Int. J. Appl. Math. and Comp. Sci., 1999, Vol.9, No.1, 207–217

D-STEP AHEAD KALMAN PREDICTOR FOR CONTROLLED AUTOREGRESSIVE PROCESSES WITH RANDOM COEFFICIENTS

NADINE HILGERT*, JEAN-PIERRE VILA*

This paper deals with prediction of controlled autoregressive processes with additive white Gaussian noise and random coefficients adapted to an observation process. Our aim is twofold. We begin by extending to the *standard Kalman predictor* a result of Chen *et al.* (1989) on the optimality of the *standard Kalman filter* when applied to linear stochastic processes with almost surely finite random coefficients. We then show on an example how some particular nonlinear autoregressive processes can be embedded in these linear processes with random coefficients. Such nonlinear processes can then benefit from this optimal prediction, which is not provided by the usual *extended Kalman predictor*.

Keywords: autoregressive processes, Kalman predictor, Kalman filter, extended Kalman filter.

1. Introduction

Let us consider the following linear system with random coefficients:

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k + D_k w_k & \forall k \geq 1 \\ y_k = C_{k-1} x_k + G_{k-1} v_{k-1} + F_{k-1} w_k & \forall k \geq 1 \end{cases} \tag{1}$$

where:

1. $x_k \in \mathbb{R}^s$ is the unknown state vector which we want to predict with d steps ahead ($d \geq 1$) and $y_k \in \mathbb{R}^p$ ($p \leq s$) is the vector of observations,
2. u_k and v_k are the input vectors which may be chosen by feedback from $(y_0, y_1, \ldots, y_{k-d})$,
3. vectors (w_k) are *i.i.d.* and $w_k \sim \mathcal{N}(0, I)$,
4. (x_1, y_0) and (w_k) are independent and x_1 is conditionally Gaussian given y_0 with conditional mean $\widehat{x}_0^{(1)}$ and conditional covariance $P_0^{(1)}$,

* INRA, Laboratoire de Biométrie, 2 Place Viala, 34060 Montpellier cedex 1, France, e-mail: {hilgert, vila}@ensam.inra.fr.

5. the matrices $A_k, B_k, C_k, D_k, F_k, G_k$ and the vectors u_k and v_k are $\mathcal{F}_{k-d}$-measurable (where $\mathcal{F}_{k-d}$ is the σ-algebra generated by $y_0, y_1, \ldots, y_{k-d}$) for $k \geq d$ (they are $\mathcal{F}_0$-measurable for $k < d$),

6. the entries of $A_k, B_k, u_k, D_k, C_k, v_k$ and F_k are finite with probability one.

This setting is close to that of Chen *et al.* (1989), in which $A_k, B_k, u_k, D_k, C_k, G_k$, v_k and F_k are $\mathcal{F}_k$-measurable (Condition A3). It is shown in that paper that Condition A4 (our Condition 6), which is much weaker than assuming $A_k, B_k, u_k, D_k, C_k, G_k$, v_k and F_k integrable, is sufficient (with Condition A3) for the conditional distribution of x_k with respect to $\mathcal{F}_k$ to be Gaussian and that the standard Kalman filtering algorithm generates the corresponding conditional expectation and conditional covariance. This result is an improvement in comparison with finite-second-moment conditions used by previous authors, e.g. Liptser and Shiryayev (1977), which can hardly be verified in practice e.g. in adaptive control by output feedback.

In Section 2, we show that this approach can easily be extended to one-step and, more generally, to d-step ahead Kalman prediction, under appropriate measurability conditions for the model coefficients and the inputs, leading to equivalent optimal estimation of the conditionally Gaussian distributed state. Moreover, this revisited standard Kalman prediction (SKP) can be applied in a straightforward manner to some nonlinear autoregressive processes which, after reparametrization, obey the linear form (1) while satisfying the appropriate measurability conditions. The same is true for the standard Kalman filtering (SKF). In that case, the well-known approximating extended Kalman procedures (EKP and EKF (Jazwinski, 1970)) can advantageously be replaced by the optimal standard ones, in spite of the fact that these processes are nonlinear. Section 3 is devoted to a comparison between the SKP and the EKP applied to a nonlinear process of this type.

Lemmas 1–3 of (Chen *et al.*, 1989; Chen and Guo, 1991) are briefly recalled in the Appendix without any proof, together with essential definitions.

