Discussion on "On quantum statistical inference" by Barndorff-Nielsen, Gill and Jupp

Luigi Accardi, Dorje C. Brody, V.P. Belavkin, Johyn T. Kent, N.H. Bingham, Jeremy G. Frey, Inge S. Helland, Jan-Åke Larsson, NK Majumdar, Marco Minozzo and J. W. Thompson

Applications: Festschrift for Constance van Eeden (eds M. Moore, S. Froda and C. Léger). Beachwood: Institute of Mathematical Statistics.
Gill, R. D. and Massar, S. (2000) State estimation for large ensembles. *Phys. Rev.* A, **61**, 2312–2327.
Gilmore, R. (1994) *Alice in Quantum Land*. Wilmslow: Sigma.
Hannemann, T., Reiss, D., Balzer, C., Neuhauser, W., Toschek, P. E. and Wunderlich, C. (2002) Self-learning estimation of quantum states. *Phys. Rev.* A, **65**, 050303-1–050303-4.
Helstrom, C. W. (1976) *Quantum Detection and Information Theory*. New York: Academic Press.
Holevo, A. S. (1982) *Probabilistic and Statistical Aspects of Quantum Theory*. Amsterdam: North-Holland.
Holevo, A. S. (2001a) Statistical structure of quantum theory. *Lect. Notes Phys. Monogr.*, **67**.
Holevo, A. S. (2001b) Lévy processes and continuous quantum measurements. In *Lévy Processes—Theory and Applications* (eds O. E. Barndorff-Nielsen, T. Mikosch and S. Resnick), pp. 225–239. Boston: Birkhäuser.
Isham, C. (1995) *Quantum Theory*. Singapore: World Scientific.
Keyl, M. and Werner, R. F. (2003) Estimating the spectrum of a density operator. *Phys. Rev.* A, **64**, 052311-1–052311-5.
Kraus, K. (1983) States, effects and operations: fundamental notions of quantum theory. *Lect. Notes Phys.*, **190**.
Leonhardt, U. (1997) *Measuring the Quantum State of Light*. Cambridge: Cambridge University Press.
Loubenets, E. R. (2001) Quantum stochastic approach to the description of quantum measurements. *J. Phys.* A, **34**, 7639–7675.
Malley, J. D. and Hornstein, J. (1993) Quantum statistical inference. *Statist. Sci.*, **8**, 433–457.
Matsumoto, K. (2002) A new approach to the Cramér-Rao-type bound of the pure state model. *J. Phys.* A, **35**, 3111–3123.
Meyer, P.-A. (1993) Quantum probability for probabilists. *Lect. Notes Math.*, **1538**.
Mølmer, K. and Castin, Y. (1996) Monte Carlo wavefunctions. *Coher. Quant. Opt.*, **7**, 193–202.
Mooij, J. E., Orlando, T. P., Levitov, L., Tian, L., van der Wal, C. H. and Lloyd, S. (1999) Josephson persistent-current qubit. *Science*, **285**, 1036–1039.
Nielsen, M. A. and Chuang, I. L. (2000) *Quantum Computation and Quantum Information*. New York: Cambridge University Press.
Ozawa, M. (1985) Conditional probability and a posteriori states in quantum mechanics. *Publ. RIMS Kyoto Univ.*, **21**, 279–295.
Percival, I. (1998) *Quantum State Diffusion*. Cambridge: Cambridge University Press.
Peres, A. (1995) *Quantum Theory: Concepts and Methods*. Dordrecht: Kluwer.
Weihs, G., Jennewein, T., Simon, C., Weinfurter, H. and Zeilinger, A. (1998) Violation of Bell's inequality under strict Einstein locality conditions. *Phys. Rev. Lett.*, **81**, 5039–5043.
Williams, D. (2001) *Weighing the Odds*. Cambridge: Cambridge University Press.
Wiseman, H. M. (1999) Adaptive quantum measurements (summary). In *Miniproc. Wrkshp Stochastics and Quantum Physics*, miscellanea, 14. Aarhus: University of Aarhus.
Young, T. Y. (1975) Asymptotically efficient approaches to quantum-mechanical parameter estimation. *Inform. Sci.*, **9**, 25–42.

Discussion on the paper by Barndorff-Nielsen, Gill and Jupp

Luigi Accardi (*Università di Roma "Tor Vergata"*)
First I thank the Royal Statistical Society for inviting me to propose the vote of thanks to Ole Barndorff-Nielsen, Richard Gill and Peter Jupp for their paper. This is something that I do with great pleasure because quantum probability has originated a major change in the landscape of modern probability theory and it is time that the implications of this radical innovation for statistics begin to be felt. The first attempts to interest statisticians in quantum probability go back to the mid-1980s and I would like to devote a thought to the memory of Jeff Watson who was historically the first statistician to become deeply involved in this field. I hope that now, after 20 years, the time is riper for this potentially historic operation.

This paper is on quantum statistics and probably most of the audience are aware of the strong and, at the same time touchy, connections between classical probability and classical statistics. Therefore some of you might be amused to confirm that the tradition continues even if you add the adjective 'quantum' to both disciplines.

I shall try to frame the present paper within the general context of quantum probability first from a conceptual point of view and then from a more mathematical point of view. Since time is short, I have chosen to proceed by short questions and schematic answers whose main role should be to stimulate further questions.

What is quantum probability?
Quantum probability is a non-Kolmogorovian probabilistic model whose role in probability is analogous to the role of non-Euclidean models in geometry. Algebraic probability theory includes both the quantum

and the classical probabilistic model, just like differential geometry includes both the Euclidean and the non-Euclidean model.

Do we need quantum probability?
We need quantum probability in the same sense as we need non-Euclidean models. If you have three non-Euclidean angles, you are not obliged to describe them by means on a non-Euclidean triangle. But in some cases this is useful both for intuition and for doing calculations. The same happens for probability.

Can we characterize the Kolmogorovianity or non-Kolmogorovianity of a given set of statistical data?
Yes, we can characterize the Kolmogorovianity by means of the statistical invariants, which are the probabilistic counterpart of the geometrical invariants, like curvature.

Beyond the quantum model are there other non-Kolmogorovian models?
There are infinitely many possibilities and a full classification is out of reach. It is just the same as in geometry.

The empirical evidence supporting the quantum model is impressive: is there any empirical evidence of emergence of non-Kolmogorovian models outside quantum physics, e.g. in biology, medicine, economics,...?
I am sure that there are but I have not yet found any.

Can you explain intuitively from where the non-Kolmogorovian statistics arise?
There are two main sources of differences: one is the existence of incompatible (non-simultaneously measurable) observables; the other is that some inductive arguments (typically the counterfactual argument) can be applied to passive but not to adaptive systems. There are many examples of incompatible events in the classical world. For a classical system the prototype of a passive property is the colour of a ball in a box and the prototype of an adaptive property is the colour of a chameleon in a box.

