Turbo Multi-User Detection for OFDM/SDMA Systems Relying on Differential Evolution Aided Iterative Channel Estimation

Jiankang Zhang, *Student Member, IEEE*, Sheng Chen, *Fellow, IEEE*,
Xiaomin Mu, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—A differential evolution (DE) algorithm aided iterative channel estimation and turbo multi-user detection (MUD) scheme is proposed for multi-user multi-input multiple-output aided orthogonal frequency-division multiplexing / space-division multiple-access (OFDM/SDMA) systems. The proposed scheme iteratively exchanges the estimated channel information and the detected data between the channel estimator and MUD employing a turbo technique, which gradually improves the accuracy of the channel estimation and the MUD, especially for the first iteration. Quadrature amplitude modulation (QAM) is employed in most wireless standards by virtue of providing a high throughput. However, the optimal maximum likelihood (ML)-MUD becomes extremely complex for employment in QAM-aided multi-user systems. Hence, two different DE aided MUD schemes, the DE aided minimum symbol error rate (MSER)-MUD as well as the discrete DE aided ML-MUD, were developed, and their achievable performance versus complexity was characterized. The simulation results demonstrate that the proposed DE aided channel estimator is capable of approaching the Cramer-Rao lower bound with just two or three iterations. The ultimate bit error rate lower-bound of the single-user additive white Gaussian noise scenario has been approached in the range of $E_b/N_0 \geq 10$ dB and $E_b/N_0 \geq 6$ dB for the DE aided MSER-MUD and the discrete DE aided ML-MUD, respectively.**

***Index Terms*—Orthogonal frequency division multiplexing (OFDM), space-division multiple-access (SDMA), minimum symbol error rate (MSER), multiuser detection (MUD), differential evolution (DE) algorithm.**

I. Introduction

THE best possible exploitation of the finite available spectrum in the light of the increasing demands for wireless services has been at the centre of wireless system optimization. Multiple antennas can be employed both at the transmitter and/or the receiver, which leads to the concept of multiple-input multiple-output (MIMO) systems, in order to attain improvements in both capacity and bit error rate (BER) [1, 2]. As one of the most wide-spread MIMO types, orthogonal frequency-division multiplexing/spatial-division multiple-access (OFDM/SDMA) systems [3, 4] exploit the advantages of both OFDM and SDMA, which increase the systems' capacity by sharing the same bandwidth and time slots by several users roaming in different geographical locations [1].

More specifically, the transmitted signals of U simultaneous single-antenna aided uplink (UL) mobile stations (MSs) are received by an array of antennas at the base station (BS). At the BS, multi-user detection (MUD) techniques are invoked for separating the signals of the different MSs. Over the past decade, a variety of SDMA MUDs have been proposed for separating the users' data on the basis of their unique, user-specific 'spatial signature', i.e. the channel impulse responses (CIRs). Naturally, for a MUD to achieve near-single-user performance, the CIRs have to be accurately estimated [1, 4]. Intensive research efforts have been devoted to developing efficient approaches for channel estimation (CE) in multi-user OFDM/SDMA systems [1, 5, 6]. In order to achieve a near-optimal performance, joint CE and signal detection schemes have recently received significant research attention [2, 7, 8]. As one of the most popular linear SDMA-receiver design strategy, minimum mean square error (MMSE) MUD [3, 4] strikes a tradeoff between the achievable multi-user interference (MUI) rejection and noise amplification. In [6], a constrained least squares (CLS) detector was designed for constant modulus signals, which exploited the constant modulus nature of the subcarrier modulation. This CLS-MUD outperformed the MMSE MUD, despite its lower computational complexity. An excellent work [9] studied various turbo MUD schemes from the viewpoint of Bayesian variational inference. It showed that various turbo MUD schemes can be derived by minimizing the *variational free energy*, including the linear turbo MMSE-MUD of [3, 4]. A robust turbo MUD design was developed in [10], which takes into account the channel estimation error in order to improve the performance of turbo MUDs. An iterative CE and turbo MMSE-MUD scheme was proposed in [11] for binary phase shift keying (BPSK) multiuser OFDM systems, where information is exchanged between the CE and the turbo MUD. The CE of [11] is based on expanding each FD channel coefficient by a Slepian basis expansion (SBE). The resultant

Paper approved by N. Benvenuto, the Editor for Modulation and Detection of the IEEE Communications Society. Manuscript received June 23, 2011; revised November 9, 2011 and January 18, 2012.

J. Zhang is with the School of Information Engineering, Zhengzhou University, 450001, China. He was a visiting student with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. from September 2009 to September 2011 (e-mail: jz09v@ecs.soton.ac.uk).

S. Chen and L. Hanzo are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: {sqc, lh}@ecs.soton.ac.uk).

X. Mu is with the School of Information Engineering, Zhengzhou University, 450001, China (e-mail: iexmmu@zzu.edu.cn).

The financial support of the China Scholarship Council (CSC) and the EPSRC under the auspices of the China-UK Science Bridge as well as of the RC-UK under the India-UK Advanced Technology centre initiative is gratefully acknowledged.

Digital Object Identifier 10.1109/TCOMM.2012.032312.110400

SBE coefficients are then estimated by the linear MMSE estimator, assuming that the maximum Doppler frequency is known. However, no evidence was offered in [11] to show that this iterative CE scheme is capable of approaching the ultimate performance lower-bound, namely, the Cramer Rao lower bound (CRLB) [5].

In general, the MMSE-MUD [3] and the CLS-MUD [6] may be viewed as linear MUDs, which minimize the mean square error (MSE) at the MUD's output. However, it is not the MSE, but the BER or symbol error rate (SER) that really matters in most communication systems. This is because minimizing the MSE doses not necessarily guarantee that the BER or SER of the communication system is also minimized, unless the MUD's output signal is strictly Gaussian [12]. This has motivated the quest for directly minimizing the system's BER or SER. In the last fifteen years, intensive research efforts have been devoted to the formulation and exploitation of the minimum BER (MBER) criterion in diverse applications [13], including MUDs in both code-division multiple-access (CDMA) systems [14] and space-time equalized SDMA systems [12, 15, 16]. Since, quadrature amplitude modulation (QAM) schemes [17] have become popular in wireless standards by virtue of providing a high throughput, direct minimum SER (MSER) detection was conceived for M-QAM systems [18]. However, perfect channel information was assumed to be available by the MBER/MSER-MUD schemes in most of the previous studies. Furthermore, there was no comparison of the achievable performance and complexity between the linear MSER-MUD scheme and the nonlinear maximum likelihood (ML)-MUD scheme.

Against this background, our new contribution is that we propose a differential evolution (DE) aided iterative CE and soft-interference-cancellation (SIC) assisted turbo MUD for multi-user OFDM/SDMA systems[1]. In particular, two types of DE aided MUD scheme are developed – the DE aided linear MSER-MUD and the discrete DE aided nonlinear ML-MUD. Their achievable performance and complexity are compared. The proposed discrete DE aided ML-MUD does not evaluate all the possible legitimate solutions in the same way as the optimal ML-MUD, it rather exploits its intrinsic evolutionary mechanism. More explicitly, it differentially evolves its population vectors by adding the scaled difference between two population vectors to a third vector of the DE algorithm, which makes it completely self-organizing and allows it to approach the optimal ML solution at a fraction of the exhaustive search complexity of the optimal ML-MUD. Furthermore, we propose a discrete DE aided turbo ML-MUD, which calculates the soft *a posteriori* information at a significantly reduced computational complexity. Finally, we propose a DE aided iterative CE and SIC assisted turbo MUD, which exploits the error correction capability of the channel code by exchanging extrinsic information between the MUD and the channel decoder. This turbo MUD can feed back more reliable detected signals to assist the CE. Likewise, more accurate channel estimates will result in a more accurate MUD output. More particularly, our simulations demonstrate that this DE aided CE is capable of approaching the CRLB using just two or three iterations.

[1]This new solution circumvents the limitations of the fundamentally hard-decision-based genetic algorithm (GA) aided OFDM/SDMA systems of [1].

The rest of this paper is organized as follows. The system model of the multi-user OFDM/SDMA UL is described in Section II. Section III is devoted to the optimization problems in the iterative CE and turbo MUD of the OFDM/SDMA systems considered. In Section IV, we will characterize the proposed discrete DE aided ML-MUD and discuss its convergence. The structure of the proposed iterative CE and turbo MUD as well as its computational complexity are illustrated in Section V. Our simulation results and discussions are presented in Section VI, while our conclusions are offered in Section VII.

II. System Model

The multi-user OFDM/SDMA UL system considered is shown in Fig. 1, where each of the U simultaneous users is equipped with a single transmission antenna, while the BS employs an array of Q antennas. All users share the same spectrum and they simultaneously transmit their independent data streams, denoted by $\mathbf{b}^u$, $u = 1, 2, \cdots, U$. The information bits $\mathbf{b}^u$ are first encoded by the independent forward error correction (FEC) encoder of each user, as seen in Fig. 1. The data stream is modulated and then the pilot symbols are embedded into the frequency domain (FD) representation of each OFDM symbol. These FD pilot symbols and their specific allocation are known at the receiver and hence can be exploited for CE. Then the signals are fed to a classic K-point inverse fast Fourier transform (IFFT) based modulator in order to generate the time-domain (TD) modulated signal. After concatenating the cyclic-prefix (CP) of N_{cp} samples, the resultant sequence is transmitted through the MIMO channel.

At the BS, the received signals $\mathbf{y}_q$ of antenna q, $q = 1, 2, \cdots, Q$, are constituted by the superposition of the independently faded TD signals of the U users sharing the same frequency resource, which are also corrupted by the Gaussian noise at the array elements. After discarding the CP and performing fast Fourier transform (FFT)-based demodulation of the received TD signals, we generate Q separate received sequences for the s-th OFDM symbol $\mathbf{Y}_q[s]$, $q = 1, 2, \cdots, Q$, which is given by the superposition of the different users' channel-impaired received signal contribution plus the additive white Gaussian noise (AWGN), formulated as:

$$\mathbf{Y}_q[s] = \sum_{u=1}^{U} \mathbf{X}^u[s]\mathbf{H}_q^u[s] + \mathbf{n}_q[s], \tag{1}$$

where $\mathbf{Y}_q[s] \in \mathbb{C}^{K\times 1}$, $\mathbf{H}_q^u[s] \in \mathbb{C}^{K\times 1}$ and $\mathbf{n}_q[s] \in \mathbb{C}^{K\times 1}$ are column vectors in Equation (1) hosting the subcarrier-related received signals $Y_q[s,k]$, the FD channel transfer factors (FD-CHTFs) $H_q^u[s,k]$ and the AWGNs $n_q[s,k]$, respectively. Each $n_q[s,k]$ has a zero mean and a variance of σ_n^2. Furthermore, $\mathbf{X}^u[s] \in \mathbb{C}^{K\times K}$ is a diagonal matrix with elements given by $X^u[s,k]$, $k = 1, 2, \cdots, K$, which represents the U users' transmitted signals, assuming values from the M-QAM symbol set of

$$\mathcal{S} \triangleq \left\{ s_{m,n} \middle| s_{m,n} = z_m + jz_n, 1 \le m, n \le \sqrt{M} \right\}, \tag{2}$$

where the real-part of the symbols is $z_m = 2m - \sqrt{M} - 1$ and the imaginary-part is $z_n = 2n - \sqrt{M} - 1$.

