

Joint Timing and Channel Estimation for Bandlimited Long-Code-Based MC-DS-CDMA: A Low-Complexity Near-Optimal Algorithm and the CRLB

Shuai Wang, Sheng Chen, *Fellow, IEEE*, Aihua Wang, *Member, IEEE*,
Jianping An, *Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

Abstract—Joint Timing and Channel Estimation (JTCE) for bandlimited long-code-aided Multi-Carrier Direct-Sequence Code Division Multiple Access (MC-DS-CDMA) systems is investigated. We establish the optimal multiuser timing and channel estimates for the uplink MC-DS-CDMA receiver by minimising a weighted least squares cost function with respect to K independent parameters, where K is the number of active users. A guided random search procedure known as Repeated Weighted Boosting Search (RWBS) is invoked for numerically solving this challenging multivariate optimisation problem, and thereby for producing near-optimal timing and channel estimates. The Cramér-Rao Lower Bound (CRLB) for the JTCE problem of interest is derived to benchmark the performance of the proposed RWBS based estimator. Quantitatively, for the scenario of $K = 10$ users, $E_b/N_0 \geq 3$ dB where E_b is the energy per bit and N_0 the single-sided noise power spectral density, and for a near-far ratio of 10 dB, the RWBS based estimator using an observation window of 20 symbols is shown to approach the CRLB at a complexity 10 orders of magnitude lower in comparison to its full maximum likelihood search based counterpart. The proposed algorithm does not require the transmission of known pilots, yet it is capable of handling time-variant channel states.

Index Terms—Joint timing and channel estimation (JTCE), multi-carrier direct-sequence code division multiple access (MC-DS-CDMA), repeated weighted boosting search (RWBS), Cramér-Rao lower bound (CRLB).

I. INTRODUCTION

AS a physical-layer enabling technique which combines the merits of both Multi-Carrier (MC) modulation and Direct-Sequence Code Division Multiple Access (DS-CDMA), MC-DS-CDMA has long been favoured by both academia and industry for its robustness and flexibility in

Manuscript received November 11, 2012; no revision. The associate editor coordinating the review of this paper and approving it for publication was H. Arslan.

S. Wang, A. Wang, and J. An are with the School of Information Science and Electronics, Beijing Institute of Technology, P. R. China (e-mail: {swang, wah, an}@bit.edu.cn).

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

This work is supported by the U.K. EPSRC under the auspices of the India-UK Advanced Technology Center Initiative, by the European Union under the Concerto project, and by the European Research Council's Advanced Fellow Grant. It is also supported by the National High Technology Research and Development Program of China under contract 2009AA011502 and by National Scientific Foundation under contract 61271258.

Digital Object Identifier 10.1109/TCOMM.2013.021913.120858

terms of supporting the next-generation wireless network [1]–[4]. In order to combat the Multiple Access Interference (MAI) and to exploit the attainable frequency diversity, the channel parameters, such as the propagation delay and the subcarrier-specific attenuation as well as phase offset, must be estimated for each user at the UpLink (UL) MC-DS-CDMA receiver [5]. Conventional timing and channel estimation schemes are often based on idealised simplifying assumptions, such as short spreading codes and (or) rectangular chip pulses [6]–[12]. In our previous work [13], channel parameter estimation designed for bandlimited long-code based MC-DS-CDMA has been studied. More particularly, in [13] a Parallel Interference Cancellation (PIC) aided *structured* scheme was conceived, which directly estimated the channel parameters by exploiting the *a priori* knowledge of the bandlimited chip waveform. This scheme was shown to outperform two of its *unstructured* counterparts, which operated in a “two-step” fashion by first estimating the Composite Channel Impulse Response (CCIR) vector and then extracting the channel parameters with the assistance of classic correlation techniques. It should be noted that all the three algorithms studied in [13] were designed for quasi-stationary channels and required the transmission of a known pilot sequence. These two properties are, however, disadvantageous when time-variant channel conditions prevail, since the pilot overhead is commensurate with the vehicular speed and hence reduces the effective throughput. More recently, in [14] we studied the timing acquisition conceived for bandlimited long-code-based Single-Carrier (SC) DS-CDMA signals communicating over *doubly selective* channels. The algorithm presented in [14] was *structured* and dispensed with pilot sequences, but it was basically considered from a single-user perspective, with the MAI treated as white or non-white Gaussian noise. As a result, its performance might be significantly degraded in the presence of strong near-far effect.

