SAGE-based Estimation Algorithms for Time-varying Channels in Amplify-and-Forward Cooperative Networks

Nico Aerts and Marc Moeneclaey
Department of Telecommunications
and Information Processing
Ghent University
9000 Gent, Belgium
Email: firstname.lastname@telin.ugent.be

***Abstract*—Cooperative communication is a technique that achieves spatial diversity by exploiting the presence of other nodes in the network. Most analyses of such networks are conducted under the simplifying assumption of perfect channel knowledge. In this paper we focus on the popular Amplify-and-forward (AF) cooperative protocol. We propose several SAGE-based iterative algorithms with different complexities for estimating the channel gain and noise variance in the case of time-varying channels. Computer simulations are provided to evaluate their performance over Rice-fading channels. We point out a low-complexity estimation algorithm yielding an error performance that (for Rayleigh fading) is only about 0.5 dB worse than in the case of perfect estimation, while outperforming a pilot-based estimation algorithm by about 1.5 dB.**

I. Introduction

Several diversity techniques [1] have been proposed to counteract the effect of fading on the performance of wireless communication systems. By exploiting the presence of other nodes in the network, the cooperative communication model [2], [3] can achieve spatial diversity, which is particularly interesting for nodes that can not be equipped with multiple antennas. The allocated time frame to transmit information to the destination, is divided into $1+n_r$ slots, of which the source uses only the first. The n_r remaining slots are used by n_r other nodes, which we will refer to as relays, to transmit to the destination signals that are related to the information sent by the source. Several cooperative protocols are proposed in literature [2]; here we will consider the Amplify-and-Forward protocol, where the relays simply retransmit an amplified version of the the signal they receive.

In literature one often assumes the channel parameters to be known by the destination when investigating the error performance of cooperative communications. However, in real-life scenarios these parameters have to be estimated. In [4] a Linear Minimum Mean Square Error (LMMSE) channel estimate based on pilot symbols is proposed for time-varying Rayleigh-fading channels in AF networks. We extend this work by (a) considering time-varying Rice-fading channels, and (b) deriving iterative estimation algorithms of various complexity that also exploit the data-part of the received signals for channel estimation.

This paper is organized as follows. The model of the AF system is described in Section II. Section III presents the derivation of the pilot-aided and of the iterative channel estimation algorithms. Numerical results regarding the estimator performance and the frame error rate are given in Section IV. Finally, conclusions are drawn in section V.

Notations

All vectors are row vectors, unless mentioned otherwise, and boldface; the Hermitian transpose, statistical expectation and estimate of the row vector $\boldsymbol{x}$ are denoted by $\boldsymbol{x}^H$, $E[\boldsymbol{x}]$ and $\hat{\boldsymbol{x}}$ respectively; $(\mathbf{X})(k,m)$ is the (k,m)th element of the matrix $\mathbf{X}$; the complex conjugate of a scalar x is denoted x^*.

II. System description

In this paper we consider a network, depicted in Fig. 1, with one relay; the extension to multiple relays is straightforward. The channels are affected by time-varying Rice

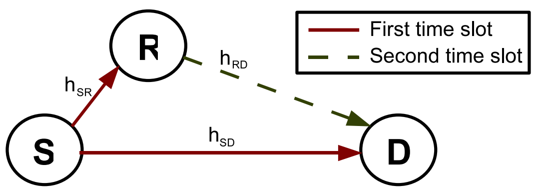

Fig. 1. Network containing a source S, a relay R and a destination D.

fading, and are characterized by the independent channel coefficient vectors $\boldsymbol{h}_i$ and the independent noise vectors $\boldsymbol{w}_i$ ($i \in \{SD, SR, RD\}$). Here the elements of the noise vectors $\boldsymbol{w}_i$ are independent zero-mean circular symmetric complex Gaussian (ZMCSCG) random variables with variances $N_{0,i}$. The subscripts SR, SD and RD refer to the source-relay, source-destination and relay-destination channels, respectively. Denoting the carrier-to-multipath ratio (C/M) of the Rice fading by $\lambda^2/(1-\lambda^2)$ with $\lambda \in [0;1]$, the channel coefficients at epoch k can be decomposed as

$$h_i(k) = \sqrt{H_i}\left(\lambda_i e^{j\theta_i} + \sqrt{1-\lambda_i^2}\, n_i(k)\right),$$

