OPTIMAL JOINT ADAPTIVE ESTIMATION
OF PARAMETER AND STATE OF A LINEAR
STOCHASTIC SYSTEM WITH APPLICATION TO TRACKING

G. SALUT J. AGUILAR G. FAVIER G. ALENGRIN

An extension is described in this paper of the optimal joint estimation of - parametres and state in stochastic linear systems, to the single input-sin-- gle output situation, which is the kind of system that appears in the trac- king of the position of an unknown flying object in one direction if interac tion between coordinates is neglected. An application to radar tracking fol- lows, with special emphasis on the description of the discrete filtering al- gorithm that processes the data given by a radar. Some results of the compu- ter program with simulated and with real data, are disclosed.

1. INTRODUCTION

The present paper starts with an abridged exposition of the joint optimal estimation of parameters and state in stochastic linear -- systems; an extensive reading of these to-- pics can be found in /5 and 6/. The particular description given here concerns the single input single output situation. This is - the class of systems that appear in the tracking of the position of an unknown flying object in one direction if interaction between coordinates is neglected.

In a second part the application to radar -- tracking is described with special emphasis for the description of the discrete filtering algorithm that processes the data given by a radar. The computer program is previously -- tested with simulated data, and finally some results with real data are given.

2. JOINT OPTIMAL ESTIMATION OF STATE AND PARAMETERS OF LINEAR DYNAMICAL SYSTEMS

2.1 Canonical form

Let us consider a linear system Σ given by - its transfer function T(s)

$$T(s) = \frac{N(s)}{D(s)}$$

with

$$N(s) = \sum_{i=0}^{n} b_i \, s^i$$

and

$$D(s) = \sum_{i=0}^{n} a_i \, s^i$$

This system is given by the block diagram -- shown in fig. 1, where $\mathcal{D}^{-1}$ is the integration operator.

In state space representation Σ is given by equations (1)

$$\frac{dx}{dt} = \begin{bmatrix} 0 & \dots & 0 & -\frac{a_0}{a_n} \\ 1 & \dots & 0 & -\frac{a_1}{a_n} \\ & \ddots & & \\ 0 & \dots & 1 & -\frac{a_{n-1}}{a_n} \end{bmatrix} x + \begin{bmatrix} b_0 & -a_0 & \frac{b_n}{a_n} \\ & & \\ b_{n-1} & -a_{n-1} & \frac{b_n}{a_n} \end{bmatrix} u \qquad (1)$$

$$y = \begin{bmatrix} 0 & \dots & 0 & \frac{1}{a_n} \end{bmatrix} x + \begin{bmatrix} 0 & \dots & 0 & \frac{b_n}{a_n} \end{bmatrix} u$$

that can be written as

$$\frac{dx}{dt} = A_* \, x + B_* \, u$$

$$y = A_0 \, x + B_0 \, u$$

- G. Salut i J. Aguilar Martin del "Laboratoire d'Automatique et d'Analyse des Systèmes du CNRS". 7, Avenue du Colonel Roche - 31400 Toulouse - France.
- G. Alengrin del "Laboratoire de Signaux et Systèmes, Université de Nice". 41, Boulevard Napoleon III - 06041 Nice Cedex - France.
- Article rebut l'Octubre del 1979.

Those equations are obviously equivalent to equations (2) in the case where $a_n=1$.

$$\dot{x} = \begin{bmatrix} 0\dots & 0 \\ 1\dots & 0 \\ \vdots & \\ 0\dots 1 & 0 \end{bmatrix} x - \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} y + \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{bmatrix} u \qquad (2)$$

$$0 = [0\dots 0 \quad 1] x - [1] y + [b_n] u$$

The essential quality of this canonical form (2) is the separation between the state and the parameters.

