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

W.J.Runggaldier†, and F.Spizzichino‡

† Dipartimento di Matematica Pura e Applicata
Universitá di Padova, Via Belzoni 7, 35131 - Padova, Italy

‡Dipartimento di Matematica
Universitá di Roma "La Sapienza", 00185 - Roma, Italy

Abstract

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, that is also constructive as shown in the examples.

Key Words: Dynamic Bayes formula, state-space models, finite dimensional filters, exponential families, infinitely divisible distributions, inverse Laplace transform.

Running title : Finite dimensional filters.

1 Introduction

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

*We wish to thank the two anonymous referees for their valuable comments and suggestions on how to improve the original version of the paper

Given two intervals I and J on a finite dimensional Euclidean space, consider a probability kernel $F(x, dy)$ 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 then a nonhomogeneous Markov chain $\{X_n, Y_n\}$ with state space $I \times J$, initial distribution $\Pi_0(dx)\delta_0(dy)$, 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 $P_n(x, dx')F(x', dy')$. The marginal process $\{X_n\}_{n \geq 0}$ forms a nonhomogeneous 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 $Y_0, Y_1, \cdots, Y_{n-1}$ 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 $y_0^n \equiv (y_0, \cdots, y_n)$. This problem is a natural dynamic generalization of the classical Bayesian inference problem, where the latter is obtained as a particular case when $P\{X_0 = X_1 = \cdots = X_n;\, n \geq 1\} = 1$. It can also be viewed as a generalization of the estimation problem in dynamic linear models (to this effect see e.g. [14]).

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

$$
\begin{array}{lcl}
\pi_0(x) & : & \text{density of } x_0 \text{ (initial density)} \\
p_n(x' \,|\, x) & : & \text{transition density} \quad (\text{in period } n) \\
f(y \,|\, x) & : & \text{observation density}
\end{array}
\tag{1}
$$

On the other hand, the filter solution corresponds, for $n = 1, 2, \cdots$, to the conditional densities of the latent state, given the observation history

$$\pi_n(x \,|\, y_0^n) \tag{2}$$

also called *filtering densities*.

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

$$\pi_n(x \,|\, y_0^n) \propto f(y_n \,|\, x) \int_I p_n(x' \,|\, x) \pi_{n-1}(x \,|\, y_0^{n-1}) dx. \tag{3}$$

To actually perform the computations required in (3), it is important to have partially observable models admitting so-called *finite-dimensional filters*, namely 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 to Bayesian statistics and generalizing the situation in the classical linear Gaussian models (Kalman filter model, see e.g. [10]), we give the following

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

$$f(y\,|\,x')\int_I p_n(x'\,|\,x)\gamma_\theta(x)\,dx \propto \gamma_{\Phi(\theta,y)}(x) \tag{4}$$

If a filter-conjugate family exists for a pair $\{f(y|\,x),\, p_n(x'|\,x)\}$, then, if the initial density $\pi_0(x)$ belongs to this filter-conjugate family, by (3) and (4) 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, i.e. 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 (see e.g. [1], [2]) 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 also the conjugate family is exponential. Correspondingly, for discrete-time filtering problems we have (see [8], [11]) that, under regularity assumptions, $\{f(y|\,x),\, p_n(x'|\,x)\}$ 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 $\{f(y|\,x),\, p_n(x'|\,x)\}$ 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 [9]). In this paper we present a method that, in addition to guaranteeing existence, allows also to determine explicitly the kernels. 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 (8) 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. Under such form our problem can also be seen as filter analogon of the Bayesian statistics problem studied in [4].