where $n_i(k)$ is ZMCSCG distributed with variance 1, θ_i is uniformly distributed in $(-\pi,\pi)$ and $E[|h_i|^2] = H_i$. The fading components are time-varying and following [5] we take

$$R_{n_i}(m) = E[n_i(k+m)n_i^*(k)] = J_0(2\pi\nu_i m),$$

with $J_0(x)$ the zeroth order Bessel function of the first kind and ν_i the normalized maximum Doppler frequency associated with $h_i(k)$. Hence the discrete-time autocorrelation between the channel coefficients $h_i(k+m)$ and $h_i(k)$ is

$$R_{h_i}(m) = H_i\left(\lambda_i^2 + (1-\lambda_i^2)J_0(2\pi\nu_i m)\right). \quad (1)$$

During the first time slot the source transmits a vector $\boldsymbol{c} = (c(1),...,c(K))$ containing K_d data symbols and K_p pilot symbols. The first and the last symbols of the frame are pilot symbols and the pilot symbols are separated by K_s data symbols (i.e., $K = K_d + K_p$), with $E_s = E[|c(k)|^2]$ denoting the average energy per symbol. Hence, $K_p = K_d/K_s + 1$ and $k_i = 1 + (K_s+1)(i-1)$ denotes the position of the i[th] pilot symbol (i=1,...,K_p). The signals received by the destination and the relay are given by

$$\boldsymbol{r}_i = \boldsymbol{h}_i \mathbf{C} + \boldsymbol{w}_i \qquad i \in \{SD, SR\} \quad (2)$$

where $\mathbf{C}$ is a $K \times K$ diagonal matrix with $c(k)$ as the k[th] diagonal element. The relay amplifies $\boldsymbol{r}_{SR}$ by a constant factor γ and transmits the result to the destination during the second time slot. Imposing the average energy transmitted by the relay at each epoch to be E_r, γ is given by

$$\gamma = \sqrt{\frac{E_r}{H_{SR}E_s + N_{0,SR}}}. \quad (3)$$

With $\odot$ denoting the Hadamard product, the signal received by the destination is given by

$$\begin{aligned}\boldsymbol{r}_{RD} &= \gamma \boldsymbol{h}_{RD} \odot \boldsymbol{r}_{SR} + \boldsymbol{w}_{RD} \\ &= \boldsymbol{h}_{SRD}\mathbf{C} + \boldsymbol{w}_{SRD},\end{aligned}$$

where $\boldsymbol{h}_{SRD} = \boldsymbol{h}_{SR} \odot \boldsymbol{h}_{RD}$ is the channel gain of the compound source-relay-destination (SRD) channel and the elements of $\boldsymbol{w}_{SRD}$ are ZMCSCG distributed with autocorrelation matrix $\mathbf{N}_{SRD}$. The matrix $\mathbf{N}_{SRD}$ is diagonal with the k[th] diagonal element given by $N_{SRD}(k) = \gamma^2|h_{RD}(k)|^2 N_{0,SR} + N_{0,RD}$. Note that the variance $N_{SRD}(k)$ of the noise on the SRD channel depends on the symbol index k through $|h_{RD}(k)|^2$.

When the channels are known by the destination, maximum-ratio combining [6] is applied to the data containing parts of the received signals in order to obtain a sufficient statistic $\boldsymbol{\eta}$ for detecting the data from the source, we have

$$\eta(k) = \frac{r_{SD}(k)h_{SD}^*(k)}{N_{0,SD}} + \frac{r_{RD}(k)h_{SRD}^*(k)}{N_{SRD}(k)}, \quad (4)$$

with k restricted to the data symbol positions.

We assume the noise variances $N_{0,i}$ ($i \in \{SD, SR, RD\}$) to be known as these are long-term properties of the channels. From (4) it is then easily seen that estimates of $\boldsymbol{h}_{SD}$, $\boldsymbol{h}_{SRD}$ and $\boldsymbol{N}_{SRD}$ are needed at the destination in order to calculate the sufficient statistic.

