Sufficient conditions for finite dimensionality of filters in discrete time: a Laplace transform-based approach

WOLFGANG J. RUNGGALDIER[1] and FABIO SPIZZICHINO[2]

[1]*Dipartimento di Matematica Pura e Applicata, Universitá di Padova, Via Belzoni 7, 35131 Padova, Italy. E-mail: runggal@math.unipd.it*
[2]*Dipartimento di Matematica, Universitá di Roma 'La Sapienza', 00185 Roma, Italy. E-mail: fabio.spizzichino@uniroma1.it*

The discrete-time filtering problem can be seen as a dynamic generalization of the classical Bayesian inference problem. For practical applications it is important to identify filtering models that, analogously to the linear Gaussian model (Kalman filter), admit a finite-dimensional filter or, equivalently, a finite-dimensional family of filter-conjugate distributions. Our main purpose here is to give sufficient conditions for the existence of finite-dimensional filters. We use a method, based on the Laplace transform, which is also constructive.

Keywords: dynamic Bayes formula; exponential families; finite-dimensional filters; infinitely divisible distributions; inverse Laplace transform; state-space models

1. Introduction

The object of our study are partially observable systems in discrete-time, namely $\{X_n, Y_n\}_{n=0,1,\ldots}$, where $\{X_n\}$ represents the unobservable component, also called a latent state process, while $\{Y_n\}$ is the observable component. More precisely, the model is as follows.

Let I and J be measurable subsets of some finite-dimensional real linear spaces with $0 \in J$. Consider a probability kernel $F(x, \mathrm{d}y)$ from I to J and a sequence $\{P_n\}_{n \geq 1}$ of transition kernels from I to I together with a probability measure Π_0 on I. Consider a non-homogeneous Markov chain $\{X_n, Y_n\}$ with state space $I \times J$, initial distribution $\Pi_0(\mathrm{d}x)\delta_0(\mathrm{d}y)$, where δ_0 denotes the point mass at zero, and transition kernel from (x, y) to (x', y'), between the times $n - 1$ and n, given by $P_n(x, \mathrm{d}x')F(x', \mathrm{d}y')$. The marginal process $\{X_n\}_{n \geq 0}$ forms a non-homogeneous Markov chain with state space I, initial distribution Π_0 and kernel P_n, and describes the evolution of the latent state process. On the other hand, the marginal process $\{Y_n\}_{n \geq 0}$, with $Y_0 = 0$, describes the observations and Y_n is, conditionally on X_n, independent of its own past $Y_0, Y_1, \ldots, Y_{n-1}$ and of that of the latent state process as well.

Given such a partially observable model, we shall consider the filtering problem that consists in determining, for each period n, the conditional distribution of the latent state X_n,

given the past and current observations $y_0^n \equiv (y_0, \ldots, y_n)$. This problem is a natural dynamic generalization of the classical Bayesian inference problem, where the latter is obtained as a particular case when $P\{X_0 = X_1 = \ldots = X_n\} = 1$ for all $n \geqslant 1$. It can also be viewed as a generalization of the estimation problem in dynamic linear models (see, for example, West *et al.* 1985).

For convenience of presentation we shall consider the case when $I, J \subseteq \mathbb{R}$ and assume that all distributions we shall be dealing with are absolutely continuous with respect to Lebesgue measure. The elements characterizing the partially observable process $\{X_n, Y_n\}$ are then: $\pi_0(x)$, the (initial) density of x_0; $p_n(x'|x)$, the transition density in period n; and $f(y|x)$, the observation density. On the other hand, the filter solution corresponds, for $n = 1, 2, \ldots$, to the conditional densities of the latent state, given the observation history $\pi_n(x|y_0^n)$, also called *filtering densities*.

