frac{1}{2}\{X_1, X_2\} \\ \frac{1}{2}\{X_2, X_1\} & X_2X_2 \end{pmatrix} = \mathbb{L}' \ , \tag{17}$$

hence

$$\min_{\mathbb{L}} \left\| \begin{pmatrix} \mathbb{L}_{11} & \mathbb{L}_{12} \\ \mathbb{L}_{12} & \mathbb{L}_{22} \end{pmatrix} - \begin{pmatrix} X_1X_1 & \frac{1}{2}\{X_1, X_2\} \\ \frac{1}{2}\{X_2, X_1\} & X_2X_2 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0 & \frac{1}{2}[X_1, X_2] \\ \frac{1}{2}[X_2, X_1] & 0 \end{pmatrix} \right\| . \tag{18}$$

We let this norm be TrAbs, which is equal to the trace for the left hand side of this equation and so the condition becomes

$$\min_{\mathbb{L}} \mathrm{Tr}[\mathbb{L}_{11} + \mathbb{L}_{22} - X_1X_1 - X_2X_2] = \mathrm{TrAbs}\, 2\mathrm{i}H \ . \tag{19}$$

Rearranging and including S_θ we arrive at

$$c_{\mathrm{NH}} = \min_{\mathbb{L},X} \mathrm{Tr}[S_\theta(\mathbb{L}_{11} + \mathbb{L}_{22})] = \min_{X} \mathrm{Tr}[S_\theta(X_1X_1 + X_2X_2)] + \mathrm{TrAbs}\, S_\theta[X_1, X_2] = c_{\mathrm{N}} \ . \tag{20}$$

Supplementary Note 3 : Generalisation to arbitrary weight matrix

We present a generalisation of our main results to an arbitrary weight matrix $\mathsf{W} \geq 0$. In the case where the weight matrix $\mathsf{W} > 0$ is full rank, it can be set to the identity after a suitable reparametrisation for the model (see for example, Sec. V of Fujiwara and Nagaoka [6]). Since we are only interested in the local bound, this reparametrisation does not matter. Specifically, we can reparametrise the model as $\varphi_j = \sum_k \mathsf{H}_{jk}\theta_k$ where $\mathsf{H} = \sqrt{\mathsf{W}}$ is a real and regular matrix. Estimating the new parameters φ is equivalent to estimating the original parameters θ with a weight matrix W.

When W is not full rank, a bit more care is required in reparametrising the model because it might be possible that some of the new parameters φ_j are exactly zero or that two of the φ_j's might be identical. This situation is common when studying parameter estimation in the presence of nuisance parameters [7–9]. Nonetheless, it is still easy to incorporate the weight matrix W into our original framework. We now wish to minimise $\mathsf{Tr}[\mathsf{W}\,\mathsf{V}_\theta]$ instead of $\mathsf{Tr}[\mathsf{V}_\theta]$. Recalling that the MSE matrix can be written as $\mathsf{V}_\theta = \mathrm{Tr}[\mathbb{S}_\theta\mathbb{L}_\theta]$, this is handled by noting the following

$$\mathsf{W}\,\mathsf{V}_\theta = \mathsf{W}\,\mathrm{Tr}[\mathbb{S}_\theta\mathbb{L}_\theta] \tag{21}$$

$$= \mathrm{Tr}[(\mathsf{W} \otimes 1)\, \mathbb{S}_\theta\mathbb{L}_\theta] \tag{22}$$

$$= \mathrm{Tr}[\mathbb{S}'_\theta\mathbb{L}_\theta] \ , \tag{23}$$

where $\mathbb{S}'_\theta = (\mathsf{W} \otimes 1)\mathbb{S}_\theta = \mathsf{W} \otimes S_\theta$ is a positive semidefinite matrix. Thus, by changing from $\mathbb{S}_\theta$ to $\mathbb{S}'_\theta$, nothing about the problem changes and it can be solved using the same SDP as in the main text.

Supplementary Note 4 : Conversion to standard SDP and complexity discussions

Here we show that the program

$$\begin{aligned} c_{\mathrm{NH}} &= \min_{\mathbb{L},\, X} \mathbb{T}\mathrm{r}[\mathbb{S}_\theta\mathbb{L}] \ , \\ \text{subject to } & \begin{pmatrix} \mathbb{L} & X \\ X^\intercal & 1 \end{pmatrix} \geq 0 \end{aligned} \tag{24}$$

with $\mathbb{L}_{jk} = \mathbb{L}_{kj}$ Hermitian and X_j Hermitian satisfying (4) in the main text can be converted to the standard SDP program

$$\begin{aligned} c_{\mathrm{NH}} &= \min_{Y \geq 0} \mathbb{T}\mathrm{r}[\boldsymbol{F}_0\boldsymbol{Y}] \\ \text{subject to } & \mathbb{T}\mathrm{r}[\boldsymbol{F}_k\boldsymbol{Y}] = c_k \ , \qquad \text{for } k = 1, \ldots, m \ , \end{aligned} \tag{25}$$

where $\boldsymbol{Y}$ is a positive-semidefinite Hermitian matrix of size $nd+d$ having the form $\boldsymbol{Y} = \begin{pmatrix} \mathbb{L} & X \\ X^\intercal & 1 \end{pmatrix}$, d is the dimension of $\mathcal{H}_q$ and m is the total number of constraints on $\boldsymbol{Y}$. The objective function to be minimised is handled with

$$\boldsymbol{F}_0 = \begin{pmatrix} \mathbb{S}_\theta & 0 \\ 0 & 0 \end{pmatrix} . \tag{26}$$

