www.ietdl.org

Published in IET Signal Processing
Received on 9th March 2010
Revised on 14th March 2011
doi: 10.1049/iet-spr.2010.0066

ISSN 1751-9675

Dual optimal filters for parameter estimation of a multivariate autoregressive process from noisy observations

A. Jamoos[1] *E. Grivel*[2] *N. Shakarneh*[1] *H. Abdel-Nour*[1]

[1]*Department of Electronic Engineering, Al-Quds University, Jerusalem, Palestine*
[2]*Equipe Signal and Image, UMR CNRS 5218 IMS - Dpt LAPS, Université Bordeaux 1, 33405 Talence Cedex, France*
E-mail: ali@eng.alquds.edu; alijamoos@ieee.org

Abstract: This study deals with the estimation of a vector process disturbed by an additive white noise. When this process is modelled by a multivariate autoregressive (M-AR) process, optimal filters such as Kalman or H_∞ filter can be used for prediction or estimation from noisy observations. However, the estimation of the M-AR parameters from noisy observations is a key issue to be addressed. Off-line or iterative approaches have been proposed recently, but their computational costs can be a drawback. Using on-line methods such as extended Kalman filter and sigma-point Kalman filter are of interest, but the size of the state vector to be estimated is quite high. In order to reduce this size and the resulting computational cost, the authors suggest using dual optimal filters. In this study, the authors propose to extend to the multi-channel case the so-called dual Kalman or H_∞ filters-based scheme initially proposed for single-channel applications. The proposed methods are first tested with a synthetic M-AR process and then with an M-AR process corresponding to a mobile fading channel. The comparative simulation study the authors carried out with existing techniques confirms the effectiveness of the proposed methods.

1 Introduction

Single-channel autoregressive (AR) model has been popular for years in signal processing applications such as in the field of speech enhancement and coding [1], biomedical signal processing [2], wireless communications [3], clutter rejection for radar processing [4] and so on. In these applications, when the observations are disturbed by an additive measurement noise, the least squares estimates of the AR parameters may be biased. In order to overcome this problem, one can use the noise-compensated Yule–Walker (NCYW) equations, which, however, require the preliminary estimation of the additive-noise variance. In order to deal with the estimations of both the AR process and the noise variance, various off-line or iterative methods have been proposed. Thus, a bias correction least squares scheme has been proposed by Zheng [5], whereas Davila [6] has presented a subspace-based method. Mahmoudi and Karimi [7] have proposed an improved least-squares-based method that combines low-order and high-order Yule–Walker (YW) equations. Bobillet *et al.* [8] have proposed an errors-in-variables-based approach, which consists in searching the noise variances that enable specific noise-compensated autocorrelation matrices of observations to be positive semi-definite. The kernel of the resulting compensated matrices corresponds to an estimation of the parameters. This method is reliable especially for low signal-to-noise ratio (SNR) and has the advantage of providing both the noise variance and the parameters. Expectation–maximisation approaches using Kalman filtering could be also considered, but the initialisation plays a key role in that case. Concerning on-line methods, they can be based on Kalman filtering. In that case, the state vector can store the AR parameters and AR process samples. This results in a non-linear state-space representation of the system. Therefore extended Kalman filter (EKF) and sigma-point Kalman filter (SPKF) can be used [9]. As an alternative, one of the authors of this paper has suggested using two mutually interactive Kalman filter-based solution to avoid a non-linear approach [10]. Once a new observation is available, the first filter uses the latest estimated AR parameters to estimate the signal, whereas the second filter uses the estimated signal to update the AR parameters. This approach can be viewed as a recursive instrumental variable-based method, and hence has the advantage of providing consistent estimates of the parameters from noisy observations. To relax Gaussian assumptions required by Kalman filtering, dual H_∞ filter-based solution was then studied [11]. Moreover, in [12], the relevance of these cross-coupled Kalman and cross-coupled H_∞ filters are investigated for the estimation of mobile fading channels.

Although scalar AR model is often used, a multivariate autoregressive (M-AR) process is more suited when correlated multi-channel signals are simultaneously processed. This is for instance the case in wireless communications, radar and sonar systems, and biomedical applications. Indeed, in biomedical signal processing, the aim is to analyse physiological signals such as the multi-channel electroencephalogram signals [13].

In addition, when dealing with the cardiovascular system, the purpose is to study the interactions between respiratory movement, heart rate fluctuations and blood pressure [14]. In radar processing, when multiple antennas are used, the sea clutter must be rejected to detect the target. Variants of the space-time adaptive processing algorithm such as the parametric adaptive matched filter and the space-time autoregressive filter [15] consist in modelling the sea clutter by an *M*-AR process. Abramovich *et al.* [16] have also focused their attention on the relevance of *M*-AR and time-varying AR process for radar processing. In the framework of wireless communication systems, multiple correlated fading processes are common in multi-carrier systems such as orthogonal frequency division multiplexing (OFDM) [17], in multiple-input multiple-output antenna systems [18], in spread spectrum systems [19] and so on. In these systems, the correlated fading channels are usually modelled as an *M*-AR process and hence can be jointly estimated.

When noise-free observations are available, the comparative study carried out by Schlögl [20] showed that the Burg-type Nuttall–Strand method [21] is the most relevant approach to estimate the *M*-AR parameters among the standard approaches (YW equations, Levinson algorithm and so on) and the so-called autoregressive fitting (ARFIT) approach [22]. However, when the *M*-AR process is disturbed by an additive white noise, the standard estimation methods, mentioned above, lead to biased estimates of the *M*-AR parameter matrices. In order to avoid this drawback, the approach proposed in [23] is based on a set of two equations that the noise variances and the coefficients of the AR matrices satisfy. The first one corresponds to the NCYW equations and the second one allows the noise variances to be expressed from the coefficients of the AR matrices and the autocorrelation of the observations filtered by the inverse filter. Therefore a Newton–Raphson algorithm is used to estimate the noise variances and the *M*-AR parameters are deduced by means of the NCYW equations. In [24], the extension of Zheng's method [5] to the multi-channel case has recently been proposed. Nevertheless, this method may lead to a set of AR parameter matrix estimates corresponding to an unstable system when the SNR is low. Petitjean *et al.* [25] have proposed to extend the method presented in [8] to the multi-channel case. Although the approach provides significant results even for low SNR, the computational cost could be reduced. Concerning on-line methods, two serially connected (SC) Kalman filters [13] or H_∞ filters [26] could be used to estimate the *M*-AR process and its parameters (see Fig. 1). The first filter aims at estimating the *M*-AR parameters, whereas the second filter uses the estimated parameters to estimate the *M*-AR process. However, these methods result in biased parameter estimates as the parameters are estimated directly from the noisy observations. In order to avoid this drawback, one can consider non-linear methods such as EKF or SPKF namely unscented Kalman filter and central difference Kalman filter [25]. Nevertheless, the size of the state vector to be estimated is quite high. Indeed, it stores both the coefficients of the AR parameter matrices and the p last values of the *M*-AR process, where p is the model order. The resulting computational cost hence becomes high. To avoid using high-dimensional matrices, we propose to extend to the multi-channel case the dual Kalman filters [10] and the dual H_∞ filters [11, 12] originally proposed for single-channel applications (see Fig. 2). The proposed methods are then compared with existing methods such as [13, 22, 23, 26] in terms of parameter estimation accuracy. All methods are first tested with synthetic *M*-AR process and then with an *M*-AR process that corresponds to correlated mobile fading channels.

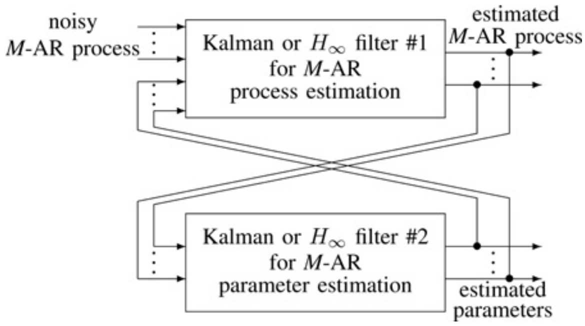

Fig. 2 *Proposed dual Kalman filters or dual H_∞ filters*

In Section 2, the problem statement is presented. In Section 3, the joint estimation of the *M*-AR process and its parameter matrices based on dual Kalman or dual H_∞ filters is introduced. The results of the comparative simulation study between dual Kalman filtering, dual H_∞ filtering and other existing approaches are given in Section 4. Conclusions are drawn in Section 5.