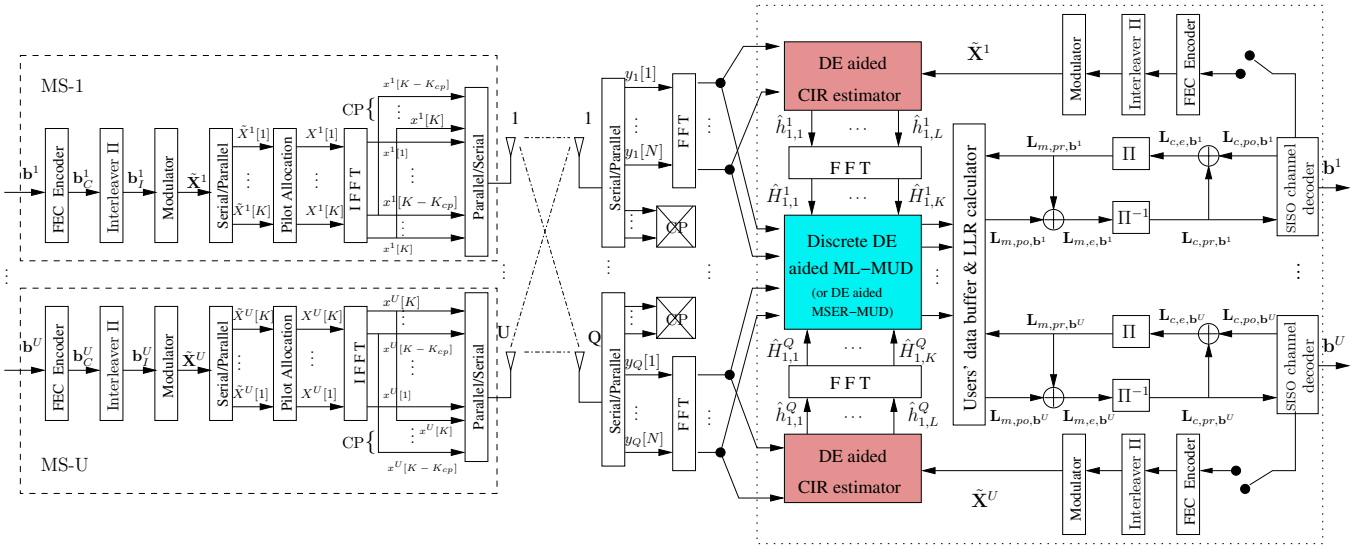


Fig. 1. Uplink system model for Multi-user MIMO OFDM/SDMA. The subscripts m and c of $\mathbf{L}$ are associated with the MUD and channel decoder, respectively, while the subscripts pr, po and e are used for representing the *a priori, a posteriori* and extrinsic information. The loop index of the superscript $^{(loop)}$ is omitted in this figure for avoiding confusion.

III. Optimization Problems in Channel Estimation and MUD

In the context of the joint CE and MUD of the OFDM/SDMA systems, the optimization problems can be based on the log-likelihood function (LLF) conditioned both on the CIR matrix $\mathbf{h}[s] = \left[\mathbf{h}_1[s]\ \mathbf{h}_2[s] \cdots \mathbf{h}_Q[s]\right]$ containing all the CIR coefficients and on the users' transmitted data matrix $\mathbf{X}[s] = \left[\mathbf{X}^1[s]\ \mathbf{X}^2[s] \cdots \mathbf{X}^U[s]\right]^T \in \mathbb{C}^{UK \times K}$, which is given by

$$J(\mathbf{h}[s], \mathbf{X}[s]) = \sum_{q=1}^{Q} \left\| \mathbf{Y}_q[s] - \mathbf{X}^T[s]\mathbf{F}\mathbf{h}_q[s] \right\|^2, \tag{3}$$

where the received data obeys $\mathbf{Y}_q[s] \in \mathbb{C}^{K \times 1}$, the FFT matrix $\mathbf{F} \in \mathbb{C}^{UK \times UL}$, the CIRs are given by $\mathbf{h}_q[s] \in \mathbb{C}^{UL \times 1}$ for $q = 1, 2, \cdots, Q$, and L is the number of CIR taps, while the FD-CHTF vector $\mathbf{F}\mathbf{h}_q[s] = \left[\left(\mathbf{H}_q^1[s]\right)^T \left(\mathbf{H}_q^2[s]\right)^T \cdots \left(\mathbf{H}_q^U[s]\right)^T\right]^T \in \mathbb{C}^{UK \times 1}$.

The joint ML optimization defined in Equation (3) may become computationally prohibitive. The complexity of this optimization process may be reduced to a tractable level by invoking an iterative search loop that first explores the entire set of possible channels $\mathbf{h}[s]$ and then searches the set of all the possible transmitted data $\mathbf{X}[s]$, which is formulated as:

$$(\hat{\mathbf{h}}[s], \hat{\mathbf{X}}[s]) = \arg \min_{\mathbf{X}[s]} \left[\min_{\mathbf{h}[s]} J\big(\mathbf{h}[s], \mathbf{X}[s]\big) \right]. \tag{4}$$

A. Channel Estimation

In multi-user OFDM/SDMA systems, the CIRs can be estimated using pilot symbols known at both the MSs and the BS. For the block-fading scenario, the pilot symbols may be assigned to the first OFDM symbol, which is also referred to as a preamble. When the channel statistics are unknown, the CIR estimate can be determined by minimizing the ML cost function (CF). More specifically, the ML estimation of the channel parameters is the solution of the following CF minimization

$$\begin{aligned} \hat{\mathbf{h}}_q &= \arg \min_{\mathbf{h}_q} J(\mathbf{h}_q) \\ &= \arg \min_{\mathbf{h}_q} \left\| \mathbf{Y}_q[1] - \mathbf{X}^T[1]\mathbf{F}\mathbf{h}_q[1] \right\|^2, \end{aligned} \tag{5}$$

where $\mathbf{Y}_q[1] \in \mathbb{C}^{K \times 1}$ is the first received OFDM symbol, while the transmitted data matrix $\mathbf{X}[1] \in \mathbb{C}^{UK \times K}$ represents the first transmitted OFDM symbols (pilot symbols) of the U simultaneous users.

The standard least squares (LS) channel estimator (CE) provides the solution of (5), which however is computationally very expensive as it requires the inverses of the Q very-large $(U \cdot L) \times (U \cdot L)$ correlation matrices in order to obtain $\hat{\mathbf{h}}_q$ for $1 \le q \le Q$. We propose an alternative DE aided CE to obtain the solution of (5), which does not require the inverse of large-dimensional matrices and is applicable in both training-based and decision-directed channel estimations.

B. The ML-MUD Scheme

As a benefit of the CP, the OFDM/SDMA symbols do not overlap and hence SDMA MUD processing can be applied on a per-carrier basis [1, 3]. Upon invoking vector notations, the set of equations constituted by Equation (1) for $q = 1, 2, \cdots, Q$ at the k-th subcarrier of the s-th OFDM symbol can be rewritten as:

$$\mathbf{Y}[s,k] = \mathbf{H}[s,k]\mathbf{X}[s,k] + \mathbf{n}[s,k], \tag{6}$$

where $\mathbf{Y}[s,k] \in \mathbb{C}^{Q \times 1}$, $\mathbf{H}[s,k] \in \mathbb{C}^{Q \times U}$, $\mathbf{X}[s,k] \in \mathbb{C}^{U \times 1}$ and $\mathbf{n}[s,k] \in \mathbb{C}^{Q \times 1}$ represent the received signals, the FD MIMO channel matrix, the transmitted signals and the AWGN noise, respectively. For notational convenience, the indices $[s,k]$ are omitted during our forthcoming discourse.

Briefly, the task of the MUD is to recover the transmitted signals $\mathbf{X} \in \mathbb{C}^{U \times 1}$ of the U users from the received signals given in Equation (6). Each element of $\mathbf{X}$, say X^u, belongs

to a finite alphabet $\mathcal{S}$ of size $|\mathcal{S}| = M$. Hence there are M^U possible candidate solutions $\hat{\mathbf{X}}$. The optimal ML-MUD exhaustively searches the full space of $\mathcal{S}^U$ to find a solution minimizing $\|\mathbf{Y} - \mathbf{HX}\|^2$, which is equivalent to

$$\hat{\mathbf{X}}_{ML-MUD} = \arg \min_{\mathbf{X} \in \mathcal{S}^U} J(\mathbf{X}) = \arg \min_{\mathbf{X} \in \mathcal{S}^U} \|\mathbf{Y} - \mathbf{HX}\|^2. \tag{7}$$

The above problem may be also viewed as a finite-alphabet-constrained LS problem [19], which is known to be a nondeterministic polynomial-time (NP)-hard problem.

C. The MSER-MUD Scheme

The estimate $\hat{\mathbf{X}}$ of the transmitted signal vector $\mathbf{X}$ of the U simultaneous users can be generated by the MUD upon linearly combining the signals received by the Q different antennas at the BS with the aid of the array weight matrix $\mathbf{W} = \left[\mathbf{W}^1\ \mathbf{W}^2 \cdots \mathbf{W}^U\right] \in \mathbb{C}^{Q \times U}$, yielding [1, 12],

$$\tilde{\mathbf{X}} \quad = \quad \mathbf{W}^H \bar{\mathbf{Y}} + \mathbf{W}^H \mathbf{n}, \tag{8}$$

where the superscript $(\cdot)^H$ denotes the Hermitian transpose and $\bar{\mathbf{Y}} = \mathbf{HX}$ represents the noise-free received data. Since the transmitted signals of different users are independent from each other, the signals outputed by the MUD can be on a per user basis, i.e. we have the u-th user's associated signal,

$$\tilde{X}^u \quad = \quad \bar{X}^u + e^u, \tag{9}$$

where $e^u = \mathbf{W}^{uH}\mathbf{n}$ represents the noise with zero mean and a variance of $\sigma_n^2 \mathbf{W}^{uH}\mathbf{W}^u$, and $\bar{X}^u = \mathbf{W}^{uH}\bar{\mathbf{Y}}$ denotes the noise-free u-th user signal outputed by the MUD.