Remark. In what follows, $(\cdot)'$ will denote matrix transposition and $(\cdot)^+$ will denote matrix pseudo-inversion as defined in the Appendix.

2. d-Step Ahead Standard Kalman Predictor

The minimum variance d-step ahead predictor of x_k is the conditional mean $\widehat{x}_k^{(d)} = \mathrm{E}(x_{k+d} \mid \mathcal{F}_k)$. For $d \geq 2$, from $\mathcal{F}_k \subset \mathcal{F}_{k+d-1}$ and the classical optimal property of the conditional expectation, we have $\mathrm{E}(x_{k+d} \mid \mathcal{F}_k) = \mathrm{E}[\mathrm{E}(x_{k+d} \mid \mathcal{F}_{k+d-1}) \mid \mathcal{F}_k]$ and then

$$\widehat{x}_k^{(d)} = \mathrm{E}\left(\widehat{x}_{k+d-1}^{(1)} \mid \mathcal{F}_k\right) \tag{2}$$

Let us now state the following result:

Theorem 1. *The quantity x_{k+d} is conditionally Gaussian with respect to $\mathcal{F}_k$ with conditional mean $\widehat{x}_k^{(d)}$ and conditional variance $P_k^{(d)} = E[(x_{k+d} - \widehat{x}_k^{(d)})(x_{k+d} - \widehat{x}_k^{(d)})' \mid \mathcal{F}_k]$, given by:*

i) for $d = 1$

$$\widehat{x}_k^{(1)} = A_k \widehat{x}_{k-1}^{(1)} + B_k u_k + K_k \left(y_k - C_{k-1} \widehat{x}_{k-1}^{(1)} - G_{k-1} v_{k-1} \right) \tag{3}$$

$$P_k^{(1)} = A_k P_{k-1}^{(1)} A_k' + D_k D_k' - K_k \left(C_{k-1} P_{k-1}^{(1)} A_k' + F_{k-1} D_k' \right) \tag{4}$$

where

$$K_k = \left(A_k P_{k-1}^{(1)} C_{k-1}' + D_k F_{k-1}' \right) \left(C_{k-1} P_{k-1}^{(1)} C_{k-1}' + F_{k-1} F_{k-1}' \right)^+ \tag{5}$$

ii) for $d \geq 2$

$$\widehat{x}_k^{(d)} = A_{k+d-1} \widehat{x}_k^{(d-1)} + B_{k+d-1} u_{k+d-1} \tag{6}$$

$$P_k^{(d)} = A_{k+d-1} P_k^{(d-1)} A_{k+d-1}' + D_{k+d-1} D_{k+d-1}' \tag{7}$$

Proof. i) Let us first consider the case $d = 1$.

We can recursively show that x_{k+1} and (x_{k+2}, y_{k+1}) are conditionally Gaussian with respect to $\mathcal{F}_k$:

- For $k = 1$:

  According to Assumption 4, x_1 is conditionally Gaussian given y_0. Moreover,

$$\begin{pmatrix} x_2 \\ y_1 \end{pmatrix} = \begin{pmatrix} A_1 & D_1 \\ C_0 & F_0 \end{pmatrix} \begin{pmatrix} x_1 \\ w_1 \end{pmatrix} + \begin{pmatrix} B_1 u_1 \\ G_0 v_0 \end{pmatrix}$$

  Here $\left(\begin{smallmatrix} x_1 \\ w_1 \end{smallmatrix}\right)$ is conditionally Gaussian with respect to $\mathcal{F}_0$ since

$$\begin{aligned} \mathrm{E}\left[\exp\left(i\lambda' x_1 + i\mu' w_1\right) \mid y_0 \right] &= \mathrm{E}\left[\exp(i\lambda' x_1) \mathrm{E}\left(\exp(i\mu' w_1) \mid x_1, y_0 \right) \mid y_0 \right] \\ &= \mathrm{E}\left[\exp(i\mu' w_1) \right] \mathrm{E}\left[\exp(i\lambda' x_1) \mid y_0 \right] \\ &= \exp\left(i\lambda' \widehat{x}_0^{(1)} - \frac{1}{2} \lambda' P_0^{(1)} \lambda - \frac{1}{2} \mu\mu' \right) \quad \text{a.s.} \end{aligned}$$

  Since the entries of $\left(\begin{smallmatrix} A_1 & D_1 \\ C_0 & F_0 \end{smallmatrix}\right)$ and of $\left(\begin{smallmatrix} B_1 u_1 \\ G_0 v_0 \end{smallmatrix}\right)$ are $\mathcal{F}_0$-measurable, Lemma 2 (Appendix) ensures that $\left(\begin{smallmatrix} x_2 \\ y_1 \end{smallmatrix}\right)$ is conditionally Gaussian with respect to $\mathcal{F}_0$, with a.s. finite conditional mean and covariance.

