Bernoulli **10**(4), 2004, 701–720

Random scale perturbation of an AR(1) process and its properties as a nonlinear explicit filter

VALENTINE GENON-CATALOT[1] and MATHIEU KESSLER[2]

[1]*Laboratoire MAP 5 (FRE-CNRS 2428), UFR de Mathématiques et Informatique, Université Paris 5 René Descartes, 45 rue des Saints Pères, 75270 Paris Cedex 06, France. E-mail: Valentine.Genon-Catalot@math-infor.univ-paris5.fr*
[2]*Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Paseo Alfonso XIII, 30203 Cartagena, Spain. E-mail: mathieu.kessler@upct.es*

We study the properties of a nonlinear model of filtering in discrete time which leads to explicit computations. The signal is a standard AR(1) process, but noises are multiplicative and non-Gaussian. If the initial distribution of the AR(1) process is taken to belong to a specified class, the prediction and optimal filters also belong to this class and the prediction and updating steps are explicit. We prove the existence of a stationary version for the prediction filter and complete the theoretical study with simulations to illustrate the behaviour of the filter.

Keywords: discrete time observations; filtering; hidden Markov models; multiplicative noise; stability of the filtering algorithm

1. Introduction

State-space models and hidden Markov models form a class of stochastic models which are a class of stochastic models traditionally used in numerous fields of applications. In these models, the process of interest is a Markov chain $(U_n,\ n \geqslant 1)$ which is not observed. Given the whole sequence of state variables (U_n), the observed variables $(Z_n,\ n \geqslant 1)$ are conditionally independent and the conditional distribution of Z_i only depends on the corresponding state variable U_i. A concrete description is often obtained as follows. Suppose that $(\varepsilon_n,\ n \geqslant 1)$ is a sequence of independent and identically distributed (i.i.d.) random variables (the noise), independent of the unobserved Markov chain $(U_n,\ n \geqslant 1)$. Now let the observed process be given, for all $n \geqslant 1$, by

$$Z_n = F(U_n, \varepsilon_n). \tag{1}$$

Then, the sequence (Z_n) satisfies the above properties. In examples where these models are involved, the function F is most often known together with the common distribution of the ε_n. (For general references, see, for example, Elliott *et al.* 1995.) One of the concrete problems raised by these observations is the prediction or filtering of (U_n). This requires the computation for each n of the conditional distribution of U_n given $Z_{n-1}, \ldots, Z_1$ (the prediction filter) and the conditional distribution of U_n given $Z_n, Z_{n-1}, \ldots, Z_1$ (the optimal filter).

Given an initial distribution for U_1, there is a well-known algorithm that allows the successive conditional distributions to be computed:

$$L(U_1) \overset{\text{updating}}{\longrightarrow} L(U_1|Z_1) \overset{\text{prediction}}{\longrightarrow} L(U_2|Z_1)$$
$$\overset{\text{updating}}{\longrightarrow} L(U_2|Z_2, Z_1) \overset{\text{prediction}}{\longrightarrow} L(U_3|Z_2, Z_1) \ldots$$

These distributions are, however, not generally explicit: the problem of exact computation quickly becomes intractable unless they all belong to the same parametric family. In this case it is enough to express the algorithm in terms of the parameters. The most popular example of such a situation is the Kalman filter that evolves in the family of Gaussian distributions. Another example is proposed in Genon-Catalot (2003) where (U_n) is a standard AR(1) process (as in the Kalman filter) but noises are multiplicative ($Z_n = U_n \varepsilon_n$) and non-Gaussian. This results in a nonlinear model with non-compact state space. Our goal in this paper is to study its theoretical properties more deeply and to illuminate the numerical behaviour of the filtering algorithm through simulations.

The study of stability properties of the filters is the subject of a huge number of contributions in which it is essentially always assumed that the state space is finite or compact (see, for example, Del Moral and Guionnet 2001, and the references therein). Very few references address stability properties for nonlinear models when the state space is non-compact. This is mainly due to the lack of explicit formulae for the steps of the algorithm. When the unobserved autoregressive process is stationary we prove, using the explicit expressions derived in Genon-Catalot (2003), stability properties of the filter in our model.

We complete the theoretical study with simulations. These provide, in our opinion, a good insight into the way the steps of the algorithm perform. The numerical implementation is easy and shows that the unobserved variables are well predicted.

In Section 2, we present the model and the formulae we need from Genon-Catalot (2003). The filtering algorithm works within a class $\mathcal{F}$ of distributions called 'serial Gaussian' (SG), which are specified by a scale and a mixture parameter. In Section 3, when (U_n) is stationary, we successively prove the stability of the scale and the mixture parameters of the prediction filter $\nu_{n|n-1:1} = \mathcal{L}(U_n|Z_{n-1}, \ldots, Z_1)$. We deduce the existence of a stationary regime for the Markov process $\left(U_n, Z_n, \nu_{n|n-1:1}\right)_{n\geqslant 1}$ with values in $\mathbb{R} \times \mathbb{R} \times \mathcal{F}$. We investigate the numerical properties of the filter Section 4. These properties are illustrated in the stationary case and in the explosive case for (U_n). The appendices contain some formulae and the proof of a technical lemma.

2. The model

This section recalls the model and filtering algorithm described in Genon-Catalot (2003).

Consider the AR(1) signal (U_n) given by the state equation

$$U_{n+1} = aU_n + \eta_{n+1}, \tag{2}$$

where a is a real number and (η_n) is a sequence of i.i.d. real-valued random variables with distribution $\mathcal{N}(0, \beta^2)$. The initial variable U_1 is assumed to be independent of $(\eta_n, n \geqslant 2)$.

Assume that we observe the random variables (Z_n) related to (U_n) through the observation equation

$$Z_n = U_n \varepsilon_n, \tag{3}$$

where (ε_n) is a sequence of i.i.d. real-valued random variables independent of the sequence (U_n).

We make the following assumptions:

Assumption 2.1. *For all n, ε_n has density*

$$f_1(z) = \frac{\lambda}{|z|^3} \exp\left(-\frac{\lambda}{z^2}\right), \tag{4}$$

where $\lambda > 0$.

Notice that the conditional distribution $L(Z_n|U_n = u) = F_u(\mathrm{d}z)$ admits, for $u \neq 0$, the density

$$f_u(z) = \frac{1}{u} f_1\left(\frac{z}{u}\right), \tag{5}$$

while for $u = 0$, $F_0(\mathrm{d}z) = \delta_0(\mathrm{d}z)$ is the Dirac mass at 0. Moreover, with statistical applications in mind, we deduce from (4) that $\lambda^{1/2}$ is a scale parameter for ε_n, which, in view of (3), implies that only $\beta\lambda^{1/2}$ can be identified from the observation of $U_1, \ldots, U_n$.

Assumption 2.2. *We choose U_1 to belong to the class $\mathcal{F}$ of serial Gaussian distributions.*

We now recall the definition of a serial Gaussian distribution.

Definition 2.1. *The class $\mathcal{F}$ consists of all the distributions $\nu = \nu_{\sigma,\alpha}$, where $\sigma \geqslant 0$ and $\alpha = (\alpha_i, i \geqslant 0)$ is a series of weight coefficients such that, for all $i \geqslant 0, \alpha_i \geqslant 0$ and $\sum_{i \geqslant 0} \alpha_i = 1$. The distributions are defined as follows:*

(a) *If $\sigma = 0$, for any α, we set $\nu_{0,\alpha}(\mathrm{d}u) = \delta_0(\mathrm{d}u)$.*
(b) *When $\sigma > 0$, $\nu(\mathrm{d}u) = \nu_{\sigma,\alpha}(\mathrm{d}u) = g(u)\mathrm{d}u$, with*

$$g(u) = \sum_{i \geqslant 0} \alpha_i \frac{1}{\sigma} g_i\left(\frac{u}{\sigma}\right), \tag{6}$$

where

$$g_i(u) = (2\pi)^{-1/2} \frac{u^{2i}}{C_{2i}} \exp\left(-\frac{u^2}{2}\right), \tag{7}$$

and $C_{2i} = \mathrm{E}(X^{2i})$, for X a standard Gaussian variable.