2 Problem statement

Let us consider multiple correlated data channels modelled by a pth-order *M*-AR process

$$\boldsymbol{h}(n) = -\sum_{l=1}^{p} \boldsymbol{A}^{(l)}\boldsymbol{h}(n-l) + \boldsymbol{u}(n) \tag{1}$$

where $\{\boldsymbol{A}^{(l)}\}_{l=1,\ \ldots,\ p}$ are the *M*-AR parameter matrices, $\boldsymbol{h}(n) = [h_1(n) \quad h_2(n) \quad \cdots \quad h_M(n)]^{\mathrm{T}}$ is the $M \times 1$ output signal vector, $\boldsymbol{u}(n) = [u_1(n) \quad u_2(n) \quad \cdots \quad u_M(n)]^{\mathrm{T}}$ is the $M \times 1$ input signal vector, and $[\cdot]^{\mathrm{T}}$ denotes the transpose operation. The driving vector $\boldsymbol{u}(n)$ is assumed to be a zero-mean white noise vector whose autocorrelation matrix $\boldsymbol{\Sigma_u}$ can be written as

$$\sum\nolimits_{\boldsymbol{u}} = \mathrm{diag}([\sigma_{u_1}^2 \quad \sigma_{u_2}^2 \quad \cdots \quad \sigma_{u_M}^2]) \tag{2}$$

where diag $([\cdot])$ denotes a diagonal matrix.

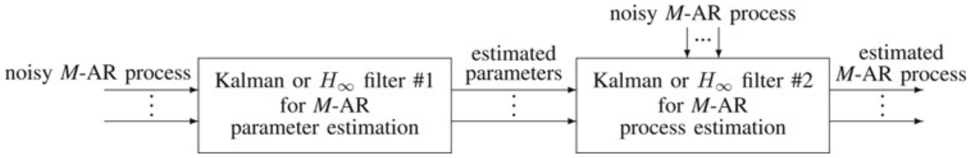

Fig. 1 *Two SC Kalman filters [13] or two SC H_∞ filters [26]*

doi: 10.1049/iet-spr.2010.0066

The M-AR parameter matrices $\{A^{(l)}\}_{l=1,\ldots,p}$ of size $M \times M$ can be expressed as follows

$$A^{(l)} = \begin{bmatrix} a_{11}^{(l)} & a_{12}^{(l)} & \cdots & a_{1M}^{(l)} \\ a_{21}^{(l)} & a_{22}^{(l)} & \cdots & a_{2M}^{(l)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1}^{(l)} & a_{M2}^{(l)} & \cdots & a_{MM}^{(l)} \end{bmatrix} \quad (3)$$

To satisfy stability conditions, the roots $\{p_i\}_{i=1,\ldots,pM}$ of

$$\det([I_M + A^{(1)}z^{-1} + A^{(2)}z^{-2} + \cdots + A^{(p)}z^{-p}]) \quad (4)$$

must lie inside the unit circle in the z-plane, where z^{-1} denotes the backward shift operator, I_M is the $M \times M$ identity matrix, and $\det([\cdot])$ is the determinant operator.

However, in practical applications, the M-AR process $h(n)$ is usually disturbed by additive zero-mean white noise vector $b(n) = [\, b_1(n) \quad b_2(n) \quad \cdots \quad b_M(n)\,]^{\mathrm{T}}$ uncorrelated with $u(n)$. Its correlation matrix satisfies

$$\sum\nolimits_b = E[b(n)b^{\mathrm{T}}(n)] = \mathrm{diag}([\, \sigma_{b_1}^2 \quad \sigma_{b_2}^2 \quad \cdots \quad \sigma_{b_M}^2\,]) \quad (5)$$

Thus, the noisy observation vector can be written as

$$y(n) = h(n) + b(n) \quad (6)$$

with $y(n) = [\, y_1(n) \quad y_2(n) \quad \cdots \quad y_M(n)\,]^{\mathrm{T}}$.

Given the noisy observation vector $y(n)$, the purpose of our approach is to estimate the M-AR parameter matrices $\{A^{(l)}\}_{l=1,\ldots,p}$ by means of mutually interactive dual optimal filters.

3 Dual H_∞ against dual Kalman filters

According to the dual optimal filter structure shown in Fig. 2, the first optimal filter uses the latest estimated AR parameter matrices to estimate the M-AR process, whereas the second optimal filter estimates the parameter matrices from the estimated process vector. These optimal filters are based on linear state-space model as described in the following subsection.

3.1 Linear state-space model

Let us first define the following state vector, whose dimension is $Mp \times 1$

$$\underline{h}(n) = [\, h(n)^{\mathrm{T}} \quad h(n-1)^{\mathrm{T}} \quad \cdots \quad h(n-p+1)^{\mathrm{T}}\,]^{\mathrm{T}} \quad (7)$$

Hence, the state-space representation of the system (1) and (6) can be expressed as

$$\begin{cases} \underline{h}(n) = \Phi \underline{h}(n-1) + \Gamma u(n) \\ y(n) = H\underline{h}(n) + b(n) \end{cases} \quad (8)$$

where the transition matrix Φ is defined as follows

$$\Phi = \begin{bmatrix} -A^{(1)} & -A^{(2)} & \cdots & -A^{(p)} \\ I_M & 0_M & \cdots & 0_M \\ 0_M & \ddots & \ddots & \vdots \\ 0_M & \cdots & I_M & 0_M \end{bmatrix} \quad (9)$$

with 0_M is the $M \times M$ zero matrix. In addition, the $M \times Mp$ output matrix H is related to the input matrix Γ as given by

$$H = \Gamma^{\mathrm{T}} = [\, I_M \quad 0_M \quad \cdots \quad 0_M\,] \quad (10)$$

When using H_∞ filters, one focuses on the estimation of a specific linear combination of the state vector components, as follows

$$z(n) = L\underline{h}(n) \quad (11)$$

where L is a $M \times Mp$ linear transformation operator. Here, as we aim at estimating the process $h(n)$, this operator is selected to be $L = H = \Gamma^{\mathrm{T}} = [\, I_M \quad 0_M \quad \cdots \quad 0_M\,]$.

3.2 Purpose of Kalman or H_∞ filtering

Based on the state-space model (8) and (11), optimal recursive filters make it possible to recursively estimate the state vector $\underline{h}(n)$. In the following, let $\hat{\underline{h}}(n/l)$ denotes the estimation of $\underline{h}(n)$ given $\{y(i)\}_{i=1,\ldots,l}$. Two kinds of approaches can be used. On the one hand, a Kalman filter provides an estimate of the M-AR process $\hat{h}(n/n) = \Gamma^{\mathrm{T}}\hat{\underline{h}}(n/n)$ by minimising the trace of the following a posteriori error covariance matrix

$$P(n/n) = E[(\underline{h}(n) - \hat{\underline{h}}(n/n))(\underline{h}(n) - \hat{\underline{h}}(n/n))^{\mathrm{T}}] \quad (12)$$

It should be noted that this covariance matrix satisfies the so-called Ricatti equation

$$P(n+1/n) = \Phi P(n/n-1)\Phi^{\mathrm{T}} + \Gamma \Sigma_u \Gamma^{\mathrm{T}} - \Phi K_{\mathrm{Kal}}(n) H P(n/n-1)\Phi^{\mathrm{T}} \quad (13)$$

where $K_{\mathrm{Kal}}(n)$ is the Kalman filter gain.