III. Estimation strategies

The Maximum a Posteriori (MAP) estimate of a vector parameter $\boldsymbol{\theta} = (\boldsymbol{h}_{SD}, \boldsymbol{h}_{SRD}, \boldsymbol{N}_{SRD})$ from an observation $\boldsymbol{r} = (\boldsymbol{r}_{SD}, \boldsymbol{r}_{RD})$ in the presence of a nuisance parameter $\boldsymbol{c}$ is obtained after convergence of the appropriate EM-algorithm [7], which produces a sequence of estimates $\hat{\boldsymbol{\theta}}^{(i)}$, $i = 1, 2, ...$:

$$\hat{\boldsymbol{\theta}}^{(i)} = \arg\max_{\boldsymbol{\theta}} \mathbb{E}\left[\ln(p(\boldsymbol{r}|\boldsymbol{\theta},\boldsymbol{c})p(\boldsymbol{\theta}))|\boldsymbol{r}, \hat{\boldsymbol{\theta}}^{(i-1)}\right], \quad (5)$$

where the conditional expectation is with respect to the nuisance parameter $\boldsymbol{c}$. We will consider the SAGE algorithm [8] instead of the EM algorithm as it replaces the multi-dimensional maximization associated with (5) by several maximizations over mutually exclusive subsets of $\boldsymbol{\theta}$: Firstly $\hat{\boldsymbol{h}}_{SD}^{(i)}$ and $\hat{\boldsymbol{h}}_{SRD}^{(i)}$ are derived with $\boldsymbol{N}_{SRD}$ fixed to its previous estimate $\hat{\boldsymbol{N}}_{SRD}^{(i-1)}$. Secondly $\hat{\boldsymbol{N}}_{SRD}^{(i)}$ is derived with $\boldsymbol{h}_{SD}$ and $\boldsymbol{h}_{SRD}$ fixed to $\hat{\boldsymbol{h}}_{SD}^{(i)}$ and $\hat{\boldsymbol{h}}_{SRD}^{(i)}$, respectively. In order to initialize the SAGE iterations an initial estimate $\hat{\boldsymbol{\theta}}^{(0)}$ is required. This initial estimate is considered in section III-A.

Considering that $p(\boldsymbol{\theta})$ and $p(\boldsymbol{r}|\boldsymbol{\theta},\boldsymbol{c})$ can be decomposed as

$$p(\boldsymbol{\theta}) = p(\boldsymbol{h}_{SD})p(\boldsymbol{h}_{SRD}|\boldsymbol{N}_{SRD})p(\boldsymbol{N}_{SRD}) \quad (6)$$
$$p(\boldsymbol{r}|\boldsymbol{\theta},\boldsymbol{c}) = p(\boldsymbol{r}_{SD}|\boldsymbol{h}_{SD},\boldsymbol{c})p(\boldsymbol{r}_{RD}|\boldsymbol{h}_{SRD},\boldsymbol{N}_{SRD},\boldsymbol{c}).$$

It follows from (5) that $\boldsymbol{h}_{SD}$ and $(\boldsymbol{h}_{SRD}, \boldsymbol{N}_{SRD})$ are estimated from $\boldsymbol{r}_{SD}$ and $\boldsymbol{r}_{RD}$, respectively. The SAGE-based estimation of $\boldsymbol{h}_{SRD}$ and $\boldsymbol{N}_{SRD}$ is detailed in sections III-B and III-C, respectively. For the sake of saving space, the estimation of $\boldsymbol{h}_{SD}$ is not considered further; the corresponding expressions can be easily derived from (5).

A. Initial estimates

In literature the estimation of the noise variance $N_{SRD}(k)$ is mostly taken equal to its average $N_{SRD,avg}$ over the fading statistics, and we will do the same for the initial estimate : $\hat{N}_{SRD}^{(0)}(k) = N_{SRD,\text{avg}}$, where $N_{SRD,\text{avg}} = \gamma^2 H_{RD} N_{0,SR} + N_{0,RD}$. The initial estimate $\hat{\boldsymbol{h}}_{SRD}^{(0)}$ is obtained in two steps. Firstly, a pilot-based ML channel coefficient estimate at the position k_i corresponding to the i[th] pilot symbol is computed [9]:

$$\hat{h}_{SRD}^{(0)}(k_i) = \frac{r_{RD}(k_i)c^*(k_i)}{|c(k_i)|^2}.$$

Secondly, by applying LMMSE filtering [10] of the estimates $\left\{\hat{h}^{(0)}_{SRD}(k_i)\right\}$, the estimate $\hat{\boldsymbol{h}}^{(0)}_{SRD}$ is obtained

