Report 99-009

An Example of Non-Attainability of Expected Quantum Information

O.E. Barndorff-Nielsen
Richard D. Gill

ISSN: 1389–2355

An Example of Non-Attainability of Expected Quantum Information

O.E. Barndorff-Nielsen[1] and R.D. Gill[2]
[1] *MaPhySto, Aarhus University, 8000 Aarhus-C, Denmark*
[2] *Mathematical Institute, University of Utrecht, Box 80010, 3508 TA Utrecht, Netherlands and Eurandom, Box 513, 5600 MB Eindhoven, Netherlands*
(31 January 99)

Braunstein and Caves showed how quantum information can be used to define a metric for quantum states, relying on attainability of a quantum information bound. We show that the bound is not generally attainable, but that a two-stage procedure of repeated measurements achieves the bound in the limit. We connect to the question of 'non-locality without entanglement': can a joint measurement on n independent copies of a quantum system yield more information than separate measurements. Though for small n generalized measurements are more informative, the gain is asymptotically negligible in the pure state, spin-half, examples studied.

PACS numbers: 03.65.Bz, 03.67.-a

1 Introduction

Braunstein and Caves [1] have clarified the relation between classical expected information $i(\theta)$, in the sense of Fisher, and the analogous concept of expected quantum information $I(\theta)$, by showing that $I(\theta)$ is an upper bound of $i(\theta;M)$ with respect to all (dominated) generalized measurements M of the state $\rho=\rho(\theta)$ where θ is an unknown parameter and $i(\theta;M)$ is the Fisher expected information for θ in the distribution of the outcome of the measurement of M. They indicate moreover that a measurement exists achieving the bound. In the present paper we show by an example, for an elementary spin-$\frac{1}{2}$ situation, that in general there does not exist a single measurement M such that $i(\theta;M)=I(\theta)$ for all θ simultaneously.

The example is presented in section 3, after a brief recapitulation in section 2 of expected classical and quantum information. Section 4 discusses the implications of the result. In the one parameter case, we show that the bound is generally asymptotically achievable, so that Braunstein and Cave's motivation for $I(\theta)$ in the definition of a statistical distinguishability metric for quantum states can be maintained. However for the multiparameter case the situation is considerably more complicated. For one special case—a completely unknown spin half pure state—it turns out again that separate measurements can asymptotically achieve anything achievable by a generalised measurement. In general, the problem remains open to characterise the classes of locally and globally achievable information matrices, both when all measurements are considered and when only separate measurements on separate particles are allowed.

2 Expected classical and quantum information

Consider a general quantum state $\rho=\rho(\theta)$ that depends on an unknown scalar parameter θ. Consider also a generalised measurement M on a measure space $(\mathcal{X},\mathcal{A})$ of the form

$$M(A)=\int_A m(x)\mu(\mathrm{d}x)$$

where the operator $m(x)$ is nonnegative and selfadjoint and μ is a real σ-finite measure on $(\mathcal{X},\mathcal{A})$. The probability density with respect to μ of the observation X arising from a single measurement by M is

$$p(x;\theta)=\mathrm{trace}\{\rho(\theta)m(x)\}.$$

The expected Fisher information from this measurement on θ is defined by

$$i(\theta;M)=\mathrm{E}\{\dot{l}_\theta^2\}=\int_{\mathcal{X}}(\dot{l}_\theta(x))^2p(x;\theta)\mu(\mathrm{d}x)$$

where

$$l=l(\theta)=\log p(x;\theta)$$

is the log likelihood function of θ and $\dot{l}_\theta=(\partial/\partial\theta)l(\theta)$. For this quantity to be statistically meaningful, it is necessary that $\{x:p(x;\theta)>0\}$ does not depend on θ.