Analogously to what is the main tool in Bayesian statistical inference, namely Bayes's formula, for a discrete-time filtering problem the main tool is the *recursive (dynamic) Bayes formula*

$$\pi_n(x'|y_0^n) \propto f(y_n|x') \int_I p_n(x'|x) \pi_{n-1}(x|y_0^{n-1}) \, dx. \tag{1}$$

To actually perform the computations required in (1), it is important to have partially observable models admitting so-called *finite-dimensional filters*, where, in every period n, the filter density $\pi_n(x|y_0^n)$ belongs to a same family of densities, parametrized by a finite-dimensional parameter. By analogy with Bayesian statistics and generalizing the situation in the classical linear Gaussian models (Kalman filter model; see, for example, Liptser and Shiryaev 1977), we give the following definition:

Definition 1.1. *Let the pair* $\{f(y|x), (p_n(x'|x)_{n \geqslant 1})\}$ *be given.* $\Gamma := \{\gamma_\theta(x)\}_{\theta \in \Theta}$ *is a family of densities* filter-conjugate *to* $\{f(y|x), (p_n(x'|x)_{n \geqslant 1})\}$ *if there exists* $\Phi : \Theta \times J \to \Theta$ *such that, for all* $\theta \in \Theta$, $y \in J$, $n = 1, 2, \ldots,$

$$f(y|x') \int_I p_n(x'|x) \gamma_\theta(x) \, dx \propto \gamma_{\Phi(\theta, y)}(x). \tag{2}$$

If a filter-conjugate family exists for a pair $\{f(y|x), (p_n(x'|x)_{n \geqslant 1})\}$ then, if the initial density $\pi_0(x)$ belongs to this filter-conjugate family, by (1) and (2) at every step the filter density $\pi_n(x|y_0^n)$ also belongs to this family. Furthermore, the updating of the filter density from one step to the next can then be accomplished by a simple transformation of the indexing parameter θ, that is, by the mapping $\Phi(\theta, y) : \Theta \times J \to \Theta$.

A finite-dimensional filter corresponds to having Θ finite-dimensional.

In Bayesian inference it is well known (Barndorff-Nielsen 1978; Barndorff-Nielsen and Pedersen 1968) that, under regularity assumptions, $\{f(y|x)\}$ admits a conjugate family of finite order if and only if it is an exponential family, and in this case the conjugate family is also exponential. Correspondingly, for discrete-time filtering problems we have (Ferrante and Runggaldier 1990; Sawitzki 1981) that, under regularity assumptions, $\{f(y|x), (p_n(x'|x)_{n \geqslant 1})\}$ admits a finite-dimensional filter-conjugate family only if $\{f(y|x)\}$ is an exponential family. In this case $\pi_n(x|y_0^n)$ is an exponential family as well.

The purpose of the paper is to study the following problem: given $f(y|x)$ belonging to an exponential class, determine transition kernels $p_n(x'|x)$ such that $\{f(y|x), (p_n(x'|x)_{n \geq 1})\}$ admits a finite-dimensional filter (a finite-dimensional filter-conjugate family).

To the best of our knowledge, only necessary conditions on the transition kernels have so far been obtained (see Ferrante 1993). In this paper we present a method that, in addition to guaranteeing existence, allows us also to determine the kernels explicitly. Our problem is thus in a sense the inverse of the usual filtering problem where, given an evolution model for (X_n, Y_n), one determines the filter distribution and its dynamics. In fact, given a model for Y_n, we first choose a family of filter distributions, namely the natural conjugate family (see (6) below). Having assigned the filter dynamics, we then determine a model for X_n so that the filtering problem for the pair (X_n, Y_n) has the solution in the given family. In this form our problem can also be seen as a filter analogue of the Bayesian statistics problem studied in Bar-Lev *et al.* (1994). Finally, we mention that a specific constructive approach, allowing finite-dimensional filters to be obtained in some cases, can be found in Ferrante and Vidoni (1998) (see also references therein).

2. Main results

Motivated by the necessary conditions for the existence of finite-dimensional filters as recalled in Section 1, we specialize to the case where $\mathscr{F} \equiv \{F(x, \mathrm{d}y)\}$ constitutes a natural exponential family (see Diaconis and Ylvisaker 1979) of order 1, that is, where the kernel $F(x, \mathrm{d}y)$ has a density of the form

$$f(y|x) = a(x)b(y)\mathrm{e}^{xy}, \tag{3}$$

in which $b(y)$ is any positive function defined on J. Consequently, we let

$$I := \left\{ x \in \mathbb{R} \,\middle|\, \int_J \mathrm{e}^{xy}b(y)\,\mathrm{d}y < +\infty \right\}. \tag{4}$$

Let, furthermore, $\overline{M}$ be the subset of points (α, z) in $\mathbb{R} \times \mathbb{R}$ such that

$$\lambda^{-1}(\alpha, z) := \int_I [a(x)]^\alpha \mathrm{e}^{xz}\,\mathrm{d}x < +\infty. \tag{5}$$

It then follows immediately that the family

$$\mathscr{G} := \{ g_{a,z}(x) = \lambda(\alpha, z)a(x)^\alpha \mathrm{e}^{xz}; \ (\alpha, z) \in \overline{M} \} \tag{6}$$

is a natural conjugate family of $\mathscr{F}$, that is, for x viewed as a parameter.