The probability density function (PDF) of the real-part $\tilde{X}_R^u$ of $\tilde{X}^u$ conditioned both on $X_R^u = z_m$ and on $\mathbf{W}^u$ is a Gaussian mixture, which may be readily formulated as [20]

$$f\left(\tilde{X}_R^u | \bar{X}_{R|X^u=z_m}^u, \mathbf{W}^u\right) = \frac{1}{\sqrt{M}} \sum_{n=1}^{\sqrt{M}} f\left(\tilde{X}_R^u | \bar{X}_{R|X^u=z_m+jz_n}^u, \mathbf{W}^u\right), \tag{10}$$

where $f\left(\tilde{X}_R^u | \bar{X}_{R|X^u=z_m+jz_n}^u, \mathbf{W}^u\right)$ represents the PDF of $\tilde{X}_R^u$ conditioned both on $X^u = z_m + jz_n$ and on $\mathbf{W}^u$, while $\bar{X}_{R|X^u=z_m+jz_n}^u$ is constituted by those $N_{sb} = M^{U-1}$ specific trial vectors, whose u-th element has a value of $(z_m + jz_n)$.

Given that $(z_1 + jz_1)$ was transmitted by user u, the probability of error for the real-part $\tilde{X}_R^u$ is simply the probability that we have $\tilde{X}_R^u < z_1 + 1$, i.e.,

$$\begin{aligned} P_{E,u,R}(\mathbf{W}^u)|_{X_R^u = z_1} &= \int_{z_1+1}^{+\infty} f\left(\tilde{X}_R^u | \bar{X}_{R|X_R^u=z_1}^u, \mathbf{W}^u\right) d\tilde{X}_R^u \\ &= \frac{1}{2N_{sb}} \sum_{i=1}^{N_{sb}} \text{erfc}\left[C_{R,i}(\mathbf{W}^u)\right], \end{aligned} \tag{11}$$

where $\text{erfc}(\cdot)$ denotes the complementary error function [20] and $C_{R,i}(\mathbf{W}^u)$ is formulated as

$$C_{R,i}(\mathbf{W}^u) = \frac{(z_1 + 1) - \bar{X}_{R,i|X_R^u=z_1}^u}{\sigma_n \sqrt{\mathbf{W}^{uH}\mathbf{W}^u}}. \tag{12}$$

Due to the symmetry of the symbol set (2) [20], the error probabilities are identical for $X_R^u = z_1$ and $X_R^u = z_{\sqrt{M}}$, while the error probabilities for $X_R^u = z_m$, $m = 2, 3, \cdots, \sqrt{M} - 1$ are twice that of $X_R^u = z_1$. Since all the legitimate M-QAM signals are equally likely to be transmitted, the total error probability of the real-part $\tilde{X}_R^u$ becomes

$$P_{E,u,R}(\mathbf{W}^u) \quad = \quad \frac{\sqrt{M} - 1}{\sqrt{M} N_{sb}} \sum_{i=1}^{N_{sb}} \text{erfc}\left[C_{R,i}(\mathbf{W}^u)\right]. \tag{13}$$

Assuming that the square-shaped M-QAM constellation is considered, the real-part and imaginary-part will be symmetric to each other. Hence, the error probability of the imaginary-part is identical to that of the real-part. Hence the total SER is given by

$$P_{E,u}(\mathbf{W}^u) \quad = \quad 2P_{E,u,R}(\mathbf{W}^u) - P_{E,u,R}^2(\mathbf{W}^u). \tag{14}$$

The MSER solution $\mathbf{W}_{MSER}^u$ is defined as the weight vector that minimizes the SER of $P_{E,u}(\mathbf{W}^u)$, which is formulated as

$$\hat{\mathbf{W}}_{MSER}^u = \arg \min_{\mathbf{W}^u} P_{E,u}(\mathbf{W}^u). \tag{15}$$

IV. Discrete DE Algorithm Aided ML-MUD

As a relatively new member in the family of evolutionary algorithms (EAs), the DE algorithm [21, 22] has its distinctive feature in that it mutates candidate-solution vectors by adding weighted, random difference-vector to them, which makes it more powerful and efficient in arriving at the globally optimal solution. However, the achievable performance is quite dependent on the setting of the algorithmic parameters, such as the mutation factor and the crossover probability according to both experimental studies and theoretical analysis [22]. In order to automatically update the control parameters and to avoid requiring a user's prior knowledge of the relationship between the algorithmic parameters and the specific characteristics of the optimization problem considered, below we will employ an adaptive DE algorithm [22] for iterative CE and MUD. Owing to space limitations, we only elaborate on the discrete DE algorithm aided ML-MUD in details, noting that similar arguments are also valid for the DE aided CE and MSER-MUD.

A. Discrete DE Algorithm Aided ML-MUD

As mentioned in *Subsection III-B*, the optimal ML-MUD constitutes an NP-hard problem, which has to exhaustively search the full space of the transmitted signals. In this treatise, the discrete DE algorithm is employed for assisting the ML-MUD as a benefit of its versatility in solving these sophisticated optimization problems. The discrete DE assisted ML-MUD may be characterized with the aid of its initialization, mutation, crossover, selection operations and parameter adaptation invoked for exploring the search space in an iterative progression, until the termination criterion is met. An adaptive discrete DE aided ML-MUD is illustrated in Fig. 2, which will be often referred to during our forthcoming discourse.

1) **Initialization.** The discrete DE algorithm commences its search from a population of P_s $(U \times A)$-element solution vectors containing logical 0/1 values. The p_s-th vector

of the population in the first generation of $g = 1$ may be readily expressed as

$$\hat{\mathbf{b}}_{1,p_s} = \Big[\hat{b}^1_{1,p_s,1}, \cdots, \hat{b}^1_{1,p_s,A}, \hat{b}^2_{1,p_s,1}, \cdots, \hat{b}^2_{1,p_s,A}, \cdots, \hat{b}^U_{1,p_s,1}, \cdots, \hat{b}^U_{1,p_s,A}\Big]^T, \quad (16)$$

where A is determined by the modulation scheme used, e.g. $A = 4$ for 16-QAM, while U is the total number of UL users. We also set the mean value of the crossover probability C_r as $\mu_{C_r} = 0.5$ and the location parameter as $\mu_\lambda = 0.5$ for the scaling factor λ of the discrete DE aided ML-MUD.

2) **Mutation.** The mutation operation allows a discrete DE algorithm to maintain the diversity of the population, while insightfully steering the optimization. More specifically, mutation is one of the distinctive features of the discrete DE algorithm, which does not use a predefined probability density function for generating the perturbed solutions. Instead, it relies upon the population itself in perturbing the candidate solutions by adding two appropriately scaled and randomly selected difference-vectors[2] to a *base population vector*. More specifically, the following equation shows how to create a mutant vector by combining three different, randomly chosen vectors and one randomly chosen 'best' vector according to

$$\hat{\mathbf{v}}_{g,i} = \hat{\mathbf{b}}_{g,i} \oplus \Big[\mathbf{z}^b_i \otimes \Big(\hat{\mathbf{b}}^p_{g,best,r_1} \oplus \hat{\mathbf{b}}_{g,i}\Big)\Big] \oplus \Big[\mathbf{z}^b_i \otimes \Big(\hat{\mathbf{b}}_{g,r_2} \oplus \hat{\mathbf{b}}_{g,r_3}\Big)\Big], \quad (17)$$

where $\hat{\mathbf{b}}_{g,i}, \hat{\mathbf{b}}_{g,r_2}$ and $\hat{\mathbf{b}}_{g,r_3}$ are selected from the current population and $\hat{\mathbf{b}}^p_{g,best,r_1}$ is randomly chosen as one of the $(100pP_s)\%$ best vector archive, as seen in Fig. 2. Here p may be interpreted as a greedy factor, which determines the greediness of the mutation strategy. Furthermore, $\oplus$ is the bit-wise exclusive-OR operation, where we have $[0110] \oplus [1111] = [1001]$, and $\otimes$ is the bit-wise exclusive-AND operation, yielding for example $[0110] \otimes [1111] = [0110]$. The bit-scaling factor $\mathbf{z}^b_i$ is a randomly generated $(U \times A)$-length 0/1 vector. More explicitly, first the elements of a $(U \times A)$-length real-valued vector obeying the Gaussian distribution of zero mean and unity variance are generated and then they are compared to the real scaling factor $\lambda_i \in (0, 1]$. If the corresponding element is less than λ_i, then it is mapped to 1, otherwise, it is mapped to 0. The scaling factor λ_i controls the rate at which the population evolves, which is randomly generated according to a Cauchy distribution[3] with location parameter μ_λ and scaling parameter of 0.1, with $\lambda_i = \text{randc}_i(\mu_\lambda, 0.1)$.

3) **Crossover.** The crossover operation generates a trial vector by replacing certain parameters of the target vector with the corresponding parameters of a randomly selected donor vector. As a significant complementarity to the above-mentioned differential mutation, the crossover operation increases the potential diversity of the population vectors. There exist diverse variants of the crossover mechanisms [21, 22]. We opt for employing the uniform crossover algorithm, where each discrete DE parameter, regardless of its location in the trial vector, has the same probability of inheriting its value from a given vector. More particularly, the j-th value of the i-th vector in the population at the g-th generation, namely $\hat{t}^u_{g,i,j}$, is given by

$$\hat{t}^u_{g,i,j} = \begin{cases} \hat{v}^u_{g,i,j}, & \text{rand}_j(0,1) \le C_{r_i} \text{ or } j = j_{rand}, \\ \hat{b}^u_{g,i,j}, & \text{otherwise}, \end{cases} \quad (18)$$

where $C_{r_i} \in [0, 1]$ represents the crossover probability, which is a problem-specific value that represents the specific weight applied to the parameter values that are copied from a previous vector to the mutant, as seen in Fig. 2. The crossover probability C_{r_i} is randomly generated according to a normal distribution of mean μ_{C_r} and standard deviation 0.1, i.e. we have $C_{r_i} = \text{randn}_i(\mu_{C_r}, 0.1)$. Furthermore, $\text{rand}_j(0, 1)$ denotes a random number generator, which returns a uniformly distributed random value from the range $[0, 1)$. The subscript bit-index $j = 1, 2, \cdots, A$ indicates that a new

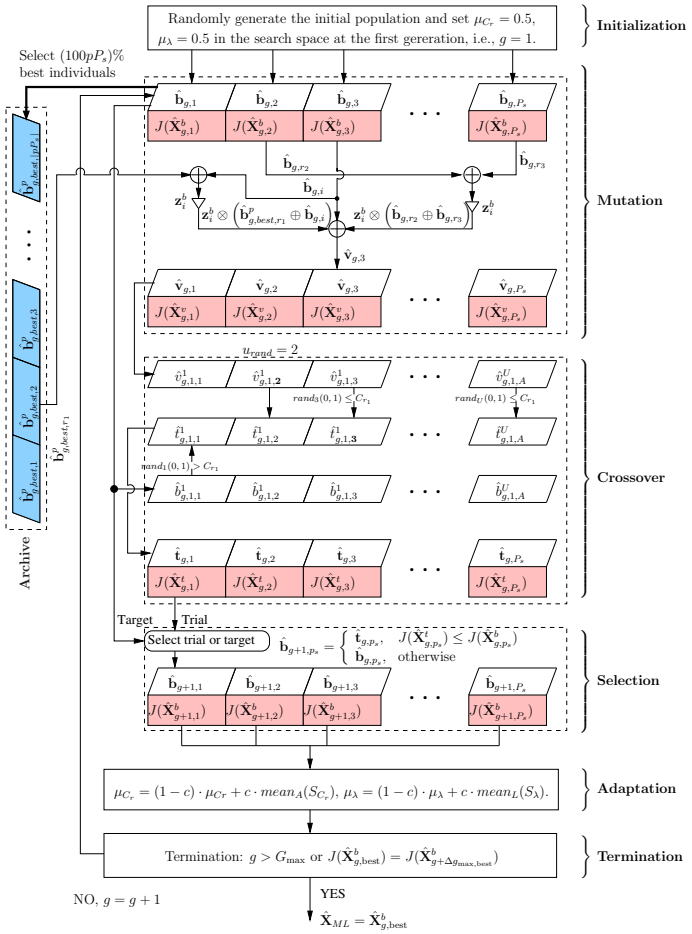


Fig. 2. Flowchart of the discrete DE algorithm.