Let us define the "parametric" vector that conveys all the internal parametric information

$$r^T = [a_0 \dots a_{n-1},\ b_0 \dots b_{n-1},\ b_n]$$

and the "external information" vector

$$z^T = [y^T \mid u^T]$$

The instantaneous relation between those two vectors can be written by equations (3)

$$\frac{dx}{dt} = E_* x + G_*(z) \cdot r$$

$$[1 \quad 0] z = E_0 x + G_0(z) \cdot r \qquad (3)$$

with

$$E_* = \begin{bmatrix} 0 & \dots\dots & 0 \\ 1 & \dots\dots & 0 \\ & \vdots & \\ 0 & \dots 1 & 0 \end{bmatrix} \qquad E_0 = [0 \dots 0 \quad 1] \qquad (4)$$

$$G_*(z) = [-y \cdot I_n \mid u \cdot I \mid 0] \quad \text{dimensión } n\ (2n+1)$$

$$G_0(z) = [0 \dots 0 \quad u] \quad \text{dimensión } 1\ (2n+1)$$

The state x appears as the dynamical link between z and r whereas matrices E_x and E_0, that are the only to multiply the state, give structural information but not parametric information, as they are made uniquely of '0' and '1' 's.

2.2 Identification problem

Let us look at the parametric identification problem and let us gather all unknown variables into a unique vector

$$s^T(t) = [x^T(t) \mid r^T(t)]$$

Vector s(t) is the state of a non linear system equivalent to Σ.

The adaptive identification problem can then be stated as to find a system $\hat{\Sigma}$ such that its state $\hat{s}(t)$ is as close as possible to s(t). The error between Σ and $\hat{\Sigma}$ will be found in 4 different places:

- ε_d = error in the dynamical equation

$$\varepsilon_d(t) = \frac{d\hat{x}}{dt} - (E_* \hat{x}(t) + G_*(z)\ \hat{r}(t)) \qquad (5)$$

- ε_m = error in the measurement equation

$$\varepsilon_m(t) = y(t) - (E_0 \hat{x}(t) + G_0(z)\ \hat{r}(t)) \qquad (6)$$

- ε_p = error in the variation of the parameters

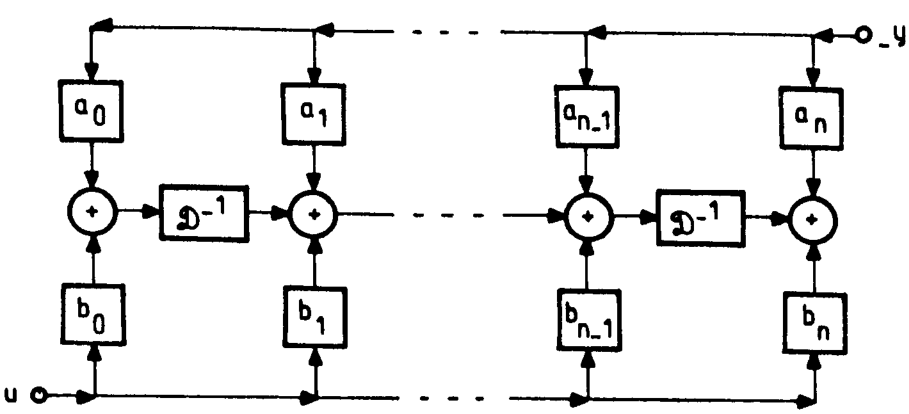

Fig. 1
Block diagram of single input single output system

$$\varepsilon_p(t) = \frac{d\hat{r}}{dt} \tag{7}$$

- $s(t_0) - \hat{s}(t_0)$ = error in the initialisation

composed of $x(t_0)-\hat{x}(t_0)$ for the initial state and of $r(t_0)-\hat{r}(t_0)$ for the a priori values of the parameters.

A quadratic penalty can be imposed as follows

$$J(t_0,t) = (s(t_0) - \hat{s}(t_0))^T \cdot Q_0(s(t_u) - \hat{s}(t_0)) + \int_{t_0}^{t} (\varepsilon_d^T Q_d \varepsilon_d + \varepsilon_m^T Q_m \varepsilon_m + \varepsilon_p^T Q_p \varepsilon_p)\, dt \tag{8}$$

where Q_0, Q_d, Q_m and Q_p are assumed non singular weighting matrices given a priori.

