Self-Adaptive Stochastic Rayleigh Flat Fading Channel Estimation

Robin Gerzaguet, Laurent Ros, Jean-Marc Brossier, Soukayna Ghandour - Haidar, Fabrice Belvèze

Self-Adaptive Stochastic Rayleigh Flat Fading Channel Estimation

Robin GERZAGUET*†, Laurent ROS*, Jean-Marc BROSSIER*, Soukayna GHANDOUR-HAIDAR*, Fabrice BELVEZE†
*GIPSA-Lab, Image and Signal Department, BP 46, 38402 Saint Martin d'Hères, France
†ST-Ericsson, Grenoble, France, robin.gerzaguet@stericsson.com

***Abstract*—This paper deals with channel estimation over flat fading Rayleigh channel with Jakes' Doppler Spectrum. Many estimation algorithms exploit the time-domain correlation of the channel by employing a Kalman filter based on a first-order (or sometimes second-order) approximation of the time-varying channel with a criterion based on correlation matching (CM), or on the Minimization of Asymptotic Variance (MAV). In this paper, we first consider a reduced complexity approach based on Least Mean Square (LMS) algorithm, for which we provide closed-form expressions of the optimal step-size coefficient versus the channel state statistic (additive noise power and Doppler frequency) and of corresponding asymptotic mean-squared-error (MSE). However, the optimal tuning of the step-size coefficient requires knowledge of the channel's statistic. This knowledge was also a requirement for the aforementioned Kalman-based methods. As a second contribution, we propose a self-adaptive estimation method based on a stochastic gradient which does not need *a priori* knowledge. We show that the asymptotic MSE of the self-adaptive algorithm is almost the same as the first order Kalman filter optimized with the MAV criterion and is better than the latter optimized with the conventional CM criterion. We finally improve the speed and reactivity of the algorithm by computing an adaptive speed process leading to a fast algorithm with very good asymptotic performance.**

***Index Terms*—Channel estimation, Rayleigh flat fading channel, Jakes' spectrum, Adaptive LMS.**

I. Introduction

Channel estimation is a fundamental task for a wireless communication receiver and Kalman filter (KF) has been mostly used in the past years, concerning a large and various range of systems from MIMO [1] to OFDM systems [2]–[4]. A classical model for wireless communication is the Rayleigh channel with Jakes' Doppler spectrum also called Clarke's model [5], [6]. In the perspective of designing a recursive algorithm, an often used approximation of the channel consists of a first order auto-regressive model (AR1) (as in [1], [2], [7], [8]), combined with a correlation matching (CM) criterion to set the AR1 coefficient. Thus for a given normalized Doppler frequency f_dT, this coefficient is equal to the standard Bessel AR1 coefficient $J_0(2\pi f_dT)$ (see section III), and the performance of this algorithm is quasi-optimal for high mobility case (see [4]). But, for most conventional Doppler frequencies, where channel variation within one symbol duration can be neglected, (*ie* $f_dT < 10^{-2}$) the performance of this estimator is relatively poor compared to less complex algorithms (see [9]), and is far from the Bayesian Cramer-Rao Bound. The analytical analysis of this poor performance is proved and explained in [10] and [11], where a more suitable criterion based on the minimisation of the asymptotic variance (MAV) is also proposed. On the other hand, other adaptive algorithms can be obtained in using constant coefficients (versus time-varying Kalman gain in Kalman filter) if an *a priori* model of the dynamic of time-varying parameters is available ([12], [13]). However, all these methods require *a priori* knowledge of the statistics of the channel (Doppler frequency and observation noise variance), at least in their optimized versions.

The goal of this paper is to address the aforementioned issues. We develop and analyse self-adaptive estimation methods for Rayleigh flat fading channel, with almost same asymptotic performance than the Kalman of the literature based on the MAV criterion, considered as an *a priori* algorithm (algorithm which works with *a priori* knowledge of the channel statistics). This paper is organized as follows. We give the system model in section II. In section III, we give a brief view of the Kalman approach with two different criteria. In section IV, we present the stochastic approach with analytic analysis and proposition of two self-adaptive algorithms. Section V validates our method by means of simulations results.

II. Model and estimation Objective

We consider the estimation of a flat Rayleigh fading channel. The normalized Doppler frequency of this channel is f_dT, where T is the symbol period. The discrete time observation is:

$$y_k = \alpha_k \cdot s_k + N_k \tag{1}$$

Where k is the symbol index, α_k is the random complex gain of the channel with unit variance ($\sigma_\alpha^2 = 1$), s_k is the data symbol and N_k is a zero mean additive white circular Gaussian noise with variance σ_N^2.

In this work, we concentrate on the performance of the channel estimator. So we treat a simplified case assuming the symbols are known (pilot-aided scenario) or perfectly decided (decision-directed scenario), thus we assume $s_k = 1$. Note that in practice, our channel estimator can easily be coupled with a detector in order to perform joint channel estimation and decision tasks, for example via the Expectation-Maximization algorithm framework (see [3]), or can be used to track the channel gain at pilot frequencies in an OFDM system as in [2].