2 Main results

Motivated by the necessary conditions for the existence of finite-dimensional filters as recalled in the Introduction, we specialize to the case when $\mathcal{F} \equiv \{F(x, dy)\}$ constitutes a natural exponential family (see [5]) of order one, i.e. when the kernel $F(x, dy)$ has a density of the form

$$f(y\,|\,x) = a(x)b(y)e^{xy} \tag{5}$$

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

$$I := \{x \in \mathbb{R} \,|\, \int_J e^{xy} b(y) dy < +\infty\} \tag{6}$$

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

$$\lambda^{-1}(\alpha, z) := \int_I [a(x)]^\alpha \, e^{xz} dx < +\infty \tag{7}$$

It then follows immediately that the family

$$\mathcal{G} := \left\{ g_{a,z}(x) = \lambda(\alpha, z) \, a(x)^\alpha e^{xz} \; ; \quad (\alpha, z) \in \bar{M} \right\} \tag{8}$$

is a natural conjugate family of $\mathcal{F}$, i.e. for x seen as a parameter.

We shall make the following

Assumption 1 : The set $\bar{M}$ contains a rectangle $M = A \times Z$ such that

- $\alpha \in A \quad \Longrightarrow \quad \alpha + 1 \in A$
- $z \in Z \,, \quad y \in J \quad \Longrightarrow \quad z + y \in Z$

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

$$\alpha_n = \phi(\alpha_{n-1}) + 1 \quad , \quad z_n = \psi(\alpha_{n-1}, z_{n-1}) + y_n \tag{9}$$

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

$$\tau_{\phi,\psi}(\alpha, z) := \frac{\lambda\left(\phi\left(\alpha\right), \psi\left(\alpha, z\right)\right)}{\lambda\left(\alpha, z\right)} \tag{10}$$

with λ as in (7).

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

$$\int_I K_\alpha(x, x') \, e^{xz} dx = \tau_{\phi,\psi}(\alpha, z) e^{x' \psi(\alpha, z)} \,, \qquad \forall x' \in I \quad z \in Z \tag{11}$$

Then

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

ii) *if the initial distribution* Π_0 *has a density* $g_{\alpha_0, z_0} \in \mathcal{G}$ *(for some* $(\alpha_0, z_0) \in M$*),* $y_1, y_2, \cdots$ *are the actual observations and* (α_n, z_n) *are as in (9) then, letting*

$$p_n(x' | \, x) = p_{\alpha_{n-1}}(x' | \, x) \tag{12}$$

the family $\mathcal{G}$ *in (8), which is the natural conjugate family of* $\mathcal{F}$*, is a filter conjugate family according to Definition 1.1. Thus,* $\forall x \in I$ *and* $\forall n \geq 1$,

$$\pi_n(x \,|\, y_1^n) = g_{\alpha_n, z_n}(x) \tag{13}$$

Proof :
i) We must show that

$$\int_I p_\alpha(x'|\, x)dx' = 1, \quad \forall \alpha \in A, \ \forall x \in I \tag{14}$$

or, equivalently,

$$\int_I [a(x')]^{\phi(\alpha)}\, K_\alpha(x,x')dx' = [a(x)]^\alpha \quad \forall x \in I\,. \tag{15}$$

By (11) and the definitions of λ in (7) and of τ in (10), the Laplace transform of the left hand side of (15) is given by

$$\int_I \left[\int_I [a(x')]^{\phi(\alpha)}\, K_\alpha(x,x')dx'\right] e^{zx}dx = \lambda^{-1}(\alpha,z) \tag{16}$$

Equation (15) now follows from the uniqueness of the inverse of the Laplace transform.
ii) From the form of $\mathcal{F}$ and $\mathcal{G}$ (see (5) and (8)), and having put (see (12)) $p_n = p_{\alpha_{n-1}}$, it immediately follows that $\mathcal{G}$ is filter-conjugate in the sense of Definition 1.1. Indeed, relation (4) holds with $\Theta = M$, $\theta = (\alpha, z)$, $\Phi(\alpha,z) = (\phi(\alpha)+1,\, \psi(\alpha,z)+y)$, $\gamma_\theta = g_{\alpha,z}$. Relation (13) then follows immediately by induction on $n = 0, 1, \cdots$, taking into account the dynamic Bayes formula (3) (see also the comment after Definition 1.1).