[2]The difference of the vectors $\hat{\mathbf{b}}_{g,r_2}$ and $\hat{\mathbf{b}}_{g,r_3}$ was defined as $\Delta_{\hat{\mathbf{b}}_{g,r_2},\hat{\mathbf{b}}_{g,r_3}} = \hat{\mathbf{b}}_{g,r_2} \oplus \hat{\mathbf{b}}_{g,r_3}$, as also seen in Fig. 2

[3]Cauchy distribution is more helpful to diversify the mutation factors and thus avoid premature convergence which often occurs in greedy mutation strategies when the mutation factors are highly concentrated around a certain value[22].

random value is generated for each of the j parameters of user $u = 1, 2, \cdots, U$.

4) **Selection.** In order to keep the population size constant for subsequent generations, the selection operator determines whether the target vector $\hat{\mathbf{b}}_{g,p_s}$ or the trial vector $\hat{\mathbf{t}}_{g,p_s}$ survives to the next generation. Before we carry out the comparison and selection operations, the 0/1 vectors of $\hat{\mathbf{b}}_{g,p_s}$ and $\hat{\mathbf{t}}_{g,p_s}$ are mapped to the M-QAM constellation. We denote the M-QAM signals mapped from the target vector $\hat{\mathbf{b}}_{g,p_s}$ and the trial vector $\hat{\mathbf{t}}_{g,p_s}$ by $\hat{\mathbf{X}}^b_{g,p_s}$ and $\hat{\mathbf{X}}^t_{g,p_s}$, respectively. Then, if the trial vector has a lower or equal CF value in comparison to the corresponding target vector, the trial vector will replace the target vector and it is allowed to proceed to the next generation. Otherwise, the target vector will remain in the population for the next generation. More specifically, the selection procedure may be described mathematically as

$$\hat{\mathbf{b}}_{g+1,p_s} = \begin{cases} \hat{\mathbf{t}}_{g,p_s}, & J(\hat{\mathbf{X}}^t_{g,p_s}) \leq J(\hat{\mathbf{X}}^b_{g,p_s}), \\ \hat{\mathbf{b}}_{g,p_s}, & \text{otherwise.} \end{cases} \tag{19}$$

5) **Adaptation.** The adaptation of μ_{C_r} and μ_λ is based on the principle that better control parameter values tend to generate high-fitness individuals that are more likely to survive and thus these values should be propagated to the following generations. The mean of crossover probability μ_{C_r} and the location parameter of scaling factor μ_λ are updated as [22]

$$\mu_{C_r} = (1-c)\cdot\mu_{C_r} + c\cdot \text{mean}_A(S_{C_r}), \tag{20}$$

$$\mu_\lambda = (1-c)\cdot\mu_\lambda + c\cdot \text{mean}_L(S_\lambda), \tag{21}$$

where the adaptive update factor $c \in (0, 1]$ is a positive value, which controls the rate of the parameter adaptation, while $\text{mean}_A(\cdot)$ is the usual arithmetic mean. Instead of an arithmetic mean, the adaptation of μ_λ augments the weight of larger successful mutation factors by using Lehmer mean [22] as $\text{mean}_L(S_\lambda) = \sum_{\lambda_i \in S_\lambda} \lambda_i^2 / \sum_{\lambda_i \in S_\lambda} \lambda_i$. Furthermore, S_{C_r} and S_λ are the set of successful crossover probabilities C_{r_i} and scaling factors λ_i in generation g, respectively.

6) **Termination.** The ultimate stopping criterion would be that of confirming that the optimal solution of the ML-MUD has been found. However, we cannot glean any proof of evidence that the ML-MUD solution has indeed been found. Hence, we opt for halting the optimization procedure, when any of the following stopping criteria are met:

• The pre-defined maximum affordable number of generations $G_{\max}$ has been exhausted.

• $\Delta g_{\max}$ generations have been explored without a trial vector being accepted, which also implies that $\Delta g_{\max}$ generations have passed without any reduction of the CF.

B. Convergence

It is quite a challenge to establish an explicit expression for the convergence speed of the DE algorithm as a function of the parameters of greedy factor p, of the adaptive update factor c as well as of the population size P_s. However, the convergence speed of the DE algorithm may be characterized by the *probability* of *convergence*, which is defined as[4] [23]

$$\lim_{g\to+\infty} \Pr\left(\left\|\hat{\mathbf{X}}^b_{g,p_s} - \mathbf{X}_{ML}\right\| > \varepsilon\right) = 0, \tag{22}$$

where $\Pr(\cdot)$ represents the probability of the given event and ε is an arbitrary positive but small value. Equation (22) suggests that the solutions are located outside $\mathbf{X}_{ML}$'s *ε-neighborhood* with a probability of zero, as the DE proceeds.

There exists a probability $p_g > 0$ at each generation that the individuals in the parental populations generate an offspring $\hat{\mathbf{X}}^t_{g,p_s}$ belong to the *ε-neighborhood* of the $\mathbf{X}_{ML}$. These $p_g, g = 1, 2, \cdots,$ values may be varying at different generations. As a benefit of the elitism, the individuals of the next generation are as good as or better than their counterparts in the current generation, which indicates that probability p_g is monotonically increasing for $g = 1, 2, \cdots$. Hence, this feature will lead to the following proposition,

$$\lim_{g\to+\infty} \Pr\left(\left|\hat{\mathbf{X}}^b_{g,p_s} - \mathbf{X}_{ML}\right| < \varepsilon\right) = 1, \tag{23}$$

which indicates that the the populations will convergence to the *ε-neighborhood* of $\mathbf{X}_{ML}$ with probability one, as DE proceeds.

V. DE-Aided Iterative Channel Estimation and Turbo MUD

A. Structure of the Iterative Channel Estimation and Turbo MUD

The structure of the proposed iterative CE and turbo MUD scheme is illustrated within the box surrounded by the dotted line at the right-hand-side of Fig. 1. The iterative CE and turbo MUD consists of three stages, namely the DE aided CE and the soft-in soft-out (SISO) interference cancellation turbo MUD, followed by U parallel single-user SISO channel decoders. The proposed iterative CE and turbo MUD scheme exploits the error correction capability by repeatedly exchanging extrinsic information between the SISO MUD and the channel decoder in order to mitigate the noise and residual multi-user interference (MUI). The enhanced estimates are fed back as 'pilot symbols' for improving the accuracy of the CE[5]. More specifically, the operation of the DE aided iterative CE and MUD is detailed as follows:

Step-1 Activate the DE aided CE using the pilot symbols.

Step-2 Activate the discrete DE aided ML-MUD (DE aided MSER-MUD) using the channel estimate output by the DE aided CE. Forward the detected users' signals to the users' data buffer and *a posteriori* information calculator, as seen in Fig. 1.

[4]Here we discuss the convergence only refer to the discrete DE assisted ML-MUD, which also makes sense to the DE aided MSER-MUD and DE aided CE.

[5]This is reminiscent of decision-directed CE, where 100% pilot symbols are available, provided that the data estimates are error-free. However, these decision-directed 'pilots' are constituted by the users' transmitted data, and the specially designed low-complexity LS CE of [24] cannot be applied here. By contrast, the DE aided CE of Fig. 1 is capable of exploiting these detected users' data for improving the accuracy of CE. This issue will be further discussed in Subsection VI-C.

Step-3 Reliable estimation of the transmitted data is achieved by exchanging extrinsic information between the MUD and the channel decoder, as seen in Fig. 1. More specifically, the SISO MUD delivers the *a posteriori* information of bit $b^u(i)$ expressed in terms of its log-likelihood ratio (LLR) as [2]

$$\begin{aligned} & L_{m,po,b^u(i)} \\ = \; & \ln \frac{Pr\big[\hat{X}^u \big| b^u(i)=0\big]}{Pr\big[\hat{X}^u \big| b^u(i)=1\big]} + \ln \frac{Pr\left[b^u(i)=0\right]}{Pr\left[b^u(i)=1\right]} \\ = \; & L_{m,e,b^u(i)} + L_{m,pr,b^u(i)}, \end{aligned} \tag{24}$$

where $b^u(i)$ represents the i-th bit mapped to the M-QAM stream of user u, for example $[0\,0\,0\,0]$ is mapped to $(-3+3j)$ in 16-QAM. Again, the indices $[s,k]$ have been omitted. The second term of Equation (24), denoted by $L_{m,pr,b^u(i)}$ represents the *a priori* LLR of the interleaved and encoded bits $b^u(i)$. The first term in Equation (24), denoted by $L_{m,e,b^u(i)}$ represents the extrinsic information delivered by the SISO MUD, based on the received signal $\mathbf{Y}$ and on the *a priori* information about the encoded bits of all users, except for the i-th bit of user u. The extrinsic information is then de-interleaved and fed into the u-th user's channel decoder, which will provide the *a priori* information for the next iteration of the turbo MUD.

Even when the turbo MUD reliably converged to a specific U-user symbol-vector, i.e. the benefits by exchanging extrinsic information between the MUD and the soft channel decoder have been fully exploited during the current loop, the bit stream output by the channel decoder is not delivered to the user before the DE aided CE has converged. Instead, it is re-encoded and re-modulated to generate $(\tilde{\mathbf{X}}^u)^{(loop)}$ of Fig. 1. More specifically, the soft bit-LLRs[6] $\mathbf{L}_{m,po,\mathbf{b}^u}$ are forwarded to the soft FEC decoder, which is capable of outputting the corresponding bit stream. The bit stream is encoded by the FEC encoder, interleaved by the interleaver and then mapped to the corresponding M-QAM signals.

Step-4 The re-encoded and re-modulated data $(\tilde{\mathbf{X}}^u)^{(loop)}$ is then used in the "feedback loop" of Fig. 1 to perform CIR estimation with the assisted of the DE algorithm.

Step-5 The CIR estimate $\hat{\mathbf{h}}^{(loop+1)}$ is then transformed to the FD by the FFT, as shown in Fig. 1. The resultant FD-CHTF $\hat{\mathbf{H}}^{(loop+1)}$ is then directly fed to the users' data buffer and to the *a posteriori* information calculator in order to activate the turbo MUD according to Step-3, so that the process may continue iteration-by-iteration, until convergence is reached.