Against this backdrop, we design a new blind Joint Timing and Channel Estimation (JTCE) scheme for bandlimited long-code-based MC-DS-CDMA systems, which is capable of handling time-variant channel conditions. We commence in Section II by outlining the system model. The organisation of the rest of this paper and its novel contribution are summarised as follows:

- 1) By generalising the results of our previous work [14],

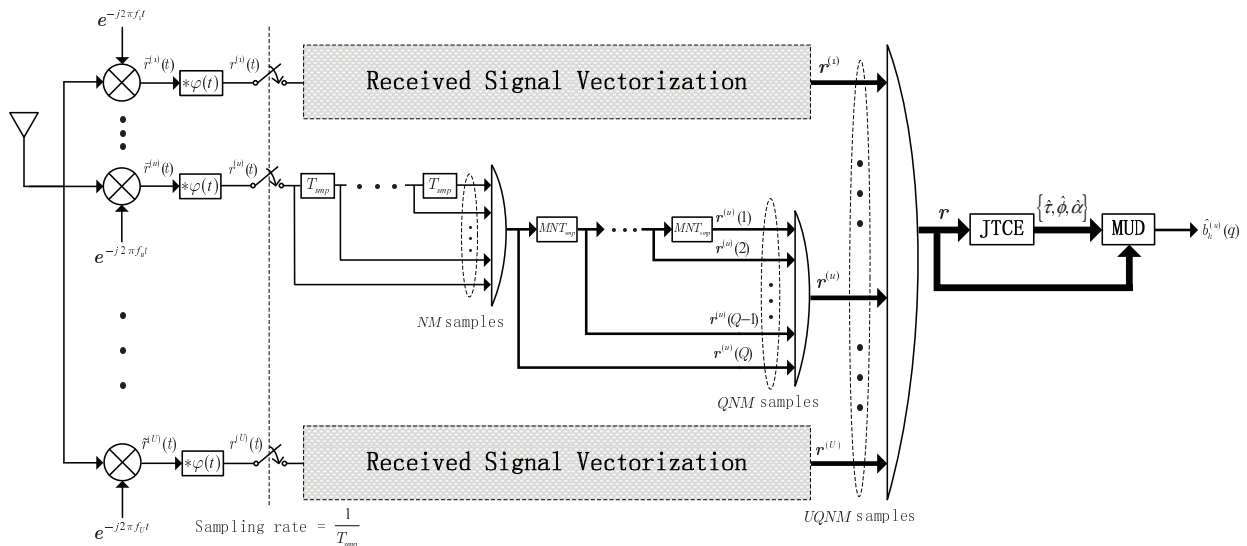


Fig. 1. Architecture of the UL MC-DS-CDMA receiver. Note the other $(U - 1)$ frequency-multiplexed branches have exactly the same signal vectorization structure as that of the u -th one, and their details are hence omitted in the diagram. The operator “*” denotes convolution, while the abbreviations JTCE and MUD stand for Joint Timing and Channel Estimation and Multi-User Detection, respectively.

in Section III we develop a new *structured* formulation of the discrete MC-DS-CDMA signal, which accounts in a collective way for many practical design considerations, such as the long cell-specific spreading codes, the bandlimited chip pulses, the time-varying channel conditions, the absence of pilot sequences, as well as the tradeoff between the attainable frequency diversity and the achievable bit rate.