THEOREM 1: The trajectory of the optimal adaptive estimator of s(t), conditionnally to the past observation of the external information vector z(t) with respect of the cost function $J(t_0,t)$ is given by the following filtering equations (9)

$$\frac{d\hat{s}}{dt} = F(z)\ \hat{s} + P\ H^T(z)\ Q_m(y-H(z)\hat{s})$$

$$\frac{dP}{dt} = F(z)\ P + P\ F^T(z) - P\ H^T(z)\ Q_m\ H(z)\ P + Q_d \tag{9}$$

with initial conditions

$\hat{s}(t_0)$ and $P(t_0)=Q_0^{-1}$

and where F(z), H(z) and Q are defined with respect to equations (3) and (8) as follows

$$F(z) = \begin{bmatrix} E_* & G_*(z) \\ 0 & I \end{bmatrix} \qquad H(z) = [0 \ldots 0\ \ 1 \mid G_0(z)] = [E_0 \mid G_0(z)]$$

$$Q = \begin{bmatrix} Q_d^{-1} & 0 \\ 0 & Q_p^{-1} \end{bmatrix}$$

Proof: Let us replace the functions of z by its actual realizations as functions of time for a given past realization of the observed variable z(t).

F(z(t)) becomes F(t)

H(z(t)) becomes H(t)

Let us define w(t) as a gaussian white noise such that

$$E\ w(t)\ w^T(\tau) = \delta(t-\tau)\ Q_d$$

and similarly v(t) such that

$$E\ v(t)\ v^T(\tau) = \delta(t-\tau)\ Q_m^{-1}$$

and let us call σ_0 a gaussian random variable such that

$$E\ \sigma_0 = s(t_0)$$

and

$$E(\sigma_0 - \hat{s}(t_0))\ (\sigma_0 - \hat{s}(t_0))^T = Q_0^{-1}$$

In that situation equations (9) are the optimal linear filtering equations for the estimation of the state σ(t) of the time varying linear system of equation (10)

$$\frac{d\sigma}{dt} = F(t)\ \sigma + w \qquad \sigma(t_0) = \sigma_0 \qquad y = H(t)\ \sigma + v \tag{10}$$

It is therefore well known that this gives the optimal trajectory with respect to the cost function stated in (8) where s(t) will be replaced by σ(t). At this stage we must only notice that system (10) is identical to system (9) for all past values of t.

2.3 Identifiability

Similarly to the condition of observability for the stability of linear filters, we may give here the general identifiability condition.

THEOREM 2: /4/. A completely observable linear system Σ is completely identifiable in the time interval

$[t_0\ t_f]$

if and only if there exists

$t \in [t_0\ t_f]$

such that the matrix $I(t_0,t)$ is definite positive:

$$I(t_0,t) = \int_{t_0}^{t} \Psi^T(t,\tau;z\overline{[\tau,t]}) H^T(z(t)) H(z(t)) \Psi(t,\tau;z\overline{[\tau,t]})\, d\tau \qquad (11)$$

where $z[\tau,t]$ stands for $\{z(\sigma) \mid \sigma \in [\tau,t]$ and Ψ is given by the matrix equation

$$\frac{\partial\, \Psi(t,\ )}{\partial t} = F(z(t))\, \Psi(t,\ ) \qquad (12)$$

$$\Psi(\tau,\tau) = I$$

<u>Proof</u>:

- Sufficiency: if condition (11) holds we have

$$S(t) = I^{-1}(t_0\ t) \int_{t_0}^{t} \Psi^T(t,\tau;z\overline{[\tau,t]}) H^T(z(\tau))\, y(\tau)\, d \qquad (13)$$

and therefore

$$y(\tau) = H(z(\tau))\, s(\tau) =$$

$$= H(z(\tau))\, \Psi(t,\tau;z\overline{[\tau,t]})\, s(t)$$

If we premultiply the second equation by Ψ and integrate it from t_0 to t we get

$$\int_{t_0}^{t} \Psi^T(t,\tau;z\overline{[\tau,t]})\, H^T(z(\tau))\, y(\tau)\, d\tau =$$

$$= I(t_0\ t)\, s(t) \qquad (14)$$

And if $I(t_0\ t) \geq 0$ it is inversible and s(t) can be reconstructed, and the system identified.

- Necessity: let us assume now that $I(t_0\ t)$ is never definite positive in the interval $[t_0\ t]$.

