EURASIP Journal on
Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH   **Open Access**

Exploiting *a priori* information for iterative channel estimation in block-fading amplify-and-forward cooperative networks

Nico Aerts[*], Iancu Avram and Marc Moeneclaey

Abstract

In an amplify-and-forward cooperative network, a closed-form expression of the *a priori* distribution of the complex-valued gain of the global relay channel is intractable, so that *a priori* information is often not exploited for estimating this gain. Here, we present two iterative channel gain and noise variance estimation algorithms that make use of *a priori* channel information and exploit the presence of not only pilot symbols but also unknown data symbols. These algorithms are approximations of maximum *a posteriori* estimation and linear minimum mean-square error estimation, respectively. A substantially reduced frame error rate is achieved as compared to the case where only pilot symbols are used in the estimation.

Keywords: Cooperative communication; MAP estimation; Amplify-and-forward; SAGE algorithm

1 Introduction

As wireless channels suffer from multipath propagation, several methods to combat the detrimental effect of fading have been proposed [1]. Cooperative communication [2] is a relatively new method where spatial diversity is achieved by exploiting the presence of other terminals in the network. The source time-shares its allocated time frame with other terminals that acts as relays. During the first slot of the frame, the source broadcasts information to the destination and the relays; the remaining slots are used by the relays to transmit to the destination information that is related to the message sent by the source. In this paper, we consider the amplify-and-forward protocol [3], and hence, each relay simply amplifies and retransmits the signal received from the source.

In real-world situations, the channels between the different terminals are unknown and must be estimated, before detection at the destination can start. The overall noise in the signal received from the relay has a variance depending on the realization of the relay-destination channel, and the overall channel gain is the product of the source-relay and relay-destination channel gains. It has been proposed (e.g., the linear minimum mean-square

error (LMMSE) cascaded channel estimation from [4,5]) that the destination estimates the overall channel gain but takes the overall noise variance equal to the variance obtained by averaging over the statistic of the relay-destination channel, whereas in [6] the relay-destination channel gain is estimated separately (and the noise variance computed accordingly) at the expense of a more sophisticated relay (that adds pilot symbols of its own). The LMMSE disintegrated estimation from [5] involves the estimation of the source-relay channel at the relay (which significantly increases relay complexity) and the relay-destination channel at the destination, whereas [7] considers maximum-likelihood (ML) estimation of both these channels at the destination. In [8], ML estimation of the overall channel gain and noise variance is performed at the destination; these results are extended in [9] to the case of a multi-antenna destination.

In this contribution, we present two pilot-based and two space-alternating generalized expectation-maximization (SAGE) [10] algorithms for estimating at the destination both the overall channel gain and (unlike the cascaded channel estimation from [5]) the overall noise variance. In contrast with [6-9], the proposed algorithms also take *a priori* channel knowledge into account. We restrict our attention to a low-complexity amplify-and-forward relay, which does not add pilot symbols of its own nor performs

*Correspondence: nico.aerts@telin.ugent.be
TELIN, UGent, Gent 9000, Belgium

channel estimation. We derive in closed form the joint *a priori* distribution of the parameters to be estimated and use this distribution to derive an approximately maximum *a posteriori* (MAP) estimate and an approximation of the LMMSE estimate. We investigate the mean-square estimation error (MSEE) and frame error rate (FER) performance resulting from these estimates and make the comparison with the performance that results from the joint ML channel gain and the noise variance estimates from [8].

Notations

All vectors are row vectors and boldface; the Hermitian transpose, statistical expectation, and estimate of the row vector x are denoted by x^H, $E[x]$, and $\hat{x}$, respectively.

2 System description

We consider a system consisting of a source S, a destination D, and a relay R. During a given frame, the channels between the terminals are characterized by a noise vector n_m with variance N_m per component and a constant channel gain h_m with variance H_m and where $m \in \{SD, SR, RD\}$. Both h_m and the elements of n_m are zero-mean circular symmetric complex Gaussian (ZMCSCG) distributed, and SD, SR, and RD refer to the source-destination, source-relay, and relay-destination path, respectively.

During the first time slot, the source transmits a sequence of $K = K_p + K_d$ complex symbols c consisting of K_p pilot symbols c_p and K_d data symbols c_d; the latter are obtained by coding and mapping a packet of information bits. We assume $E[|c|^2] = KE_s$. Throughout the paper, we use the notation x_p and x_d to indicate the part of any vector x that corresponds to the pilot and data symbols, respectively. The signals received by the relay and the destination during the first slot are

$$r_m = h_m c + n_m, \qquad m \in \{SR, SD\}. \qquad (1)$$

The relay amplifies r_{SR} using a fixed gain γ, and hence, the signal received by the destination during the second time slot is

$$r_{RD} = \gamma h_{SRD} c + n_{SRD}, \qquad (2)$$

where $h_{SRD} = h_{SR} h_{RD}$ and the elements of n_{SRD} are ZMCSCG distributed with variance

$$N_{SRD} = \gamma^2 |h_{RD}|^2 N_{SR} + N_{RD} \qquad (3)$$

when conditioned on h_{RD}. If the power constraint at the relay imposes the average energy per sent frame $E[|\gamma r_{SR}|^2]$ to be KE_r, with E_r denoting the energy per symbol transmitted by the relay, the constant gain factor follows from the relation $(H_{SR}E_s + N_{SR})\gamma^2 = E_r$.

Taking into account the energies spent by both the source and the relay, the received energy per information bit at the destination is obtained as

$$E_b = \frac{K_p + K_d}{K_d} \cdot \frac{E_s H_{SD} + E_r H_{RD}}{r_c \log_2(M)}, \qquad (4)$$

with r_c and M denoting the code rate and the constellation size, respectively. The spectral efficiency, expressed as the ratio between the information bitrate R_b and the symbol rate R_s, is given by

$$\frac{R_b}{R_s} = \frac{1}{2} \cdot \frac{K_d}{K_p + K_d} r_c \log_2(M), \qquad (5)$$

where the factor $1/2$ accounts for the fact that two time slots are used.

If the destination knows the channel gains and noise variances, maximum ratio combining [11] of the received signals yields the sufficient statistic z to be used by the decoder

$$z = \frac{r_{SD,d} h_{SD}^*}{N_{SD}} + \frac{\gamma r_{RD,d} h_{SRD}^*}{N_{SRD}}. \qquad (6)$$

The noise variances N_m ($m \in \{SD, SR, RD\}$) are long-term properties of the channels, and we consider them to be known. The three components of $v = (h_{SD}, h_{SRD}, N_{SRD})$ depend on the channel gain realizations during the considered frame; therefore, the destination must estimate v and calculate the sufficient statistic (6) with v replaced by its estimate.