$$\hat{\boldsymbol{h}}^{(0)}_{SRD} = \hat{\boldsymbol{p}}\mathbf{A}, \tag{7}$$

where $\hat{\boldsymbol{p}}$ denotes the row vector that contains the K_p estimates $\left\{\hat{h}^{(0)}_{SRD}(k_i)\right\}$, $\mathbf{A}$ is a $K_p \times K$ matrix given by

$$\mathbf{A} = \mathbf{R}^{-1}_{\hat{\boldsymbol{p}}\hat{\boldsymbol{p}}}\mathbf{R}_{\hat{\boldsymbol{p}}h},$$

while the $K_p \times K$ matrix $\mathbf{R}_{\hat{\boldsymbol{p}}h}$ and the $K_p \times K_p$ matrix $\mathbf{R}_{\hat{\boldsymbol{p}}\hat{\boldsymbol{p}}}$ are defined as $\mathbf{R}_{\hat{\boldsymbol{p}}h} = E[\hat{\boldsymbol{p}}^H\boldsymbol{h}_{SRD}]$ and $\mathbf{R}_{\hat{\boldsymbol{p}}\hat{\boldsymbol{p}}} = E[\hat{\boldsymbol{p}}^H\hat{\boldsymbol{p}}]$, respectively. It is easily verified that

$$\mathbf{R}_{\hat{\boldsymbol{p}}h}(i,j) = \mathbf{R}_{h_{SRD}}(k_i,j)$$

$$\mathbf{R}_{\hat{\boldsymbol{p}}\hat{\boldsymbol{p}}}(i,j) = \mathbf{R}_{h_{SRD}}(k_i,k_j) + \frac{\gamma^2 H_{RD}N_{0,SR} + N_{0,RD}}{E_s}\delta_{i,j},$$

with $\delta_{i,j}$ denoting the Kronecker delta, and

$$\mathbf{R}_{h_{SRD}}(k,m) = \gamma^2 R_{h_{SR}}(k-m)R_{h_{RD}}(k-m),$$

where the correlation functions $R_{h_{SR}}$ and $R_{h_{RD}}$ are according to (1).

B. SAGE estimation of $\boldsymbol{h}_{SRD}$

As the distribution $p(\boldsymbol{h}_{SRD}|\boldsymbol{N}_{SRD})$ in (6), needed in the first step of the SAGE algorithm to update the estimate of $\boldsymbol{h}_{SRD}$, is difficult to obtain, we will consider several approximation strategies.

1) Strategy S_1 : 'approximate MAP': In strategy S_1, $p(\boldsymbol{h}_{SRD}|\boldsymbol{N}_{SRD})$ is approximated by the marginal distribution $p(\boldsymbol{h}_{SRD})$, which in turn is approximated as a ZMCSCG distribution with autocorrelation matrix $\mathbf{R}_{h_{SRD}}$. In this case one can verify from (5) that $\hat{\boldsymbol{h}}^{(i)}_{SRD,S_1}$ is given by

$$\boldsymbol{r}_{RD}\mathbf{U}^{(i),H}\hat{\mathbf{N}}^{(i-1),-1}_{SRD}\mathbf{R}_{h_{SRD}}\left(\Sigma^{(i)}\hat{\mathbf{N}}^{(i-1),-1}_{SRD}\mathbf{R}_{h_{SRD}} + \mathbf{I}_K\right)^{-1}, \tag{8}$$

where $\mathbf{U}^{(i)}$ and $\Sigma^{(i)}$ are diagonal (KxK) matrices, with the k$^{\text{th}}$ diagonal element given by

$$\mu^{(i)}(k) = E\left[c(k)\left|\boldsymbol{r}_{SR},\boldsymbol{r}_{RD},\boldsymbol{\theta}^{(i-1)}\right.\right], \tag{9}$$

$$\sigma^{(i)}(k) = E\left[|c(k)|^2\left|\boldsymbol{r}_{SR},\boldsymbol{r}_{RD},\boldsymbol{\theta}^{(i-1)}\right.\right]. \tag{10}$$

Both $\boldsymbol{\mu}^{(i)}(k)$ and $\sigma^{(i)}(k)$ are easily derived from the marginal a posteriori probabilities (APPs) of the coded symbols $c(k)$, which are obtainable by message passing on a factor graph. As compared to the pilot based estimates, the computational complexity increases significantly as each iteration requires (a) a decoding step to obtain the APPs and (b) the inversion of a (KxK) matrix (see (8)).