That means that there exist $v(t) \neq 0$ such that

$$\forall t \in [t_0\ t]\ v^T(t)\, I(t_0\ t)\, v(t) \equiv 0$$

or equivalently that

$$\int_{t_0}^{t} \| H(z(t))\, \Psi(t,\tau;z\overline{[\tau,t]})\, v(t) \|^2 \equiv 0 \quad (15)$$

We may then replace s(t) by v(t) in equations (13) and we get

$$y_v(t) = H(z(t))\, \Psi(t,\tau;z\overline{[\tau,t]})\, v(t) = 0$$

for

$$\forall \tau \in [t_0,t] \qquad \text{and} \qquad \forall t \in [t_0\ t_f] \quad (16)$$

That contradicts the observability hypothesis for Σ as (16) means that two trajectories $s(t) \equiv 0$ and $v(t) \neq 0$ for $\tau \in [t_0\ t_f]$ give the same null output.

The result of this theorem is even more powerful, as it has been proved in /5/ that only completely observable linear systems with known structure can be completely identifiable.

2.4 Identification algorithm for discrete linear stochastic gaussian systems

We shall give here the discrete version of theorem 1 for stochastic gaussian systems, because it gives straightforward access to the computer program.

By analogy we will consider the following model:

$$x(t+1) = E_* x(t) - a_* y(t) + b_* u(t) + w_*(t)$$

$$0 = E_0 x(t) - 1.\ y(t) + b_n u(t) + w_n(t) \qquad (17)$$

with:

$$r(t+1) = r(t) + \pi(t) \quad , \quad (r = [a_*^T\ b_*^T\ b_n]^T)$$

It will be assumed that $\pi(.)$, $w_*(.)$, $w_n(.)$ are white gaussian noises characterized by their joint covariance, with strictly positive variance for w_n.

Note that, if the initial probability distribution $P(x_{t_0})$ of the state x is assumed to be gaussian, the output process $y(t)$, $t \geq t_0$, can be considered as a compound gaussian process (with respect to r). If, besides, the initial distribution $P(r_{t_0})$ of parameter r is also gaussian, the $y(.)$ process can be viewed as a doubly gaussian compound process.

Defining again matrices G_* and G_0 as in (3) we get

$$x(t+1) = E_* x(t) + G_*(z_t) r(t) + w_*(t)$$

$$r(t+1) = I r(t) + \pi(t) \qquad (18)$$

$$y(t) = E_0 x(t) + G_0(u_t) r(t) + w_n(t)$$

Setting

$$s = [x^T \ r^T]^T$$

as a generalized state of the system, as well as

$$\xi = [w_*^T \ \pi^T]^T$$

one may write the global representation:

$$s(t+1) = F(z_t) s(t) + \xi(t)$$

$$y(t) = H(u_t) s(t) + \chi(t) \qquad (19)$$

$$\chi(t) = w_n(t)$$

with

$$F(z_t) = \begin{bmatrix} E_* & G_*(z_t) \\ 0 & I \end{bmatrix}$$

$$H(u_t) = [E_0 \quad G_0(u_t)]$$

(.), (.) are white gaussian noises described by their covariances:

$$E[\xi(t) \ \xi^T(\tau)] = Q(t) \cdot \xi_{t,\tau}$$

$$E[\xi(t) \ \chi^T(\tau)] = S(t) \cdot \delta_{t,\tau}$$

$$E[\chi(t) \ \chi^T(\tau)] = R(t) \cdot \delta_{t,\tau} \ (\text{with } R(t)>0, \forall t)$$

From this model we may derive an optimal adaptive estimation for the following optimal joint state and parameter estimation problem:

PROBLEM

Let $Z_{t_0}^t$ be the minimum σ algebra induced by $y(\tau)$, $u(\tau+1)$ on a compact time-interval $t_0 \leq \tau \leq t$. Find among all $Z_{t_0}^t$ measurable functionals, the optimal one $(\hat{s}(t+1/t))$ such that the mean squared error:

$$E[\lambda^* \tilde{s}(t+1/t)^2]$$

is minimized for all covectors λ^* with

$$\tilde{s}(t+1/t) = s(t+1) - \hat{s}(t+1/t)$$

As it is well known, the solution to that problem is given by the conditional expectation