What is the mathematical difference between a quantum model and a Kolmogorovian model?
A quantum model is an infinite bunch of Kolmogorovian models (abelian algebras) glued together to produce a single new mathematical structure (non-abelian algebra). If you restrict a quantum model on an abelian subalgebra you obtain a Kolmogorovian model. Keeping fixed the quantum state and varying the abelian subalgebra, you obtain infinitely many Kolmogorovian models. These uniquely determine the quantum model (Gleason's theorem). Quantum tomography studies the following question: given a quantum state, what is a minimal set of its Kolmogorovian restrictions that is sufficient to determine it?

What, in classical probability, plays the role of the 'parallel axiom' in classical geometry?
Bayes's definition of conditional probability, which, in fact, is equivalent to five independent axioms, plays the role of the parallel axiom. For each of these axioms one can find counter-examples showing that there are physical situations where this axiom is not physically justified.

How does the difference between the quantum and the classical model emerge in the positive operator valued measures discussed in this paper?
The difference between the quantum and the classical model emerges through the theory of composite systems and the theory of quantum conditioning, which is the most important point where the parallelism between classical and quantum models breaks down.

How do positive operator valued measures arise?
Consider a composite system $S \times S_{\mathrm{M}}$ (where S denotes a system and S_{M} the measurement apparatus) and apply the same construction, used in the answer to the question about the mathematical difference between a quantum and a Kolmogorovian model, not to the whole composite system, but only to the apparatus. This means that a quantum conditional expectation (rather than a state) is restricted to an abelian subalgebra of the apparatus but the algebra of the system is not changed. This gives an object which is half classical and half quantum. This is a positive operator valued measure. The classical analogue of a positive operator valued measure is a familiar Markov transition probability.

Dorje C. Brody (*Imperial College, London*)
I agree with the authors concerning the importance of extending statistical data analysis to the results of quantum measurements. The fact that statisticians are now addressing this issue is quite encouraging, although I have some reservations regarding the presentation of the subject.

The authors note that the paper is intended to introduce quantum inference problems to the statistical community. However, I am slightly concerned by their starting from the relatively difficult concept of mixed state density matrices, without a clear exposition of the significance of pure states (wave functions).

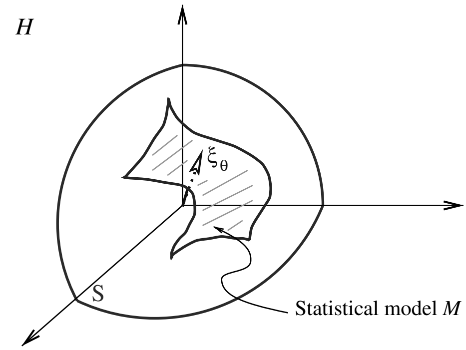

Fig. 1. Square-root density functions correspond to points in the positive orthant of the unit sphere $\mathcal{S}$ in Hilbert space $\mathcal{H}$: for a parametric density function, the state lies on a submanifold $\mathcal{M}$ of $\mathcal{S}$ with metric given by the Fisher information matrix

This might convey the impression, to readers who are unfamiliar with quantum mechanics, that the essence of quantum inference is simply the replacement of the probability density function by a density matrix. Hence, I shall, instead, focus on the nature of pure states.

Probability theory typically involves a density function $p(x|\theta)$ which may depend on a set of parameters $\{\theta\}$. In statistics, we often consider the log-likelihood function $l_\theta(x) = \ln\{p(x|\theta)\}$. However, in physics we naturally work with the square-root density function $\xi_\theta(x) = \surd p(x|\theta)$, which permits a formulation of the problem in a real Hilbert space context. This was recognized by Mahalanobis (1930, 1936) and Bhattacharyya (1943, 1946). In particular, if the state $\xi_\theta(x)$ depends on a set of parameters, then the statistical model is represented by a finite dimensional submanifold M of the unit sphere $S \subset H$, where the Fisher information determines the Riemannian metric on M (Rao, 1945, 1947). See Fig. 1. This leads to the notion of information geometry (see Burbea (1986)).

Note, however, that to recover the probability we must square the state ξ, indicating that we could have defined $\xi = -\surd p$ as the negative root of p. Thus, the relevant state space consists of equivalence classes, i.e. rays through the origin of H. Hence, all aspects of classical probability can be characterized by geometric features of real projective spaces (Brody and Hughston, 1999). This idea can be extended by allowing the state ξ to be a complex-valued function, the probability being defined by $p = |\xi|^2$. Such a complex-valued function is what physicists call a wave function or a pure state.

However, complex-valued states are considered in quantum mechanics not merely because this remains consistent with probability laws but rather because they are indispensable for the study of quantum systems. The role of complex numbers in quantum mechanics is twofold: first, a complex number rotates a state by the angle $\frac{1}{2}\pi$. Second, it determines the 'direction of time' via Schrödinger's equation. The latter point implies that quantum mechanics is not a static but a dynamic theory. These two features are relevant to the estimation problem that is discussed below.

Now, if the quantum state depends on a set of unknown parameters, then the various problems of statistical inference can be cast in a Hilbert space setting *à la* Efron. The inference problem for pure states has been extensively studied in the literature (see, for example, Burbea and Rao (1984), Caianiello (1983), Caianiello and Guz (1988) and Brody and Hughston (1996, 1997, 1998); see also Yuen *et al.* (1975), Helstrom (1982) and Brody and Meister (1996) for further work on quantum hypothesis testing that extends earlier results by Helstrom and Holevo). Citation of these results would make the paper more informative to statisticians who are familiar with information geometry and asymptotic inference or standard hypothesis testing.

To illustrate the subtlety of pure state estimation, consider the inference problem in Section 6 (example 7). Given a pair of spin half eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$, a generic pure state is expressible as $|\xi(\theta,\phi)\rangle = \cos(\frac{1}{2}\theta)|\uparrow\rangle + \sin(\frac{1}{2}\theta)\exp(\mathrm{i}\phi)|\downarrow\rangle$. The value of θ is assumed fixed, and we seek to estimate the unknown parameter ϕ. However, unlike the problems in classical estimation theory, where the state remains static, the complex square-root density function $|\xi\rangle$ *evolves* in time (the second role of the complex number). More precisely, only an energy eigenstate can be static in quantum physics.

In this example, unitary evolution consists in rigid rotation of the state space around the axis defined by the spin eigenstates (Fig. 2). Therefore, the value of θ will remain fixed, whereas the value of ϕ, to be estimated, varies according to $\phi = \omega t \pmod{2\pi}$, where ω is the difference in energy of the two eigenstates.