Strategy S_2 : 'reduced complexity approximate MAP': The complexity associated with calculating the inverse of a (KxK) matrix is of the order K^3. To reduce this complexity associated with strategy S_1, we divide the slot of K symbols into N_{it} subslots of K_{it} symbols each (with $K = N_{it}K_{it}$) and assume that the channel gains in a subslot are independent from the gains in other subslots. Application of the approximate MAP strategy to each of the individual subslots requires N_{it} inversions of a $K_{it} \times K_{it}$ matrix, yielding a complexity reduction by a factor $(N_{it})^2$ as compared to strategy S_1. Note that the calculation of $\mu^{(i)}(k)$ and $\sigma^{(i)}(k)$ is still according to (9) and (10).

Strategy S_3 : 'ML': Removing from (5) the a priori distribution $p(\boldsymbol{\theta})$, the resulting EM/SAGE algorithm converges to the maximum-likelihood (ML) estimate. In the i$^{\text{th}}$ iteration, the estimate is given by

$$\hat{\boldsymbol{h}}^{(i)}_{SRD,ML} = \boldsymbol{r}_{RD}\mathbf{U}^{(i),H}\Sigma^{(i),-1}. \tag{11}$$

Note that (11) does not exploit the correlation of the channel coefficients. This ML estimate can be calculated element wise, because $\Sigma^{(i)}$ is diagonal.

Strategy S_4 : 'Filtered ML': Similar to the case of pilot based ML estimates, LMMSE filtering can be applied to the 'ML' estimate from (11) to take the time-correlation of the channel gains into account. This requires the calculation of the correlation matrices $E[\hat{\boldsymbol{h}}^{(i),H}_{SRD,ML}\boldsymbol{h}_{SRD,ML}]$ and $E[\hat{\boldsymbol{h}}^{(i),H}_{SRD,ML}\hat{\boldsymbol{h}}_{SRD}]$. However, due to the presence of $\mu^{(i)}(k)$ and $\sigma^{(i)}(k)$, it is not possible to obtain closed form expressions. We will therefore calculate these correlation matrices under the simplifying assumption that $\mu^{(i)}(k)$ and $\sigma^{(i)}(k)$ can be replaced by $c(k)$ and E_s, respectively; this yields

$$\hat{\boldsymbol{h}}^{(i)}_{SRD,S_4} = \hat{\boldsymbol{h}}^{(i)}_{SRD,ML}\left(\mathbf{R}_{h_{SRD}} + \frac{N_{SRD,\text{avg}}}{E_s}\mathbf{I}_K\right)^{-1}\mathbf{R}_{h_{SRD}}. \tag{12}$$

To interpret the result (12), we rewrite the estimate $\hat{\boldsymbol{h}}^{(i)}_{SRD,S_1}$ from (8) as

$$\hat{\boldsymbol{h}}^{(i)}_{SRD,S_1} = \hat{\boldsymbol{h}}^{(i)}_{SRD,ML}\left(\mathbf{R}_{h_{SRD}} + \hat{\mathbf{N}}^{(i-1)}_{SRD}\Sigma^{(i),-1}\right)^{-1}\mathbf{R}^{-1}_{h_{SRD}}.$$

From the above expressions it follows that the 'filtered ML' estimator is obtained from the 'approximated MAP' estimator when making the assumption that $\hat{\mathbf{N}}^{(i-1)}_{SRD}$ is equal to $N_{SRD,\text{avg}}\Sigma^{(i)}/E_s$.

It is now very important to notice that for computing (12) only the diagonal matrix $\Sigma^{(i)}$ must be inverted at each iteration (see (11)). Hence, the computational complexity of the 'filtered ML' estimate is significantly lower than when using the 'approximated MAP' or 'reduced complexity approximate MAP'estimator.

