Joint Channel Estimation, Partial Successive Interference Cancellation, and Data Decoding for DS-CDMA Based on the SAGE Algorithm

Alexander Kocian, Ingmar Land, and Bernard H. Fleury

Abstract—This paper deals with the derivation and optimization of an iterative receiver architecture performing joint multiuser-decoding and channel-estimation. We consider an asynchronous multi-rate convolutional coded DS-CDMA system that communicates over quasi-static flat Rayleigh fading channels. The proposed receiver is derived within the Space-Alternating Generalized Expectation-Maximization (SAGE) framework in connection with the noise-splitting approach. The used theoretical framework guarantees convergence of the receiver, as opposed to many other iterative receiver structures. Furthermore, the noise-splitting approach provides a set of noise-weighting coefficients that can be optimized under weak constraints. The inputs to the single-user decoders are linear combinations of two kinds of soft values with weights determined by the noise-weighting coefficients. These two kinds of soft values can be interpreted as a-priori information and extrinsic information, respectively, if the channels are known. In the case of unknown channels, they are asymptotically a-priori and asymptotically extrinsic, i.e., they become a-priori and extrinsic when the length of the observed frame tends to infinity. In most cases, the optimum coefficients lead to extrinsic or asymptotically extrinsic values fed to the input of the single-user decoders. Monte-Carlo simulations show that the proposed receiver is resistant to channel estimation errors and supports high system loads.

Index Terms—Multiuser decoding, Iterative information processing, SAGE algorithm, Iterative joint data decoding and channel estimation.

I. INTRODUCTION

For a convolutional coded multiuser system, the optimum multiuser maximum likelihood (ML) sequence decoder exhibits a computational complexity that is exponential in both the number of users and the constraint length of the convolutional codes [1]. This makes a real-time implementation prohibitive even for a small number of users and a small constraint length. Furthermore, the mobile channels have to be estimated. Thus, algorithms for joint multiuser decoding and channel estimation are required that have feasible complexities. An often used approach is to iterate between channel estimation (CE), interference cancellation (IC), and single-user (SU) data decoding. Generally spoken, convergence cannot be guaranteed in such a setting.

The Expectation-Maximization (EM) algorithm [2], [3] is an iterative method to approximate the maximum-likelihood solution at polynomial computational complexity. At the same time, convergence of the algorithm is guaranteed. A variety of techniques have been proposed for accelerating the convergence of the EM algorithm such as Louis’ Turbo EM [4], which performs a Newton-Raphson step, or the Space-Alternating Generalized Expectation-maximization (SAGE) algorithm [5]. Nelson and Poor [6] extend the EM and SAGE algorithms to a so-called Space-Alternating Missing-Parameter EM (MPEM) algorithm: at each iteration, a certain subset of parameters is updated while the other parameters are considered as probabilistic missing data. Furthermore, the EM, SAGE, and MPEM algorithms are applied to derive various multiuser detectors for the additive white Gaussian noise (AWGN) channel. In these detectors hard or soft decisions computed in the single-user decoders to the interference canceler. Monte-Carlo simulations show the near-far resistance of these schemes.

Feder and Weinstein [7] apply the EM algorithm to the problem of estimating parameters of superimposed signals by assigning a fraction of total noise to each signal component, referred to as noise-splitting. Following this approach, Borran et al. propose a computationally feasible multiuser detector for the AWGN channel [8]. For multi-rate transmission, de Maio and Episopo [9] have developed a SAGE multiuser detector that exploits the fact that different users transmit at different rates. Monte-Carlo simulations of a near-far scenario with dual-rate transmission in AWGN demonstrate that both low-rate and high-rate users achieve similar multi-user efficiency. The appealing properties of the EM-based multiuser detectors remain when the channel is frequency-selective [10], [11].

The above multiuser detectors are derived under the assumption that the channel is known to the receiver. In practical wireless communication systems, however, the users’ channel need to be estimated. A pragmatic approach consists in supplementing the above multi-user (MU) detectors with separate channel estimators. We refer to this method as Separate multiuser Detection and channel Estimation (SDE) [12], [13]. These SDE schemes prove, however, to be highly sensitive to channel estimation errors. The solution to this sensitivity problem is Joint multiuser Detection and channel Estimation (JDE) [14]. At each iteration the JDE scheme computes more
reliable estimates of the users’ data symbols based on more accurate channel estimates and vice versa. Kocian and Fleury derive in [15] two EM-based schemes for JDE. Monte Carlo simulations in flat Rayleigh fading show that the proposed schemes are robust against channel estimation errors and have excellent multi-user efficiency. Q. Li et al. have independently developed in [16] two sequential EM-based JDE schemes for time-varying frequency-selective channels subject to known path-delays and time-bandwidth product of the channel.

Since practical CDMA systems employ channel coding, recent research has focused on incorporating data decoding in JDE. The approach can be either heuristic or systematic. In heuristic approaches, an IC device and a bank of SU decoders (one for each user) exchange soft information on the coded symbols [17]. Monte Carlo simulations in various scenarios have shown that a performance close to that in a SU scenario can be achieved even if the system load is large. In [18], Kobayashi et al. extend the heuristic approach to include channel estimation, using the EM algorithm. In a similar scheme, Chiavaccini and Vietta [19] iterate between Bayesian EM-based JDE and data decoding. This algorithm, however, has exponential complexity in the number of users, preventing practical implementation.

A crucial issue determining the efficiency of iterative multiuser decoding is the kind of information exchanged between the IC device and the SU decoders and vice versa. Boutros and Caire show in [20] that a soft-input soft-output (SISO) SU decoder requires extrinsic values on the code symbols [21] in the large-system limit when the channel is known to the receiver, and so does the IC device. Theoretical results are still lacking for the optimal information to be exchanged between the SU decoders and the IC device when the system comprises a finite number of users and finite-size interleavers.

In this paper, we derive an iterative receiver architecture for Joint multiuser Decoding, interference Cancellation, and channel Estimation (JDCE) in a systematic way by employing the SAGE method in connection with noise-splitting. The resulting SAGE-JDCE receiver scheme comprises a channel estimator, an interference cancellation device, and a soft-input/hard output (SIHO) SU decoder. Notice that only one SU decoder is required, as only one user is decoded in each iteration. A similar scheme was proposed in [15]. The present paper extends [15] by including channel coding, asynchronous signal reception, and Variable signature Sequence Length (VSL) multirate transmission.

The idea for deriving the algorithm is as follows: In each iteration, the received signal is considered as a sum of two terms. The first term contains the signal of the User-Of-Interest (UOI) and a fraction of the noise, while the second term contains the Multiple Access Interference (MAI) and the remaining of the noise. Thus, the overall noise is (conceptually) split into two parts determined by so-called noise-weighting coefficients.