On the other hand, the a posteriori H_∞ filter aims at estimating $\hat{h}(n) = L\hat{\underline{h}}(n)$ by minimising the H_∞ norm of the transfer function that maps the noise vectors $u(n)$, $b(n)$, and the initial state error $e_0 = \underline{h}(0) - \hat{\underline{h}}(0)$ to the estimation error $e(n) = h(n) - \hat{h}(n)$, as follows

$$J_\infty = \sup_{\mathbf{u}(n),\mathbf{b}(n),\underline{\mathbf{h}}(0)} J \quad (14)$$

where

$$J = \frac{\sum_{n=0}^{N-1} e(n)^{\mathrm{T}} e(n)}{e_0^{\mathrm{T}} P_0^{-1} e_0 + \sum_{n=0}^{N-1} (u(n)^{\mathrm{T}} Q_u^{-1} u(n) + b(n)^{\mathrm{T}} R_b^{-1} b(n))} \quad (15)$$

with N the number of available data samples, Q_u and R_b are weighting positive matrices that are tuned by the practitioner to achieve performance requirements. In addition, P_0 denotes a positive matrix that reflects how small the initial state error $e_0 = \underline{h}(0) - \hat{\underline{h}}(0)$ is.

However, as a closed-form solution to the above optimal H_∞ estimation problem does not always exist, the following suboptimal design strategy is usually considered

$$J_\infty < \gamma^2 \quad (16)$$

where $\gamma > 0$ is a prescribed level of disturbance attenuation.

Following the method presented in [27], there exists an H_∞ estimator $\hat{\boldsymbol{h}}(n)$ for a given $\gamma > 0$ if there exists a stabilising symmetric positive definite solution $\boldsymbol{P}_\infty(n) > 0$ to the following Riccati-type equation

$$\boldsymbol{P}_\infty(n+1) = \boldsymbol{\Phi}\boldsymbol{P}_\infty(n)\boldsymbol{D}(n)^{-1}\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{Q}_{\boldsymbol{u}}\boldsymbol{\Gamma}^{\mathrm{T}} \quad (17)$$

where

$$\boldsymbol{D}(n) = \boldsymbol{I}_{Mp} - \gamma^{-2}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{P}_\infty(n) + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{b}}^{-1}\boldsymbol{H}\boldsymbol{P}_\infty(n) \quad (18)$$

This leads to the following constraint

$$\boldsymbol{P}_\infty(n)\boldsymbol{D}(n)^{-1} > \boldsymbol{0} \quad (19)$$

It should be noted that the level attenuation factor should be carefully selected to satisfy the condition in (19) as proposed by Shen and Deng in [28]

$$\gamma^2 > \max(\mathrm{eig}[\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}[\boldsymbol{P}_\infty(n)^{-1} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{b}}^{-1}\boldsymbol{H}]^{-1}]) \quad (20)$$

where $\max(\mathrm{eig}[\boldsymbol{F}])$ is the maximum eigenvalue of the matrix $\boldsymbol{F}$.

At that stage, one can either set γ^2 to a specific constant value that is high enough to satisfy (20) or adjust it according to (20) as follows

$$\gamma^2(n) = \zeta\max(\mathrm{eig}[\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}[\boldsymbol{P}_\infty(n)^{-1} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{b}}^{-1}\boldsymbol{H}]^{-1}]) \quad (21)$$

with $\zeta > 2$.

3.3 Kalman or H_∞ filtering for state-vector estimation

For both Kalman or H_∞ filtering, the state vector and the M-AR process are, respectively, estimated as follows

$$\hat{\underline{\boldsymbol{h}}}(n/n) = \boldsymbol{\Phi}\hat{\underline{\boldsymbol{h}}}(n-1/n-1) + \boldsymbol{K}(n)\boldsymbol{\upsilon}(n) \quad (22)$$

and

$$\hat{\boldsymbol{h}}(n/n) = \boldsymbol{\Gamma}^{\mathrm{T}}\hat{\underline{\boldsymbol{h}}}(n/n) \quad (23)$$

where the so-called innovation process $\upsilon(n)$ is given by

$$\boldsymbol{\upsilon}(n) = \boldsymbol{y}(n) - \boldsymbol{H}\boldsymbol{\Phi}\hat{\underline{\boldsymbol{h}}}(n-1/n-1) \quad (24)$$

However, the way the gain $\boldsymbol{K}(n)$ in (22) is defined depends on the kind of filtering. Thus, when a Kalman filter is used, the gain, now noted $\boldsymbol{K}_{\mathrm{Kal}}(n)$, is given by

$$\boldsymbol{K}_{\mathrm{Kal}}(n) = \boldsymbol{P}(n/n-1)\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{P}(n/n-1)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{b}}]^{-1} \quad (25)$$

where the covariance matrix is updated by using the following set of relations

$$\boldsymbol{P}(n/n-1) = \boldsymbol{\Phi}\boldsymbol{P}(n-1/n-1)\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\Gamma}^{\mathrm{T}} \quad (26)$$

$$\boldsymbol{P}(n/n) = \boldsymbol{P}(n/n-1) - \boldsymbol{K}_{\mathrm{Kal}}(n)\boldsymbol{H}\boldsymbol{P}(n/n-1) \quad (27)$$

In addition, the covariance matrix of the innovation process $\boldsymbol{\upsilon}(n)$ satisfies

$$\boldsymbol{C}(n) = \boldsymbol{H}\boldsymbol{P}(n/n-1)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{b}} \quad (28)$$

When an H_∞ filter is used, the gain is denoted as $\boldsymbol{K}_\infty(n)$ and is given by

$$\boldsymbol{K}_\infty(n) = \boldsymbol{P}_\infty(n)\boldsymbol{D}(n)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{b}}^{-1} \quad (29)$$

According to Yaesh and Shaked [29], the matrix $\boldsymbol{P}_\infty(n)$ can be seen as an upper bound of the error covariance matrix in the Kalman filter theory, that is

$$\begin{aligned}\boldsymbol{P}_\infty(n) &\geq \boldsymbol{P}(n/n)\\ &= E[(\underline{\boldsymbol{h}}(n) - \hat{\underline{\boldsymbol{h}}}(n/n))(\underline{\boldsymbol{h}}(n) - \hat{\underline{\boldsymbol{h}}}(n/n))^{\mathrm{T}}]\end{aligned} \quad (30)$$

Considering the equations of Kalman and H_∞ filters, three remarks can be drawn.

Remark 1: Owing to (18), the H_∞ filter has a computational cost slightly higher than Kalman's one.

Remark 2: Unlike Kalman filter, the H_∞ filter needs to adjust the level attenuation factor γ such as to satisfy (20).

Remark 3: If the weighting matrices $\boldsymbol{Q}_{\boldsymbol{u}}$, $\boldsymbol{R}_{\boldsymbol{b}}$ and $\boldsymbol{P}_0$ are, respectively, chosen to be $\boldsymbol{\Sigma}_{\boldsymbol{u}}$, $\boldsymbol{\Sigma}_{\boldsymbol{b}}$ and the initial error covariance matrix of $\underline{\boldsymbol{h}}(0)$ then the H_∞ filter reduces to the Kalman one when $\gamma \to \infty$.

Kalman or H_∞ filter can be carried out provided that the parameter matrices $\{\boldsymbol{A}^{(l)}\}_{l=1,\ldots,p}$ are available. They will be estimated using the method presented in the following subsection.