$$E[s(t+1)/Z_{t_0}^t].$$

THEOREM 3: Let the initial probability density $p(s_{t0})$ of the generalized state s be gaussian with mean $\hat{s}_0$ and covariance P_0. The conditional probability density $P(s(t+1)/Z_{t0}^t)$ remains gaussian and its mean $\hat{s}_{t+1/t}$ and covariance $P_{t+1/t}$ are given by the following set of recursive equations:

$$\hat{s}_{t+1/t} = F(z_t) \hat{s}_{t/t-1} + K_t[y_t - H(u_t)\hat{s}_{t/t-1}],$$

$$\hat{s}_{t0/t0} = \hat{s}_0 \qquad (20)$$

$$P_{t+1/t} = [F(z_t)P_{t/t-1} F^T(z_t) + Q_t] - K_t T_{t/t-1} K_t^T, \ P_{t0/t0} = P_0$$

with

$$K_t = [F(z_t) P_{t/t-1} H^T(u_t) + S_t] T_{t/t-1}^{-1}$$

$$T_{t/t-1} = [H(u_t) P_{t/t-1} H^T(u_t) + R_t]$$

Proof:

It suffices to apply the optimal linear filter to the time varying system (19) as it was done for the continuous case (10).

3. APPLICATION TO RADAR TRACKING

3.1 Modelisation

We shall now apply the theoretical results - given in section 1 to the particular case of a flying object characterized for a single - coordinate by its position X_1 and its speed X_2 and controlled by an unknown force or acceleration X_3. The basic equation is then

$$\begin{bmatrix} \frac{dX_1}{dt} \\ \frac{dX_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} X_3$$

the control variable X_3 is considered as a - random variable whose correlation function - is

$$\rho(\tau) = q\, e^{-\alpha|\tau|}$$

where q represents its dispersion and $1/\alpha$ -- the mean time between manoeuvres. We may therefore add a dynamic equation for X_3

$$\frac{dX_3}{dt} = -\alpha\, X_3 + w(t)$$

with w(t): white gaussian noise such that

$$E[w(t)\; w(\tau)] = \delta(t-\tau)\; q$$

The discretization of that model gives for

$$\Delta = t_{i+1} - t_i$$

$$X(t+1) = \begin{bmatrix} 1 & \Delta & \frac{1}{\alpha^2}\left(-1+\alpha\Delta+e^{-\alpha\Delta}\right) \\ 0 & 1 & \frac{1}{\alpha}\left(1 - e^{-\alpha\Delta}\right) \\ 0 & 0 & e^{-\alpha\Delta} \end{bmatrix} X(t)+\varepsilon(t)$$

The random variable $\varepsilon(t)$ is given by the stochastic integral:

$$\varepsilon(t) = \int_t^{t+\Delta} e^{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha \end{bmatrix}(\tau-t)} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} w(\tau)\; d\tau$$

It has been show in /3/ that for $\alpha\Delta$ small we get

$$\Omega_\varepsilon = E[\varepsilon(t)\; \varepsilon^T(\tau)] \simeq \delta_t,$$

$$\cdot\; 2\alpha q \begin{bmatrix} \frac{\Delta^5}{2} & \frac{\Delta^4}{8} & \frac{\Delta^3}{6} \\ \cdot & \frac{\Delta^3}{3} & \frac{\Delta^2}{2} \\ \cdot & \cdot & \Delta \end{bmatrix}$$

The parameters q and α are evaluated beforehand by the admissible characteristics of -- the aircraft taken as target.

We shall therefore identify such a system -- using a canonical model:

$$x(t+1) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x(t) - \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} y(t) + \begin{bmatrix} d_0 \\ d_1 \\ d_2 \end{bmatrix} v(t)$$

$$y(t) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} x(t) + d_3\; v(t)$$

d_0, d_1, d_2, d_3 and the covariance of v(t) -- are known and directly related to Ω_ε and to the error attached to the radar.

3.2 Simulated tests of the algorithm

A first study of the computer algorithm im-- plementing equation (2) is made for different values of the noise characteristics allowing differences between the correct values and - the ones used in the program.

The characteristics of the simulated system are:

- dynamical parameters

$$a^T = [.24,\;\; -1.18,\;\; 1.9]$$

- noise parameters (true values)

$$d^T = [-.056,\;\; .5,\;\; -1.3]$$

The initial values of parameter state estimates are null.

