Global and Stochastic Analysis
Vol. 1, No. 1, January-June, 2014

www.mukpublications.com

Filtering with exponential criteria via linear observation channels

M.L. Kleptsyna[a], A. Le Breton[b] and M. Viot[c]

[a] Laboratoire de Statistique et Processus / Université du Maine
Av. Olivier Messiaen, 72085 Le Mans, Cedex 9, France
e-mail: Marina.Kleptsyna@univ-lemans.fr

[b] Laboratoire Jean Kuntzmann / Université J. Fourier
BP 53, 38041 Grenoble Cedex 9, France
e-mail: Alain.Le-Breton@imag.fr

[c] Laboratoire Jean Kuntzmann / Université J. Fourier
BP 53, 38041 Grenoble Cedex 9, France

Abstract

The explicit solution of the discrete time filtering problems with exponential criteria for a general Gaussian signal is obtained through an approach based on a conditional Cameron-Martin type formula. This key formula is derived for conditional expectations of exponentials of some quadratic forms of Gaussian sequences. The formula involves conditional expectations and conditional covariances in some auxiliary optimal risk-neutral filtering problem which is used in the proof. Closed form recursions of Volterra type for these ingredients are provided. Particular cases for which the results can be further elaborated are investigated.

Key words: Gaussian process, optimal filtering, filtering error, Riccati-Volterra equation, risk-sensitive filtering, exponential criteria

1 Introduction

The linear exponential Gaussian (LEG for short) filtering problem, *i.e.*, with an exponential of integral performance index (see the definition (2.3) below), and the so called risk-sensitive (RS for short) filtering problem (see [2] and the statement (4.1) below) have been given a great deal of interest over the last decades. Numerous results have been already reported in specific models, specially around Markov models, but, as far as we know, without exhibiting the relationship between these two problems. See, *e.g.*, Whittle [19], Speyer *et al.* [16], Elliott *et al.*, [4], [6] and Bensoussan and van Schuppen [1] for contributions on this subject and related LEG and RS control problems. Therein the notion of "information state" has been introduced without any clear probabilistic meaning for auxiliary processes which are involved, even in the Gauss-Markov case. Moreover, the method proposed in [6] does not work in a non Markovian situation. On the other hand, the general solution for the optimal risk-neutral linear filtering problem and a Cameron-Martin type formula for a general Gaussian process have been obtained in [8], and [9]. It seems natural to use the approach proposed in [8] and [9] to derive the solution of the LEG and RS filtering problems for a general Gaussian observation signal and to precise their link. For details, see Section where in particular we prove that the LEG and RS filtering problems have the same solution and we propose an example to show that in a bit more general setting, two similar problems may have different solutions. The exponential criterion type problems in continuous time setting have been analyzed in [10]-[12].

In this paper we are interested in the explicit solution of the Linear Exponential Gaussian (LEG) and Risk Sensitive (RS) filtering problems for general Gaussian signals. Namely we deal with a signal-observation model $(X_t, Y_t)_{t\geq 1}$, where the signal $X = (X_t)_{t\geq 1}$ is an arbitrary Gaussian sequence with mean $m = (m_t, t \geq 1)$ and covariance $K = (K(t,s), t \geq 1, s \geq 1)$, *i.e.*,

$$\mathbb{E}X_t = m_t, \quad \mathbb{E}(X_t - m_t)(X_s - m_s) = K(t,s)\,, \quad t \geq 1\,,\ s \geq 1\,,$$

and, for some sequence $A = (A_t,\, t \geq 1)$ of the real numbers, the observation process $Y = (Y_t,\, t \geq 1)$ is given by

$$Y_t = A_t X_t + \varepsilon_t, \tag{1.1}$$

where $\varepsilon = (\varepsilon_t)_{t\geq 1}$ is a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables and ε and X are independent.

Suppose that only Y is observed and for a given real number μ and a sequence $(Q_t)_{t\geq 1}$ of nonnegative real numbers, one wishes to minimize with respect to $h:\ h_t \in \mathcal{Y}_t, t \geq 1$ the quantity:

$$\mathbb{E}\,\mu \exp\left\{\frac{\mu}{2}\sum_{t=1}^{T}(X_t - h_t)^2 Q_t\right\}, \tag{1.2}$$

where $(\mathcal{Y}_t)$ is the natural filtration of Y, *i.e.*, $\mathcal{Y}_t = \sigma(\{Y_u\,,\, 1 \leq u \leq t\})$ and $h_t \in \mathcal{Y}_t$ means that h_t is $\mathcal{Y}_t$-measurable.

Note that, according to the sign of the real parameter μ, there are two different cases for this linear exponential Gaussian (LEG) filtering problem (the terminology is taken from the linear exponential Gaussian optimal control problem) :

- $\mu < 0$, called *risk-preferring* filtering problem,
- $\mu > 0$, called the *risk-averse* filtering problem.

It is well known (see, *e.g.*, [16] for the discrete time Markov case setting) that the solution to this problem is not the conditional expectation of X_t given the σ-field $\mathcal{Y}_t$. Our first aim is to show that the solution can be completely explicited : the characteristics of the optimal solution are obtained as the solution of a closed form system of Volterra type equations which actually reduce to the equations known also for the RS setting when the signal process X is Gauss-Markov (see, *e.g.*, [15]). Our second aim is to give the probabilistic interpretation of this optimal solution in terms of an auxiliary *risk-neutral* filtering problem. Actually, we extend the filtering approach initiated in [8] and [10] for one-dimensional processes, to obtain a conditional Cameron-Martin type formula for the *conditional Laplace transform* of a quadratic functional of the involved process. Namely, we give an explicit representation for the random variable

$$\mathcal{I}_T = \mathbb{E}\left(\exp\left\{ \frac{\mu}{2} \sum_{s=1}^{T} (X_s - h_s)^2 Q_s \right\} \Big/ \mathcal{Y}_T \right), \tag{1.3}$$

where $h_s \in \mathcal{Y}_s$, $s \geq 1$.

The paper is organized as follows. In Section we derive the solution of the LEG filtering problem : explicit recursive equations, involving the covariance function of the filtered process, are obtained. In particular, in Section , an appropriate auxiliary risk-neutral filtering problem is matched to that of deriving the key Cameron-Martin type formula. The solution of this auxiliary filtering problem is discussed in Section . In Section we investigate some specific cases where the results can be further elaborated. In Section we discuss the relationship between LEG and RS filtering problems.

2 Solution of the LEG filtering problem

Let us introduce the following condition (C_μ):

(C_μ) the equation

$$\overline{\gamma}(t,s) = K(t,s) - \sum_{l=1}^{s-1} \overline{\gamma}(t,l)\overline{\gamma}(s,l)\, \frac{S_l}{1 + S_l \overline{\gamma}_l}, \quad S_l = A_l^2 - \mu Q_l \tag{2.1}$$

has a unique and bounded solution on $\{(t,s): 1 \le s \le t \le T\}$, such that $\overline{\gamma}_l = \overline{\gamma}(l,l) \ge 0$, $l \ge 1$ and moreover

$$1 + S_l \overline{\gamma}_l > 0,\ l \ge 1.$$

Remark 1. *Notice that for all μ* ***negative*** *the condition (C_μ) is satisfied and if μ is* ***positive****, the condition (C_μ) is satisfied for μ sufficiently small, for example, those such that for any $t \le T$ $A_t^2 - \mu Q_t$ is nonnegative.*

The first result is the following

Theorem 1. *Suppose that the condition (C_μ) is satisfied. Let $(\overline{h}_t)_{t\ge 1}$ be the solution of the following equation:*

$$\overline{h}_t = m_t + \sum_{l=1}^{t} A_l \overline{\gamma}(t,l)(Y_l - A_l \overline{h}_l), \tag{2.2}$$

where $\overline{\gamma} = (\overline{\gamma}(t,s), 1 \le s \le t \le T)$ is the unique solution of equation (2.1)*.*

Then $(\overline{h}_t)_{t\ge 1}$ is the solution of the LEG filtering problem, i.e.,

$$\overline{h} = \underset{h:\, h_t \in \mathcal{Y}_t, t\ge 1}{\operatorname{argmin}} \mathbb{E}\, \mu \exp\left\{ \frac{\mu}{2} \sum_{t=1}^{T} (X_t - h_t)^2 Q_t \right\}. \tag{2.3}$$

Moreover, the corresponding optimal risk is given by

$$\mathbb{E}\, \mu \exp\left\{ \frac{\mu}{2} \sum_{t=1}^{T} (X_t - \overline{h}_t)^2 Q_t \right\} = \mu \prod_{t=1}^{T} \left[\frac{1 + S_t \overline{\gamma}_t}{1 + A_t^2 \overline{\gamma}_t} \right]^{-1/2}$$

Theorem 1 is a direct consequence of results of Section . Its proof will be given at the end of Section .