C. Estimation of $\mathbf{N}_{SRD}$

During the SAGE iterations we also assume the noise variance to be time invariant, but we will exploit the newly obtained knowledge on the unknown data symbols, i.e. (9) and (10). One can verify that in this case the SAGE-based ML estimation of the diagonal elements of $\hat{\mathbf{N}}^{(i)}_{SRD}$, using (5) with $p(\boldsymbol{\theta})$ removed, gives rise to

$$\frac{|\boldsymbol{r}_{RD} - \hat{\boldsymbol{h}}^{(i)}_{SRD}\mathbf{U}^{(i)}|^2 + \hat{\boldsymbol{h}}^{(i)}_{SRD}\left(\Sigma^{(i)} - \mathbf{U}^{(i)}\mathbf{U}^{(i),H}\right)\hat{\boldsymbol{h}}^{(i),H}_{SRD}}{K}, \quad (13)$$

where $\hat{\boldsymbol{h}}^{(i)}_{SRD}$ is the channel gain estimate according to the considered strategy (S_1,S_2,S_3,S_4). For the ML estimation strategy S_3 we obtain from (see (11))

$$\begin{aligned}|\boldsymbol{r}_{RD} - \hat{\boldsymbol{h}}^{(i)}_{SRD}\mathbf{U}^{(i)}|^2 &= |\boldsymbol{r}_{RD} - \boldsymbol{r}_{RD}\Sigma^{(i),-1}\mathbf{U}^{(i)}\mathbf{U}^{(i),H}|^2 \\ &= |\boldsymbol{r}_{RD}\Sigma^{(i),-1}\left(\Sigma^{(i)} - \mathbf{U}^{(i)}\mathbf{U}^{(i),H}\right)|^2.\end{aligned}$$

Taking into account that $(\Sigma^{(i)} - \mathbf{U}^{(i)}\mathbf{U}^{(i),H})$ converges to $\mathbf{0}$ in the high SNR region, estimating $\mathbf{N}_{SRD}$ according to (13) yields a poor performance. Therefor, for strategy S_3 we will not update the estimate of $\mathbf{N}_{SRD}$, i.e. $\hat{\mathbf{N}}^{(i)}_{SRD} = \hat{\mathbf{N}}^{(0)}_{SRD}$.

IV. Simulation Results

In this section we invesstigate the performance of the different estimation strategies in terms of mean-square estimation error (MSEE) and frame error rate (FER). Defining $\text{SNR}_{SD} = E_s H_{SD}/N_{0,SD}$, $\text{SNR}_{SR} = E_s H_{SR}/N_{0,SR}$ and $\text{SNR}_{RD} = E_r H_{RD}/N_{0,RD}$, we assume that (a) the maximum Doppler frequency is the same in all channels and equal to 0.01 and (b) SNR_{SR}=SNR_{RD}=2 SNR_{SD}, indicating that both SNR_{SR} and SNR_{RD} are 3dB higher than SNR_{SD}. It is to be understood that when SNR_{SD} changes, SNR_{SR} and SNR_{RD} are modified accordingly, maintaining the 3 dB difference with SNR_{SD}.

Unless mentioned otherwise, the data is encoded at the source using a non-recursive rate 1/2 convolutional code with generating polynomials $(15, 17)_8$ [11] and is decoded at the destination by means of the Viterbi algorithm [12] after convergence of the SAGE iterations. At each SAGE iteration, the symbol a posteriori probabilities needed to compute the moments (9) and (10) are obtained from the BCJR algorithm [13], which is about three times as complex as the Viterbi algorithm.

A. 'reduced complexity approximation MAP' : Influence of N_{it}

As argued before, the computational complexity of the SAGE algorithm can be reduced by dividing the RD slot in N_{it} subslots, which the estimator considers uncorrelated. Fig. 2 depicts the FER for several values of N_{it}. Observe that reducing the computational complexity in (8) by a factor 64 (N_{it}=8) results in a loss of about only 0.6 dB as compared to 'approximate MAP' estimation (N_{it}=1).

