

Coherent states and Bayesian duality

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 J. Phys. A: Math. Theor. 41 365302

(<http://iopscience.iop.org/1751-8121/41/36/365302>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.150

The article was downloaded on 03/06/2010 at 07:10

Please note that [terms and conditions apply](#).

Coherent states and Bayesian duality

S Twareque Ali¹, J-P Gazeau² and B Heller³

¹ Department of Mathematics and Statistics, Concordia University, Montréal, Québec, Canada H3G 1M8

² Astroparticules et Cosmologie (APC, UMR 7164), Université Paris Diderot Paris 7, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

³ Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA

E-mail: stali@mathstat.concordia.ca, gazeau@apc.univ-paris7.fr, heller@iit.edu and effe@midway.uchicago.edu

Received 1 May 2008, in final form 14 July 2008

Published 5 August 2008

Online at stacks.iop.org/JPhysA/41/365302

Abstract

We demonstrate how large classes of discrete and continuous statistical distributions can be incorporated into coherent states, using the concept of a reproducing kernel Hilbert space. Each family of coherent states is shown to contain, in a sort of duality, which resembles an analogous duality in Bayesian statistics, a discrete probability distribution and a discretely parametrized family of continuous distributions. It turns out that nonlinear coherent states, of the type widely studied in quantum optics, are a particularly useful class of coherent states from this point of view, in that they contain many of the standard statistical distributions. We also look at vector coherent states and multidimensional coherent states as carriers of mixtures of probability distributions and joint probability distributions.

PACS numbers: 04.20.Ex, 03.65.Ta

1. Introduction

In a series of recent papers, [17–19], an intimate connection between certain families of coherent states and statistical distributions has been demonstrated and studied. The coherent states discussed in these papers all have group theoretical origins and the Haar measure on the group has then been shown to induce a prior measure on the statistical parameters entering the definition of the discrete distributions. In this paper we look at a broader class of coherent states, which do not necessarily have their origins in group representations. In particular we show how, under certain technical restrictions, we can start with a discrete probability distribution, depending on a single real parameter, and associate coherent states to it. In the process we obtain a natural family of discretely indexed continuous distributions, which are then in a sort of *duality* with the original discrete distribution, via the coherent states. This duality is

highly reminiscent of a similar duality observed in the theory of Bayesian statistics, since the resolution of the identity condition, which we impose on the coherent states, introduces a preferred *prior measure* on the parameter space of the discrete distribution, with this distribution itself playing the role of the *likelihood function*. The associated discretely indexed continuous distributions become the related *conditional posterior distributions*. Alternatively, one can also start with a discretely parametrized family of continuous distributions, and under a certain convergence assumption, once more build coherent states. These coherent states then again give rise to a dual discrete distribution or likelihood function. We illustrate the theory by looking at a few examples of well-known statistical distributions (additional examples may be found in [12]). Although most of these examples have been studied earlier, in the context of Glauber–Klauder–Sudarshan or Gilmore–Perelomov coherent states [17–19], we analyze them here from the present perspective, i.e., without invoking any group property.

We take the discussion further by studying the relevance of vector coherent states and multidimensional coherent states when mixtures of probability distributions or joint distributions are considered. As far as we are aware, this is the first time that such vector coherent states have been studied in connection with statistical distributions.

It ought to be pointed out here that the classical Bayes theorem has been extended and used in the context of quantum probability before (see, for example, [26] and references cited therein) but these were in the context of quantum-conditional expectations, as related to quantum measurement theory. What we point out here is the appearance of classical probability distributions, in a form dictated by the classical Bayesian theorem, but related to certain quantum state vectors.

2. Experimental model context

In the following paragraphs, using simple experimental setups, we try to motivate the simultaneous appearance of a family of discrete probability distributions and a family of continuous distributions in the sort of duality referred to earlier. First we describe the classical inference procedure known as Bayesian inference in the experimental context indicated below. Then, as indicated above, we will consider a relationship between our subsequent mathematical analysis and this classical procedure (see the appendix).

2.1. Discrete case

Suppose we have an experimental setup for which we have an ‘experimental model’ in the form of a family of discrete probability distributions $n \mapsto P(n, \lambda)$ relating to a discrete set of possible experimental outcomes. That is, we do not know the preparation exactly, only to the extent of a family of states, indexed, say by the parameter λ which takes (continuous) values in some parameter space. The parameter usually represents a quantitative property of interest. In fact, the whole idea of the experiment, presumably, is to obtain data with which to estimate this physical property represented by the parameter. As an elementary example, let us think in terms of setting up an experiment to toss a coin N times and count the total number, k , of heads. Now perform the experiment and designate the observed value of k as k_{obs} . Then use k_{obs} to estimate the bias of the coin. The statistical model would be a family of binomial distributions indexed by a parameter p with ‘true’ but unknown parameter value p_0 . One can estimate the value of p_0 as $p_{\text{est}} = k_{\text{obs}}/N$. But conditionally upon the observed value, k_{obs} , one may consider p as a random variable and construct a certain conditional probability distribution over the parameter space which we now treat as a measurable space. The motivation for this inference procedure is that, for example, one could then find subsets of the parameter space

for which one could make statements such as ‘given the result of the experiment, there is a 99% chance that the true value p_0 lies within that subset’. (Think of an experiment where one tossed a coin 1000 times and got 999 heads.)

In [23, 30], the classical Bayesian method of inference is used in a quantum probability context in a form similar to the present paper, in that there appears both a *predictive* stochastic model in which the preparation is not fully determined and an inferred *retrodictive* probability distribution on the parameter space which indexes a family of possible preparation states. Here, we focus upon this duality of conditional probability distributions (one discrete and the other continuous), which may be obtained not only by the classically originated Bayesian method but by a consequent duality of two families of quantum probability distributions, constructed via the use of coherent states. The extension of the classical Bayesian method to quantum probability has been developed for use in other related quantum contexts as, for example, in [9, 26, 32].

2.2. The duality

In the Bayesian context, both the quantity to be observed and the unknown parameter are considered to be random quantities, playing a dual role. We consider two conditional probability distributions. Before performing the random experiment, the experimental model in the form of a family $P(y, \lambda)$ of discrete probability distributions is viewed as a conditional distribution of the random variable Y given the parameter value, say λ . After performing the experiment, we have an observed value, say y_{obs} , and we compute the conditional probability density function of the parameter λ given y_{obs} , obtaining a posterior conditional probability distribution. But, of course, we need to choose a prior measure $P(d\lambda)$. Suppose we have a probability density function where $P(d\lambda) = \Pi(\lambda) d\lambda$. The posterior probability density function is then given by [8, 30] (see also the appendix at the end),

$$f(\lambda, y_{\text{obs}}) = \frac{P(y_{\text{obs}}, \lambda)\Pi(\lambda)}{\int P(y_{\text{obs}}, \lambda')\Pi(\lambda') d\lambda'}. \quad (2.1)$$

A prototype classical example of the binomial distribution is the coin tossing experiment mentioned above and given in the appendix. In that classical context, the posterior conditional probability density function for the parameter p would be obtained according to (2.1).

An example of a Bayesian approach involving the binomial distribution in a quantum context is given in [30]. A thought experiment is described involving a count of photons which are passed through a polarizer, a pinhole and a calcite crystal, eventually triggering a detector as (+) or (−). In that context, a posterior distribution is obtained via (2.1) for the binomial parameter θ , the direction of the polarizer.

In [30], the family of probability distributions which we have called the stochastic model for the experiment is designated as *predictive*. The conditional probability distribution for the parameter that we have called Bayesian posterior is there designated as *retrodictive*.

3. A general setting for statistical distributions and coherent states

Let $\{X, \mu\}$ be a measure space. X could, for example, be the space of some statistical parameters or a larger space containing such parameters. Consider the Hilbert space $\mathfrak{H} = L^2(X, \mu)$ and suppose that it contains a reproducing kernel subspace \mathfrak{H}_K . This means that for any orthonormal basis, $\{\Phi_k\}_{k=0}^N$ of \mathfrak{H}_K (where N could be finite or infinite) the following is true:

- (1) $\sum_{k=0}^N |\Phi_k(x)|^2 < \infty$, for almost all $x \in X$ and in fact, it is possible to define the functions $\Phi_k(x)$ in a way so that this convergence condition holds everywhere.

(2) The function

$$K(x, y) = \sum_{k=0}^N \Phi_k(x) \overline{\Phi_k(y)} \tag{3.1}$$

defines a *reproducing kernel* i.e., $K(x, y)$ satisfies the properties,

$$\begin{aligned} K(x, y) &= \overline{K(y, x)}, & K(x, x) &> 0, & \text{for all } x \in X; \\ \int_X K(x, z) K(z, y) d\mu(z) &= K(x, y), & \text{for all } x, y \in X. \end{aligned} \tag{3.2}$$

It turns out that the kernel is independent of the orthonormal basis chosen to represent it.

For such a Hilbert space \mathfrak{H}_K , we can define a set of vectors, $|x\rangle$, labeled by the points of X in the manner:

$$|x\rangle = \mathcal{N}(x)^{-\frac{1}{2}} K(., x) = \mathcal{N}(x)^{-\frac{1}{2}} \sum_{k=0}^N \overline{\Phi_k(x)} \Phi_k, \quad \mathcal{N}(x) = K(x, x) = \sum_{k=0}^N |\Phi_k(x)|^2. \tag{3.3}$$

The normalization factor $\mathcal{N}(x)$ is chosen in order to ensure that $\langle x|x\rangle = 1$. In view of (3.2), these vectors are then immediately seen to satisfy the *resolution of the identity*,

$$\int_X |x\rangle \langle x| \mathcal{N}(x) d\mu(x) = I_{\mathfrak{H}_K}, \tag{3.4}$$

This condition implies that the vectors $|x\rangle$ form an *overcomplete* set in \mathfrak{H}_K , so that any vector in it can be written as a linear combination, either as a sum of or an integral over these. Very often such a set of vectors is associated with a unitary representation of some group, and are constructed by letting the representation operators act on a fixed vector in \mathfrak{H}_K . At other times such vectors are obtained by exploiting analytic properties of vectors in \mathfrak{H}_K . But at this point, we prefer to adopt a more general point of view and to just focus on the reproducing kernel Hilbert space structure. We shall call the vectors $|x\rangle$ (*generalized*) *coherent states* (see, for example [4], for a detailed discussion).