The serial Gaussian distribution with parameters σ and $\alpha = (\alpha_i, i \geqslant 0)$ will be denoted by $\mathrm{SG}(\sigma, \alpha)$.

Thus, for positive σ, an $\mathrm{SG}(\sigma, \alpha)$ distribution is a mixture distribution which is specified by a scale parameter σ and a mixture parameter α. Each SG distribution is symmetric and $\mathcal{F}$ contains the centred Gaussian laws. By standard series expansion, it can also be seen to contain the symmetric mixture of Gaussian distributions, $\frac{1}{2}\mathcal{N}(-m, \sigma^2) + \frac{1}{2}\mathcal{N}(m, \sigma^2)$.

As stated in the Introduction, we are interested, at stage n, in the prediction filter, that is, the conditional distribution of U_n given $Z_{n-1}, \ldots, Z_1$, and in the optimal filter, that is, the conditional distribution of U_n given $Z_n, \ldots, Z_1$. Let $\nu_1(\mathrm{d}u)$ denote the distribution of U_1 and, for $n \geqslant 1$,

$$\nu_{n|n:1}(\mathrm{d}u) = L(U_n | Z_n, \ldots, Z_1), \tag{8}$$

$$\nu_{n+1|n:1}(\mathrm{d}u) = L(U_{n+1} | Z_n, \ldots, Z_1). \tag{9}$$

Notice that the joint process (U_n, Z_n) is Markov with transition probability

$$p(u, u') f_{u'}(z') \mathrm{d}u' \, \mathrm{d}z', \tag{10}$$

where $p(u, u')$ is the transition density of (2), that is, the Gaussian density with mean au and variance β^2, and $f_{u'}(z')$ is given in (5).

The updating and prediction steps of the filtering algorithm can be described by introducing the following operators. Let $\mathcal{P}(\mathbb{R})$ denote the set of probability measures on $\mathbb{R}$. For $\nu \in \mathcal{P}(\mathbb{R})$, the probability measure $\varphi_z(\nu)$ is defined, for any bounded Borel function h on $\mathbb{R}$, by

$$\varphi_z(\nu) h = \frac{\nu(f_{.}(z) h)}{\nu(f_{.}(z))}, \tag{11}$$

with the convention that $0/0 = 0$. On the other hand, for $\mu \in \mathcal{P}(\mathbb{R})$, the probability measure $\psi(\mu) = \mu P$ is defined by

$$\psi(\mu)(h) = \mu P h = \int_{\mathbb{R} \times \mathbb{R}} p(u, u') h(u') \mu(\mathrm{d}u) \mathrm{d}u'. \tag{12}$$

For the updating step of the filtering algorithm we then have, for $n \geqslant 1$,

$$\nu_{n|n:1} = \varphi_{Z_n}(\nu_{n|n-1:1}),$$

and for the prediction step,

$$\nu_{n+1|n:1} = \psi(\nu_{n|n:1}).$$

Both steps can be expressed in terms of the parameters of the corresponding SG distributions (see Propositions 2.2 and 3.2 in Genon-Catalot 2003). Let us denote by $\Phi_z = \psi \circ \varphi_z$ the resulting composition, which will be extensively used in what follows. We shall use the same notation Φ_z to denote the mapping

$$(\sigma^2, \alpha) \rightarrow \Phi_z(\sigma^2, \alpha) = (\bar{\sigma}^2(z), \bar{\alpha}(z)) \tag{13}$$

which specifies the parameters. However, we introduce a special notation for the scale parameter which is ruled by an autonomous algorithm and set

$$\overline{\sigma}^2(z) = F_z(\sigma^2) = \beta^2 + a^2 \frac{\sigma^2 z^2}{z^2 + 2\lambda\sigma^2}. \tag{14}$$

The evolution of the mixture parameters is given by somewhat more intricate relations that we recall in Appendix A.

Let us stress that if $\nu = \mathrm{SG}(\sigma, \alpha)$ has a finite-length mixture parameter ($\alpha_i = 0$ for i greater than some integer l), then $\varphi_z(\nu)$, $\psi(\nu)$ and $\Phi_z(\nu)$ all have a finite-length mixture parameter, as can be seen from formulae (A.2)–(A.5).

3. Stability properties

In this section we assume that $a^2 < 1$. We denote by σ_{s}^2 the stationary variance of $(U_n)_{n\geqslant 1}$,

$$\sigma_{\mathrm{s}}^2 = \frac{\beta^2}{1 - a^2}. \tag{15}$$

Henceforth we assume that U_1 has distribution

$$\nu_{\mathrm{s}} = \mathcal{N}(0, \sigma_{\mathrm{s}}^2), \tag{16}$$

so that the process (U_n) is strictly stationary and ergodic. The joint process (U_n, Z_n) inherits the stationarity and ergodicity of the hidden chain. This property may easily be checked directly in this model. For the study of stability, we shall consider the strictly stationary and ergodic process (U_n, Z_n), $n \in \mathbb{Z}$, which is the extension indexed by $\mathbb{Z}$ of the stationary process (U_n, Z_n), $n \geqslant 1$, with the same finite-dimensional distributions. Let us denote by $\mathbb{P}$ the distribution of (U_n, Z_n), $n \in \mathbb{Z}$, on the canonical space $\Omega = \mathbb{R}^{\mathbb{Z}} \times \mathbb{R}^{\mathbb{Z}}$ endowed with its usual Borel σ-field $\mathcal{A}$, and denote also by (U_n, Z_n), $n \in \mathbb{Z}$, the canonical process.

3.1. Stability for the scale parameter

In this subsection, to study the scale parameter, we use the approach presented in Bougerol (1993) to obtain the stability by a general theorem on iterations of Lipschitz random functions on a complete separable metric space. We first study some elementary properties of the function (see (14))

$$F_z(v) = \beta^2 + a^2 \frac{v z^2}{z^2 + 2\lambda v} \tag{17}$$

defined on $[0, +\infty)$. Then, we will look for strictly stationary solution processes $(V_n, n \in \mathbb{Z})$, defined on $(\Omega, \mathcal{A}, \mathbb{P})$, of the recursive equation

$$V_n = F_{Z_n}(V_{n-1}), \tag{18}$$

which is based on the strictly stationary and ergodic process $(Z_n, n \in \mathbb{Z})$. Define the closed interval

$$I = [\beta^2, \sigma_{\mathrm{s}}^2]. \tag{19}$$

Lemma 3.1. *For* $z \neq 0$, *the function* F_z *is increasing from* I *onto* I *and is Lipschitz with* $0 \leqslant \mathrm{d}F_z/\mathrm{d}v \leqslant a^2$. *For* $z = 0$, $F_0(v) = \beta^2$ *for all* v.