We shall make the following assumption:

Assumption 1. *The set $\overline{M}$ contains a rectangle $M = A \times Z$ such that*

- $\alpha \in A \Rightarrow \alpha + 1 \in A$,
- $z \in Z, \ y \in J \Rightarrow z + y \in Z$.

Fix two functions $\phi : A \to A$, $\psi : M \to Z$ and, given $(\alpha_0, z_0) \in M$ as well as a sequence $y_1, y_2, \ldots \in J$, define recursively, for $n = 1, 2, \ldots,$

$$\alpha_n = \phi(\alpha_{n-1}) + 1, \qquad z_n = \psi(\alpha_{n-1}, z_{n-1}) + y_n. \tag{7}$$

By Assumption 1 we have $\alpha_n \in A$, $z_n \in Z$, for all n. Define, moreover,

$$\tau_{\phi,\psi}(\alpha, z) := \frac{\lambda(\phi(\alpha), \psi(\alpha, z))}{\lambda(\alpha, z)}, \tag{8}$$

with λ as in (5).

Theorem 2.1. *Suppose that, for all $\alpha \in A$, $K_\alpha : I \times I \to \mathbb{R}^+$ is a kernel satisfying*

$$\int_I K_\alpha(x, x')\, e^{xz}\, dx = \tau_{\phi,\psi}(\alpha, z) e^{x'\psi(\alpha,z)}, \qquad \text{for all } x' \in I, z \in Z. \tag{9}$$

Then:

(i) $p_\alpha(x'|x) := ([a(x')]^{\phi(\alpha)}/[a(x)]^\alpha)K_\alpha(x, x')$ *is a transition kernel from I to I for all $\alpha \in A$.*

(ii) *If the initial distribution Π_0 has a density $g_{\alpha_0,z_0} \in \mathscr{G}$ (for some $(\alpha_0, z_0) \in M$), $y_1, y_2, \ldots$ are the actual observations and (α_n, z_n) are as in (7) then, letting*

$$p_n(x'|x) = p_{\alpha_{n-1}}(x'|x), \tag{10}$$

the family $\mathscr{G}$ in (6), which is the natural conjugate family of $\mathscr{F}$, is a filter-conjugate family according to Definition 1.1. Thus, for all $x \in I$ and for all $n \geqslant 1$,

$$\pi_n(x|y_1^n) = g_{\alpha_n,z_n}(x). \tag{11}$$

Proof. (i) We must show that

$$\int_I p_\alpha(x'|x)\, dx' = 1, \qquad \forall \alpha \in A, \qquad \forall x \in I, \tag{12}$$

or, equivalently,

$$\int_I [a(x')]^{\phi(\alpha)} K_\alpha(x, x')\, dx' = [a(x)]^\alpha, \qquad \forall x \in I. \tag{13}$$

By (9) and the definitions of λ in (5) and of τ in (8), the Laplace transform of the left-hand side of (13) is given by

$$\int_I \left[\int_I [a(x')]^{\phi(\alpha)} K_\alpha(x, x')\, dx' \right] e^{zx}\, dx = \lambda^{-1}(\alpha, z). \tag{14}$$

Equation (13) now follows from the uniqueness of the inverse of the Laplace transform.

(ii) From the form of $\mathscr{F}$ and $\mathscr{G}$ (see (3) and (6)), and having put (see (10)) $p_n = p_{\alpha_{n-1}}$, it immediately follows that $\mathscr{G}$ is filter-conjugate in the sense of Definition 1.1. Indeed, relation (2) holds with $\Theta = M$, $\theta = (\alpha, z)$, $\Phi(\alpha, z) = (\phi(\alpha) + 1, \psi(\alpha, z) + y)$, $\gamma_\theta = g_{\alpha,z}$.