3 Joint *a priori* distribution

The SD channel gain h_{SD} is ZMCSCG distributed with variance H_{SD} and is independent of (h_{SRD}, N_{SRD}). When conditioned on h_{RD}, (a) the variables h_{SRD} and N_{SRD} are statistically independent, (b) h_{SRD} is ZMCSCG distributed with variance $H_{SR}|h_{RD}|^2$, and (c) the probability density function (pdf) of N_{SRD} is a Dirac delta function located at $\gamma^2 |h_{RD}|^2 N_{SR} + N_{RD}$. Then, $p(h_{SRD}, N_{SRD})$ is obtained by averaging $p(h_{SRD}|h_{RD})p(N_{SRD}|h_{RD})$ over h_{RD}, which is ZMCSCG distributed with variance H_{RD}. As outlined in Appendix 1, this yields

$$p(h_{SRD}, N_{SRD}) = Cy^{-1} \exp\left(-Ay - B|h_{SRD}|^2 y^{-1}\right), \qquad (7)$$

where $A = (\gamma^2 N_{SR} H_{RD})^{-1}$, $B = \gamma^2 N_{SR}/H_{SR}$, $C = (\pi H_{SR} H_{RD})^{-1}$, and y is a short-hand notation for $N_{SRD} - N_{RD}$.

4 Estimation

In order to obtain the MAP estimate of v, which is defined by:

$$\hat{v}_{\text{MAP}} = \arg \max_v \ln p(r, v), \qquad (8)$$

with $r = (r_{SD}, r_{RD})$ and $p(r, v) = p(r|v)p(v)$, one has to compute $p(r|v)$ by averaging over the unknown data symbols c, i.e.,

$$p(r|v) = \sum_c p(r|v, c) Pr[c], \qquad (9)$$

with $Pr[c]$ the *a priori* probability of c. As the summation in (9) runs over all possible codewords, the computation of $p(r|v)$ is not feasible because its complexity increases exponentially with the data sequence length.

The expectation-maximization (EM) algorithm [12], which is tailored to deal with a nuisance parameter such as c, produces a sequence of estimates $\hat{v}^{(i)}$, $i = 1,2,...$ according to

$$\hat{v}^{(i)} = \arg \max_{v} E\left[\ln \left(p(r|v,c)p(v) \right) | r, \hat{v}^{(i-1)} \right], \quad (10)$$

that, when properly initialized, converges to the MAP estimate (8); the conditional expectation in (10) is with respect to the nuisance parameter c. We will use the SAGE algorithm [10] instead of the EM algorithm in order to avoid the complexity of the multidimensional maximization associated with (10). The SAGE algorithm replaces the multidimensional maximization in (10) by several lower dimensional maximizations over mutually exclusive subsets of v. The SAGE algorithm needs an initial estimate $\hat{v}^{(0)}$ to start the iterations; this initial estimate is derived from the pilot symbols. Using (1) and (2), it follows from (10) that h_{SD} and (h_{SRD}, N_{SRD}) can be estimated independently

$$\hat{h}_{SD}^{(i)} = \arg \max_{h_{SD}} E\left[\ln \left(p(r_{SD}|h_{SD},c)p(h_{SD}) \right) | r, \hat{v}^{(i-1)} \right]$$

$$(11)$$

and $\left(\hat{h}_{SRD}^{(i)}, \hat{N}_{SRD}^{(i)} \right)$ is given by

$$\arg \max_{(h_{SRD},N_{SRD})} E\left[\ln \left(p(r_{RD}|h_{SRD},N_{SRD},c)p(h_{SRD},N_{SRD}) \right) | r, \hat{v}^{(i-1)} \right].$$

$$(12)$$

As the estimation of h_{SD} is well documented [13], we will present the estimates of (h_{SRD}, N_{SRD}) only.

4.1 Iterative approximate MAP estimation
Let us first derive a pilot-based initial MAP estimate $\left(h_{SRD}^{(0)}, N_{SRD}^{(0)} \right)$. When conditioned on h_{SRD} and N_{SRD}, $r_{RD,p} - h_{SRD}c_p$ is ZMCSCG distributed with covariance matrix $N_{SRD}I_{K_p}$. Taking (7) into account, obtaining the pilot-based MAP estimate of (h_{SRD}, N_{SRD}) that maximizes $p(r_{RD,p}, h_{SRD}, N_{SRD})$ is, as shown in Appendix 2, intractable. To circumvent this problem, we firstly compute from $r_{RD,p}$ the ML estimate of h_{SRD}, which is given by [8]

$$\hat{h}_{SRD,ML}^{(0)} = \frac{r_{RD,p}c_p^H}{\gamma |c_p|^2}. \quad (13)$$

Next, exploiting (7), a numerical search algorithm is used to obtain an initial estimate of N_{SRD} according to

$$\hat{N}_{SRD}^{(0)} = \arg \max_{\substack{N_{SRD} \\ N_{SRD} \geq N_{RD}}} \ln p\left(r_{RD,p}, \hat{h}_{SRD,ML}^{(0)}, N_{SRD} \right). \quad (14)$$

Finally, we maximize $\ln p\left(r_{RD,p}, h_{SRD}, \hat{N}_{SRD}^{(0)} \right)$ over h_{SRD}; according to (31) in Appendix 2, this yields the closed-form expression

$$\hat{h}_{SRD}^{(0)} = \left(\frac{|c_p|^2 H_{SR} \left(\hat{N}_{SRD}^{(0)} - N_{RD} \right)}{|c_p|^2 H_{SR} \left(\hat{N}_{SRD}^{(0)} - N_{RD} \right) + \hat{N}_{SRD}^{(0)} N_{SR}} \right) \hat{h}_{SRD,ML}^{(0)}. \quad (15)$$