- For $k > 1$:

  Suppose that x_k and $\binom{x_{k+1}}{y_k}$ are conditionally Gaussian with respect to $\mathcal{F}_{k-1}$, with a.s. finite conditional means and covariances. Then by Lemma 3 (Appendix) x_{k+1} is conditionally Gaussian with respect to $\mathcal{F}_k$.

  Now

$$\begin{pmatrix} x_{k+2} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} A_{k+1} & D_{k+1} \\ C_k & F_k \end{pmatrix} \begin{pmatrix} x_{k+1} \\ w_{k+1} \end{pmatrix} + \begin{pmatrix} B_{k+1}u_{k+1} \\ G_k v_k \end{pmatrix}$$

  where A_{k+1}, B_{k+1}, u_{k+1} and D_{k+1} are $\mathcal{F}_k$-measurable. C_k, F_k, v_k and G_k are $\mathcal{F}_{k-1}$-measurable and then $\mathcal{F}_k$-measurable. As was in case $k = 1$, one can show that $\binom{x_{k+1}}{w_{k+1}}$ is conditionally Gaussian with respect to $\mathcal{F}_k$, with a.s. finite conditional mean and covariance. Lemma 2 ensures then that $\binom{x_{k+2}}{y_{k+1}}$ is conditionally Gaussian with respect to $\mathcal{F}_k$. The conditional mean and covariance are finite a.s.

Let us now turn towards the calculation of the conditional mean $\widehat{x}_k^{(1)} = \mathrm{E}(x_{k+1} \mid \mathcal{F}_k)$ and the conditional variance $P_k^{(1)} = \mathrm{E}[(x_{k+1} - \widehat{x}_k^{(1)})(x_{k+1} - \widehat{x}_k^{(1)})' \mid \mathcal{F}_k]$. By the conditional Gaussianity of $\binom{x_{k+1}}{y_k}$ with respect to $\mathcal{F}_{k-1}$ Lemma 3 ensures that x_{k+1} is conditionally Gaussian with respect to $\mathcal{F}_k$ and that

$$\widehat{x}_k^{(1)} = \mathrm{E}\big(x_{k+1} \,\big|\, \mathcal{F}_{k-1}\big) + P_c P_b^{+} \big(y_k - \mathrm{E}(y_k \,|\, \mathcal{F}_{k-1})\big) \quad \text{a.s.}$$

$$P_k^{(1)} = P_a - P_c P_b^{+} P_c' \quad \text{a.s.}$$

where

$$P_a = \mathrm{E}\Big[\big(x_{k+1} - \mathrm{E}(x_{k+1} \,|\, \mathcal{F}_{k-1})\big)\big(x_{k+1} - \mathrm{E}(x_{k+1} \,|\, \mathcal{F}_{k-1})\big)' \,\Big|\, \mathcal{F}_{k-1}\Big]$$

$$P_b = \mathrm{E}\Big[\big(y_k - \mathrm{E}(y_k \,|\, \mathcal{F}_{k-1})\big)\big(y_k - \mathrm{E}(y_k \,|\, \mathcal{F}_{k-1})\big)' \,\Big|\, \mathcal{F}_{k-1}\Big]$$

$$P_c = \mathrm{E}\Big[\big(x_{k+1} - \mathrm{E}(x_{k+1} \,|\, \mathcal{F}_{k-1})\big)\big(y_k - \mathrm{E}(y_k \,|\, \mathcal{F}_{k-1})\big)' \,\Big|\, \mathcal{F}_{k-1}\Big]$$