Remark 2.

- *Note that equation* (2.2) *is really recursive equation and it can be rewritten in the equivalent form:*

$$\overline{h}_t = \frac{1}{1 + A_t^2 \overline{\gamma}_t} \left[m_t + \sum_{l=1}^{t-1} A_l \overline{\gamma}(t,l)(Y_l - A_l \overline{h}_l) + A_t \overline{\gamma}_t Y_t \right],$$

- *It is worth emphasizing that taking $\mu = 0$ in equation (2.1), one gets through equation (2.2) the solution $\overline{h}$ of the risk-neutral filtering problem of the signal X given the observation Y, i.e., $\overline{h}_t = \mathbb{E}(X_t/\mathcal{Y}_t)$ (see, e.g., [8]).*

2.1 Conditional version of a Cameron-Martin formula

The proof of Theorem 1 is based on the conditional version of the Cameron–Martin formula provides the conditional expectation $\mathcal{I}_t$ defined by (1.3). Let

$$J_t = \exp\left\{-\frac{1}{2}\sum_{s=1}^{t}(X_s - h_s)^2 Q_s\right\}. \tag{2.4}$$

Then $\mathcal{I}_t = \pi_t(J_t)$, where for any random variable η such that $\mathbb{E}|\eta| < +\infty$, the notation $\pi_t(\eta)$ is used for the conditional expectation of η given the σ-field $\mathcal{Y}_t = \sigma(\{Y_s\,,\, 1 \le s \le t\})$,

$$\pi_t(\eta) = \mathbb{E}(\eta/\mathcal{Y}_t)\,.$$

Proposition 2. *Suppose that the condition* (C_μ) *is satisfied. Let* $(\overline{\gamma}(t,s),\ 1 \le s \le t \le T)$ *be the solution of equation* (2.1) *and* $(Z_t^h,\, t \ge 1)$ *be the solution of the following equation*

$$Z_t^h = m_t - \sum_{l=1}^{t-1}\overline{\gamma}(t,l)\frac{\mu Q_l}{1+S_l\overline{\gamma}_l}(h_l - Z_l^h) + \sum_{l=1}^{t-1}\overline{\gamma}(t,l)\frac{A_l}{1+S_l\overline{\gamma}_l}(Y_l - A_l Z_l^h). \tag{2.5}$$

Then the following representation of the random variable $\mathcal{I}_t$ *defined by* (1.3) *holds for any* $T \ge 1$*:* $\mathcal{I}_T =$

$$\prod_{t=1}^{T}\left[\frac{1+S_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}\right]^{-1/2} \times \exp\left\{\frac{\mu}{2}Q_t\frac{1+A_t^2\overline{\gamma}_t}{1+S_t\overline{\gamma}_t} \times \left[h_t - \frac{Z_t^h + A_t\overline{\gamma}_t Y_t}{1+A_t^2\overline{\gamma}_t}\right]^2\right\} \times \mathcal{M}_T,$$

where $(\mathcal{M}_T)_{T\ge1}$ *is a martingale defined by :*

$$\mathcal{M}_T = \prod_{t=1}^{T}\left[\frac{(1+A_t^2\gamma_t)}{1+A_t^2\overline{\gamma}_t}\right]^{1/2} \exp\left\{\frac{A_t}{1+A_t^2\overline{\gamma}_t}\,(Z_t^h - \pi_{t-1}(X_t))\nu_t - \right.$$
$$\left. -\frac{1}{2}\cdot\frac{A_t^2}{1+A_t^2\overline{\gamma}_t}\,(Z_t^h - \pi_{t-1}(X_t))^2 - \frac{1}{2}\cdot\frac{A_t^2(\gamma_t - \overline{\gamma}_t)\cdot\nu_t^2}{(1+A_t^2\overline{\gamma}_t)(1+A_t^2\gamma_t)}\right\}, \tag{2.6}$$

in terms of the innovation sequence $(\nu_t)_{t\ge1}$*:*

$$\nu_t = Y_t - A_t\pi_{t-1}(X_t); \quad \pi_{t-1}(X_t) = \mathbb{E}(X_t/\mathcal{Y}_{t-1}),$$

and of the variances of one-step prediction errors $(\gamma_t)_{t\ge1}$*:*

$$\gamma_t = \mathbb{E}(X_t - \pi_{t-1}(X_t))^2.$$

Remark 3.

1. *The probabilistic interpretation the auxiliary processes* (Z_t^h) *and* $(\overline{\gamma}_t)_{t\ge1}$ *appearing in the Proposition 2 will be clarified below.*

2. *Proposition 2 reduces to the ordinary Cameron-Martin type formula (c.f. Theorem 1 [8] for* $h \equiv 0$ *when* $A_t = 0$, $l \ge 1$ *and hence* X *and* Y *are independent.*

Proof of Proposition 2 We will prove Proposition 2 for $\mu < 0$, namely $\mu = -1$. Then we can replace Q by $-\mu Q$ and the statement of Proposition 2 is still valid because of the analytical properties of the involved functions.

The proof of Proposition 2 for $\mu = -1$ will be separated into two steps.

I. (Actually it is the discrete time analog for the general filtering theorem.) Since $h_t \in \mathcal{Y}_t$, $t \geq 1$, in the proof we can suppose that h is a deterministic function. First of all, we claim that for J_t, defined by (2.4)

$$\pi_t(J_t) = \left.\frac{\pi_{t-1}(J_t \beta_t^y)}{\pi_{t-1}(\beta_t^y)}\right|_{y=Y_t}, \tag{2.7}$$

where $\beta_t^y = \exp(A_t X_t y - \frac{1}{2} A_t X_t^2)$.

Indeed, let us introduce the new probability measure $\hat{\mathbb{P}}$, defined by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \exp(-A_t X_t \varepsilon_t - \frac{1}{2} A_t^2 X_t^2).$$

The classical Bayes formula gives that

$$\pi_t(J_t) = \frac{\hat{\pi}_t(J_t \exp(A_t X_t \varepsilon_t + \frac{1}{2} A_t^2 X_t^2))}{\hat{\pi}_t(\exp(A_t X_t \varepsilon_t + \frac{1}{2} A_t^2 X_t^2))} = \frac{\hat{\pi}_t(J_t \exp(A_t X_t Y_t - \frac{1}{2} A_t^2 X_t^2))}{\hat{\pi}_t(\exp(A_t X_t Y_t - \frac{1}{2} A_t^2 X_t^2))},$$

where $\hat{\pi}_t(\cdot)$ denotes a conditional expectation with respect to $\mathcal{Y}_t$ under $\hat{\mathbb{P}}$. Note that under $\hat{\mathbb{P}}$ the distribution of $(X_s, Y_r)_{s \leq t, r \leq t-1}$ is the same as under $\mathbb{P}$ and Y_t is a $\mathcal{N}(0,1)$ random variable independent of $(X_s, Y_r)_{s \leq t, r \leq t-1}$.