Proof. We have

$$\frac{\mathrm{d}F_z}{\mathrm{d}v}(v) = \frac{a^2 z^4}{(z^2 + 2\lambda v)^2}.$$

For all $v \in I$,

$$\beta^2 \leqslant F_z(v) \leqslant F_z\left(\frac{\beta^2}{1-a^2}\right) \leqslant \frac{\beta^2}{1-a^2}.$$

The proof is thus complete. □

Remark. From this lemma, we see that, for all $v \in I$, and in particular for $v = \beta^2/(1-a^2)$,

$$\frac{\beta^2}{F_z(v)} \in [1-a^2, 1] \quad \text{and} \quad 1 - \frac{\beta^2}{F_z(v)} \in [0, a^2]. \tag{20}$$

The above values are precisely the mixture parameter after one iteration of the prediction filter starting from the stationary distribution for U_1. We can see that they both belong to very small length intervals when a is close to 0, which is a situation close to independence.

For $z = (z_n, n \in \mathbb{Z}) \in \mathbb{R}^{\mathbb{Z}}$ and $n \in \mathbb{Z}$, let us denote by $\underline{z}_n = (z_n, z_{n-1}, \ldots)$ the element of $\mathbb{R}^{\mathbb{N}}$ which is the infinite past of z starting from n.

Proposition 3.1. *There exists, on* $(\Omega, \mathcal{A}, \mathbb{P})$, *a measurable function* $V(\underline{Z}_n)$, *such that the process* $(\overline{V}_n = V(\underline{Z}_n), n \in \mathbb{Z})$ *is the unique strictly stationary solution process of* (18). *This process takes values in* I, *is ergodic and has the following properties:*

(i) *For all* $v \in I$, $n \in \mathbb{Z}$,

$$V_{n:n-p}(v) = F_{Z_n} \circ F_{Z_{n-1}} \circ \ldots \circ F_{Z_{n-p}}(v) \to V(\underline{Z}_n),$$

almost surely under $\mathbb{P}$ *as* p *tends to infinity.*

(ii) *Moreover,* $V_{n:1}(v) = F_{Z_n} \circ F_{Z_{n-1}} \circ \ldots \circ F_{Z_1}(v)$ *satisfies*

$$V_{n:1}(v) - V(\underline{Z}_n) \to 0,$$

almost surely under $\mathbb{P}$ *as* n *tends to* $+\infty$.

(iii) *Let* $\mu(\mathrm{d}u, \mathrm{d}z, \mathrm{d}v')$ *denote the distribution of* $(U_1, Z_1, V(\underline{Z}_0))$. *Then* $(U_n, Z_n, V_{n-1:1}(v))$ *converges in distribution to* $\mu(\mathrm{d}u, \mathrm{d}z, \mathrm{d}v')$ *as* n *tends to* $+\infty$.

Proof. We apply Bougerol's (1993) Theorem 3.1 and Corollaries 3.2–3.3. By our Lemma 3.1, (F_{Z_n}) is an ergodic sequence of Lipschitz maps on the separable complete metric space I, and the Lipschitz coefficient of (F_{Z_n}) is $\rho(F_{Z_n}) = a^2$. To obtain our result, we only need to check that, for some $v_0 \in I$, $\mathrm{E}\log^+|F_{Z_1}(v_0) - v_0|$ is finite. This follows from the fact that the interval I is bounded from below by $\beta^2 > 0$. Moreover,

$$|V_{n:1}(v) - V(\underline{Z}_n)| \leqslant a^{2n}|v - V(\underline{Z}_0)| \leqslant a^{2n}\frac{2\beta^2}{1-a^2} \tag{21}$$

since $V(\underline{Z}_n) = F_{Z_n} \circ F_{Z_{n-1}} \circ \ldots \circ F_{Z_1}(V(\underline{Z}_0))$. The stated results follow. □

Remark. By (iii), we have obtained the stability for the Markov process $(U_n, Z_n, V_{n-1:1}(v))$ with respect to the initial condition v of its third component.

3.2. Stability of the mixture parameter

Because of the intricate expression for the mapping Φ_z that describes the action of the algorithm on the mixture parameter (see Appendix A), we cannot use the same reasoning as for the scale parameter. This implies that the study of stability for the mixture parameter is significantly harder. Actually we only prove a weak form of stability. More precisely, we consider the recursive equation

$$\nu_n = \Phi_{Z_n}(\nu_{n-1}),$$

defined on $\mathcal{F}$, and prove the existence of a strictly stationary and ergodic process solution $(\tilde{\nu}_n, n \in \mathbb{Z})$. For each n, $\tilde{\nu}_n$ has scale parameter $V(\underline{Z}_n)$ obtained in Section 3.1 and mixture parameter $\alpha(\underline{Z}_n)$ which we obtain in this subsection.

We first prove two properties of Φ_z as a function on $\mathcal{F}$, which we will need later.

Lemma 3.2. *(i) For all $\nu \in \mathcal{F}$ and all z, $\Phi_z(\nu) \in \mathcal{F}$ and has a positive scale parameter, hence this distribution has density.*

(ii) The mapping Φ_z is continuous with respect to the topology of weak convergence.

Proof. (i) For all $\nu \in \mathcal{F}$ and all z, the scale parameter of $\Phi_z(\nu)$ is greater than or equal to β^2.

(ii) Since, for all z, the function $u \to f_u(z)$ is continuous and bounded, if ν_n weakly converges to $\nu \neq \delta_0$, then $\varphi_z(\nu_n)$ (see (11)) weakly converges to $\varphi_z(\nu)$. We have $\varphi_z(\delta_0) = \delta_0$, and from direct computations it is easily seen that if ν_n weakly converges to δ_0, then $\varphi_z(\nu_n)$ weakly converges to δ_0. Hence, φ_z is continuous. Moreover, the transition operator P of (2) is Feller. Therefore, if h is continuous and bounded, so is Ph. This implies the continuity of ψ with respect to the topology of weak convergence of probability measures. The result follows for the composition $\Phi_z = \psi \circ \varphi_z$. □

Remark. Let us denote by $\mathcal{F}_I$ the subset of SG distributions with scale parameter belonging to I. The mapping Φ_z maps $\mathcal{F}_I$ onto itself.

Due to the strict stationarity of $((U_n, Z_n), n \in \mathbb{Z})$, the conditional distribution of U_1 given $Z_0, Z_{-1}, \ldots, Z_{-n+2}$ is

$$\nu_{1|0:-n+2} = \Phi_{Z_0} \circ \Phi_{Z_{-1}} \circ \ldots \circ \Phi_{Z_{-n+2}}(\nu_{\mathrm{s}}). \tag{22}$$

We consider the asymptotic behaviour of these distributions as random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$ with values in the subset $\mathcal{F} \subset \mathcal{P}(\mathbb{R})$ of SG distributions endowed with the (induced) Borel σ-field associated with the topology of weak convergence. We will show that, almost surely, this sequence converges weakly to an SG distribution which we will denote by $\nu_{1|0:-\infty}$. We first characterize this distribution and then prove the weak convergence property.

Proposition 3.2. *There exists a regular version of the conditional distribution of* U_1*, given the infinite past* $\underline{Z}_0$*, of the form*

$$\tilde{\nu}_{1|0:-\infty}(\mathrm{d}u) = \tilde{g}(u; \underline{Z}_0)\,\mathrm{d}u$$

where (see (10)*):*

(i) *for all* $u \in \mathbb{R}$, $\mathrm{E}(p(U_0, u)|\underline{Z}_0) = \tilde{g}(u; \underline{Z}_0)$, $\mathbb{P}$-*a.s.*.;
(ii) $(u, \underline{Z}_0(\omega)) \to \tilde{g}(u; \underline{Z}_0(\omega))$ *is measurable;*
(iii) $\mathbb{P}$-*a.s.,* $\tilde{g}(\cdot; \underline{Z}_0)$ *is continuous on* $\mathbb{R}$.