3.4 Kalman or H_∞ filtering for M-AR parameter estimation

In this subsection, we propose to estimate the M-AR parameter matrices $\{\boldsymbol{A}^{(l)}\}_{l=1,\ldots,p}$ from the estimated process $\hat{\boldsymbol{h}}(n/n)$. To this end, (22) and (23) are first combined to express the estimated process as a function of the parameter matrices

$$\begin{aligned}\hat{\boldsymbol{h}}(n/n) &= \Gamma^{\mathrm{T}}\boldsymbol{\Phi}\hat{\underline{\boldsymbol{h}}}(n-1/n-1) + \boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{K}(n)\boldsymbol{\upsilon}(n)\\ &= -\boldsymbol{\Theta}\hat{\underline{\boldsymbol{h}}}(n-1/n-1) + \nu(n)\end{aligned} \quad (31)$$

where the parameter matrix

$$\boldsymbol{\Theta} = [\boldsymbol{A}^{(1)} \quad \boldsymbol{A}^{(2)} \quad \cdots \quad \boldsymbol{A}^{(p)}] = \left[\begin{bmatrix} a_{11}^{(1)} & \cdots & a_{1M}^{(1)} \\ \vdots & \ddots & \vdots \\ a_{M1}^{(1)} & \cdots & a_{MM}^{(1)} \end{bmatrix} \cdots \begin{bmatrix} a_{11}^{(p)} & \cdots & a_{1M}^{(p)} \\ \vdots & \ddots & \vdots \\ a_{M1}^{(p)} & \cdots & a_{MM}^{(p)} \end{bmatrix}\right] \quad (32)$$

and the noise vector

$$\nu(n) = \boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{K}(n)\boldsymbol{\upsilon}(n) \quad (33)$$

When a Kalman filter is used, the covariance matrix of $\nu(n)$ is

doi: 10.1049/iet-spr.2010.0066

equal to

$$\boldsymbol{\Sigma}_\nu(n) = \boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{Kal}}(n)\boldsymbol{C}(n)\boldsymbol{K}_{\mathrm{Kal}}(n)^{\mathrm{T}}\boldsymbol{\Gamma} \quad (34)$$

By stacking the columns of the matrix $\boldsymbol{\Theta}^{\mathrm{T}}$ on top of each others, the resulting $M^2p \times 1$ state vector can be expressed as

$$\boldsymbol{\theta}(n) = [[a_{11}^{(1)} \ \cdots \ a_{1M}^{(1)}] \ \cdots \ [a_{11}^{(p)} \ \cdots \ a_{1M}^{(p)}] \ \cdots \ [a_{M1}^{(1)} \ \cdots \ a_{MM}^{(1)}] \ \cdots \ [a_{M1}^{(p)} \ \cdots \ a_{MM}^{(p)}]]^{\mathrm{T}} \quad (35)$$

Hence, (31) can be rewritten as follows

$$\hat{\boldsymbol{h}}(n/n) = \boldsymbol{H}_{\boldsymbol{\theta}}(n)\boldsymbol{\theta}(n) + \nu(n) \quad (36)$$

where

$$\boldsymbol{H}_{\boldsymbol{\theta}}(n) = -\boldsymbol{I}_M \otimes \underline{\hat{\boldsymbol{h}}}(n-1/n-1)^{\mathrm{T}} \quad (37)$$

with $\otimes$ denotes the matrix Kronecker product.

When the M-AR process is assumed stationary, the AR parameters are time-invariant and, hence, satisfy the following relationship

$$\boldsymbol{\theta}(n) = \boldsymbol{\theta}(n-1) \quad (38)$$

Thus, (36) and (38) define a state-space representation for the estimation of the AR parameters. A second optimal filter is then used to recursively estimate $\boldsymbol{\theta}(n)$. If a second H_∞ filter is chosen to recursively estimate $\boldsymbol{\theta}(n)$, the AR parameter estimation error is defined as $\boldsymbol{e}_{\boldsymbol{\theta}} = \boldsymbol{H}_{\boldsymbol{\theta}}(n)\ (\boldsymbol{\theta}(n) - \hat{\boldsymbol{\theta}}(n))$. This second H_∞ filter requires two weighting positive matrices $\boldsymbol{R}_{\boldsymbol{\nu}} > 0$ and $\boldsymbol{P}_{\theta 0} > 0$ that can be tuned by the designer. In addition, it needs the disturbance attenuation level $\gamma_{\boldsymbol{\theta}}$, which should be selected in the same manner as γ in (20).

The dual Kalman and H_∞ filtering algorithms are summarised respectively in Tables 1 and 2. The estimation and tuning of other parameters required by the dual optimal filters are addressed in the following subsection.

3.5 Other parameters to be estimated or tuned

Labarre *et al.* [10] have proposed an iterative procedures for the estimation of the variances of both the driving process

Table 1 Dual Kalman filtering algorithm

First Kalman filter: M-AR process estimation
$\boldsymbol{P}(n/n-1) = \boldsymbol{\Phi}\boldsymbol{P}(n-1/n-1)\ \boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\Gamma}^{\mathrm{T}}$
$\upsilon(n) = \boldsymbol{y}(n) - \boldsymbol{H}\boldsymbol{\Phi}\underline{\hat{\boldsymbol{h}}}(n-1/n-1)$
$\boldsymbol{C}(n) = \boldsymbol{H}\boldsymbol{P}(n/n-1)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{b}}$
$\boldsymbol{K}_{\mathrm{Kal}}(n) = \boldsymbol{P}(n/n-1)\boldsymbol{H}^{\mathrm{T}}\boldsymbol{C}(n)^{-1}$
$\underline{\hat{\boldsymbol{h}}}(n/n) = \boldsymbol{\Phi}\underline{\hat{\boldsymbol{h}}}(n-1/n-1) + \boldsymbol{K}_{\mathrm{Kal}}(n)\upsilon(n)$
$\hat{\boldsymbol{h}}(n/n) = \boldsymbol{\Gamma}^{\mathrm{T}}\underline{\hat{\boldsymbol{h}}}(n/n)$
$\boldsymbol{P}(n/n) = \boldsymbol{P}(n/n-1) - \boldsymbol{K}_{\mathrm{Kal}}(n)\boldsymbol{H}\boldsymbol{P}(n/n-1)$
Second Kalman filter: M-AR parameter estimation
$\boldsymbol{H}_{\boldsymbol{\theta}}(n) = -\boldsymbol{I}_M \otimes \underline{\hat{\boldsymbol{h}}}(n-1/n-1)^{\mathrm{T}}$
$\boldsymbol{P}_\theta(n/n-1) = \boldsymbol{P}_\theta(n-1/n-1)$
$\boldsymbol{C}_\theta(n) = \boldsymbol{H}_\theta(n)\boldsymbol{P}_\theta(n/n-1)\boldsymbol{H}_\theta(n)^{\mathrm{T}} + \boldsymbol{\Sigma}_\nu(n)$
$\boldsymbol{K}_\theta(n) = \boldsymbol{P}_\theta(n/n-1)\boldsymbol{H}_\theta(n)^{\mathrm{T}}\boldsymbol{C}_\theta(n)^{-1}$
$\boldsymbol{\theta}(n) = \boldsymbol{\theta}(n-1) + \boldsymbol{K}_{\boldsymbol{\theta}}(n)(\hat{\boldsymbol{h}}(n/n) - \boldsymbol{H}_{\boldsymbol{\theta}}(n)\boldsymbol{\theta}(n-1))$
$\boldsymbol{P}_\theta(n/n) = \boldsymbol{P}_\theta(n/n-1) - \boldsymbol{K}_\theta(n)\boldsymbol{H}_\theta(n)\boldsymbol{P}_\theta(n/n-1)$