We consider a Jakes' Doppler Spectrum for the channel gain, also called Clarke's model given by:

$$\Gamma_\alpha(f) = \begin{cases} \frac{\sigma_\alpha^2}{\pi f_d \sqrt{1-\left(\frac{f}{f_d}\right)^2}} & \text{if } |f| < f_d \\ 0 & \text{if } |f| > f_d \end{cases} \quad (2)$$

The autocorrelation coefficient $R_\alpha[m]$ of the stationary complex gain α is then defined for lag m by:

$$R_\alpha[m] = \mathrm{E}\{\alpha_{(n)}.\alpha^*_{(n-m)}\} = \sigma_\alpha^2 J_0(2\pi f_d T.m) \quad (3)$$

where J_0 is the zeroth-order Bessel function of the first kind.

Given the observation model (1) and the Doppler spectrum (2) for the dynamic evolution of the gain, we look for an estimation $\hat{a}_{(n)}$ of $\alpha_{(n)}$. The estimation error is defined by

$$e_{(n)} \stackrel{def}{=} \alpha_{(n)} - \hat{a}_{(n)} \quad (4)$$

and we will study the asymptotic mean square error (MSE):

$$\sigma_e^2 \stackrel{def}{=} E\left\{|e_{(n)}|^2\right\} \quad (5)$$

III. Review of the Kalman filter with AR_1 Model

The first approach widely investigated in the literature consists in approximating the complex gain evolution by a first order auto-regressive model [1], [2], [4], [8], [10], [11] defined by:

$$\alpha_k^{AR1} = \gamma \alpha_{k-1}^{AR1} + b_k \quad (6)$$

where k is the symbol index, $0 < \gamma < 1$ and b_k a circular complex white additive noise with $\sigma_b^2 = (1-\gamma^2)\sigma_\alpha^2$.
The γ coefficient is obtained by [10]:

$$\gamma = \frac{R_{\alpha^{AR_1}}[1]}{R_{\alpha^{AR_1}}[0]} \quad (7)$$

From (6), a Kalman filter can be designed [10]:

$$K_k = \frac{\gamma^2 P_{k-1} + \sigma_b^2}{\gamma^2 P_{k-1} + \sigma_b^2 + \sigma_N^2} \quad (8)$$

$$P_k = (1-K_k)(\gamma^2 P_{k-1} + \sigma_b^2) \quad (9)$$

$$\hat{a}_k = \gamma \hat{a}_{k-1} + K_k(y_k - \gamma \hat{a}_{k-1}) \quad (10)$$

where K_k is the Kalman gain and P_k is the prediction error variance.
The main issue is to link the algorithm to a value of γ. A usual choice in literature for an AR_1 model is to use *Correlation Matching* criterion [1], [2], [4], [7] to match real autocorrelation function and AR autocorrelation function for lag 0 and 1: $R_{\alpha^{AR_1}}[0] = R_\alpha[0] = \sigma_\alpha^2$ and $R_{\alpha^{AR_1}}[1] = R_\alpha[1]$. With (7) and (3) we obtain:

$$\gamma_{CM} = J_0(2\pi f_d T) \quad (11)$$

This algorithm is called $\mathrm{AR1_{CM}}$-KF [14]. For conventional Doppler, its asymptotic performance is not good enough and [10] and [11] propose to change the CM criterion for the minimum of asymptotic variance (MAV) criterion. The optimal γ and its corresponding variance are then [10], [14]:

$$\gamma_{MAV} = \sqrt{1 - 4\sqrt[3]{(\pi f_d T)^4 \frac{\sigma_N^2}{\sigma_\alpha^2}}} \quad (12)$$

$$\sigma_e^2(AR1_{MAV}) \simeq \frac{3}{2}\left(\sigma_{\alpha^2}\right)^{\frac{1}{3}}\left(\pi f_d T \sigma_N^2\right)^{\frac{2}{3}} \quad (13)$$

The Kalman tuned with this MAV criterion, called $\mathrm{AR1_{MAV}}$-KF, offers better asymptotic performance than the previous one (cf [10]).

But it should be noted that the use of Kalman filters exhibits a certain complexity, and requires the knowledge of the channel statistics (see (11) or (12)). It is why we propose alternative methods in the next section.

IV. Stochastic gradient approach

A. LMS approach: $\mathrm{O1_{MAV}}$-F algorithm

In this part we do not consider an AR_1 model combined with a Kalman filter anymore, but a Least Mean Square approach to get $\hat{a}_k$, estimate of α_k [15]. The chosen cost function is:

$$J(\hat{a}) = E\left[|y_k - \hat{a}|^2\right] \quad (14)$$

A stochastic gradient approach using (14) leads to the update rule for the estimate:

$$\hat{a}_k = \hat{a}_{k-1} + \mu(y_k - \hat{a}_{k-1}) \quad (15)$$

The algorithm is denoted $\mathrm{O1_{MAV}}$-F.