Now, let $\mathbb{D}\rho$ denote the symmetric logarithmic derivative or *quantum score* of ρ with respect to θ, that is, the self-adjoint operator $\mathbb{D}\rho$ given implicitly by

$$\dot{\rho}_\theta=\frac{1}{2}(\rho\,\mathbb{D}\rho+\mathbb{D}\rho\,\rho). \qquad (1)$$

The expected quantum information on θ is defined by

$$I(\theta)=\mathrm{trace}\{\rho(\mathbb{D}\rho)^2\}.$$

Note that this quantity does not depend on M. It is possible to express the information $i(\theta;M)$ in terms of the quantum score for ρ, namely as

$$i(\theta;M)=\int_{\mathcal{X}}p(x;\theta)^{-1}\left[\mathrm{Re}\,\mathrm{trace}\{\rho\mathbb{D}\rho m(x)\}\right]^2\mu(\mathrm{d}x).$$

This follows on noting that

$$\begin{aligned}\dot{l}_\theta(x)&=p(x;\theta)^{-1}\mathrm{trace}\{\dot{\rho}_\theta m(x)\}\\&=p(x;\theta)^{-1}\frac{1}{2}\mathrm{trace}\{(\rho\mathbb{D}\rho+\mathbb{D}\rho\rho)m(x)\}\\&=p(x;\theta)^{-1}\mathrm{Re}\,\mathrm{trace}\{\rho\mathbb{D}\rho m(x)\}.\end{aligned}$$

As shown by Braunstein and Caves [1], it follows from the Cauchy-Schwarz inequality for traces of operators on the underlying Hilbert space $\mathcal{H}$, that

$$i(\theta; M) \leq I(\theta). \tag{2}$$

Necessary and sufficient conditions for equality in (2) are that for (μ almost) all x we have

$$\mathrm{Im}\,\mathrm{trace}\{\rho\mathbb{D}\rho m(x)\} = 0 \tag{3}$$

and

$$m(x)^{1/2}\{\kappa(x;\theta)^{1/2}\mathbf{1} - \mathbb{D}\rho\}\rho^{1/2} = \mathbf{0} \tag{4}$$

where $\kappa(x;\theta) = \mathrm{trace}\{\rho\mathbb{D}\rho m(x)\mathbb{D}\rho\}/\mathrm{trace}\{\rho m(x)\}$.

3 Spin-$\frac{1}{2}$ example

For a single spin-$\frac{1}{2}$ particle the pure states have density matrices of the form $\rho = |\psi\rangle\langle\psi|$ where

$$|\psi\rangle = |\psi(\eta,\phi)\rangle = \begin{bmatrix} e^{-i\phi/2}\cos(\eta/2) \\ e^{i\phi/2}\sin(\eta/2) \end{bmatrix}$$

and hence $\rho = \rho(\eta,\phi)$ is given by

$$\rho = \begin{bmatrix} \cos^2(\frac{1}{2}\eta) & e^{-i\phi}\cos(\frac{1}{2}\eta)\sin(\frac{1}{2}\eta) \\ e^{i\phi}\cos(\frac{1}{2}\eta)\sin(\frac{1}{2}\eta) & \sin^2(\frac{1}{2}\eta) \end{bmatrix}.$$

In the following we consider $\eta \in (0,\pi)$ as known and $\phi \in [0, 2\pi)$ as the unknown parameter. (For $\eta = 0$ and $\eta = \pi$ the parameter ϕ is meaningless).

Our first step is to determine the quantum score for ρ. The derivative of ρ itself with respect to ϕ is

$$\dot{\rho}_\phi = \begin{bmatrix} 0 & -ie^{-i\phi}\cos(\frac{1}{2}\eta)\sin(\frac{1}{2}\eta) \\ ie^{i\phi}\cos(\frac{1}{2}\eta)\sin(\frac{1}{2}\eta) & 0 \end{bmatrix}.$$

Since ρ is pure, $\rho^2 = \rho$ and hence $\rho\dot{\rho}_\phi + \dot{\rho}_\phi\rho = \dot{\rho}_\phi$. Comparing with the defining relation (1) shows us that $\mathbb{D}\rho = 2\dot{\rho}_\phi$ and hence

$$\mathbb{D}\rho = \sin\eta \begin{bmatrix} 0 & e^{-i(\phi+\pi/2)} \\ e^{i(\phi+\pi/2)} & 0 \end{bmatrix} = \sin\eta\,\sigma_{\pi/2,\phi+\pi/2}$$

where $\sigma_{\eta,\phi} = \sin\eta\cos\phi\,\sigma_x + \sin\eta\sin\phi\,\sigma_y + \cos\eta\,\sigma_z$ denotes the Pauli spin matrix (with eigenvalues ± 1) for the direction, in polar cordinates, (η,ϕ).