Relation (11) then follows immediately by induction on $n = 0, 1, \ldots$, taking into account the dynamic Bayes formula (1) (see also the comment after Definition 1.1). $\qquad\square$

The process $\Theta_n := (\alpha_n, Z_n)$, where α_n and Z_n satisfy (7) with the random sequence $Y_1, Y_2, \ldots$ replacing a specific realization $y_1, y_2, \ldots$, will be called (see Van Schuppen 1979) a filter process. This process depends on the choice of the functions ϕ and ψ which (see Bather 1965) we shall also call *connecting functions*.

Given the family $\mathscr{F}$ and, therefore, its natural conjugate family $\mathscr{G}$, every choice of the connecting functions determines a specific sequence of transition densities $p_n = p_{\alpha_{n-1}}$ according to Theorem 2.1. Notice that this implies that the Markov chain $\{X_n\}$ is in general non-homogeneous. Note also that there may not exist for every choice of the connecting functions a solution of the crucial equation (9). The solvability of (9) hinges upon sufficient conditions on the connecting functions such that the right-hand side in (9) is indeed the Laplace transform of a positive measure for any $\alpha \in A$, $x' \in I$. In what follows we study the solvability of (9) in each of the two special cases (a) $\psi(\alpha, z) = c(\alpha)z$, (b) $I = [0, +\infty)$. The peculiarity of these two cases stems from the following fact. The kernels $K_\alpha(x, x')$ in the left-hand side of (9) can be obtained by means of a convolution integral, having applied the inverse Laplace transform separately to each of the two factors $\tau(\alpha, z)$ and $\exp\{\psi(\alpha, z)x'\}$ in the right-hand side of (9). The two special cases allow the invertibility of the second factor $\exp\{\psi(\alpha, z)x'\}$ to be studied more easily.

For case (a) we have:

Proposition 2.2. *Given $\phi(\alpha)$ and $c(\alpha)$, assume that, for all $\alpha \in A$, we can find v_α: $\mathbb{R} \to [0, +\infty)$ satisfying the following conditions:*

(i) *Denoting $V_v := \{x \in \mathbb{R}|v(x) > 0\}$, we have $V_{v_\alpha} + c(\alpha) \cdot I \subseteq I$.*
(ii) *$\int_{V_{v_\alpha}} v_\alpha(x)e^{xz}\,dx = \tau_{\phi,\psi}(\alpha, z)$, for all $\alpha \in A$, $z \in Z$.*

Then, for all $\alpha \in A$,

$$K_\alpha(x, x') = v_\alpha(x - c(\alpha)x') \tag{15}$$

satisfies equation (9).

Proof. We have

$$\int_I K_\alpha(x, x')e^{xz}\,dx = \int_{I \cap \{V_{v_\alpha} + c(\alpha)x'\}} v_\alpha(x - c(\alpha)x')e^{xz}\,dx$$

$$= \int_{I \cap \{V_{v_\alpha} + c(\alpha)x'\} - c(\alpha)x'} v_\alpha(\xi)\exp\{(\xi + c(\alpha)x')z\}\,d\xi$$

$$= \exp\{c(\alpha)x' \cdot z\}\int_{I \cap \{V_{v_\alpha} + c(\alpha)x'\} - c(\alpha)x'} v_\alpha(\xi)e^{\xi z}\,d\xi.$$

Now, from condition (i), $I \cap \{V_{v_\alpha} + c(\alpha)x'\} = V_{v_\alpha} + c(\alpha)x'$, for any $x' \in I$, whence $I \cap \{V_{v_\alpha} + c(\alpha)x'\} - c(\alpha)x' = V_{v_\alpha}$ (see also Spizzichino 1990, Lemma 2.2).

The proposition then follows by condition (ii) and by taking into account that, in the case $\psi(\alpha, z) = c(\alpha)z$, the right-hand side of equation (9) is just $\tau_{\phi,\psi}(\alpha, z)e^{x'c(\alpha)z}$.   □

For case (b), since $I = [0, +\infty)$, we assume $Z \subset (-\infty, 0]$, $\psi(\alpha, 0) = 0$, for all $\alpha \in A$, and use the notation (where, for simplicity we drop, the subscripts ϕ and ψ from $\tilde{\tau}$)

$$\tilde{\psi}(\alpha, z) := -\psi(\alpha, -z), \qquad \tilde{\tau}(\alpha, z) := \frac{\lambda(\phi(\alpha), \tilde{\psi}(\alpha, z))}{\lambda(\alpha, -z)}.$$

Proposition 2.3. *Let $\phi(\alpha)$ and $\tilde{\psi}(\alpha, z)$ be such that:*