There are five groups of constraints on $\boldsymbol{Y}$ that have to be implemented through $\boldsymbol{F}_k$ and c_k. Denoting $S_j = \frac{\partial S_\theta}{\partial \theta_j}$, the constraints are:

1. $\mathrm{Tr}[S_\theta X_j] = \theta_j$.
2. $\mathrm{Tr}[S_j X_k] = \delta_{jk}$.
3. X_j Hermitian.
4. $\mathbb{L}_{jk} = \mathbb{L}_{kj}$ Hermitian.
5. The lower n-by-n block of $\boldsymbol{Y}$ equals the identity operator.

In the following, we set $n = 3$ to simplify the notations. The group 1 constraints are achieved with the n matrices and constants

$$\begin{aligned}
\boldsymbol{F}_1^{(1)} &= \begin{pmatrix} 0 & \begin{pmatrix} S_\theta \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} S_\theta & 0 & 0 \end{pmatrix} & 0 \end{pmatrix} , \qquad c_1^{(1)} = 2\theta_1 , \\
\boldsymbol{F}_2^{(1)} &= \begin{pmatrix} 0 & \begin{pmatrix} 0 \\ S_\theta \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & S_\theta & 0 \end{pmatrix} & 0 \end{pmatrix} , \qquad c_2^{(1)} = 2\theta_2 , \\
\boldsymbol{F}_3^{(1)} &= \begin{pmatrix} 0 & \begin{pmatrix} 0 \\ 0 \\ S_\theta \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & S_\theta \end{pmatrix} & 0 \end{pmatrix} , \qquad c_3^{(1)} = 2\theta_3 .
\end{aligned} \tag{27}$$

The group 2 constraints are achieved with the $n{\times}n$ matrices and constants

$$\begin{aligned}
\boldsymbol{F}_{1j}^{(2)} &= \begin{pmatrix} 0 & \begin{pmatrix} S_j \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} S_j & 0 & 0 \end{pmatrix} & 0 \end{pmatrix} , \qquad c_{1j}^{(2)} = 2\delta_{1j} , \\
\boldsymbol{F}_{2j}^{(2)} &= \begin{pmatrix} 0 & \begin{pmatrix} 0 \\ S_j \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & S_j & 0 \end{pmatrix} & 0 \end{pmatrix} , \qquad c_{2j}^{(2)} = 2\delta_{2j} , \\
\boldsymbol{F}_{3j}^{(2)} &= \begin{pmatrix} 0 & \begin{pmatrix} 0 \\ 0 \\ S_j \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & S_j \end{pmatrix} & 0 \end{pmatrix} , \qquad c_{3j}^{(2)} = 2\delta_{3j} ,
\end{aligned} \tag{28}$$

for $j=1,\dots,n$. To implement the rest of the constraints, we introduce d^2 Hermitian basis-operators B_j for $\mathcal{L}(\mathcal{H}_q)$ where $\mathcal{L}(\mathcal{H}_q)$ denote the space of Hermitian operators in $\mathcal{H}_q$, $\mathrm{Tr}[B_jB_k]=\delta_{jk}$ and B_1 proportional to the identity [10–12]. If S_θ is not full rank, the number of basis operators can be reduced by $(d-r)^2$ where r is the rank of S_θ by restricting B_j to the quotient space $\mathcal{L}(\mathcal{H}_q)/\mathcal{L}(\ker(S_\theta))$. See for example the discussions in [13, Sec. 2.10] or [14]. The group 3 constraints are then implemented by $n\times d^2$ matrices and constants

$$\boldsymbol{F}^{(3)}_{1j}=\begin{pmatrix} 0 & \begin{pmatrix}\mathrm{i}B_j\\0\\0\end{pmatrix}\\ \begin{pmatrix}-\mathrm{i}B_j & 0 & 0\end{pmatrix} & 0\end{pmatrix}, \qquad c^{(3)}_{1j}=0\,,$$
$$\boldsymbol{F}^{(3)}_{2j}=\begin{pmatrix} 0 & \begin{pmatrix}0\\\mathrm{i}B_j\\0\end{pmatrix}\\ \begin{pmatrix}0 & -\mathrm{i}B_j & 0\end{pmatrix} & 0\end{pmatrix}, \qquad c^{(3)}_{2j}=0\,, \tag{29}$$
$$\boldsymbol{F}^{(3)}_{3j}=\begin{pmatrix} 0 & \begin{pmatrix}0\\0\\\mathrm{i}B_j\end{pmatrix}\\ \begin{pmatrix}0 & 0 & -\mathrm{i}B_j\end{pmatrix} & 0\end{pmatrix}, \qquad c^{(3)}_{3j}=0\,,$$