Table 2 Dual H_∞ filtering algorithm

First H_∞ filter: M-AR process estimation
$\boldsymbol{D}(n) = \boldsymbol{I}_{Mp} - \gamma^{-2}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{P}_\infty(n) + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{b}}^{-1}\boldsymbol{H}\boldsymbol{P}_\infty(n)$
$\boldsymbol{K}_\infty(n) = \boldsymbol{P}_\infty(n)\boldsymbol{D}(n)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{b}}^{-1}$
$\upsilon(n) = \boldsymbol{y}(n) - \boldsymbol{H}\boldsymbol{\Phi}\underline{\hat{\boldsymbol{h}}}(n-1)$
$\underline{\hat{\boldsymbol{h}}}(n) = \boldsymbol{\Phi}\underline{\hat{\boldsymbol{h}}}(n-1) + \boldsymbol{K}_\infty(n)\upsilon(n)$
$\hat{\boldsymbol{h}}(n) = \boldsymbol{L}\underline{\hat{\boldsymbol{h}}}(n)$
$\boldsymbol{P}_\infty(n+1) = \boldsymbol{\Phi}\boldsymbol{P}_\infty(n)\boldsymbol{D}(n)^{-1}\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{Q}_{\boldsymbol{u}}\boldsymbol{\Gamma}^{\mathrm{T}}$
Second H_∞ filter: M-AR process estimation
$\boldsymbol{H}_{\boldsymbol{\theta}}(n) = -\boldsymbol{I}_M \otimes \hat{\boldsymbol{h}}(n-1)^{\mathrm{T}}$
$\boldsymbol{D}_{\boldsymbol{\theta}}(n) = \boldsymbol{I}_{M^2p} - \gamma_{\boldsymbol{\theta}}^{-2}\boldsymbol{H}_{\boldsymbol{\theta}}(n)^{\mathrm{T}}\boldsymbol{H}_{\boldsymbol{\theta}}(n)\underline{\boldsymbol{P}}_{\boldsymbol{\theta}}(n) + \boldsymbol{H}_{\boldsymbol{\theta}}(n)^{\mathrm{T}}\boldsymbol{R}_\nu(n)^{-1}\boldsymbol{H}_{\boldsymbol{\theta}}(n)\underline{\boldsymbol{P}}_{\boldsymbol{\theta}}(n)$
$\underline{\boldsymbol{K}}_{\boldsymbol{\theta}}(n) = \underline{\boldsymbol{P}}_{\boldsymbol{\theta}}(n)\boldsymbol{D}_{\boldsymbol{\theta}}(n)^{-1}\boldsymbol{H}_{\boldsymbol{\theta}}(n)^{\mathrm{T}}\boldsymbol{R}_\nu(n)^{-1}\mathrm{v}$
$\boldsymbol{\theta}(n) = \boldsymbol{\theta}(n-1) + \underline{\boldsymbol{K}}_{\boldsymbol{\theta}}(n)(\hat{\boldsymbol{h}}(n) - \boldsymbol{H}_{\boldsymbol{\theta}}(n)\boldsymbol{\theta}(n-1))$
$\underline{\boldsymbol{P}}_{\boldsymbol{\theta}}(n+1) = \underline{\boldsymbol{P}}_{\boldsymbol{\theta}}(n)\boldsymbol{D}_{\boldsymbol{\theta}}(n)^{-1}$

and additive noise for single-channel applications. Here, we propose to extend these results to the multi-channel case. Thus, the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}}$ of the driving noise vector $\boldsymbol{u}(n)$ can be iteratively estimated by using the Kalman filtering (25)–(28) as follows

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}(n) = \lambda\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}(n-1) + (1-\lambda)\boldsymbol{F}\boldsymbol{M}(n)\boldsymbol{F}^{\mathrm{T}} \quad (39)$$

where the matrix $\boldsymbol{M}(n) = \underline{\boldsymbol{P}}(n/n) - \boldsymbol{\Phi}\boldsymbol{P}(n-1/n-1)\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{K}(n)\upsilon(n)\upsilon(n)^{\mathrm{T}}\boldsymbol{K}(n)^{\mathrm{T}}, \boldsymbol{F} = [\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}]^{-1}\boldsymbol{\Gamma}^{\mathrm{T}} = [\boldsymbol{I}_M \ \ \boldsymbol{0}_M \ \ \cdots \ \ \boldsymbol{0}_M]$ and λ is the forgetting factor.

In addition, the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{b}}$ of the driving noise vector $\boldsymbol{b}(n)$ can be iteratively estimated based on (25)–(28) in the following manner

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{b}}(n) = \lambda\hat{\boldsymbol{\Sigma}}_{\boldsymbol{b}}(n-1) + (1-\lambda)(\upsilon(n)\upsilon(n)^{\mathrm{T}} - \boldsymbol{H}\boldsymbol{P}(n/n-1)\boldsymbol{H}^{\mathrm{T}}) \quad (40)$$

Instead of manually tuning the weighting matrices $\boldsymbol{Q}_{\boldsymbol{u}}$ and $\boldsymbol{R}_{\boldsymbol{\nu}}$ in the H_∞ filters, a recursive and heuristic approach is here presented. Thus, by analogy with the Kalman filter theory, the weighting matrix $\boldsymbol{Q}_{\boldsymbol{u}}$ in the first H_∞ filter can be recursively tuned as follows [11]

$$\hat{\boldsymbol{Q}}_{\boldsymbol{u}}(n) = \lambda\hat{\boldsymbol{Q}}_{\boldsymbol{u}}(n-1) + (1-\lambda)\boldsymbol{L}\boldsymbol{M}_\infty(n)\boldsymbol{L}^{\mathrm{T}} \quad (41)$$

where $\boldsymbol{M}_\infty(n) = \boldsymbol{P}_\infty(n) - \boldsymbol{\Phi}\boldsymbol{P}_\infty(n-1)\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{K}_\infty(n)\upsilon(n)\upsilon(n)^{\mathrm{T}}\boldsymbol{K}_\infty(n)^{\mathrm{T}}$. Furthermore, the weighting matrix $\boldsymbol{R}_{\boldsymbol{\nu}}$ in the second H_∞ filter could be tuned in the following manner

$$\boldsymbol{R}_\nu = \boldsymbol{L}\boldsymbol{K}_\infty(n)\upsilon(n)\upsilon(n)^{\mathrm{T}}\boldsymbol{K}_\infty(n)^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} \quad (42)$$

Here, the weighting matrix $\boldsymbol{R}_{\boldsymbol{b}}$ is assigned to $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{b}}(n)$. Moreover, as there is no a priori knowledge about the initial state error, the weighting matrices $\boldsymbol{P}_0$ and $\boldsymbol{P}_{\boldsymbol{\theta}0}$ are assigned to the identity matrices $\boldsymbol{I}_{Mp}$ and $\boldsymbol{I}_M$, respectively.

4 Simulation results

In this section, we carry out a comparative simulation study on the estimation of M-AR parameter matrices between several methods:

- the proposed dual H_∞ filters,
- the proposed dual Kalman filters,
- the two SC H_∞ filters [26],

- the two SC Kalman filters [13],
- the so-called ARFIT method [22],
- Hasan's method [23],
- the NCYW equations.

We have tested two kinds of M-AR processes. The first one is a synthetic M-AR process, whereas the second one corresponds to correlated mobile fading channels.

4.1 Synthetic M-AR process

Here, we have considered a second-order ($p = 2$) two-channel ($M = 2$) AR process

$$\boldsymbol{h}(n) = -\boldsymbol{A}^{(1)}\boldsymbol{h}(n-1) - \boldsymbol{A}^{(2)}\boldsymbol{h}(n-2) + \boldsymbol{u}(n) \qquad (43)$$

where the M-AR parameter matrices are those defined by Hasan in [23]

$$\boldsymbol{A}^{(1)} = \begin{bmatrix} -0.71 & 0.32 \\ -0.88 & -0.24 \end{bmatrix}, \quad \boldsymbol{A}^{(2)} = \begin{bmatrix} 0.57 & -0.15 \\ -0.49 & -0.30 \end{bmatrix}$$

In that case, the M-AR parameter matrices lead to four roots of $\det([\boldsymbol{I}_M + \boldsymbol{A}^{(1)}z^{-1} + \boldsymbol{A}^{(2)}z^{-2}])$, namely $p_1 = 0.941 \times e^{j1.125}$, $p_2 = 0.941 \times e^{-j1.125}$, $p_3 = 0.599$ and $p_4 = -0.461$. In addition, $\boldsymbol{u}(n)$ is the two-channel stationary Gaussian white noise, uncorrelated between channels and with unit variance on each channel. The additive noise $\boldsymbol{b}(n)$ is also a two-channel stationary Gaussian white noise, uncorrelated with $\boldsymbol{u}(n)$.

In all of our simulations, the results are averaged over 1000 realisations. The criterion we considered is the mean square error (MSE) of the estimated modulus and the estimated argument of the roots.

According to Table 3, the ARFIT, the YW equations and the SC Kalman filters result in large MSE and thus yield biased estimates. Hasan algorithm leads to a smaller MSE but might diverge in some cases. Our approaches and the NCYW equations with true noise variances provide quite similar results and lead to the smallest MSE.

Figs. 3 and 4 show the MSE of the estimated modulus of the first pole p_1 for different number of samples and different SNR, respectively. According to these figures, increasing the number of samples or SNR will decrease the MSE. In addition, the SC-Kalman filters and the SC-H_∞ filters yield large MSE when compared with the other methods and hence lead to biased estimates. Moreover, the proposed approaches and the NCYW equations with true noise covariance matrices provide quite similar results and yield smallest MSE.