During the ith SAGE iteration, we use (12) to firstly estimate h_{SRD} with N_{SRD} fixed to $\hat{N}_{SRD}^{(i-1)}$; according to Appendix 3, this yields the closed-from expression

$$\hat{h}_{SRD}^{(i)} = \left(\frac{\sigma^{(i)} H_{SR} \left(\hat{N}_{SRD}^{(i-1)} - N_{RD} \right)}{\sigma^{(i)} H_{SR} \left(\hat{N}_{SRD}^{(i-1)} - N_{RD} \right) + \hat{N}_{SRD}^{(i-1)} N_{SR}} \right) \hat{h}_{SRD,ML}^{(i)}, \quad (16)$$

where

$$\hat{h}_{SRD,ML}^{(i)} = \frac{r_{RD}\mu^{(i),H}}{\gamma \sigma^{(i)}} \quad (17)$$

is the SAGE-based ML estimate of h_{SRD} [8]. The quantities

$$\mu^{(i)} = E\left[c|r, \hat{v}^{(i-1)} \right], \quad (18)$$

$$\sigma^{(i)} = E\left[|c|^2|r, \hat{v}^{(i-1)} \right] \quad (19)$$

are the *a posteriori* expectations (APEs) of c and $|c|^2$ and are easily derived from the marginal *a posteriori* probabilities of the coded symbols, which are obtainable by message passing on a factor graph [14]. Next, we look for the estimate $\hat{N}_{SRD}^{(i)}$ by using (12) with h_{SRD} fixed to $\hat{h}_{SRD}^{(i)}$. As it was the case with the pilot-based estimate, it is not possible to obtain a closed-form expression for $\hat{N}_{SRD}^{(i)}$. As a result, a numerical search is used to obtain $\hat{N}_{SRD}^{(i)}$ from (12), with h_{SRD} replaced by $\hat{h}_{SRD}^{(i)}$, i.e.,

$$\hat{N}_{SRD}^{(i)} = \arg \max_{\substack{N_{SRD} \\ N_{SRD} \geq N_{RD}}} E\left[\ln p(r_{RD}, \hat{h}_{SRD}^{(i)}, N_{SRD})|r, \hat{v}^{(i-1)} \right]. \quad (20)$$

4.2 Iterative approximate LMMSE estimation
A drawback of the approximate MAP estimation in the previous section is the numerical search required to obtain the noise variance estimate. This can be avoided by considering a LMMSE estimate for h_{SRD}, which is obtained independently of the noise variance estimate. If c was known at the destination, we could take a linear estimate of the form $\hat{h}_{SRD} = r_{RD}u^H$ and determine u such that $E\left[|\hat{h}_{SRD} - h_{SRD}|^2 \right]$ is minimum, with E[.] denoting the expectation over the *a priori* statistics of the noise and the fading. As shown in Appendix 4, this yields $u = \alpha c/(\gamma |c|^2)$ with $\alpha = \kappa/(1+\kappa)$, $\kappa = \gamma^2|c|^2 H_{SR}H_{RD}/N_{SRD,avg}$,

and $N_{SRD,\text{avg}} = \gamma^2 H_{RD} N_{SR} + N_{RD}$. In order to (approximately) apply LMMSE estimation in the case where c is not known to the destination, we simply replace c and $|c|^2$ by their *a posteriori* expectations $\mu^{(i)}$ and $\sigma^{(i)}$, respectively, yielding

$$\hat{h}_{SRD}^{(i)} = \left(\frac{\sigma^{(i)} H_{SR} \left(N_{SRD,\text{avg}} - N_{RD} \right)}{\sigma^{(i)} H_{SR} (N_{SRD,\text{avg}} - N_{RD}) + N_{SRD,\text{avg}} N_{SR}} \right) \hat{h}_{SRD,ML}^{(i)},$$

(21)

with $\hat{h}_{SRD,ML}^{(i)}$ given by (17). This iterative estimate is initialized by means of the pilot-based LMMSE estimate given by

$$\hat{h}_{SRD}^{(0)} = \left(\frac{|c_p|^2 H_{SR} \left(N_{SRD,\text{avg}} - N_{RD} \right)}{|c_p|^2 H_{SR} (N_{SRD,\text{avg}} - N_{RD}) + N_{SRD,\text{avg}} N_{SR}} \right) \hat{h}_{SRD,ML}^{(0)},$$

(22)

with $\hat{h}_{SRD,ML}^{(i)}$ given by (13). This yields the estimates from Section 4.1, but with $\hat{N}_{SRD}^{(i)}$ in (15) and $\hat{N}_{SRD}^{(0)}$ in (16) replaced by $N_{SRD,\text{avg}}$. The LMMSE channel gain estimate coincides with the MAP channel gain estimate for the (fictitious) case where h_{SRD} is ZMCSCG distributed with variance $H_{SR} H_{RD}$ and N_{SRD} is replaced by $N_{SRD,\text{avg}}$.

In order to obtain a closed-form expression for the noise variance estimate, we ignore the *a priori* distribution (7) of (h_{SRD}, N_{SRD}). The pilot-based initial estimate is given by

$$\hat{N}_{SRD}^{(0)} = \max \left(\frac{|r_{RD,p} - \gamma \hat{h}_{SRD}^{(0)} c_p|^2}{K_p - 1}, N_{RD} \right).$$

(23)

Note that the first argument in the max(.) function in (23) is an unbiased estimate of N_{SRD}; the (biased) ML estimate of N_{SRD} is obtained when replacing $K_p - 1$ by K_p. The restriction that $\hat{N}_{SRD}^{(0)}$ should not be smaller than N_{RD} arises naturally from (3). Similarly, during the SAGE iterations, we ignore in (12) the *a priori* distribution of (h_{SRD}, N_{SRD}) and perform the maximization over N_{SRD} with h_{SRD} fixed to $\hat{h}_{SRD}^{(i)}$, yielding

$$\hat{N}_{SRD}^{(i)} = \max \left(\frac{|r_{RD} - \gamma \hat{h}_{SRD}^{(i)} \mu^{(i)}|^2 + \gamma^2 |\hat{h}_{SRD}^{(i)}|^2 \left(\sigma^{(i)} - |\mu^{(i)}|^2 \right)}{K}, N_{RD} \right),$$

(24)

which is the ML estimate from [8].

4.3 MSEE performance analysis
Here, we present closed-form expressions for the MSEE resulting from some of the considered algorithms. We make the comparison with the Cramer-Rao bound (CRB) and the modified CRB (MCRB), which are fundamental lower bounds on the MSEE of unbiased estimates. Note that these bounds do not take into account the

a priori distribution of the parameter to be estimated. For the derivation of the (M)CRBs in the context of an amplify-and-forward relaying system, we refer to [15].

Let us first consider the pilot-based estimation of h_{SRD}, which uses only r_p. The corresponding CRB and MCRB both equal $N_{SRD,\text{avg}}/|\gamma c_p|^2$, which was shown in [8] to coincide with the MSEE of the pilot-based ML estimate from (13). A closed-form expression for the MSEE of the pilot-based MAP channel gain estimate from (15) is hard to find, because this estimate is coupled with the (numerically obtained) noise variance estimate. In contrast, the MSEE of the pilot-based LMMSE channel gain estimate from (22) is, as illustrated in Appendix 4, easily found

$$E \left[\left| \hat{h}_{SRD}^{(0)} - h_{SRD} \right|^2 \right] = \left((H_{SR} H_{RD})^{-1} + \left(\frac{\gamma^2 |c_p|^2}{N_{SRD,\text{avg}}} \right) \right)^{-1}$$

(25)

Note that this MSEE is always below the corresponding (M)CRB, because the (biased) LMMSE channel gain estimate takes *a priori* information into account. This effect is more pronounced in the low SNR region, whereas at high SNR, the MSEE (25) converges to the (M)CRB.