for $j=1,\dots,d^2$. The group 4 constraints are implemented with $\dfrac{n^2-n}{2}\times d^2$ matrices and constants

$$\boldsymbol{F}^{(4)}_{1,2,j}=\begin{pmatrix}\begin{pmatrix}0 & \mathrm{i}B_j & 0\\ -\mathrm{i}B_j & 0 & 0\\ 0 & 0 & 0\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}, \qquad c^{(4)}_{1,2,j}=0\,,$$
$$\boldsymbol{F}^{(4)}_{1,3,j}=\begin{pmatrix}\begin{pmatrix}0 & 0 & \mathrm{i}B_j\\ 0 & 0 & 0\\ -\mathrm{i}B_j & 0 & 0\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}, \qquad c^{(4)}_{1,3,j}=0\,, \tag{30}$$
$$\boldsymbol{F}^{(4)}_{2,3,j}=\begin{pmatrix}\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & \mathrm{i}B_j\\ 0 & -\mathrm{i}B_j & 0\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}, \qquad c^{(4)}_{2,3,j}=0\,.$$

for $j=1,\dots,d^2$. Finally, the group 5 constraints are implemented with d^2 matrices and constants

$$\boldsymbol{F}^{(5)}_{1}=\begin{pmatrix}0 & 0\\ 0 & B_1\end{pmatrix}, \qquad c^{(5)}_{1}=\sqrt{d}, \qquad \text{and} \qquad \boldsymbol{F}^{(5)}_{j}=\begin{pmatrix}0 & 0\\ 0 & B_j\end{pmatrix}, \qquad c^{(5)}_{j}=0 \tag{31}$$

for $j=2,3,\dots,d^2$.

The worst-case time complexity for solving the SDP (24) or (25) to a desired accuracy ϵ is $O(\sqrt{N}\log(1/\epsilon))$ where $N=(n+1)d$ is the size of the matrix $\boldsymbol{F}_0$ [15, 16]. However in our simulations, we observed that the time complexity is independent of N. This is consistent with reports in the literature that in practice, the SDP algorithms perform much better than its worst-case bound [16]. Each time step requires solving a system of linear equations with a computational complexity of $O(N^3)$. Therefore, the overall worst-case computational complexity is $O\left(N^{3/2}\log(1/\epsilon)\right)$.

Supplementary Note 5 : Estimation of qubit rotations with a two-qubit probe—analytic POVM saturating the Nagaoka–Hayashi bound

We now present an analytic measurement strategy that saturates the Nagaoka–Hayashi bound for the qubit rotation estimation problem. We first define the four sub-normalised projectors

$$\left.\begin{matrix}|\phi_1\rangle\\|\phi_2\rangle\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix}1\\ \pm a\mathrm{i}\\ \pm a\mathrm{i}\\ 1\end{pmatrix} \qquad \text{and} \qquad \left.\begin{matrix}|\phi_3\rangle\\|\phi_4\rangle\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix}1\\ \mp b\\ \mp b\\ -1\end{pmatrix} \tag{32}$$

where a and b are two non-zero real parameters satisfying $a^2+b^2 \leq 1$. An optimal strategy that saturates the Nagaoka bound for estimating θ_x and θ_y consists of measuring the five-outcome POVM with $\Pi_j = |\phi_j\rangle\langle\phi_j|$ for $j = 1, 2, 3, 4$ and $\Pi_5 = 1 - (\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4)$. The probability for each POVM outcome is

$$\begin{aligned} \left.\begin{matrix}p_1\\p_2\end{matrix}\right\} &= \frac{1}{4}a(2-\epsilon)(a \pm \theta_x)\ , \\ \left.\begin{matrix}p_3\\p_4\end{matrix}\right\} &= \frac{1}{4}b(2-\epsilon)(b \pm \theta_y)\ , \\ p_5 &= 1 - \frac{1}{2}(2-\epsilon)(a^2+b^2)\ . \end{aligned} \tag{33}$$

We can use this to construct unbiased estimators for θ_x and θ_y with

$$\begin{aligned} \xi_{x,1} = -\xi_{x,2} &= \frac{2}{(2-\epsilon)a}\ , \qquad \xi_{x,3} = \xi_{x,4} = \xi_{x,5} = 0\ , \\ \xi_{y,3} = -\xi_{y,4} &= \frac{2}{(2-\epsilon)b}\ , \qquad \xi_{y,1} = \xi_{y,2} = \xi_{y,5} = 0\ . \end{aligned} \tag{34}$$

In this construction, the fifth outcome Π_5 does not give any additional information about θ_x or θ_y. Nonetheless, it is still necessary to be included so that the POVM outcomes sum up to 1. For a finite sample, to have a better estimate of θ_x and θ_y, it is thus beneficial to have both a and b large so the outcomes Π_1 to Π_4 occur more often. However, in the asymptotic limit, the variances in our estimate of θ_x and θ_y are

$$\begin{aligned} v_x &= \xi_{x,1}^2\, p_1 + \xi_{x,2}^2\, p_2 = \frac{4(p_1+p_2)}{(2-\epsilon)^2 a^2} = \frac{2}{2-\epsilon}\ , \\ v_y &= \xi_{y,3}^2\, p_3 + \xi_{y,4}^2\, p_4 = \frac{4(p_3+p_4)}{(2-\epsilon)^2 b^2} = \frac{2}{2-\epsilon} \end{aligned} \tag{35}$$

which do not depend on a or b. The sum $v_x + v_y = 4/(2-\epsilon)$ saturates the Nagaoka bound as claimed.