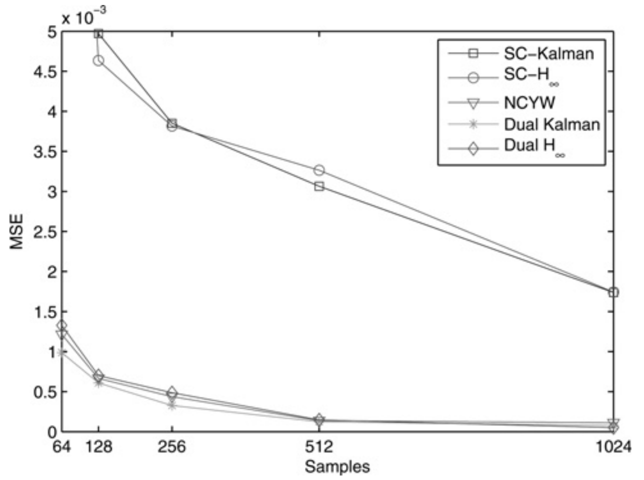

Fig. 3 *MSE of the modulus of p_1 for different number of samples with SNR = 10 dB over each channel*

True values of noise covariance matrices are used

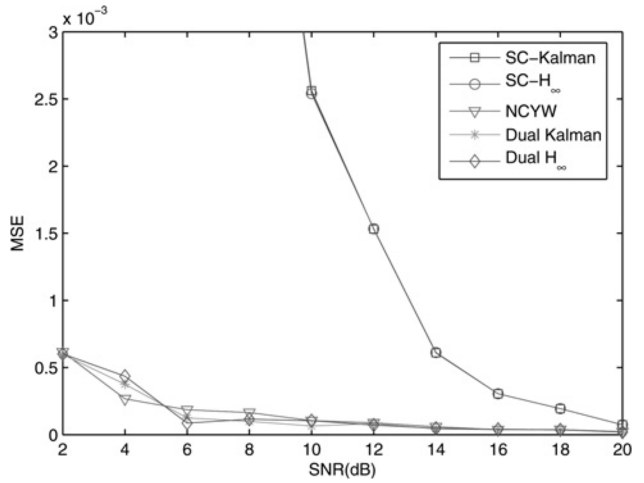

Fig. 4 *MSE of the modulus of p_1 for different SNR with 1024 samples*

True values of noise covariance matrices are used

In order to study the performance of our approaches when only estimated values of noise variances are available, let us introduce β whose ith element is defined as follows

$$\beta^{(i)} = \frac{\hat{\sigma}_{u_i}^2}{\sigma_{u_i}^2} = \frac{\hat{\sigma}_{b_i}^2}{\sigma_{b_i}^2}, \quad i = 1, 2, \ldots, M \qquad (44)$$

Table 3 MSE of the estimated modulus and argument of the poles. SNR = 10 dB over each channel, 300 samples and 1000 realisations

	MSE of p_1		MSE of p_2		MSE of p_3		MSE of p_4	
	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-3}$)	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-3}$)	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-3}$)	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-3}$)
ARFIT [22]	2.74	0.30	2.74	0.30	136.04	0	42.81	0
SC-Kalman [13]	2.81	0.31	2.81	0.31	127.15	0	41.36	0
YW	3.16	0.30	3.16	0.30	129.82	0	43.44	0
Hasan [23]	1.04	0.25	1.04	0.25	252.68	0	16.36	0
NCYW	0.23	0.28	0.23	0.28	3.67	0	10.00	0
Dual Kalman	0.21	0.29	0.21	0.29	4.05	0	21.99	0
Dual H_∞	0.22	0.27	0.23	0.27	2.89	0	11.86	0

doi: 10.1049/iet-spr.2010.0066

Table 4 MSE of the estimated modulus and argument of the pole p_1 for different $\beta^{(i)}$. SNR = 10 dB for the first channel and SNR = 5 dB for the second channel, 1024 samples and 1000 realisations

$\beta^{(i)}$	MSE of estimated mod. of p_1 ($\times 10^{-6}$)			MSE of estimated arg. of p_1 ($\times 10^{-6}$)		
	NCYW	Dual Kalman	Dual H_∞	NCYW	Dual Kalman	Dual H_∞
0.01	3618.10	164.30	138.71	160.85	161.42	152.21
0.10	2897.50	132.21	113.68	141.07	145.40	138.61
0.80	445.06	116.51	107.93	106.56	113.27	124.05
0.90	197.06	106.78	92.60	87.92	119.22	123.86
0.95	191.88	92.23	83.77	79.13	96.47	98.08
1.00	91.82	52.98	60.19	81.93	83.89	95.88
1.05	91.62	83.92	89.97	96.20	88.70	74.36
1.10	188.25	101.85	96.73	112.79	124.27	117.90
1.20	219.21	119.96	112.52	147.42	124.49	117.10
10.00	775711.10	144.60	135.72	692295.60	143.01	133.37

where $\hat{\sigma}^2_{u_i}$ and $\hat{\sigma}^2_{b_i}$ denote, respectively, the estimated driving and additive noise variances over the ith channel. Different values of $\beta^{(i)}$ are used to study the influence of under estimate ($\beta^{(i)} < 1$) and over estimate ($\beta^{(i)} > 1$) the noise variances on the MSE of the estimated modulus and argument of the poles. For the sake of clarity, we only provide simulation results dedicated to the first pole p_1. The results of the other poles are omitted for convenience, as they produce approximately the same kind of results. According to Table 4 and for every method, the MSE of the argument and the modulus of the estimated pole increase when increasing $\beta^{(i)}$ above 1 or decreasing it below 1. For $\beta^{(i)} \geq 1.1$ and $\beta^{(i)} \leq 0.1$, the dual H_∞ filtering algorithm yields less MSE than the dual Kalman filtering algorithm, whereas the NCYW equations provide the worst performance. Thus, the dual H_∞ filtering algorithm is more robust to driving and additive noise variance deviations than the dual Kalman filtering algorithm.

In order to study the convergence characteristics of the dual Kalman filtering algorithm and the dual H_∞ filtering algorithm, we present in Fig. 5 the estimated parameters of the first matrix $\boldsymbol{A}^{(1)}$ against the number of samples. From this figure, one can notice that the two algorithms provide approximately the same convergence characteristics and converge to the true parameter values after $\simeq$200 samples.

4.2 Correlated mobile fading channels

In this subsection, we use the method presented by Baddour and Beaulieu in [30] to generate M-AR processes that corresponds to correlated mobile fading channels. More particularly, we generate two correlated fading channels $\boldsymbol{h}(n) = [h_1(n) \quad h_2(n)]^{\mathrm{T}}$ from a second-order ($p = 2$) two-channel ($M = 2$) AR process. The generated channels are based on the following correlation matrix

$$\boldsymbol{R}_{\boldsymbol{hh}}(k) = E[\boldsymbol{h}(n+k)\boldsymbol{h}(n)^{\mathrm{H}}] = \begin{bmatrix} J_0(2\pi f_m|k|) & 0.6 \times J_0(2\pi f_m|k|) \\ 0.6 \times J_0(2\pi f_m|k|) & J_0(2\pi f_m|k|) \end{bmatrix} \quad (45)$$

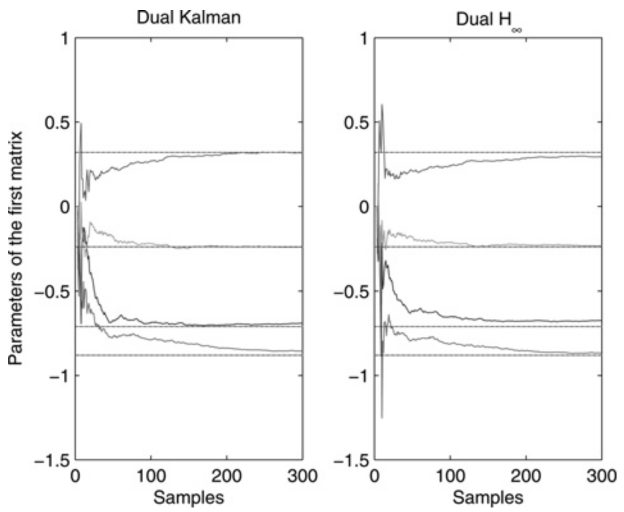

Fig. 5 *Convergence characteristics of the estimated parameters of the first matrix* $\boldsymbol{A}^{(1)}$
a Dual Kalman algorithm
b Dual H_∞ algorithm