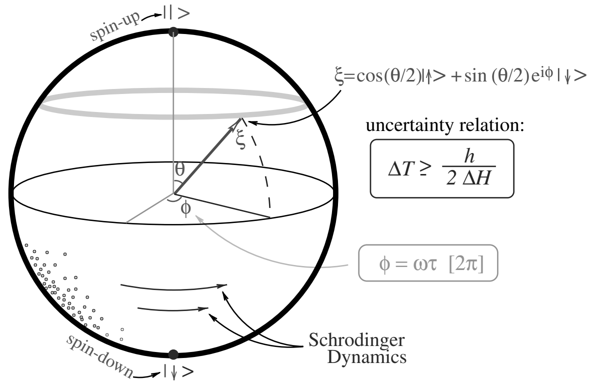

Fig. 2. Spin half state space: a generic state is expressible as a superposition of the spin-up and spin-down eigenstates; Schrödinger evolution corresponds to the rigid rotation of the sphere around the poles determined by the eigenstates; the 'speed' of the Schrödinger trajectory is determined by the Anandan–Aharanov relation; owing to the energy–time relation, the lower bound on time precision precludes the 'preparation' of a state with a definite relative phase ϕ

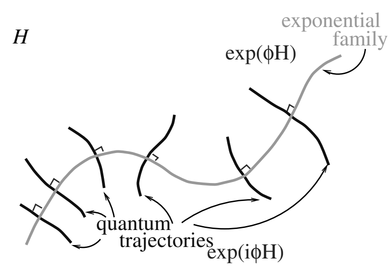

Fig. 3. One-parameter family of exponential curves and one-parameter families of unitary trajectories: the unitary trajectories are everywhere orthogonal to the corresponding exponential family; consequently, the Cramér–Rao lower bound is never attained for the estimation of the unitary trajectory

In this situation, the proper formulation of an estimation problem appears obscure. This can be remedied by prescribing that the time parameter t is the object of estimation, but then the results concerning quantum exponential families and the attainability of the Cramér–Rao bound no longer apply.

To see this, apply the formulation of Efron (1975) to the square-root density function rather than to the log-likelihood function. Then, as illustrated in Fig. 3 (see Brody and Hughston (1996) for the proof), if the exponential family admits the form $\exp(\phi H)$ where H is a random variable, the trajectories of the one-parameter family of unitary transformations $\exp(\mathrm{i}\phi H)$ are everywhere orthogonal to the corresponding exponential family (first role of the complex number). Therefore, although the quantum exponential families that were investigated by the authors presumably are relevant in some physical contexts, the authors have yet to demonstrate this in specific examples.

As for mixed states, these are just (classical) probability distributions $\rho(\xi)$ over the space of complex square-root likelihood functions. The density matrix $\hat{\rho}$ is defined by the first quadratic moment of the distribution: $\hat{\rho} = \int |\xi\rangle\langle\xi|\rho(\xi)\,\mathrm{d}\xi$, where the integration extends over the pure state space. For standard linear observables, the information contained in $\hat{\rho}$ suffices to determine all statistical properties of the system. Thus, an alternative approach to quantum inference would be to take the classical results that are readily applicable to $\rho(\xi)$, and then to project out the component of the first moment to restore the corresponding quantum results. An apparent paradox arises in that, if $\rho(\xi)$ is a classical exponential family, then the corresponding density matrix $\hat{\rho}$ is *not* an exponential family as defined in the paper. I hope that the authors will further investigate this issue in their forthcoming book on the subject.

The vote of thanks was passed by acclamation.

V. P. Belavkin (*University of Nottingham*)
The development in the early 1970s of the statistical inference theory initiated in Helstrom (1967), which is based on the non-Kolmogorov quantum probability, not only opened a new direction in statistics but also helped to clarify some problems of quantum theory. It made it possible to introduce the concept of an approximate measurement of incompatible observables described by non-commuting operators and enabled the solution of various problems in the quantum theory of information and communication (Belavkin and Grishanin, 1973; Belavkin, 1975a), e.g. to give for pure states a precise formulation of a generalized Heisenberg uncertainty principle for quantities such as the time and energy, or phase and number of quanta (Helstrom, 1973, 1974), and to define precisely what is a measurement of the time and phase in quantum mechanics (Belavkin, 1975b). The present paper gives a nice elementary introduction to the subject of quantum statistical inference which is particularly suitable for statisticians as well as for physicists and it reviews mostly recent advances into this subject. However, it leaves many interesting questions raised by that development untouched; see for example Gill and Massar (2000) and Petz (2003). It would also be useful to have an overview of the achievements of the earlier development of quantum estimation and hypotheses testing theory (Belavkin, 1972, 1975c) and to review formulations and results obtained from the more contemporary point of view of quantum probability and mathematical quantum theory. Some of the references on quantum statistics that have been omitted are Stratonovich (1973), Yuen and Lax (1973) and Belavkin (1976), and these may to some extent complement this review.

John T. Kent (*University of Leeds*)
The paper has offered a fascinating glimpse into the world of quantum probability and associated questions of statistical inference. I shall explore some links to other branches of statistics, especially shape analysis. For background information, see, for example, Dryden and Mardia (1998), especially chapters 4, 6 and 9. The basic building-block in quantum probability is the set of $d \times d$ Hermitian, positive definite complex matrices ρ with trace 1. In addition we can impose a rank restriction, $\text{rank}(\rho) = m$, for some fixed m, $1 \leqslant m \leqslant d$.

When $m = 1$, there is a close connection to the representation of the shape of a configuration of $d + 1$ points or landmarks in two-dimensional real space. Such a configuration can be represented as a $(d + 1)$-dimensional complex vector. Recall that the shape of a configuration consists of the information invariant under changes in location, scale and rotation. Location effects can be removed by taking a set of d orthonormal contrasts between the landmarks, yielding a d-dimensional complex vector $\boldsymbol{\mu}$, say. Size effects can be removed by scaling $\boldsymbol{\mu}$ to have unit size, $\boldsymbol{\mu}^*\boldsymbol{\mu} = 1$. Rotation, i.e. the equivalence of $\boldsymbol{\mu}$ and $\exp(\mathrm{i}\psi)\boldsymbol{\mu}$, $0 \leqslant \psi < 2\pi$, can be removed by focusing on the rank 1 matrix $\rho = \boldsymbol{\mu}\boldsymbol{\mu}^*$.

A similar representation holds for $m \geqslant 1$ if we shift attention to the slightly modified concept of 'reflection shape'. In this case the shape information can be coded by a *real* $d \times d$ matrix ρ of rank m. This representation forms the basis for classical multidimensional scaling.

In the setting $m = 1$ and ρ complex, we can ask whether it makes sense to consider inference from a Bayesian point of view, i.e. to put a prior on ρ or equivalently on $\boldsymbol{\mu}$. If so, there are several choices from shape analysis obtained by either conditioning or projecting suitable multivariate normal distributions on shape space, e.g. the complex Bingham distribution, the complex angular central Gaussian distribution and the Mardia–Dryden distribution. In principle, these models can be extended to the cases $m \geqslant 1$ and/or ρ real, and some work has already been done in the reflection shape context.