According to (1), we have

$$\mathrm{E}\big(x_{k+1} \,|\, \mathcal{F}_{k-1}\big) = A_k \mathrm{E}\big(x_k \,|\, \mathcal{F}_{k-1}\big) + B_k u_k$$

$$\mathrm{E}\big(y_k \,|\, \mathcal{F}_{k-1}\big) = C_{k-1} \mathrm{E}\big(x_k \,|\, \mathcal{F}_{k-1}\big) + G_{k-1} v_{k-1}$$

Then

$$P_a = \mathrm{E}\Big[\Big(A_k\big(x_k - \widehat{x}_{k-1}^{(1)}\big) + D_k w_k\Big)\Big(A_k\big(x_k - \widehat{x}_{k-1}^{(1)}\big) + D_k w_k\Big)' \,\Big|\, \mathcal{F}_{k-1}\Big]$$

$$= A_k P_{k-1}^{(1)} A_k' + D_k D_k'$$

since $\mathrm{E}(x_k w_k' \mid \mathcal{F}_{k-1}) = 0$ and $\mathrm{E}(w_k w_k' \mid \mathcal{F}_{k-1}) = \mathrm{E}(w_k w_k') = I$.

$$P_b = \mathrm{E}\Big[\Big(C_{k-1}\big(x_k - \widehat{x}_{k-1}^{(1)}\big) + F_{k-1}w_k\Big)\Big(C_{k-1}\big(x_k - \widehat{x}_{k-1}^{(1)}\big) + F_{k-1}w_k'\Big) \mid \mathcal{F}_{k-1}\Big]$$

$$= C_{k-1}P_{k-1}^{(1)}C_{k-1}' + F_{k-1}F_{k-1}'$$

$$P_c = \mathrm{E}\Big[\Big(A_k\big(x_k - \widehat{x}_{k-1}^{(1)}\big) + D_k w_k\Big)\Big(C_{k-1}\big(x_k - \widehat{x}_{k-1}^{(1)}\big) + F_{k-1}w_k\Big)' \mid \mathcal{F}_{k-1}\Big]$$

$$= A_k P_{k-1}^{(1)} C_{k-1}' + D_k F_{k-1}'$$

Finally,

$$\widehat{x}_k^{(1)} = A_k\widehat{x}_{k-1}^{(1)} + B_k u_k + K_k\Big(y_k - C_{k-1}\widehat{x}_{k-1}^{(1)} - G_{k-1}v_{k-1}\Big)$$

$$P_k^{(1)} = A_k P_{k-1}^{(1)} A_k' + D_k D_k' - K_k\Big(A_k P_{k-1}^{(1)} C_{k-1}' + D_k F_{k-1}'\Big)'$$

$$K_k = \Big(A_k P_{k-1}^{(1)} C_{k-1}' + D_k F_{k-1}'\Big)\Big(C_{k-1}P_{k-1}^{(1)}C_{k-1}' + F_{k-1}F_{k-1}'\Big)^{+}$$

ii) The case $d \geq 2$ is a straightforward generalization of part *(i)*. ■

3. Application

Some nonlinear dynamical systems of high practical value can easily be embedded in the general linear system (1) with random coefficients, while satisfying the related set of assumptions. The optimal standard Kalman prediction (SKP) can then be applied to these nonlinear systems, instead of the approximative extended Kalman prediction (EKP) procedure which relies on an implicit linearization of the models. Obviously, the same remark is true in the filtering context of Chen *et al.* (1989), for the SKF with respect to the EKF.

We now present an application from the biotechnological field, which is well-adapted to the one-step ahead prediction context. To strengthen the relevance of the SKP method, we shall then adjust the example to the two-step ahead prediction.

3.1. One-Step Ahead Prediction

Let us consider the following two-dimensional nonlinear autoregressive process (x_k), whose second component $(x_k^{(S)})$ is exactly known through the observation variable y_k:

$$\begin{cases} x_{k+1}^{(B)} = \big(1 + T(\mu_k - u_k)\big)x_k^{(B)} + \xi_k^{(B)} \\ x_{k+1}^{(S)} = x_k^{(S)} - T\mu_k x_k^{(B)}/\tau + Tu_k\big(S_0 - x_k^{(S)}\big) + \xi_k^{(S)} \\ y_k = x_k^{(S)} \end{cases} \tag{8}$$

This system is characteristic of a basic biotechnological reaction. It describes a microbial growth in a stirred tank reactor, in the case of one population of microorganisms on a single limiting substrate, in continuous operating mode (Bastin and Dochain, 1990).