[6]Note that the bit-LLRs $\mathbf{L}_{m,po,\mathbf{b}^u}$ and their corresponding bits can be assumed to be statistically independent, when a sufficiently long interleaver is employed.

B. *The Soft* a Posteriori *Information of the Turbo MUD*

The soft *a posteriori* information of the optimal turbo ML-MUD associated with bit $b^u(i)$ is given by

$$\begin{aligned} & L^{Opt-ML}_{m,po,b^u(i)} \\ = \; & \ln \frac{Pr\left[\mathbf{Y}, b^u(i)=0\right]}{Pr\left[\mathbf{Y}, b^u(i)=1\right]} \\ = \; & \ln \frac{\sum\limits_{\forall \mathbf{X}\in\mathcal{S}^U: b^u(i)=0} \exp\big(-\frac{\|\mathbf{Y}-\mathbf{H}\mathbf{X}\|^2}{2\sigma_n^2}\big) \prod\limits_{u=1}^{U}\prod\limits_{j=1}^{A} Pr\left[b^u(j)\right]}{\sum\limits_{\forall \mathbf{X}\in\mathcal{S}^U: b^u(i)=1} \exp\big(-\frac{\|\mathbf{Y}-\mathbf{H}\mathbf{X}\|^2}{2\sigma_n^2}\big) \prod\limits_{u=1}^{U}\prod\limits_{j=1}^{A} Pr\left[b^u(j)\right]}, \end{aligned} \tag{25}$$

where the probability $Pr\left[b^u(j)\right]$ of $b^u(j)$ is given by

$$Pr\big[b^u(j)\big] = \frac{1}{2}\Big[1+\mathrm{sgn}\big(\frac{1}{2}-b^u(j)\big)\tanh\Big(\frac{L^{Opt-ML}_{m,pr,b^u(j)}}{2}\Big)\Big]. \tag{26}$$

By observing Equation (25) we can see that the $M^U=|\mathcal{S}|^U$ legitimate candidate solutions of the U users are partitioned into two subsets corresponding to $b^u(i)=0$ or $b^u(i)=1$. This is a challenging task, especially for high-order M-QAM aided multi-user systems. By contrast, the discrete DE aided turbo ML-MUD is capable of reducing the complexity of the soft *a posteriori* information calculation to that of a near-single-user scenario, once the transmitted data $\mathbf{X}$ was detected by the discrete DE aided ML-MUD. More specifically, the soft *a posteriori* information of the discrete DE aided ML-MUD associated with bit $b^u(i)$ may be readily written as

$$\begin{aligned} & L^{DE-ML}_{m,po,b^u(i)} \\ = \; & \ln \frac{\sum\limits_{\forall X^u\in\mathcal{S}: b^u(i)=0} \exp\big(-\frac{\|\mathbf{Y}-\mathbf{H}\tilde{\mathbf{X}}\|^2}{2\sigma_n^2}\big) \prod\limits_{j=1}^{A} Pr\left[b^u(j)\right]}{\sum\limits_{\forall X^u\in\mathcal{S}: b^u(i)=1} \exp\big(-\frac{\|\mathbf{Y}-\mathbf{H}\tilde{\mathbf{X}}\|^2}{2\sigma_n^2}\big) \prod\limits_{j=1}^{A} Pr\left[b^u(j)\right]}, \end{aligned} \tag{27}$$

where $Pr\left[b^u(j)\right]$ is given by Equation (26) and $\tilde{\mathbf{X}} = \big[\hat{X}^1,\cdots,\hat{X}^{u-1},X^u,\hat{X}^{u+1},\cdots,\hat{X}^U\big]^T$, while X^u is taken from the set of M legitimate M-QAM symbols, with $\hat{X}^v, v=1,\cdots,u-1,u+1,\cdots,U$ being acquired by the discrete DE aided ML-MUD at the first turbo iteration. Then, following the first turbo iteration they are given by[7]

$$\hat{X}^v = \max_{X^v\in\mathcal{S}} Pr\big\{X^v\big\} = \max_{X^v\in\mathcal{S}} \prod_{j=1}^{A} Pr\big[b^v(j)\big]. \tag{28}$$

Observe in Equation (27) that the number of legitimate candidate solutions for the U users has been reduced to $M=|\mathcal{S}|$ for each user, since the transmitted signal of user v ($v\neq u$) is given by Equation (28). Then the computational complexity of the soft *a posteriori* information's calculation has been

[7]Note that the computational complexity of the optimal ML-MUD can also be reduced using Equation (27) after the first iteration of the turbo MUD. Also note that Note that at low E_b/N_0, i.e. $E_b/N_0<10$ dB observed from our simulations, suboptimal $\hat{X}^{v'}$ that $\{\hat{X}^{v'}\big|Pr\big\{X^{v'}\big\}\geq \max\limits_{X^v\in\mathcal{S}} Pr\{X^v\}/D\big\}$ may be as well considered also, since then the noise interference is grievous. It is no doubt that computational complexity will increase then, but still very less than the optimal ML-MUD. D is a positive integer, bigger values of D indicate that more suboptimal $\hat{X}^{v'}$ are considered, which will result in more computational complexity.

reduced to $M \cdot U$, compared to the complexity of M^U for the optimal ML-MUD.

Once the MSER-MUD weight vector[8] $\hat{\mathbf{W}}^u$ has been acquired, the soft *a posteriori* information of the turbo MSER-MUD associated with bit $b^u(i)$ can be written as [25]: where $b^u(i) \mapsto \Re\{\hat{X}^u\}$ indicates that $b^u(i)$ is mapped to the real part of $\hat{X}^u$. The means and variances of the real-part and imaginary-part components of $\hat{X}^u$ are given by

$$\begin{cases} \mu_R^u &= \Re\{X^u \hat{\mathbf{W}}^u \mathbf{H}^u\}, \\ \mu_I^u &= \Im\{X^u \hat{\mathbf{W}}^u \mathbf{H}^u\}, \end{cases} \tag{30}$$

and

$$\begin{cases} (\sigma_R^u)^2 \\ = \Re\{\hat{\mathbf{W}}^{uH}\mathbf{H}\}\underline{\mathbf{V}}_R^u \Re\{\mathbf{H}^H \hat{\mathbf{W}}^u\} \\ - \Im\{\hat{\mathbf{W}}^{uH}\mathbf{H}\}\underline{\mathbf{V}}_I^u \Im\{\mathbf{H}^H \hat{\mathbf{W}}^u\} + \sigma_n^2 \hat{\mathbf{W}}^{uH}\hat{\mathbf{W}}^u, \\ (\sigma_I^u)^2 \\ = \Re\{\hat{\mathbf{W}}^{uH}\mathbf{H}\}\underline{\mathbf{V}}_I^u \Re\{\mathbf{H}^H \hat{\mathbf{W}}^u\} \\ - \Im\left\{\hat{\mathbf{W}}^{uH}\mathbf{H}\right\}\underline{\mathbf{V}}_R^u \Im\{\mathbf{H}^H \hat{\mathbf{W}}^u\} + \sigma_n^2 \hat{\mathbf{W}}^{uH}\hat{\mathbf{W}}^u. \end{cases} \tag{31}$$

In Equation (31) we have $\underline{\mathbf{V}}_R^u = \text{diag}\{v_R^1, \cdots, v_R^{u-1}, 0, v_R^{u+1}, \cdots, v_R^U\}$ with $v_R^{u'} = E\{\Re^2\{X^{u'}\}\} - E^2\{\Re\{X^{u'}\}\}$. Similar relationship is valid for $\underline{\mathbf{V}}_I^u$.

C. Computational Complexity

A low-complexity termination criterion is constituted by the number of CF evaluations (*CF-Evals.*), which may be readily used for evaluating the loose computational complexity imposed. For a given population size P_s terminated after G generations, the number of *CF-Evals.* employed by the DE algorithm for finding the weight vectors $\hat{\mathbf{W}}_{MSER}^u$ representing the MSER solution is equal to $(P_s \times G)$. Hence the total number of *CF-Evals.* for the U-user scenario of the proposed DE aided MSER-MUD is equal to $U(P_s \times G)$. Similarly, the total number of *CF-Evals.* for the discrete DE aided ML-MUD is $(P_s \times G)$. By contrast, the number of *CF-Evals.* of the optimum ML-MUD using exhaustive search is equivalent to M^U. It is worth noting that the weight vectors $\hat{\mathbf{W}}_{MSER}$ of the MSER-MUD solutions acquired may be employed for prolonged time-intervals in the scenarios, when the channels are slowly block-fading. However, the discrete DE aided ML-MUD and the optimal ML-MUD must evaluate the candidate solutions for every OFDM symbol. The total computational complexity of the SIC assisted turbo MUD includes the weight calculations and the output extrinsic LLR calculations for both the MMSE-MUD and for the DE aided MSER-MUD. However, the total computational complexity is determined by the calculation of the output extrinsic LLRs for both the DE aided turbo ML-MUD and for the optimal ML-MUD. For a specific example, the computational complexity of the DE aided ML-MUD is mainly determined[9] by the mutation operations of the discrete DE algorithm in Equation (17) and by the calculation of the soft *a posteriori* information in Equation (27). More specifically, in Table I we summarize the total computational complexity comparison of the SIC assisted turbo MUD schemes for a block-fading scenario, where the U UL-users simultaneously transmit their M-QAM data to the BS over K subcarriers. The channel is assumed to be time-invariant over $S = 100$ consecutive OFDM symbols. Furthermore, N_{loop} and N_{Iter} are the number of outer loop iterations and the number of turbo MUD iterations, respectively. In order to have a straightforward comparison between them, the specific values of N_{loop} and N_{Iter} used in our simulations at $E_b/N_0 = 10$ dB are also included in Table I. The total computational complexity is specified in the last column, while the values of the other parameters are specified in Table II.

[8]Investigations not included here owing to lack of space reveal that the MSER weight matrix can be only calculated once during the first iteration and then it is used during the following iterations. This will lead to a significantly reduced complexity at the cost of approximating the soft *a posteriori* information, which slightly degrades the achievable performance.

[9]Naturally, there are some additional calculations.

VI. Simulation Results

In this section, our simulation results are presented in order to illustrate the attainable performance of the proposed iterative CE and turbo MUD in the context of multi-user OFDM/SDMA systems. It was assumed that the UL multi-user OFDM/SDMA system was equipped with Q antennas at the BS, and it supported U MSs simultaneously transmitting their data in the UL to the BS. A summary of the default values for the various parameters used in our simulations is provided in Table II. Unless otherwise specified, these default parameter values were used throughout, where all the users employed 16-QAM.

A. Characterizing the Performance of the Discrete DE Aided ML-MUD

Due to space limitations, we only illustrate the influences of the discrete DE's parameters with regard to the discrete DE aided ML-MUD, but similar trends may be observed for the DE aided MSER-MUD and for the DE aided CE.