(i) *for all $\alpha \in A$, $\tilde{\psi}(\alpha, z)$ possesses partial derivatives $\tilde{\psi}^{(n)}(\alpha, z)$ of all orders with respect to z and the partial derivative $\tilde{\psi}^{(1)}(\alpha, z)$ is completely monotone, that is,*

$$(-1)^{n+1}\tilde{\psi}^{(n)}(\alpha, z) \geq 0, \qquad n \geq 1;$$

(ii) *$\tilde{\tau}(\alpha, z)$ possesses partial derivatives $\tilde{\tau}^{(n)}(\alpha, z)$ of all orders with respect to z, for all $\alpha \in A$, and, for some $C > 0$,*

$$0 \leq \frac{(-1)^n\tilde{\tau}^{(n)}(\alpha, z)}{n!} \leq C.$$

Then, for all $\alpha \in A$, there exists a probability measure P_α over $[0, +\infty)$ and a bounded function v_α, such that

$$K_\alpha(x, x') = \int_0^\infty v_\alpha(x - u)P_\alpha^{x'*}(\mathrm{d}u) \tag{16}$$

satisfies equation (9)*, where P^{r*} denotes the rth convolution power of P.*

Proof. It is a well-known result (see Feller 1970, Ch. 13) that condition (i), together with $\psi(\alpha, 0) = 0$, characterizes $\exp\{-\tilde{\psi}(\alpha, t)\}$ as the Laplace transform of an infinitely divisible probability distribution P_α over $[0, +\infty)$. Then, under (i),

$$\exp\{-\tilde{\psi}(\alpha, z) \cdot x'\} = \exp\{\psi(\alpha, -z) \cdot x'\}$$

is the Laplace transform of the probability distribution $(P_\alpha)^{x'*}$.

On the other hand (Feller 1970, Ch. 13), (ii) is a necessary and sufficient condition for $\tilde{\tau}(\alpha, z)$ to be the ordinary Laplace transform of a non-negative bounded function v_α: $[0, +\infty) \to (0, +\infty)$. Then, for all $\alpha \in A$ and for all $z \in Z$, the Laplace transform of the function $K_\alpha(x, x')$ defined by (16) is given by the product $\tilde{\tau}(\alpha, z) \cdot \exp\{\psi(\alpha, z) \cdot x'\}$, that is, equation (9) holds.   □

When both cases (a) and (b) hold simultaneously, one can obviously obtain the kernels $K_\alpha(x, x')$ (for a given choice of the functions ϕ and c) by means of either Proposition 2.2 or Proposition 2.3. In this case, conditions (i) in Proposition 2.2 is trivially satisfied for any function v_α vanishing outside $(0, +\infty)$. Condition (ii) in the same Proposition holds (with $Z \subset (-\infty, 0]$) if and only if $\tilde{\tau}(\alpha, z)$ is completely monotone (Feller 1970, Ch. 13).

Furthermore, condition (i) in Proposition 2.3 is trivially verified in that $\tilde{\psi}(\alpha, z) = c(\alpha) \cdot z$ and, indeed, P_α is then the Dirac measure with mass concentrated in $c(\alpha)$.

3. Examples

Example 1. As an example of the application of Proposition 2.2 (case (a)), consider the observation model

$$Y_n = X_n + W_n \tag{17}$$

with $\{W_n\}$ independent and identically distributed standard Gaussian, so that $\mathscr{F}$ is defined by

$$f(y|x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \exp\left\{-\frac{y^2}{2}\right\} \exp\{-xy\}, \tag{18}$$

whence $a(x) = \exp\{-x^2/2\}$ and the natural conjugate family is that of the Gaussian densities. We therefore take $I = J = \mathbb{R}$, $A = (0, +\infty)$, $Z = \mathbb{R}$, which trivially ensures the validity of Assumption 1.