When also incorporating r_d, it is not possible to obtain expressions for the MSEEs resulting from the considered estimation algorithms nor for the CRB, because of the presence of the unknown data symbols; however, the corresponding MCRB is easily evaluated and is given by $N_{SRD,\text{avg}}/E \left[|\gamma c|^2 \right]$. In the high SNR region, where the approximation $\mu^{(i)} \approx c$ is valid, the MSEE of the ML estimate coincides with the MCRB, and the MSEE of the LMMSE estimate for given c is given by (25) but with c_p replaced by c; again, the LMMSE estimate outperforms the ML estimate because the former exploits *a priori* information. For a constant-magnitude constellation, the MSEE of the LMMSE and the ML channel gain estimates (21) and (16) both converge to the MCRB with increasing SNR.

5 Numerical performance results
In this section, we study the MSEE and FER performance of the presented MAP and LMMSE algorithms by means of Monte Carlo simulations. We use a non-recursive rate-1/2 convolutional code [16] with generating polynomials $(15, 17)_8$ to encode 180 information bits per frame at the source; the resulting 360 coded bits are mapped to 4-quadrature amplitude modulation (QAM) symbols, yielding $K_d = 180$. We select $K_p = 2$, so that $K = 182$. We take $E_s = E_r$ and assume $N_{SR} = N_{SD} = N_{RD} \stackrel{\triangle}{=} N_0$, $H_{SR} = H_{RD} = 1$, and $H_{SD} = 0.5$. Defining the signal-to-noise ratios on the *SD*, *SR*, and *RD* channels as $SNR_{SD} = E_s H_{SD}/N_{SD}$, $SNR_{SR} = E_s H_{SR}/N_{SR}$, and $SNR_{RD} = E_r H_{RD}/N_{RD}$, we obtain $2 SNR_{SD} = SNR_{SR} = SNR_{RD}$. A comparison is made with both the pilot-based and the

SAGE-based ML estimates (derived in [8]), which do not exploit *a priori* information; as far as the FER performance is concerned, also a reference system is considered, where the destination knows all channel parameters and no pilot symbols are transmitted. For the MAP, LMMSE and ML algorithms considered, the SAGE algorithm converges after only two iterations.

The behaviour of the MSEE related to estimating h_{SRD} as a function of SNR_{SD} is depicted in Figure 1. Considering estimation based only on pilot symbols, we observe that the MAP and LMMSE estimates slightly outperform the ML estimate at low SNR, and their MSEE converges to the MMSE for ML estimation (which equals $N_{SRD,\text{avg}}/|\gamma \boldsymbol{c}_p|^2$) with increasing SNR. As far as the iterative SAGE-based estimates that exploit the entire received signal is concerned, the MAP, LMMSE, and ML estimates yield essentially the same MSEE, which is considerably less than when only the pilot symbols are exploited; at high SNR, the MSEE converges to the MCRB (which equals $N_{SRD,\text{avg}}/E\left[|\gamma \boldsymbol{c}|^2\right]$), whereas at low SNR, the MSEE is considerably larger than the MCRB because the *a posteriori* expectation $\boldsymbol{\mu}^{(i)}$ significantly deviates from $\boldsymbol{c}$.

For the same strategies, Figure 2 depicts the corresponding FER as a function of E_b/N_0, where E_b denotes the received energy per information bit at the destination (see (4)). When exploiting only the pilot symbols for channel estimation, the MAP approach slightly (by about 0.2 dB) outperforms the LMMSE and ML approaches in terms of FER; as compared to the reference system, the FER resulting from MAP estimation is about 2 dB worse. The iterative SAGE-based strategies (MAP, LMMSE, and ML) that exploit the entire received signal yield essentially the

same FER, which is only about 0.3 dB worse than the reference system.

Figure 3 depicts the FER performance of the LMMSE estimates for several iterations. We observe that the performance improvement beyond two iterations becomes negligibly small. We have verified that also the FER performance of the ML and MAP estimates has essentially converged after two iterations (results not displayed to avoid overloading the figure).

The considered algorithms make use of pilot symbols to assist the estimation. However, the transmission of pilot symbols reduces both the power efficiency (part of the transmit energy is devoted to pilot symbols that carries no information) and the spectral efficiency (see (5)). Figure 4 shows the effect of the number of pilot symbols K_p on the FER, for E_b/N_0 = 8 dB. The general trend is that (a) for small K_p, the FER decreases with increasing K_p, because the effect of improving estimation quality is more important than the reduction of the power efficiency; and (b) for large K_p, the FER increases with increasing K_p, because the reduction of the power efficiency is more important than the (slightly) improving estimation quality. For the given K_p, the SAGE-based algorithms outperform the pilot-based algorithms in terms of FER, because the former also exploits the data-bearing part of the received signal during the estimation process. We observe that the FER for the SAGE-based algorithms is only weakly dependent on K_p in the interval $2 \leq K_p \leq 10$ and achieves a broad minimum at $K_p = 5$; for the pilot-based algorithms, the dependence of the FER on K_p is stronger, because only the pilot symbols contribute to the estimation.