Proof. Let $\hat{\nu}(\mathrm{d}u_0; \underline{Z}_0(\omega))$ be a regular version of the conditional distribution under of U_0 given $\underline{Z}_0$. Set

$$\tilde{g}(u; \underline{Z}_0(\omega)) = \int_{\mathbb{R}} p(u_0, u)\hat{\nu}(\mathrm{d}u_0; \underline{Z}_0(\omega)) \tag{23}$$

so that (i) holds. The above function is a probability density. It is easily seen that (ii)–(iii) also hold.

Now we must prove that, for all $\varphi : \mathbb{R} \to [0, 1]$ Borel,

$$\mathrm{E}(\varphi(U_1)|\underline{Z}_0) = \int \varphi(u)\tilde{g}(u; \underline{Z}_0)\mathrm{d}u. \tag{24}$$

By the Markov property of (U_n, Z_n) and the special form of its transition probability (see (10)), we have

$$\mathrm{E}(\varphi(U_1)|\underline{U}_0, \underline{Z}_0) = \mathrm{E}(\varphi(U_1)|U_0, Z_0) = \mathrm{E}(\varphi(U_1)|U_0). \tag{25}$$

Hence,

$$\mathrm{E}(\varphi(U_1)|\underline{Z}_0) = \int \hat{\nu}(\mathrm{d}u_0; \underline{Z}_0(\omega)) \int \varphi(u)p(u_0, u)\mathrm{d}u, \tag{26}$$

which gives (24). □

We now turn to the weak convergence of (22).

Proposition 3.3. *The sequence of probability measures* $(\nu_{1|0:-n+2}(\mathrm{d}u))$ *weakly converges, as* n *tends to* $+\infty$*,* $\mathbb{P}$*-almost surely, to* $\nu_{1|0:-\infty}$ *given in Proposition 3.2.*

Proof. For $x \in \mathbb{R}$, consider the random distribution functions

$$F_n(x) = \int_{-\infty}^{x} \nu_{1|0:-n+2}(\mathrm{d}u) \tag{27}$$

and

$$F(x) = \int_{-\infty}^{x} \nu_{1|0:-\infty}(\mathrm{d}u). \tag{28}$$

From (22), Lemma 3.2, and Proposition 3.2, these distribution functions are continuous in x almost surely, since their corresponding distributions have densities. For all $x \in \mathbb{R}$, $\mathbb{P}$-almost surely, we have

$$F_n(x) = \mathrm{E}\big(1_{(-\infty,x]}(U_1)|Z_0, \ldots, Z_{-n+2}\big) \tag{29}$$

and

$$F(x) = \mathrm{E}\big(1_{(-\infty,x]}(U_1)|\underline{Z}_0\big). \tag{30}$$

By the martingale convergence theorem, we obtain

$$\forall x \in \mathbb{R},\ \mathbb{P}\text{-a.s.,} \quad F_n(x) \to F(x),$$

as n tends to infinity. Therefore, there exists a null set N in $\mathcal{A}$ ($\mathbb{P}(N) = 0$) such that

$$\forall \omega \in N^c, \forall r \in Q,\ F_n(r, \omega) \to F(r, \omega).$$

Now, fix $\omega \in N^c$ and $x \in \mathbb{R}$. For all $\varepsilon > 0$, there exist $r', r'' \in Q$ such that

$$r' \leqslant x \leqslant r'' \quad \text{and} \quad F(x, \omega) - \varepsilon \leqslant F(r', \omega) \leqslant F(r'', \omega) \leqslant F(x, \omega) + \varepsilon$$

because of the continuity of $F(\cdot, \omega)$. The inequality

$$F_n(r', \omega) \leqslant F_n(x, \omega) \leqslant F_n(r'', \omega)$$

implies

$$F(r', \omega) \leqslant \liminf_n F_n(x, \omega) \leqslant \limsup_n F_n(x, \omega) \leqslant F(r'', \omega).$$

Hence $F_n(x, \omega) \to F(x, \omega)$, and we have shown that, for all $\omega \in N^c$, the weak convergence of $\nu_{1|0:-n+2}(\mathrm{d}u, \omega)$ to $\nu_{1|0:-\infty}(\mathrm{d}u, \omega)$ holds, which completes the proof. □

We are now in a position to prove:

Proposition 3.4. *There exists, $\mathbb{P}$-almost surely, a unique random mixture parameter $\alpha(\underline{Z}_0)$ which is a measurable function of $\underline{Z}_0$ such that $\nu_{1|0:-\infty}(\mathrm{d}u)$ is the SG distribution with parameters $V(\underline{Z}_0)$ and $\alpha(\underline{Z}_0)$.*

Proof. By Proposition 3.3, the distribution $\nu_{1|0:-\infty}$ is the limit of the sequence of SG distributions $\nu_{1|0:-n+2}$ which have $V_{0|0:-n+2}(\sigma_{\mathrm{s}}^2)$ as scale parameters. By Proposition 3.1, the sequence $V_{0|0:-n+2}(\sigma_{\mathrm{s}}^2)$ converges almost surely to $V(\underline{Z}_0) > 0$. We are finished if we prove the following technical lemma. □

Lemma 3.3. *Let $(U_n)_{n\geqslant 1}$ be a sequence of random variables with, for all n, $L(U_n) = \mathrm{SG}(\sigma_n, \alpha^{(n)})$, such that*

(i) $\sigma_n \underset{n\to\infty}{\longrightarrow} \sigma > 0$,

(ii) $U_n \underset{n\to\infty}{\longrightarrow} U$ *in distribution.*

Then, for all $i \geqslant 0$, $\alpha_i^{(n)} \to \alpha_i$, *where* $\alpha_i \geqslant 0$ *and* $\sum_{i\geqslant 0}\alpha_i = 1$; *moreover,* $L(U) = \mathrm{SG}(\sigma, \alpha)$.

The proof of Lemma 3.3 is postponed to Appendix B.

Let us now set

$$\tilde{\nu}_n = \nu_{n|n-1:-\infty}. \tag{31}$$

We have the following result.

Proposition 3.5. *For all* $n \in \mathbb{Z}$,

$$\tilde{\nu}_{n+1} = \Phi_{Z_n}(\tilde{\nu}_n), \tag{32}$$

$\mathbb{P}$*-almost surely.*

Proof. First note that, by the strict stationarity of (U_n, Z_n), the conditional distribution of U_2 given $\underline{Z}_1$ is $\nu_{2|1:-\infty}$. Moreover, by the same proof as for Proposition 3.3, the weak convergence of $\nu_{2|1:-n+2}$ to $\nu_{2|1:-\infty}$ holds $\mathbb{P}$-almost surely. We also have

$$\nu_{2|1:-n+2} = \Phi_{Z_1}(\nu_{1|0:-n+2}).$$

Hence, by Lemma 3.2 and Proposition 3.3, the weak convergence of $\nu_{2|1:-n+2}$ to $\Phi_{Z_1}(\nu_{1|0:-\infty})$ holds $\mathbb{P}$-almost surely. Thus, we obtain

$$\nu_{2|1:-\infty} = \Phi_{Z_1}(\nu_{1|0:-\infty}).$$

The proposition follows □

Remark. For each n, $\tilde{\nu}_n$ is the conditional distribution of U_n given $\underline{Z}_{n-1}$. On $(\Omega, \mathcal{A}, \mathbb{P})$, the process $(U_n, Z_n, \tilde{\nu}_n(\mathrm{d}u))_{n\in\mathbb{Z}}$ with state space $\mathbb{R}\times\mathbb{R}\times\mathcal{F}$ is strictly stationary and ergodic. So, by Proposition 3.5, we have obtained a stationary regime for the Markov process $(U_n, Z_n, \Phi_{Z_n} \circ \Phi_{Z_{n-1}} \circ \ldots \circ \Phi_{Z_1}(\nu_\mathrm{s}), n \geqslant 1)$.