In the resulting algorithm, in each iteration the input to the SU decoder is a weighted sum of soft values from the previous iteration, that can be conceived as a priori information on the UOI’s code symbols, and values from the current iteration, that can be conceived as extrinsic information on the UOI’s code symbols. These soft values represent a priori soft values and extrinsic soft values in the sense usually used in Turbo-code literature [21], [22] if the channels are known. For the case of unknown channels, these soft values become a priori and extrinsic, respectively, if the length of the observed frame tends to infinity. Therefore, we coin these two types of soft values *asymptotically a priori* information and *asymptotically extrinsic* information.

These two contributions are weighted by the above noise-weighting coefficients. Thus, the input of the single-user decoder is a weighted sum of a (asymptotically) a priori information and (asymptotically) extrinsic information on the code symbols. The noise-weighting coefficients can be selected almost freely under only weak constraints guaranteeing convergence of the algorithm. Thus, they can be optimized to yield maximum performance.

In the present paper, we propose an optimization criterion for the noise-weighting coefficients, and we derive analytical expressions for the optimum coefficients. Notice that the optimized weighting coefficients found in [15] violate the EM framework and thus do not guarantee convergence of the algorithm, as opposed to the optimum coefficients considered in the present paper. Simulation results show that the optimum coefficients lead in most cases to solely extrinsic or asymptotically extrinsic information at the input of the SU decoder. This result for finite number of users and finite-size interleavers agrees with the result for infinite number of user users and infinite-length interleavers in [20]. The proposed method may be extended to time-varying fading or frequency-selective fading.

The paper is organized as follows: Section II establishes the notation of the coded CDMA system over block Rayleigh fading channels. In Section III we apply the SAGE algorithm to JDCE and optimize its performance. Furthermore, we discuss implementation issues such as computational complexity and initialization. The performance of the proposed receiver is evaluated by means of Monte-Carlo simulations in Section IV.

II. **System Description**

We consider an asynchronous multi-rate convolutional-coded CDMA system with $K$ active users, employing long spreading codes with variable signature sequence length (VSL) access at individual normalized transmission rate $R_k \in \mathbb{Q}$, $k = 1, \ldots, K$, i.e. the fastest transmitting user has rate one.

The information stream $\{b_k[m]\}$, transmitted by user $k$, comprises $\hat{L}_{b,k} \in \mathbb{Z}^+$ binary bits per frame. The information bit sequence $b_k \triangleq \text{col}\{b_k[0], \ldots, b_k[\hat{L}_{b,k} - 1]\}$ is encoded by a user-independent encoder of rate $R$, mapped into the BPSK modulated data sequence $\hat{c}_k \triangleq \text{col}\{\zeta_k[0], \ldots, \zeta_k[\hat{L}_{d,k} - 1]\}$, where $\hat{L}_{d,k} = \hat{L}_{b,k}/R \in \mathbb{Z}^+$, and fed into user specific random channel interleaver $\Pi_k$.

The overall frame comprises $F \in \mathbb{Z}^+$ blocks, and each block consists of $L_{d,k} \in \mathbb{Z}^+$ data symbols multiplexed with $L_{p,k} \in \mathbb{Z}^+$ random preamble symbols. Hence, the users’ block and frame lengths equal respectively $L_k = L_{d,k} + L_{p,k}$ and $\hat{L}_k = L_k F$ code symbols. Notice further that the average
number of information bits per code symbol for user $k$ is $\hat{L}_{b,k}/\hat{L}_k = R(L_k - L_{D,k})/L_k$. Real-valued signature waveforms $s_{k,l}(t)$, having duration $T_k$ and spreading factor $N_k$, are assigned to user $k$ during signaling interval $l$, $k = 1, \ldots, K$, $l = 0, \ldots, \hat{L}_k - 1$. The waveforms are normalized to have unit energy, i.e. $\int_{T_k}^{(l+1)T_k} |s_{k,l}(t)|^2 \, dt = 1$. Hence, the signature waveform of the fastest transmitting users has duration $T_k = R_k T_h$. Note that the signature waveform of user $k$ has power proportional to $R_k$. The modulated waveforms are transmitted over frequency-nonselective Rayleigh fading channels, having gain $\alpha_1(t), \ldots, \alpha_K(t)$, that are assumed to be constant over each block interval. The received complex baseband signal reads

$$r(t) = \sum_{k=1}^{K} \sum_{l=0}^{\hat{L}_k-1} \alpha_{k,l} \zeta_k[l] s_{k,l}(t - lT_k - \tau_k) + w(t). \quad (1)$$

In the above expression $\tau_k$ is the transmission delay of user $k$. We assume that the transmission delays of the $K$ users are independent and Uniform distributed over $[0,T_h)$. Without loss of generality, the users are ranked according to their relative delays, i.e. $0 \leq \tau_1 \leq \cdots \leq \tau_K \leq T_h$. The complex Gaussian random variable $\alpha_{k,l} = \alpha_k([l/L_k]L_kT_k)$ with zero mean and variance $\sigma_\alpha^2$ denotes the channel gain of the signal transmitted by user $k$ for data block $[l/L_k)$. The symbol $\lfloor \cdot \rfloor$ designates the greatest integer not larger than the argument. Furthermore, $w(t)$ represents Additive complex-valued White Gaussian Noise (AWGN) with single-sided spectral height $N_0$. For the sake of convenience, we define $\alpha \triangleq \text{col}\{\alpha_1, \ldots, \alpha_K, \zeta_k \}$ as the column vector with the channel coefficients as its entries; $\zeta \triangleq \text{col}\{\zeta_1, \ldots, \zeta_K\}$ as the channel vector incorporating the data symbols of all users, and $\zeta_k \triangleq \text{col}\{\zeta_1, \ldots, \zeta_{k-1}, \zeta_{k+1}, \ldots, \zeta_K\}$ the column vector containing the data symbols of all users but those of user $k$.

At the receiver site, we model the multi-rate system as equivalent single-rate system, and we use the notation introduced in [23]. For each user $k$, the output of a filter matched to the signature waveform $s_{k,l}(t)$ is sampled at time instant $(l'+1)T_k + \tau_k$ with $k = 1, \ldots, K$, $l' = 0, \ldots, \hat{L}_{\text{max}} - 1$ where $\hat{L}_{\text{max}} = \max_k \hat{L}_k$. The number of samples per signaling interval $T_k$ is $R_k^{-1}$. The channel vector $z[l'] \triangleq \text{col}\{z_1[l'], \ldots, z_K[l']\} \in C^K$ contain the output samples of the $K$ Matched Filters (MF) in bin $l'$. Notice that bin $l'$ corresponds to the signaling interval $I_{kh}[l'] = [R_k l']$ of user $k$. It follows from (1) that


Here, the symbol $(\cdot)^T$ denotes transposition. The entries of the matrix $R[l', 0] \in C^{K \times K}$ are the zero-offset cross-correlations between the signature waveforms $s_{j,l}(t)$ and $s_{k,l}(t)$ in bin $l'$, i.e.

$$[R[l', 0]]_{jk} \triangleq \begin{cases} \rho_{jkh}[l', 0]; & j < k \\ R_k; & j = k \\ \rho_{jkh}[l', 0]; & j > k \end{cases} \quad (3)$$

with

$$\rho_{jkh}[l', 0] = \int_{T_k}^{(l'+1)T_k + \tau_k} s_{j,l}(t - l_j[l']T_k - \tau_j) \times s_{k,l}(t - l_k[l']T_k - \tau_k) \, dt.$$  