Table 5 MSE of the estimated modulus and argument of the poles. SNR = 10 dB for each channel, 300 samples and 1000 realisations

	MSE of p_1		MSE of p_2		MSE of p_3		MSE of p_4	
	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-6}$)	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-6}$)	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-6}$)	mod. ($\times 10^{-3}$)	arg. ($\times 10^{-6}$)
ARFIT [22]	62.52	9.57	67.51	101.17	181.26	15.31	789.88	2589.1
SC-Kalman [13]	62.61	9.65	67.28	101.21	181.25	15.36	790.13	2592.1
YW	63.97	9.95	68.39	101.50	188.02	15.85	792.98	2541.6
Hasan [23]	8.71	24.21	8.23	31.97	15.61	119.05	18.20	83.83
NCYW	15.65	5.09	18.23	4.26	10.28	1.56	9.75	1.33
dual Kalman	0.51	0.25	0.30	0.15	3.092	0.77	5.14	4.46
dual H_∞	0.46	0.31	0.57	1.50	26.27	1.09	19.58	18.66

doi: 10.1049/iet-spr.2010.0066

where $(\cdot)^{\mathrm{H}}$ denotes the Hermitian operation, $\boldsymbol{J}_0(\cdot)$ is a zero-order Bessel function of the first kind and f_m denotes the Doppler rate. When the Doppler rate is set to $f_m = 0.1$ then the corresponding M-AR parameter matrices $\boldsymbol{A}^{(1)}$ and $\boldsymbol{A}^{(2)}$, obtained from the YW equations, satisfy

$$\boldsymbol{A}^{(1)} = \begin{bmatrix} -1.7625 & 0 \\ 0 & -1.7625 \end{bmatrix}, \quad \boldsymbol{A}^{(2)} = \begin{bmatrix} 0.9503 & 0 \\ 0 & 0.9503 \end{bmatrix}$$

In addition, the driving process covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}}$ satisfies

$$\boldsymbol{\Sigma}_{\boldsymbol{u}} = \begin{bmatrix} 0.0178 & 0.0124 \\ 0.0124 & 0.0178 \end{bmatrix}$$

The M-AR parameter matrices $\boldsymbol{A}^{(1)}$ and $\boldsymbol{A}^{(2)}$ lead to the following poles

$$p_1 = p_3 = 0.9748 \times e^{j0.4417}, \quad p_2 = p_4 = 0.9748 \times e^{-j0.4417}$$

Note that the parameter matrices are diagonal, whereas the driving process covariance matrix is not diagonal. In that case, the correlation between the generated channels is because of the two-channel driving vector $\boldsymbol{u}(n) = [\, u_1(n) \quad u_2(n)\,]^{\mathrm{T}}$.

According to Table 5, Figs. 6 and 7, the proposed approaches result in lower MSE than the NCYW equations, whereas the other approaches yield much larger MSE.

From Table 6 and for $\beta^{(i)} \geq 1.05$ and $\beta^{(i)} \leq 0.1$, the dual H_∞ filtering algorithm provides less MSE than the dual Kalman filtering algorithm, although the NCYW equations yield the largest MSE. Thus, the dual H_∞ filtering algorithm is more robust to driving and additive noise variance deviations than the dual Kalman filtering algorithm.

The order of computational complexity per data sample for the investigated methods are summarised in Table 7. The EKF and the SPKF approaches have the highest computational cost as the size of the state vector $(Mp + M^2p) \times 1$ to be estimated is quite high. Indeed, it stores both the p last values of the M-AR process and the coefficients of the AR parameter matrices. The dual filters as well as the SC ones have lower computational cost than the EKF. In these filters, the size of the state vector for estimating the M-AR process is $Mp \times 1$, whereas that for estimating the coefficients of the AR parameter matrices is $M^2p \times 1$. Thus, they require a complexity of the order of $O((Mp)^3)$ and $O((M^2p)^2)$ for estimating the M-AR process and its parameters, respectively. Hasan's method requires

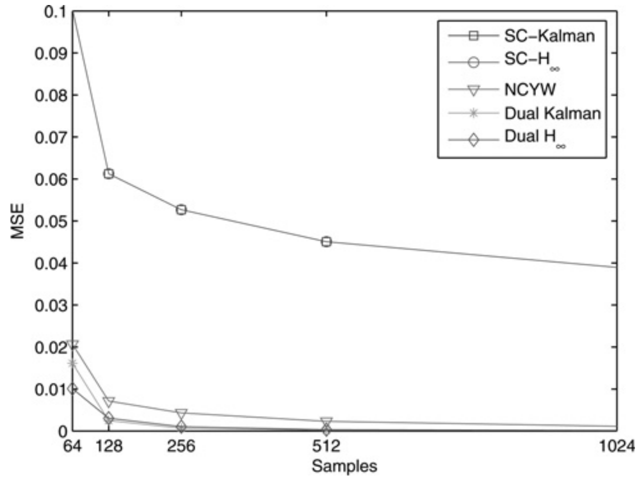

Fig. 6 *MSE of the modulus of p_1 for different number of samples with SNR = 10 dB over each channel*

True values of noise covariance matrices are used

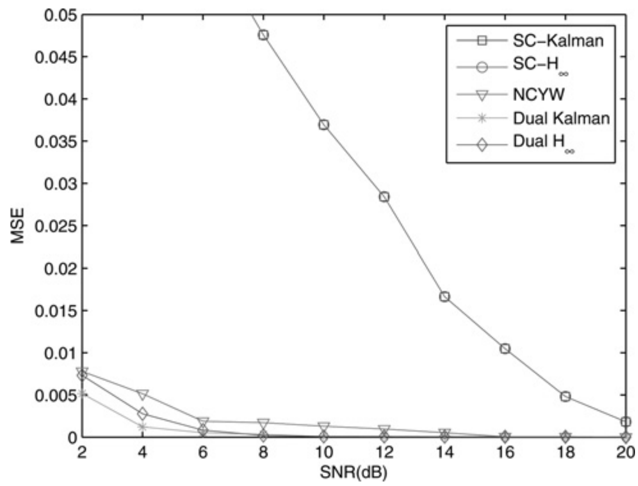

Fig. 7 *MSE of the modulus of p_1 for different SNR with 1024 samples*

True values of noise covariance matrices are used

Table 6 MSE of the estimated modulus and argument of the pole p_1 for different $\beta^{(i)}$. SNR = 10 dB for the first channel and SNR = 5 dB for the second channel, 1024 samples and 1000 realisations

$\beta^{(i)}$	MSE of estimated mod. of p_1 ($\times 10^{-6}$)			MSE of estimated arg. of p_1 ($\times 10^{-6}$)		
	NCYW	Dual Kalman	Dual H_∞	NCYW	Dual Kalman	Dual H_∞
0.01	117 059.6	1750.20	1255.90	97 072.7	2224.30	377.53
0.10	92 132.7	1116.70	883.80	54 750.6	2023.60	517.57
0.80	12 320.5	611.26	729.92	1916.50	1159.20	924.38
0.90	2469.7	595.60	591.33	1624.80	1130.00	512.56
0.95	1944.2	396.90	413.73	1445.10	494.19	678.77
1.00	1559.3	489.01	537.05	214.15	266.99	276.66
1.05	1347.2	667.98	376.30	6264.60	535.65	297.33
1.10	9360.5	524.79	500.60	11 249.4	646.91	366.94
1.20	23 984.9	578.30	534.67	69 000.5	537.46	451.66
10.00	171 722.9	905.84	636.11	102 701.3	1065.50	970.22
20.00	408 634.8	766.65	619.68	986 772.7	1745.00	1017.10

doi: 10.1049/iet-spr.2010.0066

Table 7 Order of computational complexity per data sample for the various methods

Method	Computational complexity
EKF or SPKF [9]	$O((Mp+M^2p)^3)$
dual Kalman or dual H_∞	$O((Mp)^3)+O((M^2p)^2)$
SC-Kalman [13] or SC-H_∞ [26]	$O((Mp)^3)+O((M^2p)^2)$
ARFIT [22]	$O((Mp)^2)$
Hasan [23]	$O((Mp)^2)$
NCYW	$O((Mp)^2)$

solving the NCYW equations, which can be done using recursive techniques such as the Levinson algorithm with complexity of the order of $O((Mp)^2)$. Although solving the NCYW equations requires the lowest computational cost, the dual filter approaches have the advantage of not only estimating the coefficients of the AR parameter matrices but also providing an estimate of the M-AR process.