It is possible to associate two types of probability distributions to the basis vectors in a reproducing kernel Hilbert space. First, writing

$$P(n, x) = \frac{|\Phi_n(x)|^2}{\mathcal{N}(x)}, \quad n = 0, 1, 2, \dots, N, \tag{3.5}$$

we see that $\sum_{n=0}^N P(n, x) = 1$. Thus, $P(n, x)$ can be looked upon as a *discrete probability distribution* with parameter x . For instance, it can be based upon some experimental setup and then might be viewed as a stochastic model. Second, if $X \subset \mathbb{R}^m$, and if $d\mu$ has a Radon–Nikodym density with respect to the Lebesgue measure dx (on \mathbb{R}^m), then the functions,

$$\Psi_n(x) = |\Phi_n(x)|^2 \frac{d\mu(x)}{dx} = P(n, x) \mathcal{N}(x) \frac{d\mu(x)}{dx}, \quad n = 0, 1, 2, \dots, N, \tag{3.6}$$

define, for each n a *continuous probability density* on X , since $\int_X \Psi_n(x) dx = 1$. In the context of Bayesian statistics, this could be thought of as a conditional probability density for x , given n . If $P(n, x)$ is a statistical distribution, corresponding to some physical situation, which depends on the parameter x , the measure

$$d\bar{\kappa}(x) = \mathcal{N}(x) d\mu(x) \tag{3.7}$$

can be interpreted as a prior measure on the parameter space X and then the $\Psi_n(x)$ become the associated posterior distributions, in conformity with (2.1). In [17–19], a group

theoretical argument, exploiting the invariant measure and coherent states related to a particular representation of the group on a Hilbert space, was invoked to obtain the prior measure. Here we see that the appearance of a discrete probability distribution $P(n, x)$ and the continuous probability distributions $\Psi_n(x)$ in this dual relationship is embodied in the structure of the coherent states $|x\rangle$, independently of any group action.

3.1. A generic example

As a particular example, of the above situation, which will be useful for the purposes of the present paper, and which will turn out to have rich applications to statistical distributions encountered in extensive physical contexts, we introduce a family of the so-called *nonlinear coherent states*. These are built by taking an abstract, complex, separable Hilbert space \mathfrak{H} , of dimension N (finite or infinite), choosing an orthonormal basis $\phi_k, k = 0, 1, 2, \dots, N$, of it and defining on it the vectors

$$|z\rangle = \mathcal{N}(|z|^2)^{-\frac{1}{2}} \sum_{k=0}^N \frac{z^k}{[x_k!]^{\frac{1}{2}}} \phi_k, \tag{3.8}$$

where z is a parameter drawn from some appropriate open subset of \mathbb{C} and x_1, x_2, x_3, \dots is a conveniently chosen positive sequence of numbers for which we define the generalized factorial, $x_k! = x_1 x_2 \dots x_k$, with $x_0! = 1$, by definition. The normalization factor in this case is $\mathcal{N}(|z|^2) = \sum_{k=0}^N \frac{|z|^{2k}}{x_k!}$ and of course, $\langle z|z\rangle = 1$. In order to ensure that these coherent states form an overcomplete set of vectors in the Hilbert space \mathfrak{H} , one requires the resolution of the identity,

$$\int_{\mathcal{D}} |z\rangle \langle z| \mathcal{N}(|z|^2) d\nu(z, \bar{z}) = I_{\mathfrak{H}}, \tag{3.9}$$

to hold, where $I_{\mathfrak{H}}$ is the identity operator on the Hilbert space \mathfrak{H} and \mathcal{D} is an appropriate domain of the complex plane (usually the open unit disc or an open annulus, but which could also be the entire plane). It is not hard to see that the resolution of the identity (3.9) will hold if the measure $d\nu$, which is usually of the type $dQ(r) d\theta$ (for $z = r e^{i\theta}$), is such that dQ is related to the $x_k!$ through the following moment condition (see, for example, [33] for a discussion of the moment problem):

$$\frac{x_k!}{2\pi} = \int_0^{\sqrt{L}} r^{2k} dQ(r), \quad k = 0, 1, 2, \dots, \tag{3.10}$$

with L being the radius of convergence of the series $\sum_{k=0}^N \frac{|z|^{2k}}{x_k!}$ (considered as a series in $\lambda = |z|^2$). This means that once the sequence x_1, x_2, x_3, \dots is specified, the measure dQ is to be determined by solving the moment problem (3.10). There is an extensive literature on the construction of coherent states of this type (see, for example, [11, 24, 25, 28]). On the other hand, if the moment problem has no solution or, it has a solution but the corresponding measure is not explicitly known, there exists an alternative constructive procedure which allows one to build nonlinear coherent states, again resolving the identity [5].

We proceed now to analyze the discrete and continuous probability distributions, in the sense of the previous section, associated with these coherent states.

3.2. Discrete distribution associated with $|z\rangle$

With $\lambda = |z|^2$, define the discrete probability distribution $P(n, \lambda), n = 0, 1, 2, \dots, N$, by

$$P(n, \lambda) = \frac{\lambda^n}{x_n!} \mathcal{N}(\lambda)^{-1}. \tag{3.11}$$

The normalization condition $\langle z|z \rangle = 1$ is seen to imply that

$$\sum_{n=0}^N P(n, \lambda) = 1. \tag{3.12}$$

In the special case, where $x_n = n$, this distribution is just the well-known Poisson distribution, for then $x_n! = n!$, $N(\lambda) = e^\lambda$ and $L = \infty$. We shall see later that many of the well-known discrete statistical distributions are related to nonlinear coherent states in this manner. Note that if Y denotes the discrete random variable, $Y(n) = x_n$, then taking $x_0 = 0$, we obtain its expectation value

$$\langle Y \rangle = \sum_{n=0}^N x_n P(n, \lambda) = \lambda. \tag{3.13}$$

Thus for each λ we get a discrete probability distribution, which is some sort of a generalized Poisson distribution. In general, the sort of distributions given by (3.11) is of the power series type, well known in statistics (see, for example, [22]).

3.3. Continuous distributions associated with $|z\rangle$

We next note that in view of (3.10),

$$2\pi \int_0^L P(n, \lambda) \mathcal{N}(\lambda) d\bar{q}(\lambda) = 1, \quad n = 0, 1, 2, \dots, N,$$

where we have written

$$d\bar{q}(\lambda) = d\varrho(r), \quad r^2 = \lambda. \tag{3.14}$$

Thus, the functions,

$$\Psi_n(\lambda) = 2\pi P(n, \lambda) \mathcal{N}(\lambda) \frac{d\bar{q}(\lambda)}{d\lambda} = 2\pi \frac{\lambda^n}{x_n!} \frac{d\bar{q}(\lambda)}{d\lambda}, \quad n = 0, 1, 2, \dots, \tag{3.15}$$

define, for each n , a continuous probability density over the parameter space $0 \leq \lambda \leq L$. Here, $\frac{d\bar{q}(\lambda)}{d\lambda}$ denotes the Radon–Nikodym derivative of the measure $d\bar{q}$ with respect to the Lebesgue measure $d\lambda$, provided it exists. Clearly,

$$\int_0^L \Psi_n(\lambda) d\lambda = 1, \quad n = 0, 1, 2, \dots \tag{3.16}$$

From (3.12) it follows that

$$\sum_{n=0}^N \Psi_n(\lambda) = 2\pi \mathcal{N}(\lambda) \frac{d\bar{q}(\lambda)}{d\lambda} < \infty, \tag{3.17}$$

for almost all $\lambda \in [0, L]$. Also, if Λ is the continuous random variable over the parameter space $[0, L]$, such that $\Lambda(\lambda) = \lambda$, then

$$\langle \Lambda \rangle_n = \int_0^L \lambda \Psi_n(\lambda) d\lambda = x_{n+1}, \tag{3.18}$$

which is a dual relation to (3.13).

Finally note, that in terms of the discrete and continuous probability distributions themselves, the coherent states (3.8) may be written as

$$\begin{aligned} |z\rangle &= \sum_{n=0}^N [P(n, \lambda)]^{\frac{1}{2}} e^{-in\theta} \phi_n \\ &= \left[2\pi \mathcal{N}(\lambda) \frac{d\bar{q}(\lambda)}{d\lambda} \right]^{-\frac{1}{2}} \sum_{n=0}^N [\Psi_n(\lambda)]^{\frac{1}{2}} e^{-in\theta} \phi_n, \quad z = \sqrt{\lambda} e^{-i\theta}, \end{aligned} \tag{3.19}$$

and which satisfy the resolution of the identity,

$$\int_0^L \int_0^{2\pi} |z\rangle\langle z| \mathcal{N}(\lambda) d\bar{q}(\lambda) d\theta = I_{\mathfrak{H}}. \tag{3.20}$$

Comparing (2.1) and (3.15) we see that the measure

$$d\bar{\kappa}(\lambda) = 2\pi \mathcal{N}(\lambda) d\bar{q}(\lambda), \tag{3.21}$$

gives a prior measure on the parameter space $[0, L]$. Furthermore, these results give us a hint as to how one might construct coherent states starting from families of probability distributions.