The entries of the upper-triangular matrix $R[l', 1] \in C^{K \times K}$ are the one-offset cross-correlations between the signature waveforms $s_{j,l}(t)$ and $s_{k,l}(t)$ in bin $l'$, i.e.,

$$[R[l', 1]]_{jk} \triangleq \begin{cases} \rho_{jkh}[l', 1]; & j < k \\ 0; & j \geq k \end{cases} \quad (4)$$

with

$$\rho_{jkh}[l', 1] = \int_{T_k}^{(l'+1)T_k + \tau_k} s_{j,l}(t - l_j[l']T_k - \tau_j) \times s_{k,l}(t + T_k - l_k[l']T_k - \tau_k) \, dt.$$  

The symbol diagonal matrix $C[l'] \in C^{K \times K}$ contains

$$c_k[l'] = \begin{cases} \zeta_k[l'][l'] & 0 \leq l' \leq \hat{L}_{\text{max}} - 1 \\ 0 & \text{elsewhere} \end{cases} \quad (5)$$

as its diagonal elements. The channel column vector $a[l'] \in C^K$ is defined as $a[l'] \triangleq \text{col}\{a_1[l'], \ldots, a_K[l']\}$.

$$a_k[l'] = \begin{cases} \alpha_k([l'/L_{\text{max}}]L_{\text{max}}T_k); & 0 \leq l' \leq \hat{L}_{\text{max}} - 1 \\ 0 & \text{elsewhere} \end{cases} \quad (6)$$

With the assumptions made above, $a[l']$ has the bin-independent covariance matrix $\Sigma_a \triangleq \text{diag}\{\sigma_1^2, \ldots, \sigma_K^2\}$. Finally, the column vector $n[l'] \in C^K$ contains zero-mean colored Gaussian noise [23]. Notice that the shortcuts $c_k[l']$ in (5) and $a_k[l']$ in (6) have only be introduced to describe the equivalent single-rate system. Clearly, $c_k[l']$ and $a_k[l']$ remain constant within, respectively, one signaling interval of user $k$, and one data block.

To ease the derivation of the SAGE-JDCE scheme, $z[l']$ is processed by subsequent noise whitening (anti-causal) filter. We start with the (non-singular) correlation matrix $R[l', z]$ of (2) in the $z$-transform domain at bin $l'$ that can be factored as [24]


The first and the second factors are given by, respectively,

$$F[l', z]^T = (F[l', 0] + F[l' + 1, 1]z)^T, \quad (7)$$

and

$$F[l', z^{-1}] = F[l', 0] + F[l', 1]z^{-1}. \quad (8)$$

The matrix $F[l', 0] \in C^K \times K$ is lower left triangular and the matrix $F[l', 1] \in C^K \times K$ is upper right triangular with zero diagonal, satisfying [25]

$$R[l', 0] = F[l', 0]^T F[l', 0] + F[l' + 1, 1]^T F[l' + 1, 1], \quad (9)$$

and

$$R[l', 1] = F[l', 0]^T F[l', 1], \quad (10)$$

where $F[0, 1] = F[\hat{L}_{\text{max}} - 1, 1] = R[0, 1] = R[\hat{L}_{\text{max}} - 1, 1] = 0$. If $z[l']$ is fed to the filter $(F[l', z]^T)^{-1}$, the output correlation matrix in the $z$-transform domain simply
reads $F(l', z^{-1})$ in (8), and thus, the corresponding output sequence $y[l'] \in \mathbb{C}^K$ yields

$$y[l'] = \sum_{\Delta l' = 0}^l F[l', \Delta l'] C[l' - \Delta l'] a[l' - \Delta l'] + w[l'], \quad (11)$$

The symbol $\Delta l'$ denotes the time-offset. The noise vector $w[l']$ is white, i.e., has covariance matrix $N_0 I$ with $I$ denoting the $K \times K$ identity matrix [26]. Conversely, from (7) it can be directly seen that the input $z[l']$ is related to the output $y[l']$ of the whitening filter according to

$$z[l'] = \sum_{\Delta l' = 0}^l F[l', \Delta l']^T y[l' + \Delta l'], \quad (12)$$

In the sequel the information bit sequences and the channel gains are unknown to the receiver and hence, require estimation.

III. THE SAGE-JDCE RECEIVER

A. The SAGE Algorithm

The Space Alternating Generalized Expectation-Maximization (SAGE) algorithm [5] can be used to iteratively approximate the maximum-likelihood estimate of $\zeta$ with respect to the observed (incomplete) data $y$ when a direct calculation of this estimate is computationally intractable.

We sketch the principle of the SAGE algorithm within the application context of the paper. At iteration $i$, only the symbols of user $k, \zeta_k$, are updated, while the symbols of the other users, $\zeta_i$, are kept fixed. The SAGE algorithm relies on the concept of a hypothetical (admissible hidden) data $\mathbf{X}_k$ with respect to $\zeta_k$ to which the incomplete data $y$ is related through a possibly non-deterministic mapping $\mathbf{X}_k \mapsto y(\mathbf{X}_k)$. In our application the admissible hidden data $\mathbf{X}_k$ incorporates the channel gains, which are nuisance parameters that would aid in the estimation of $\zeta$, but cannot be observed. In a first step called expectation step (E-step), the algorithm computes an estimate

$$Q(\zeta_k | \zeta_i[i]) \triangleq E \left\{ \Lambda (\mathbf{X}_k; \zeta_k, \zeta_i[i]) | y, \zeta_i[i] \right\}$$

of the log-likelihood function $\Lambda(\cdot)$ of $\zeta_k$ for the observation $\mathbf{X}_k$ based on $y$ and a current estimate $\zeta^{[i]}_k$ of $\zeta_k$. Then in a second step called maximization step (M-step) $\zeta_k$ is updated as the value that maximizes this estimated log-likelihood function. Starting from an initial estimate $\zeta^{[0]}_k$, the E-step and the M-step are successively repeated in subsequent iterations until convergence is achieved. The SAGE algorithm exhibits the monotonicity property: the sequence $\left\{ \Lambda (y[k^{[i]}]) \right\}_{i=0}^\infty$ is non-decreasing, i.e., the likelihood of the symbol estimates is non-decreasing during the iterative processing.