For estimating all three parameters θ_x, θ_y and θ_z, one measurement strategy is to use the same POVM outcomes for estimating θ_x and θ_y but splitting Π_5 to get some information on θ_z. Ideally, we would like to use these four projectors we get when setting $a = b = 0$,

$$\Pi_1 = \Pi_2 = \frac{1}{4}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix}, \qquad \Pi_3 = \Pi_4 = \frac{1}{4}\begin{pmatrix}1&0&0&-1\\0&0&0&0\\0&0&0&0\\-1&0&0&1\end{pmatrix}, \qquad \Pi_5 = \frac{1}{2}\begin{pmatrix}0&0&0&0\\0&1&\mathrm{i}&0\\0&-\mathrm{i}&1&0\\0&0&0&0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0&0&0&0\\0&1&-\mathrm{i}&0\\0&\mathrm{i}&1&0\\0&0&0&0\end{pmatrix} \tag{36}$$

to obtain the most information on θ_z without affecting the estimate of θ_x and θ_y. But the problem is that at this singular point, the first four outcomes Π_1, Π_2, Π_3 and Π_4 do not give any information on θ_x and θ_y. To fix this, we need both a and b to be close to but not exactly zero. Writing $\delta = (a^2+b^2)/2$, we can split Π_5 as

$$\Pi_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1-\delta & -\delta & 0 \\ 0 & -\delta & 1-\delta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{37}$$

$$= \underbrace{\delta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{\Pi_5^{(3)}} + \underbrace{\frac{1-2\delta}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -\mathrm{i} & 0 \\ 0 & \mathrm{i} & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{\Pi_6^{(3)}} + \underbrace{\frac{1-2\delta}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \mathrm{i} & 0 \\ 0 & -\mathrm{i} & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{\Pi_7^{(3)}} \tag{38}$$

which has outcome probabilities

$$\begin{aligned} p_5 &= \delta\,\epsilon\,, \\ \left.\begin{matrix} p_6 \\ p_7 \end{matrix}\right\} &= \frac{1}{2}(1-2\delta)\left(1 \pm (1-\epsilon)\theta_z\right)\ . \end{aligned} \tag{39}$$

This together with

$$\xi_{z,1} = \xi_{z,2} = \xi_{z,3} = \xi_{z,4} = \xi_{z,5} = 0\,, \qquad \text{and} \qquad \xi_{z,6} = -\xi_{z,7} = \frac{1}{(1-\epsilon)(1-2\delta)}\,, \tag{40}$$

give a variance for estimating θ_z as $v_z = \frac{1}{(1-\epsilon)^2(1-2\delta)}$ which approaches $v_z = \frac{1}{(1-\epsilon)^2}$ as δ tends to zero.

Supplementary Note 6 : Phase and transmissivity estimation in interferometry—analytic POVM saturating the Holevo Cramér–Rao bound for 1 photon state

Consider the 1 photon state $|\psi_{\mathrm{in}}\rangle = |01\rangle\, a_0 + |10\rangle\, a_1$ where a_0 and a_1 are positive coefficients. This state transforms through the lossy interferometer with transmissivity η and a phase shift ϕ to the state with matrix representation

$$S_\theta = \begin{pmatrix} (1-\eta)a_1^2 & 0 & 0 \\ 0 & a_0^2 & \sqrt{\eta}a_0a_1e^{-\mathrm{i}\phi} \\ 0 & \sqrt{\eta}a_0a_1e^{\mathrm{i}\phi} & \eta a_1^2 \end{pmatrix} \tag{41}$$

whose derivatives evaluated at $\phi = 0$ are

$$\frac{\partial S_\theta}{\partial \eta} = \begin{pmatrix} -a_1^2 & 0 & 0 \\ 0 & 0 & \frac{a_0a_1}{2\sqrt{\eta}} \\ 0 & \frac{a_0a_1}{2\sqrt{\eta}} & a_1^2 \end{pmatrix} \qquad \text{and} \qquad \frac{\partial S_\theta}{\partial \phi} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i}\sqrt{\eta}a_0a_1 \\ 0 & \mathrm{i}\sqrt{\eta}a_0a_1 & 0 \end{pmatrix}\,, \tag{42}$$

where the matrix basis is $\{|00\rangle\,, |01\rangle\,, |10\rangle\}$. The Holevo bound for this model was computed by Albarelli et al. [14] to be

$$c_{\mathrm{H}} = \begin{cases} \dfrac{1+3\eta-4\eta^3}{4\eta a_1^2} & \text{for } a_1 < \dfrac{1}{\sqrt{2}} \text{ and } \eta < \dfrac{a_0^2-a_1^2}{2a_0^2}\,, \\ \dfrac{\left(a_0^2+\eta a_1^2\right)\left(1+4\eta(1-\eta)a_0^2\right)}{4\eta a_0^2a_1^2} & \text{otherwise.} \end{cases} \tag{43}$$