It can also be regarded as a simplification of the Kalman estimation (10) in setting $\gamma = 1$, which means replacing the AR_1 model (6) by a Brownian model (also called random walk model) for the evolution of α (an integrated Brownian model could also be used as in [12], [13] or in [16]-section 4.1). But here a fixed gain $K_k = \mu$ is used as in a steady-state mode of the time varying Kalman filter. Indeed, it is well known that the LMS algorithm can be derived as a steady-state version of a Kalman Filter based on a Brownian model [12]. We get then in eq. (15) a simple first-order time-invariant filter, such that in Z-domain:

$$\hat{a}(z) = L(z)y(z) \quad (16)$$

where L is the low pass filter with its cut-off normalized frequency $f_c T = \frac{\mu}{2\pi}$:

$$L(z) = \frac{\mu}{1-(1-\mu)z^{-1}} \quad (17)$$

Using (1) and (4) with (16) and (17), we can express the estimation error in Z domain as

$$e(z) = (1 - L(z)) \cdot \alpha(z) - L(z) \cdot N(z) \quad (18)$$

from which we can derive the asymptotic MSE in frequency domain as:

$$\begin{aligned} \sigma_e^2 &= \int_{-\frac{1}{2T}}^{\frac{1}{2T}} |1 - L(e^{2j\pi fT})|^2 \cdot \Gamma_\alpha(f) df \\ &+ \sigma_N^2 T \int_{-\frac{1}{2T}}^{\frac{1}{2T}} |L(e^{2j\pi fT})|^2 df \end{aligned}$$

Using (17), the second term on the right-hand side, only due to observed white noise, is equal to $\sigma_N^2 \cdot \frac{\mu}{2}$. Using (2) and by assuming that $f_d T \ll 1$ and that the cut-off frequency is properly adjusted ($2\pi f_d T \leq \mu \ll 1$), the overall estimation MSE can be approximated by:

$$\sigma_e^2 \approx \frac{1}{2} \cdot \left(\frac{2\pi f_d T}{\mu}\right)^2 \cdot \sigma_\alpha^2 \quad + \quad \sigma_N^2 \cdot \frac{\mu}{2} \tag{19}$$

Minimizing (19) we obtain the expression of the optimal gain μ_{MAV} and the corresponding minimal variance:

$$\mu_{MAV} = 2(\pi f_d T)^{\frac{2}{3}} \cdot \left(\frac{\sigma_\alpha^2}{\sigma_N^2}\right)^{\frac{1}{3}} \tag{20}$$

$$\sigma_e^2(\mathrm{O1_{MAV}}) = \frac{3}{2}(\sigma_{\alpha^2})^{\frac{1}{3}}(\pi f_d T \sigma_N^2)^{\frac{2}{3}} \tag{21}$$

It is noticeable that the expression of the variance in tracking mode is the same as the $\mathrm{AR1_{MAV}}$-KF (12) reported in the literature, assuming the same *a priori* knowledge of the channel statistics.

B. Self Adaptive algorithm: $\mathrm{O1_{AUTO}}$-F

The three algorithms $\mathrm{AR1_{CM}}$-KF , $\mathrm{AR1_{MAV}}$-KF and $\mathrm{O1_{MAV}}$-F are called *a priori algorithms* as they depend on the knowledge of statistics of the channel (which might require an additional estimation process). In this part, we focus on a self-adaptive estimation.
A stochastic descent on (15) is performed. We choose

$$J_2(\mu) = \lim_{k\to\infty} E\left[|y_k - \hat{a}_k|^2\right] \tag{22}$$

as cost function, the gradient descent in the stochastic approximation is :

$$\mu_k = \mu_{k-1} - \epsilon \cdot \nabla_\mu J_2'(\mu, k)_{|k-1} \tag{23}$$

where $J_2'(k,\mu) = |y_k - \hat{a}_k|^2$ and the gain update is as follow (cf appendix A):

$$\mu_k = \mu_{k-1} + 2\epsilon \cdot \Re\left[(y_k - \hat{a}_k)\frac{\delta \hat{a}_k^*}{\delta \mu}\right]_{|k-1} \tag{24}$$

with k the symbol index, μ_k the gain at iteration k, ϵ a constant step, $()^*$ the conjugate and $\Re()$ the real part of complex value. If we denote $G_k = \frac{\partial \hat{a}_k}{\partial \mu}$, the G_k coefficient can be obtained by deriving (15):

$$G_k = (1 - \mu_{k-1})G_{k-1} + (y_k - \hat{a}_{k-1}) \tag{25}$$

The stochastic self-adaptive algorithm, called $\mathrm{O1_{AUTO}}$-F, becomes:

$$\begin{aligned} \hat{a}_k &= \hat{a}_{k-1} + \mu_{k-1}(y_k - \hat{a}_{k-1}) \\ G_k &= (1 - \mu_{k-1})G_{k-1} + (y_k - \hat{a}_{k-1}) \\ \mu_k &= \mu_{k-1} + \epsilon \cdot \Re\left[(y_k - \hat{a}_{k-1})G_{k-1}^*\right] \end{aligned} \tag{26}$$

C. Self Adaptive speed algorithm: $\mathrm{O1_{AUTO2}}$-F

One of the drawback with stochastic gradient approach is that the speed-accuracy compromise for μ is set by ϵ. If having a too important value of ϵ will seriously decrease asymptotic accuracy, dimensioning a too small value of ϵ creates two problems:

- A long transitional regime before convergence,
- A strong inertia: $\mathrm{O1_{AUTO}}$-F will not react to a modification of statistics of the channel (modification of Doppler frequency, SNR...).