B. Receiver Design

In our application of the SAGE algorithm to JDCE, we make use of the noise splitting approach [7] to decompose the output signal vector of the whitened matched filter in (11) into the sum of two terms

$$y[l'] = x_k[l'] + \bar{x}_k[l'], \quad (13)$$

The first term represents a noisy version of the $k$th user’s signal:

$$x_k[l'] = \sum_{\Delta l' = 0}^l F_k[l', \Delta l'] c_k[l' - \Delta l'] a_k[l' - \Delta l'] + w_k[l'], \quad (14)$$

$k = 1 \ldots K$, with $F_k[l', \Delta l']$ denoting the $k$th column of $F[l', \Delta l']$. The white Gaussian noise vector $w_k[l']$ has variance $\beta_k N_0$. The second term in (13) corresponds to the $k$th user’s MAI plus the vector $\bar{w}_k[l']$ containing the remaining noise with variance $(1 - \beta_k) N_0$. The weighting coefficient $\beta_k$ is a free parameter. It will be used to optimize the performance of the receiver in Section III-C. The only constraint is $0 \leq \beta_k \leq 1$. Notice that if $\alpha$ was known, then $x_k$ would be a natural admissible data for $\zeta_k$ under the assumption that the other users’ symbols are known. Since $\alpha$ is not known, we apply a standard approach to incorporate $\alpha$ into the admissible data $\mathbf{X}_k$. Hence, $\mathbf{X}_k = \{x_k, \alpha\}$, where $x_k \in \co \{x_k[0], \ldots, x_k [L_{\text{max}} - 1]\}$. In this case, the E-step of the SAGE algorithm is given by

$$Q(\zeta_k | \zeta_i[i]) = E \left\{ \Lambda(x_k | \alpha, \zeta_k, \zeta_i[i]) | y, \zeta_i[i] \right\}.$$  (15)

Inserting (14) into (15) and neglecting terms independent of $\zeta_k$ yields

$$Q(\zeta_k | \zeta_i[i]) \propto \sum_{l=0}^{L-1} \zeta_k[l] \sum_{l' = R_k^{-1} l + 1}^{R_k^{-1} l + 1 - 1} 1 \sum_{\Delta l' = 0}^l \text{Re} \left\{ F_k[l' + \Delta l', \Delta l']^T (a_k[l'] \cdot x_k[l' + \Delta l']) | \zeta_k[l] \right\}.$$  (16)

By iterating conditional expectations, the joint estimate $(a_k[l'] \cdot x_k[l'])^{[i]} = E \left\{ a_k[l'] \cdot x_k[l'] | y, \zeta_i[i] \right\}$ can be computed as

$$(a_k[l'] \cdot x_k[l'])^{[i]} = E \left\{ a_k[l'] \cdot E \left\{ x_k[l'] | y, \zeta_i[i], \alpha \right\} \right\} | y, \zeta_i[i]. \quad (17)$$

The symbol $(\cdot)^*$ denotes conjugate complex of the argument. The conditional expectation of $x_k[l']$ in the right hand term in (17) reads [13]

$$E \left\{ x_k[l'] | y, \zeta_i[i], \alpha \right\} = \sum_{\Delta l' = 0}^1 F_k[l', \Delta l'] c_k[l' - \Delta l']^{[i]} a_k[l' - \Delta l'] + \beta_k \left( y[l'] \right) - \sum_{j=1}^K \sum_{\Delta m' = 0}^1 F_j[l', \Delta m'] c_j[l' - \Delta m']^{[i]} a_j[l' - \Delta m'] \right\}, \quad (18)$$

where the estimate of the data symbol $c_k[l'^{[i]}]$ in the equivalent single-rate system and the corresponding estimate $\zeta_k[l_k[l'^{[i]}]]$ in the multi-rate system are related by (5). Inserting (18) into (17), we obtain (19), shown on the top of the next page. Since the channel is quasi-static, the conditional channel estimate $a_k[l'^{[i]}]$ in data block $f, f = 0, \ldots, F - 1$, and its covariance
\[
(a_k[l']^i \cdot x_k[l'])^i = \sum_{\Delta l'=0}^{1} F_k[l', \Delta l'] c_k[l' - \Delta l']^i (a_k[l']^i a_k[l' - \Delta l'])^i + \beta_k \left( a_k[l']^i \cdot y[l'] - \sum_{j=1}^{K} \sum_{\Delta m'=0}^{1} F_j[l', \Delta m'] c_j[l' - \Delta m']^i (a_k[l']^i a_j[l' - \Delta m'])^i \right).
\] (19)

Matrix can be computed to be [13]

\[
a_k[f L_{\text{max}}]^i = \left[ E \{ a[f L_{\text{max}}] y, \zeta[i] \} \right]_k = \left[ N_0^{-1} \Sigma_{a[f L_{\text{max}}]}^{(f+1) L_{\text{max}} - 1} \sum_{m'=f L_{\text{max}}}^{(f+1) L_{\text{max}} - 1} C[m']^i z[m'] \right]_k,
\] (20)

and

\[
\Sigma_{a[f L_{\text{max}}]}^{(f+1) L_{\text{max}} - 1} = N_0 \left( N_0^{-1} \sum_{m'=f L_{\text{max}}}^{(f+1) L_{\text{max}} - 1} R_{\text{ca}[m']} \right)^{-1},
\]

respectively. The short-cut symbol \( R_{\text{ca}[m']} \) is assigned to

\[
R_{\text{ca}[m']} \triangleq C[m']^i \left( \sum_{\Delta m'=0}^{1} R[m', \Delta m'] C[m' - \Delta m']^i + R[m' + 1, 1]^T C[m' + 1]^i \right),
\]

and \( L_{\text{max}} = \max_k L_k \).

Without loss of generality, we have considered the 0th bin of each data block \( f \) in (20). Notice that in the equivalent single-rate system, the conditional channel estimates in the other \( L_{\text{max}} - 1 \) bins of each block \( f \) are identical. Inserting (19) into (16), while considering the identities (9), (10), and (12), we obtain after some algebraic manipulations

\[
Q \left( \zeta_k l^i \right) = \sum_{l=0}^{L_k - 1} \zeta_k[l] \lambda_k(l, \beta_k)
\] (21)

with the branch definition

\[
\lambda_k(l, \beta_k) \triangleq \begin{cases} R_k^{(1+1)-1} - \beta_k \sum_{l'=R_k^{(1+1)}}^{R_k^{(1+1)-1}} \Re \left\{ \rho_{kk}[l'] Q c_k[l']^i (a_k[l']^i)^i \right\} & \text{asymptotically a priori information} \\ + \beta_k \sum_{m'=R_k^{(1+1)}}^{R_k^{(1+1)-1}} \Re \left\{ a_k[m']^i z_k[m'] - \sum_{j \neq k} I_j[m']^i \right\} & \text{asymptotically extrinsic information} \end{cases}
\] (22)

and an estimate of the interference term

\[
I_j[l']^i \triangleq \sum_{\Delta l'=0}^{1} \rho_{jk}[l', \Delta l']' c_j[l' - \Delta l']^i (a_k[l']^i a_j[l' - \Delta l'])^i + \rho_{jk}[l' + 1, 1] c_j[l' + 1]^i (a_k[l']^i a_j[l' + 1])^i.
\] (23)

The M-step of the SAGE algorithm updates a new fit of the data sequence estimate according to

\[
\zeta_k^{[i+1]} = \arg \max_{\zeta_k} Q \left( \zeta_k l^i \right), \quad \zeta_k^{[i+1]} = \zeta_k^{[i]}.
\] (24)