4. Behaviour of the filter on numerical simulations

4.1. Behaviour of the algorithm

Let us first notice that in the nonlinear filtering literature, one is interested in finding finite-dimensional filters, that is, filters that evolve in a family of distributions that can be parametrized by a finite number of parameters as for the Kalman filter. However, finite-dimensional filters are difficult to obtain (see Sawitzki 1981), and generally they induce a time-inhomogeneous signal (see Runggaldier and Spizzichino 2001), which excludes the stationary processes. Our filter is not finite-dimensional because of the mixture parameters;

Table 1. Behaviour of the SG parameters of the update filter $\nu_{i|i:1} = L(U_i \mid Z_i, \ldots, Z_1)$ and prediction filter $\nu_{i+1|i:1} = L(U_{i+1} \mid Z_i, \ldots, Z_1)$ for $a = 0.3$, $\beta = 1$ and $\lambda = 1$

i	Step	z_i	σ	α_0	α_1	α_2	α_3	α_4	α_5	α_6	α_7
1	Update	−2.84	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	Prediction		1.04	0.61	0.32	0.06	0.01	0.00	0.00	0.00	0.00
2	Update	−1.68	0.78	0.00	0.48	0.42	0.09	0.01	0.00	0.00	0.00
	Prediction		1.03	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00
3	Update	1.62	0.76	0.00	0.87	0.13	0.00	0.00	0.00	0.00	0.00
	Prediction		1.03	0.94	0.06	0.00	0.00	0.00	0.00	0.00	0.00
4	Update	−2.65	0.90	0.00	0.88	0.12	0.00	0.00	0.00	0.00	0.00
	Prediction		1.04	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00
5	Update	0.71	0.45	0.00	0.96	0.04	0.00	0.00	0.00	0.00	0.00
	Prediction		1.01	0.98	0.02	0.00	0.00	0.00	0.00	0.00	0.00
6	Update	−0.17	0.12	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	Prediction		1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	Update	−0.27	0.19	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	Prediction		1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	Update	−0.45	0.30	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	Prediction		1.00	0.99	0.01	0.00	0.00	0.00	0.00	0.00	0.00
9	Update	−1.72	0.77	0.00	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	Prediction		1.03	0.95	0.05	0.00	0.00	0.00	0.00	0.00	0.00
10	Update	−0.28	0.20	0.00	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	Prediction		1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	Update	0.05	0.03	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	Prediction		1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	Update	0.50	0.33	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	Prediction		1.00	0.99	0.01	0.00	0.00	0.00	0.00	0.00	0.00
13	Update	−4.94	0.97	0.00	0.97	0.03	0.00	0.00	0.00	0.00	0.00
	Prediction		1.04	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00
14	Update	−1.67	0.78	0.00	0.87	0.13	0.00	0.00	0.00	0.00	0.00

Table 2. Behaviour of the SG parameters of the update filter $\nu_{i|i:1} = L(U_i \mid Z_i, \ldots, Z_1)$ and prediction filter $\nu_{i+1|i:1} = L(U_{i+1} \mid Z_i, \ldots, Z_1)$ for $a = 0.8$, $\beta = 1$ and $\lambda = 1$

i	Step	z_i	σ	α_0	α_1	α_2	α_3	α_4	α_5
1	Update	3.12	1.24	0.00	0.00	0.00	0.00	1.00	
	Prediction		1.41	0.06	0.25	0.37	0.25	0.06	
2	Update	−1.57	0.87	0.00	0.09	0.39	0.37	0.13	0.02
	Prediction		1.22	0.38	0.42	0.17	0.03	0.00	0.00
3	Update	−1.12	0.66	0.00	0.45	0.45	0.09	0.01	0.00
	Prediction		1.13	0.67	0.29	0.03	0.00	0.00	0.00
4	Update	−2.58	0.96	0.00	0.48	0.45	0.06	0.00	0.00
	Prediction		1.26	0.50	0.42	0.08	0.00	0.00	0.00
5	Update	−1.14	0.68	0.00	0.55	0.41	0.04	0.00	0.00
	Prediction		1.14	0.69	0.29	0.03	0.00	0.00	0.00
6	Update	−0.22	0.15	0.00	0.98	0.02	0.00	0.00	0.00
	Prediction		1.01	0.99	0.01	0.00	0.00	0.00	0.00
7	Update	−0.08	0.06	0.00	1.00	0.00	0.00	0.00	0.00
	Prediction		1.00	1.00	0.00	0.00	0.00	0.00	0.00
8	Update	18.92	1.00	0.00	0.99	0.01	0.00	0.00	0.00
	Prediction		1.28	0.61	0.39	0.00	0.00	0.00	0.00
9	Update	−4.32	1.18	0.00	0.38	0.62	0.00	0.00	0.00
	Prediction		1.38	0.37	0.49	0.14	0.00	0.00	0.00

however, the simulation study we now present, shows that the number of significant components of the mixture is of the order of 2 or 3, which is remarkably small.

We begin by illustrating the evolution of the parameters of the prediction filter $\nu_{n+1|n:1}$ and of the optimal filter $\nu_{n|n:1}$ as SG distributions for several non-negative values of a. Recall that if we start the algorithm with a finite-length mixture parameter then all steps yield a finite-length mixture parameter. To avoid unnecessary computations we have set the mixture components to zero if they turn out smaller than 10^{-9}.

The numerical results presented in Tables 1–3 illustrate the relations between $\overline{\sigma}$, $\hat{\sigma}$, $\overline{\alpha}$, $\hat{\alpha}$ and z_i described in equations (A.2)–(A.5). In particular, they confirm that when z_i is close to zero, $\hat{\sigma}$ is close to zero as well, and $\hat{\alpha}$ is essentially equal to $(0, 1, \ldots,)$; see relations (A.3)–(A.5). As a consequence, we deduce from relations (14) and (A.2) that, if z_i is close to zero, $\overline{\sigma}$ is close to β and $\overline{\alpha}$ is approximately $(1, 0, 0, \ldots)$, which means that $\nu_{i+1|i:1}$ is almost a centred Gaussian distribution.

Thus, if $a < 1$, which corresponds to the stationary case for (U_n), even if we begin with a long mixture parameter, after a few iterations, only the first components of the mixture parameter of the prediction filter $\nu_{i+1|i:1}$ are significant. In particular, the distributions $\nu_{i+1|i:1}$ are unimodal and are similar to Gaussian distributions.

On the other hand, if $a > 1$, which corresponds to the explosive case, the scale parameter as well as the length of the mixture parameter seem to explode. Notice, however, that in Table 3 the number of significant components of the mixture parameter is small.