In the following, we show that this bound can be saturated by a separable measurement. There exist a family of measurements that can saturate the Holevo bound. One of them is the four-outcome POVM

$$\Pi_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \qquad \Pi_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 - \dfrac{a_0^2}{(1-\eta)(1+2\eta)a_0^2 - \eta a_1^2} \end{pmatrix} ,$$
$$\left.\begin{matrix}\Pi_3\\ \Pi_4\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \mp\dfrac{\mathrm{i}a_0}{\sqrt{(1-\eta)(1+2\eta)a_0^2 - \eta a_1^2}} \\ 0 & \pm\dfrac{\mathrm{i}a_0}{\sqrt{(1-\eta)(1+2\eta)a_0^2 - \eta a_1^2}} & \dfrac{a_0^2}{(1-\eta)(1+2\eta)a_0^2 - \eta a_1^2} \end{pmatrix} \tag{44}$$

together with the estimation coefficients

$$\xi_{\eta,1} = -\frac{1+2\eta}{2a_1^2} , \qquad \xi_{\eta,2} = \frac{(1-\eta)(1+2\eta)}{2\eta a_1^2} , \qquad \xi_{\eta,3} = \xi_{\eta,4} = \frac{1}{2a_0^2} ,$$
$$\xi_{\phi,1} = \xi_{\phi,2} = 0 , \qquad \text{and} \qquad \xi_{\phi,3} = -\xi_{\phi 4} = \frac{\sqrt{(1-\eta)(1+2\eta)a_0^2 - \eta a_1^2}}{2\sqrt{\eta}a_0^2 a_1} . \tag{45}$$

One can verify that when $\eta < (a_0^2 - a_1^2)/2$, these outcomes are non-negative operators that satisfy $\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4 = 1$. The estimator matrices $X_\eta = \xi_{\eta,1}\Pi_1 + \xi_{\eta,2}\Pi_2 + \xi_{\eta,3}\Pi_3 + \xi_{\eta,4}\Pi_4$ and $X_\phi = \xi_{\phi,3}\Pi_3 + \xi_{\phi,4}\Pi_4$ satisfy the unbiased conditions, (4) in the main text. The probability for each outcome to occur is

$$\begin{aligned} p_1 &= (1-\eta)a_1^2 , \\ p_2 &= \eta a_1^2 - \frac{\eta a_0^2 a_1^2}{(1-\eta)(1+2\eta) - \eta a_1^2} , \\ p_3 = p_4 &= \frac{a_0^2}{2}\left(1 + \frac{\eta a_1}{(1-\eta)(1+2\eta)a_0^2 - \eta a_1^2}\right) . \end{aligned} \tag{46}$$

The variances of these two estimators are

$$\begin{aligned} v_\eta &= \xi_{\eta,1}^2\, p_1 + \xi_{\eta,2}^2\, p_2 + \xi_{\eta,3}^2\, p_3 + \xi_{\eta,4}^2\, p_4 = \frac{1+\eta-2\eta^2}{2a_1^2} , \\ v_\phi &= \xi_{\phi,3}^2\, p_3 + \xi_{\phi,4}^2\, p_4 = \frac{1+\eta-2\eta^2}{4\eta a_1^2} , \end{aligned} \tag{47}$$

which together gives $v_\eta + v_\phi = (1 + 3\eta - 4\eta^3)/4\eta a_1^2$ saturating the Holevo bound (43) as claimed.

At the boundary $\eta = (a_0^2 - a_1^2)/2a_0^2$, the POVM outcome $\Pi_2 = 0$ while the remaining three reduce to a projective measurement on the eigenstate of the SLD operator [14]

$$\Pi_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \qquad \left.\begin{matrix}\Pi_3\\ \Pi_4\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \mp\mathrm{i} \\ 0 & \pm\mathrm{i} & 1 \end{pmatrix} . \tag{48}$$

In this case, the estimator coefficients are

$$\xi_{\eta,1} = -\frac{a_0^2 + a_1^2\eta}{a_1^2} , \qquad \xi_{\eta,3} = \xi_{\eta,4} = 1-\eta , \qquad \xi_{\phi,1} = 0 \qquad \text{and} \qquad \xi_{\phi,3} = -\xi_{\phi,4} = \frac{1}{2\sqrt{\eta}a_0 a_1} . \tag{49}$$

This measurement scheme remains optimal even when $\eta > (a_0^2 - a_1^2)/2a_0^2$. Comparing the 4-outcome POVM (44) to the 3-outcome POVM (48), we see that the role played by Π_2 is to obtain a better estimate of η, but at the expense of a worse estimate of ϕ. Whether this trade-off improves the overall sum of the MSE depends on the exact form of the probe and the value of η. We note that the estimators presented here depend on the unknown parameter η. Although this would be an issue if we were interested in global parameter estimation, for local estimation this is not an issue, as we are only interested in estimating η in the local neighbourhood of some a priori known value, η_0.