Conversely, the quantum exponential and transformation models of Section 4 may have a useful role in providing models for systematic changes of shape in shape analysis. So far, the available models have either been limited to simple changes (e.g. a great circle path on a sphere) or to small scale changes, in which case standard multivariate techniques can be used on a tangent plane to shape space.

Finally, it would be interesting to know the status of the statistical models in the paper. To what extent are they *empirical* (i.e. chosen for their tractability and convenience) and to what extent are they *scientific* (i.e. governed by some underlying physical principles)?

Dorje C. Brody (*Imperial College, London*)
Professor Kent has pointed out an apparent resemblance between the statistical theory of shape and certain aspects of quantum theory. This can be made more precise by noting that the shape space of k labelled points on a plane can be identified with the space of pure quantum states of dimension $k - 2$. However, this correspondence applies only to planar configurations, and not those in higher dimensions. Nevertheless, this suggests the possibility of applying quantum theoretical methods to the statistical theory of shapes, or conversely. Further details on this correspondence are investigated elsewhere (Brody, 2003).

Professor Kent has also queried the relevance of mixed states in this context. Each mixed state density matrix in quantum mechanics represents an equivalence class of distributions of different point configurations. For example, for a Markovian diffusion of shapes, this leads to a Lindblad-type equation for the deterministic evolution of the density matrix. Professor Kent's comment thus leads to an important question, namely to what extent would density matrices suffice in representing statistical information concerning the distribution of different shapes?

N. H. Bingham (*Brunel University, Uxbridge*)
I very much welcome this paper, with its avowed intent of introducing the statistical community to quantum statistical inference. To my knowledge, the first such attempt at cross-fertilization between quantum physics (an intrinsically statistical theory) and statistics was the paper of Jose E. Moyal (1910–1998) in the 'Symposium on stochastic processes' on June 7th, 1949, organized by Moyal, Bartlett and D. G. Kendall (Moyal, 1949). Incidentally, Moyal's work of this period has undergone a recent revival (Moyal quantization), in contexts such as quantum groups and non-commutative geometry (Chari and Pressley (1994), pages 2–3, Connes (1994) and Gani (1998)). It would be interesting to see such work related to this paper.

The theory of Section 6, on quantum and classical Fisher information, is very attractive. Ideas of information in statistics go back to Fisher, to Cramér and Rao, to Kullback and others (Barndorff-Nielsen (1978), which is modestly uncited); see Barndorff-Nielsen and Cox (1994) and Severini (2000). A textbook exposition of the extension of this classical (and modern) theory to the quantum–non-commutative context would be valuable, and I look forward to the authors' forthcoming book on this score.

In Sections 7 and 8.2, the authors allude to some of the extensive literature on quantum and free probability. It would be a valuable service to the probability community to see more details on the relationship between the ideas of this paper and those of, *inter alia*, Meyer (1995) and Voiculescu (2000).

Quantum stochastic calculus is now recognized as a field in its own right (classification 81S25 in the *Mathematical Reviews* classification scheme). One wonders what it may have to offer here.

The following contributions were received in writing after the meeting.

Jeremy G. Frey (*University of Southampton*)
As a Physical Chemist involved in experimental and theoretical spectroscopy the phrase quantum statistics triggers ideas of Bose–Einstein and Fermi–Dirac statistics, and statistical mechanics. As practising chemists we usually view the rules of quantum mechanics as a recipe for deriving a probability, which I think we then frequently treat somewhat classically unless we are involved in a series of experiments on the same object, in which case as the paper makes clear the result of the measurement in collapsing the system state makes for classically unexpected results.

The significant developments in quantum information often pass us by other than interests in the Einstein–Podolsky–Rosen 'paradox' and thus Bell's inequalities and the Aspect experiments. The mathematical developments in quantum probability are beyond our normal purview but the interaction of quantum probabilities with classical statistics is a significant issue as most of our experiments and apparatus are not 100% efficient and suffer from background and other effects. As I understand it this means that the experiments that test Bell's inequalities, and other photonics experiments that were referred to in the paper and the presentation, are not 'perfect' but are statistical demonstrations of the 'truth' of the underlying quantum relationship. If this is so then it would have been very useful to have shown some data from these 'simple' two-state experiments to provide more insight into the working of the theoretical framework that is provided in the paper.

The rise of quantum computing (and the use of nuclear magnetic resonance in this regard) is bringing this area into much higher view. In this area the limitations in the application of different experimental systems to quantum computing seem to arise from the loss of coherence in the systems, due to interactions with the 'outside'. Can this loss of coherence and the way that it influences the carriage of quantum information be viewed as applying or supplying an increased level of uncertainty to the quantum system, i.e. as an externally supplied almost classical randomness and thus amenable to standard statistical treatment (to investigate repeats of the experiment), or should or must it be included as part of the quantum level description? Another area of practical concern to spectroscopists is how much of the statistics that is typically applied to a large ensemble of molecules can be applied to the work undertaken on single molecules? I wonder whether this work can help in this regard.

Inge S. Helland (*University of Oslo*)
The authors are to be congratulated on a very interesting paper on an extremely interesting topic: the inter-

play between quantum theory and ordinary statistical theory. This area has moved forward considerably since Max Born first suggested his statistical interpretation of quantum mechanics, which was nothing but the now well-known probabilistic interpretation of the modulus squared of the wave function. As I see it, it should be possible to move even further along this route: quantum theory *is* a statistical theory. It is not 'mechanics' limited to describing the movement of particles. It is capable of doing probabilistic prediction of essentially every phenomenon in the microworld, just as ordinary statistics is capable of modelling and predicting essentially every phenomenon in the macroworld.

The relationship between quantum theory and statistics has been discussed by many, most recently by Loubenets (2003) in an abstract common modelling and inference framework. To develop this link further in a more direct setting, an obvious task would be to extend the ordinary statistical paradigm, based as it is essentially only on models as classes of probability measures. This has already been proposed for completely different reasons by McCullagh (2002). Some obvious candidates for such extensions are symmetry by using group theory, diversity of evidence (Barndorff-Nielsen, 1995) by using complementary parameters and complementary experiments and model reduction. In Helland (2003a, b) attempts are made to take quantum theory along such directions. Although there are still unsettled questions, the results definitely seem sufficiently promising to deserve further discussion. In my view, it should be possible ultimately to find a quantum theory where formal elements are derived, not assumed in the axiomatic basis. This may also give some unified theory of inference, where parts can be useful also for other applied disciplines.