Table 3. Behaviour of the SG parameters of the update filter $\nu_{i|i:1} = L(U_i \mid Z_i, \ldots, Z_1)$ and prediction filter $\nu_{i+1|i:1} = L(U_{i+1} \mid Z_i, \ldots, Z_1)$ for $a = 1.5$, $\beta = 1$ and $\lambda = 1$

i	Step	z_i	σ	α_0	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9
1	Update	−1.84	1.19	0.00	1.00								
	Prediction		2.05	0.24	0.76								
2	Update	−1.69	1.03	0.00	0.29	0.71							
	Prediction		1.84	0.15	0.50	0.35							
3	Update	−4.70	1.61	0.00	0.06	0.49	0.44						
	Prediction		2.62	0.02	0.20	0.50	0.28						
4	Update	6.12	2.24	0.00	0.01	0.17	0.52	0.30					
	Prediction		3.50	0.00	0.04	0.26	0.48	0.21					
5	Update	−22.85	3.42	0.00	0.00	0.02	0.21	0.50	0.27				
	Prediction		5.23	0.00	0.00	0.04	0.26	0.48	0.22				
6	Update	33.40	5.11	0.00	0.00	0.00	0.03	0.21	0.49	0.27			
	Prediction		7.73	0.00	0.00	0.00	0.04	0.24	0.48	0.24			
7	Update	−20.12	6.79	0.00	0.00	0.00	0.00	0.04	0.25	0.48	0.22		
	Prediction		10.23	0.00	0.00	0.00	0.00	0.06	0.27	0.46	0.21		
8	Update	−19.50	8.22	0.00	0.00	0.00	0.00	0.01	0.09	0.33	0.43	0.14	
	Prediction		12.37	0.00	0.00	0.00	0.00	0.01	0.10	0.33	0.42	0.14	
9	Update	−101.75	12.19	0.00	0.00	0.00	0.00	0.00	0.01	0.08	0.31	0.44	0.16
	Prediction		18.31	0.00	0.00	0.00	0.00	0.00	0.01	0.08	0.32	0.44	0.15

4.2. Empirical distribution of the parameters of $\nu_{n+1|n:1}$

In this subsection, as a support to the theoretical results on the stability of the filtering algorithm proved in Section 3, we have performed, in the stationary case ($a = 0.3$), 200 simulations of the algorithm and present the histograms of the relevant parameters of the SG distribution $\nu_{n+1|n:1} = L(U_{n+1} \mid Z_n, \ldots, Z_1)$ for $n = 1000$. The initial random variable U_1 was chosen to have the stationary law ν_s defined in (16).

Figures 1 and 2 contain the histograms of the first two components $\overline{\alpha}_0$ and $\overline{\alpha}_1$ of the mixture parameter $\overline{\alpha}$ of $\nu_{n+1|n:1}$, while Figure 3 describes the contribution of the remaining components $\sum_{i\geqslant 2}\overline{\alpha}_i$. These representations confirm the fact that in the stationary case the prediction filter $\nu_{n+1|n:1}$ is essentially a finite-component mixture with two or three components, it is generally close to a centred Gaussian law.

Figure 4 presents the histogram of the 200 realizations of the scale parameter $\overline{\sigma}$. Notice that, since $a = 0.3$, $\beta = 1$ and $\lambda = 1$, the interval I, (see (19)), which turns out to be relevant in the study of the stability of the scale parameter, is equal to [1,1.0989]. One may verify that the support of the histogram is contained in I.

4.3. Prediction properties

To conclude the simulation study, we are interested in the prediction properties of the algorithm. We simulate one trajectory of $(U_n, Z_n)_n$ and plot the density of $L(U_{n+1}^2 \mid Z_n, \ldots, Z_1)$ for $n = 49$ together with the density of $\mathcal{L}(U_{n+1}^2 \mid U_n)$, which would

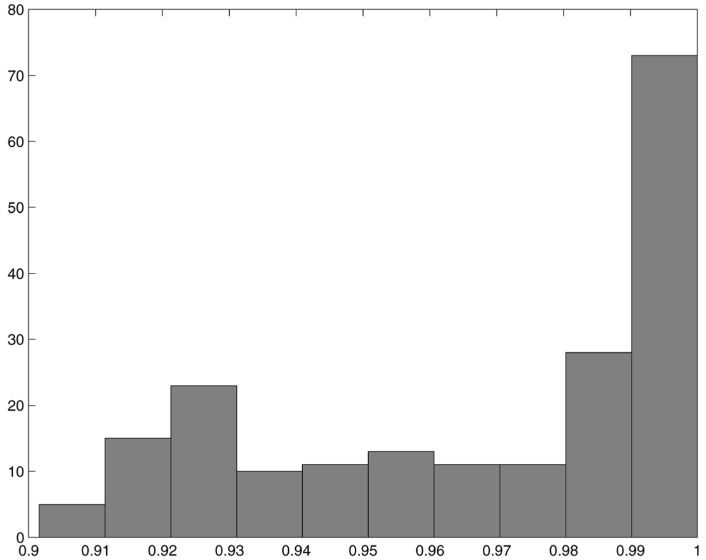

Figure 1. Two hundred realizations of $\bar{\alpha}_0$ for $\nu_{n+1|n:1} = L(U_{n+1} \mid Z_n, \ldots, Z_1)$ for $n = 1000$, $a = 0.3$, $\beta = 1$ and $\lambda = 1$

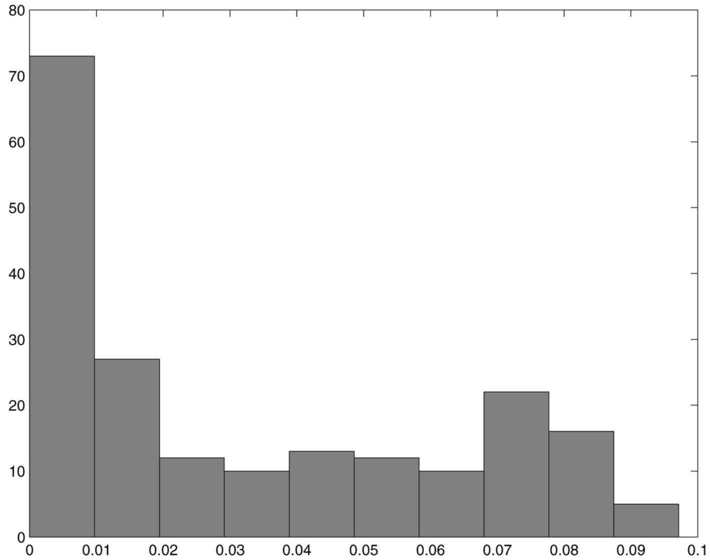

Figure 2. Two hundred realizations of $\bar{\alpha}_0$ for $\nu_{n+1|n:1} = L(U_{n+1} \mid Z_n, \ldots, Z_1)$ for $n = 1000$, $a = 0.3$, $\beta = 1$ and $\lambda = 1$

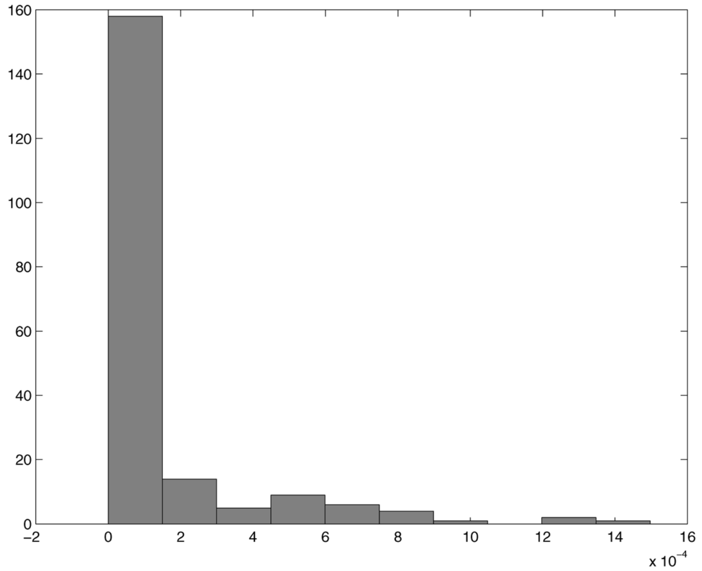

Figure 3. Two hundred realizations of $\sum_{i \geqslant 2} \bar{\alpha}_0$ for $\nu_{n+1|n:1} = L(U_{n+1} \mid Z_n, \ldots, Z_1)$ for $n = 1000$, $a = 0.3$, $\beta = 1$ and $\lambda = 1$