Specifically, the effects of the adaptive update factor c and the greedy factor p on the proposed discrete DE aided ML-MUD are investigated in Fig. 3. Observe in Fig. 3 that the proposed scheme requires an increasing number of generations for satisfying the termination criterion defined in Subsection IV-A, when increasing the value of the greedy factor p and using a fixed adaptive update factor c. Moreover, we can see from Fig. 3 that the number of generations required for achieving convergence slightly varied upon increasing the adaptive update factor c for a fixed greedy factor p. However, observe in Fig. 4 that larger greedy factors tend to achieve a better BER performance. The reason may be that larger greedy factors reserve more candidate individuals, which may increase the potential diversity of the population.

In order to evaluate the 'limited performance' of the non-turbo discrete DE aided ML-MUD and the non-turbo DE aided MSER-MUD where the channel decoder and the MUD do not exchange information, we plot their BER performance in Fig. 5 for a fixed population size P_s upon increasing the number of generations G. We can see from Fig. 5 that the BER remains almost the same upon increasing G for $G \geq 26$, when we have $E_b/N_0 < 12$ dB for the discrete DE aided ML-MUD and $E_b/N_0 < 22$ dB for the DE aided MSER-MUD, respectively. However, when we experience $E_b/N_0 > 12$ dB for the discrete DE aided ML-MUD and $E_b/N_0 > 22$ dB

$$L_{m,po,b^u(i)}^{MSER} = \ln \frac{Pr\{\hat{X}^u, b^u(i)=0\}}{Pr\{\hat{X}^u, b^u(i)=1\}} = \begin{cases} \ln \dfrac{\sum\limits_{\forall X^u : b^u(i)=0} \exp\left(-\frac{(\Re\{\hat{X}^u\}-\mu_R^u)^2}{2(\sigma_R^u)^2}\right) Pr\{X^u\}}{\sum\limits_{\forall X^u : b^u(i)=1} \exp\left(-\frac{(\Re\{\hat{X}^u\}-\mu_R^u)^2}{2(\sigma_R^u)^2}\right) Pr\{X^u\}}, & \text{for } b^u(i) \mapsto \Re\left\{\hat{X}^u\right\}, \\ \ln \dfrac{\sum\limits_{\forall X^u : b^u(i)=0} \exp\left(-\frac{(\Im\{\hat{X}^u\}-\mu_I^u)^2}{2(\sigma_I^u)^2}\right) Pr\{X^u\}}{\sum\limits_{\forall X^u : b^u(i)=1} \exp\left(-\frac{(\Im\{\hat{X}^u\}-\mu_I^u)^2}{2(\sigma_I^u)^2}\right) Pr\{X^u\}}, & \text{for } b^u(i) \mapsto \Im\left\{\hat{X}^u\right\}, \end{cases} \tag{29}$$

TABLE I
COMPARISON OF THE COMPUTATIONAL COMPLEXITY

MUD schemes	Operations	Computational complexity	N_{loop}	N_{Iter}	Total	Percentage
MMSE-MUD	Multiplications	$N_{loop}UK\Big\{Q\big(4U^2+4UQ+9Q+4\big)+ N_{Iter}S\Big[\big(M\log_2 M+12Q+16M+12\big)\log_2 M-M\Big]\Big\}$	3	13	1.454×10^9	0.32%
	Additions	$N_{loop}UK\Big[2Q\big(2U^2+2UQ+4Q-U-1\big)+ 2N_{Iter}S\big(6Q+6M-1\big)\log_2 M\Big]$			902843904	0.20%
	Matrix inversion	$N_{loop}UK$ times $(Q\times Q)$ complex-valued matrix inversion			768	–
DE aided MSER-MUD	Multiplications	$N_{loop}UK\Big\{P_sG\big(4Q+3\big)+2\big(2QU+10Q+4U+4M+ 1\big)+N_{Iter}S\left[\big(M\log_2 M+7M+1\big)\log_2 M-M\right]\Big\}$	3	10	762063360	0.17%
	Additions	$N_{loop}UK\Big[P_sG\big(M^{U-1}+4Q\big)+\big(4QU+20Q+ 4U+4M-13\big)+2N_{Iter}S\big(2M-1\big)\log_2 M\Big]$			6.3289×10^{10}	13.91%
	erfc$(\cdot)$	$N_{loop}P_sGUKM^{U-1}$			6.2915×10^{10}	–
Discrete DE aided ML-MUD	Multiplications	$KS(2AU+QU+Q)P_sG+N_{loop}N_{Iter}KS\cdot \Big[U\big(6M\log_2 M+4QMU+3M\big)+1\Big]$	3	6	6.5691×10^9	1.45%
	Additions	$KS(4AU+QU-1)P_sG+N_{loop}N_{Iter}KS\cdot \Big[U\big(2M\log_2 M+4QUM-2QM+5M\big)-1\Big]$			1.0005×10^{10}	2.88%
Optimal ML-MUD	Multiplications	$N_{loop}N_{Iter}KS\Big[MU\big(4\log_2 M-1\big)+ 2M^U\big(2QU+\log_2 M+2\big)+1\Big]$	3	6	4.5310×10^{11}	100%
	Additions	$2N_{loop}N_{Iter}KS\Big[MU\log_2 M+M^U\left(2QU-Q+2\right)\Big]$			3.4735×10^{11}	100%

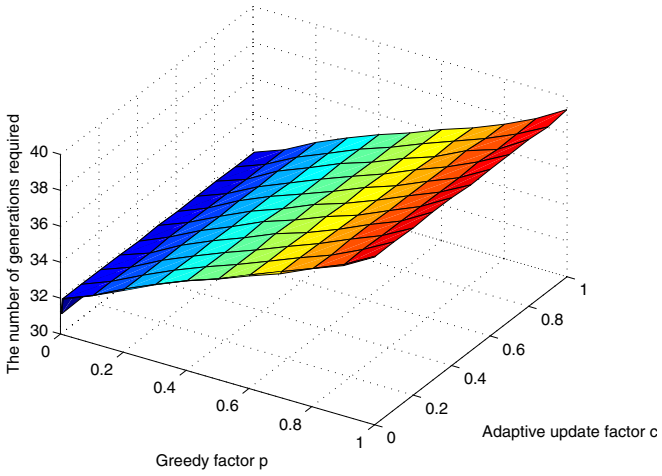


Fig. 3. Average number of evolution generations required for convergence of the proposed discrete DE aided ML-MUD scheme versus the adaptive update factor c and the greedy factor p at $E_b/N_0 = 15$ dB. All other parameters are given in Table II.

TABLE II
BASIC SIMULATION PARAMETERS USED IN OUR SIMULATIONS

FEC encoder and decoder	Type	RSC
	Code rate	1/2
	Constraint length	3
	Polynomial	$(g_0, g_1) = (7, 5)$
Channel	Number of paths L	4
	Delay	$0, 1, \cdots, 3$
	Average path gains	$[0; -5; -10; -15]$ (dB)
	MSs U	4
	Receiver antennas Q	3
	Subcarriers K	64
	Cyclic prefix N_{cp}	16
DE assisted MSER-MUD	Initialization of the population	Randomly generated
	Population size P_s	100
	Greedy factor p	0.7
	Adaptive update factor c	0.8
	Maximum No. of generations $G_{\max}$	200
	$\Delta g_{\max}$	20

for the DE aided MSER-MUD, respectively, the BER is reduced upon increasing the affordable complexity, i.e. the number of generations G. Furthermore, observe from Fig. 5 that the proposed discrete DE aided ML-MUD significantly outperforms the DE aided MSER-MUD and the MMSE-MUD, since it is capable of approaching the performance of the optimal ML-MUD, especially for $E_b/N_0 < 14$ dB. Moreover, the complexity of the discrete DE aided ML-MUD is a small fraction of the complexity imposed by the optimal ML-MUD. As an example, at $E_b/N_0 = 16$ dB, the complexity of discrete DE aided ML-MUD is $\frac{P_s\times G}{M^U} = \frac{100\times200}{16^4} \approx 0.3052$ or $\frac{P_s\times G}{M^U} = \frac{100\times180}{16^4} \approx 0.2747$ of the complexity required

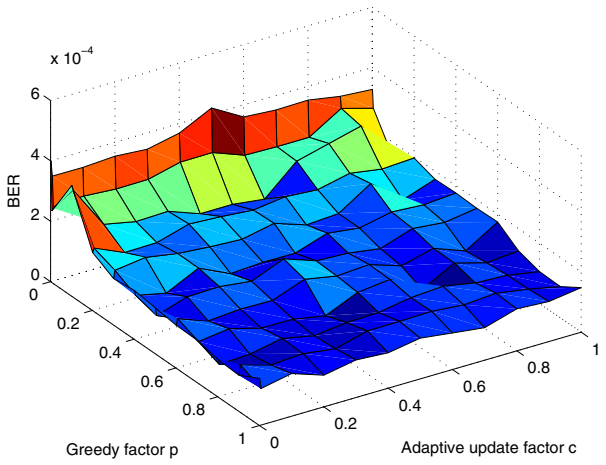


Fig. 4. BER performance of the proposed discrete DE aided ML-MUD scheme versus the adaptive update factor c and the greedy factor p at $E_b/N_0 = 15$. dB The number of evolution generations for different combination of (p, c) are illustrated in Fig. 3. All other parameters are given in Table II.

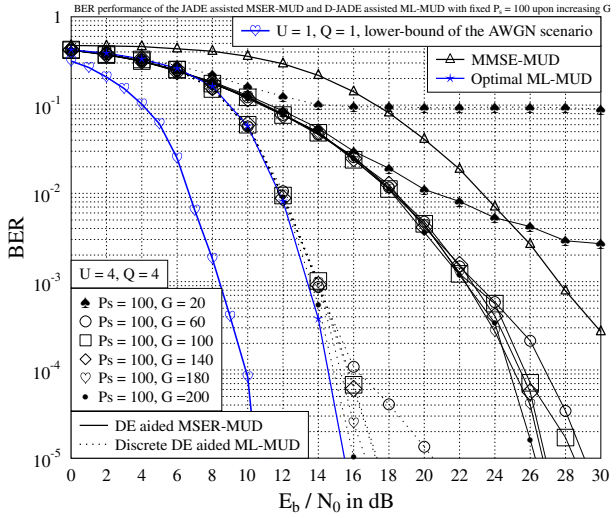


Fig. 5. BER performance of the non-turbo DE aided MSER-MUD and the non-turbo discrete DE aided ML-MUD with fixed population size $P_s = 100$ upon increasing the number of evolution generations. All other parameters are given in Table II.

by the optimal ML-MUD, at the cost of about 0.5 dB or 1 dB performance penalty, respectively. Despite all this, there exits still a wide gap between the attainable 'limited performance' and the ultimate single-user lower-bound over AWGN channels. Below, we will illustrate how to approach the ultimate lower-bound by the turbo MUDs with the aid of EXIT chart analysis.