We emphasize again that the duality appearing here, between the family of discrete probability distributions, $n \mapsto P(n, \lambda)$, parametrized by λ , and the family of continuous distributions $\lambda \mapsto \Psi_n(\lambda)$, parametrized by n , is analogous to the *Bayesian duality*, that we already referred to at the end of section 2.2, between a discrete probabilistic model $P(n, \lambda)$ and the continuous probability density function (see also the appendix to this paper), and which is captured in the relation

$$f(\lambda, n) = \frac{P(n, \lambda)\Pi(\lambda)}{\int_0^\infty P(n, \lambda)\Pi(\lambda) d\lambda}, \tag{3.22}$$

where n represents an experimentally realized value of the discrete random variable and this conditional density function (*Bayesian posterior* density function) is obtained using the prior measure $\Pi(\lambda) d\lambda$ (see, for example, [6]).

It is interesting to note that the coherent states $|z\rangle$, which are unit vectors in the Hilbert space \mathfrak{H} , may be thought of as being *square roots* of the discrete probability distribution function $n \mapsto P(n, \lambda)$, in the sense that $\| |z\rangle \|^2 = \sum_{n=0}^N P(n, \lambda) = 1$.

The probability distribution $P(n, \lambda)$ can be extracted from the coherent state $|z\rangle$ by taking the trace

$$P(n, \lambda) = \text{Tr}[|z\rangle\langle z| \mathbb{P}_n] = |\langle \phi_n | z \rangle|^2, \tag{3.23}$$

where $\mathbb{P}_n = |\phi_n\rangle\langle \phi_n|$. In a quantum-mechanical interpretation, this $P(n, \lambda)$ is the probability of measuring the physical quantity encoded by the state ϕ_n when the system under observation had been prepared in the state $|z\rangle$.

3.4. Coherent states from discrete statistical distributions

Suppose now that we start with a discrete probability distribution, $P(n, \lambda)$, where again $n = 0, 1, 2, \dots, N$, with N being either finite or infinite and λ is a parameter drawn from the interval $[a, b] \subset [0, \infty)$. Of course, $\sum_{n=0}^N P(n, \lambda) = 1$ and we further assume that $P(n, \lambda)$ satisfies the conditions:

- (1) There exists a measure $d\kappa$ on $[a, b]$, absolutely continuous with respect to the Lebesgue measure $d\lambda$ and such that

$$\int_a^b P(n, \lambda) d\kappa(\lambda) := c_n < \infty, \quad n = 0, 1, 2, \dots, N. \tag{3.24}$$

- (2) For all $\lambda \in [a, b]$,

$$\sum_{n=0}^N \frac{P(n, \lambda)}{c_n} < \infty. \tag{3.25}$$

On the interval $[a, b]$, let us define the functions

$$\Psi_n(\lambda) = \frac{1}{c_n} P(n, \lambda) \frac{d\kappa(\lambda)}{d\lambda}, \tag{3.26}$$

for which we note that

$$\int_a^b \Psi_n(\lambda) d\lambda = 1, \quad n = 0, 1, 2, \dots, N, \tag{3.27}$$

and using them we define on the open annulus,

$$\mathcal{D} = \{z = \sqrt{\lambda} e^{-i\theta} | a < \lambda < b, 0 \leq \theta < 2\pi\} \subset \mathbb{C}, \tag{3.28}$$

the functions

$$\Phi_n(z) = \frac{1}{\sqrt{2\pi}} [\Psi_n(\lambda)]^{\frac{1}{2}} e^{-in\theta}. \tag{3.29}$$

Note that the range of values of the index n need not be constrained to lie among the non-negative integers only. It could also be a subset of \mathbb{Z} or all of it.

It is worthwhile pointing out that the measure $d\kappa$ postulated in (3.24) is not necessarily unique, which leaves the possibility of there being several such measures which could be acceptable. In the case of the discrete distributions arising from nonlinear coherent states, the requirement of the resolution of the identity, i.e., the moment condition (3.10) fixes the measure $d\kappa$. Also the functions (3.26) are exactly like the $f(\lambda, n)$ in (3.22), appearing in the duality studied in Bayesian statistics [6, 30] although, unlike in that case, we have here the additional restriction (3.25).

Clearly, the functions $\{\Phi_n\}_{n=0}^N$ form an orthonormal set,

$$\int_{\mathcal{D}} \overline{\Phi_m(z)} \Phi_n(z) d\lambda d\theta = \delta_{mn}. \tag{3.30}$$

Let \mathfrak{H} denote the Hilbert subspace of $L^2(\mathcal{D}, d\lambda d\theta)$ generated by these functions. Since,

$$\sum_{n=0}^N |\Phi_n(z)|^2 = \frac{1}{2\pi} \frac{d\kappa(\lambda)}{d\lambda} \sum_{n=0}^N \frac{P(n, \lambda)}{c_n} < \infty, \tag{3.31}$$

by virtue of (3.25), \mathfrak{H} is a reproducing kernel Hilbert space. From the discussion at the beginning of this section (see (3.3)), we can then define coherent states in \mathfrak{H} as

$$|z\rangle = [\mathcal{N}(\lambda)]^{-\frac{1}{2}} \sum_{n=0}^N \left[\frac{P(n, \lambda)}{c_n} \right]^{\frac{1}{2}} e^{-in\theta} \Phi_n, \quad \mathcal{N}(\lambda) = \sum_{n=0}^N \frac{P(n, \lambda)}{c_n}, \tag{3.32}$$

which now satisfy the resolution of the identity

$$\frac{1}{2\pi} \int_a^b \int_0^{2\pi} |z\rangle \langle z| \mathcal{N}(\lambda) d\kappa(\lambda) d\theta = I_{\mathfrak{H}}. \tag{3.33}$$

Note that from (3.24) and (3.26), we get

$$\Psi_n(\lambda) = \frac{P(n, \lambda)\Pi(\lambda)}{\int_a^b P(n, \lambda)\Pi(\lambda) d\lambda}, \quad \text{where} \quad \Pi(\lambda) = \frac{d\kappa(\lambda)}{d\lambda}, \tag{3.34}$$

so that $d\kappa$ can be thought of (see (3.22)) as a *prior measure* on the parameter space $a < \lambda < b$ and the Ψ_n as the associated Bayesian posteriors.

To make the connection with (3.3) and (3.4), we easily see that the coherent states (3.32) can also be written as

$$|z\rangle = \tilde{\mathcal{N}}(|z|^2)^{-\frac{1}{2}} \sum_{n=0}^{\infty} \overline{\Phi_n(z)} \Phi_n, \quad \tilde{\mathcal{N}}(|z|^2) = \sum_{n=0}^{\infty} |\Phi_n(z)|^2, \tag{3.35}$$

and the resolution of the identity as

$$\frac{1}{2\pi} \int_a^b \int_0^{2\pi} |z\rangle\langle z| \tilde{\mathcal{N}}(\lambda) d\lambda d\theta = I_{\mathfrak{H}}. \tag{3.36}$$

3.5. Coherent states from continuous statistical distributions

We now proceed to construct analogous families of coherent states from sets of continuous probability distributions. Suppose that $\Psi_n(\lambda), n = 0, 1, 2, \dots, N$, is a set of continuous probability densities defined over the set $I \subset \mathbb{R}$. Evidently, they satisfy

$$\int_I \Psi_n(\lambda) d\lambda = 1, \quad n = 0, 1, 2, \dots, N.$$

We assume in addition that

$$\tilde{\mathcal{N}}(\lambda) := \frac{1}{2\pi} \sum_{n=0}^N \Psi_n(\lambda) < \infty, \quad \lambda \in I. \tag{3.37}$$

Then, as before we construct the set of functions on $X = I \times [0, 2\pi)$,

$$\Phi_n(\lambda, \theta) = \frac{1}{\sqrt{2\pi}} [\Psi_n(\lambda)]^{\frac{1}{2}} e^{-in\theta}, \quad n = 0, 1, 2, \dots, N, \tag{3.38}$$

and note that they form an orthonormal set in $L^2(X, d\lambda d\theta)$. Let \mathfrak{H} be the Hilbert subspace of $L^2(X, d\lambda d\theta)$ generated by these vectors. Then once again, following (3.3) we construct the coherent states in \mathfrak{H} ,

$$|\lambda, \theta\rangle = \tilde{\mathcal{N}}(\lambda)^{-\frac{1}{2}} \sum_{n=0}^N \overline{\Phi_n(\lambda, \theta)} \Phi_n, \tag{3.39}$$

with $\tilde{\mathcal{N}}(\lambda)$ as in (3.37). These coherent states satisfy the resolution of the identity,

$$\int_I \int_0^{2\pi} |\lambda, \theta\rangle\langle\lambda, \theta| \tilde{\mathcal{N}}(\lambda) d\lambda d\theta = I_{\mathfrak{H}}. \tag{3.40}$$

Clearly, the discrete distribution function this time is

$$P(n, \lambda) = \frac{\Psi_n(\lambda)}{\tilde{\mathcal{N}}(\lambda)}, \tag{3.41}$$

with $\tilde{\mathcal{N}}(\lambda) d\lambda$ the prior measure.

3.6. A quick recapitulation

In the preceding subsections we developed a correspondence between two apparently disparate mathematical constructions originating in differing contexts. From the field of coherent states, we have the duality of two types of probability distributions as described in section 3: on the one hand, we have a discrete family of states indexed by some parameter(s), while on the other hand, we have a continuous probability distribution over the parameter space provided by the resolution of the identity associated with the coherent state family.

As described in the introduction, sections 1, 2 and the appendix, this duality is mirrored by the Bayes method of statistical inference which takes place in an experimental context, modeled by a parametric family of probability distributions. The duality here is between the original stochastic model and a conditional probability distribution over the parameter space, given observed results. Thus, the field of coherent states within a quantum physics

context provides a method for performing statistical inference within a classical experimental context. The Bayes method of inference pertains to classical probability distributions but the method has been adopted for use in quantum experimental situations as, for example, in [9, 23, 26, 32].