Since $\sigma^2_{\eta,\phi} = \mathbf{1}$ and $\mathrm{trace}(\rho) = 1$ it follows that the expected quantum information on ϕ is

$$I(\phi) = \sin^2\eta.$$

In terms of the basis $|\uparrow\rangle = |\psi(\eta,\phi)\rangle$ and $|\downarrow\rangle = |\psi(\pi - \eta, \phi + \pi)\rangle$, we find

$$\rho = |\uparrow\rangle\langle\uparrow|, \quad \mathbb{D}\rho = \sin\eta(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|).$$

Conditions (3) and (4) for equality of $I(\phi)$ and $i(\phi)$ therefore take the form

$$\mathrm{Im}\, m_{\downarrow\uparrow} = 0$$

and

$$\kappa^{1/2}m_{\uparrow\uparrow} = \sin\eta\, m_{\uparrow\downarrow}, \qquad \kappa^{1/2}m_{\downarrow\uparrow} = \sin\eta\, m_{\downarrow\downarrow}.$$

The first and second of these three equalities tell us $m_{\downarrow\uparrow} = m_{\uparrow\downarrow} = (\kappa^{1/2}/\sin\eta)m_{\uparrow\uparrow}$ (real), and together with the third we find $m_{\downarrow\downarrow} = (\kappa^{1/2}/\sin\eta)^2 m_{\uparrow\uparrow}$. Define the real numbers $\alpha = m^{1/2}_{\uparrow\uparrow}$, $\beta = (\kappa^{1/2}/\sin\eta)m^{1/2}_{\uparrow\uparrow}$, so that $m_{\uparrow\uparrow} = \alpha^2$, $m_{\downarrow\uparrow} = m_{\uparrow\downarrow} = \alpha\beta$, and $m_{\downarrow\downarrow} = \beta^2$. Then $m = |\xi\rangle\langle\xi|$ with $|\xi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ with α and β real. As ϕ varies, α and β may vary too but m must remain constant. In the original coordinate system

$$\xi = \begin{bmatrix} (\alpha\cos(\frac{1}{2}\eta) - i\beta\sin(\frac{1}{2}\eta))e^{-i\phi/2} \\ (\alpha\sin(\frac{1}{2}\eta) + i\beta\cos(\frac{1}{2}\eta))e^{i\phi/2} \end{bmatrix}.$$

This vector must be constant as ϕ varies, up to an arbitrarily varying phase. Therefore $\|\xi\|^2 = |\xi_1|^2 + |\xi_2|^2 = \alpha^2 + \beta^2$ is constant, and $|\xi_1|^2 = \alpha^2\cos^2(\frac{1}{2}\eta) + \beta^2\sin^2(\frac{1}{2}\eta)$ is constant. This implies, as long as $\eta \neq \pi/2$ so that these two equations are linearly independent, that α^2 and β^2 are constant. Since α and β are real, this implies that α and β are constant as ϕ takes on at least several different values. Thus ξ varies (by more than a phase change) as ϕ varies. Consequently, for $\eta \neq \pi/2$, no measurement M exists with Fisher information $i(\phi; M)$ equal to the quantum information $I(\phi)$ whatever the value of the unknown parameter ϕ.

In the exceptional case $\eta = \pi/2$ it is possible to achieve the bound uniformly in ϕ. Any measurement with all components proportional to projector matrices for spin directions in the x-y plane will do the job.

4 Discussion

We have shown, for the case of a one-dimensional parameter as considered by Braunstein and Caves, that there need not exist a measurement M such that $i(\theta; M) = I(\theta)$ for all parameter values θ simultaneously. It is on the other hand possible to find a measurement M such that at a *given* parameter-value, $i(\theta; M) = I(\theta)$, as Braunstein and Caves indicate. They do not remark on the possible dependence of M on θ. As we explain below it is vital for their arguments to come up with a measurement which achieves the bound independently of θ. A similar lack of distinction between measurements optimal at a single point in the parameter space and global optimality occurs elsewhere in the literature; see Fujiwara and Nagaoka [2] (section 4, formula (17)) for another instance.