B. EXIT chart analysis of the turbo MUD

In Fig. 6, we plot the EXIT chart of the MUD schemes considered using recursive systematic convolutional codes

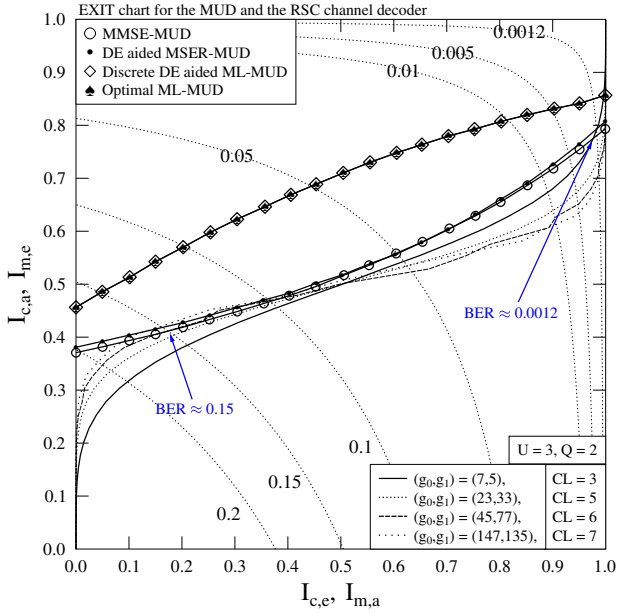


Fig. 6. EXIT chart for the MUDs and the RSC channel decoder, given $E_b/N_0 = 8$ dB. All other parameters are given in Table II. The values on the contour plot are the estimated BER calculated using Equation (32).

(RSCs), given $E_b/N_0 = 8$ dB. The generator polynomials (g_0, g_1) are taken from [26], which have been shown to be the best rate-1/2 RSC polynomials for iterative decoding. The contour plot of the RSC decoder's estimated coded BER is also included as a reference. The estimated BER[10] is formulated as [27, 28]

$$P_b \approx \frac{1}{2}\text{erfc}\left(\frac{\sigma_p}{2\sqrt{2}}\right), \tag{32}$$

where $\sigma_p^2 = \sigma_e^2 + \sigma_a^2$ is the variance of the decoder's soft output $\mathbf{L}_{c,po}$, as seen in Fig. 1. It can be observed from Fig. 6 that there exists a broader EXIT tunnel between the MUD and the RSC decoder EXIT curves for shorter constraint lengths (CLs), when $\mathbf{I}_{c,e}, \mathbf{I}_{m,a} < 0.5$. More specifically, there is an open EXIT tunnel between the MSER-MUD and the RSC code having $(g_0, g_1) = (7, 5)$ and CL $= 3$ at $E_b/N_0 = 8$ dB, which allows the turbo MUD to converge to BER $\approx 1.2 \times 10^{-3}$. However, if the system employs the RSC code of $(g_0, g_1) = (45, 77)$ and CL $= 6$, the turbo MUD only converges to BER $\approx 1.5 \times 10^{-1}$ when $E_b/N_0 = 8$ dB. These results indicate that having shorter CLs for the RSC code will allow the users to transmit their data at a reduced transmit power.

Observe in Fig. 7 that at $E_b/N_0 = 8$ dB, the tunnels between the EXIT curve of the MSER/ML-MUD and the EXIT curve of the RSC decoder become narrower upon increasing the number of supported users, which limits the maximum number of users supported by the multi-user OFDM/SDMA system using a particular MUD scheme at this E_b/N_0. More

[10]Note that the EXIT based BER estimation is reliable when the channels fading are independent, although there is slight difference between the estimated BER and the actual simulated systems' BER [27, 28].

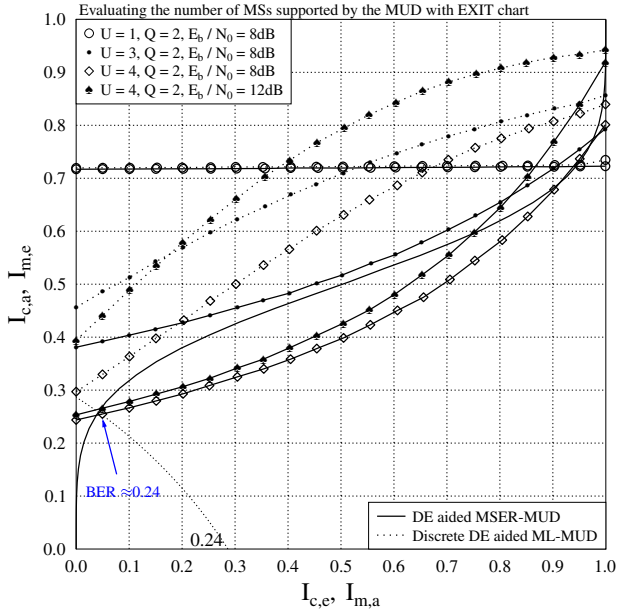


Fig. 7. EXIT chart for the DE aided MSER-MUD and discrete DE aided ML-MUD as well as the channel decoder. The BS is equipped with two antennas. All other parameters are given in Table II.

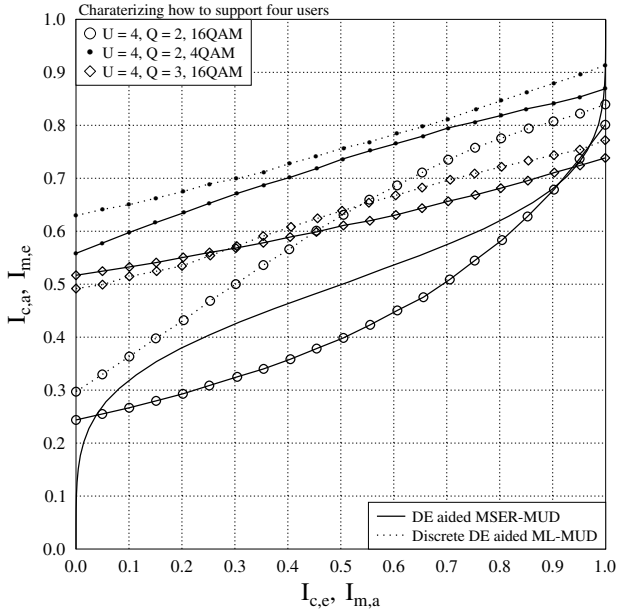


Fig. 8. EXIT chart analysis of the solutions for solving the problem of how to support four users at a specific $E_b/N_0 = 8$ dB. All other parameters are given in Table II.

specifically, the system supports $U = 4$ users simultaneously transmitting their data to the BS and the BS can reliably separate the users' signals, hence approaching the single-user performance, if the BS employs the proposed discrete DE aided ML-MUD at $E_b/N_0 = 8$ dB. By contrast, the DE aided MSER-MUD can only support $U = 3$ users at $E_b/N_0 = 8$ dB, since the BS fails to successfully separate the users' data, yielding $BER \approx 0.24$, as seen in Fig. 7.

Furthermore, even if we increase the transmit power to $E_b/N_0 = 12$ dB, the system fails to support $U = 4$ users, if the BS employs the DE aided MSER-MUD scheme. There are two solutions to assist the system in supporting $U = 4$ users to simultaneously transmit their data, if the BS employs the DE aided MSER-MUD, which are 1) reducing the order of QAM; and 2) increasing the number of antennas at the BS. Observe in Fig. 8 that there is an open tunnel between the EXIT curve of the MSER-MUD and that of the RSC decoder, regardless of which solution is employed by the multi-user OFDM/SDMA system. However, the transmission rate will be reduced and all of the users must alter their system configuration, if the first solution is employed. This may not be acceptable, since future communication systems have to be more user-centric. By contrast, the second solution only influences the BS's configuration, which might be more acceptable, although the achievable performance only has a slight difference.

C. Overall Performance of the Iterative CE and turbo MUD

In Fig. 9 we characterize the overall MSE performance of the DE aided CE at different outer iterations, which is termed as 'loop' in Fig. 9. The Cramer Rao lower bound (CRLB) [5] and the MSE performance of the low-complexity LS CE of [24] are also included as benchmarks. Observe in Fig. 9 that the simplified LS CE relying on optimally designed pilots [24] has the same MSE performance as the DE aided CE at $loop = 0$. However, observe in the figure that this simplified LS CE performs very poorly when using the error-free users' transmitted data for training (marked by ♣). The reason for this observation is that the low-complexity LS CE of [24] requires the "optimal pilots" as discussed in Section III of [24], where the relative phases of the training sequences (pilots) for the different users (transmit antennas) must be carefully designed so that each individual CIR (linking i-th transmit antenna to j-th receive antenna) can be estimated separately. However, the users' transmitted data do not meet this requirement for "optimal pilots". Hence, this simplified LS CE cannot benefit from the iterative CE using the detected users' data. In fact, as clearly seen in Fig. 9, it cannot be used at all even when the error-free users' transmitted data are applied for training[11]. By contrast, the estimation errors of the CIRs are significantly reduced in the ranges of $E_b/N_0 \geq 12$ dB, $E_b/N_0 \geq 10$ dB and $E_b/N_0 \geq 8$ dB, as seen in Fig. 9, when the DE aided CE is combined iteratively with the MMSE-MUD, DE aided MSER-MUD and discrete DE aided ML-MUD, respectively. Specifically, the performance of these iterative schemes approach the CRLB as a benefit of employing the detected data as pilots. In particular, this remarkable improvement is achieved at $loop = 1$ in the ranges of $E_b/N_0 \geq 12$ dB, $E_b/N_0 \geq 10$ dB and $E_b/N_0 \geq 8$ dB, when the DE aided CE is combined iteratively with the MMSE-MUD, DE aided MSER-MUD and discrete DE aided ML-MUD, respectively. This indicates that a single outer iteration

[11]We therefore did not combine the low-complexity LS CE of [24] with different MUD, as it cannot benefit from the iterative CE using the detected users' data at all.

is sufficient for approaching the CRLB. However, there are no useful performance improvements for $E_b/N_0 < 11$ dB, $E_b/N_0 < 8$ dB and $E_b/N_0 < 6$ dB, when the DE aided CE is iteratively combined with the MMSE-MUD, DE aided MSER-MUD and discrete DE aided ML-MUD, respectively. Quite the contrary, the performance are degraded because the detected data are unreliable, as demonstrated in Fig. 10.

In order to portray the overall system's achievable performance, we provide the attainable BER of the turbo MUD schemes considered for $loop = 0, 1, 2, 3$ in Fig. 10. The MSE performance of the DE aided CE at $loop = 0$ is identical to that of the low-complexity LS CE [24] based on the optimally designed pilots, which can be seen in Fig. 9. Hence the BER performance of the different MUD schemes at $loop = 0$ will remain the same, when using the CIR estimated by the low-complexity LS CE of [24] based on the optimally designed pilots[12]. The BER performance of a single user ($U = 1$) recorded for the AWGN channel is also included as the ultimate BER lower-bound benchmark. Observe in Fig. 10 that the proposed iterative DE aided CE combined with the turbo MUD schemes is capable of approaching the ultimate BER lower-bound in the ranges of $E_b/N_0 \geq 10$ dB and $E_b/N_0 \geq 6$ dB for the DE aided MSER-MUD and for the discrete DE aided ML-MUD, respectively. As a special case, the discrete DE aided ML-MUD at $loop = 3$ is capable of approaching the ultimate BER lower-bound of the single-user scenario for the AWGN channel for $E_b/N_0 \geq 6$ dB. The same discrete DE aided ML-MUD at $loop = 0$ becomes capable of approaching the ultimate BER lower-bound for $E_b/N_0 \geq 11$ dB. These significant improvements accrue from the more accurate CE arising from employing the turbo technique by exchanging soft extrinsic information between the soft MUD and the soft channel decoder.