Jan-Åke Larsson (*Linköping University*)
Regarding Section 7, 'Classical *versus* quantum', it is true that from an abstract point of view a basic structure in probability theory is isomorphic to a special case of a basic structure in quantum probability, and that this entails a rather narrow (functional analytic) view of classical probability. It is also true that the basic structure in quantum probability is isomorphic to a particular case of the basic structure of classical statistical inference. I would like to argue (tongue in cheek) that this claim entails a rather narrow (parameter-model-oriented) view of quantum probability. Moreover, many quantum probabilists will feel that *only* using a family of classical probability models with sufficiently many parameters is 'discarding the key feature of ordinary probability theory' since this broader mathematical structure has no analogue of the Hilbert space, and hence no opportunity for a quantum probabilist's beloved co-ordinate transformations.

The parameter-model-oriented mathematical structure contains quantum probability and much more; it is, in its pure form, too general. Of course, Barndorff-Nielsen, Gill and Jupp *are* aware of this; the paper uses the Hilbert space formalism to find appropriate model families and notions that are useful for statistical inference. Here, the parameter-model-oriented approach is *the* tool and therefore from this point of view quantum probability *does* belong to the field of classical statistics. However, Sections 2–6 suggest that this claim is true only after having used the quantum language to establish *classical* parameters in some family of quantum models.

Recall that the 'difference' between quantum and classical is only visible in situations where the influence of the parameters is restricted for some physical or philosophical reason. For instance, the Bell inequality that is mentioned in the paper requires that the parameter influence is spatially local. This imposes certain restrictions on any classical probability model for the system, which is violated by the quantum mechanical predictions. These restrictions are not violated by current experiments because of practical limitations, so the experimental data allow a classical (*ad hoc*) probabilistic model; see Larsson (1998, 1999). An example of a consequence is the existence of a Trojan horse in the supposedly secure Bell-inequality-based quantum cryptography (Larsson, 2002).

In any case, restrictions of this type are not to be found in the parameter models that are discussed in the paper. Thus, regardless of any 'difference' between quantum and classical, only classical statistical properties emerge in this context. I would argue that this is a property of the formulation of the problem rather than of the underlying system.

N. K. Majumdar (*London*)
I would like to quote Karl Popper's statement that quantum theory is a statistical subject, which supports the paper. I also support the idea that quantum theory does not necessarily need a new definition of probability which is different from what the statistician is used to. All quantities of interest such as the probability density represented by $|\psi|^2$ or even by the probability current density can be fully understood by classical probability.

My next remark is concerned with the fact that the wave function formulation of quantum mechanics is more familiar to many. The wave function not only gives the position probability $|\psi|^2$ but can also be expanded into a probability distribution of eigenfunctions for each observable. So, we can find a probability distribution for each observable. A statistical study may conceivably be made of how the wave function combines, the probabilistic distribution of eigenvalues and the position probability densities as well as the deterministic evolution of ψ in time.

Marco Minozzo (*University of Perugia*)
My comments will focus on some foundational questions on the violation of Bell inequalities which although touched on marginally in this paper are of great concern also to the authors (Barndorff-Nielsen *et al.*, 2001; Gill, 2003a, b). Since Bell (1964), physicists are still looking for a loophole-free correlation experiment to settle definitively the questions surrounding the experimental violations that have been reported in the literature. Leaving aside the experiment by Aspect *et al.* (1982) where the periodic transition probabilities were a major concern (Aspect, 1999; Minozzo, 2000), here I would like to stress that for virtually all other experiments it seems to have been the (more or less explicit) assumption of 'rotational invariance' of the correlation measurements that led to the violations claimed. To my knowledge, this assumption has never been proved and the only empirical evidence that is available so far for the correlation measurements is for an (approximate) cosine-like law for some (not all) absolute orientations.

Let us consider the ideal setting of the Einstein–Podolsky–Rosen–Bohm *Gedankenexperiment* and assume that, for the gathering of the correlation measurements, four runs are performed in which the source is rotated alongside with the orientation of analyser A. To model this we need a family of stochastic processes $(\Lambda_{\varphi_s,n})$ for the source and two other families $(A_{\varphi_a,\varphi_s,n})$ and $(B_{\varphi_b,\varphi_s,n})$ for the analysers, with $\varphi_s, \varphi_a, \varphi_b \in [0, 2\pi)$, $n \in \mathbb{N}$. For the source in orientation φ_s, and analysers in orientation φ_a and φ_b, let, forgetting time, $\mu_{\varphi_s}(\varphi_a, \varphi_b) = E(A_{\varphi_a,\varphi_s,n} B_{\varphi_b,\varphi_s,n})$, $n = 1, 2, \ldots$. In Minozzo (2000) a classical (purely particle) toy model is considered for which $\mu_0(0, \varphi_b) = \cos(2\varphi_b)$, for $\varphi_s = \varphi_a = 0$, $\varphi_b \in [0, \pi]$, $\mu_{\varphi_s}(\varphi_a, \varphi_b) = 1$, for $\varphi_a = \varphi_b$, and $\mu_{\varphi_s}(\varphi_a, \varphi_b) = -1$, for $\varphi_b = \varphi_a + \pi/2$. This model, although agreeing with quantum mechanical predictions for some absolute orientations, is not rotationally invariant. For instance, for $\varphi = \varphi_b - \varphi_a \in [0, \pi]$, $\mu_0(0, \varphi) \neq \mu_0(\pi/4, \pi/4 + \varphi)$. (However, it holds that $\mu_{\varphi_s}(\varphi_a, \varphi_b) = \mu_{\varphi_s+\theta}(\varphi_a + \theta, \varphi_b + \theta)$, for $\theta \in [0, 2\pi)$.) For the Bell-type quantity of Clauser *et al.* (1969) we have

$$\begin{aligned} S &= \mu_{\varphi_a}(\varphi_a, \varphi_b) - \mu_{\varphi_a}(\varphi_a, \varphi'_b) + \mu_{\varphi'_a}(\varphi'_a, \varphi_b) + \mu_{\varphi'_a}(\varphi'_a, \varphi'_b) \\ &= \mu_0(0, \varphi_b - \varphi_a) - \mu_0(0, \varphi'_b - \varphi_a) + \mu_0(0, \varphi_b - \varphi'_a) + \mu_0(0, \varphi'_b - \varphi'_a), \end{aligned}$$

and, for orientations $\varphi_a = 0$, $\varphi'_a = \pi/4$, $\varphi_b = \pi/8$ and $\varphi'_b = 3\pi/8$, the model gives $2\sqrt{2}$ (the maximal violation that is expected by quantum mechanics). In other words, from a probabilistic point of view, we could say that the Bell theorem is just showing that the quantum mechanical cosine-like (rotationally invariant) law for the correlation measurements does not specify, in the probability framework of Kolmogorov, a coherent set of correlations.