1st test: d is fixed at its true value

$$\hat{a}^T(1000) = [.18,\;\; -1.05,\;\; 1.83]$$

the state components are all well estimated.

2nd test: d is taken very far from its true value

$$d^T = [.5, \ .06, \ .5]$$

The two first components of the state are -- poorly estimated the estimator $\hat{a}$ does not -- settle to a fixed value. The position component of the state is nevertheless well estimated, this is because of the filtering effect of the system.

3rd test: d is taken closer to the true values than in the precedent test

$$d^T = [-.125, \ .75, \ -1.5]$$

then

$$\hat{a}^T(1000) = [.33, \ -1.39, \ 2.02]$$

and all components of state are well estimated.

4th test: the identification by a lower order system has been attempted d has been replaced by

$$d^T = [.25, \ -1]$$

The estimation of the position is always good whereas the estimators of the parameters do not stop fluctuating.

5th test: estimation of a deterministic system is experienced setting d=0 but using non zero initial state conditions

$$x^T(t_0) = [-5, \ 2, \ 10]$$

Then, in a finite number of steps we reach

$$a^T(10) = [.19, \ -1.01, \ 1.77]$$

this is the best estimator as after that time the algorithm amplifies round off errors.

6th test: estimation of a deterministic -- unstable system has been done in the same situation as in the precedent test. The dynamical parameters are now:

$$a^T = [.24, \ -1.19, \ .484]$$

and as previously, in a finite number of - steps we get

$$\hat{a}^T(10) = [.225, \ -1.19, \ .484]$$

Those test enable us to evaluate the robustness of the estimators reached with this algorithm.

3.3 Real data experiment

Fig. 2 represents real data radar measurements of the position (coordinate Z) of an aircraft, for a short period of 50 sample times.

The joint state and dynamical parameter filtering was applied to it, in place of a tra-

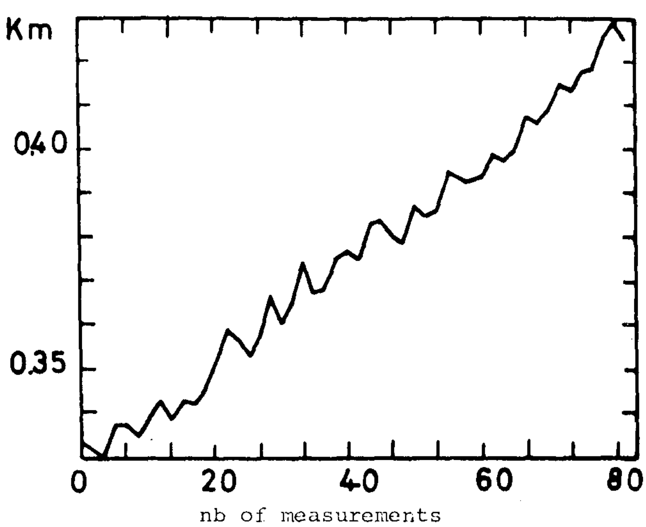

Fig. 2
Real noisy radar measurements

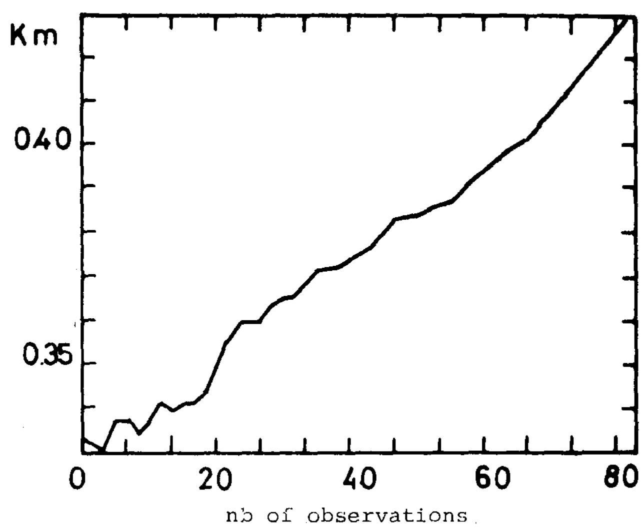

Fig. 3
Estimate of the position

ditionnal linear filtering with arbitrarily fixed dynamical parameters, for a model of - order n=3.