To understand this point it is sufficient to write the following equality for the mutual characteristic function with arbitrary real numbers (α_j, λ_j):

$$\begin{aligned}
&\hat{\mathbb{E}} \exp\left\{ i \sum_{j=1}^t \alpha_j X_j + i \sum_{j=1}^t \lambda_j Y_j \right\} = \\
&\quad \mathbb{E} \exp\left\{ i \sum_{j=1}^t \alpha_j X_j + i \sum_{j=1}^{t-1} \lambda_j Y_j + i\lambda_t Y_t - A_t X_t \varepsilon_t - \frac{1}{2} A_t^2 X_t^2 \right\} = \\
&\mathbb{E}\left(\mathbb{E} \exp\left\{ i \sum_{j=1}^t \alpha_j X_j + i \sum_{j=1}^{t-1} \lambda_j Y_j + i\lambda_t Y_t - A_t X_t \varepsilon_t - \frac{1}{2} A_t^2 X_t^2 \right\} \Big/ \mathcal{X}_t \right) = \\
&\mathbb{E} \exp\left\{ i \sum_{j=1}^t \alpha_j X_j + i \sum_{j=1}^{t-1} \lambda_j Y_j + i\lambda_t A_t X_t - \frac{1}{2} A_t^2 X_t^2 + \frac{1}{2}(i\lambda_t - A_t X_t)^2 \right\} = \\
&\qquad\qquad e^{-\frac{1}{2}\lambda_t^2}\, \mathbb{E} \exp\left\{ i \sum_{j=1}^t \alpha_j X_j + i \sum_{j=1}^{t-1} \lambda_j Y_j \right\},
\end{aligned}$$

where $\mathcal{X}_t$ is the σ-field $\mathcal{X}_t = \sigma(\{X_s\,,\, 1 \le s \le t\})$. Hence,

$$\hat{\pi}_t(J_t \exp(A_t X_t Y_t - \frac{1}{2}A_t^2 X_t^2)) =$$
$$= \pi_{t-1}(J_t \exp(A_t X_t y - \frac{1}{2}A_t^2 X_t^2))|_{y=Y_t} =$$
$$= \pi_{t-1}(J_t \beta_t^y)|_{y=Y_t}.$$

Similarly,

$$\hat{\pi}_t \left(\exp(A_t X_t y - \frac{1}{2}A_t^2 X_t^2) \right) = \pi_{t-1}(\beta_t^y)\big|_{y=Y_t}\ ,$$

and hence (2.7) holds.

II. In the second step we will calculate the ratio $\frac{\mathcal{I}_t}{\mathcal{I}_{t-1}}$ which, due to (2.7) can be rewritten as

$$\frac{\mathcal{I}_t}{\mathcal{I}_{t-1}} = \frac{\pi_t(J_t)}{\pi_{t-1}(J_{t-1})} = \left.\frac{\pi_{t-1}(J_t \beta_t^y)}{\pi_{t-1}(J_{t-1})\pi_{t-1}(\beta_t^y)}\right|_{y=Y_t}. \tag{2.8}$$

For this aim similarly to what we proposed in [8] and [10] we introduce the auxiliary processes $(Y_t^2)_{t\ge1}$ and $(\xi_t)_{t\ge1}$. Let $\bar{\varepsilon} = (\bar{\varepsilon}_t)_{t\ge1}$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables independent of X and define $(Y_t^2, \xi_t)_{t\ge1}$ by:

$$Y_t^2 = Q_t(X_t - h_t) + \sqrt{Q_t}\bar{\varepsilon}_t, \tag{2.9}$$

$$\xi_t = \sum_{s=1}^{t} (X_s - h_s) Y_s^2. \tag{2.10}$$

Now the following equality holds:

$$\left.\frac{\pi_{t-1}(J_t\beta_t^y)}{\pi_{t-1}(J_{t-1})}\right|_{y=Y_t} = \left.\frac{\overline{\pi}_{t-1}(\exp\{-\frac{1}{2}Q_t(X_t-h_t)^2 - \xi_{t-1}\}\beta_t^y)}{\overline{\pi}_{t-1}(\exp(-\xi_{t-1}))}\right|_{y=Y_t},$$

where $\overline{\pi}_t(\cdot)$ stands for a conditional expectation w.r.t. to the σ-field $\bar{\mathcal{Y}}_t = \sigma(\{Y_s, Y_s^2, s \le t\})$ under the initial measure $\mathbb{P}$.

Again the proof of this equality is based on the Bayes formula. Namely, let $\tilde{\mathbb{P}}$ be the new probability measure defined by

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \rho_{t-1} = \exp\left\{ -\frac{1}{2}\sum_{1}^{t-1} Q_s(X_s-h_s)^2 - \sum_{1}^{t-1}\sqrt{Q_s}(X_s-h_s)\bar{\varepsilon}_s \right\}. \tag{2.11}$$

Then $J_t\rho_{t-1} = \exp\{-\xi_{t-1} - \frac{1}{2}Q_t(X_t - h_t)^2\}$ and $J_{t-1}\rho_{t-1} = \exp\{-\xi_{t-1}\}$. Thus

$$\left.\frac{\overline{\pi}_{t-1}(\exp(-\xi_t - \frac{1}{2}Q_t(X_t-h_t)^2)\beta_t^y)}{\overline{\pi}_{t-1}(\exp(-\xi_{t-1}))}\right|_{y=Y_t} =$$

$$= \left.\frac{\mathbb{E}(J_t\beta_t^y\rho_{t-1}/\bar{\mathcal{Y}}_{t-1})}{\mathbb{E}(\rho_{t-1}/\bar{\mathcal{Y}}_{t-1})} \cdot \frac{\mathbb{E}(\rho_{t-1}/\bar{\mathcal{Y}}_{t-1})}{\mathbb{E}\exp(J_{t-1}\rho_{t-1})/\bar{\mathcal{Y}}_{t-1})}\right|_{y=Y_t} =$$

$$= \left.\frac{\tilde{\mathbb{E}}(J_t\beta_t^y/\bar{\mathcal{Y}}_{t-1})}{\tilde{\mathbb{E}}(J_{t-1}/\bar{\mathcal{Y}}_{t-1})}\right|_{y=Y_t} = \left.\frac{\pi_{t-1}(J_t\beta_t^y)}{\pi_{t-1}(J_{t-1})}\right|_{y=Y_t},$$

where the last equality holds because under the probability measure $\tilde{\mathbb{P}}$ the distribution of $(X_s, Y_s)_{s\leq t}$ is the same as under the initial measure $\mathbb{P}$ and $(X_s, Y_s)_{s\leq t-1}$ is independent of $(Y_s^2)_{s\leq t-1}$.

Finally we have proved the following:

$$\frac{\pi_t(J_t)}{\pi_{t-1}(J_{t-1})} = \left.\frac{\overline{\pi}_{t-1}(\exp\left[-\xi_{t-1} + A_tX_ty - \frac{1}{2}Q_t(X_t-h_t)^2 - \frac{1}{2}A_t^2X_t^2\right])}{\overline{\pi}_{t-1}(\exp(-\xi_{t-1}))\pi_{t-1}(\beta_t^y)}\right|_{y=Y_t}. \tag{2.12}$$

At this point we will use the conditionally Gaussian properties of (X_t, ξ_{t-1}) w.r.t. $\bar{\mathcal{Y}}_{t-1}$ and Lemma 11.6 [13] which says that for a Gaussian pair (U, V) with mean values m_U, m_V, variances γ_U, γ_V and covariance γ_{UV}

$$\begin{aligned}\mathbb{E}\exp\left\{-\frac{1}{2}DU^2 + \lambda_1U - \lambda_2V\right\} &= (1+D\gamma_U)^{-1/2}\times \\ \times\exp\Bigg\{-\lambda_2m_V + \frac{\lambda_2^2}{2}\gamma_V - \frac{1}{2}\cdot\frac{D}{1+D\gamma_U}&(m_U - \lambda_2\gamma_{UV})^2 + \\ &+ \frac{\lambda_1^2\gamma_U + 2\lambda_1(m_U - \lambda_2\gamma_{UV})}{2(1+D\gamma_U)}\Bigg\},\end{aligned} \tag{2.13}$$

for any real numbers λ_1, λ_2 and $D \geq 0$. Indeed, in (2.12) we will apply this formula to $(U, V) = (X_t, \xi_{t-1})$ given $\bar{\mathcal{Y}}_{t-1}$ with

$$D = S_t = Q_t + A_t^2, \quad \lambda_2 = 1, \quad \lambda_1 = A_ty + Q_th_t,$$

in the numerator and $D = \lambda_1 = 0, \quad \lambda_2 = 1$ in the first factor of the denominator and again to $(U, V) = (X_t, \xi_{t-1})$ given $\mathcal{Y}_{t-1}$ with

$$D = A_t^2, \quad \lambda_2 = 0, \quad \lambda_1 = A_ty,$$

in the second factor of the denominator.