Finally, we display in Figure 5 the FER performance for a 16-QAM constellation; we still have $K_d = 180$ and $K_p = 2$
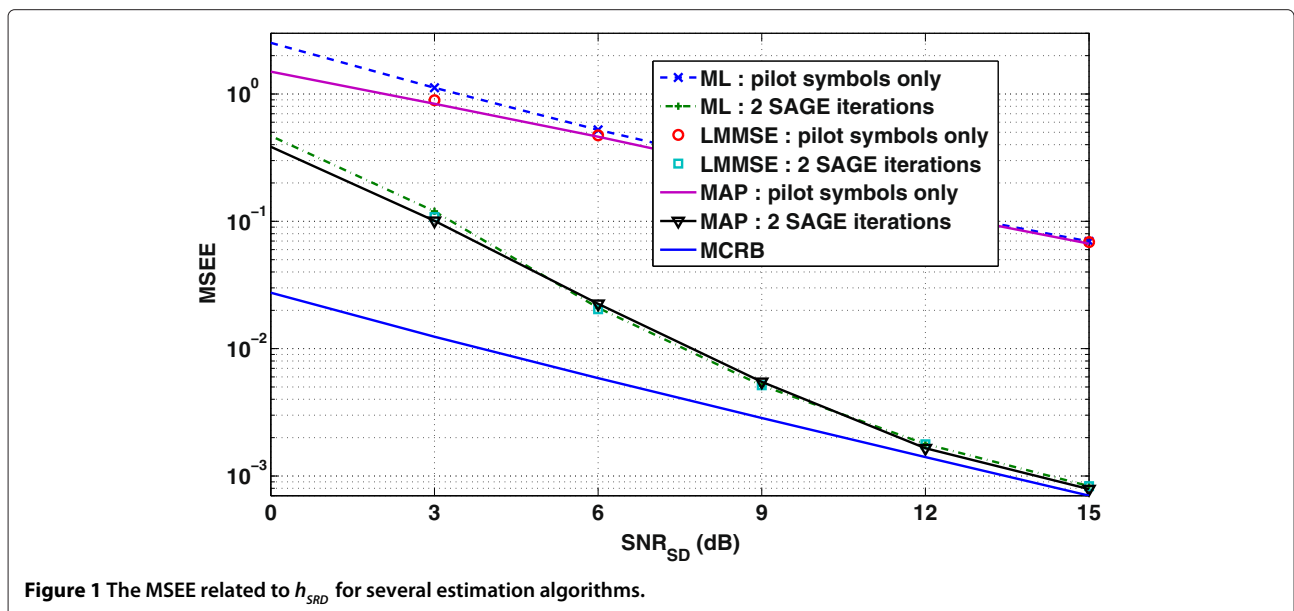

Figure 1 The MSEE related to h_{SRD} for several estimation algorithms.
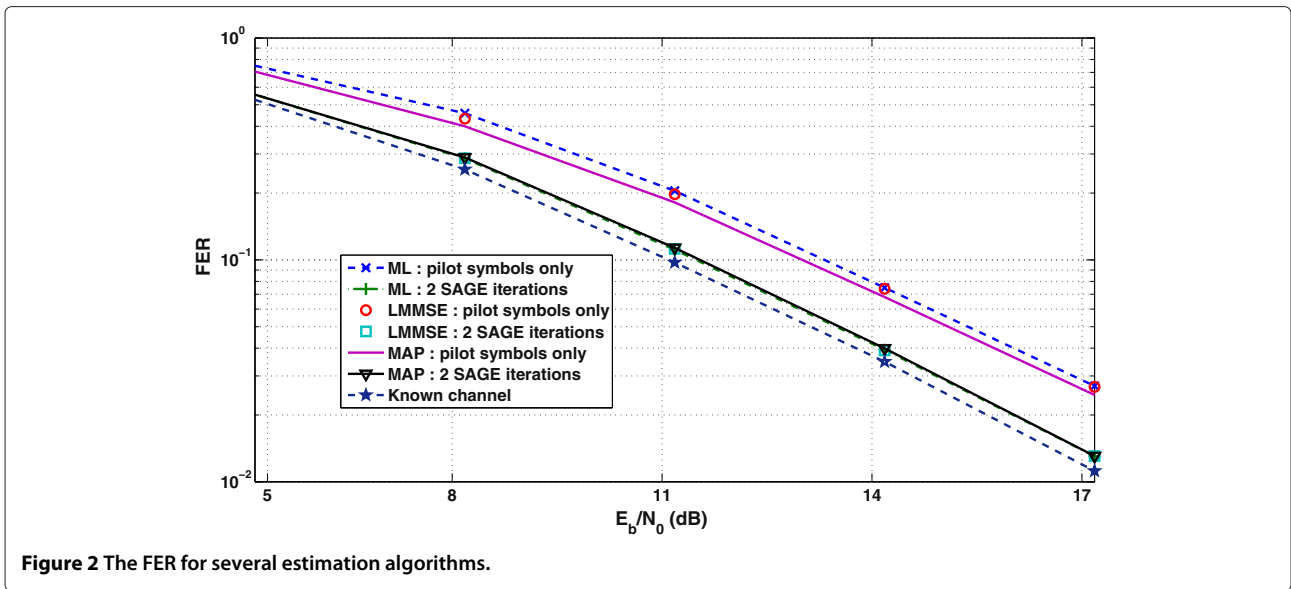
Figure 2 The FER for several estimation algorithms.

and use the same rate 1/2 code as for 4-QAM, implying that each frame now contains 360 information bits. In order to achieve a given FER, using 16-QAM requires a larger value of E_b/N_0 as compared to 4-QAM, because the distance between constellation points is smaller for the former. For pilot-based estimation, the MAP algorithm is slightly better than the LMMSE algorithm in terms of FER and outperforms the ML algorithm by about 0.8 dB; as compared to the reference system, the FER resulting from MAP estimation is about 2.2 dB worse. The iterative SAGE-based strategies (MAP, LMMSE, and ML) that exploit the entire received signal yield essentially the same FER, which is about 1.2 dB worse than the reference system.

6 Complexity considerations

Let us consider the computational complexity of the receiver as the sum of the complexities of the estimation algorithm and the decoding algorithm. As we have considered convolutional encoding, the decoder makes use of the Viterbi algorithm, which efficiently determines the shortest past in a trellis [16]. The complexity of the Viterbi algorithm is proportional to the number K_d of data symbols and the number of trellis states.