Examples relating to several commonly used discrete stochastic models are given in section 4. From section 5, we see that inference in the case of mixture stochastic models and multivariate models can also be constructed from the duality provided by vector coherent states.

4. Some illustrative examples

In this section we construct coherent states for some standard statistical distributions, following the general procedure outlined above. These coherent states have been obtained before, using group theoretical arguments [17–19] and we shall indicate, in each case, the group theoretic relevance of the coherent states. Moreover, in each case the interplay between the dual system of discrete and continuous distributions, embodied in the coherent states will be explicitly demonstrated.

4.1. Coherent states from the Poisson distribution

For the Poisson distribution, the probability of n successes, given that the average number of successes is $\lambda > 0$, is

$$P(n, \lambda) = \frac{e^{-\lambda} \lambda^n}{n!} \quad \text{and} \quad \sum_{n=0}^{\infty} P(n, \lambda) = 1. \quad (4.1)$$

Once again we would like to relate these to a family of coherent states. Also, thinking of λ itself as a random variable, we would like to obtain a distribution function for it. We start by introducing the complex variable, $z = \sqrt{\lambda} e^{-i\theta}$, and since $\int_0^{\infty} P(n, \lambda) d\lambda = 1$ for all n , we define the functions (see (3.29))

$$\Phi_n(z) = \frac{1}{\sqrt{2\pi}} [P(n, \lambda)]^{\frac{1}{2}} e^{-in\theta} = \frac{1}{\sqrt{2\pi}} \left[\frac{\lambda^n e^{-\lambda}}{n!} \right]^{\frac{1}{2}} e^{-in\theta}, \quad n = 0, 1, 2, \dots, \infty. \quad (4.2)$$

These functions are clearly orthonormal with respect to the measure $d\lambda d\theta$,

$$\int_0^{\infty} \int_0^{2\pi} \overline{\Phi_m(z)} \Phi_n(z) d\lambda d\theta = \delta_{mn}.$$

Let $\mathfrak{H} \subset L^2(\mathbb{C}, d\lambda d\theta)$ be the (infinite-dimensional separable) Hilbert space generated by them. Next we see that conditions (3.24) and (3.25) are satisfied with $d\kappa = d\lambda$ and $c_n = 1$ for all n . Thus, following (3.32) we may define coherent states on \mathfrak{H} as

$$|z\rangle = \sum_{n=0}^{\infty} \sqrt{P(n, \lambda)} e^{-in\theta} \Phi_n = e^{-\frac{|z|^2}{2}} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} \Phi_n, \quad (4.3)$$

so that

$$\langle z|z\rangle = \sum_{n=0}^{\infty} P(n, \lambda) = 1.$$

Again, the coherent states $|z\rangle$, may be thought of as being square roots of the discrete Poisson distribution, $n \mapsto P(n, \lambda)$.

These coherent states also satisfy a resolution of the identity,

$$\frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} |z\rangle\langle z| d\lambda d\theta = I_{\mathcal{H}}. \tag{4.4}$$

It is clear that this time the prior measure on the parameter space $0 \leq \lambda < \infty$ is just the uniform distribution $d\lambda$, with the Bayesian posteriors being given by $\Psi_n(\lambda) = P(n, \lambda)$. The coherent states (4.3) are the canonical coherent states, well known in the physical literature (see, e.g., [4]). Moreover, these coherent states are associated with a unitary representation of the Weyl–Heisenberg group and the prior measure $d\lambda$ is also obtainable from the Haar measure of this group [19].

Finally, it ought to be pointed out that the continuous distribution given by the function $\Psi_n(\lambda) = P(n, \lambda)$ is just a γ -distribution, for each n . In other words, the discrete Poisson distribution and the continuous γ -distributions (which may now be thought of as being conditional distributions for the average number of success λ , given n successes) are in duality through the canonical coherent states. Moreover, had we started with the γ -distribution functions, $\gamma_n(\lambda) = \frac{\lambda^{n-1} e^{-\lambda}}{\Gamma(n)}$, defined $\Psi_n = \gamma_{n+1}$, $n = 0, 1, 2, \dots, \infty$ and followed through the steps in section 3.5, we would have arrived at the same coherent states (4.3). In the field of statistics, the gamma distribution is said to be a *natural conjugate* to the Poisson sampling process [7].

4.2. Coherent states from the binomial distribution

Consider the binomial distribution for N independent trials, each having a probability of success p and of failure $q = 1 - p$. The probability of getting n successes in these N trials is

$$P(n, p) = \binom{N}{n} p^n q^{N-n} = \frac{N!}{(N-n)!n!} p^n q^{N-n}, \quad n = 0, 1, 2, \dots, N, \tag{4.5}$$

and of course

$$\sum_{n=0}^N P(n, p) = (q + p)^N = 1.$$

As before, we treat the parameter p itself also as a random variable and then use our general construction in order to: (1) obtain coherent states representing this distribution and (2) find a posterior distribution for p . This case has also been worked out in [18], using coherent states of the rotation group and we shall indicate the connection to this approach in what follows. Let us first introduce a new parameter λ , which will be more convenient for our purposes,

$$\lambda = \frac{p}{q} \implies q = \frac{1}{1+\lambda} \quad \text{and} \quad 0 \leq \lambda < \infty. \tag{4.6}$$

Using this we introduce the complex variable $z = \sqrt{\lambda} e^{-i\theta}$ and note that in terms of λ , the probability distribution (4.5) can be rewritten as

$$P(n, \lambda) = \frac{N!}{(N-n)!n!} \cdot \frac{\lambda^n}{(1+\lambda)^N} = \frac{\Gamma(N+1)}{\Gamma(N-n+1)\Gamma(n+1)} \cdot \frac{|z|^{2n}}{(1+|z|^2)^N} \tag{4.7}$$

Since

$$(N+1) \int_0^\infty \frac{\lambda^n}{(1+\lambda)^{N+2}} d\lambda = (N+1) \int_0^1 q^{N-n} (1-q)^n dq = \frac{n!(N-n)!}{N!},$$

we take (see (3.24))

$$d\kappa = \frac{(N+1)}{(1+\lambda)^2} d\lambda, \quad c_n = 1. \tag{4.8}$$

Since N is finite, (3.25) is trivially satisfied. Thus, we take

$$\Psi_n(\lambda) = P(n, \lambda) \frac{d\kappa(\lambda)}{d\lambda} = \frac{(N+1)!}{(N-n)!n!} \cdot \frac{\lambda^n}{(1+\lambda)^{N+2}} \tag{4.9}$$

and

$$\Phi_n(z) = \left[\frac{(N+1)!}{2\pi(N-n)!n!} \right]^{\frac{1}{2}} \frac{z^n}{(1+|z|^2)^{\frac{N}{2}+1}}, \quad z \in \mathbb{C}, \quad n = 0, 1, 2, \dots, N. \tag{4.10}$$

Clearly, these vectors are orthonormal,

$$\int_0^\infty \int_0^{2\pi} \overline{\Phi_m(z)} \Phi_n(z) d\lambda d\theta = \delta_{mn}$$

and we denote by \mathfrak{H} the $(N+1)$ -dimensional Hilbert space generated by these vectors. On this space we then have the coherent states,

$$|z\rangle = \sqrt{P(n, \lambda)} e^{-in\theta} \Phi_n = \frac{1}{(1+|z|^2)^{\frac{N}{2}}} \sum_{n=0}^N \frac{\sqrt{\Gamma(N+1)} z^n}{\sqrt{\Gamma(N-n+1)\Gamma(n+1)}} \Phi_n. \tag{4.11}$$

Note again, that since

$$\langle z|z\rangle = 1 = \sum_{n=0}^N P(n, \lambda),$$

for each $\lambda = |z|^2$, the coherent state $|z\rangle$ is sort of a vectorial square root of the probability distribution $P(n, \lambda), n = 0, 1, 2, \dots, N$. These coherent states satisfy the resolution of the identity,

$$\frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} |z\rangle\langle z| d\kappa(\lambda) d\theta = \frac{N+1}{2\pi} \int_0^\infty \int_0^{2\pi} |z\rangle\langle z| \frac{d\lambda d\theta}{(1+\lambda)^2} = I_{\mathfrak{H}}. \tag{4.12}$$

Next, introducing the new labels $N = 2j, k = n - j$, we write

$$|z\rangle = (1+|z|^2)^{-\frac{N}{2}} \sum_{k=-j}^j \frac{\sqrt{\Gamma(2j+1)} z^{k+j}}{\sqrt{\Gamma(j-k+1)\Gamma(j+k+1)}} \Phi_k, \tag{4.13}$$

which are immediately recognized as being the Gilmore–Perelomov–Radcliffe-type coherent states [1, 4, 29, 31] for the $(2j+1)$ -representation of $SU(2)$. Indeed, the vectors $|z\rangle$ may be rewritten in terms of the $SU(2)$ generators J_\pm, J_3 and the lowest basis vector Φ_{-j} as

$$|z\rangle = e^{zJ_+} e^{\eta J_3} e^{-\bar{z}J_-} \Phi_{-j} = e^{\xi J_+ - \bar{\xi} J_-} \psi_{-j} := D(\xi) \Phi_{-j}, \tag{4.14}$$

where writing $z = -\tan \frac{\vartheta}{2} e^{-i\gamma}$,

$$\xi = i \frac{\vartheta}{2} e^{i\gamma} \quad \text{and} \quad \eta = \log(1+|z|^2) = 2 \log \sec \frac{\vartheta}{2}.$$

Finally, note that by virtue of (4.7) and (4.9), the measure

$$d\kappa(\lambda) = \frac{N+1}{(1+\lambda)^2} d\lambda \quad \text{or equivalently,} \quad d\kappa(p) = (N+1) dp \tag{4.15}$$

gives in this case the prior measure (again uniform) of the parameter p over the interval $[0, 1]$.