We now choose the (non-negative) function $\phi(\alpha)$ arbitrarily and put $\psi(\alpha, z) = c(\alpha) \cdot z$, with $c(\cdot): A \to \mathbb{R}$ an arbitrary function satisfying (see (20) below)

$$|c(\alpha)| < \sqrt{\frac{\phi(\alpha)}{\alpha}}. \tag{19}$$

With such a choice, we have $\lambda(\alpha, z) = (\alpha/2\pi)^{1/2} \exp\{-z^2/2\alpha\}$, whence

$$\tau_{\phi,\psi}(\alpha, z) = \sqrt{\frac{\phi(\alpha)}{\alpha}} \exp\{-\tfrac{1}{2}Q(\alpha) \cdot z^2\},$$

having put

$$Q(\alpha) := \frac{\phi(\alpha) - \alpha c^2(\alpha)}{\alpha\phi(\alpha)}, \tag{20}$$

which is positive by (19). It is thus easily seen that

$$v_\alpha(x) := \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\phi(\alpha)}{\alpha Q(\alpha)}} \exp\left\{-\frac{x^2}{2Q(\alpha)}\right\} \tag{21}$$

satisfies condition (ii) of Proposition 2.2, whereas condition (i) is trivially satisfied in that $I = \mathbb{R}$. Then $K_\alpha(x, x') := v_\alpha(x - c(\alpha)x')$ satisfies equation (9) and from Theorem 2.1 we obtain, as the corresponding transition kernel,

$$p_\alpha(x'|x) = \frac{[a(x')]^{\phi(\alpha)}}{a^\alpha(x)} v_\alpha(x - c(\alpha)x')$$

$$= \sqrt{\frac{\phi(\alpha)}{2\pi\alpha Q(\alpha)}} \exp\left\{-\frac{x'^2}{2}\left(\phi(\alpha) + \frac{c^2(\alpha)}{Q(\alpha)}\right)\right\} \exp\left\{-\frac{x^2}{2}\left(\frac{1}{Q(\alpha)} - \alpha\right)\right\} \exp\left\{\frac{x \cdot x' \cdot c(\alpha)}{Q(\alpha)}\right\}.$$

Taking into account the definition of $Q(\alpha)$ in (20), after some manipulations we obtain that $p_\alpha(x'|x)$ coincides with the density function of a Gaussian distribution with mean value $(\alpha c(\alpha)/\phi(\alpha))x$ and variance $\alpha Q(\alpha)/\phi(\alpha)$, which corresponds to a state-evolution model of the form

$$X_{n+1} = C(\alpha_n)X_n + \sigma(\alpha_n)V_{n+1} \tag{22}$$

with

$$C(\alpha_n) = \frac{\alpha_n c(\alpha_n)}{\phi(\alpha_n)}, \qquad \sigma(\alpha_n) = \sqrt{\frac{\alpha_n Q(\alpha_n)}{\phi(\alpha_n)}},$$

where $\{V_n\}$ are independent and identically distributed standard Gaussian random variables.

The filtering distribution at time n is $\pi_n = g_{\alpha_n, z_n}$, which coincides with the Gaussian density having mean and variance respectively given by

$$m_n = \frac{z_n}{\alpha_n}, \qquad s_n^2 = \frac{1}{\alpha_n}. \tag{23}$$

As a special case note that, letting, for some positive constants C and σ,

$$c(\alpha) = \frac{C}{\alpha\sigma^2 + C^2}, \qquad \phi(\alpha) = \frac{\alpha}{\alpha\sigma^2 + C^2},$$

relation (22) becomes

$$X_{n+1} = CX_n + \sigma V_{n+1},$$

which, together with the observation model (17), gives the classical linear Gaussian filtering model (in its time-homogeneous form). The corresponding evolution of the filter is then given by

$$\alpha_n = \frac{\alpha_{n-1}}{\alpha_{n-1}\sigma^2 + C^2} + 1, \qquad z_n = \frac{C}{\alpha_{n-1}\sigma^2 + C^2} \cdot z_{n-1} + y_n,$$

which corresponds to the familiar Kalman filter, where (see Liptser and Shiryaev 1977)

$$m_n = \frac{C \cdot m_{n-1} + (\sigma^2 + C^2 s_{n-1}^2)y_n}{1 + \sigma^2 + C^2 s_{n-1}^2}, \qquad s_n^2 = \frac{\sigma^2 + C^2 s_{n-1}^2}{1 + \sigma^2 + C^2 s_{n-1}^2}.$$

The above results can be generalized (see D'Ambrosio *et al.* 1998) to partially observable models of the form

$$\begin{aligned}
X_{n+1} &= a_n(Y_0^n)X_n + b_n(Y_0^n)V_{n+1}, \\
Y_n &= \eta_n(Y_0^n)X_n + \zeta_n(Y_0^n)W_n,
\end{aligned} \tag{24}$$

which are known from the literature as *conditionally Gaussian models* (see Liptser and Shiryaev 1977).