In the case of pilot-based estimation, the complexity of the considered estimation algorithms is roughly proportional to the number K_p of pilot symbols. The complexity of these algorithms is negligible as compared to that of the Viterbi decoder because $K_p \ll K_d$. Hence, for pilot-based
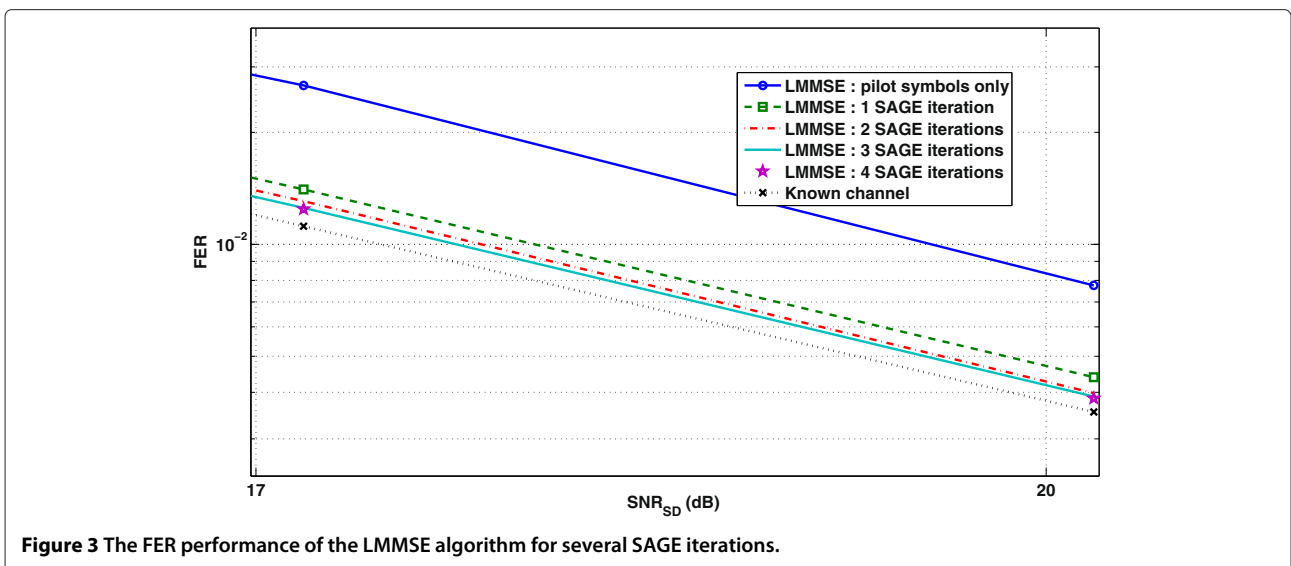

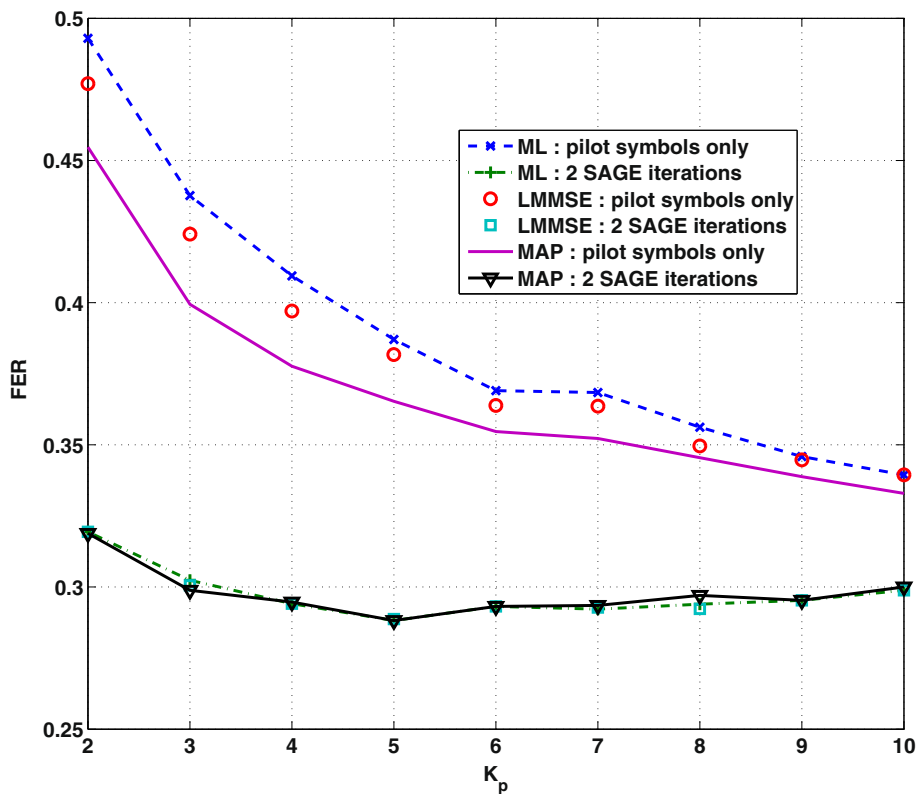**Figure 3 The FER performance of the LMMSE algorithm for several SAGE iterations.**

Figure 4 The FER as function of K_p for several estimation algorithms, with $E_b/N_0 = 8$ dB.

estimation, the receiver complexity is dominated by that of the Viterbi decoder.

In the case of SAGE-based estimation, at each iteration, the *a posteriori* probabilities of all K_d data symbols are needed in order to obtain the *a posteriori* expectations

(18) and (19). For a convolutional code, these probabilities are efficiently computed by means of the Bahl, Cocke, Jelinek, and Raviv (BCJR) algorithm [16], whose computational complexity is about three times as large as that of the Viterbi algorithm. The complexity of the considered
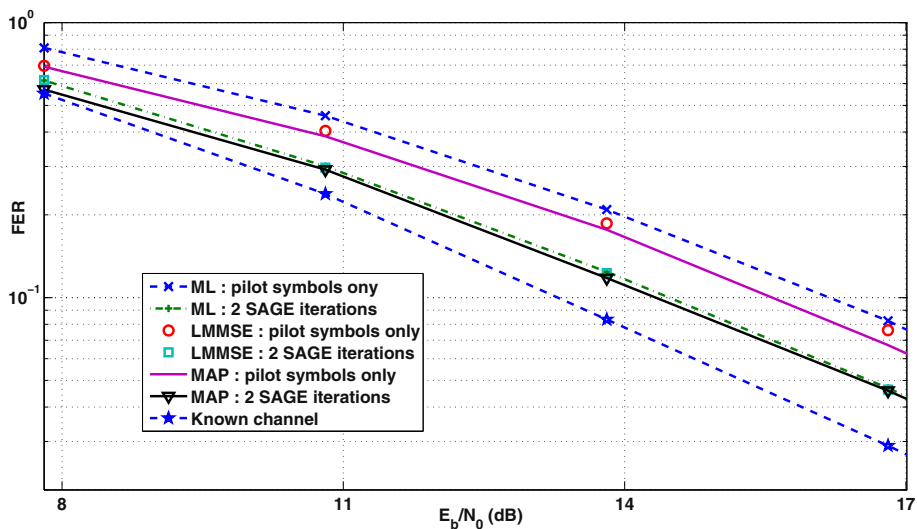

Figure 5 The FER performance for several estimation algorithms when using a 16-QAM constellation.

estimation algorithms is dominated by that of the BCJR algorithm. Hence, when I iterations are executed, the receiver complexity is about $3I + 1$ times the complexity of the Viterbi algorithm.