5 Conclusion

In this paper, we present a method for the joint estimation of a multivariate AR process and its parameter matrices from noisy observations. In particular, we propose to extend to the multi-channel case the so-called dual Kalman or H_∞ filters based scheme initially proposed for single-channel applications. According to the comparative simulation study we carried out on M-AR parameter estimation, the proposed dual optimal filters outperform the existing two SC Kalman or H_∞ filter-based methods. In addition, the dual H_∞ filtering algorithm is robust to noise variance deviations and outperforms the dual Kalman filtering algorithm and the NCYW equations in that case.

6 References

1 Chen, J.H., Cox, R.V., Lin, Y.C., Jayant, N., Melchner, M.J.: 'A low-delay CELP coder for the CCITT 16 kb/s speech coding standard', *IEEE J. Sel. Areas Commun.*, 1992, **10**, pp. 830–849

2 Guler, I., Kiymik, M.K., Akin, M., Alkan, A.: 'AR spectral analysis of EEG signals by using maximum likelihood estimation', *Comput. Biol. Med.*, 2001, **31**, pp. 441–450

3 Baddour, K.E., Beaulieu, N.C.: 'Autoregressive modeling for fading channel simulation', *IEEE Trans. Wirel. Commun.*, 2005, **4**, pp. 1650–1662

4 Wensink, H.: 'On parametric detection of small targets in sea clutter'. Proc. Third Int. Conf. Information Fusion, Paris, France, 10–13 July 2000, pp. 17–24

5 Zheng, W.X.: 'Fast identification of autoregressive signals from noisy observations', *IEEE Trans. Circuits Syst.-II: Express Briefs*, 2005, **52**, pp. 43–48

6 Davila, C.E.: 'A subspace approach to estimation of autoregressive parameters from noisy measurements', *IEEE Trans. Signal Process.*, 1998, **46**, pp. 531–534

7 Mahmoudi, A., Karimi, M.: 'Parameter estimation of autoregressive signals from observations corrupted with colored noise', *Signal Process.*, 2010, **90**, pp. 157–164

8 Bobillet, W., Diversi, R., Grivel, E., Guidorzi, R., Najim, M., Soverini, U.: 'Speech enhancement combining optimal smoothing and errors-in-variables identification of noisy AR processes', *IEEE Trans. Signal Process.*, 2007, **55**, pp. 5564–5578

9 Haykin, S.: 'Kalman filtering and neural networks' (John Wiley & Sons, Inc., NY, USA, 2001)

10 Labarre, D., Grivel, E., Berthoumieu, Y., Todini, E., Najim, M.: 'Consistent estimation of autoregressive parameters from noisy observations based on two interacting Kalman filters', *Signal Process.*, 2006, **86**, pp. 2863–2876

11 Labarre, D., Grivel, E., Najim, M., Christov, N.: 'Dual H_∞ algorithms for signal processing – application to speech enhancement', *IEEE Trans. Signal Process.*, 2007, **55**, pp. 5195–5208

12 Jamoos, A., Grivel, E., Christov, N., Najim, M.: 'Estimation of autoregressive fading channels based on two cross-coupled H_∞ filters', *Signal, Image Video Process. J. (SIViP), Springer*, 2009, **3**, (4), pp. 209–216

13 Arnold, M., Milner, W.H.R., Witte, H., Bauer, R., Braun, C.: 'Adaptive AR modeling of nonstationary time series by means of Kalman filtering', *IEEE Trans. Biomed. Eng.*, 1998, **45**, pp. 553–562

14 Korhonen, I., Takalo, R., Turjanmaa, V.: 'Multivariate autoregressive model with immediate transfer paths for assessment of interactions between cardiopulmonary variability signals', *Med. Biol. Eng. Comput.*, 1996, **34**, pp. 199–206

15 Russ, J., Casbeer, D., Swindlehurst, A.L.: 'STAP detection using space-time autoregressive filtering'. Proc. IEEE-Radar Conf., Philadelphia, USA, 2004, pp. 541–545

16 Abramovich, Y.I., Johnson, B.A., Spencer, N.K.: 'Two-dimensional multivariate parametric models for radar applications-part i: Maximum-entropy extensions for toeplitz-block matrices', *IEEE Trans. Signal Process.*, 2008, **56**, pp. 5509–5526

17 Chen, W., Zhang, R.: 'Estimation of time and frequency selective channels in OFDM systems: a Kalman filter structure'. Proc. IEEE-Globecom, TX, USA, 2004, pp. 800–803

18 Komninakis, C., Fragouli, C., Sayed, A.H., Wesel, R.D.: 'Multi-input multi-output fading channel tracking and equalization using Kalman estimation', *IEEE Trans. Signal Process.*, 2002, **50**, pp. 1065–1076

19 Gao, W., Tsai, S., Lehnert, J.S.: 'Diversity combining for DS/SS systems with time-varying, correlated fading branches', *IEEE Trans. Commun.*, 2003, **51**, pp. 284–295

20 Schlögl, A.: 'A comparaison of multivariate AR estimators', *Signal Process.*, 2006, **86**, pp. 2426–2429

21 Marple, S.L., Nuttall, A.H.: 'Experimental comparison of three multichannel linear prediction spectral estimators'. IEE Proceedings-F, 1983, vol. 130, pp. 218–229

22 Neumaier, A., Schneider, T.: 'Estimation of parameters and eigenmodes of multivariate autoregressive models', *ACM Trans. Math. Softw.*, 2001, **27**, pp. 27–57

23 Hasan, K., Hossain, J., Haque, A.: 'Parameter estimation of multichannel autoregressive processes in noise', *Signal Process.*, 2003, **83**, pp. 603–610

24 Mahmoudi, A., Karimi, M.: 'Estimation of the parameters of multichannel autoregressive signals from noisy observations', *Signal Process.*, 2008, **88**, pp. 2777–2783

25 Petitjean, J., Grivel, E., Bobillet, W., Roussilhe, P.: 'Multichannel AR parameter estimation from noisy observations as an errors-in-variables issue', *Signal Image Video Process. J. (SIViP), Springer*, 2010, **4**, (2), pp. 209–220

26 Cai, J., Shen, X., Mark, J.W.: 'Robust channel estimation for OFDM wireless communication systems – an H_∞ approach', *IEEE Trans. Wirel. Commun.*, 2004, **3**, pp. 2060–2071

27 Shen, X., Deng, L.: 'Game theory approach to discrete H_∞ filter design', *IEEE Trans. Signal Process.*, 1997, **45**, pp. 1092–1095

28 Shen, X., Deng, L.: 'A dynamic system approach to speech enhancement using the H_∞ filtering algorithm', *IEEE Trans. Speech Audio Process.*, 1999, **7**, pp. 391–399

29 Yaesh, I., Shaked, U.: 'Game theory approach to optimal linear state estimation and its relation to the minimum H_∞ – norm estimation', *IEEE Trans. Autom. Control*, 1992, **37**, pp. 828–831

30 Baddour, K.E., Beaulieu, N.C.: 'Accurate simulation of multiple cross-correlated rician fading channels', *IEEE Trans. Commun.*, 2004, **52**, pp. 1650–1662

doi: 10.1049/iet-spr.2010.0066