The construction of M at a specified value of θ is as follows. Supposing for simplicity that $\mathbb{D}\rho$ has discrete spectrum, let $m(x)$ be the projector onto the eigenspace of $\mathbb{D}\rho$ with eigenvalue x, and let μ be counting measure on the eigenvalues so that $\int m(x)\mu(dx) = \mathbf{1}$. Then for each x, $\mathbb{D}\rho$ and $m(x)$ commute and their product equals $xm(x)$. We find that (2.3) and (2.4) are satisfied with $\kappa(x,\theta) = x^2$. However the eigenspace decomposition of $\mathbb{D}\rho(\theta)$ generally depends on θ so this does not define a measurent M which achieves the bound uniformly in θ.

Braunstein and Caves' aim was to define a statistical distinguishability metric between quantum states. The reason that $i(\theta; M)$ is relevant in this context is (as those authors explain) that based on n independent measurements M of identical copies of the given quantum system, there typically exists an asymptotically unbiased estimator of θ with *asymptotic* variance $(ni(\theta; M))^{-1}$ and moreover no estimator based on the same measurements can do better. That estimator—the maximum likelihood estimator—works for all values of θ. On the other hand it is only for special types of models (so-called exponential families), and particular parameters (the so-called mean parameter) in those models, for which the Cramér-Rao lower bound $(ni(\theta; M))^{-1}$ for the variance of an unbiased estimator can be achieved exactly (for fixed n, in particular, $n = 1$), uniformly in θ. In fact a similar result can be proved for quantum models; see Barndorff-Nielsen, Gill and Jupp (in preparation).

Note that the classical information based on n independent and identically distributed realisations from a given density $p(x, \theta)$ is equal to n times the information for one realisation. Similarly, the quantum information in the state $\rho(\theta)^{\otimes n}$ corresponding to n identical particles each in state $\rho(\theta)$ is easily found to be n times the quantum information for one particle.

In view of these facts the question is therefore: does there exist a measurement procedure (not depending on θ) on the state $\rho^{\otimes n}$, on the basis of which an estimator of θ can be constructed having asymptotic variance $(nI(\theta))^{-1}$? If the answer is 'yes', then $I(\theta)$ is not just an upper bound on $i(\theta; M)$ but in an asymptotic sense the least upper bound, hence Braunstein and Caves' proposed role for the quantum information $I(\theta)$ in defining a statistical distinguishability metric is well motivated.

It seems rather natural to try the two-stage procedure: first estimate the parameter using a perhaps inefficient procedure on a vanishing proportion of the particles, say $\log n$ or n^α $(0 < \alpha < 1)$ out of the total of n; now carry out the 'estimated optimal measurement' on the remaining ones. In both stages only simple measurements (i.e., measurements of classical observables) on separate particles are needed.

In our example this would reduce to the following. Measure the spin σ_x on $k = \frac{1}{2}\log n$ of the copies. The number of '+1' observed is binomially distributed with parameters k and $p = \frac{1}{2}(\sin\eta\cos\phi + 1)$. Similarly for another k measurements of the spin σ_y we get a binomial number of '+1' with parameters k and $p = \frac{1}{2}(\sin\eta\sin\phi + 1)$. This allows us consistent estimation of both $\sin\phi$ and $\cos\phi$ and hence of $\phi \in [0, 2\pi)$. Denote such an estimator by $\widetilde{\phi}$. We saw that $\mathbb{D}\rho$ in this example was proportional to the spin in the direction $\pi/2, \phi + \pi/2$. Let us use the remaining $n - \log n$ particles to measure this spin with ϕ replaced by $\widetilde{\phi}$. Given $\widetilde{\phi}$, this results in a binomial number X of '+1' with parameters $n' = n - \log n$ and $p = \frac{1}{2}(1 - \sin\eta\sin(\phi - \widetilde{\phi}))$. Let

$$\widehat{\phi} = \widetilde{\phi} + \arcsin((n' - 2X)/(n'\sin\eta)).$$

Analysis of this 'final' estimator shows that $\widehat{\phi}$ has asymptotically the $\mathcal{N}(\phi, (n\sin^2(\eta))^{-1})$ distribution (the normal distribution with indicated mean and variance), whatever ϕ, so that the quantum information bound is asymptotically achievable by our two stage procedure.