A Maximum Likelihood Sequence Decoder (MLSD) such as the Viterbi algorithm can be used to evaluate (24) within the code \( C \). The objective function \( Q \left( \zeta_k l^i \right) \) in (21) acts as path metric in the Viterbi algorithm. If coding is not present, the M-step of the SAGE algorithm simplifies to [13]

\[
b_k[m]^{[i+1]} = \text{sgn} \left\{ \sum_{\nu=H_k^{-1}+m}^{R_k^{-1}(m+1)-1} \Re \left\{ (1 - \beta_k) R_k c_k[l']^i (a_k[l']^i)^i \right\} + \beta_k \left( a_k[l']^i \cdot z_k[l'] - \sum_{j \neq k} I_j[l']^i \right) \right\}.
\] (25)

Inspecting (21), we can see how the SAGE-JDCE scheme operates at each iteration: first, the users’ channel coefficients \( a_k[l] \) in (20) are updated in parallel block by block. Based on these new channel estimates, contributions to the \( k \)th user’s MAI in (23) are formed at a common rate, i.e., the rate of the fastest transmitting users. The cleaned signal components from each bin \( l' \) are then soft-combined to form the branch values \( \lambda_k(l, \beta_k) \) in (22) at symbol rate and fed into a MLSD that decodes the \( k \)th user’s data sequence. A block diagram of the SAGE-JDCE receiver is shown in Fig. 1. Notice that the whitening filter (see (12)) has only been used to simplify the derivation of the SAGE-JDCE scheme. Indeed, computation of the E-step (15) and the M-step (24) of the SAGE-JDCE scheme do not require knowledge of this whitening filter. For each bin \( l' \), a new fit of the \( k \)th user’s signal sequence is computed using a fraction of its current estimate based on an estimate of MAI from the current iteration and a part of its tentative bit decision obtained from the previous iteration cf. (21) and Fig. 2. The weighting coefficients \( \beta_k \) determine the balance. Notice that this operation corresponds to partial successive interference cancellation.

Looking at the same problem from a different point of view, it can be seen that the amount of noise, split between user \( k \) and its MAI, is governed by \( \beta_k \). Furthermore, the factor associated with \( (1 - \beta_k) \) and with \( \beta_k \) in (22) can be identified with a priori information and extrinsic information, respectively, as known from iterative decoding of channel codes [21]. For known channels, we observe the following: If \( \beta_k = 0 \), according to (21) and (22) the SAGE-JDCE scheme passes a priori values on the \( k \)th user’s code symbols from
the interference cancellation (IC) device to the MLSD. On the other hand, if $\beta_k = 1$, extrinsic values on the $k^{th}$ user’s code symbols are forwarded. In the case of unknown channels, the situation is slightly different. The channel estimate is influenced by the previous estimates of all symbols, and thus, the information passed to the MLSD is neither pure extrinsic nor pure a priori. However, the influence of the $k^{th}$ user’s $l^{th}$ symbol on its channel estimate vanishes as $L_{\text{max}} \to \infty$. Therefore, we refer to the two types of soft information occurring in (22) as asymptotically a priori and asymptotically extrinsic, respectively.

C. Noise Splitting

As earlier mentioned the weighting coefficient vector $\beta_k$ can be used to split noise between the $k^{th}$ user and its MAI. This splitting can be optimized in such a way that the joint bit-error-probability is minimized as the number of iterations tend to infinity. This task, however, is difficult to meet due to the feedback and the non-linear structure of the SAGE-JDCE scheme on the one hand, and the code constraints on the other hand. Hence, we rely on sub-optimal methods for optimizing the weighting coefficients. In [15], the performance of an EM-based JDE scheme is enhanced using two optimization criteria. The first one is within, the second one outside the EM framework. In the sequel, the latter criterion is adapted to the SAGE-JDCE scheme within the SAGE framework so that the monotonicity property holds and therefore convergence is ensured. As a result we will get new insight into joint partial interference cancellation and data decoding.

Taking into account that the $k^{th}$ user’s signal energy is spread over several adjacent signaling intervals, a suitable sub-optimal optimization criterion for the weighting coefficient is to project different time-shifted versions of the estimation error $(a_k[l']^*x_k[l'] - (a_k[l']^*x_k[l'])^{|i|})$ belonging to the $k^{th}$ user’s $l^{th}$ symbol on the corresponding versions of $F_k[l', \cdot]$, take the conditional expectation of the mean squared sum with respect to the estimated symbols of all but the $k^{th}$ user, $S_k^{[i]}$, and minimize the result according to

$$\hat{\beta}_k^{[i]} = \arg \min_{\beta_k} E \left\{ \frac{1}{|J_k|} \sum_{l' = 0}^{L_{\text{max}}} F_k[l' + \Delta l', \Delta l']^T (a_k[l']^*x_k[l'] + \Delta l') \right\}$$

By invoking the relations (9) and (10), it follows from (14) for the projected true signal vector in (26)

$$a_k[l']^* \sum_{\Delta l' = 0}^{L_{\text{max}}} F_k[l' + \Delta l', \Delta l']^T x_k[l' + \Delta l'] = R_k c_k[l'] (a_k[l']^*)^2$$

Equation (27) and (28) for (26), making use of $E \{ a_k[l']^*n_k[l'] \} = 0$, $j \neq k$ and the identity $\sqrt{\beta_k} n_k[l'] = \sum_{\Delta l' = 0}^{L_{\text{max}}} F_k[l' + \Delta l', \Delta l']^T w_k[l' + \Delta l']$, and dropping terms independent of $\beta_k$, the minimization problem is equivalent to

$$\hat{\beta}_k^{[i]} = \arg \min_{\beta_k} \left\{ \frac{\beta_k \sigma_k^2}{2} E \left\{ n_k[l']^*n_k[l'] | \xi_k^{[i]} \right\} - 2\beta_k \left( E \left\{ |T_k[l'] - T_k[l']|^{[i]} | \xi_k^{[i]} \right\} \right) + \beta_k \sigma_k^2 E \left\{ n_k[l']^*n_k[l'] \right\} \right\} \right\}$$