In [17], an approximate computation of the *a posteriori* symbol probabilities based on the Viterbi decoder metrics has been presented, with a complexity similar to that of the Viterbi algorithm. When using the algorithm from [17] during I SAGE iterations, the receiver complexity is about $I + 1$ times the complexity of the Viterbi algorithm.

In summary, the performance improvement resulting from the application of SAGE-based estimation comes at the cost of an increase of the receiver complexity as compared to the use of pilot-based estimation only. This increase is $(3I+1)$-fold for the BCJR algorithm and $(I+1)$-fold for the algorithm from [17]; in the case at hand, we have $I = 2$, which corresponds to a complexity increases by a factor of 7 and 3, respectively.

7 Conclusions

In the context of the cooperative amplify-and-forward protocol, we proposed approximations of MAP and LMMSE channel estimations, exploiting the *a priori* distribution of the channel parameters, and compared these algorithms to the ML channel estimation, which does not make use of *a priori* information. When using pilot symbols only, both the MAP and LMMSE algorithms outperform the ML algorithm in terms of FER; the MAP algorithm yields the better FER performance, providing a gain of about 0.2 dB (0.8 dB) compared to the ML algorithm, in the case of a 4-QAM (16-QAM) constellation and frame parameters, $K_p = 2$ and $K_d = 180$. Performance of the MAP, LMMSE, and ML algorithms is substantially improved by exploiting also the presence of the unknown data symbols; the resulting iterative SAGE-based algorithms converge after only two iterations and yield nearly identical FER performance for the MAP, LMMSE, and ML algorithms, which for a 4-QAM (16-QAM) constellation is only about 0.3 dB (1.2 dB) worse than in the case where the destination knows all channel parameters.

Appendix 1

The density $p(h_{SRD}, N_{SRD})$ is expressed as $p(h_{SRD}, N_{SRD}) = E\left[p(h_{SRD}|h_{RD}) \; p(N_{SRD}|h_{RD}) \right]$, where the averaging is w.r.t. h_{RD}. We have $p(N_{SRD}|h_{RD}) = \delta(N_{SRD} - N_{RD} - |h_{RD}|^2 \gamma^2 N_{SR})$ and

$$p(h_{SRD}|h_{RD}) = \frac{1}{\pi |h_{RD}|^2 H_{SR}} \exp\left(\frac{-|h_{SRD}|^2}{|h_{RD}|^2 H_{SR}} \right). \quad (26)$$

Noting that $p(N_{SRD}|h_{RD})$ and $p(h_{SRD}|h_{RD})$ depend on $|h_{RD}|^2$ but not on $\arg(h_{RD})$, we introduce $x = |h_{RD}|^2 \gamma^2 N_{SR}$, which has an exponential density: $p(x) = A \exp(-Ax)$ for $x \geq 0$ and $p(x) = 0$ otherwise, with $A = (\gamma^2 N_{SR} H_{RD})^{-1}$.

Expressing $p(h_{SRD}|h_{RD})$ and $p(N_{SRD}|h_{RD})$ as a function of x yields $p(N_{SRD}|h_{RD}) = \delta(y - x)$ and

$$p\left(h_{SRD}|h_{RD}\right) = \frac{B}{\pi x} \exp\left(\frac{-B|h_{SRD}|^2}{x} \right), \quad (27)$$

where $y = N_{SRD} - N_{RD}$ and $B = \gamma^2 N_{SR}/H_{SR}$. The density $p(h_{SRD}, N_{SRD})$ is then obtained as the expectation of $p(h_{SRD}|h_{RD}) p(N_{SRD}|h_{RD})$ w.r.t. x, i.e.,

$$p\left(h_{SRD}, N_{SRD}\right) = \int_0^{+\infty} \delta(y-x) \frac{AB}{\pi x} \exp\left(\frac{-B|h_{SRD}|^2}{x} - Ax \right) dx \quad (28)$$

$$= \frac{C}{y} \exp\left(\frac{-B|h_{SRD}|^2}{y} - Ay \right), \quad (29)$$

where $C = AB/\pi = (\pi H_{SR} H_{RD})^{-1}$.

Appendix 2

In order to obtain the pilot-based MAP estimate of (h_{SRD}, N_{SRD}), based on the observation $\boldsymbol{r}_{RD,p}$ only, we have to jointly maximize $\ln p(\mathbf{r}_{RD,p}, h_{SRD}, N_{SRD}) = \ln p(\mathbf{r}_{RD,p}|h_{SRD}, N_{SRD}) + \ln p(h_{SRD}, N_{SRD})$. Taking into account (7) and (2), we have (within terms not depending on (h_{SRD}, N_{SRD}))

$$\ln p\left(\mathbf{r}_{RD,p}, h_{SRD}, N_{SRD}\right) = -\frac{|\boldsymbol{r}_{RD,p} - \gamma h_{SRD} \boldsymbol{c}_p|^2}{N_{SRD}} - K_p \ln N_{SRD}$$
$$- \ln\left(N_{SRD} - N_{RD}\right) - \frac{B|h_{SRD}|^2}{N_{SRD} - N_{RD}} - A\left(N_{SRD} - N_{RD}\right). \quad (30)$$

Taking into account that (30) is quadratic in h_{SRD}, the value of h_{SRD} that maximizes (30) for given N_{SRD} is obtained as

$$\hat{h}_{SRD}(N_{SRD}) = \frac{\gamma \boldsymbol{r}_{RD,p} \boldsymbol{c}_p^H}{N_{SRD}} \left(\frac{\gamma^2 |\boldsymbol{c}_p|^2}{N_{SRD}} + \frac{\gamma^2 N_{SR}}{(N_{SRD} - N_{RD}) H_{SR}} \right)^{-1}. \quad (31)$$

Substituting (31) back into (30) yields the function $\ln p\left(\mathbf{r}_{RD,p}, \hat{h}_{SRD}(N_{SRD}), N_{SRD}\right)$; the value of N_{SRD} which maximizes this function cannot be obtained in closed form.