J. W. Thompson (*University of Hull*)
Most, perhaps all, measurements that can be made in experiments give results which have a familiar form whether the experiment is intended to verify predictions from classical Newtonian mechanics or from quantum theory. We can record whether specified events occur or not and the values that quantities of interest have attained and we can code them for electronic storage. Whenever the theory investigated predicts a random outcome, then the experiment must be repeated so that, for example, relative frequencies can be compared with predicted probabilities. In short, the space of outcomes that are considered possible (the sample space) obeys Kolmogorov's finite axioms, at least. One of the most interesting aspects of quantum models is that the non-classical probability systems within them induce classical probability measures on the observable sample space. It is from this basic classical structure that inference about the underlying quantum probabilities must be drawn.

Of course, the space of quantum measures is very different. The probabilities are no longer classical and they evolve in a predictable fashion with time. Furthermore, potentially new issues of coherence between the outcome of an inference procedure and the 'true' underlying model generating the data arise strongly. In classical inference coherence was measured by a loss function, or equivalent, and then by consideration of the average behaviour of a proposed procedure when exposed to particular cases of the general model (the risk function). The authors have made a good start by proposing quantum analogues of concepts of information which have been very productive in classical inference. However, a re-examination of the notion of 'closeness' between inference and the underlying model seems necessary, particularly in the case of the testing of the adequacy of quantum predictions.

The role of small predicted probabilities could prove particularly important. Although some strange, even disastrous, event might be extremely unlikely in a single evaluation in a relatively simple quantum system, it could become almost certain in a complex linking of a very large number of such simple systems. Of course such possibly unpleasant eventualities can be reduced by introducing redundancy but the *scale* of the 'tail' probabilities will be crucial.

Finally, I congratulate the authors on such a most interesting and stimulating paper.

The **authors** replied later, in writing, as follows.

We are grateful to the discussants for their interesting and stimulating contributions. We find little to disagree with, but much worth commenting on.

Road-maps

An issue that is raised by several discussants (in particular, Professor Accardi and Dr Frey) is where our subject should be located on the map of various disciplines.

The first step is to translate key notions. As Dr Frey remarks, there are already problems with the title. 'Quantum statistics' could be the Bose–Einstein and Fermi–Dirac distributions of quantum ensembles, as opposed to classical multinomial (balls in boxes). Our focus is statistical inference when the *stochastic* contribution from quantum mechanics is important: a finite number of measurements on individual systems.

We are not finished after translating key terms and locating our problem area on an existing map. A scientific field has not just a theoretical framework but also motivations and aims. Professor Accardi rightly locates quantum statistical inference as the study of certain completely positive maps on certain abstract probability-like spaces. From his point of view, we are studying a special case of systems of the kind one studies in much more generality in quantum probability.

However, one mathematical framework being formally contained in another does not mean that one poses, in the smaller field, the same questions as in the larger. Typically, interesting questions for the 'smaller' field are meaningless in the larger. Consequently, the tools and results of the larger field, though possibly relevant, are not prerequisites for entering the smaller field, though it is useful to know that they are there (cf. probability as part of measure theory).

We appreciate, in this light, Dr Larsson's reaction to our suggestion that quantum probability is a special case of classical statistics. Our subject is multifaceted and one pair of spectacles does not suffice.

Eventually, quantum statistical inference might change the maps of adjacent and enveloping disciplines, as well as being fed by them. We hope that some of the experts from quantum probability would consider quantum statistical inference; they can contribute. Professor Accardi's map may help them to take their bearings.

Missing topics

Many topics had to be omitted from our paper: it is not a complete survey, but the outline of a kernel. As Professor Bingham and Professor Accardi emphasize, we have left out the rich connections to quantum probability and quantum stochastic processes, free probability, hypothesis testing (Professor Belavkin), the relationships between quantum statistics and quantum information theory, Bayesian approaches, nonparametric models and foundations (Professor Helland). Some of these topics figure in Barndorff-Nielsen *et al.* (2001) and they will reappear in our forthcoming monograph. Above all, we keenly miss data examples (Dr Frey). They will come.

Regarding nonparametric quantum statistics we would like to mention Gill and Guţă (2003).

Quantum probability is important for quantum physics, and deeply connected to classical probability. Quantum stochastic analysis has applications in the continuous time evolution of quantum systems, quantum field theory, classical stochastic processes and quantum statistical inference (continuous time data).

Professor Belavkin cites early papers which explore further the connections between exponential families and information bounds. Much more besides can be found in the work of the 1970s and 1980s and needs to be re-evaluated in the light of modern statistics.

Professor Kent and Dr Brody elaborate on the geometric content of quantum statistical models. We should mention that the Japanese school (Hayashi, Fujiwara, Matsumoto and Nagaoka) is making rapid progress in exploiting quantum statistical geometry. The connection to shape analysis is exciting.

One of the fascinations of quantum statistics is the foundational turmoil. As Professor Helland points out, it is difficult to accept a theory of the world which posits an abstract mathematical structure 'in the background', laying down axiomatically its impact on experience. His point of departure is the structure of statistical information in potential experiments. The Hilbert space structure emerges from intuitive invariance properties (geometry again). Others are working on similar lines, but physicists do not know much statistics, whereas we do not know much physics, which means that there is a real chance that connecting these fields could lead to significant progress.

We mention Belavkin's (2002) recent approach to the 'measurement problem' which considers both measurement and unitary evolution as a special case of a single, intrinsically stochastic, continuous time, evolution of a physical system. Professor Belavkin, like Professor Helland, adheres to Bohr's standpoint, that there is nothing (but paradoxes) to be gained from trying to find a classical reality behind quantum randomness. This is echoed in Dr Thompson's remark that all data (all experience) are classical; quantum statistical inference is concerned with the surprising and beautiful structure behind, connecting the probabilities.

Bell experiments

Much more could have been written about Bell inequalities. We reintroduce the topic here by describing a little simulation experiment to generate ± 1-valued variables X and Y, whose distribution depends on two parameters $\vec{a}$ and $\vec{b}$. Specifically, $\Pr(X = x, Y = y) = p(x, y; \vec{a}, \vec{b}) = (1 - xy\vec{a} \cdot \vec{b})/4$, where $x, y = \pm 1$, and $\vec{a}$ and $\vec{b}$ are unit vectors in $\mathbb{R}^3$. This is the distribution of measurement of spins of two spin half particles in the Bell singlet state.

The problem is to come up with a random variable Z and transformations such that (up to an arbitrarily good approximation)

from Z and $\vec{a}$ one constructs X and S,
from Z and $\vec{b}$ one constructs Y and T,
(X, Y), conditional on $S = 0$ and $T = 0$, is distributed according to p.