Collecting the terms as coefficients for h_t^2 and h_t, we obtain that

$$\frac{\mathcal{I}_t}{\mathcal{I}_{t-1}} = \frac{(1+S_t\overline{\gamma}_t)^{-1/2}}{(1+A_t^2\gamma_t)^{-1/2}} \cdot \exp\left\{-\frac{Q_t}{2}\frac{1+A_t^2\overline{\gamma}_t}{1+S_t\overline{\gamma}_t} \times \left[h_t - \frac{Z_t^h + A_t\overline{\gamma}_t Y_t}{1+A_t^2\overline{\gamma}_t}\right]^2\right\} \times$$

$$\times \exp\left\{-\frac{A_t^2(Z_t^h)^2 - A_t^2\overline{\gamma}_t Y_t^2}{2(1+A_t^2\overline{\gamma}_t)} + \frac{Y_t Z_t^h A_t}{1+A_t^2\overline{\gamma}_t}\right.$$

$$\left.+\frac{1}{2}\cdot\frac{A_t^2\pi_{t-1}^2(X_t) - 2A_t\pi_{t-1}(X_t)Y_t - A_t^2 Y_t^2\gamma_t}{1+A_t^2\gamma_t}\right\},$$

where $Z_t^h = \overline{\pi}_{t-1}(X_t) - \overline{\gamma}_{X\xi}(t)$ with

$$\overline{\gamma}_{X\xi}(t) = \mathbb{E}[(X_t - \overline{\pi}_{t-1}(X_t))(\xi_{t-1} - \overline{\pi}_{t-1}(\xi_{t-1}))/\bar{\mathcal{Y}}_{t-1}],\ t \geq 2\,;\ \overline{\gamma}_{X\xi}(1) = 0\,. \tag{2.14}$$

To finish the proof we just replace Y_t by $\nu_t + A_t\pi_{t-1}(X_t)$. Thus in the last exponential term we find:

$$\exp\left\{-\frac{\nu_t^2 A_t^2(\gamma_t - \overline{\gamma}_t)}{2(1+A_t^2\overline{\gamma}_t)(1+A_t^2\gamma_t)} + \frac{Z_t^h - \pi_{t-1}(X_t)}{1+A_t^2\overline{\gamma}_t}A_t\nu_t - \right.$$

$$\left.-\frac{1}{2}\cdot\frac{A_t^2}{1+A_t^2\overline{\gamma}_t}(Z_t^h - \pi_{t-1}(X_t))^2\right\},$$

which gives the Proposition.

Remark 4. 1. *Note that now the probabilistic interpretation of the ingredients $\overline{\gamma}_t$ and Z_t^h is clarified for* ***negative*** *μ. Namely, $\overline{\gamma}_t = \mathbb{E}(X_t - \overline{\pi}_{t-1}(X_t))^2$, and $Z_t^h = \overline{\pi}_{t-1}(X_t) - \overline{\gamma}_{X\xi}(t)$, but when μ is* ***positive****, there is no such connection anymore.*

2. *Observe that actually $\overline{\pi}_{t-1}(X_t)$ and $\overline{\gamma}_{X\xi}(t)$ are $\bar{\mathcal{Y}}_{t-1}$-measurable, but Z_t^h is $\mathcal{Y}_{t-1}$ measurable.*

Proof of Theorem 1 The statement of Theorem 1 is the direct consequence of Proposition . Indeed, we claim that the following chain of inequalities holds for any $h:\ h_t \in \mathcal{Y}_t, t \geq 1$:

$$\mathbb{E}\,\mu\exp\left\{\frac{\mu}{2}\sum_{t=1}^{T}(X_t - h_t)^2 Q_t\right\}$$

$$= \mathbb{E}\left[\mathbb{E}\,\mu\left(\exp\left\{\frac{\mu}{2}\sum_{t=1}^{T}(X_t - h_t)^2 Q_t\right\}\Big/\ \mathcal{Y}_T\right)\right]$$

$$= \mu\,\mathbb{E}\prod_{t=1}^{T}\left[\frac{1+S_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}\right]^{-1/2} \times \exp\left\{\frac{\mu}{2}Q_t\frac{1+A_t^2\overline{\gamma}_t}{1+S_t\overline{\gamma}_t} \times \left[h_t - \frac{Z_t^h + A_t\overline{\gamma}_t Y_t}{1+A_t^2\overline{\gamma}_t}\right]^2\right\} \times$$

$$\times \mathcal{M}_T, \overset{(a)}{\geq} \prod_{t=1}^{T}\left[\frac{1+S_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}\right]^{-1/2} \mu\,\mathbb{E}\,\mathcal{M}_T$$

$$\overset{(b)}{=} \mu \prod_{t=1}^{T}\left[\frac{1+S_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}\right]^{-1/2}.$$

Of course under condition (C_μ), since the term in the last line is finite, it is sufficient to consider the case:

$$\mathbb{E}\,\mu\exp\left\{\frac{\mu}{2}\sum_{t=1}^{T}(X_t - h_t)^2 Q_t\right\} < \infty, \tag{2.15}$$

which gives the first equality. Inequality (a) follows directly from Proposition . Equality (b) is a direct consequence of (2.13) which gives that $\mathbb{E}\,\mathcal{M}_T = 1$. Now, to obtain the lower bound we must take

$$\overline{h}_t = \frac{Z_t^{\overline{h}} + A_t\overline{\gamma}_t Y_t}{1+A_t^2\overline{\gamma}_t},\ t \geq 1,$$

or equivalently

$$\overline{h}_t = Z_t^{\overline{h}} + \frac{A_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}(Y_t - A_t Z_t^{\overline{h}}),\ t \geq 1,$$

where Z^h is the solution of equation (2.5), which means that

$$Z_t^{\overline{h}} = m_t + \sum_{l=1}^{t-1}\frac{\overline{\gamma}(t,l)A_l}{1+A_l^2\overline{\gamma}_l}[Y_l - A_l Z_l^{\overline{h}}],$$

and hence

$$\overline{h}_t = m_t + \sum_{l=1}^{t}\frac{\overline{\gamma}(t,l)A_l}{1+A_l^2\overline{\gamma}_l}[Y_l - A_l Z_l^{\overline{h}}] = m_t + \sum_{l=1}^{t} A_l\overline{\gamma}(t,l)(Y_l - A_l\overline{h}_l).$$

Thus $\bar{h}$ is the unique solution of equation (2.2). Finally for $\bar{h}$ the lower bound is attained.

2.2 Solution of the auxiliary filtering problems

Here, for an arbitrary Gaussian sequence X, we deal with the one-step prediction and filtering problems of the signals X and ξ given by (2.10) respectively from the observation of $\bar{Y} = (Y, Y^2)$ defined in (1.1) and (2.9). Actually, we follow the ideas proposed in our paper [8]. Recall that the solutions can be reduced to equations for the conditional moments. The following statement provides the equations for the characteristics which give the solution of the prediction problem and the equation for the other quantity $\overline{\pi}_{t-1}(X_t) - \overline{\gamma}_{X\xi}(t)$ appearing in Proposition 2 for $\mu = -1$.