To ensure a fast and reactive algorithm with less asymptotic error, we propose to add an adaptive speed to the previous algorithm (26). The same cost function as previously is used:

$$J_3(\epsilon) = \lim_{k\to\infty} E\left[|y_k - \hat{a}_k|^2\right] \tag{27}$$

However, it should be noted that the adaptive algorithms deriving from (22) and (27) will not have the same trajectory, as we consider only an adaptive gain versus an adaptive gain and an adaptive speed in the second case. By adding an other adaptive system, some additional derivative components appear, all initialized to zero, and we define:

$$N_k = \frac{\delta \hat{a}_k}{\delta \epsilon},\ M_k = \frac{\delta G_k}{\delta \epsilon},\ L_k = \frac{\delta \mu_k}{\delta \epsilon} \tag{28}$$

We propose a original update for the speed coefficient to have a better convergence of the gain μ_k. This update is a combination of a multiplicative procedure [17], [18] and the use of an exponential forgotten factor ζ with $\zeta < 1$ and ζ near of 1 [19].

$$\epsilon_k = \epsilon_{k-1} \cdot (\zeta - \lambda \Re((y_k - \hat{a}_{k-1})N_{k-1}^*) \tag{29}$$

Where λ is a constant step. For convergence reason we commonly bound ϵ factor: $0 < \epsilon_{min} \leq \epsilon_k \leq \epsilon_{max}$. The choice of ϵ_{min} and ϵ_{max} is discussed in section V. Notice that in asymptotic mode, the cost function represented by $\Re((y_k - \hat{a}_{k-1})N_{k-1}^*)$ is close to zero, and then the adaptive coefficient converges to ϵ_{min} with a geometric descent ($\epsilon_k \approx \zeta \cdot \epsilon_{k-1}$).

Considering derivation of the three equations of (26), we obtain the adaptive algorithm with adaptive speed $\mathrm{O1_{AUTO2}}$-F:

Channel estimation

$$\hat{a}_k = \hat{a}_{k-1} + \mu_{k-1}(y_k - \hat{a}_{k-1})$$

Gain estimation

$$G_k = (1 - \mu_{k-1})G_{k-1} + (y_k - \hat{a}_{k-1})$$
$$\mu_k = \mu_{k-1} + \epsilon_{k-1} \cdot \Re\left[(y_k - \hat{a}_{k-1})G_{k-1}^*\right]$$

Adaptive speed (30)

$$\begin{aligned} N_k &= N_{k-1} + L_{k-1}(y_k - \hat{a}_{k-1}) - \mu_{k-1}N_{k-1} \\ L_k &= L_{k-1} + \Re((y_k - \hat{a}_{k-1})G_{k-1}^*) \\ &\quad + \epsilon_{k-1}\Re((y_k - \hat{a}_{k-1})M_{k-1}^*) - \epsilon_{k-1}\Re(G_{k-1}^* N_{k-1}) \\ M_k &= (1 - \mu_{k-1})M_{k-1} - L_{k-1}G_{k-1} - N_{k-1} \\ \epsilon_k &= \left[\epsilon_{k-1} \cdot (\zeta - \lambda \Re((y_k - \hat{a}_{k-1})N_{k-1}^*)\right]_{\epsilon_{min}}^{\epsilon_{max}} \end{aligned}$$

Where $[\epsilon]_{\epsilon_{min}}^{\epsilon_{max}}$ means that the ϵ value is bound between ϵ_{min} and ϵ_{max}. Asymptotically, due to the geometric multiplicative update of the adaptive speed, $\epsilon_k = \epsilon_{min}$, and then the algorithm described in (30) is equivalent to (26).

The three algorithms ($\mathrm{O1}_{\mathrm{MAV}}$-F, $\mathrm{O1}_{\mathrm{AUTO}}$-F and $\mathrm{O1}_{\mathrm{AUTO2}}$-F) have the same structure, but only the last two are totally self-adaptive without *a priori* knowledge. Besides $\mathrm{O1}_{\mathrm{AUTO2}}$-F is an improvement of $\mathrm{O1}_{\mathrm{AUTO}}$-F as the speed adaptivity is an overlay which aims at increasing the speed and reactivity of the stochastic algorithm.