The process $\Theta_n := (\alpha_n, Z_n)$, where α_n and Z_n satisfy (9) with the random sequence $Y_1, Y_2, \cdots$ replacing a specific realization $y_1, y_2, \cdots$ will be called (see [12]) filter process. This process depends on the choice of the functions ϕ and ψ that (see [3]) we shall also call *connecting functions*.

Given the family $\mathcal{F}$ and, therefore, its natural conjugate family $\mathcal{G}$, 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 not for any choice of the connecting functions may there exist a solution of the crucial equation (11). The solvability of (11) hinges upon sufficient conditions on the connecting functions such that the right hand side in (11) 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 (11) 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 (11) can be obtained by means of a convolution integral, after 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 (11). The two special cases allow to study more easily the invertibility of the second factor $\exp\{\psi(\alpha,z)x'\}$.

For case a) we have

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

i) *Denoting* $V_v := \{x \in \mathbb{R} \,|\, v(x) > 0\}$ *, we have* $V_{v_\alpha} + c(\alpha) \cdot I \subseteq I$

ii) $\displaystyle\int_{V_{v_\alpha}} v_\alpha(x) e^{xz} dx = \tau_{\phi,\psi}(\alpha, z)$ *,* $\forall \alpha \in A,\ z \in Z.$

Then, $\forall \alpha \in A$,

$$K_\alpha(x, x') = v_\alpha(x - c(\alpha)x') \tag{17}$$

satisfies equation (11).

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 Lemma 2.2 in [13]).

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 (11) just reads $\tau_{\phi,\psi}(\alpha, z) e^{x' c(\alpha) z}$.

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

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

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

i) $\forall \alpha \in A$, $\tilde{\psi}(\alpha, z)$ *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, i.e.*

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

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

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

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

$$K_\alpha(x, x') = \int_0^\infty v_\alpha(x - u) P_\alpha^{x' \star}(du) \tag{18}$$

satisfies equation (11), where $P^{r\star}$ *denotes the* $r-th$ *convolution power of* P.

Proof : It is a well-known result (see e.g. [7], Ch. XIII) that condition i), together with $\psi(\alpha, 0) = 0$, characterizes $\exp\{-\widetilde{\psi}(\alpha, t)\}$ as the Laplace transform of an infinitely divisible probability distribution P_α over $[0, +\infty)$. Then, under i),

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

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

On the other hand (see again [7], Ch. XIII), ii) is a necessary and sufficient condition for $\widetilde{\tau}(\alpha, z)$ being the ordinary Laplace transform of a non-negative bounded function $v_\alpha : [0, +\infty) \to (0, +\infty)$. Then, $\forall \alpha \in A$, $\forall z \in Z$, the Laplace transform of the function $K_\alpha(x, x')$ defined by (18) is given by the product $\widetilde{\tau}(\alpha, z) \cdot \exp\{\psi(\alpha, z) \cdot x'\}$, i.e. equation (11) 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, condition i) in Proposition 2.2 is trivially satisfied for any function v_α vanishing outside $(0, +\infty)$. Condition ii) in that same Proposition holds (with $Z \subset (-\infty, 0]$) if and only if $\widetilde{\tau}(\alpha, z)$ is completely monotone (see again [7], Ch. XIII). Furthermore, condition i) in Proposition 2.3 is trivially verified in that $\widetilde{\psi}(\alpha, z) = c(\alpha) \cdot z$ and, indeed, P_α is then the Dirac measure with mass concentrated in $c(\alpha)$.