A more difficult problem is to find an exact construction with *certain* success, i.e. without conditioning (rejection sampling, post-selection). Bell (1964) tells us that it is impossible. The easier problem can be solved in many ways; see for instance Larsson (1998) and Accardi *et al.* (2002) ('the chameleon effect'). Bell (1981) knew that experimental post-selection must be prohibited; otherwise the experiment cannot rule out a classical explanation like our simulation.

As Dr Minozzo and Dr Larsson mention, no Bell experiment has yet been done which does not have an alternative classical explanation, e.g. our rejection sampling story. For different experiments, the 'explanation' must be rather different (compare Minozzo's story with ours). Often it is rather artificial.

The physicists do their best, and the Weihs *et al.* (1998) version of Aspect's experiment seems free of both defects described by Dr Minozzo, though it still involves post-selection on coincidences. Most physicists believe that a loophole-free experiment can and will be done, maybe even within a couple of years. However, some consider it a real possibility that quantum mechanics itself will always force loopholes.

We would like to know for what distributions p can we do the simulation exactly, with probability of success bounded away from 0, uniformly in the parameters? Does it help when, as in quantum mechanics, we have no action at a distance, i.e. the marginal of X does not depend on $\vec{b}$ and the marginal of Y does not depend on $\vec{a}$?

Bell experiments form a rich field for classical statistics. Very recently, van Dam *et al.* (2003) have used methodology from missing data theory to compare strengths of different Bell-type experiments.

Miscellany

Our decision to concentrate on mixed states was largely pedagogical. (This is in answer to Dr Brody and Dr Majumdar.) They are central in modern quantum information theory and the reader must become used to them. Mixed states are mixtures of pure states, so the pure case is included. We agree with Dr Brody that there can be physical problems when considering our toy models in their most immediate context. However, quantum statistical models for spin half apply also to the polarization of photons, as well as to many other physical systems. Ballester (2003) has new results on the quantum tomography of operations which involves a pure state model that is only marginally more complex than those which we mentioned as toys, and the experiment is being done right now in a quantum optics laboratory.

Professor Kent asks whether our models are empirical or (physics) theoretical. Well, some do appear in theory in physics, but a main motivation is just as in classical statistics: models which allow elegant inference are studied in their own right as paradigms of statistical–theoretical modelling.

Dr Thompson refers to coherence, and Dr Frey to decoherence. Dr Thompson is referring to loss functions. A characteristic of *quantum* statistical inference is that, the more we can learn about one parameter, the less we can learn about another. This means that a consideration of loss functions is necessary at the design stage. As Dr Thompson emphasizes, we need to think carefully what the meaning of the parameters is.

Decoherence (Dr Frey) refers to the decay of quantum entanglement through interaction with an environment. This process is part of quantum physics and part of actual laboratory measurement. Deciding at which point the quantum modelling can be replaced by classical (often forced by complexity as systems grow increasingly large) is one of the most difficult jobs of real quantum physics. It is part of the modelling of 'what measurement is actually being performed' on the system of interest.

Professor Bingham's nice comments on Moyal point both to the past and to the future. Concerning the connections between statistics and physics, we mention Edwards (2002) on the enormous significance of Fisher's background in physics for his work in statistics and genetics.

Many physicists have only hazy ideas of what statistics (in our sense) can mean for them. Ernest Rutherford once said 'if you need statistics, you did the wrong experiment'. However, he was not always as right as he was influential. His prediction concerning the possibility ('inconceivable') of using the structure of the atomic nucleus as a practically useful source of energy is a case in point: what about the use of statistics in exploring the quantum nature of the world as a source of computational power?

References in the discussion

Accardi, L., Imafuku, K. and Regoli, M. (2002) On the EPR-chameleon experiment. *Infin. Dimen. Anal. Quant. Probab. Reltd Flds*, **5**, 1–20.

Aspect, A. (1999) Bell's inequality test: more ideal than ever. *Nature*, **398**, 189–190.

Aspect, A., Dalibard, J. and Roger, G. (1982) Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.*, **49**, 1804–1807.

Ballester, M. A. (2003) Estimation of unitary quantum operations. *Preprint*. Mathematical Institute, University of Utrecht, Utrecht. (Available from `http://arxiv.org/abs/quant-ph/0305104`.)

Barndorff-Nielsen, O. E. (1978) *Information and Exponential Families in Statistical Theory*. Chichester: Wiley.

Barndorff-Nielsen, O. E. (1995) Diversity of evidence and Birnbaum's theorem. *Scand. J. Statist.*, **22**, 513–522.

Barndorff-Nielsen, O. E. and Cox, D. R. (1994) *Inference and Asymptotics*. London: Chapman and Hall.

Barndorff-Nielsen, O. E., Gill, R. D. and Jupp, P. E. (2001) On quantum statistical inference. *Research Report 2001-19*. Centre for Mathematical Physics and Stochastics, University of Aarhus, Aarhus. (Available from `http://arxiv.org/abs/quant-ph/0307189`.)

Belavkin, V. P. (1972) Optimal estimation of noncommuting quantum variables. *Radiotekh. Elektron.*, **17**, 2527–2532.

Belavkin, V. P. (1975a) Optimal observation of Boson signals in quantum Gaussian channels. *Prob. Control Inf. Theory*, **4**, 241–257.

Belavkin, V. P. (1975b) Optimal discrimination of non-orthogonal quantum signals. *Radiotekh. Elektron.*, **20**, 1177–1185.

Belavkin, V. P. (1975c) Optimal multiple quantum statistical hypothesis testing. *Stochastics*, **1**, 315–345.

Belavkin, V. P. (1976) Generalized uncertainty relations and efficient measurements in quantum systems (in Russian). *Theor. Mat. Fiz.*, **26**, 316–329.

Belavkin, V. P. (2002) Quantum causality, stochastics, trajectories and information. *Rep. Prog. Phys.*, **65**, 353–420.

Belavkin, V. P. and Grishanin, B. A. (1973) Investigation of the problem of optimal estimation in quantum channels by the methods of generalized Heisenberg inequality. *Prob. Pered. Inf.*, **4**, 44–52.

Bell, J. S. (1964) On the Einstein Podolsky Rosen paradox. *Physics*, **1**, 195–200.

Bell, J. S. (1981) Bertlmann's socks and the nature of reality. *J. Phys.*, **42**, C2, 41–61.

Bhattacharyya, A. (1943) On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calc. Math. Soc.*, **35**, 99–109.

Bhattacharyya, A. (1946) On a measure of divergence between two multinomial populations. *Sankhya*, **7**, 401–406.

Brody, D. C. (2003) Shapes of quantum states. Submitted to *J. R. Statist. Soc.* B.

Brody, D. C. and Hughston, L. P. (1996) Geometry of quantum statistical inference. *Phys. Rev. Lett.*, **77**, 2851–2854.