V. SIMULATION

In this section the proposed algorithms of the section IV are compared in terms of asymptotic performance, convergence speed and convergence accuracy to the algorithms reported into literature ($\mathrm{AR1}_{\mathrm{CM}}$-KF [1]–[4] and $\mathrm{AR1}_{\mathrm{MAV}}$-KF [10], [11]) and in section III. For all our simulations, the channel autocorrelation function is assumed to be given by the widely accepted Jakes' model, as stated in section II. Excepted for the last figure 5, the channel estimator is tested assuming the knowledge of the data (pilot aided scenario, with $s_k = 1$).

First, for standard SNR (20dB) and normalized frequency f_dT (10^{-3}), fig.1 shows that $\mu(k)$ converges to μ_{MAV} and that adaptive speed algorithm is faster and more accurate than the constant step solutions. The bound of ϵ can be chosen in a wide range and ϵ_{max} is designed to ensure a fast initial convergence (ϵ_{max}=1). On the other hand, ϵ_{min} can be very small to reduce convergence error, and in the simulation, $\epsilon_{min} = 10^{-5}$.

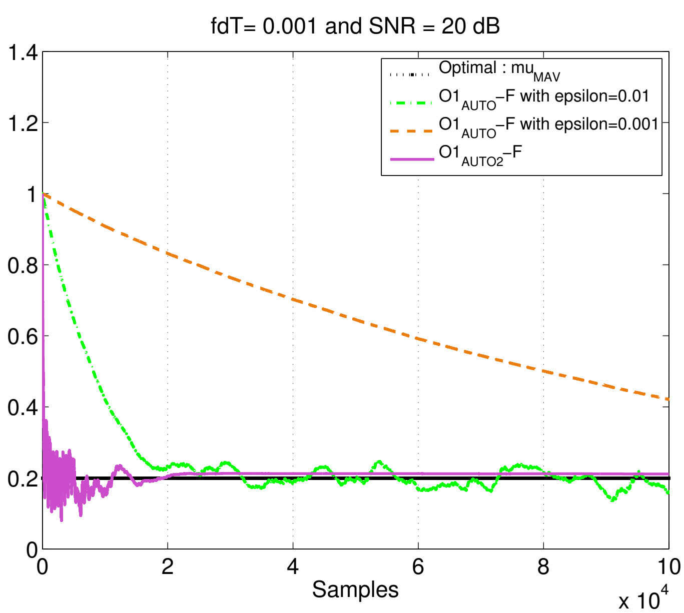

Fig. 1. *Step-size μ for SNR= 20dB and $f_dT = 10^{-3}$*

Fig.2 gives the comparison between asymptotic MSE versus SNR for a normalized Doppler frequency $f_dT = 10^{-3}$ and the Bayesian Cramer Rao Bound (BCRB) ([20]). It shows that the MAV-based algorithms have better performance than the Kalman filter using CM criterion [1], [2] which corroborates the results of [10], [11] and [16]. It is above all observed that self-adaptive algorithms have almost the same asymptotic performance as the *a priori* algorithms. Thus, using self-adaptive algorithms without *a priori* information does not decrease the method efficiency. Besides, the asymptotic performance of the optimal step size algorithm $\mathrm{O1}_{\mathrm{MAV}}$-F directly validates the theoretical expressions (20) and (21) as it reaches the bound of the Kalman-MAV Filter expressed in (13) for slow normalized Doppler frequency.

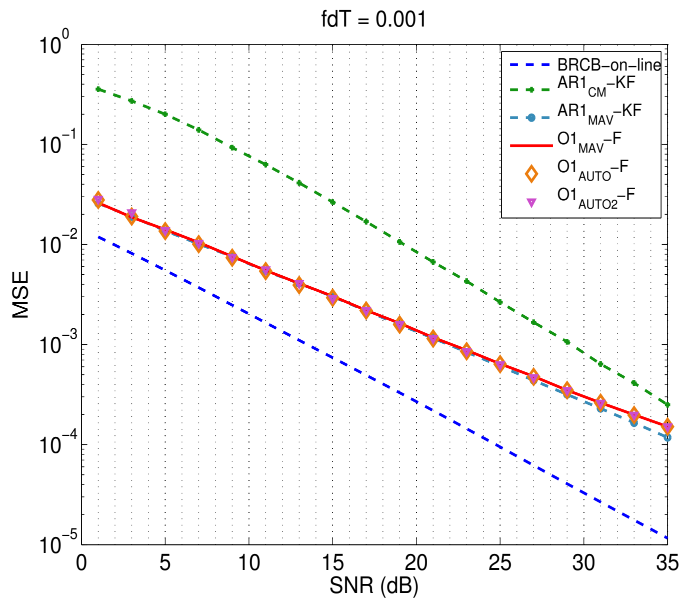

Fig. 2. *MSE versus SNR, for $f_dT = 10^{-3}$*

Fig.3 compares MSE versus normalized Doppler frequency f_dT for a given SNR of 20dB. As described in [10] and [16], for low normalized Doppler frequency, the MAV algorithms have better behaviour than the Kalman filter with CM criterion. Besides, the self-adaptive algorithms maintains the same asymptotic performance as the *a priori* algorithm.