This approach will work in wide generality in problems with a one-dimensional parameter θ. Suppose one has a consistent estimator $\widetilde{\theta}$ based on measurements on a vanishing proportion of the particles, which will typically be possible. Compute the quantum score at this point, and measure it on each of the remaining particles. Compute the maximum likelihood estimator $\widehat{\theta}$ of θ based on the new data, whose probability distribution depends on the unknown θ and the known $\widetilde{\theta}$. We argue that $\widehat{\theta}$ has approximately the $\mathcal{N}(\theta, (nI(\theta))^{-1})$ distribution, thus this estimator asymptotically achieves the quantum information bound. Let $i(\theta; \widetilde{\theta})$ denote the Fisher information for θ in a measurement on one particle of the quantum score at $\widetilde{\theta}$; thus $i(\widetilde{\theta}; \widetilde{\theta}) = I(\widetilde{\theta})$ for all values of $\widetilde{\theta}$, but generally $i(\theta; \widetilde{\theta}) < I(\theta)$. Now for n large, $\widetilde{\theta}$ is close to θ. By the classical results for maximum likelihood estimators, given $\widetilde{\theta}$, $\widehat{\theta}$ has approximately the $\mathcal{N}(\theta, (ni(\theta; \widetilde{\theta}))^{-1})$ distribution. So if ρ depends on θ smoothly enough that $i(\theta; \widetilde{\theta})$ is close to $i(\theta; \theta) = I(\theta)$ for $\widetilde{\theta}$ close to θ, we have that unconditionally $\widehat{\theta}$ has approximately the $\mathcal{N}(\theta, (nI(\theta))^{-1})$ distribution, hence asymptotically achieves the bound.

The situation is rather unclear when there are several unknown parameters. However in the spin-$\frac{1}{2}$ situation with both η and ϕ unknown, the same appears to hold: asymptotically, a two-stage procedure of measurements of classical observables on separate particles can achieve maximum information.

The quantum scores for the two parameters η and ϕ are $\sigma_{\eta+\pi/2,\phi}$ and $\sin\eta\,\sigma_{\pi/2,\phi+\pi/2}$ respectively. After a small proportion of measurements we know roughly the location of the parameter, and it is sufficient to investigate optimal measurement at a 'known' parameter value.

More specifically, consider (essentially without loss of generality) the special point $\eta = \pi/2$, $\phi = 0$. At this point the quantum scores are σ_y and $-\sigma_z$, and the quantum information matrix is the identity $\mathbf{1}$. The arguments of Braunstein and Caves do not appear to extend to the multi-parameter case, but the quantum Cramér-Rao bound does hold also for the multi-parameter case, [3], [4], with the inverse of the quantum information matrix being a lower bound to the variance-covariance matrix of an unbiased estimator of (η, ϕ) based on the outcome of a single measurement M. However there is not a single measurement whose probability distribution has Fisher information matrix for (η, ϕ) equal to the quantum information matrix, since by our results it would have to be of the form $m(x) = |\xi\rangle\langle\xi|$ with $|\xi\rangle$, up to a phase, equal to

$\alpha|\uparrow\rangle+\beta|\downarrow\rangle$ with α and β real and α non-zero for attainability of the ϕ component of the information, while by a similar calculation for η (for which the quantum score is $i|\uparrow\rangle\langle\downarrow| - i|\downarrow\rangle\langle\uparrow|$) m should again be rank-one but now with $|\xi\rangle$, up to a phase, of the form $\alpha'|\uparrow\rangle+i\beta'|\downarrow\rangle$ with α' and β' real and α' non-zero, which is only possible if $\beta=\beta'=0$. Though $m(x)$ can have this form for some x it is impossible for it to be true for all, since $\int m(x)\mu(\mathrm{d}x)=\mathbf{1}$.