Theorem 3. *The conditional mean $\overline{\pi}_{t-1}(X_t)$ and the variance of the one-step prediction error $\overline{\gamma}_t = \mathbb{E}[X_t - \overline{\pi}_{t-1}(X_t)]^2$ are given by the equations*

$$\overline{\pi}_{t-1}(X_t) = m_t + \sum_{s=1}^{t-1} \frac{\overline{\gamma}(t,s)}{1 + (A_s^2 + Q_s)\overline{\gamma}_s} [A_s(Y_s - A_s\overline{\pi}_{s-1}(X_s)) \\ + Q_s(Y_s^2 - Q_s(\overline{\pi}_{s-1}(X_s) - h_s)] , \quad t \geq 1 , \quad (2.16)$$

$$\overline{\gamma}_t = \overline{\gamma}(t,t) , \quad t \geq 1 . \quad (2.17)$$

where $\overline{\gamma} = (\overline{\gamma}(t,s),\, 1 \leq s \leq t)$ is the unique solution of equation (2.1). Moreover, with $\overline{\gamma}_{X\xi}(t)$ defined by (2.14), the difference $\overline{\pi}_{t-1}(X_t) - \overline{\gamma}_{X\xi}(t)$ is the solution Z_t^h of equation (2.5).

Proof Note that since $h_t \in \mathcal{Y}_t$ and that the joint distribution of $(X_r, Y_s, Y_s^2 + Q_s h_s)$ for any $r\,, s$ is Gaussian we can apply the Note following Theorem 13.1 in [14]. For any $k \leq t$ we can write

$$\begin{cases} \overline{\pi}_k(X_t) = \overline{\pi}_{k-1}(X_t) + [\mathrm{cov}(X_t, \overline{\nu}_k)]' \,\mathrm{var}(\overline{\nu}_k)^{-1} \overline{\nu}_k \\ \overline{\pi}_0(X_t) = m_t, \end{cases} \quad (2.18)$$

where

$$\overline{\nu}_k = \bar{Y}_k - \mathbb{E}(\bar{Y}_k / \overline{\mathcal{Y}}_{k-1}) = \begin{pmatrix} Y_k - A_k \overline{\pi}_{k-1}(X_k) \\ Y_k^2 + Q_k h_k - Q_k \overline{\pi}_{k-1}(X_k) \end{pmatrix}$$

is the innovation with covariance matrices

$$\mathrm{var}(\overline{\nu}_k) = \begin{pmatrix} 1 + A_k^2 \overline{\gamma}_k & A_k Q_k \overline{\gamma}_k \\ A_k Q_k \overline{\gamma}_k & Q_k + Q_k^2 \overline{\gamma}_k \end{pmatrix}, \quad (2.19)$$

and

$$\mathrm{cov}(X_t, \overline{\nu}_k) = \overline{\gamma}(t,k) \begin{pmatrix} A_k \\ Q_k \end{pmatrix} \quad (2.20)$$

with

$$\overline{\gamma}(t,k) = \mathbb{E}(X_t - \overline{\pi}_{k-1}(X_t))(X_k - \overline{\pi}_{k-1}(X_k)) . \quad (2.21)$$

By the definition (2.21), we see for $k = t$ that the variance $\overline{\gamma}_t$ is given by (2.17). Now, equality (2.18) implies

$$\overline{\pi}_k(X_t) = m_t + \sum_{l=1}^{k} \overline{\gamma}(t,l) \begin{pmatrix} A_l & Q_l \end{pmatrix} (\operatorname{var} \overline{\nu}_l)^{-1} \overline{\nu}_l =$$

$$= m_t + \sum_{s=1}^{k} \frac{\overline{\gamma}(t,s)}{1+(A_s^2+Q_s)\overline{\gamma}_s} [A_s(Y_s - A_s \overline{\pi}_{s-1}(X_s)) +$$
$$+ Q_s(Y_s^2 - Q_s(\overline{\pi}_{s-1}(X_s) - h_s)] \quad (2.22)$$

and putting $k = t-1$ we get nothing but equation (2.16). Concerning the solution of the one-step prediction problem, it just remains to show that the covariance $\overline{\gamma}(t,s)$ satisfies equation (2.1).

Let us define

$$\delta_X(t,l) = X_t - \overline{\pi}_l(X_t)\,.$$

According to (2.18) we can write

$$\delta_X(t,l) = \delta_X(t,l-1) - \overline{\gamma}(t,l) \begin{pmatrix} A_l & Q_l \end{pmatrix} (\operatorname{var} \overline{\nu}_l)^{-1} \overline{\nu}_l$$

and so

$$\mathbb{E}\, \delta_X(t_1,l)\delta_X(t_2,l) = \mathbb{E}\, \delta_X(t_1,l-1)\delta_X(t_2,l-1) -$$
$$- \overline{\gamma}(t_1,l)\overline{\gamma}(t_2,l) \begin{pmatrix} A_l \\ Q_l \end{pmatrix}' \operatorname{var}(\bar{\nu}_l)^{-1} \begin{pmatrix} A_l \\ Q_l \end{pmatrix}$$

or

$$\mathbb{E}\delta_X(t^1,l)\delta_X(t^2,l) = \mathbb{E}\delta_X(t^1,0)\delta_X(t^2,0) -$$
$$- \sum_{r=1}^{l} \overline{\gamma}(t,r)\overline{\gamma}(s,r) \frac{A_r^2 + Q_r}{1+(A_r^2+Q_r)\overline{\gamma}_r}. \quad (2.23)$$

Taking $t^1 = t\,, t^2 = s\,, l = s-1$ in (2.23), it is readily seen that equation (2.1) holds for $\overline{\gamma}(t,s)$.

Now we analyze the difference $\overline{\pi}_{t-1}(X_t) - \overline{\gamma}_{X\xi}(t)$. Using the representation $\xi_t = \sum_{s=1}^{t}(X_s - h_s)Y_s^2$ we can rewrite $\overline{\pi}_{t-1}(\xi_{t-1})$ in the following form

$$\overline{\pi}_{t-1}(\xi_{t-1}) = \sum_{s=1}^{t-1} (\pi_{t-1}(X_s) - h_s)Y_s^2,$$

which implies that

$$\xi_{t-1} - \overline{\pi}_{t-1}(\xi_{t-1}) = \sum_{s=1}^{t-1} (X_s - \overline{\pi}_{t-1}(X_s))Y_s^2.$$

And so we have

$$\overline{\gamma}_{X\xi}(t) = \sum_{s=1}^{t-1} \overline{\pi}_{t-1}[(X_s - \overline{\pi}_{t-1}(X_s))(X_t - \overline{\pi}_{t-1}(X_t))]Y_s^2 =$$

$$= \sum_{s=1}^{t-1} \mathbb{E}(X_s - \overline{\pi}_{t-1}(X_s))(X_t - \overline{\pi}_{t-1}(X_t))Y_s^2 = \sum_{s=1}^{t-1} \widetilde{\gamma}(t,s)Y_s^2, \quad (2.24)$$

where

$$\widetilde{\gamma}(t,s) = \mathbb{E}(X_s - \overline{\pi}_{t-1}(X_s))(X_t - \overline{\pi}_{t-1}(X_t)) = \overline{\gamma}(s,t). \quad (2.25)$$

Using the definitions (2.21) and (2.25) we can write

$$\widetilde{\gamma}(t,s) - \overline{\gamma}(t,s) = -\,\mathbb{E}\,X_t(\overline{\pi}_{t-1}(X_s) - \overline{\pi}_{s-1}(X_s)).$$

Again, applying the Note following Theorem 13.1 in [14], we can write also

$$\overline{\pi}_l(X_r) = \overline{\pi}_{l-1}(X_r) + \overline{\gamma}(t,l)\begin{pmatrix} A_l & Q_l \end{pmatrix}(\operatorname{var}\overline{\nu}_l)^{-1}\overline{\nu}_l.$$