The state variables $x_k^{(B)}$ and $x_k^{(S)}$ are the biomass and substrate concentrations, respectively. Here u_k, the dilution rate, is the control variable. S_0 is the substrate concentration in the influent, τ the yield coefficient of the substrate consumption by the biomass, and T the sampling period. S_0, τ and T are known constants. ξ_k is a white Gaussian noise.

The quantity μ_k represents the microbial growth rate which is a nonlinear function of the substrate concentration. It is supposed to correspond to the Sokol-Howell law (1981), adapted to Condition 5 of Section 1 with $d = 1$:

$$\mu_k = \alpha \frac{x_{k-1}^{(S)}}{\gamma + \left(x_{k-1}^{(S)}\right)^{\beta}}$$

where α, β and γ are appropriate constants. From (8) it is obvious that the nonlinearity of the system depends on the variations of μ_k. In a simulation perspective, this nonlinearity can then be stressed by acting on the coefficients γ and β, in an adapted way with regard to the domain of variation of $x^{(S)}$.

The control u_k is assumed to be defined by feedback from the past observations $y_0, \ldots, y_{k-1}$.

Let us now rewrite model (8) in order to reduce it to the linear structure of system (1):

$$\begin{cases} x_{k+1} = \begin{pmatrix} x_{k+1}^{(B)} \\ x_{k+1}^{(S)} \end{pmatrix} = \begin{pmatrix} 1 + T(\mu_k - u_k) & 0 \\ -T\mu_k/\tau & 1 - Tu_k \end{pmatrix} x_k + \begin{pmatrix} 0 \\ TS_0 \end{pmatrix} u_k + D_k w_k \\ y_k = x_k^{(S)} \end{cases} \tag{9}$$

Here ξ_k is replaced by $D_k w_k$, where w_k is a two-dimensional $\mathcal{N}(0, I)$ noise and D_k an appropriate scaling matrix. Under some good initial conditions, Theorem 1 can be applied with $d = 1$, and then it gives the standard one-step ahead Kalman prediction of x: $\widehat{x}_{(\mathrm{SKP}),k} = E(x_k \mid y_0, \ldots, y_{k-1})$.

On the other hand, applying the extended Kalman predictor to system (8) requires addition of a new state variable to the system, in order to perform linearizations: $z_{k+1} = x_k^{(S)}$. This procedure is sufficiently well-known and does not deserve a further description (Goodwin and Sin, 1984). Let $\widehat{x}_{(\mathrm{EKP}),k}$ be the related prediction, given the past values $y_0, \ldots, y_{k-1}$.

We now present some simulation results in order to compare the one-step ahead SKP and the EKP. Here γ is fixed to a constant value and β remains the only parameter to characterize the system nonlinearity. Decreasing the values of β increases this nonlinearity. Many simulations trials were carried out for the same conditions.

For each value of β, three quantities are computed. The first two ones are measures of efficiency of both methods, SKP and EKP, in predicting $x^{(B)}$. These are the mean-squared errors of prediction:

$$\mathrm{MS}_{\mathrm{SKP}} = \frac{1}{n-1}\sum_{k=1}^{n}\left(x_k^{(B)} - \widehat{x}^{(B)}_{(\mathrm{SKP}),k}\right)^2, \qquad \mathrm{MS}_{\mathrm{EKP}} = \frac{1}{n-1}\sum_{k=1}^{n}\left(x_k^{(B)} - \widehat{x}^{(B)}_{(\mathrm{EKP}),k}\right)^2$$

where n is the simulated number of steps, each of duration T.