Notice that $E \{ n_k[l']^*n_k[l'] | \xi_k^{[i]} \} = R_k N_0$. Long codes can be modeled as random codes with Uniform distributed delay $\tau_n \in [0, T_c]$, $k = 1, \ldots, K$, if the period of the pseudo-noise waveforms is much larger than the symbol period $T_c$. Hence, $E \{ \rho_{j,k}[l', 0] | \rho_{j,k}[l', 1] \} = 0$. The remaining expectations in (29) can be computed as

$$E \left\{ n_k[l']^*n_k[l'] | \xi_k^{[i]} \right\} = \begin{cases} \sigma_k^2 \left( 1 - E \left\{ c_j[l'] c_j[l']^{[i]} \right\} \right) = 12 \sigma_k^2 P_{j,j} \left| c_j[l']^{[i]} \right| ; j = k \\
0 \quad ; j \neq k \end{cases}$$

with the error-probability $P_{j,j} \left| c_j[l']^{[i]} \right| \triangleq P \{ c_j[l']^{[i]} \neq c_j[l'] \}$ of the re-encoded bits. Following this approach, our optimization
problem reduces to

$$\hat{\beta}_k^{[i]} = \arg\min_{\beta_k} \{ - \beta_k (2\Gamma_k - 1) + \beta_k^2 (\Gamma_k - 1) \},$$

where $\Gamma_k$ is the short-cut for $\Gamma_k \triangleq 12 R_k \gamma_{b,k} P_{c,k} |P_{[\cdot]}^{[\cdot]}$ and $\gamma_{b,k} \triangleq \frac{\gamma_k}{\gamma_0}$ denotes the $k$th user’s average-signal-to-noise ratio (SNR). It is evident that the objective function $f(\beta_k)$ in (31) might range from convex ($\Gamma_k > 1$), through linear ($\Gamma_k = 1$), to concave ($\Gamma_k < 1$) in $\beta_k$. In the first case, the minimum of (31) is achieved by differentiating, setting the result equal zero, and solving with respect to $\beta_k$, within its constraint interval $[0, 1]$. Otherwise, evaluation of $f(\beta_k)$ within $[0, 1]$ is required. With the definition

$$R \triangleq \frac{2\Gamma_k - 1}{2(\Gamma_k - 1)},$$

we readily obtain the solution to our optimization problem:

$$\hat{\beta}_k^{[i]} = \begin{cases} R(u\{R\} - u\{R - 1\}) + u\{R - 1\} & : \Gamma_k > 1 \\ u\{-R + \frac{1}{2}\} & : \Gamma_k < 1, R \neq \frac{1}{2} \\ \in \{0, 1\} & : \Gamma_k < 1, R = \frac{1}{2} \end{cases}$$

In the above expression $u\{\cdot\}$ denotes the unit-step function. Fig. 3 illustrates the behavior of the weighting coefficients. It can be seen that for finite values of $\Gamma_k$, i.e. finite average SNR, the objective function $f(\beta_k)$ has one minimum at $\hat{\beta}_k = 1$ regardless of the individual bit-error-probability $P_{c,k}[P_{[\cdot]}^{[\cdot]}]$, implying that the optimized SAGE-JDCE scheme forwards asymptotically extrinsic values on all user’s symbols in consecutive iterations. Notice that this behavior is independent of the number of users $K$ and interleaver size $L_k$ in contrast to the approach followed by Boutros and Caire in [20]. It is worth mentioning that no knowledge of the whitening filter is required to calculate $\hat{\beta}_k^{[i]}$.

**D. Initialization**

The MMSE estimate of $\mathbf{a}$ given the MF-output $\mathbf{z}$ and the user specific (uncoded) pilot symbols yields the initial estimate $\hat{\mathbf{a}}^{[0]}$. Notice that the first and last bin in each preamble block have been omitted from initial channel estimation due to unknown MAI arising from the adjacent bins to the preamble. To keep the computational complexity low, the output signal of the $k$th user’s MF, weighted by $\hat{\mathbf{a}}^{[0]}$, $k = 1, \ldots, K$ forms the initial symbol estimate $\hat{\mathbf{z}}^{[0]}$.

**E. Computational Complexity**

The time complexity of the proposed SAGE-JDCE scheme can be quantified by its number of Floating Point OPerations (FLOPs) per user and information bit. To improve efficiency of the scheme, complexity reducing means have been incorporated such as exploiting matrix structure, reusing mathematical operations, and replacing power series $^1$ by their respective sum. In addition, matrix inversion has been realized by LU decomposition [26]. Following this approach, Table I reveals the contribution of the key constituent blocks to the time complexity of the SAGE-JDCE scheme assuming that convergence is achieved after $S$ stages. Within one stage every user’s data sequence is updated once. Hence, one stage corresponds to $K$ iterations. It can be readily seen that first, re-estimation of the channel coefficients in the E-step dominates the time complexity of the SAGE-JDCE scheme and secondly, the SAGE-JDCE scheme works more efficiently the longer the length of the information block is. In the limiting case of $L_k \rightarrow \infty$ the time complexity per user and information bit of the SAGE-JDCE scheme is bounded by $O\{SK^2\}$.

**IV. Numerical Examples**

The feedback structure of EM-based JDE schemes makes a theoretical performance analysis cumbersome. We therefore evaluate the performance of the SAGE-JDCE scheme by means of Monte-Carlo simulations. The system parameters are chosen as follows: All users employ a rate $R = 1/2$ convolutional code with generator matrix $G = [5_b, \gamma_8]$. Consecutive segments of a user-specific Gold sequence [27] of length $2^{20} - 1$ form the individual signature waveform during consecutive signallng intervals.

For comparison purpose, the performance of the initial MF and a windowed MMSE linear MU detector [28] with window size of 20 bins including 2 excess bins on each side have also been included in subsequent plots. Both schemes are supplemented with a separate MMSE channel estimator as described in Section III-D and a separate MLSD. We refer to these decoding methods as respectively MF separate decoding and channel estimation (MF-SDCE) and MMSE-SDCE.

The average bit-error-rate $BER$ of the proposed receiver is plotted in Fig. 4 versus the average effective SNR $\gamma = \frac{L_p}{L_k} \gamma_{b,k}$, $k = 1, \ldots, K$, for a scenario with $K = 8$ and $K = 32$ equal average power and equal rate users, spreading factor $N_c \triangleq N_1 = \ldots = N_K = 8$, $L_p = 4$ pilot symbols, block length $L = 40$ and frame length $\hat{L} = 800$. Notice that the effective SNR also accounts for the energy of the transmitted pilot bits. The number of stages in the receiver is limited to 20. Notice that averaging is performed over both the realizations of the channel coefficients and the users. Hence, all users perform on average the same regardless of their transmission rate. For $BER = 10^{-2}$, the performance loss of the SAGE-JDCE scheme w.r.t to the SU in known channel is 1.1 dB and 2.8 dB for an effective system load $\vartheta \triangleq K R / N_c = 0.5$ ($K = 8$) and $\vartheta = 2$ ($K = 32$) respectively. In the latter case, however, the SAGE-JDCE scheme has difficulties with handling MAI and hence, the $BER$ curve floors out as the average SNR tends to infinity. Notice that the effective system load incorporates the rate of the channel codes as well. In contrast, the loss in performance of the MMSE-SDCE scheme w.r.t to the SU-case in known channel is 5.6 dB and larger than 10 dB, respectively.