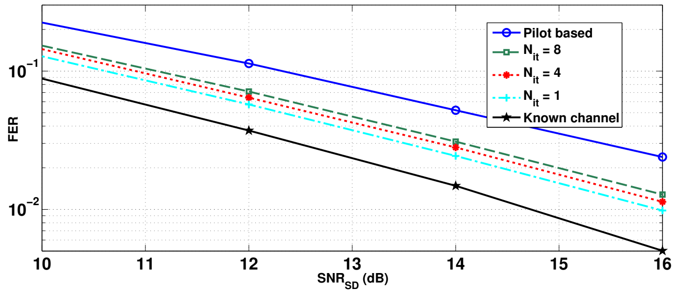

Fig. 2. C/M=$-\infty$ dB, 2 SAGE iterations, K_d=200, K_s=10, 4QAM

B. Estimation of h_{SRD}

Fig. 3 depicts a segment of a realization of the real part of h_{SRD} along with its estimates corresponding to the different strategies, as a function of the symbol index. The corresponding normalized MSEE (w.r.t. $E[|\boldsymbol{h}_{SRD}|^2]$) is shown in Fig. 4. Observe that, (i) the most complex strategy (S_1) yields the smallest MSEE, (ii) the estimate resulting from strategy S_2 is discontinuous at the edges of the subblocks (hence, $N_{it}-1$ discontinuities occur in each slot), (iii) the ML estimate (strategy S_3) exhibits large fluctuations and the largest MSEE, because the estimate (11) does not exploit the time-correlation of h_{SRD} and (iv) by exploiting this time-correlation, strategy S_4 results in a significant improvement, but the more complex strategies S_1-S_2 yield a slightly lower MSEE.

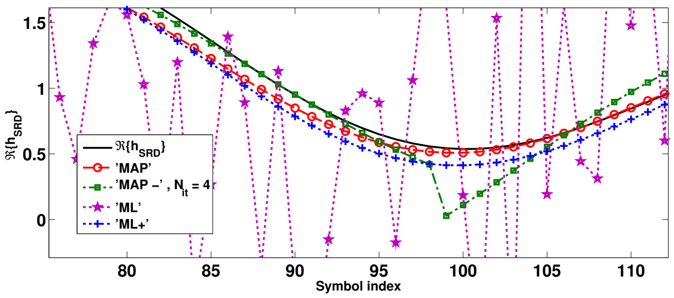

Fig. 3. C/M=$-\infty$ dB, 1 SAGE iterations, K_d=180, K_s=12, 16QAM

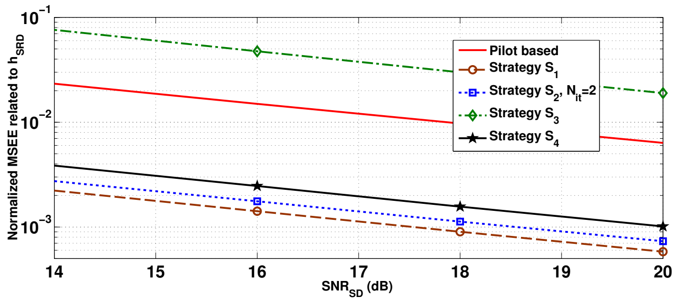

Fig. 4. C/M=$-\infty$ dB, 3 SAGE iterations, K_d=180, K_s=12, 16QAM

C. Estimation of N_{SRD}

Regarding the estimation of $\boldsymbol{N}_{SRD}$, Fig. 5 depicts the normalized MSEE (w.r.t. $E[|\boldsymbol{N}_{SRD}|^2]$). Except for strategy

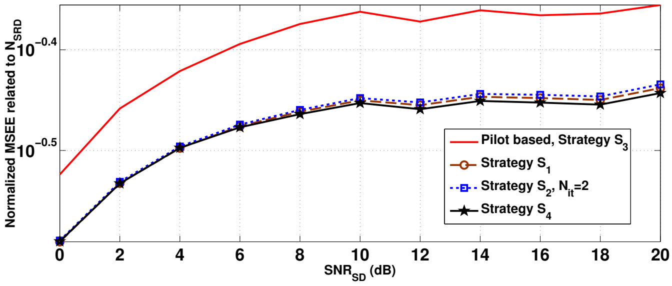

Fig. 5. C/M=$-\infty$ dB, 3 SAGE iterations, K_d=180, K_s=12, 16QAM

S_3 which does not update the noise variance estimate, all other algorithms estimate $\boldsymbol{N}_{SRD}$ according to (13) and similar MSEEs are observed.