VII. Conclusions

A DE aided iterative CE and turbo MUD scheme has been proposed for multi-user OFDM/SDMA systems, which consists of three stages, namely the DE aided CE, soft MUD and soft channel decoder. The proposed scheme exploits the error correction capability of the turbo MUD and iteratively exchanges information between the CE and MUD. The attainable BER performance has been analyzed with the aid of EXIT charts and the achievable performance has been investigated using Monte Carlo simulations. It has been demonstrated that the proposed DE aided MSER-MUD and the discrete DE aided ML-MUD are capable of approaching the ultimate single-user BER lower-bound for the AWGN channel when we have $E_b/N_0 \geq 10$ dB and $E_b/N_0 \geq 6$ dB, respectively. As a specific example, the number of multiplications and additions required by the discrete DE aided ML-MUD at $E_b/N_0 = 10$ dB are as low as about 1.45% and 2.88% of those of the optimal ML-MUD, respectively.

[12]As demonstrated in Fig. 9, this low-complexity LS CE cannot be applied even when the error-free users' data are used for training. Therefore, this low-complexity LS CE was not combined with various MUD schemes to form different iterative CE and turbo MUD schemes, as we have done here for the DE aided CE.

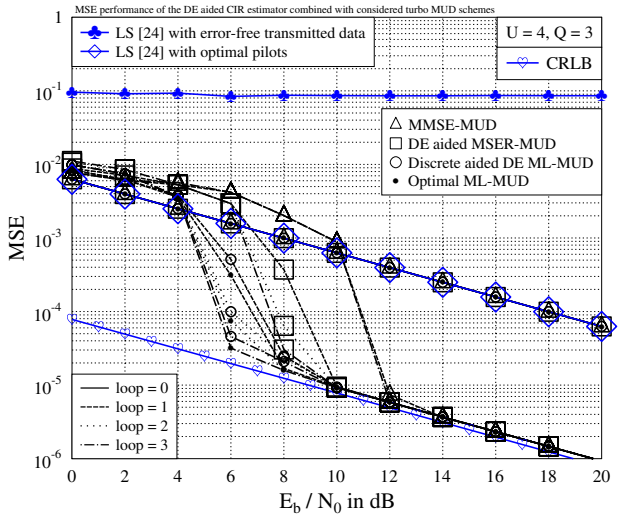


Fig. 9. MSE performance of the DE aided CIR estimator, which uses the detected users' data at the iterations $loop = 1$ to 3, when combined with different turbo MUD schemes. All other parameters are given in Table II. MSE performance of the low-complexity LS CE [24] based on the optimal pilots and the error-free users' transmitted data, respectively, are also shown for comparison.

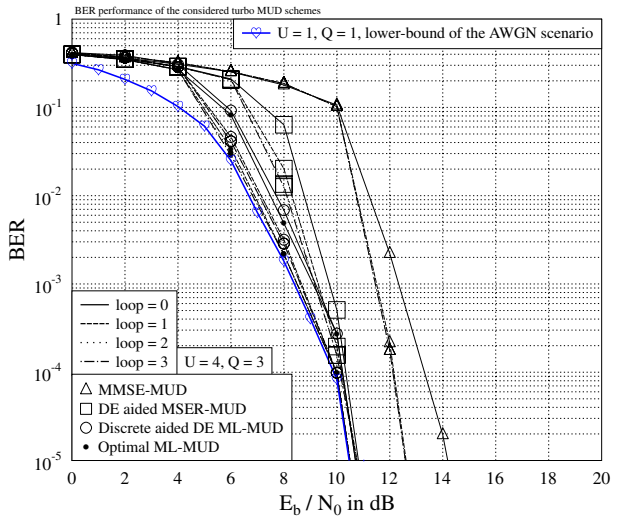


Fig. 10. Comparison of the BER performance for the considered iterative DE aided CE and turbo MUD schemes. All the other parameters are given in Table II.

Acknowledgment

The financial support of the National Natural Science Foundation of China under Grant 61172086 is gratefully acknowledged.

References

[1] M. Jiang and L. Hanzo, "Multiuser MIMO-OFDM for next-generation wireless systems," *Proc. IEEE*, vol. 95, no. 7, pp. 1430–1469, 2007.

[2] L. Hanzo, Y. Akhtman, L. Wang, and M. Jiang, *MIMO-OFDM for LTE, WIFI and WIMAX: Coherent versus Non-Coherent and Cooperative Turbo-Transceivers*. Wiley, 2010.

[3] P. Vandenameele, L. Van Der Perre, M. Engels, B. Gyselinckx, and H. De Man, "A combined OFDM/SDMA approach," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2312–2321, 2000.

[4] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs, and Broadcasting*. IEEE Press, 2003.

[5] J. Zhang, L. Hanzo, and X. Mu, "Joint decision-directed channel and noise-variance estimation for MIMO OFDM/SDMA systems based on expectation-conditional maximization," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2139–2151, 2011.

[6] S. Thoen, L. Deneire, L. Van der Perre, M. Engels, and H. De Man, "Constrained least squares detector for OFDM/SDMA-based wireless networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 129–140, 2003.

[7] M. Jiang, J. Akhtman, and L. Hanzo, "Iterative joint channel estimation and multi-user detection for multiple-antenna aided OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2904–2914, 2007.

[8] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in turbo coded MIMO-OFDM," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1784–1796, 2009.

[9] D. Lin and T. J. Lim, "A variational inference framework for soft-in soft-out detection in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2345–2364, 2009.

[10] S. M.-S. Sadough, M.-A. Khalighi, and P. Duhamel, "Improved iterative MIMO signal detection accounting for channel-estimation errors," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3154–3167, 2009.

[11] P. S. Rossi and R. R. Müller, "Joint twofold-iterative channel estimation and multiuser detection for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4719–4729, 2008.

[12] M. Alias, S. Chen, and L. Hanzo, "Multiple-antenna-aided OFDM employing genetic-algorithm-assisted minimum bit error rate multiuser detection," *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1713–1721, 2005.

[13] B. Mulgrew and S. Chen, "Adaptive minimum-BER decision feedback equalisers for binary signalling," *Signal Process.*, vol. 81, no. 7, pp. 1479–1489, 2001.

[14] R. de Lamare and R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels," *IEEE Commun. Lett.*, vol. 7, no. 2, pp. 73–75, 2003.

[15] S. Chen, L. Hanzo, and A. Livingstone, "MBER space-time decision feedback equalization assisted multiuser detection for multiple antenna aided SDMA systems," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3090–3098, 2006.

[16] D. Gesbert, "Robust linear MIMO receivers: A minimum error-rate approach," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2863–2871, 2005.

[17] L. Hanzo, S. Ng, T. Keller, and W. Webb, *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded Turbo-Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*. John Wiley & Sons, Ltd., 2004.

[18] S. Chen, A. Livingstone, H. Du, and L. Hanzo, "Adaptive minimum symbol error rate beamforming assisted detection for quadrature amplitude modulation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1140–1145, 2008.

[19] E. Larsson, "MIMO detection methods: How they work [lecture notes]," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 91–95, 2009.

[20] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill Companies, Inc., 2001.

[21] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Springer Verlag, 2005.

[22] A. Qin, V. Huang, and P. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evolutionary Computation*, vol. 13, no. 2, pp. 398–417, 2009.

[23] T. Hanne, "On the convergence of multiobjective evolutionary algorithms," *European J. Operational Research*, vol. 117, no. 3, pp. 553–564, 1999.

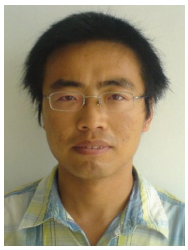
[24] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, 2002.

[25] S. Tan, S. Chen, and L. Hanzo, "Iterative multiuser minimum symbol error rate beamforming aided QAM receiver," *IEEE Signal Process. Lett.*, vol. 15, pp. 301–304, 2008.

[26] M. Ho, S. Pietrobon, and T. Giles, "Improving the constituent codes of turbo encoders," in *Proc. 1998 IEEE Global Telecommun. Conf.*, vol. 6, pp. 3525–3529.

[27] K. Li and X. Wang, "Exit chart analysis of turbo multiuser detection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 300–311, 2005.

[28] L. Xu, S. Chen, and L. Hanzo, "EXIT chart analysis aided turbo MUD designs for the rank-deficient multiple antenna assisted OFDM uplink," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2039–2044, 2008.



Jiankang Zhang (S'08) received the B. S. degree in Mathematics and Applied Mathematics from Beijing University of Posts and Telecommunications, in 2006. He was currently working toward the Ph. D degree in Zhengzhou University, China.

From September 2009 to September 2011, he was a Visiting Student in the School of Electronics and Computer Science, University of Southampton, Southampton, U. K.. His research interests are in the areas of wireless communications and signal processing, including Space Division Multiple Access (SDMA), Orthogonal Frequency-Division Multiplexing(OFDM), channel estimation and multiuser detection.

Sheng Chen (M'90-SM-'97-F'08) obtained his BEng degree from the East China Petroleum Institute, Dongying, China, in January 1982, and his PhD degree from the City University, London, in September 1986, both in control engineering. In 2005, he was awarded the Doctor of Sciences (DSc) degree from the University of Southampton, Southampton, UK. From 1986 to 1999, He held research and academic appointments at the Universities of Sheffield, Edinburgh and Portsmouth, all in UK. Since 1999, he has been with the School of Electronics and Computer Science, the University of Southampton, UK. Professor Chen's recent research interests include adaptive signal processing, wireless communications, modelling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods and optimization. He has published over 450 research papers. In the database of the world's most highly cited researchers in various disciplines, compiled by Institute for Scientific Information (ISI) of the USA, Dr Chen is on the list of the highly cited researchers in the engineering category, see http://www.ISIHighlyCited.com.

Xiaomin Mu received her B. E. degree form Beijing Institute of Technology, Beijing, China in 1982. She is currently a Full Professor with the School of Information Engineering, Zhengzhou University. She has published various paper in the field of signal processing and co-authored two books. Her research interests include signal processing in communications system, wireless communications and cognitive radio.

Lajos Hanzo (M'91-SM-'92-F'04) received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded the honorary doctorate "Doctor Honaris Causa" by the Technical University of Budapest. During his 35-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has co-authored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published in excess of 1000 research entries at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European IST Programme and the Mobile Virtual Centre of Excellence (VCE), UK. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. Since 2008 he has been the Editor-in-Chief of the IEEE Press and since 2009 a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to http://www-mobile.ecs.soton.ac.uk