The average bit-error-rate in a scenario with triple-rate transmission according to $N_1 = \ldots = N_{K/3} = 16$ ($R_1, \ldots, R_{K/3} = 0.25$), $N_{K/3+1} = \ldots = N_{K/2} = 8$ ($R_{K/3+1}, \ldots, R_{K/2} = 0.5$), and $N_{K/2+1} = \ldots , N_K = 4$ ($R_{K/2+1}, \ldots , R_K = 1$), is depicted in Fig. 5. For a fair comparison, the system parameters

$^1$We need $\sum_{n=1}^{N} n = (N^2 + N)/2$ and $\sum_{n=1}^{N} n^2 = N(N + 1)(2N + 1)/6$. 

by LU decomposition [26]. Following this approach, Table I reveals the contribution of the key constituent blocks to the time complexity of the SAGE-JDCE scheme assuming that convergence is achieved after $S$ stages. Within one stage every user’s data sequence is updated once. Hence, one stage corresponds to $K$ iterations. It can be readily seen that first, re-estimation of the channel coefficients in the E-step dominates the time complexity of the SAGE-JDCE scheme and secondly, the SAGE-JDCE scheme works more efficiently the longer the length of the information block is. In the limiting case of $L_k \rightarrow \infty$ the time complexity per user and information bit of the SAGE-JDCE scheme is bounded by $O\{SK^2\}$.

**IV. Numerical Examples**

The feedback structure of EM-based JDE schemes makes a theoretical performance analysis cumbersome. We therefore evaluate the performance of the SAGE-JDCE scheme by means of Monte-Carlo simulations. The system parameters are chosen as follows: All users employ a rate $R = 1/2$ convolutional code with generator matrix $G = [5_b, \gamma_8]$. Consecutive segments of a user-specific Gold sequence [27] of length $2^{20} - 1$ form the individual signature waveform during consecutive signallng intervals.

For comparison purpose, the performance of the initial MF and a windowed MMSE linear MU detector [28] with window size of 20 bins including 2 excess bins on each side have also been included in subsequent plots. Both schemes are supplemented with a separate MMSE channel estimator as described in Section III-D and a separate MLSD. We refer to these decoding methods as respectively MF separate decoding and channel estimation (MF-SDCE) and MMSE-SDCE.

The average bit-error-rate $BER$ of the proposed receiver is plotted in Fig. 4 versus the average effective SNR $\gamma = \frac{L_p}{L_k} \gamma_{b,k}$, $k = 1, \ldots, K$, for a scenario with $K = 8$ and $K = 32$ equal average power and equal rate users, spreading factor $N_c \triangleq N_1 = \ldots = N_K = 8$, $L_p = 4$ pilot symbols, block length $L = 40$ and frame length $\hat{L} = 800$. Notice that the effective SNR also accounts for the energy of the transmitted pilot bits. The number of stages in the receiver is limited to 20. Notice that averaging is performed over both the realizations of the channel coefficients and the users. Hence, all users perform on average the same regardless of their transmission rate. For $BER = 10^{-2}$, the performance loss of the SAGE-JDCE scheme w.r.t to the SU in known channel is 1.1 dB and 2.8 dB for an effective system load $\vartheta \triangleq K R / N_c = 0.5$ ($K = 8$) and $\vartheta = 2$ ($K = 32$) respectively. In the latter case, however, the SAGE-JDCE scheme has difficulties with handling MAI and hence, the $BER$ curve floors out as the average SNR tends to infinity. Notice that the effective system load incorporates the rate of the channel codes as well. In contrast, the loss in performance of the MMSE-SDCE scheme w.r.t to the SU-case in known channel is 5.6 dB and larger than 10 dB, respectively.

The average bit-error-rate in a scenario with triple-rate transmission according to $N_1 = \ldots = N_{K/3} = 16$ ($R_1, \ldots, R_{K/3} = 0.25$), $N_{K/3+1} = \ldots = N_{K/2} = 8$ ($R_{K/3+1}, \ldots, R_{K/2} = 0.5$), and $N_{K/2+1} = \ldots , N_K = 4$ ($R_{K/2+1}, \ldots , R_K = 1$), is depicted in Fig. 5. For a fair comparison, the system parameters
are chosen as $L_{\text{p,max}} = 8$, $L_{\text{max}} = 80$, and $\tilde{l}_{\text{max}} = 1600$. Then for $K = 6$ and $K = 24$, the system loads again yield $\vartheta = R \sum_k 1/N_k = 0.5$ and $\vartheta = 2$ respectively. It can be seen that on average the SAGE-JDCE scheme copes with triple-rate transmission almost as good as it does with single-rate transmission. The small difference in performance can be explained by the fact that low-rate users are supported by less pilot symbols than high-rate users. On top of that, low-rate users employ smaller interleaver sizes than high-rate users do. Consequently, the SAGE-JDCE scheme converges to a local maximum of the likelihood function with larger probability. In particular, the performance loss of the SAGE-JDCE scheme w.r.t. to the SU-case in known channel is $\approx 0.5$ dB for the same settings as in the previous examples. For single-rate transmission, $\vartheta = 2$ (high load), and low effective SNR, the SAGE-JDCE schemes requires roughly 14 stages to converge. Due to the high error rate, the point of convergence corresponds obviously to a local maximum of the likelihood function in most of the cases. With increasing effective SNR, the convergence curve has a maximum around an effective SNR of $6$ dB. This maximum corresponds to the region of the BER curve, in which the iterative receiver starts to operate properly. This effect is well-known for Turbo codes operating at the decoding threshold [29]. We can conclude that the iterative receiver converges now more often to the global maximum of the likelihood function (corresponding to a low BER) but requires many stages to achieve it. When the effective SNR is further increased, the convergence speed of the SAGE-JDCE scheme increases. In the asymptotic case, i.e., when the effective SNR approaches infinity, one stage is sufficient to achieve convergence.

A similar effect can also be observed for $\vartheta = 0.5$, corresponding to a low load: the convergence curve exhibits a maximum at about $-4$ dB. For the triple-rate scenario, a similar behavior is expected, but cannot be observed as the maximal number of stages was limited to 20 in the simulations.

V. SUMMARY

An iterative receiver for coded DS/CDMA has been derived based on the SAGE algorithm and the noise splitting approach. This receiver performs joint channel estimation, partial successive interference cancellation, and data decoding. Soft-values on the coded symbols computed by the interference cancellation device are forwarded to the decoder, and hard values on the coded symbols are returned to the channel estimation/interference cancellation device. The receiver design leads to noise-weighting coefficients that can be chosen under weak constraints within the SAGE framework to control the quality of the soft values passed to the decoder. We have formulated an optimization criterion, and we have derived an analytical expression for the optimal weighting coefficients. In the optimized receiver, the values computed in the interference cancellation device and passed to the single-user decoder are asymptotically extrinsic values in most cases. Remarkably, this holds independently of the interleaver size, and thus provides a theoretical argument for the use of extrinsic values, as done in many heuristic approaches for multiuser decoding. Monte-Carlo simulations for synchronous coded DS/CDMA over flat Rayleigh fading channels show that the proposed SAGE-based receiver can support a large system load of $\vartheta = R \sum_k 1/N_k \leq 2$ (taking into account the rate of the channel codes), which justifies the proposed receiver design.

VI. ACKNOWLEDGEMENT

The authors want to thank Dr. Joachim Wehinger, NewLogic Technologies, Sophia-Antipolis, France, for the valuable discussions.