Example 2. We now turn to an application of Proposition 2.3, considering the case when $I = [0, +\infty)$ and the connecting function $\psi(\alpha, z)$ is nonlinear. We start from an observation model defined by

$$f(y|x) = \frac{1}{\Gamma(c)} x^c y^{c-1} \exp\{-y \cdot x\}, \qquad y > 0, \ c > 0, \tag{25}$$

and, since this implies $a(x) = x^c$, the natural conjugate family $\mathscr{G}$ is the family of gamma distributions

$$g_{a,z}(x) \propto x^{c\alpha} \exp\{zx\}, \qquad \alpha \in A = [0, +\infty), \ z \in Z = (-\infty, 0).$$

As a possible function $\tilde{\psi}(\alpha, z)$ satisfying condition (i) of Proposition 2.3, we take

$$\tilde{\psi}(\alpha, z) = (-2z)^{1/2}.$$

With this choice, letting for simplicity $c = 1$ and putting $\phi(\alpha) = 2\alpha - 1$, we have

$$\tilde{\tau}_{\phi,\psi}(\alpha, z) = \frac{\Gamma(\alpha + 1)}{\Gamma(2\alpha)} 2^\alpha (-z)^{-1} \tag{26}$$

(details of the calculations here and below can be found in D'Ambrosio *et al.* 1998). This shows that condition (ii) of Proposition 2.3 is satisfied as well.

The kernels $K_\alpha(x, x')$ are then given by (16), where the measure P_α does not depend on α and is the infinitely divisible distribution admitting the Laplace transform $\exp\{-(2t)^{1/2}\}$, namely it is the stable distribution with characteristic exponent $\frac{1}{2}$. On the other hand, v_α must be the inverse Laplace transform of $\tilde{\tau}$ in (26), which gives $v_\alpha(\xi) = \Gamma(\alpha + 1)/\Gamma(2\alpha)2^\alpha 1_{\{\xi > 0\}}$ and therefore (see Feller 1970, p. 173)

$$K_\alpha(x, x') = \int_0^\infty v_\alpha(x - u) P_\alpha^{x'*}(\mathrm{d}u)$$

$$= \frac{\Gamma(\alpha + 1)}{\Gamma(2\alpha)} \cdot 2^\alpha \int_0^x \frac{1}{\sqrt{2\pi}} \frac{x'}{u^{3/2}} \cdot \exp\left\{-\frac{(x')^2}{2u}\right\} \mathrm{d}u$$

$$= \frac{\Gamma(\alpha + 1)}{\Gamma(2\alpha)} 2^{\alpha+1} \left[1 - \Phi\left(\sqrt{\frac{x'}{x}}\right)\right], \tag{27}$$

where Φ is the standard Gaussian distribution function. By Theorem 2.1, and recalling that we have put $c = 1$, $\phi(\alpha) = 2\alpha - 1$,

$$p_\alpha(x'|x) = \frac{(x')^{2\alpha-1}}{x^\alpha} K_\alpha(x, x') = \frac{\Gamma(\alpha + 1)}{\Gamma(2\alpha)} \frac{2^{\alpha+1}(x')^{2\alpha-1}}{x^\alpha} \left[1 - \Phi\left(\sqrt{\frac{x'}{x}}\right)\right]. \tag{28}$$

An example, involving the same observation model (25) but with a linear connecting function $\psi(\alpha, z) = c(\alpha) \cdot z$ and leading to a transition density $p_\alpha(x'|x)$ of the beta type can be found in D'Ambrosio *et al.* (1998, Section 4.2.1). This latter result is therefore an example where both cases (a) and (b) hold simultaneously.

We remark, finally, that completely different observation models may lead to the same filter-conjugate family and thus to the same possible families of transition densities $p_\alpha(x'|x)$. In D'Ambrosio *et al.* (1998, Remark 4.1) an example is given of an observation model, different from (25), leading also to transition densities $p_\alpha(x'|x)$ of the beta type.