This means that

$$\pi_{t-1}(X_r) - \pi_{r-1}(X_r) = \sum_{l=r}^{t-1} \overline{\gamma}(t,l)\begin{pmatrix} A_l & Q_l \end{pmatrix}(\operatorname{var}\overline{\nu}_l)^{-1}\overline{\nu}_l\,,$$

or equivalently

$$\pi_{t-1}(X_r) - \pi_{r-1}(X_r) = \sum_{l=r}^{t-1} \widetilde{\gamma}(l,t)\begin{pmatrix} A_l & Q_l \end{pmatrix}(\operatorname{var}\overline{\nu}_l)^{-1}\overline{\nu}_l\,.$$

Then, multiplying by X_t and taking expectations in both sides, we get

$$\mathbb{E}X_t(\pi_{t-1}(X_r) - \pi_{r-1}(X_r)) = \sum_{l=r}^{t-1} \widetilde{\gamma}(l,r)\begin{pmatrix} A_l & Q_l \end{pmatrix}(\operatorname{var}\overline{\nu}_l)^{-1}\operatorname{cov}(X_t, \overline{\nu}_l) =$$

$$= \sum_{l=s}^{t-1} \widetilde{\gamma}(l,s)\overline{\gamma}(t,l)\frac{A_l^2 + Q_l}{1 + (A_l^2 + Q_l)\overline{\gamma}_l}.$$

Hence we have proved the following relation

$$\widetilde{\gamma}(t,s) - \overline{\gamma}(t,s) = -\sum_{l=s}^{t-1} \widetilde{\gamma}(l,s)\overline{\gamma}(t,l)\frac{A_l^2 + Q_l}{1 + (A_l^2 + Q_l)\overline{\gamma}_l}. \quad (2.26)$$

Now we can show that the difference $Z_t^h = \overline{\pi}_{t-1}(X_t) - \overline{\gamma}_{X\xi}(t)$ satisfies the equation (2.5). Using (2.22) and (2.24), we can write

$$
\begin{aligned}
Z_t^h = m_t + \sum_{l=1}^{t-1} \overline{\gamma}(t,l) \left(\begin{array}{cc} A_l & Q_l \end{array} \right) (\operatorname{var} \overline{\nu}_l)^{-1} \overline{\nu}_l - \sum_{s=1}^{t-1} \overline{\gamma}(t,s) Y_s^2 = \\
= m_t + \sum_{l=1}^{t-1} \frac{A_l \overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} (Y_l - A_l \overline{\pi}_{l-1}(X_l)) + \\
+ \sum_{l=1}^{t-1} \frac{\overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} (Y_l^2 - Q_l(\overline{\pi}_{l-1}(X_l) - h_l)) - \sum_{l=1}^{t-1} \widetilde{\gamma}(t,l) Y_l^2 = \\
= m_t + \sum_{l=1}^{t-1} \frac{A_l \overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} Y_l + \sum_{l=1}^{t-1} \frac{\overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} Q_l h_l - \\
- \sum_{l=1}^{t-1} \overline{\gamma}(t,l) \frac{A_l^2+Q_l}{1+(A_l^2+Q_l)\overline{\gamma}_l} \overline{\pi}_{l-1}(X_l) + \\
+ \sum_{l=1}^{t-1} [\frac{\overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} - \widetilde{\gamma}(t,l)] Y_l^2. \quad (2.27)
\end{aligned}
$$

Now we can rewrite the last term in (2.27) using the equality (2.26). We have

$$
\begin{aligned}
\sum_{l=1}^{t-1} [\frac{\overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} - \widetilde{\gamma}(t,l)] Y_l^2 = \sum_{l=1}^{t-1} \overline{\gamma}(t,l) (\frac{1}{1+(A_l^2+Q_l)\overline{\gamma}_l} - 1) Y_l^2 + \\
+ \sum_{l=1}^{t-1} \sum_{r=l}^{t-1} \overline{\gamma}(t,r) \widetilde{\gamma}(r,l) \frac{A_r^2+Q_r}{1+(A_r^2+Q_r)\overline{\gamma}_r} Y_l^2 = \\
= \sum_{r=1}^{t-1} \overline{\gamma}(t,r) \left[\sum_{l=1}^{r} \widetilde{\gamma}(r,l) Y_l^2 \right] \frac{A_r^2+Q_r}{1+(A_r^2+Q_r)\overline{\gamma}_r} = \\
= \sum_{r=1}^{t-1} \overline{\gamma}(t,r) \overline{\gamma}_{X\xi}(r) \frac{A_r^2+Q_r}{1+(A_r^2+Q_r)\overline{\gamma}_r}, \quad (2.28)
\end{aligned}
$$

with, in the last step, the use of the equality (2.24).

Finally (2.27)-(2.28) imply:

$$
\begin{aligned}
Z_t^h = m_t + \sum_{l=1}^{t-1} \frac{A_l \overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} Y_l + \sum_{l=1}^{t-1} \frac{Q_l \overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} h_l - \\
- \sum_{l=1}^{t-1} \overline{\gamma}(t,l) \frac{A_l^2+Q_l}{1+(A_l^2+Q_l)\overline{\gamma}_l} [\overline{\pi}_{l-1}(X_l) - \overline{\gamma}_{X\xi}(l)] = \\
m_t + \sum_{l=1}^{t-1} \frac{A_l \overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} Y_l + \sum_{l=1}^{t-1} \frac{Q_l \overline{\gamma}(t,l)}{1+(A_l^2+Q_l)\overline{\gamma}_l} h_l - \\
- \sum_{l=1}^{t-1} \overline{\gamma}(t,l) \frac{A_l^2+Q_l}{1+(A_l^2+Q_l)\overline{\gamma}_l} Z_l^h \,,
\end{aligned}
$$

which is nothing else but equation (2.5) with $\mu = -1$.

3 Particular cases and applications

Here we deal with some specific cases where the results can be further elaborated.

3.1 LEG filtering of Gauss-Markov sequences

In this part we concentrate on the case of a Gaussian AR(1) process X, *i.e.*, a Gauss-Markov process driven by

$$
X_t = a_t X_{t-1} + D_t^{\frac{1}{2}} \widetilde{\varepsilon}_t \,,\ t \geq 1 \,; \quad X_0 = x \,, \tag{3.1}
$$

where $(\widetilde{\varepsilon}_t,\ t = 1, 2, \dots)$ is a sequence of i.i.d. standard Gaussian random variables and $(D_t,\, t \geq 1)$ is a (deterministic) sequence of real numbers such that $D_t \geq 0$ for $t \geq 1$. In this setting, it is easy to check that the mean and covariance functions of X are given by

$$
m_t = [\prod_{u=1}^{t} a_u] x = \Lambda_t x \,; \quad K(t,s) = [\prod_{u=s+1}^{t} a_u] k_s = \frac{\Lambda_t}{\Lambda_s} k_s \,,\ 1 \leq s \leq t \,,
$$

where $\Lambda_t = \prod_{u=1}^{t} a_u$ and $k_t = a_t^2 k_{t-1} + D_t$, $t \geq 1$, $k_0 = 0$. Suppose that the following the Riccati type equation

$$
\overline{\gamma}_s = D_s + \frac{a_s^2 \overline{\gamma}_{s-1}}{1+(A_{s-1}^2 - \mu Q_{s-1})\overline{\gamma}_{s-1}},\ s \geq 1,\ \overline{\gamma}_0 = 0 \tag{3.2}
$$

has the unique nonnegative solution.