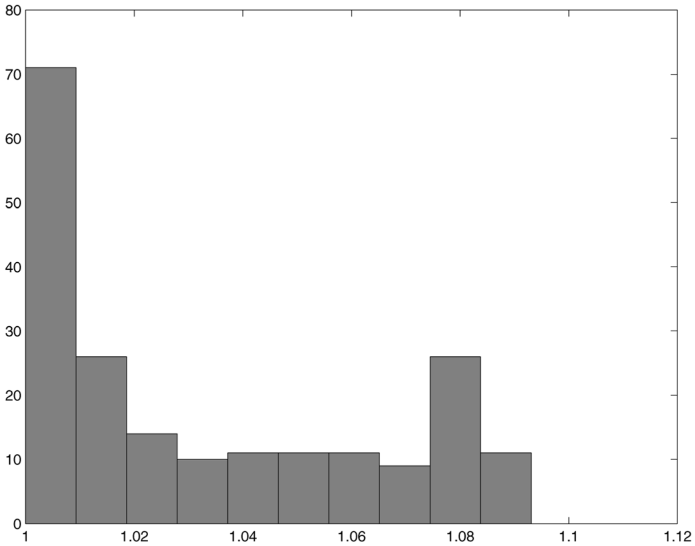

Figure 4. Two hundred realizations of $\overline{\sigma}^2$ for $\nu_{n+1|n:1} = L(U_{n+1} \mid Z_n, \ldots, Z_1)$ for $n = 1000$, $a = 0.3$, $\beta = 1$ and $\lambda = 1$

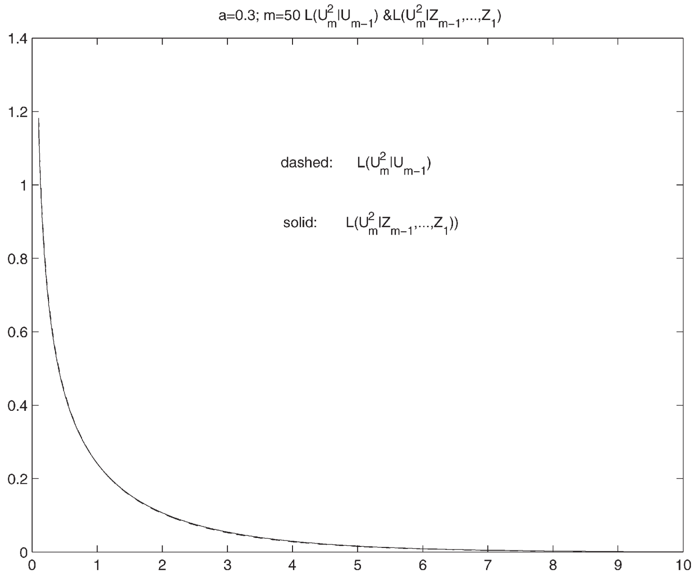

Figure 5. The densities of one realization of $L(U_{50}^2 \mid Z_{49}, \ldots, Z_1)$ and of $L(U_{50}^2 \mid U_{49}^2)$ for $a = 0.3$, $\beta = 1$ and $\lambda = 1$

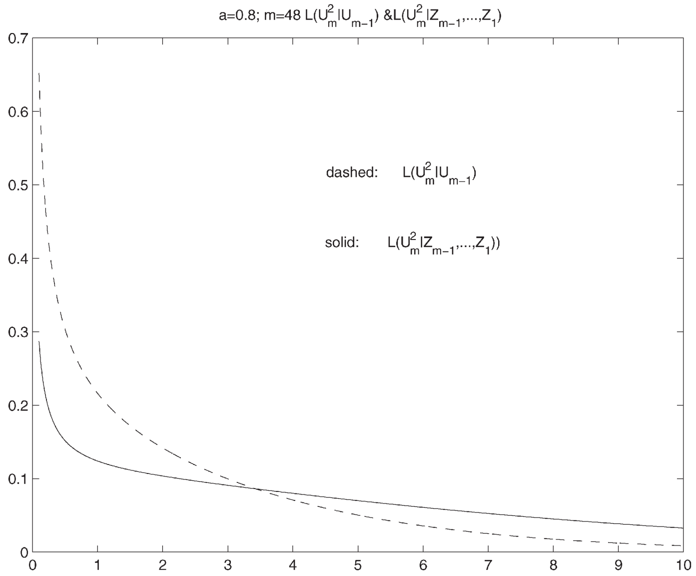

Figure 6. The densities of one realization of $L(U_{50}^2 \mid Z_{49}, \ldots, Z_1)$ and of $L(U_{50}^2 \mid U_{49}^2)$ for $a = 0.8$, $\beta = 1$ and $\lambda = 1$

correspond to the optimal predictor for U_{n+1}^2 in the absence of noise. We have chosen to consider the prediction of the square of U_{n+1}, because it is of course impossible from the observation of Z to make any inference about the sign of U.

Figures 5 and 6 show these comparative plots for $a = 0.3$ and $a = 0.8$, respectively. In both cases the prediction properties are similar, and for $a = 0.3$ the two prediction densities are even superimposed. In the classical linear Kalman filter model (where $Z_n = U_n + \varepsilon'_n$, with (ε'_n) i.i.d. centred Gaussian and U_n as here), the prediction and optimal filters are Gaussian and it is traditional to consider as best predictors of the value U_n the expectations of the filters (the conditional expectations given the set of observations). These are the L^2 predictors. In our model, the variables we want to predict are U_n^2 and the conditional expectations turn out to be inadequate because of the skewness of the filters. On the other hand, if we consider L^1 predictors, then we obtain the conditional medians. The simulation results show that these medians are good predictors that can be used to infer the unobserved values U_n^2, U_{n+1}^2 from $Z_n, \ldots, Z_1$.

Appendix A. Formulae for the filtering algorithm

Recall that Φ_z denotes the mapping

$$(\sigma^2, \alpha) \to \Phi_z(\sigma^2, \alpha) = (\overline{\sigma}^2(z), \overline{\alpha}(z)), \tag{A.1}$$

where $\overline{\sigma}^2(z)$ was given in (14). The mixture coefficients $\overline{\alpha}(z)$ depend on both σ^2 and α as follows.

For $k \geqslant 0$,

$$\overline{\alpha}_k(z) = \left(1 - \frac{\beta^2}{\overline{\sigma}^2(z)}\right)^k \sum_{i \geqslant k} \binom{i}{k} \left(\frac{\beta}{\overline{\sigma}(z)}\right)^{2(i-k)} \hat{\alpha}_i(z), \tag{A.2}$$

where if $\sigma(z) \neq 0$,

$$\hat{\alpha}_0(z) = 0, \qquad \hat{\alpha}_i(z) = \alpha_{i-1} \frac{h_{i-1}(z/\lambda^{1/2}\sigma)}{h(z/\lambda^{1/2}\sigma)}, \quad i \geqslant 1, \tag{A.3}$$

in which

$$h(z) = \sum_{i \geqslant 0} \alpha_i h_i(z), \quad h_i(z) = \frac{(2i+1)z^{2i}}{(z^2+2)^{i+3/2}}. \tag{A.4}$$

If $\sigma(z) = 0$, $\overline{\alpha}_0(z) = 1$.