Brody, D. C. and Hughston, L. P. (1997) Generalised Heisenberg relations for quantum statistical estimation. *Phys. Lett.* A, **236**, 257–262.

Brody, D. C. and Hughston, L. P. (1998) Statistical geometry in quantum mechanics. *Proc. R. Soc. Lond.* A, **454**, 2445–2475.

Brody, D. C. and Hughston, L. P. (1999) Geometrization of statistical mechanics. *Proc. R. Soc. Lond.* A, **455**, 1683–1715.

Brody, D. and Meister, B. (1996) Minimum decision cost for quantum ensembles. *Phys. Rev. Lett.*, **76**, 1–5.

Burbea, J. (1986) Informative geometry of probability spaces. *Exp. Math.*, **4**, 347.

Burbea, J. and Rao, C. R. (1984) Differential metrics in probability spaces. *Prob. Math. Statist.*, **3**, 241–258.

Caianiello, E. R. (1983) Geometrical identification of quantum and information theories. *Lett. Nuov. Cim.*, **38**, 539–543.
Caianiello, E. R. and Guz, W. (1988) Quantum Fisher metric and uncertainty relations. *Phys. Lett.* A, **126**, 223–225.
Chari, V. and Pressley, A. (1994) *A Guide to Quantum Groups*. Cambridge: Cambridge University Press.
Clauser, J. F., Horne, M. A., Shimony, A. and Holt, R. A. (1969) Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, **23**, 880–884.
Connes, A. *Non-commutative Geometry*. San Diego: Academic Press.
van Dam, W., Gill, R. D. and Grünwald, P. D. (2003) The statistical strength of nonlocality proofs. *Preprint*. (Available from `http://arxiv.org/abs/quant-ph/0307125`.)
Dryden, I. L. and Mardia, K. V. (1998) *Statistical Shape Analysis*. Chichester: Wiley.
Edwards, A. W. F. (2002) Fisher information and the fundamental theorem of natural selection. *Ist. Lomb. Rend. Sci.* B, **134**, 3–17.
Efron, B. (1975) Defining curvature on statistical model. *Ann. Statist.*, **3**, 1189–1242.
Gani, J. (1998) Obituary, Jose Enrique Moyal. *J. Appl. Probab.*, **35**, 1012–1017.
Gill, R. D. (2003a) Accardi contra Bell (cum mundi): the impossible coupling. In *Mathematical Statistics and Applications: Festschrift for Constance van Eeden* (eds M. Moore, S. Froda and C. Léger). Beachwood: Institute of Mathematical Statistics.
Gill, R. D. (2003b) Time, finite statistics, and Bell's fifth position. In *Proc. Foundations of Probability and Physics 2*, pp. 179–206. Växjö: Växjö University Press.
Gill, R. D. and Guţă, M. (2003) An invitation to quantum tomography. *Preprint*. Mathematical Institute, University of Utrecht, Utrecht. (Available from `http://arxiv.org/abs/quant-ph/0303020`.)
Gill, R. D. and Massar, S. (2000) State estimation for large ensembles. *Phys. Rev.* A, **61**, 2312–2327.
Helland, I. S. (2003a) Extended statistical modelling under symmetry: the link towards quantum mechanics. To be published. (Available from `http://folk.uio.no/ingeh/publ.html`.)
Helland, I. S. (2003b) Quantum theory as a statistical theory under symmetry and complementarity. *Preprint*. (Available from `http://folk.uio.no/ingeh/publ.html`.)
Helstrom, C. W. (1967) Minimum mean-square error of estimate in quantum statistics. *Phys. Lett.* A, **25**, 101–102.
Helstrom, C. W. (1973) Cramer-Rao inequalities for operator-valued measures in quantum mechanics. *Int. J. Theoret. Phys.*, **8**, 361–376.
Helstrom, C. W. (1974) Estimation of a displacement parameter in quantum detection and estimation theory. *Int. J. Theoret. Phys.*, **11**, 357–378.
Helstrom, C. W. (1982) Bayes-cost reduction algorithm in quantum hypothesis testing. *IEEE Trans. Inform. Theory*, **28**, 359–366.
Larsson, J.-Å. (1998) The Bell inequality and detector inefficiency. *Phys. Rev.* A, **87**, 3304–3308.
Larsson, J.-Å. (1999) Modeling the singlet state with local variables. *Phys. Lett.* A, **256**, 245–252.
Larsson, J.-Å. (2002) A practical Trojan horse for Bell-inequality-based quantum cryptography. *Quant. Inform. Computn*, **2**, 434–442.
Loubenets, E. (2003) General probabilistic framework of randomness. Centre for Mathematical Physics and Stochastics, University of Aarhus, Aarhus. (Available from `http://arxiv.org/abs/quant-ph/0305126`.)
Mahalanobis, P. C. (1930) On tests and measures of groups divergence. *J. Asiat. Soc. Bengal*, **26**, 541–588.
Mahalanobis, P. C. (1936) On the generalised distance in statistics. *Proc. Natn. Inst. Sci.* A, **2**, 49–55.
McCullagh, P. (2002) What is a statistical model? *Ann. Statist.*, **30**, 1225–1310.
Meyer, P.-A. (1995) Quantum probability for probabilists. *Lect. Notes Math.*, **1538**.
Minozzo, M. (2000) Bell inequalities and correlation experiments: a purely particle statistical investigation. In *The Foundations of Quantum Mechanics: Historical Analysis and Open Questions* (eds C. Garola and A. Rossi), pp. 307–318. Singapore: World Scientific.
Moyal, J. E. (1949) Stochastic processes and statistical physics. *J. R. Statist. Soc.* B, **11**, 150–210.
Petz, D. (2003) Covariance and Fisher information in quantum mechanics. *Report quantph/0106125*.
Rao, C. R. (1945) Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calc. Math. Soc.*, **37**, 81–91.
Rao, C. R. (1947) The problem of classification and distance between two populations. *Nature*, **159**, 30–31.
Severini, T. A. (2000) *Likelihood Methods in Statistics*. Oxford: Oxford University Press.
Stratonovich, R. L. (1973) The quantum generalization of optimal statistical estimation and hypothesis testing. *J. Stochast.*, **1**, 87–126.
Voiculescu, D. V. (2000) Lectures on free probability. *Lect. Notes Math.*, **1738**.
Weihs, G., Jennewein, T., Simon, C., Weinfurter, H. and Zeilinger, A. (1998) Violation of Bell's inequality under strict Einstein locality conditions. *Phys. Rev. Lett.*, **81**, 5039–5043.
Yuen, H. P., Kennedy, R. S. and Lax, M. (1975) *IEEE Trans. Inform. Theory*, **21**, 125.
Yuen, H. and Lax, M. (1973) Multi-parameter quantum estimation and measurement of nonselfadjoint observables. *IEEE Trans. Inform. Theory*, **19**, 740–750.