REFERENCES


Bernard H. Fleury received the diploma in electrical engineering and mathematics in 1978 and 1990 respectively, and the doctoral degree in electrical engineering in 1990 from the Swiss Federal Institute of Technology Zurich (ETHZ), Switzerland.

Since 1997 Bernard H. Fleury has been with the Department of Communication Technology, Aalborg University, Denmark, where he is Professor in Digital Communications. He has also been affiliated with the Telecommunication Research Center, Vienna (ftw.) since April 2006. Bernard H. Fleury is presently Chairman of Department 2 Radio Channel Modelling for Design Optimisation and Performance Assessment of Next Generation Communication Systems of the on-going FP6 network of excellence NEWCOM (Network of Excellence in Communications). During 1978-85 and 1988-92 he was Teaching Assistant and Research Assistant, respectively, at the Communication Technology Laboratory and at the Statistical Seminar at ETHZ. In 1992 he joined again the former laboratory as Senior Research Associate. In 1999 he was elected IEEE Senior Member.

Bernard H. Fleury’s general fields of interest cover numerous aspects within Communication Theory and Signal Processing mainly for Wireless Communications. His current areas of research include stochastic modelling and estimation of the radio channel, characterization of multiple-input multiple-output (MIMO) channels, and iterative (turbo) techniques for joint channel estimation and data detection/decoding in multi-user communication systems.

Alexander Kocian received the Dipl. Ing. degree (with distinction) from Vienna University of Technology, Austria, in 1997, and the Ph. D. degree from Aalborg University (AAU), Denmark, in 2003, both in electrical engineering.

From 1997-1999 he was with the Spread Spectrum Team at the Communication Technology Laboratory at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. From 1999-2005 he served as Teaching Assistant and later on as Assistant Professor at the Digital Communications Laboratory at AAU. In 2001, AAU sponsored him a 6-month Visiting Research Scholarship at the Wireless Systems Laboratory, Georgia Institute of Technology, Atlanta, GA, USA. Since 2005 he has been Reader at the Department of Electronics Engineering at Birla Institute of Technology (BIT), Muscat Center, Sultanate of Oman with a collaboration program between AAU and BIT.

His research interests include iterative information processing in multiple-access communication systems and characterization of multiple-input multiple-output (MIMO) channels.

Ingmar Land is Assistant Professor for communication theory at Aalborg University, Denmark. His research topics are channel coding, iterative decoding, and information theory. He received his Dr.-Ing. degree at the University of Kiel, Germany, in 2004, where he was employed as research and teaching assistant. He studied electrical engineering at the Universities of Ulm and Erlangen-Nürnberg, Germany, where he received his Dipl.-Ing. degree in 1999.
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Time complexity of the constituent blocks of the SAGE-JDCE scheme</td>
<td>11</td>
</tr>
</tbody>
</table>

LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Block diagram of the SAGE-JDCE scheme</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>Illustration of (22): Joint channel estimation and partial interference cancellation.</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>Behavior of the objective function $f(\beta_k)$ in (31); $\Gamma_k = 12 R_k \frac{\sigma_k^2}{N_0} P_c[k][l]$ and $R_k \triangleq \frac{P}{L}$</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>Performance of a single-rate system in quasi-static flat Rayleigh fading for various system loads $\vartheta$: $N_c = 8$, $L_d = 4$, $L = 40$, $\hat{L} = 800$.</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>Performance of a triple-rate system in quasi-static flat Rayleigh fading for various system loads $\vartheta$: $N_1, \ldots, N_{K/3} = 16$, $N_{K/3+1}, \ldots, N_{K/2} = 8$, $N_{K/2+1}, \ldots, N_K = 4$, $L_{p,max} = 8$, $L_{max} = 80$, $\hat{L}_{max} = 1600$.</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>Convergence speed of the SAGE-JDCE scheme for various system loads $\vartheta$.</td>
<td>16</td>
</tr>
<tr>
<td>Block</td>
<td>FLOPs per user $k$ and information bit</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------</td>
<td></td>
</tr>
<tr>
<td>$\alpha^{[\ell]}$</td>
<td>$\frac{2^K}{T_{\text{max}}} + \mathcal{O}{K}$</td>
<td></td>
</tr>
<tr>
<td>$\beta^{[\ell]}$</td>
<td>$\mathcal{O}{K}$</td>
<td></td>
</tr>
<tr>
<td>E-step (20)</td>
<td>$\frac{2^K}{T_{\text{max}}} + \frac{K}{2} \left( \frac{M_{\text{max}} + 4}{T_{\text{max}} - T_{p,\ell}} \right) + \mathcal{O}{K}$</td>
<td></td>
</tr>
<tr>
<td>M-step (24)</td>
<td>$\mathcal{O}{K}$</td>
<td></td>
</tr>
<tr>
<td>$\gamma^{[\ell]}$</td>
<td>$\frac{2^K}{T_{\text{max}}} + \frac{K}{2} \left( \frac{M_{\text{max}} + 4}{T_{\text{max}} - T_{p,\ell}} \right) + \mathcal{O}{K}$</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE I**

Time complexity of the constituent blocks of the SAGE-JDCE scheme.

---

**Fig. 1.** Block diagram of the SAGE-JDCE scheme.
Fig. 2. Illustration of (22): Joint channel estimation and partial interference cancellation.
Fig. 3. Behavior of the objective function $f(\beta_k)$ in (31); $\Gamma_k \triangleq 12 R_k \frac{B^2}{\gamma_0} P_{c,k} |\tilde{h}|^4$ and $\mathcal{R} \triangleq \frac{\Gamma_k - 1}{\gamma_k (\gamma - 1)}$. 

a) $\Gamma_k > 1$

b) $\Gamma_k = 1$

c) $\Gamma_k < 1$
Fig. 4. Performance of a single-rate system in quasi-static flat Rayleigh fading for various system loads $\phi$: $N_c = 8$, $I_p = 4$, $L = 40$, $\hat{L} = 800$. 

\[ \frac{L_1 - L_{n1}}{L_1} \gamma_{b,1} = \cdots = \frac{L_K - L_{n,K}}{L_K} \gamma_{b,K} \text{ [dB]} \]
Fig. 5. Performance of a triple-rate system in quasi-static flat Rayleigh fading for various system loads $\vartheta$: $N_1, \ldots, N_{K/3} = 36$, $N_{K/3+1}, \ldots, N_K/2 = 8$, $N_{K/2+1}, \ldots, N_K = 4$, $L_{\text{D,min}} = 8$, $L_{\text{D,max}} = 80$, $L_{\text{D,ave}} = 1600$. 

$$\frac{L_1 - L_{b,1}}{L_1} \gamma_{b,1} = \cdots = \frac{L_K - L_{b,K}}{L_K} \gamma_{b,K} \text{ [dB]}$$
Fig. 6. Convergence speed of the SAGE-JDCE scheme for various system loads $\vartheta$. 

$$\frac{L_1 - L_{b,1}}{L_1} \tilde{\gamma}_{b,1} = \ldots = \frac{L_K - L_{b,K}}{L_K} \tilde{\gamma}_{b,K} \text{ [dB]}$$