D. FER

With C/M=$-\infty$dB (Rayleigh fading) and C/M=5dB, the FER for several estimation strategies is depicted in Fig. 6 and 7, respectively. In successive SAGE iterations smaller FER values are obtained, and convergence is achieved after only 3 iterations. The 'approximate MAP' estimator (strategy S_1) performs slightly better than the 'filtered ML' estimator (strategy S_4), but the latter has a much smaller complexity. Both algorithms yield an improvement of over 1.5dB as compared to the system (proposed in [4]) that uses pilot-based estimates only. Note that although the ML estimate is clearly worse than the pilot-based estimate in terms of MSEE, the ML estimator slightly outperforms the pilot-based estimator in terms of FER.

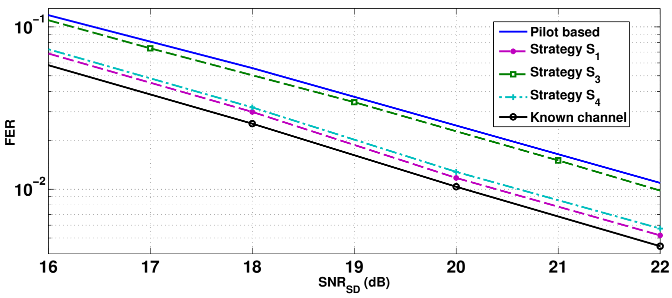

Fig. 6. C/M=$-\infty$dB, 3 SAGE iterations, K_d=180, K_s=10, 16QAM

V. Conclusions

In this contribution we derived several SAGE-based iterative channel estimation algorithms for time-varying Rice fading AF networks. We showed that a low-complexity filtering of the ML channel coefficient estimate yields a performance very close to the considerably more complex 'approximately MAP' estimate, and clearly outperforms the system that uses pilot-based estimates.

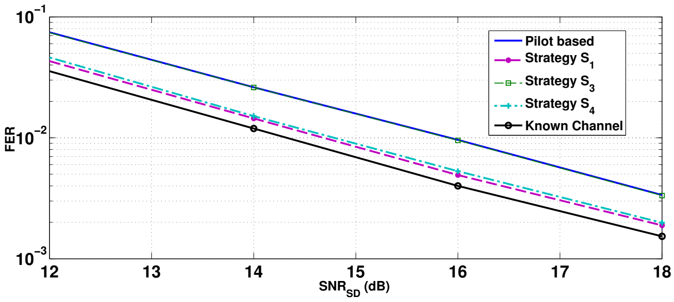

Fig. 7. C/M= 5dB, 3 SAGE iterations, K_d=180, K_s=10, 16QAM

References

[1] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 1995.

[2] J. N. Laneman, "Cooperative Diversity in Wireless networks: Algorithms and Architectures," Ph.D. dissertation, Massachusetts Institute of Technology, Massachusetts, USA, 2002

[3] A. Sendonaris, E. Erkip and B. Aazhang, "User Cooperation Diversity-Part I: System Description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927-1938, Nov. 2003

[4] C. Patel and G. Stüber, "Channel Estimation for Amplify and Forward Relay Based Cooperation Diversity Systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2348-2356, June 2007

[5] W. C. Jakes, *Microwave Mobile Communications,* 2nd ed. Piscataway, NJ: IEEE PRess, 1994

[6] D. Brennan, "Linear Diversity Combining Techniques," *Proceedings of the IRE*, vol. 47, no. 6, pp. 1075-1102, June 1959

[7] A. Dempster, N. Laird and D. B. Rubin, "Maximum Like-lihood From Incomplete Data Via the EM Algorithm," *J. Royal Statistical Soc.*, vol. 39, no. 1, pp. 1-38, 1977

[8] H. Guenach, M. Wymeersch and M. Moeneclaey, "On the Channel Parameter Estimation in a Space Time Bit-interleaved-coded Modulation System for Multipath DS-CDMA Uplink with Receive Diversity," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1747-1758, Sept. 2005

[9] N. Aerts, I. Avram, D. Duck and M. Moeneclaey "EM Based Channel Estimation in an Amplify-and-Forward Relaying Network" *IEEE 21th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2010, pp. 724-729

[10] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time-series Analysis.* Reading, MA: Addison-Wesley, 1991

[11] S. Lin and D. Costello *Error Control Coding* Pearson Education Inc., second edition, 2004

[12] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions Information Theory*, vol. 13, pp. 260-269, April 1967

[13] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv *Optimal Decoding of Linear Codes for minimizing symbol error rate* IEEE Trans. on Inform. Theory, vol. IT-20(2), pp.284-287, March 1974