3 Examples

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

$$Y_n = X_n + W_n \tag{19}$$

with $\{W_n\}$ i.i.d. standard Gaussian, so that $\mathcal{F}$ is defined by

$$f(y|x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} \exp\{-\frac{y^2}{2}\} \exp\{-xy\} \tag{20}$$

whence $a(x) = \exp\{-\frac{x^2}{2}\}$ 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 (22) below)

$$|c(\alpha)| < \sqrt{\frac{\phi(\alpha)}{\alpha}}. \tag{21}$$

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

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

having put

$$Q(\alpha) := \frac{\phi(\alpha) - \alpha c^2(\alpha)}{\alpha\phi(\alpha)} \tag{22}$$

which is positive by (21). It is thus easily seen that

$$v_\alpha(x) := \frac{1}{\sqrt{2\pi}}\sqrt{\frac{\phi(\alpha)}{\alpha Q(\alpha)}}\exp\{-\frac{x^2}{2Q(\alpha)}\} \tag{23}$$

satisfies the condition ii) of Proposition 2.2, whereas the condition i) is trivially satisfied in that $I = \mathbb{R}$. Then $K_\alpha(x, x') := v_\alpha(x - c(\alpha)x')$ satisfies equation (11) and from Theorem 2.1 we obtain, as the corresponding transition kernel,

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

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

Taking into account the definition of $Q(\alpha)$ in (22), after some manipulations we obtain that $p_\alpha(x'|\,x)$ coincides with the density function of a Gaussian distribution with mean value $\frac{\alpha c(\alpha)}{\phi(\alpha)}x$ and variance $\frac{\alpha Q(\alpha)}{\phi(\alpha)}$, which corresponds to a state-evolution model of the form

$$X_{n+1} = C(\alpha_n)X_n + \sigma(\alpha_n)V_{n+1} \tag{24}$$

with

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

where $\{V_n\}$ are standard Gaussian i.i.d. 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

$$m_n = \frac{z_n}{\alpha_n}, \quad s_n^2 = \frac{1}{\alpha_n}. \tag{25}$$

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

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

relation (24) becomes

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

which, together with the observation model (19), gives the classical linear Gaussian filtering model (in its time-homogeneous form). The corresponding evolution of the filter is then given by

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

that corresponds to the familiar Kalman filter, where (see e.g. [10])

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

The above results can be generalized (see [6]) to partially observable models of the form

$$\begin{aligned} X_{n+1} &= a_n(Y_0^n) X_n + b_n(Y_0^n) V_{n+1} \\ Y_n &= \eta_n(Y_0^n) X_n + \zeta_n(Y_0^n) W_n \end{aligned} \tag{26}$$

which are known from the literature as *conditionally Gaussian models* (see [10]).

Example 2 : We turn now 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

$$f(y|x) \propto x^c y^{c-1} \exp\{-y \cdot x\}\,; \quad y > 0, \quad c > 0 \tag{27}$$

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

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

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

$$\widetilde{\psi}(\alpha, z) = (-2z)^{\frac{1}{2}}$$

With this choice, letting for simplicity $c = 1$ and putting $\phi(\alpha) = 2\alpha - 1$, we have (details of the calculations here and below can be found in [6]),

$$\widetilde{\tau}_{\phi,\psi}(\alpha, z) = \frac{\Gamma(\alpha+1)}{\Gamma(2\alpha)}\, 2^\alpha (-z)^{-1} \tag{28}$$

This shows that condition ii) of Proposition 2.3 is satisfied as well.

The kernels $K_\alpha(x, x')$ are then given by (18), where the measure P_α does not depend on α and is the infinitely divisible distribution admitting the Laplace transform $\exp\left\{-(2t)^{\frac{1}{2}}\right\}$, namely it is the stable distribution with characteristic exponent 1/2. On the other hand, v_α must be the inverse Laplace transform of $\widetilde{\tau}$ in (28), which gives $v_\alpha(\xi) = \frac{\Gamma(\alpha+1)}{\Gamma(2\alpha)} 2^\alpha 1_{\{\xi>0\}}$ and therefore (see pg.173 in [7])

$$\begin{aligned} K_\alpha(x, x') &= \int_0^\infty v_\alpha(x-u) P_\alpha^{x'\star}(du) \\ &= \tfrac{\Gamma(\alpha+1)}{\Gamma(2\alpha)} \cdot 2^\alpha \int_0^x \frac{1}{\sqrt{2\pi}} \frac{x'}{u^{\frac{3}{2}}} \cdot \exp\left\{-\frac{(x')^2}{2u}\right\} du = \frac{\Gamma(\alpha+1)}{\Gamma(2\alpha)} 2^{\alpha+1} \left[1 - \Phi\left(\sqrt{\frac{x'}{x}}\right)\right] \end{aligned} \tag{29}$$