4. Additional remarks

Our attention focuses on the case when the filter-conjugate family coincides with the family $\mathscr{G}$ in (6), which is the natural conjugate of $\mathscr{F}$ given by the densities in (3). This simplifies the treatment of the main problem. In particular, it implies that the prediction densities, namely those given by

$$q_{\alpha,z}(\cdot) = \int_I p_n(\cdot|x)\, g_{\alpha,z}(x)\, \mathrm{d}x, \tag{29}$$

also belong to the family $\mathscr{G}$. This is, however, by no means the only possible case where one can obtain finite-dimensional filters (see Bather 1965; Spizzichino 1990). On the other hand, our approach leads to a workable methodology, namely one that is based on the use of the inverse Laplace transform.

In our treatment the role of the two parameters α and z in the conjugate family $\mathscr{G}$ is not symmetrical. In fact, z plays the role of the argument in the Laplace transform. Note, however, that we have chosen natural exponential families of distributions for the observations, namely $f(y|x) \propto a(x)\mathrm{e}^{x \cdot y}$ (see (3)); all our results can straightforwardly be extended to generalized exponential families of the form $f(y|x) \propto a(x)\mathrm{e}^{q(x) \cdot y}$. This would allow the roles of α and z in the conjugate family $\mathscr{G}$ to be reversed. In fact, putting $\tilde{a}(x) := \exp\{q(x)\}$, $\tilde{q}(x) := \log\{a(x)\}$, one has equality of the familes $\{g_{\alpha,z}(x) \propto a(x)^{\alpha}\mathrm{e}^{z \cdot q(x)}\}$ and $\{\tilde{g}_{\alpha,z}(x) \propto \tilde{a}(x)^z\mathrm{e}^{\alpha \cdot \tilde{q}(x)}\}$. As a consequence, for each model admitting a finite-dimensional filter one can construct a dual model with the same property. An example can be found in D'Ambrosio *et al.* (1998, Section 4.3).

Acknowledgements

References

Bar-Lev, S., Enis, P. and Letac, G. (1994) Sampling models which admit a given general exponential family as a conjugate family of priors. *Ann. Statist.*, **22**, 1555–1586.

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

Barndorff-Nielsen, O. and Pedersen, K. (1968) Sufficient data reduction and exponential families. *Math. Scand.*, **22**, 197–202.

Bather, J.A. (1965) Invariant conditional distributions. *Ann. Math. Statist.*, **36**, 829–846.

D'Ambrosio, E., Runggaldier, W.J. and Spizzichino, F. (1998) Construction of discrete time models admitting a finite dimensional filter: an approach based on the inverse Laplace transform. Report 32/98, Dipartimento di Matematica 'Guido Castelnuovo', Universitá di Roma 'La Sapienza',

September. Available at *http://www.math.unipd.it/∼probab/home/wolfgang/wolfgangpubl.html.*

Diaconis, P. and Ylvisaker, D. (1979) Conjugate priors for exponential families. *Ann. Statist.*, **7**, 269–281.

Feller, W. (1970) *An Introduction to Probability Theory and Its Applications, Vol. II*. New York: Wiley.

Ferrante, M. (1993) On the existence of finite-dimensional filters in discrete time. *Stochastics Stochastics Rep.*, **40**, 169–179.

Ferrante, M. and Runggaldier, W.J. (1990) On necessary conditions for the existence of finite-dimensional filters in discrete time. *Systems Control Lett.*, **14**, 63–69.

Ferrante, M. and Vidoni, P. (1998) Finite dimensional filters for nonlinear stochastic difference equations with multiplicative noises. *Stochastic Process. Appl.*, **77**, 69–81.

Liptser, R.S. and Shiryaev, A.N. (1977) *Statistics of Random Processes I.* New York: Springer-Verlag.

Sawitzki, G. (1981) Finite-dimensional filter systems in discrete time. *Stochastics*, **5**, 107–114.

Spizzichino, F. (1990) Finite dimensional stochastic filtering in discrete time: the role of convolution semigroups, In A. Bensoussan and P.L. Lions (eds), *Analysis and Optimization of Systems*, Lecture Notes in Control and Inform. Sci. 144, pp. 238–248. Berlin: Springer-Verlag.

Van Schuppen, J.H. (1979) Stochastic filtering theory: a discussion on concepts, methods and results. In M. Kohlmann and W. Vogel (eds), *Stochastic Control Theory and Stochastic Differential Systems*, Lecture Notes in Control and Inform. Sci. 16, pp. 209–226. Berlin: Springer-Verlag.

West, M., Harrison, P.J. and Migon, H.S. (1985) Dynamic generalized linear models and Bayesian forecasting. *J Amer. Statist. Assoc.*, **80**, 73–83.