We finally consider the transitional Mean Squared Error (in dB) of the estimate of the channel in Figure 4, for a normalized Doppler frequency of 10^{-3} and a 20dB SNR, for the 500 first samples. The asymptotic performance of the different methods are almost the same, which is coherent with figure 3 except for the $\mathrm{O1}_{\mathrm{AUTO}}$-F which is still in transitional mode. The Kalman Filter based on the MAV criterion has best performance in terms of speed and accuracy, due to its *a priori* knowledge. The $\mathrm{O1}_{\mathrm{AUTO2}}$-F is very fast at the beginning of the simulation (faster than the optimal step size algorithm and comparable to the Kalman Filter) but then reached the asymptotic level with a slower speed due to the oscillations of the step size (cf figure 1). The design of the adaptive speed algorithm joins initial fast speed and good asymptotic mode, and this without any *a priori* knowledge.

Until now, all the performances are dressed in terms of MSE. Finally, we have a look at the consequences in terms

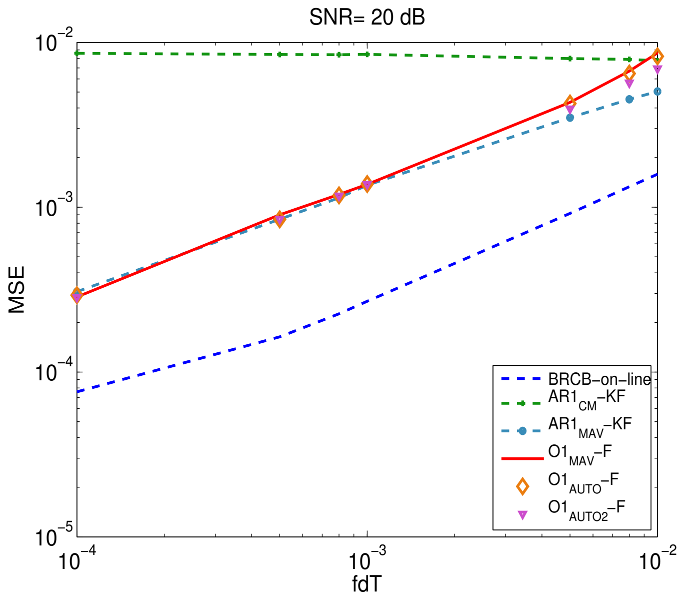

Fig. 3. *MSE versus f_dT, for SNR=20dB*

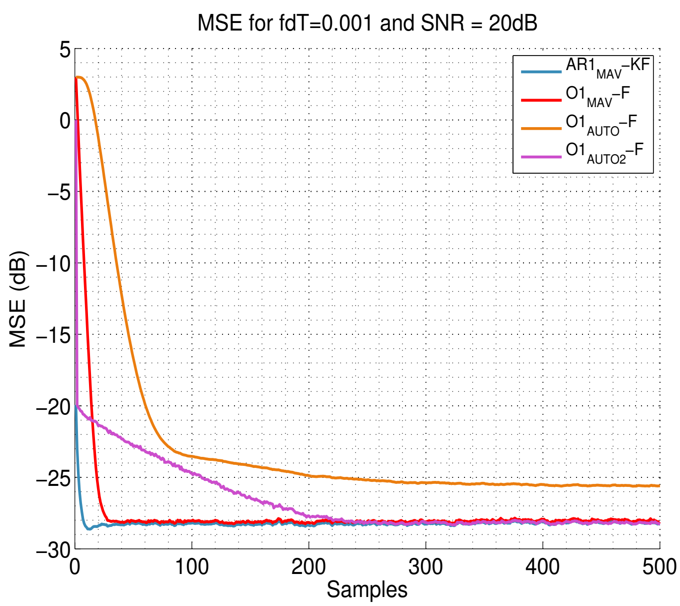

Fig. 4. *Evolution of MSE for the 500 first samples and different algorithms: The MAV-Kalman filter AR1$_{MAV}$-KF, the optimal step size algorithm O1$_{MAV}$-F, the constant speed algorithm O1$_{AUTO}$-F and the adaptive speed algorithm O1$_{AUTO2}$-F.*

of Bit Error Rate (BER) in figure 5. This figure refers to a binary PSK (BPSK) transmitted over a Rayleigh fading channel with $f_dT = 10^{-3}$ versus different values of the SNR. In order to ensure convergence of the iterative detection and decoding algorithm, 10 pilot symbols (known at the receiver) are inserted every 100 transmitted symbols [1]. So we use a semi-blind pilot assisted channel estimation. The BER results agree with the previous MSE performance[2]. We show that self adaptive methods have same BER performance as *a priori* methods optimized with the MAV criterion and outperform the conventional Kalman filter when the latter is based on the CM criterion.