Appendix 3

Computation of the function to be maximized in (12) yields (within terms not depending on (h_{SRD}, N_{SRD}))

$$E\left[\ln\left(p\left(\boldsymbol{r}_{RD}|h_{SRD}, N_{SRD}, \boldsymbol{c}\right) p\left(h_{SRD}, N_{SRD}\right)\right) | \boldsymbol{r}, \hat{\boldsymbol{v}}^{(i-1)} \right]$$
$$= -\frac{1}{N_{SRD}} E\left[\gamma^2 |h_{SRD}|^2 |\mathbf{c}|^2 - 2\gamma \Re\left[\mathbf{r}_{RD} \mathbf{c}^H h_{SRD}^* \right] | \boldsymbol{r}, \hat{\boldsymbol{v}}^{(i-1)} \right]$$
$$- K \ln N_{SRD} + \ln p(h_{SRD}, N_{SRD}) \quad (32)$$
$$= -\frac{1}{N_{SRD}} \left(\gamma^2 |h_{SRD}|^2 \sigma^{(i)} - 2\gamma \Re\left[\mathbf{r}_{RD} \boldsymbol{\mu}^{(i),H} h_{SRD}^* \right] \right)$$
$$- K \ln N_{SRD} + \ln p\left(h_{SRD}, N_{SRD}\right) \quad (33)$$

where $\ln p(h_{SRD}, N_{SRD})$ follows from (7). In a similar way as for the pilot-based estimate, it can be verified that

$$\hat{h}_{SRD}^{(i)} = \frac{\gamma \boldsymbol{r}_{RD} \boldsymbol{\mu}^{(i),H}}{N_{SRD}^{(i-1)}} \left(\frac{\gamma^2 \sigma^{(i)}}{N_{SRD}^{(i-1)}} + \frac{\gamma^2 N_{SR}}{\left(N_{SRD}^{(i-1)} - N_{RD}\right) H_{SR}} \right)^{-1} \tag{34}$$

maximizes (33) for $N_{SRD} = \hat{N}_{SRD}^{(i-1)}$. As the maximization of (33) w.r.t. N_{SRD} for $h_{SRD} = \hat{h}_{SRD}^{(i)}$ is intractable analytically, we resort to a numerical search.

Appendix 4

In this appendix, we will determine the vector $\boldsymbol{u}$ that minimizes $E\left[|\hat{h}_{SRD} - h_{SRD}|^2\right]$, where $\hat{h}_{SRD} = \boldsymbol{r}_{RD}\boldsymbol{u}^H$. It follows from (2) that this MSEE is given by

$$E\left[|\boldsymbol{r}_{RD}\boldsymbol{u}^H - h_{SRD}|^2\right] = H_{SR}H_{RD}|\gamma \boldsymbol{u}\boldsymbol{c}^H - 1|^2 + N_{SRD,\text{avg}}|\boldsymbol{u}|^2. \tag{35}$$

It follows from (35) that in order to minimize the MSEE, $\boldsymbol{u}$ should be proportional to $\boldsymbol{c}/(\gamma|\boldsymbol{c}|^2)$. With $\boldsymbol{u} = \alpha\boldsymbol{c}/(\gamma|\boldsymbol{c}|^2)$, (35) yields

$$\alpha = \frac{\gamma^2|\boldsymbol{c}|^2 H_{SR}H_{RD}}{\gamma^2|\boldsymbol{c}|^2 H_{SR}H_{RD} + N_{SRD,\text{avg}}}. \tag{36}$$

The corresponding minimum value of (36) is obtained by substituting (36) into (35), which yields

$$E\left[|\boldsymbol{r}_{RD}\boldsymbol{u}^H - h_{SRD}|^2\right]_{\min} = \frac{H_{SR}H_{RD}N_{SRD,\text{avg}}}{\gamma^2|\boldsymbol{c}|^2 H_{SR}H_{RD} + N_{SRD,\text{avg}}} \tag{37}$$

The results for pilot-based LMMSE estimation are obtained by simply replacing in the above equations $\mathbf{r}_{RD}$ and $\mathbf{c}$ by $\mathbf{r}_{RD,p}$ and $\mathbf{c}_p$, respectively.

Competing interests
The authors declare that they have no competing interests.

References
1. JG Proakis, *Digital Communications*. (McGraw-Hill, New York, 1995)
2. A Sendonaris, E Erkip, B Azhang, User cooperation diversity - part I: system description. IEEE Trans. Commun. **51**(11), 1927–1938 (2003)
3. JN Laneman. Cooperative Diversity in Wireless networks: Algorithms and Architectures Ph.D. dissertation, Massachusetts Institute of Technology, Massachusetts, USA, 2002
4. B Gedik, M Uysal, in *Canadian conference on Electrical and Computer Engineering (CCECE)*. Two channel estimation methods for amplify-and-forward relay networks (IEEE Niagara Falls, 2008), pp. 615–618
5. O Amin, B Gedik, M Uysal, Channel estimation for amplify-and-forward relaying: cascaded against disintegrated estimators. IET Commun. **4**(10), 1207–1216 (2010)
6. M Liu, J Zhang, Y Zhang, Y Liu, in *IEEE 70th Vehicular Technology Conference Fall (VTC-Fall)*. A channel estimation scheme for amplify-and-forward OFDM relay networks (IEEE Anchorage, 2009), pp. 1–5
7. JS Sheu, WH Sheen, in *IEEE International Conference on Communications (ICC)*. An EM algorithm-based channel estimation for OFDM amplify and-forward relaying systems (IEEE Cape Town, South Africa, 2010), pp. 1–5
8. N Aerts, I Avram, D Duck, M Moeneclaey, in *IEEE 21th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. EM based channel estimation in an amplify-and-forward relaying network (Instanbul, 2010), pp. 724–729
9. C Zhang, S Tan, P Ren, Iterative receiver for amplify-and-forward relay networks with unknown noise correlation. Wirel. Commun. Mob. Comput. **14**, 1601–1612 (2012)
10. JA Fessler, AO Hero, Space-alternating generalized expectation-maximization algorithm. IEEE Trans. Signal Process. **42**(10), 2664–2677 (1994)
11. D Brennan, Linear diversity combining techniques. Proc. IRE. **47**(2), 1075–1102 (1959)
12. A Dempster, N Laird, DB Rubin, Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. **39**(1), 1–38 (1977)
13. J Choi, *Adaptive and Iterative Signal Processing in Communications*. (Cambridge University Press, Cambridge, 2006)
14. FR Kschischang, BJ Frey, HA Loeliger, Factor graphs and the sum-product algorithm. IEEE Trans. Inform. Theory. **47**(10), 498–519 (2001)
15. N Aerts, I Avram, M Moeneclaey, in *2010 IEEE 11th International Symposium on Spread Spectrum Techniques and Applications (ISITA)*. 'The Cramer-Rao bound for channel estimation in block fading amplify-and-forward relaying networks', (IEEE Taichung, 2010), pp. 12–16
16. S Lin, D Costello, *Error Control Coding*, 2nd edn. (Pearson Education Inc., New Jersey, 2004)
17. N Aerts, I Avram, M Moeneclaey, Low-complexity a posteriori probability approximation in EM-based channel estimation for trellis-coded systems. EURASIP J. Wireless Commun. Netw. (2014)