Once again, it is clear that had we started with the continuous distributions (4.9), which are β -distributions of the first kind, and followed through with the procedure in section 3.5, we would also have arrived at the coherent states (4.11). Thus, the continuous β -distributions of the first kind and the discrete binomial distribution (statistical *conjugate pair*) are in duality through the $SU(2)$ coherent states.

4.3. Coherent states from the negative binomial and β -distributions

The negative binomial and the β -distributions have a dual relationship through the coherent states arising from the discrete series representations of the $SU(1, 1)$ group. Recall that, for a fixed integer $m \geq 1$, the negative binomial distribution is given by

$$P(m, n; \lambda) = \frac{\Gamma(m+n)}{\Gamma(n+1)\Gamma(m)} \lambda^m (1-\lambda)^n, \quad n = 0, 1, 2, \dots, \infty, \quad (4.16)$$

where the parameter λ lies in the interval $(0, 1)$. The quantity $P(m, n, \lambda)$ can be thought of as being the probability that $m+n$ is the number of independent trials that are necessary to obtain the result of m successes (the $(m+n)$ th trial being a success) when λ is the probability of success in a single trial. The term negative binomial stems from the fact that

$$(1-\lambda)^{-k} = \sum_{n=0}^{\infty} \frac{\Gamma(k+n)}{\Gamma(n+1)\Gamma(k)} \lambda^n,$$

from which it also follows that

$$\sum_{n=0}^{\infty} P(m, n; \lambda) = 1. \quad (4.17)$$

The β -distribution is a continuous distribution, in the variable $\lambda \in [0, 1]$, with discrete parameters $m, n = 1, 2, 3, \dots, \infty$,

$$\beta(\lambda; m, n) = \frac{1}{B(m, n)} \lambda^{m-1} (1-\lambda)^{n-1}, \quad \int_0^1 \beta(\lambda; m, n) d\lambda = 1, \quad (4.18)$$

where

$$B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = \int_0^1 t^{m-1} (1-t)^{n-1} dt.$$

We note that

$$\beta(\lambda; m+1, n+1) = \frac{P(m, n; \lambda)}{c_{m,n}}, \quad \text{with} \quad c_{m,n} = \frac{m}{(m+n+1)(m+n)}, \quad (4.19)$$

implying, by virtue of (4.18),

$$\int_0^1 P(m, n; \lambda) d\lambda = c_{m,n} \quad \text{and} \quad d\kappa(\lambda) = d\lambda. \quad (4.20)$$

Thus, (3.24) is satisfied, with $c_n = c_{m,n}$ and (3.25) is also satisfied since

$$\sum_{n=0}^{\infty} \frac{P(m, n; \lambda)}{c_{m,n}} = \frac{m+1}{\lambda^2} \sum_{n=0}^{\infty} P(m+2, n; \lambda) = \frac{m+1}{\lambda^2} < \infty \quad (4.21)$$

by virtue of (4.17).

Thus, for fixed $m \geq 1$ and $n = 0, 1, 2, \dots, \infty$, we define, using (3.26) and (4.19), the continuous distributions,

$$\Psi_{m,n}(\lambda) = \frac{P(m, n; \lambda) d\kappa(\lambda)}{c_{m,n} d\lambda} = \beta(\lambda; m+1, n+1) = \frac{1}{B(m+1, n+1)} \lambda^m (1-\lambda)^n, \quad (4.22)$$

and the associated functions in the complex variable $\zeta = \sqrt{\lambda} e^{-in\theta}$, $0 \leq \lambda < \infty$, $0 \leq \theta < 2\pi$,

$$\Phi_{m,n}(\zeta) = \frac{1}{\sqrt{2\pi}} [\Psi_{m,n}(\lambda)]^{\frac{1}{2}} e^{-in\theta},$$

which satisfy the orthonormality condition

$$\int_0^\infty \int_0^{2\pi} \overline{\Phi_{m,n}(\zeta)} \Phi_{m,k}(\zeta) d\lambda d\theta = \delta_{nk}.$$

Denoting by \mathfrak{H} the (infinite-dimensional separable) Hilbert space spanned by these vectors, and noting that by (4.21),

$$\mathcal{N}(\lambda) = \sum_{n=0}^\infty \frac{P(m, n; \lambda)}{c_{m,n}} = \frac{m+1}{\lambda^2},$$

we define the coherent states associated with the discrete negative binomial and continuous β -distributions, on this space using (3.32),

$$\begin{aligned} |\zeta; m\rangle &= \mathcal{N}(\lambda)^{-\frac{1}{2}} \sum_{n=0}^\infty \left[\frac{P(m, n; \lambda)}{c_{m,n}} \right]^{\frac{1}{2}} e^{-in\theta} \Phi_{m,n} \\ &= \sum_{n=0}^\infty \left[\frac{\Gamma(m+n+2)}{\Gamma(m+2)\Gamma(n+1)} \right]^{\frac{1}{2}} \lambda^{\frac{m}{2}+1} (1-\lambda)^{\frac{n}{2}} e^{-in\theta} \Phi_{m,n}. \end{aligned} \tag{4.23}$$

These satisfy the resolution of the identity,

$$\frac{m+1}{2\pi} \int_0^1 \int_0^{2\pi} |\zeta, m\rangle \langle \zeta, m| \frac{d\lambda d\theta}{\lambda^2} = I_{\mathfrak{H}}. \tag{4.24}$$

while from (3.22), (4.20) and (4.22) we obtain the prior measure on the parameter space $[0, 1]$,

$$d\kappa(\lambda) = d\lambda. \tag{4.25}$$

Note that this measure is different from the one obtained in [17], which was derived using a group theoretical argument. However, in the present case, $m = 1, 2, 3, \dots$, while in [17] the value $m = 1$ was excluded. The associated Bayesian posteriors this time are the $\Psi_{m,n}, n = 0, 1, 2, \dots, \infty$.

Once again it is clear that if we start with the continuous β -distributions (4.18), and construct coherent states following section 3.5, with $\Psi_n(\lambda) = \beta(\lambda; m+1, n+1)$, we arrive at these same coherent states.

To make contact with the coherent states of the $SU(1, 1)$ group let us introduce the new complex variable $z = (1 - |\zeta|^2)^{\frac{1}{2}} e^{-i\theta} = (1 - \lambda)^{\frac{1}{2}} e^{-i\theta}$ and write $m+2 = 2j$. Then in terms of this variable we get the coherent states

$$|z; j\rangle = (1 - |z|^2)^j \sum_{n=0}^\infty \left[\frac{\Gamma(2j+n)}{\Gamma(2j)\Gamma(n+1)} \right]^{\frac{1}{2}} z^n \Phi_{2j,n}, \quad j = \frac{3}{2}, 2, \frac{5}{2}, \dots \tag{4.26}$$

These are the Gilmore–Perelomov-type coherent states arising from the discrete series representations [4, 17, 29] of $SU(1, 1)$. Since we are assuming that $m \geq 1$, the representation corresponding to $j = 1$ does not appear here. We observe that in the mathematical literature, these coherent states are usually written without the factor of $(1 - |z|^2)^j$ appearing before the sum on the right-hand side of (4.26). This is because, unlike in our case, the Hilbert space for the discrete series representations of $SU(1, 1)$ is taken to be the one consisting of all holomorphic functions on the open unit disc of \mathbb{C} , which are square integrable with respect to the measure $\frac{(2j-1)}{\pi}(1 - |z|^2)^{2j-2} dx dy$, where $z = x + iy$, and the factor is absorbed into the measure.

Note finally, that all three examples discussed here lead to coherent states of the nonlinear type (see (3.8)). To summarize, we have seen that the canonical coherent states combine in duality the continuous γ -distributions with the Poisson distribution, the coherent states of

the $SU(2)$ group so combine the continuous β -distributions of the first kind with the discrete binomial distribution and the coherent states obtained from the discrete series representations of the $SU(1, 1)$ group combine in duality the continuous β -distributions with the discrete negative binomial distribution.

5. Vector and multidimensional coherent states from probability distributions

So far we have considered only single discrete probability distributions and constructed coherent states from them. We now look at a situation where several independently distributed random variables are at play. It will turn out that the appropriate type of coherent states to associate with such situations are vector coherent states (VCS) of the type discussed in [3, 34] or multidimensional coherent states of the type studied in [27].

Let us take a discrete probability distribution $P(n, \lambda), n = 0, 1, 2, \dots, N$ (finite or infinite). This is the probability distribution of the discrete random variable \mathfrak{N} such that $\mathfrak{N}(n) = n$ and assume that it is of the type (3.11), i.e., the associated coherent states are of the nonlinear type. Assume now that we have M such independent, random variables, distributed with parameters $\lambda_1, \lambda_2, \dots, \lambda_M$, respectively, each drawn from the interval $[0, L]$. Then

$$P(\lambda_1, \lambda_2, \dots, \lambda_M; n) = \frac{1}{M} \sum_{i=1}^M P(n, \lambda_i) \tag{5.1}$$

is the probability of n ‘successes’ coming from any one of these processes when we are indifferent to which one it comes from. We now ask if there is a natural set of coherent states that could incorporate such a system of distributions, along the lines of what we saw earlier. It will turn out that a Hilbert space over a matrix domain, consisting of normal matrices, will be appropriate for the construction of such coherent states. Recall that a normal matrix \mathfrak{Z} is defined by the condition $\mathfrak{Z}^* \mathfrak{Z} = \mathfrak{Z} \mathfrak{Z}^*$ and if \mathfrak{Z} is an $M \times M$ matrix, it can be diagonalized by means of a unitary matrix, i.e.,

$$\mathfrak{Z} = U \text{diag}[z_1, z_2, \dots, z_M] U^* \tag{5.2}$$

where $U \in U(M)$ and the elements $z_i, i = 1, 2, 3, \dots, M$, of the diagonal matrix are complex numbers. Writing $z_i = \sqrt{\lambda_i} e^{-i\theta_i}$, let Ω denote the set of all such matrices for which $0 \leq \lambda_i < L, i = 1, 2, 3, \dots, M$. We next define the matrix valued functions on the domain Ω ,

$$\Phi_n(\mathfrak{Z}) = \frac{\mathfrak{Z}^n}{\sqrt{x_n!}}, \quad n = 0, 1, 2, \dots, N, \tag{5.3}$$

and on Ω we define the measure

$$d\Omega(\mathfrak{Z}, \mathfrak{Z}^*) = dU \prod_{i=1}^M d\bar{\varrho}(\lambda_i) d\theta_i, \quad \int_{\Omega} d\Omega(\mathfrak{Z}, \mathfrak{Z}^*) = 1, \tag{5.4}$$

where dU is the (normalized) invariant measure of $U(M)$ and $d\bar{\varrho}$ is the measure introduced into (3.10) and (3.14).