The third test compares the relative efficiency of the SKP with that of the EKP:

$$\Delta = \frac{\mathrm{MS}_{\mathrm{EKP}} - \mathrm{MS}_{\mathrm{SKP}}}{\mathrm{MS}_{\mathrm{SKP}}} \times 100$$

Model (8) was simulated with $S_0 = 50\,\mathrm{mg{\cdot}l^{-1}}$, $\tau = 1$, $T = 1\,\mathrm{h}$, $\alpha = 0.5$, $\gamma = 125$, $x_1^{(B)} = 25\,\mathrm{mg{\cdot}l^{-1}}$, $x_1^{(S)} = 30\,\mathrm{mg{\cdot}l^{-1}}$, $D_k = \left(\begin{smallmatrix} 0.3 & 0.1 \\ 0.1 & 0.1 \end{smallmatrix}\right)$ and the feedback control law $u_k = 0.05/(y_{k-1} + 1)$, which is a relatively good stabilizer and regulator of both concentrations of biomass $x^{(B)}$ and substrate $x^{(S)}$. Here n is equal to 150, an observation window which approximately encloses the stabilization of $x^{(B)}$ and $x^{(S)}$. Both SKP and EKP predictors are initialized with a reasonable bias: $\widehat{x}_1^{(B)} = 10\,\mathrm{mg{\cdot}l^{-1}}$ and $\widehat{x}_1^{(S)} = 45\,\mathrm{mg{\cdot}l^{-1}}$. The initial conditional variances are

$$P_1^{(\mathrm{SKP})} = \begin{pmatrix} 80 & 500 \\ 500 & 80 \end{pmatrix}, \qquad P_1^{(\mathrm{EKP})} = \begin{pmatrix} 80 & 500 & 20 \\ 500 & 80 & 50 \\ 20 & 50 & 100 \end{pmatrix}$$

The simulations trials revealed an increasing superiority of the SKP when the nonlinearity of the model increases. Table 1 reports an extract of about ten different trials with respective values of $\mathrm{MS}_{\mathrm{SKP}}$, $\mathrm{MS}_{\mathrm{EKP}}$ and Δ. The trajectories of $x_k^{(B)}$, $\widehat{x}^{(B)}_{(\mathrm{SKP}),k}$ and $\widehat{x}^{(B)}_{(\mathrm{EKP}),k}$ are given in Fig. 1, for the intermediate β value equal to 0.5.

Table 1.

β	$\mathrm{MS}_{\mathrm{SKP}}$	$\mathrm{MS}_{\mathrm{EKP}}$	Δ
$\beta = 2$	65.38	66.22	1.28 %
$\beta = 1$	70.52	79.03	12.06 %
$\beta = 0.8$	72.82	83.81	15.08 %
$\beta = 0.5$	74.58	87.2	16.93 %
$\beta = 0.1$	75.37	88.59	17.55 %
$\beta = 0.05$	75.41	88.66	17.57 %
$\beta = 0.01$	75.44	88.71	17.59 %
$\beta = 0.005$	75.44	88.72	17.59 %
$\beta = 0.001$	75.45	88.72	17.59 %
$\beta = 0$	75.45	88.72	17.59 %

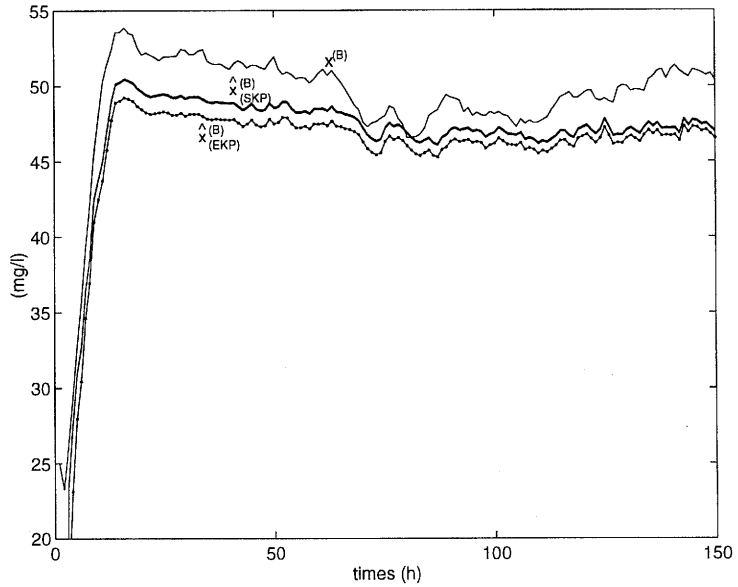

Fig. 1. One-step ahead predictions of $x^{(B)}$ with SKP and EKP ($\beta = 0.5$).

According to the simulation results, the SKP proved to be uniformly significantly better than the EKP. Moreover, it is important to point out the simplicity of the SKP computations with respect to those of the EKP.