Supplementary Note 7 : Dual solutions for the semidefinite program

From the constructed POVM, we can arrive at a candidate for the optimal X and $\mathbb{L}$ matrices which gives an upper bound to the primal solution. In this supplementary note, we write down the dual problem and provide its solution which gives a lower bound to the primal solution. One can easily check that the lower and upper bounds coincide which implies that the candidate solution is indeed an optimal solution for the Nagaoka–Hayashi bound.

The dual problem is

$$\begin{aligned} \tilde{c}_{\mathrm{NH}} &= \max_{y} \sum_k y_k\, c_k \\ \text{subject to } & \sum_k y_k \boldsymbol{F}_k \leq \boldsymbol{F}_0\ , \end{aligned} \tag{50}$$

where the matrices $\boldsymbol{F}_k$ and constants c_k implements the constraints on the primal SDP as defined in Supplementary Note 4.

We first present the dual solution for the qubit rotation estimation problem. In order to write down the dual solutions, we need to choose a representation for the set of basis matrices $\{B_j\}$ in Supplementary Note 4. We use the following 16 matrices:

$$\begin{aligned}
B_1 &= \frac{1}{2}\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{pmatrix}, &
B_2 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&0\\0&0&0&0 \end{pmatrix}, &
B_3 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&1&0\\0&0&0&0\\1&0&0&0\\0&0&0&0 \end{pmatrix}, &
B_4 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0 \end{pmatrix}, \\
B_5 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&0 \end{pmatrix}, &
B_6 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&0\\0&0&0&1\\0&0&0&0\\0&1&0&0 \end{pmatrix}, &
B_7 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&1&0 \end{pmatrix}, &
B_8 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&-\mathrm{i}&0&0\\\mathrm{i}&0&0&0\\0&0&0&0\\0&0&0&0 \end{pmatrix}, \\
B_9 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&-\mathrm{i}&0\\0&0&0&0\\\mathrm{i}&0&0&0\\0&0&0&0 \end{pmatrix}, &
B_{10} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&-\mathrm{i}\\0&0&0&0\\0&0&0&0\\\mathrm{i}&0&0&0 \end{pmatrix}, &
B_{11} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&0\\0&0&-\mathrm{i}&0\\0&\mathrm{i}&0&0\\0&0&0&0 \end{pmatrix}, &
B_{12} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&0\\0&0&0&-\mathrm{i}\\0&0&0&0\\0&\mathrm{i}&0&0 \end{pmatrix}, \\
B_{13} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&-\mathrm{i}\\0&0&\mathrm{i}&0 \end{pmatrix}, &
B_{14} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1&0&0&0\\0&-1&0&0\\0&0&0&0\\0&0&0&0 \end{pmatrix}, &
B_{15} &= \frac{1}{\sqrt{6}}\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&-2&0\\0&0&0&0 \end{pmatrix}, &
B_{16} &= \frac{1}{2\sqrt{3}}\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-3 \end{pmatrix}.
\end{aligned} \tag{51}$$

One can then check that for estimating a single parameter, the following dual solution coincides with the primal candidate:

$$\begin{aligned} y^{(2)}_{1,1} = 1,\ y^{(3)}_{1,2} = y^{(3)}_{1,7} = \frac{1}{\sqrt{8}},\ y^{(3)}_{1,3} = y^{(3)}_{1,6} = \frac{1-\epsilon}{\sqrt{8}}, \\ y^{(5)}_{1,1} = -\frac{1}{2},\ y^{(5)}_{1,4} = -\frac{1-\epsilon}{\sqrt{2}},\ y^{(5)}_{1,14} = -\frac{1}{\sqrt{8}},\ y^{(5)}_{1,15} = -\frac{1}{\sqrt{24}},\ y^{(5)}_{1,16} = \frac{1}{\sqrt{12}}\ , \end{aligned} \tag{52}$$

and all other y_k zero. For estimating two parameters:

$$y^{(2)}_{1,1} = y^{(2)}_{2,2} = \frac{2}{2-\epsilon},\ y^{(3)}_{1,2} = y^{(3)}_{1,7} = y^{(3)}_{2,8} = y^{(3)}_{2,13} = \frac{1}{\sqrt{2}(2-\epsilon)},\ y^{(3)}_{1,3} = y^{(3)}_{1,6} = y^{(3)}_{2,9} = y^{(3)}_{2,12} = \frac{1-\epsilon}{\sqrt{2}(2-\epsilon)},$$
$$y^{(4)}_{1,2,14} = -\frac{\epsilon}{\sqrt{8}},\ y^{(4)}_{1,2,15} = \epsilon\sqrt{\frac{3}{8}},\ y^{(5)}_{1} = -\frac{2}{2-\epsilon},\ y^{(5)}_{14} = -\frac{\sqrt{2}}{2-\epsilon},\ y^{(5)}_{15} = -\frac{\sqrt{2}}{\sqrt{3}(2-\epsilon)},\ y^{(5)}_{16} = \frac{2}{\sqrt{3}(2-\epsilon)}, \tag{53}$$