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

$$p_\alpha(x' \mid x) = \frac{(x')^{2\alpha-1}}{x^\alpha} K_\alpha(x, x') = \frac{\Gamma(\alpha+1)}{\Gamma(2\alpha)} \; \frac{2^{\alpha+1}(x')^{2\alpha-1}}{x^\alpha} \left[1 - \Phi(\sqrt{\frac{x'}{x}}) \right] \tag{30}$$

An example, involving the same observation model (27) 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 section 4.2.1 of [6]. This latter result is therefore an example, where both cases a) and b) hold simultaneously.

We finally remark 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 Remark 4.1 of [6] an example is given of an observation model, different from (27), 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 $\mathcal{G}$ in (8), which is the natural conjugate of $\mathcal{F}$ given by the densities in (5). This simplifies the treatment of the main problem. In particular it implies that also the prediction densities, namely those given by

$$q_{\alpha,z}(\cdot) = \int_I p_n(\cdot | x) \, g_{\alpha,z}(x) \, dx, \tag{31}$$

belong to the family $\mathcal{G}$. This is, however, by no means the only possible case where one can obtain finite-dimensional filters (see [3], [13]). 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 $\mathcal{G}$ 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 (see (5)), $f(y \mid x) \propto a(x)e^{x \cdot y}$; all our results can straightforwardly be extended to generalized exponential families of the form $f(y \mid x) \propto a(x)e^{q(x) \cdot y}$. This would allow to invert the roles of α and z in the conjugate family $\mathcal{G}$. In fact, putting $\tilde{a}(x) := \exp\{q(x)\}$, $\tilde{q}(x) := \log\{a(x)\}$, one has equality of the families $\{g_{\alpha,z}(x) \propto a(x)^\alpha e^{z \cdot q(x)}\}$ and $\{\tilde{g}_{\alpha,z}(x) \propto \tilde{a}(x)^z e^{\alpha \cdot \tilde{q}(x)}\}$. 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 section 4.3 of [6].

References

[1] Barndorff-Nielsen, O. (1978) *Information and Exponential Families*, John Wiley, New York.

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

[3] Bather, J.A. (1965) Invariant conditional distributions, *Ann.Math.Stat.* **36**, 829-846.

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

[5] Diaconis, P. & Ylvisaker, D. (1979) Conjugate Priors for Exponential Families, *The Annals of Statistics* **7**, 269-281.

[6] D'Ambrosio, E. & Runggaldier, W.J. & 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 1998. Available under `http://www.uniroma1.it/preprint/archive/98032.ps`

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

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

[9] Ferrante, M. (1993) On the existence of finite-dimensional filters in discrete time, *Stochastics and Stochastics Reports* **40**, 169-179.

[10] Liptser, R.Sh & Shiryaev, A.N. (1977) *Statistics of Random processes*, Springer-Verlag, New York.

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

[12] Van Schuppen, J.H. (1979) Stochastic filtering theory : a discusison on concepts, methods and results, *In : M.Kohlmann and W.Vogel (eds.), Stochastic Control Theory and Stochastic Differential Systems*, LN in Control and Info. Sci. No. 16, Springer, Berlin, 209-226.

[13] 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*, LN in Control and Info. Sci., Springer, Berlin, 238-248.

[14] West, M. & Harrison, P.J. & Migon, H.S. (1985) Dynamic Generalized Linear Models and Bayesian Forecasting, *Journal of the American Statistical Ass.* **80**, 73-83.