Since no measurement attains simultaneously full quantum information for η and ϕ, at a given parameter point, but separately this is possible, we see that the class of Fisher information matrices for an arbitrary measurement on the spin-$\frac{1}{2}$ system does not include its least upper bound (the identity matrix $\mathbf{1}$). This means that for different loss functions, different repeated measurements will be optimal. An appealing loss function is one minus the squared inner-product between the true state vector and its estimate. This equals one minus the squared cosine of half the angle between the points on the Poincaré sphere representing the two states. At the special point under consideration therefore, the loss function is asymptotically equivalent to one quarter times the sum of the squares of the errors in η and ϕ.

Massar and Popescu [5], in response to a problem posed by Peres and Wootters [6], exhibited a measurement, optimal in the Bayes sense, with respect to this loss function and a uniform prior distribution. It had an asymptotic mean square error $4/n$. This was a genuine generalised measurement of the composite system $\rho^{\otimes n}$. They showed that for the case of $n=2$ there were no measurement methods of the two particles separately which were as good as the optimal method.

However, consider taking with probability half measurements of σ_y and σ_z, independently on each particle. We find that the Fisher information matrix for η,ϕ at $\eta=\pi/2$, $\phi=0$ is $\frac{1}{2}\mathbf{1}$. Therefore $(\frac{1}{2}\mathbf{1})^{-1}/n=2\mathbf{1}/n$ is an asymptotically achievable lower bound, at the point under consideration, for the covariance matrix of (asymptotically unbiased) estimators of η,ϕ based on n of such measurements. The maximum likelihood method would provide estimators asymptotically achieving this bound. The sum of the variances is $4/n$. This strongly suggests that a two-stage procedure similar to what we proposed in the one-parameter case can asymptotically achieve Massar and Popescu's mean square error of $4/n$, using simple measurements of single particles only. More explicitly, first carry out measurements of each of σ_x, σ_y and σ_z on a small proportion of separate particles. Compute from the results a consistent estimate of η,ϕ. Now rotate the coordinate system so that the estimated value is at $\eta=\pi/2$, $\phi=0$, and measure alternately σ_y and σ_z on the remaining particles. Estimate η,ϕ (new coordinate system) by the method of maximum likelihood using the second stage observations, and finally transform back to the original coordinate system using the first estimates.

The quantum information matrix was in this case equal to $\mathbf{1}$ itself. This implies a lower bound to the covariance matrix of unbiased estimators of η,ϕ of $\mathbf{1}^{-1}/n=\mathbf{1}/n$ (at the special point under consideration before). This bound applies to estimators based on arbitrary measurements of the n particles as a single system. The sum of the mean square errors cannot be less than the trace of this matrix, $2/n$. By analogy with classical statistical theory one should expect the same bound to apply to the wider class of asymptotically unbiased estimators. However the Massar and Popescu results suggest that the actual asymptotic bound is $4/n$, asymptotically achieved by their optimal Bayes procedure based on a generalised measurement of the n particles as a single system, but also by our two-stage procedure of separate measurements of classical observables. Recent work of Gill and Massar has confirmed this bound, and extended the results to arbitrary loss functions and, to some extent, more general estimation problems.

To conclude, in the multiparameter case, the bound implied by the quantum information matrix is not even asymptotically achievable. Hence the role of the quantum information matrix in multiparameter problems is rather less fundamental than in the one-parameter case. The derivation in more general examples of an asymptotically attainable lower bound for any given loss function is a challenging open problem.

Acknowledgements

We are grateful for useful discussions with Peter E. Jupp and Klaus Mølmer. This research was supported by the Danish National Research Foundation and the Oberwolfach Mathematical Institute's Research in Pairs programme. Richard Gill is also grateful for the hospitality of the Department of Mathematics and Statistics, University of Western Australia.

References

1. S.L. Braunstein and C.M. Caves, Phys. Review Letters **72**, 3439 (1994).
2. A. Fujiwara and H. Nagaoka, Phys. Letters A **201**, 119 (1995).
3. C.W. Helstrom, *Quantum Detection Theory* (Academic, New York, 1976).
4. A.S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam, North-Holland, 1982).
5. S. Massar and S. Popescu, Phys. Review Letters **74**, 1259 (1995).
6. A. Peres and W.K. Wootters, Phys. Review Letters **66**, 1119 (1991).