and all other y_k zero. For estimating three parameters:

$$y^{(2)}_{1,1} = y^{(2)}_{2,2} = \frac{2}{2-\epsilon},\ y^{(2)}_{3,3} = \frac{1}{(1-\epsilon)^2},\ y^{(3)}_{1,2} = y^{(3)}_{1,7} = y^{(3)}_{2,8} = y^{(3)}_{2,13} = \frac{1}{\sqrt{2}(2-\epsilon)},$$
$$y^{(3)}_{1,3} = y^{(3)}_{1,6} = y^{(3)}_{2,9} = y^{(3)}_{2,12} = \frac{1-\epsilon}{\sqrt{2}(2-\epsilon)},\ y^{(3)}_{3,14} = -\frac{1}{\sqrt{8}},\ y^{(3)}_{3,15} = \frac{\sqrt{3}}{\sqrt{8}},\ y^{(4)}_{1,2,14} = -\frac{\epsilon}{\sqrt{8}},\ y^{(4)}_{1,2,15} = \epsilon\sqrt{\frac{3}{8}}, \tag{54}$$
$$y^{(5)}_{1} = -\frac{2}{2-\epsilon} - \frac{1}{2(1-\epsilon)^2},\ y^{(5)}_{1} = \frac{1}{\sqrt{2}(1-\epsilon)},\ y^{(5)}_{14} = -\frac{\sqrt{2}}{2-\epsilon} + \frac{\sqrt{2}}{4(1-\epsilon)^2},\ y^{(5)}_{15} = \frac{1}{\sqrt{3}}y^{(5)}_{14},\ y^{(5)}_{16} = -\sqrt{\frac{2}{3}}y^{(5)}_{14},$$

and all other y_k zero.

We now write down the dual solution to the Nagaoka–Hayashi bound for the second example, phase and transmissivity estimation in an interferometer, when $N = 1$. To do this, we use the following 9 matrices as basis matrices:

$$B_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\quad B_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},\quad B_3 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix},$$
$$B_4 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix},\quad B_5 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&-\mathrm{i}&0\\\mathrm{i}&0&0\\0&0&0\end{pmatrix},\quad B_6 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&-\mathrm{i}\\0&0&0\\\mathrm{i}&0&0\end{pmatrix}, \tag{55}$$
$$B_7 = \frac{1}{\sqrt{2}}\begin{pmatrix}0&0&0\\0&0&-\mathrm{i}\\0&\mathrm{i}&0\end{pmatrix},\quad B_8 = \frac{1}{\sqrt{2}}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix},\quad B_9 = \frac{1}{\sqrt{6}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix}.$$

When $a_1 < 1/\sqrt{2}$ and $\eta < (a_0^2 - a_1^2)/2a_0^2$, one solution to the dual problem is:

$$y^{(2)}_{1,1} = \frac{(1-\eta)(1+2\eta)}{2a_1^2},\ y^{(2)}_{2,2} = \frac{(1-\eta)(1+2\eta)}{4\eta a_1^2},$$
$$y^{(3)}_{1,7} = -y^{(3)}_{2,4} = \frac{a_0(1+\eta-2\eta^2)}{2\sqrt{2\eta}a_1},\ y^{(3)}_{2,8} = \frac{1+\eta-2\eta^2}{2\sqrt{2}},\ y^{(3)}_{2,9} = \sqrt{\frac{3}{8}}(1+\eta-2\eta^2),$$
$$y^{(4)}_{1,2,1} = -\frac{1}{\sqrt{3}},\ y^{(4)}_{1,2,4} = -\sqrt{2\eta}a_0a_1,\ y^{(4)}_{1,2,8} = \frac{a_0^2-(1-\eta)a_1^2}{\sqrt{2}},\ y^{(4)}_{1,2,9} = -\frac{1}{\sqrt{6}} + \sqrt{\frac{3}{2}}\eta a_1^2, \tag{56}$$
$$y^{(5)}_{1} = -\frac{(1-\eta)(1+2\eta)^2}{4\sqrt{3}\eta a_1^2},\ y^{(5)}_{8} = -\frac{(1-\eta)(1+2\eta)^2}{4\sqrt{2}a_1^2},\ y^{(5)}_{9} = \frac{(1-\eta)(2-3\eta)(1+2\eta)^2}{4\sqrt{6}\eta a_1^2},$$

and all other y_k zero. When the condition $a_1 < 1/\sqrt{2}$ and $\eta < (a_0^2 - a_1^2)/2a_0^2$ is not satisfied, one solution to the dual problem is given by:

$$\begin{gathered} y_{1,1}^{(2)} = \frac{(1-\eta)(a_0^2+\eta a_1^2)}{a_1^2},\ y_{2,2}^{(2)} = \frac{a_0^2+\eta a_1^2}{4\eta a_0^2 a_1^2},\ y_{1,7}^{(3)} = \frac{(1-\eta)a_0}{\sqrt{2\eta}a_1}(a_0^2+\eta a_1^2), \\ y_{2,4}^{(3)} = \frac{a_1^2\eta - a_0^2 - 8\eta(1-\eta)^2 a_1^2 a_0^4}{2\sqrt{2\eta}a_0 a_1},\ y_{2,8}^{(3)} = -\frac{1}{2\sqrt{2}} + \sqrt{2}a_0^2(1-\eta)^2(2a_0^2+\eta a_1^2),\ y_{2,9}^{(3)} = \sqrt{\frac{3}{8}} + \sqrt{6}\eta(1-\eta)^2 a_0^2 a_1^2, \\ y_{1,2,1}^{(4)} = -\frac{2}{\sqrt{3}}(1-\eta)a_0^2,\ y_{1,2,4}^{(4)} = -2\sqrt{2\eta}(1-\eta)a_0^3 a_1,\ y_{1,2,8}^{(4)} = \sqrt{2}(1-\eta)a_0^2(a_0^2-(1-\eta)a_1^2), \\ y_{1,2,9}^{(4)} = -\sqrt{\frac{2}{3}}(1-\eta)a_0^2(1-3\eta a_1^2),\ y_1^{(5)} = -\frac{(a_0^2+\eta a_1^2)(1+4\eta(1-\eta)a_0^2)}{4\sqrt{3}\eta a_0^2 a_1^2}, \\ y_4^{(5)} = \frac{1-4(1-\eta)^2 a_0^4}{2\sqrt{2\eta}a_0 a_1},\ y_8^{(5)} = \frac{a_1^2 - 4a_0^2(1-\eta)(a_0^2+a_0^2 a_1^2\eta + a_1^4\eta^2)}{4\sqrt{2}a_0^2 a_1^2}, \\ y_9^{(5)} = \frac{-\eta + (2+\eta+8\eta^2-20\eta^3+12\eta^4)a_0^2 + 4\eta(1-\eta)^2(5-6\eta)a_0^4 - 12\eta(1-\eta)^2(2-\eta)a_0^6}{4\sqrt{6}\eta a_0^2 a_1^2}, \end{gathered} \tag{57}$$

and all other y_k zero. One can check that these solutions coincide with the primal solution in Supplementary Note 6.

Supplementary References

[1] Hayashi, M. On simultaneous measurement of noncommutative observables. In *Development of infinite-dimensional non-commutative anaysis*, 96–188 (Surikaisekikenkyusho (RIMS), Kyoto Univ., Kokyuroku No. 1099, In Japanese, 1999).

[2] Nagaoka, H. A new approach to Cramér–Rao bounds for quantum state estimation. In *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, 100–112 (World Scientific, 2005). Originally published as IEICE Technical Report, 89, 228, IT 89-42, 9-14, (1989).

[3] Nagaoka, H. A generalization of the simultaneous diagonalization of Hermitian matrices and its relation to quantum estimation theory. In *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, 133–149 (World Scientific, 2005). Originally published as Trans. Jap. Soc. Indust. Appl. Math., 1, 43-56, (1991) in Japanese. Translated to English by Y.Tsuda.

[4] Lin, M. & Wolkowicz, H. Hiroshima's theorem and matrix norm inequalities. *Acta Sci. Math.(Szeged)* **81**, 45–53 (2015).

[5] Bhatia, R. *Positive definite matrices* (Princeton university press, 2009).

[6] Fujiwara, A. & Nagaoka, H. An estimation theoretical characterization of coherent states. *J. Math. Phys.* **40**, 4227–4239 (1999).

[7] Suzuki, J. Nuisance parameter problem in quantum estimation theory: Tradeoff relation and qubit examples. *J. Phys. A Math. Theor.* **53**, 264001 (2020).

[8] Suzuki, J., Yang, Y. & Hayashi, M. Quantum state estimation with nuisance parameters. *J. Phys. A Math. Theor.* **53**, 453001 (2020).

[9] Tsang, M., Albarelli, F. & Datta, A. Quantum semiparametric estimation. *Phys. Rev. X* **10**, 031023 (2020).

[10] Hioe, F. T. & Eberly, J. H. N-level coherence vector and higher conservation laws in quantum optics and quantum mechanics. *Phys.Rev. Lett.* **47**, 838–841 (1981).

[11] Kimura, G. The Bloch vector for N-level systems. *Phys. Lett. A* **314**, 339–349 (2003).

[12] Bertlmann, R. A. & Krammer, P. Bloch vectors for qudits. *J. Phys. A Math. Theor.* **41**, 235303 (2008).

[13] Holevo, A. S. *Probabilistic and statistical aspects of quantum theory*, vol. 1 (Springer Science & Business Media, 2011).

[14] Albarelli, F., Friel, J. F. & Datta, A. Evaluating the Holevo Cramér-Rao bound for multiparameter quantum metrology. *Phys. Rev. Lett.* **123**, 200503 (2019).

[15] Boyd, S. & Vandenberghe, L. *Convex optimization* (Cambridge university press, 2004).

[16] Vandenberghe, L. & Boyd, S. Semidefinite programming. *SIAM review* **38**, 49–95 (1996).