3.2. Two-Step Ahead Prediction

To present a simulation example adapted to a two-step ahead prediction, we consider the system (8) with a slight modification: the microbial growth rate μ_k is now supposed to be a function of $x_{k-2}^{(S)}$ (instead of $x_{k-1}^{(S)}$) in order to verify Assumption 5 of the Introduction,

$$\mu_k = \alpha \frac{x_{k-2}^{(S)}}{\gamma + \left(x_{k-2}^{(S)}\right)^{\beta}}$$

with α, β and γ determined as in the last subsection. This modification is purely formal and performed for the sake of illustration. It does not pretend to an experimental reality.

The control variable is defined by feedback from the past observations $y_0, \ldots, y_{k-2}$. Under some good initial conditions, Theorem 1 can be applied with $d = 2$, which gives the standard two-step ahead Kalman prediction of x: $\widehat{x}_{(\mathrm{SKP2}),k} = E(x_k \mid y_0, \ldots, y_{k-2})$.

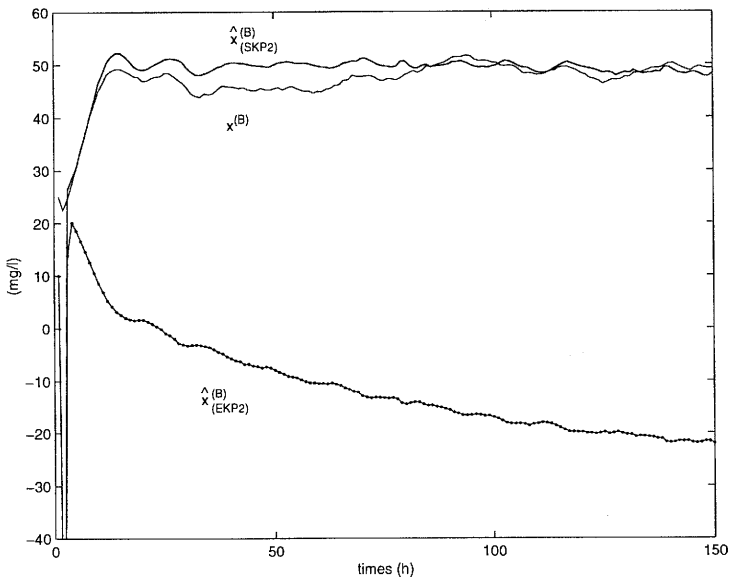

Fig. 2. Two-step ahead predictions of $x^{(B)}$ with SKP and EKP ($\beta = 0.5$).

This two-step ahead SKP predictor of x, $\widehat{x}_{(\mathrm{SKP2}),k}$, was compared by many simulation trials to the two-step ahead EKP predictor $\widehat{x}_{(\mathrm{EKP2}),k}$, following the same procedure as previously for the one-step ahead EKP. The SKP method proved to be more appropriate than the EKP one, as illustrated by Fig. 2. In this example, the simulation conditions are the same as those used for the experiment described by Fig. 1, except for the control law $u_k = 0.05/(y_{k-2} + 1)$.

4. Conclusion

We have extended to d-step ahead Kalman prediction the optimal result obtained by Chen *et al.* (1989) for the Kalman filtering of an autoregressive process with random coefficients, under almost sure finiteness assumptions and measurability assumptions for these coefficients. Both of the results allow for the use of the standard Kalman procedures in some nonlinear context without requiring any linearization and with optimal performances and reduced computations.

Appendix

On the Conditionally Gaussian Situation

One can consult classic probability sources as (Neveu, 1975) in addition to (Chen *et al.*, 1989; Chen and Guo, 1991). The proofs of the lemmas can be found in (Chen and Guo, 1991).