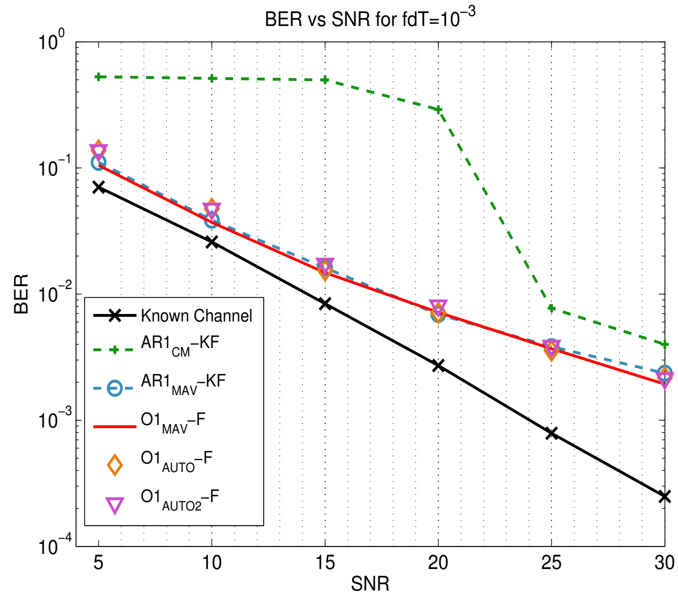

Fig. 5. *BER in function of SNR for $f_dT = 10^{-3}$ for the different algorithms : The CM-Kalman Filter AR1$_{CM}$-KF, The MAV Kalman filter AR1$_{MAV}$-KF, the optimal step size algorithm O1$_{MAV}$-F, the constant speed algorithm O1$_{AUTO}$-F and the adaptive speed algorithm O1$_{AUTO2}$-F.*

VI. Conclusion

This study focuses on the estimation methods of the complex gain of a flat fading Rayleigh channel with Jakes' Doppler spectrum based on first order approximation models. We first have checked that first order auto-regressive Kalman filter, tuned by the traditional Correlation Matching criterion (leading to the AR1$_{CM}$-KF) can be significantly improved for low normalized Doppler frequency by using a more appropriate *Minimum Asymptotic Variance* criterion (AR1$_{MAV}$-KF). Next, we showed that a LMS approach can have the same asymptotic performance than the AR1$_{MAV}$-KF when the LMS is tuned with the optimal step-size for which we provide analytical expression. Besides,we have designed a self-adaptive stochastic gradient algorithm. This algorithm is recursive and has almost

[1]To use the KF equations in (8), (9), (10), the O1$_{MAV}$-F in (15) the O1$_{AUTO}$-F in (26) and the O1$_{AUTO\text{-}2}$-F in (30) which are given for known symbols case, i.e. for $s_k = 1$ in (1), and since we have $s_k \in \{1;+1\}$, the different equations are modified by substituting y_k with $y_k \times \hat{s}^*_{k|k-1}$ where $\hat{s}^*_{k|k-1} = s_k$ if s_k is known (pilot) or $\hat{s}^*_{k|k-1} = sgn\{\Re(\alpha^*_{k-1} \times y_k)\}$ if s_k is unknown (data). In this case, $\hat{s}^*_{k|k-1}$ represents the *a priori* decision and the final decision will be $\hat{s_k} = sgn\{\Re(\hat{\alpha}^*_k \times y_k)\}$.

[2]In this semi-blind scenario, we may have error propagation effect, leading to MSE performance different from formulae established for pilot-aided scenario.

the same asymptotic performance as the AR1 $_{mboxMAV}$-KF. As this solution has a very good asymptotic mode but a low transitional speed, we add an adaptive speed factor ϵ_k to our algorithm leading to a fast and accurate solution. Thus the two last algorithms are self-adaptive and work without any knowledge of the channel statistics.

APPENDIX A
GRADIENT CALCULATION IN (24)

Considering the cost function defined in (22), a gradient is applied versus a real variable, equal to the step μ which belongs to $\mathbb{R}$. We use the stochastic approximation [17] (*ie* we use the instantaneous value of the estimation error $J_2'(\mu,k) = |y_k - \hat{a}_k|^2$ instead of $J_2(\mu) = \lim_{k\to\infty} E\left[|y_k - \hat{a}_k|^2\right]$), and the update of the gain is expressed as :

$$\begin{aligned} \mu_k &= \mu_{k-1} - \epsilon \cdot \nabla_\mu J_2'(\mu,k)_{|k-1} \\ \mu_k &= \mu_{k-1} - \epsilon \cdot \nabla_\mu \left(|y_k - \hat{a}_k|^2\right)_{|k-1} \end{aligned} \tag{31}$$