It then follows that the functions Φ_n satisfy the matrix orthogonality condition

$$\int_{\Omega} \Phi_m(\mathfrak{Z}) \Phi_n(\mathfrak{Z})^* d\Omega(\mathfrak{Z}, \mathfrak{Z}^*) = \mathbb{I}_M \delta_{mn}, \tag{5.5}$$

where \mathbb{I}_M is the $M \times M$ identity matrix. Let $\{\chi^i\}_{i=1}^M$ be an orthonormal basis of \mathbb{C}^M and define the \mathbb{C}^M -valued functions,

$$\Phi_n^i(\mathfrak{Z}^*) = \Phi_n(\mathfrak{Z}^*) \chi^i. \tag{5.6}$$

Note that

$$\text{Tr}[\mathfrak{Z}\mathfrak{Z}^*] = \sum_{i=1}^M |z_i|^2.$$

Also, the series,

$$\sum_{n=0}^N \text{Tr}[\Phi_n(\mathfrak{Z})^* \Phi_n(\mathfrak{Z})] = \sum_{n=0}^N \sum_{i=1}^M \Phi_n^i(\mathfrak{Z})^\dagger \Phi_n^i(\mathfrak{Z}) = \sum_{n=0}^N \sum_{i=1}^M \frac{\lambda_i^n}{x_n!} \quad (5.7)$$

converges for all $\lambda_i \in [0, L)$, which following the discussion at the beginning of section 3, is the condition for building reproducing kernel Hilbert spaces, which we now proceed to do.

Consider the Hilbert space $\tilde{\mathfrak{H}} = L^2_{\mathbb{C}^M}(\Omega, d\Omega)$ of square integrable, M -component vector-valued functions on Ω . The vectors $\Phi_k^i, i = 1, 2, \dots, M, k = 0, 1, 2, \dots, N$ are elements of this Hilbert space and in fact, by virtue of (5.5), they form an orthonormal set in it,

$$\langle \Phi_m^i | \Phi_n^j \rangle = \int_{\Omega} \Phi_m^i(\mathfrak{Z})^\dagger \Phi_n^j(\mathfrak{Z}) d\Omega(\mathfrak{Z}, \mathfrak{Z}^*) = \delta_{mn} \delta_{ij}.$$

Denote by \mathfrak{H}_K the Hilbert subspace of $\tilde{\mathfrak{H}}$ generated by this set of vectors. Then, in view of the convergence of the series in (5.7),

$$\sum_{i,k} \|\Phi_k^i(\mathfrak{Z}^*)\|^2 < \infty, \quad \forall \mathfrak{Z}^* \in \Omega.$$

Thus, \mathfrak{H}_K is a reproducing kernel Hilbert space of analytic functions in the variable \mathfrak{Z}^* , with matrix-valued kernel $K : \Omega \times \Omega \mapsto \mathbb{C}^{M \times M}$, given by (see (3.1))

$$\begin{aligned} K(\mathfrak{Z}^*, \mathfrak{Z}) &= \sum_{i,k} \Phi_k^i(\mathfrak{Z}') \Phi_k^i(\mathfrak{Z}^*)^\dagger = \sum_{i,k} \frac{\mathfrak{Z}^{*k} \chi^i \chi^{i\dagger} \mathfrak{Z}^k}{x_k!} \\ &= \sum_k \frac{\mathfrak{Z}^{*k} \mathfrak{Z}^k}{x_k!}. \end{aligned} \quad (5.8)$$

When $M = 1, \mathfrak{Z} = z, \Omega = \mathbb{C}$ and $x_k! = k!$, we get the well-known Bargmann kernel

$$K(\bar{z}', z) = e^{\bar{z}'z},$$

and \mathfrak{H}_K is the Hilbert space of entire analytic functions in the variable \bar{z} . This is the kernel associated with the canonical coherent states (4.3).

The vector coherent states associated with the reproducing kernel K are (see (3.3)) the vectors $|\mathfrak{Z}; i\rangle \in \mathfrak{H}_K$,

$$|\mathfrak{Z}; i\rangle(\mathfrak{Z}') = \mathcal{N}(\mathfrak{Z}^*, \mathfrak{Z})^{-\frac{1}{2}} K(\mathfrak{Z}^*, \mathfrak{Z}) \chi^i, \quad \mathcal{N}(\mathfrak{Z}^*, \mathfrak{Z}) = \frac{K(\mathfrak{Z}, \mathfrak{Z}^*)}{M} \quad (5.9)$$

defined for each $\mathfrak{Z} \in \Omega$ and $i = 1, 2, \dots, M$. Note that since $K(\mathfrak{Z}^*, \mathfrak{Z})$ is a strictly positive-definite matrix,

$$K(\mathfrak{Z}^*, \mathfrak{Z}) = U \text{diag}[\mathcal{N}(\lambda_1), \mathcal{N}(\lambda_2), \dots, \mathcal{N}(\lambda_M)] U^*, \quad (5.10)$$

where for each $i, \mathcal{N}(\lambda_i) = \sum_{k=0}^N \frac{\lambda_i^k}{x_k!}$ is the same normalization factor as in (3.8), the negative square root makes sense. The vector coherent states (5.9) satisfy the resolution of the identity (compare with (3.20)),

$$\sum_{i=1}^M \int_{\Omega} |\mathfrak{Z}; i\rangle \langle \mathfrak{Z}; i| \mathcal{N}(\mathfrak{Z}^*, \mathfrak{Z}) d\Omega(\mathfrak{Z}, \mathfrak{Z}^*) = I_K, \quad (5.11)$$

and the normalization condition

$$\sum_{i=1}^M \langle \mathfrak{z}; i | \mathfrak{z}; i \rangle = 1. \tag{5.12}$$

The kernel K has the matrix elements

$$K(\mathfrak{z}^*, \mathfrak{z})_{ij} = \chi^{i\dagger} K(\mathfrak{z}^*, \mathfrak{z}) \chi^j.$$

But also, in view of (5.5),

$$\begin{aligned} \langle \mathfrak{z}'; i | K(\mathfrak{z}^*, \mathfrak{z}) | \mathfrak{z}; j \rangle &= \int_{\Omega} \chi^{i\dagger} K(\mathfrak{z}^*, \mathfrak{x})^* K(\mathfrak{x}^*, \mathfrak{z}) \chi^j d\Omega(\mathfrak{x}, \mathfrak{x}^*) \\ &= \chi^{i\dagger} K(\mathfrak{z}^*, \mathfrak{z}) \chi^j = K(\mathfrak{z}^*, \mathfrak{z})_{ij}. \end{aligned} \tag{5.13}$$

Using (5.8) the VCS can alternatively written as

$$| \mathfrak{z}; i \rangle (\mathfrak{z}^*) = \mathcal{N}(\mathfrak{z}^*, \mathfrak{z})^{-\frac{1}{2}} \sum_k \frac{\mathfrak{z}^{*k} \mathfrak{z}^k \chi^i}{x_k!} = \mathcal{N}(\mathfrak{z}^*, \mathfrak{z})^{-\frac{1}{2}} \sum_{j,k} \frac{\mathfrak{z}^{*k} \chi^j}{\sqrt{x_k!}} \cdot \frac{\chi^{j\dagger} \mathfrak{z}^k \chi^i}{\sqrt{x_k!}},$$

so that

$$| \mathfrak{z}; i \rangle = \mathcal{N}(\mathfrak{z}^*, \mathfrak{z})^{-\frac{1}{2}} \sum_{j=1}^M \sum_{k=0}^N \Phi_k^j \frac{\chi^{j\dagger} \mathfrak{z}^k \chi^i}{\sqrt{x_k!}}. \tag{5.14}$$

Let \mathfrak{H} be an N -dimensional (complex, separable) Hilbert space and let $\{\phi_k\}_{k=0}^N$ be an orthonormal basis for it. Then the vectors $\chi^i \otimes \phi_k, i = 1, 2, \dots, M, k = 0, 1, 2, \dots, N$, form an orthonormal basis of $\mathbb{C}^M \otimes \mathfrak{H}$. We make a unitary transformation, $V : \mathfrak{H}_K \rightarrow \mathbb{C}^M \otimes \mathfrak{H}$, by the basis change $\Phi_k^i \mapsto \chi^i \otimes \phi_k$. Under this map, the VCS $| \mathfrak{z}; i \rangle$ transform to the vectors

$$\begin{aligned} | \mathfrak{z}, i \rangle^{\sim} := V | \mathfrak{z}, i \rangle &= \mathcal{N}(\mathfrak{z}^*, \mathfrak{z})^{-\frac{1}{2}} \sum_{j=1}^M \sum_{k=0}^N \chi^j \otimes \phi_k \frac{\chi^{j\dagger} \mathfrak{z}^k \chi^i}{\sqrt{x_k!}} \\ &= \mathcal{N}(\mathfrak{z}^*, \mathfrak{z})^{-\frac{1}{2}} \sum_{k=0}^N \frac{\mathfrak{z}^k \chi^i}{\sqrt{x_k!}} \otimes \phi_k \in \mathbb{C}^M \otimes \mathfrak{H}, \end{aligned} \tag{5.15}$$

which are exactly the VCS defined (over matrix domains) in [3]. Also, in this form the VCS resemble the nonlinear coherent states (3.8) more closely. The inverse of the map V is then easily seen to be given by

$$(V^{-1} \Phi)(\mathfrak{z}^*) = \sum_{i=1}^M \langle \mathfrak{z}, i | \Phi \rangle \chi^i, \quad \Phi \in \mathbb{C}^M \otimes \mathfrak{H}. \tag{5.16}$$