Moreover, for the updating step $\varphi_z(\sigma^2, \alpha) = (\hat{\sigma}^2(z), \hat{\alpha}(z))$ with $\hat{\alpha}(z)$ as given above and

$$\hat{\sigma}^2(z) = \sigma^2 \frac{z^2}{z^2 + 2\lambda\sigma^2}. \tag{A.5}$$

Appendix B. Proof of Lemma 3.3

Under assumptions (i) and (ii) of the lemma, we deduce that

$$U'_n = \frac{U_n}{\sigma_n\sqrt{2}} \xrightarrow[n\to\infty]{\mathcal{L}} U' = \frac{U}{\sigma\sqrt{2}},$$

where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution. Let us introduce, for all n, $V_n = U_n'^2$. It is easy to see that the law of V_n is a mixture of gamma laws and admits the density

$$f_n(v) = \sum_{i \geqslant 0} \alpha_i^{(n)} \mathrm{e}^{-v} v^{i-1/2} \frac{\mathrm{d}v}{\Gamma(i+1/2)}, \quad \text{for } v > 0.$$

Since the laws of U'_n and U' are symmetric around zero, we are done if we prove that, if V_n is a sequence of random variables with, for all n, $\mathcal{L}(V_n) = f_n(v)\mathrm{d}v$ that converges in distribution to a random variable V with law μ, then, for all $i \geqslant 0$, $\alpha \underset{n\to\infty}{\longrightarrow} \alpha_i$, where $\alpha_i \geqslant 0$ and $\sum_{i \geqslant 0} \alpha_i = 1$; moreover, $\mu(\mathrm{d}v) = f(v)\mathrm{d}v$, with

$$f(v) = \sum_{i \geqslant 0} \alpha_i \mathrm{e}^{-v} v^{i-1/2} \frac{\mathrm{d}v}{\Gamma(i+1/2)}, \quad \text{for } v > 0.$$

Let $\Omega = \{z : \mathrm{Re}(z) < 0\}$, where $\mathrm{Re}(z)$ denotes the real part of the complex number z. We define, for $z \in \overline{\Omega}$, $\varphi_n(z) = \mathrm{E}[\mathrm{e}^{zV_n}]$ and $\varphi(z) = \mathrm{E}[\mathrm{e}^{zV}]$. Functions φ_n and φ are continuous on $\overline{\Omega}$ and holomorphic on Ω. It is straightforward to check that

$$\varphi_n(z) = \sum_{i \geqslant 0} \alpha_i^{(n)} \frac{1}{(1-z)^{i+1/2}}, \quad \text{for } z \in \overline{\Omega}. \tag{B.1}$$

Moreover since $V_n \xrightarrow{\mathcal{L}} V$, we have that

$$\varphi_n(z) \underset{n \to \infty}{\longrightarrow} \varphi(z), \quad \text{for } z \in \overline{\Omega}. \tag{B.2}$$

Let introduce, for z in $\overline{\Omega}$, the function

$$\chi(z) = \frac{1}{1-z}.$$

χ is continuous on $\overline{\Omega}$, holomorphic on Ω and one to one from Ω to $D = \{z \in \mathbb{C} : |z - \frac{1}{2}| < \frac{1}{2}\}$. Its inverse is $\chi^{-1} : D \to \Omega$, $\chi^{-1}(z) = 1 - 1/z$. Define, moreover, for $z \in D$, the functions

$$\psi_n(z) = z^{-1/2} \varphi_n(\chi^{-1}(z)) = \sum_{i \geqslant 0} \alpha_i^{(n)} z^i, \tag{B.3}$$

$$\psi(z) = z^{-1/2} \varphi(\chi^{-1}(z)). \tag{B.4}$$

From (B.3) we deduce that, for all n, ψ_n is holomorphic on $U = \{z : |z| < 1\}$ and continuous on $\overline{U}$. Moreover, $|\psi_n| \leqslant 1$ on U. By Montel's theorem (see, for example, Yger 2001), we deduce that we can extract a subsequence (ψ_{n_k}) from (ψ_n) that converges uniformly on any compact of U to a function g holomorphic on U. Now from (B.2) and (B.3) we know that $\psi_n(z) \to \psi(z)$, for z in D. Thus $g = \psi$ on D, and g is an extension of ψ to U. Moreover, any convergent subsequence (ψ_{m_k}) converges to a function h which by Montel's theorem is holomorphic, and coincides with g on D, therefore $h = g$ on U. Thus any convergent subsequence (ψ_{m_k}) converges to g on U. Taking into account the fact that $|\psi_n| \leqslant 1$ on U, we deduce that the sequence (ψ_n) converges uniformly to g on any compact of U. Since (ψ_n) and g are holomorphic on U, we also have the convergence on U of the derivatives of any order of ψ_n to the corresponding derivative of g. In particular, for any $i \geqslant 0$,

$$\alpha_i^{(n)} = \frac{1}{n!} \psi_n^{(i)}(0) \underset{n \to \infty}{\longrightarrow} \alpha_i := \frac{1}{n!} g^{(i)}(0).$$

Moreover, since $\sum_{i \geqslant 0} \alpha_i^{(n)} = 1$, for all n, Fatou's lemma implies that $\sum_{i \geqslant 0} \alpha_i \leqslant 1$. Consider, for $x \in [0, 1]$, $g(x) = \sum_{i \geqslant 0} \alpha_i x^i$, where g is continuous on $[0, 1]$ and $g(1) = \psi(1) = \varphi(0) = 1$, since φ is the Laplace transform of a probability distribution. Hence we have proved that, for all $z \in U$, $g(z) = \sum_{i \geqslant 0} \alpha_i z^i$, where $\alpha_i \geqslant 0$ and $\sum_{i \geqslant 0} \alpha_i = 1$. In particular, we deduce that, for all $z \in \overline{\Omega}$, $g(\chi(z)) = (\chi(z))^{-1/2} \varphi(z)$, which implies that

$$\varphi(z) = \sum_{i \geqslant 0} \alpha_i \frac{1}{(1-z)^{i+1/2}}, \quad \text{for } z \in \overline{\Omega},$$

which ends the proof. □

Acknowledgements

This work was partially supported by the Research Training Network 'Statistical Methods for Dynamical Stochastic Models' under the Improving Human Potential Program, financed by the Fifth Framework Program of the European Commission. The authors are grateful to Abdellah Youssfi and Wolfgang Runggaldier for helpful discussions and wish to thank two anonymous referees for comments that helped us to significantly improve the paper.

References

Bougerol, P. (1993) Kalman filtering with random coefficients and contractions. *SIAM. J. Control Optim.*, **31**, 942–959.

Del Moral, P. and Guionnet, A. (2001) On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. H. Poincaré Probab. Statist.*, **37**(2), 155–194.

Elliott, R.J., Lakhdar, A. and Moore, J.B. (1995) *Hidden Markov Models: Estimation and Control*. Berlin: Springer-Verlag.

Genon-Catalot, V. (2003) A non-linear explicity filter. *Statist. Probab. Lett.*, **61**, 145–154.

Runggaldier, W.J. and Spizzichino, F. (2001) Sufficient conditions for finite dimensionality for filters in discrete time: a Laplace transform-based approach. *Bernoulli*, **7**, 211–221.

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

Yger, A. (2001) *Analyse complexe et distributions: cours et exercices corrigés.* Paris: Ellipses.

Received October 2002 and revised January 2004