Fig. 3 shows how fast the adaptive filtering works for the useful component $E_0\ \hat{x}$.

Fig. 4-6 give the values to which parameter estimates $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$ converge shen they have been assumed constant.

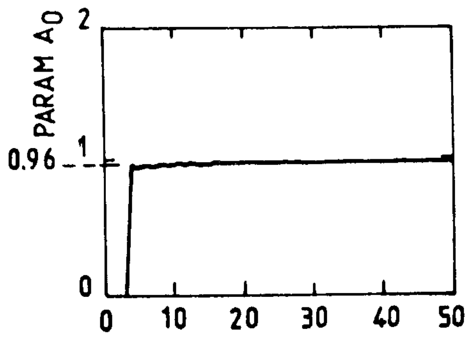

Fig.4
Estimate of the a_0 parameter

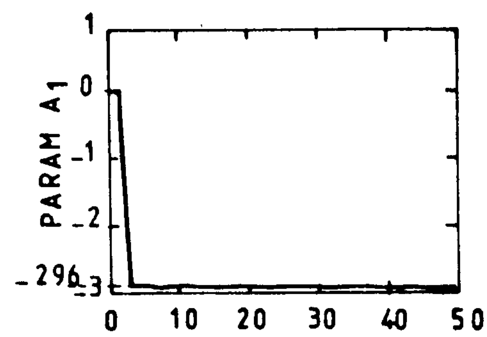

Fig. 5
Estimate of the a_1 parameter

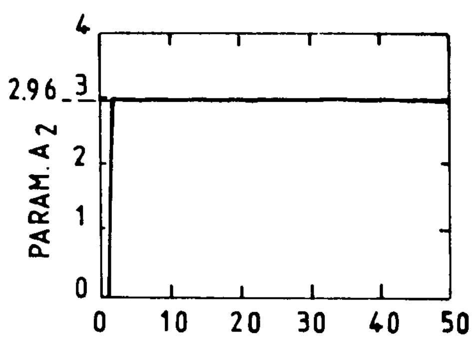

Fig. 6
Estimate of the a_2 parameter

It should be noticed that convergence rate - for the parameter estimates is particularly fast.

4. CONCLUSIONS

Getting to a joint optimal estimation of parameters and state when the state equations are expressed in canonical form, has allowed the estimation of the state and of the three parameters of the system dynamics in the radar tracking problem.

The results obtained in simulated tests and real experiments show that the algorithm converges even with unstable systems provided - that the characteristics of the noise are -- well known.

Besides, the filtering behaviour in front of real data is quite good. The convergence in the parameter estimation is very fast, and a correct state tracking is thus possible.

5. ACKNOWLEDGMENTS

The authors are pleased to acknowledge the - useful help given by Dr. C. BOZZO, Directeur du CAPCA-TOULON, particularly concerning the modelization and collection of real data for the application presented here. This research has been partially supported by the DRM under contract nº 731564 and by the NSF grant INT 75-0437-A01 US-France ERL Berkely / LAAS Toulouse cooperative program.

6. REFERENCES

/1/ LEFEBVRE, S. "Identification des systèmes linéaires. Application aux réseaux - électriques". Rapport technique EP 78-R-31. Ecole Polytechnique de Montréal. 18 Septembre 1978.

/2/ LEFEBVRE, S., AGUILAR MARTIN, J., SALUT, G., DE SANTIS, R.M. "An application of - SALUT canonical form to the identifica-- tion of a two area power system". 16th - Annual Allerton Conference on Communication Control and Computing. Oct. 1978.

/3/ LEGRAND, W. "Analyse des problèmes de mo délisation et de sensibilité en filtrage linéaire. Application à l'étude des problèmes de poursuite par radar". Thèse de Docteur-Ingénieur, Université Paul Sabatier. Toulouse, 1973.

/4/ SALUT, G. "Identifiabilité d'un système dynamique linéaire invariant". C.R. Acad. Sc. Paris, t. 278, série A, 1974, p.181.

/5/ SALUT, G. "Identification optimale des systèmes linéaires stochastiques". Thèse d'Etat. Université Paul Sabatier. Toulouse, 1976.