To return to the discussion of the probability distribution $P(\lambda_1, \lambda_2, \dots, \lambda_M; n)$ in (5.1), we first rewrite the VCS (5.15) explicitly in matrix form as

$$\begin{aligned} &| \mathfrak{z}, i \rangle^{\sim} \\ &= \frac{1}{\sqrt{M}} \sum_{k=0}^N U \begin{pmatrix} \sqrt{P(k, \lambda_1)} e^{-ik\theta_1} & 0 & \dots & 0 \\ 0 & \sqrt{P(k, \lambda_2)} e^{-ik\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{P(k, \lambda_M)} e^{-ik\theta_M} \end{pmatrix} U^* \chi^i \otimes \phi_k. \end{aligned} \tag{5.17}$$

Again, let $\mathbb{P}_n = |\phi_n\rangle\langle\phi_n|$ and define

$$\mathcal{P}(\mathfrak{Z}, \mathfrak{Z}^*; n) = \text{Tr}_{\mathfrak{H}} \left[\sum_{i=1}^M |\mathfrak{Z}, i\rangle\langle\mathfrak{Z}, i| \mathbb{I}_M \otimes \mathbb{P}_n \right], \tag{5.18}$$

where $\text{Tr}_{\mathfrak{H}}$ denotes a partial trace in \mathfrak{H} . Clearly, $\mathcal{P}(\mathfrak{Z}, \mathfrak{Z}^*; n)$ is an $M \times M$ matrix and it is not hard to see that

$$\mathcal{P}(\mathfrak{Z}, \mathfrak{Z}^*; n) = \frac{1}{M} U \begin{pmatrix} P(n, \lambda_1) & 0 & \dots & 0 \\ 0 & P(n, \lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P(n, \lambda_M) \end{pmatrix} U^*. \tag{5.19}$$

Now taking the trace in \mathbb{C}^M we immediately see that

$$\begin{aligned} \text{Tr}_{\mathbb{C}^M} [\mathcal{P}(\mathfrak{Z}, \mathfrak{Z}^*; n)] &= \text{Tr}_{\mathbb{C}^M \otimes \mathfrak{H}} \left[\sum_{i=1}^M |\mathfrak{Z}, i\rangle\langle\mathfrak{Z}, i| \mathbb{I}_M \otimes \mathbb{P}_n \right] \\ &= P(\lambda_1, \lambda_2, \dots, \lambda_M; n), \end{aligned} \tag{5.20}$$

which should be compared with (3.23). Finally, the determinant

$$\det [M\mathcal{P}(\mathfrak{Z}, \mathfrak{Z}^*; n)] = P(n, \lambda_1)P(n, \lambda_2) \dots P(n, \lambda_M) \tag{5.21}$$

denotes the joint probability of getting n ‘successes’ from each distribution.

Before leaving this topic of matrix-valued distributions, let us point out that more general situations than envisaged by (5.1) can also be treated using similar techniques. For example, instead of attaching the same weight, $\frac{1}{M}$, to each component $P(n, \lambda_i)$ of the mixture, we could also attach different weights μ_i to them (with $\mu_i > 0$ for all i and $\sum_{i=1}^M \mu_i = 1$). Examples of this type will be dealt with in a future publication, where we shall also allow the possibility of M being infinite.

To treat general joint probabilities of the type,

$$P(n_1, \lambda_1; n_2, \lambda_2; \dots; n_M, \lambda_M) = P(n_1, \lambda_1)P(n_2, \lambda_2) \dots P(n_M, \lambda_M), \tag{5.22}$$

it is necessary to go to multidimensional coherent states. We intend to treat this in greater detail in a future publication, but here we briefly indicate the main idea. Consider again a discrete distribution $P(n, \lambda)$ of the type (3.11), i.e., such that it has associated coherent states of the type (3.19). These coherent states $|z\rangle$ are defined on a Hilbert space \mathfrak{H} . Let $\mathfrak{H}^M = \mathfrak{H} \otimes \mathfrak{H} \otimes \dots \otimes \mathfrak{H}$ be the M -fold tensor product of \mathfrak{H} with itself. On \mathfrak{H}^M we define the vectors,

$$\begin{aligned} |z_1, z_2, \dots, z_M\rangle &= |z_1\rangle|z_2\rangle \dots |z_M\rangle \\ &= \sum_{n_1=0, n_2=0, \dots, n_M=0}^N [P(n_1, \lambda_1; n_2, \lambda_2; \dots; n_M, \lambda_M)]^{\frac{1}{2}} \\ &\quad \times e^{-i(n_1\theta_1+n_2\theta_2+\dots+n_M\theta_M)} \phi_{n_1, n_2, \dots, n_M}, \end{aligned} \tag{5.23}$$

where the vectors

$$\phi_{n_1, n_2, \dots, n_M} = \phi_{n_1} \otimes \phi_{n_2} \otimes \dots \otimes \phi_{n_M}, \quad 0 \leq n_1, n_2, \dots, n_M \leq N,$$

form an orthonormal basis for \mathfrak{H}^M . We call the vectors (5.23) *multidimensional coherent states*. Such coherent states have been studied in different contexts before (see, for example, [27]). These vectors are normalized,

$$\langle z_1, z_2, \dots, z_M | z_1, z_2, \dots, z_M \rangle = 1,$$

and they satisfy the resolution of the identity (compare with (3.20)),

$$\int_{\mathcal{D}^M} |z_1, z_2, \dots, z_M\rangle \langle z_1, z_2, \dots, z_M| \prod_{i=1}^M \mathcal{N}(\lambda_i) d\bar{\varrho}(\lambda_i) d\theta_i = I_{\mathcal{H}^M},$$

where $\mathcal{D}^M = \mathcal{D} \times \mathcal{D} \times \dots \times \mathcal{D}$ is the M -fold Cartesian product of the domain $\mathcal{D} = \{\sqrt{\lambda} e^{i\theta} \in \mathbb{C} | \lambda \in [0, L), \theta \in [0, 2\pi)\}$ over which the coherent states $|z\rangle$ in (3.19) are defined.

Once again these coherent states appear as ‘generalized square-roots’ of the joint probability distribution $P(n_1, \lambda_1; n_2, \lambda_2; \dots; n_M, \lambda_M)$, $0 \leq n_1, n_2, \dots, n_M \leq N$, and just as in (3.23),

$$\begin{aligned} P(n_1, \lambda_1; n_2, \lambda_2; \dots; n_M, \lambda_M) &= \text{Tr}[|z_1, z_2, \dots, z_M\rangle \langle z_1, z_2, \dots, z_M| \mathbb{P}_{n_1, n_2, \dots, n_M}] \\ &= |\langle \phi_{n_1, n_2, \dots, n_M} | z_1, z_2, \dots, z_M \rangle|^2, \end{aligned} \tag{5.24}$$

with

$$\mathbb{P}_{n_1, n_2, \dots, n_M} = |\phi_{n_1, n_2, \dots, n_M}\rangle \langle \phi_{n_1, n_2, \dots, n_M}|.$$

Recently (see [10]), this formalism has been applied to the construction of vector coherent states for the quantum motion of a particle in an infinite square well, enabling one to define in an unambiguous way the momentum operator. The construction in [10] is based on Gaussian probability distributions but it can be carried out using a large class of distributions.

6. Conclusion

As mentioned in the introduction, the relationship between coherent states and statistical distributions has been studied before. We have tried to demonstrate here the deeper connection between such distributions, both continuous and discrete, and reproducing kernel Hilbert spaces, in so far as the latter are the carriers of generalized coherent states. Moreover, taking this point of view, it has been possible to connect vector coherent states to mixtures of probability distributions and multi-dimensional coherent states to joint probability distributions. The posterior distribution, appearing on the parameter space of a discrete distribution, is clearly seen to be a consequence of the resolution of the identity satisfied by the coherent states. Again this has been noted earlier, but here we are able to put it in a more general context.

While we do have in mind possible implications to quantum physics of the connection between Bayesian duality and coherent states, we defer a detailed discussion of that point to a future publication. The specific examples, chosen for the sake of illustration in section 4, were dictated by the fact that they involve, in a sense, the most commonly used statistical distributions as well as the most important coherent states from a physical point of view. However, the one intriguing question that arises from the general discussion is the following: as has been demonstrated, a discrete statistical distribution, or a family of discretely parametrized continuous distributions, satisfying certain technical conditions, lead to the existence of coherent states on an associated Hilbert space. These coherent states, in turn, can be shown to lead to quantum probabilities, embodied in a positive operator valued measure, on the parameter space. The nature of classical (commutative) and quantum (non-commutative) probability are intrinsically different, yet it seems to be possible to make a smooth transition from one to the other. This is reminiscent of the process of quantization, i.e., the passage from a classical-mechanical system to its quantum counterpart, and in particular, coherent state quantization (see, for example, [2] for a review of the theory of quantization and [10, 13–16] for a series of examples). So one might ask the question as to whether the procedure

described above could be considered as constituting a quantization of the underlying classical probability theory. In this connection, it would also be interesting to study more closely the duality appearing between the discrete and continuous distributions incorporated in the coherent states and the analogous duality familiar from Bayesian statistics.