Definition 1. Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space, $\mathcal{G}$ a sub-σ-algebra of $\mathcal{A}$ and x a not necessarily integrable random variable. If either $\mathrm{E}(\max(x,0)\,|\,\mathcal{G}) < \infty$ a.s. or $\mathrm{E}(\max(-x,0)\,|\,\mathcal{G}) < \infty$ a.s., then one can define the conditional expectation of x with respect to $\mathcal{G}$ as

$$\mathrm{E}(x\,|\,\mathcal{G}) = \mathrm{E}\big(\max(x,0)\,|\,\mathcal{G}\big) - \mathrm{E}\big(\max(-x,0)\,|\,\mathcal{G}\big) \quad \text{a.s.}$$

Definition 2. Let $\sigma(Y)$ be the σ-algebra generated by the random vector Y. Let X be a random vector such that $|X| < \infty$ a.s. X is said to be conditionally Gaussian with respect to $\sigma(Y)$ if there exists a $\sigma(Y)$-measurable random vector $\widehat{X}$ and a $\sigma(Y)$-measurable random symmetric semi positive-definite matrix P such that

$$\mathrm{E}\big[\exp(i\lambda' X)\,|\,\sigma(Y)\big] = \exp\Big(i\lambda'\widehat{X} - \frac{1}{2}\lambda' P\lambda\Big) \quad \text{a.s.}$$

for every constant vector λ.

Lemma 1. *If X is conditionally Gaussian with respect to $\sigma(Y)$, then $\widehat{X} = E[X\,|\,\sigma(Y)]$ a.s. and $P = E[(X-\widehat{X})(X-\widehat{X})'\,|\,\sigma(Y)]$ a.s.*

Lemma 2.

(i) If X is conditionally Gaussian with respect to $\sigma(Z)$ and $A(\cdot)$ and $b(\cdot)$ are measurable functions with $\|A(Z)\| < \infty$ a.s. and $\|b(Z)\| < \infty$ a.s., then $A(Z)X + B(Z)$ is conditionally Gaussian with respect to $\sigma(Z)$.

(ii) If $\binom{X}{Y}$ is conditionally Gaussian with respect to $\sigma(Z)$, then X and Y are conditionally independent with respect to $\sigma(Z)$ if and only if

$$E\Big[\Big(X - E\big(X\,|\,\sigma(Z)\big)\Big)\Big(Y - E\big(Y\,|\,\sigma(Z)\big)\Big)'\,|\,\sigma(Z)\Big] = 0 \quad \text{a.s.}$$

Lemma 3. *If $\binom{X}{Y}$ is conditionally Gaussian with respect to $\sigma(Z)$ with conditional variance*

$$\begin{pmatrix} P_{xx\,|\,z} & P_{xy\,|\,z} \\ P_{yx\,|\,z} & P_{yy\,|\,z} \end{pmatrix} \quad a.s.$$

then X is conditionally Gaussian with respect to $\sigma(Z,Y)$ with conditional mean

$$E[X\,|\,\sigma(Z,Y)] = E[X\,|\,\sigma(Z)] + P_{xy\,|\,z}P^{+}_{yy\,|\,z}\big(Y - E(Y\,|\,Z)\big) \quad a.s.$$

and conditional variance $P_{xx\,|\,zy} = P_{xx\,|\,z} - P_{xy\,|\,z}P^{+}_{yy\,|\,z}P_{yx\,|\,z}$. Here P^{+} is the pseudo-inverse of P uniquely defined by $P^{+} = V'(VV')^{-1}(U'U)^{-1}U'$ whatever U and V such that $P = UV$, where U and V are full-rank $p \times r$ and $r \times p$ matrices, respectively, with $p = dim(Y)$ and $r = rank(P) \le p$.

References

Bastin G. and Dochain D. (1990): *On-Line Estimation and Adaptive Control of Bioreactors.* — Amsterdam: Elsevier.

Chen H.F., Kumar P.R. and Van Schuppen J.H. (1989): *On Kalman filtering for conditionally Gaussian systems with random matrices.* — Syst. Contr. Lett., Vol.13, pp.397–404.

Chen H.F. and Guo L. (1991): *Identification and Stochastic Adaptive Control.* — Basel: Birkhäuser.

Goodwin G.C. and Sin K.S. (1984): *Adaptive Filtering, Prediction and Control.* — Englewood Cliffs: Prentice Hall.

Jazwinski A.H. (1970): *Stochastic Processes and Filtering Theory.* — San Diego: Academic Press.

Liptser R.S. and Shiryayev A.N. (1977): *Statistics of Random Processes, II.* — New-York: Springer.

Neveu J. (1975): *Discrete Parameter Martingales.* — Amsterdam: North-Holland.

Sokol W. and Howell J.A. (1981): *Kinetics of phenol oxidation by washed cells.* — Biotechnol. Bioeng., Vol.23, pp.2039–2049.