From the classical filtering theory it is well-known that (for $\mu < 0$) $\overline{\gamma}_s$ is nothing but the variance of the error of the one-step prediction problem of the signal X given by the auxiliary observation $\bar{Y}$ defined by equations (1.1) and (2.9). Then, it is readily seen that the function $\overline{\gamma}(t,s)$, where $\overline{\gamma}(t,s) = \dfrac{\Lambda_t}{\Lambda_s}\overline{\gamma}_s$ is the solution of equation (2.1) and that moreover equation (2.2) for the solution $\overline{h}$ of the LEG filtering problem (2.3) can be reduced to the following one:

$$\overline{h}_t = \frac{a_t}{1+A_t^2\overline{\gamma}_t}\overline{h}_{t-1} + \frac{A_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}Y_t,\ t \geq 1,\ \overline{h}_0 = x, \tag{3.3}$$

or, equivalently:

$$\overline{h}_t = a_t\overline{h}_{t-1} + \frac{A_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}[Y_t - a_tA_t\overline{h}_{t-1}],\ t \geq 1,\ \overline{h}_0 = x.$$

Actually equation (3.3) can also be obtained directly from the general filtering theory (for $\mu = -1$ and replacing Q by $-\mu Q$). For arbitrary $(h_t \in \mathcal{Y}_t, t \geq 1)$ the Note following Theorem 13.1 in [14] gives the equation for Z^h:

$$Z_t^h = a_tZ_{t-1}^h + a_t\overline{\gamma}_t\frac{Q_{t-1}}{1+S_{t-1}\overline{\gamma}_t}[h_{t-1} - Z_{t-1}^h] + \\ + a_t\overline{\gamma}_t\frac{A_{t-1}}{1+S_{t-1}\overline{\gamma}_t}[Y_{t-1} - A_{t-1}Z_{t-1}^h],\ t \geq 1,\ Z_0^h = x.$$

Hence, again the solution $\overline{h}_t = \dfrac{Z_t^{\bar{h}} + A_t\overline{\gamma}_tY_t}{1+A_t^2\overline{\gamma}_t}$, $t \geq 1$, of the LEG filtering problem (2.3) is given by (3.3).

Let us emphasize that these equations are nothing but those given in Speyer *et al.* [16].

It is interesting to note that in the case $a_t = 0$ (i.i.d. signal) the solution of the LEG filtering problem is nothing else but the solution of the risk neutral filtering problem *i.e.* $\overline{h}_t = \pi_t(X_t)$.

3.2 LEG filtering of moving averages of order 1

Here we consider the case of a MA(1) process, *i.e*, a non Markovian process X defined by

$$X_t = \widetilde{\varepsilon}_t + \lambda\widetilde{\varepsilon}_{t-1}\,;t \geq 1\,,$$

where $(\widetilde{\varepsilon}_0, \widetilde{\varepsilon}_1, \dots)$ is a sequence of i.i.d. standard Gaussian variables and λ is a real number. Of course X is centered and has the covariance function

$K(t,s) = 1+\lambda^2$ if $s=t$, λ if $s=t-1$ and 0 if $s<t-1$. In order to solve equation (2.1) we can take

$$\overline{\gamma}(t,s)=0\,,\ s<t-1\,;\quad \overline{\gamma}(t,t-1)=\lambda\,,\ t\geq 1\,,$$

and $\overline{\gamma}(t,t)=\overline{\gamma}_t$ where $\overline{\gamma}_t$ is the solution of the equation:

$$\overline{\gamma}_t = 1+\lambda^2-\lambda\frac{A_{t-1}^2-\mu Q_{t-1}}{1+(A_{t-1}^2-\mu Q_{t-1})\overline{\gamma}_{t-1}}\ ,\ t\geq 1\,;\quad \gamma_0=1+\lambda^2,$$

provided that this equation has the unique nonnegative solution.

Moreover equation (2.2) for the solution $\overline{h}$ of the LEG filtering problem (2.3) can be reduced to the following one:

$$\overline{h}_t = \lambda\frac{A_{t-1}}{1+A_t^2\overline{\gamma}_t}[Y_{t-1}-A_{t-1}\overline{h}_{t-1}]+\frac{A_t\overline{\gamma}_t}{1+A_t^2\overline{\gamma}_t}Y_t,\ t\geq 1,\ \overline{h}_0=x.$$

Again, it is interesting to note that for $\lambda=0$ (i.i.d. signal) the solution of LEG filtering problem is nothing else but the solution of the risk neutral filtering problem *i.e.* $\overline{h}_t=\pi_t(X_t)$.

4 LEG and RS filtering problems

Here, at first we show that actually the LEG and RS filtering problems have the same solution. Then we dive an example which show that in a more general context similar problems may have a different solutions.

4.1 Equivalence of LEG and RS filtering problems

Let $\bar{h}=(\bar{h}_s)_{s\geq 1}$ be the solution of the LEG filtering problem (2.3) given by equation (2.2). For any fixed $t\leq T$, let us denote by $\hat{g}_t$:

$$\hat{g}_t \;=\; \arg\min_{g\in\mathcal{Y}_t}\mathbb{E}\Big[\mu\exp\Big\{\frac{\mu}{2}(X_t-g)^2Q_t\ +\ \frac{\mu}{2}\sum_{s=1}^{t-1}(X_s-\bar{h}(s))^2Q_s\Big\}\Big/\ \mathcal{Y}_t\Big],$$

where $g\in\mathcal{Y}_t$ means that g is a $\mathcal{Y}_t$-measurable variable. It follows directly from Proposition 2 that, provided that $(1+S_t\overline{\gamma}_t)>0$, the equality $\hat{g}_t=\dfrac{Z_t^{\bar{h}}+A_t\overline{\gamma}_tY_t}{1+A_t^2\overline{\gamma}_t}$, $t\geq 1$ holds. Since it was noted in the proof of Theorem 1 that $\overline{h}_t=\dfrac{Z_t^{\bar{h}}+A_t\overline{\gamma}_tY_t}{1+A_t^2\overline{\gamma}_t}$, $t\geq 1$, hence we have also $\hat{g}_t=\bar{h}_t$. It means that for

$t \geq 1$ the solution $\bar{h}$ of the LEG filtering problem satisfies the following recursive equation:

$$\hat{g}_t = \arg\min_{g\in\mathcal{Y}_t}\mathbb{E}\Big[\mu \exp\Big\{\frac{\mu}{2}(X_t - g)^2 Q_t + \frac{\mu}{2}\sum_{s=1}^{t-1}(X_s - \bar{h}(s))^2 Q_s\Big\} \Big/ \mathcal{Y}_t\Big], \quad (4.1)$$

Indeed, in the literature, the recursion (4.1) is the basic **definition** of the so-called risk-sensitive (RS) filtering problem which was introduced in [5]. Therefore we have also proved the following statement

Theorem 4. *Assume that the condition* (C_μ) *is satisfied. Let* $\overline{h} = (\overline{h}_t)_{t\geq 1}$ *be the unique solution of equation (2.2), i.e.,* $\overline{h}$ *is the solution of the LEG filtering problem* (2.3)*. Then* $\overline{h}$ *is the solution of the RS filtering problem* (4.1)*.*

4.2 Discrepancy between LEG and RS type filtering problems: an example

Actually, we did not find in the literature any trace of the discussion about the relationship between the LEG filtering problem (2.3) and the RS filtering problem (4.1) even in a Gauss-Markov case. As a complement to our observation that these two problems have the same solution, we propose an example to show that in a bit more general setting, two similar problems may have different solutions.

For given positive symmetric deterministic 2×2 matrices $\Lambda_s, 1 \leq s \leq T$, let us set $\Phi_t(h) = (X_t\, h_t)\Lambda_t \begin{pmatrix} X_t \\ h_t \end{pmatrix}$. We can define $\bar{h}_t \in \mathcal{Y}_t\, t \geq 1$ as a solution of a *LEG type filtering problem* :

$$\overline{h} = \arg\min_{h_t\in\mathcal{Y}_t,\, t\geq 1} \mathbb{E}\left[\mu \exp\left\{\frac{\mu}{2}\sum_{1}^{T}\Phi_s(h)\right\}\right]. \quad (4.2)$$

We can also define $\hat{h}$ as the solution of the following recursive equation (*RS type filtering problem*):

$$\hat{h}_t = \arg\min_{g\in\mathcal{Y}_t}\mathbb{E}\Big[\mu \exp\Big\{\frac{\mu}{2}\Phi_t(g) + \frac{\mu}{2}\sum_{1}^{t-1}\Phi_s(\hat{h})\Big\} \Big/ \mathcal{Y}_t\Big], \quad (4.3)$$

where $g \in \mathcal{Y}_t$ means that g is a $\mathcal{Y}_t$ measurable variable.