Acknowledgments

The work of one of the authors (STA) was partially supported through grants from the Natural Sciences and Engineering Research Council (NSERC), Canada and the Fonds québécois de la recherche sur la nature et les technologies (FQRNT), Québec. The authors would also like to acknowledge useful discussions with M Wang.

Appendix. Some elements of Bayesian inference

In this appendix, we put together some notions from Bayesian statistical inference that have been used in this paper. Some relevant references are [6, 9, 20, 21, 23, 26, 30, 32]

A.1. Event space background

The context is the setup and subsequent performance of an experiment where there is a random component to the results and where the set U of possible results is known. In the field of statistics, the experiment is called a ‘random experiment’. Events are identified with measurable subsets of U . That is, we say that event E has occurred if the observed result u_{obs} is in the subset E . One ‘experiment’, of course, could be an amalgam of a whole set of sub-experiments, sometimes called ‘trials’.

A.2. Classical conditional probabilities

Let $P(E|B)$ designate the classical conditional probability that event E occurs given that event B has occurred. Then

$$P(E|B) = \frac{P(E \cap B)}{P(B)},$$

where the numerator stands for the joint probability of occurrence of events E and B and the denominator is the unconditional probability of occurrence of event B (to ensure normalization). Consider the conditional probability the other way around $P(B|E)$,

$$P(B|E) = \frac{P(E \cap B)}{P(E)}.$$

Suppose that we do not know the joint probability and in fact we only know the first conditional probability $P(E|B)$ and the two unconditional probabilities, then we can write

$$P(B|E) = \frac{P(E|B)P(B)}{P(E)}.$$

The probability $P(B|E)$ is called the *posterior conditional probability* for B given E and $P(B)$ is called the *prior probability* of B . Sometimes we compute several of these posterior probabilities in the cases where the set of events $\{B_1, B_2, \dots, B_n\}$ is a partition of U and the events B_i are in the nature of possible causal hypotheses for the subsequent occurrence of event E . Suppose that we know the conditional probabilities $P(E|B_i)$ and the unconditional (prior) probabilities $P(B_i)$ for each B_i . Then one chooses a likely hypothesis by computing each of the posterior probabilities.

Quantum analogues are given, for example, in [9, 23, 26, 32].

A.3. The case of a continuous family of discrete probability distributions

Consider the performance of a classical experiment in which the outcome has a random component within the following context. Let $n = 0, 1, 2, \dots, N$ index the (discrete) set of possible outcomes of the experiment, where N is a positive integer or ∞ . For real parameter $\lambda \in \Lambda$, let $P(n, \lambda)$ be a family of classical discrete probability distributions indexed by λ , which serves as a stochastic model for the experiment. We suppose that λ is unknown and the object of the experiment is to obtain data with which to infer a probability distribution on the parameter space Λ . After the performance of the experiment, let k indicate the observed outcome. Then construct a conditional probability density function f for λ , given k , in the form

$$f(\lambda, k) = \frac{P(k, \lambda)\Pi(\lambda)}{\int_{\Lambda} P(k, \lambda')\Pi(\lambda') d\lambda'},$$

where $\Pi(\lambda)$ is an unconditional probability measure on the parameter space Λ , arbitrary, subject to the integrability of the denominator. The measure Π is called the *prior measure* on Λ and the conditional probability density function f is called the density function of the *posterior probability distribution* on Λ .

Example. Toss a coin N times observing n , the number of occurrences of heads. Let the parameter p be the probability of obtaining heads on one toss. Supposing that p is unknown, the object is to use the outcome of the experiment to obtain a probability distribution on the parameter space $(0, 1)$. The stochastic model is the binomial family,

$$P(n, p) = \frac{N!}{(N-n)!n!} p^n (1-p)^{(N-n)}, \quad \text{for } n = 0, 1, 2, \dots, N,$$

where N is a positive integer. After the performance of the experiment, having obtained k heads, with choice of prior measure $\Pi(p)$, the posterior distribution on $(0, 1)$ is given by the conditional probability density function,

$$f(p, k) = \frac{p^k (1-p)^{(N-k)} \Pi(p)}{\int_0^1 p'^k (1-p')^{(N-k)} \Pi(p') dp'}.$$

References

- [1] Arrecchi F T, Courtens E, Gilmore R and Thomas H 1972 Atomic coherent states in quantum optics *Phys. Rev. A* **6** 2211–37
- [2] Ali S T and Engliš M 2005 Quantization methods: a guide for physicists and analysts *Rev. Math. Phys.* **17** 391–490
- [3] Ali S T, Engliš M and Gazeau J-P 2004 Vector coherent states from Plancherel's theorem, Clifford algebras and matrix domains *J. Phys. A: Math. Gen.* **37** 6067–89
- [4] Ali S T, Antoine J-P and Gazeau J-P 2000 *Coherent States, Wavelets and their Generalizations* (New York: Springer)
- [5] Ali S T, Balkova L, Curado E M F, Gazeau J-P, Rego-Monteiro M A, Rodrigues L M C S and Sekimoto K Non-commutative reading of the complex plane through Delaunay sequences (in preparation)
- [6] Kotz S, Balakrishnan N, Read C B and Vidakovic B (editors-in-chief) 2006 *Encyclopaedia of Statistical Sciences* 2nd edn, vol 1 (Hoboken, NJ: Wiley-Interscience) (under the entry 'Bayesian Inference')
- [7] Kotz S, Balakrishnan N, Read C B and Vidakovic B (editors-in-chief) 2006 *Encyclopaedia of Statistical Sciences* 2nd edn, vol 2 (Hoboken, NJ: Wiley-Interscience) (under the entry 'Conjugate Families of Distributions')
- [8] Box G E P and Tiao G C 1973 *Bayesian Inference in Statistical Analysis* (Wiley Classics Library)
- [9] Caves C M, Fuchs C A and Schack R 2002 Unknown quantum states: the quantum de Finetti representation *J. Math. Phys.* **43** 4537–59
- [10] Garcia de Leon P, Gazeau J-P and Quéva J 2008 The infinite well revisited: coherent states and quantization *Phys. Lett. A* **372** 3597–607

- [11] Gazeau J-P and Klauder J R 1999 Coherent states for systems with discrete and continuous spectrum *J. Phys. A: Math. Gen.* **32** 123–32
- [12] Gazeau J-P 2005 *Quantization is Just a Certain Regard to...* (*The Old and New Concepts of Physics* vol 2) pp 1–37 http://www.conceptsofphysics.net/cphys_II.1-2.htm
- [13] Gazeau J-P and Piechocki W 2004 Coherent states quantization of a particle in de Sitter space *J. Phys. A: Math. Gen.* **37** 6977–86
- [14] Gazeau J-P, Josse-Michaux F-X and Monceau P 2006 Finite dimensional quantization of the (q, p) -plane: new space and momentum inequalities *Int. J. Mod. Phys. B* **20** 1778–91
- [15] Gazeau J-P and García de León P L 2007 Coherent state quantization and phase operator *Phys. Lett. A* **361** 301–4
- [16] Gazeau J-P, Huguet E, Lachièze Rey M and Renaud J 2007 Fuzzy spheres from inequivalent coherent states quantizations *J. Phys. A: Math. Theor.* **40** 10225–49
- [17] Heller B and Wang M 2007 Posterior distribution for negative binomial parameter p using a group invariant prior *Stat. Prob. Lett.* **77** 1542–8
- [18] Heller B and Wang M 2006 Group invariant inferred distributions via noncommutative probability *Recent Developments in Nonparametric Inference and Probability (IMS Lecture Notes—Monograph Series* vol 50) pp 1–19
- [19] Heller B and Wang M 2004 Posterior distributions on certain parameter spaces by using group theoretic methods adopted from quantum mechanics *Technical Report No 546* (Univ. of Chicago, Dept. of Statistics Technical Report Series) 26 pp
- [20] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic) (chapters II and VIII)
- [21] Holevo A S 2001 *Statistical Structure of Quantum Theory* (Berlin: Springer) p 46
- [22] Johnson N L and Kotz S 1969 *Distributions in Statistics, Discrete Distributions* (New York: Wiley)
- [23] Jones K F W 1991 Principles of quantum inference *Ann. Phys., NY* **207** 140–70
- [24] Klauder J R, Penson K A and Sixdeniers J-M 2001 Constructing coherent states through solutions of Stieltjes and Hausdorff moment problems *Phys. Rev. A* **64** 013817-(1–18)
- [25] Man'ko V I, Marmo G, Sudarshan E C G and Zaccaria F 1997 f -oscillators and non-linear coherent states *Phys. Scr.* **55** 528–41
- [26] Neri F 2005 Quantum Bayesian methods and subsequent measurements *Phys. Rev. A* **72** 062306-(1–6)
- [27] Novaes M and Gazeau J-P 2003 Multidimensional generalized coherent states *J. Phys. A: Math. Gen.* **36** 199–212
- [28] Odziejewicz A 1998 Quantum algebras and q -special functions related to coherent states maps of the disc *Commun. Math. Phys.* **192** 183–215
- [29] Perelomov A M 1986 *Generalized Coherent States and their Applications* (Berlin: Springer)
- [30] Peres A 1995 *Quantum Theory: Concepts and Methods* (Dordrecht: Kluwer) p 45
- [31] Radcliffe J M 1971 Some properties of coherent spin states *J. Phys. A: Gen. Phys.* **4** 313–23
- [32] Schack R, Brun T A and Caves C M 2001 Quantum Bayes rule *Phys. Rev. A* **64** 014305-(1–4)
- [33] Simon B 1998 The classical moment problem as a self-adjoint finite difference operator *Adv. Math.* **137** 82–203
- [34] Thirulogasanthar K and Ali S T 2003 A class of vector coherent states defined over matrix domains *J. Math. Phys.* **44** 5070–83