We denote $\nabla_\mu(\cdot) = \frac{\partial(\cdot)}{\partial\mu}$, $c(k) = y(k) - \hat{a}(k)$, with $c_I(k)$ the real part and $c_Q(k)$ the imaginary part of $c(k)$, and we express the gradient of $c(k)$ respect to μ:

$$\begin{aligned} \nabla_\mu |c(k)|^2 &= \nabla_\mu c(k) \cdot c(k)^* \\ &= c(k)\nabla_\mu c(k)^* + c(k)^*\nabla_\mu c(k) \\ &= (c_I(k) + jc_Q(k))\nabla_\mu(c_I(k) - jc_Q(k)) \\ &+ (c_I(k) - jc_Q(k))^*\nabla_\mu((c_I(k) - jc_Q(k)) \\ &= 2c_I(k)\nabla_\mu c_I(k) + 2c_Q(k)\nabla_\mu c_Q(k) \end{aligned}$$

Besides,

$$\begin{aligned} 2\Re((c(k)\nabla_\mu c^*(k)) &= 2\Re\left((c_I(k) + jc_Q(k))\nabla_\mu(c_I(k) - jc_Q(k))\right) \\ &= 2c_I(k)\nabla_\mu c_I(k) + 2c_Q(k)\nabla_\mu c_Q(k) \end{aligned}$$

Which leads to the expression of the gradient of a modulus

$$\nabla_\mu |c(k)|^2 = 2\Re\left(c(k)\nabla_\mu c^*(k)\right) \tag{32}$$

which, in our case as the observation is independent from the step leads to the iterative expression of the step :

$$\mu_k = \mu_{k-1} + 2\epsilon \cdot \Re\left[(y_k - \hat{a}_k)\frac{\delta\hat{a}_k^*}{\delta\mu}\right]_{|k-1} \tag{33}$$

REFERENCES

[1] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using kalman estimation," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[2] W. Chen and R. Zhang, "Kalman-filter channel estimator for OFDM systems in time and frequency-selective fading environment," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 4, may 2004, pp. iv–377 – iv–380 vol.4.

[3] T. Y. Al-Naffouri, "An EM-Based Forward-Backward kalman filter for the estimation of Time-Variant channels in OFDM," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3924–3930, Jul. 2007.

[4] H. Hijazi and L. Ros, "Joint data QR-detection and kalman estimation for OFDM time-varying rayleigh channel complex gains," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 170–178, Jan. 2010.

[5] W. Jakes, "Microwave mobile communications, 1994."

[6] R. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J*, vol. 47, no. 6, pp. 957–1000, 1968.

[7] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1650–1662, Jul. 2005.

[8] C. C. Tan and N. C. Beaulieu, "On first-order markov modeling for the rayleigh fading channel," *IEEE Transactions on Communications*, vol. 48, no. 12, pp. 2032–2040, Dec. 2000.

[9] L. Ros, H. Hijazi, and E. P. Simon, "Paths complex gain tracking algorithms for OFDM receiver in slowly-varying channels," in *2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Limassol, Cyprus, Mar. 2010.

[10] S. Ghandour-Haidar, L. Ros, and J. M. Brossier, "On the use of first-order autoregressive modeling for rayleigh flat fading channel estimation with kalman filter," *Signal Processing*, vol. 92, no. 2, pp. 601–606, Feb. 2012.

[11] A. Barbieri, A. Piemontese, and G. Colavolpe, "On the ARMA approximation for fading channels described by the clarke model with applications to kalman-based receivers," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 535–540, Feb. 2009.

[12] L. Lindbom, A. Ahlen, M. Sternad, and M. Falkenstrom, "Tracking of time-varying mobile radio channels. ii. a case study," *Communications, IEEE Transactions on*, vol. 50, no. 1, pp. 156 –167, jan 2002.

[13] Q. Nasir, "Wireless channel tracking based on optimum predictive lms," *Wirel. Pers. Commun.*, vol. 48, no. 4, pp. 511–519, Mar. 2009. [Online]. Available: http://dx.doi.org/10.1007/s11277-008-9535-9

[14] L. Ros, J. M. Brossier, and S. Ghandour-Haidar, "Sur l'utilisation des filtres du premier-ordre pour l'estimation d'un canal radio-mobile de rayleigh à évanouissement plat," in *23eme colloque Gretsi*, Bordeaux, France, 2011.

[15] J. M. Brossier, *Signal et communication numérique, égalisation et synchronisation*. Hermes, 1997.

[16] L. Ros and E. P. Simon, "Second-order modeling for rayleigh flat fading channel estimation with kalman filter," in *2011 17th International Conference on Digital Signal Processing (DSP)*, Corfu, Greece, Jul. 2011.

[17] A. Benveniste and P. Metivier, *Adaptive algorithms and stochastic approximations*. Springer, 1991.

[18] W. A. and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 805–810, Apr. 2001.

[19] S.-L. Goh and D.-P. Mandic, "Stochastic Gradient-Adaptive Complex-Valued nonlinear neural adaptive filters with a Gradient-Adaptive step size," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1511–1516, Sep. 2007.

[20] H. Hijazi and L. Ros, "Bayesian Cramer-Rao bounds for complex gain parameters estimation of slowly varying rayleigh channel in ofdm systems," *Signal Processing*, vol. 89, no. 1, pp. 111 – 115, 2009.