The question which we discuss now is the following: does the equality $\bar{h} = \hat{h}$ hold?

As we have just proved, the answer is positive for singular matrices Λ, namely, when $\Lambda_{11} = \Lambda_{22} = -\Lambda_{12} = Q$. But in the general situation the answer may be negative. Actually it is sufficient to consider the following example: $\Lambda = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$, $A_t = 1$, $\mu = -1$ and $X_t = X_{t-1} + \widetilde{\varepsilon}_t$, where $(\widetilde{\varepsilon}_t,\ t = 1, 2, \dots)$ is a sequence of i.i.d. standard Gaussian random variables. Even in this Markov case $\hat{h} \neq \overline{h}$. More explicitly let us introduce the new probability measure $\hat{\mathbb{P}}$:

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \frac{\exp\left[-\frac{1}{2}\sum_{i=1}^{T} X_i^2\right]}{\mathbb{E}\exp\left[-\frac{1}{2}\sum_{i=1}^{T} X_i^2\right]}.$$

One can check that with respect to $\hat{\mathbb{P}}$ the observation model $(X_t, Y_t)_{t\geq 1}$ can be written in the following form:

$$X_t = a_t X_{t-1} + D_t^{\frac{1}{2}} \hat{\varepsilon}_t\,,\ t \geq 1\,; \quad X_0 = x\,, Y_t = X_t + \varepsilon_t\,,$$

where $(\hat{\varepsilon}_t)_{t\geq 1}$ is a sequence of i.i.d. standard Gaussian random variables independent of the sequence ε,

$$a_t = D_t = \frac{1}{1 + \Gamma(T, t)}\,,$$

and $\Gamma(T, \cdot)$ is the solution of the backward Riccati equation

$$\Gamma(T, t) = 1 + \frac{\Gamma(T, t+1)}{1 + \Gamma(T, t+1)},\ \Gamma(T, T) = 0.$$

So that

$$\Gamma(T, t) = 10\frac{\lambda^T - \lambda^t}{(1-\sqrt{5})\lambda^T - (1+\sqrt{5})\lambda^t},\ \lambda = \frac{(3-\sqrt{5})}{(3+\sqrt{5})}.$$

Indeed to explain this change of the observation model it is sufficient to calculate the conditional characteristic function:

$$\hat{\mathbb{E}}\left[\exp(i\lambda X_t)/\,\mathcal{X}_{t-1}\right] = \frac{\mathbb{E}\left[\exp\left[i\lambda X_t - \frac{1}{2}\sum_{i=1}^{T} X_i^2\right]\Big/\,\mathcal{X}_{t-1}\right]}{\mathbb{E}\left[\exp\left[-\frac{1}{2}\sum_{i=1}^{T} X_i^2\right]\Big/\,\mathcal{X}_{t-1}\right]},$$

where $\mathcal{X}_{t-1}$ is the σ-field $\mathcal{X}_{t-1} = \sigma(\{X_s\,,\ 1 \leq s \leq t-1\})$. But it follows directly from the equation (19)-(20) in [8] and from (2.13) that

$$\hat{\mathbb{E}}\left[\exp(i\lambda X_t)/\,\mathcal{X}_{t-1}\right] = \exp\left\{\frac{i\lambda}{1+\Gamma(T,t)}X_{t-1} - \frac{\lambda^2}{2(1+\Gamma(T,t))}\right\}.$$

Since the density $\frac{d\hat{\mathbb{P}}}{d\mathbb{P}}$ does not depend on h the initial LEG filtering problem (4.2) can be rewritten as:

$$\overline{h} = \arg\min_{h \in \mathcal{Y}} \hat{\mathbb{E}} \left[- \exp \left\{ -\frac{1}{2} \sum_{1}^{T} (X_s - h_s)^2 \right\} \right].$$

Hence we can apply Theorem 1 or in particular (3.2) and (3.3). Clearly, $\overline{h}$ depends on T and $\hat{h}$ does not depend on T by the definition. A bit more explicitly we have for example for $t = 1$: $\overline{h}_1 = \frac{1 + \Gamma(T,1)}{2 + \Gamma(T,1)} Y_1$ and obviously $\hat{h}_1 = \frac{\pi_1(X_1)}{1 + \gamma_1} = \frac{1}{4} Y_1$ and clearly they are different.

References

[1] A. Bensoussan and J.H. van Schuppen, "Optimal control of partially observable stochastic systems with an exponential of integral performance index", SIAM J. Optimization and Control, Vol. 23, pp. 599-613, 1985

[2] I. Collings, M. James and J.B. Moore, "An Information State Approach to Risk-Sensitive Tracking Problems", Journal of Mathematical Systems, Estimation and Control, Vol. 6, No 3, pp. 1-24, 1996

[3] S. Dey and J.B. Moore, "Risk sensitive filtering and smoothing for hidden Markov models", Systems Control Lett., Vol. 25, No 5, pp. 361-366, 1995

[4] S. Dey and J.B. Moore, "Risk-sensitive filtering and smoothing via reference probability methods", IEEE Trans. Automat. Control, Vol. 42, No 11 pp. 1587-1591, 1997

[5] S. Dey, R.J. Elliott and J.B. Moore, "Finite dimensional risk-sensitive estimation for continuous time nonlinear systems", Proceedings of the European Control Conference, Brussels, 1997.

[6] R.J. Elliott, L. Aggoun and J.B. Moore, "Hidden Markov Models: Estimation and Control", Springer, Berlin, 1994

[7] M.L. Kleptsyna and A. Le Breton, "Optimal linear filtering of general multidimensional Gaussian processes - Application to Laplace transforms of quadratic functionals", Journal of Applied Mathematics Stochastic Analysis, Vol. 14, No. 3, pp. 215-226, 2001

[8] M.L. Kleptsyna, A. Le Breton and M. Viot, "New formulas around Laplace transforms of quadratic forms for general Gaussian sequences", Journal of

Applied Mathematics Stochastic Analysis, Vol. 15, No 4, pp. 323-339, 2002

[9] M.L. Kleptsyna and A. Le Breton, “A Cameron-Martin type formula for general Gaussian processes – A filtering approach”, Stochastics and Stochastics Reports, Vol. 72, No 3-4, pp. 229-250, 2002

[10] M.L. Kleptsyna, A. Le Breton and M. Viot, “On the linear-exponential filtering problem for general Gaussian processes”,

[11] M.L.Kleptsyna, A. Le Breton and M. Viot, “Exponential type filtering problems for general Gaussian processes”, Proceedings of 48th IEEE Conference on Decision and Control, Shanghai, pp. 2646-2651, 2009

[12] M.L.Kleptsyna, A.Le Breton, M.Viot, “The Risk Sensitive and LEG filtering problems are not equivalent”, Systems and Control Letters, Vol. 59, pp. 484-490, 2010

[13] R.S. Liptser and A.N. Shiryaev, “Statistics of Random Processes I - General Theory”, Springer-Verlag, New-York, 1977

[14] Liptser, R. S. and Shiryaev, A. N., “Statistics of Random Processes II - Applications”, Springer-Verlag, New-York, 1978

[15] J.B. Moore, R.J. Elliott and S. Dey, “Risk-sensitive generalization of minimum variance estimation and control”, IFAC Symposium on Nonlinear Control Systems Design, pp. 465-470, 1995

[16] J.L. Speyer, C. Fan and R.N. Banavar, “Optimal Stochastic Estimation with Exponential Criteria”, Proceedings of the 31st Conference on Decision and Control, Vol. 2, pp. 2293-2298, 1992

[17] P. Whittle, “0ptimization over time: Dynamic programming and Stochastic Control”, Wiley series in Probability and Mathematical Statistics, 11, John Wiley and Sons, New York, 1982

[18] P. Whittle, “Risk-sensitive optimal control”, John Wiley and Sons, New York, 1990

[19] P. Whittle, “A risk-sensitive maximum principle: the case imperfect state observations”, IEEE Trans. Automat. Control, Vol. 36, No 7, pp. 793-801, 1991