2) Given this potent formulation, in Section IV, we start by demonstrating that the optimal Maximum Likelihood (ML) multiuser timing and channel estimates may be obtained by minimising a Weighted Least-Squares (WLS) Cost Function (CF) with respect to K independent real-valued parameters, where K is the number of active users. However, solving this optimal JTCE problem by full ML search imposes an excessive complexity and this naturally motivates us to develop near-optimal solutions operating at an affordable computational complexity. Thus, Section IV is mainly devoted to the presentation of a Guided Random Search (GRS) procedure, known as the Repeated Weighted Boosting Search (RWBS) [15], [16], invoked for solving the above-mentioned ML optimisation. It is found that the *filling algorithm* or the generation initialisation scheme adopted by the RWBS procedure plays a key part in avoiding the local minima of the CF. Hence, we propose an ad-hoc filling algorithm termed as *Parallel One-Dimensional Randomisation* (PODR), which turns out to be far more efficient for the problem at hand than other existing generation initialisation alternatives, such as *Pure Randomisation* (PR) [15] or *Mutation* (Mut) [17].

3) In order to benchmark the performance of the proposed RWBS-PODR based estimator, we derive the Cramér-Rao Lower Bound (CRLB) [18] for the blind JTCE problem of interest in Section V, while in Section VI we briefly outline the differential De-Correlation (DC) and Minimum Mean Square Error (MMSE) receivers relying on the proposed low-complexity near-optimal RWBS based blind JTCE algorithm.

4) Our simulation results and complexity analysis are provided in Section VII. Quantitatively, for the case of $K = 10$, $E_b/N_0 \geq 3$ dB, where E_b is the energy emitted per bit and

N_0 the single-sided noise Power Spectral Density (PSD), and a 10 dB near-far ratio, the RWBS-PODR based estimator using an observation window of 20 symbols is shown to reach the CRLB, while imposing a complexity that is 10 orders of magnitude lower than its full ML search based counterpart. We also investigate the BER performance of the differential DC and differential MMSE receivers based on the estimate provided by the blind RWBS-PODR estimator, in comparison with that of an ideal differential MMSE detector which relies on the perfect Channel State Information (CSI).

Notations: Throughout the discussion, $\mathbb{R}^{M \times N}$ and $\mathbb{C}^{M \times N}$ stand for the $M \times N$ real- and complex-valued spaces, respectively. Boldface capital and lower-case letters (like \mathbf{A} and $\boldsymbol{\alpha}$) stand for matrices and vectors, respectively. Given $\mathbf{A} \in \mathbb{R}^{M \times N}$ or $\mathbb{C}^{M \times N}$ and $\boldsymbol{\alpha} \in \mathbb{R}^{M \times 1}$ or $\mathbb{C}^{M \times 1}$, $\mathbf{A}(i, j)$, $\boldsymbol{\alpha}(i)$ and $\mathbf{A}(:, j)$ denote the (i, j) -th entry of \mathbf{A} , the i -th entry of $\boldsymbol{\alpha}$ and the j -th column of \mathbf{A} , respectively. The $(M \times M)$ -element identity matrix and the N -dimensional zero vector are written as \mathbf{I}_M and $\mathbf{0}_N$, respectively. Moreover, the operators $\mathbb{E}(\cdot)$, $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent the expectation, conjugate, transpose and conjugate transpose operations. Finally, $\mathbb{N}^+ \triangleq \{1, 2, 3, \dots\}$, $\mathbf{j} = \sqrt{-1}$, while \odot denotes the element-wise Hadamard product operator.

II. SYSTEM MODEL

We consider the asynchronous MC-DS-CDMA UL shared by K active subscribers, all of whom transmit on the same U subcarriers with the center frequencies being $\{f_1, f_2, \dots, f_U\}$. The architecture of a typical uplink MC-DS-CDMA receiver is illustrated in Fig. 1, and the signals at each point of the receiver are defined in Eqs. (1) to (10b) given below. As in [1] and [6], the U frequency-multiplexed subchannels are assumed to be sufficiently far apart so that they do not overlap with each other. Hence, the equivalent baseband signal received from the u -th subchannel at the UL

receiver can be formulated as:

$$\tilde{r}^{(u)}(t) = \sum_{k=1}^K \sum_{q=0}^{+\infty} \dot{b}_k^{(u)}(q) \int_{-\infty}^{+\infty} \tilde{s}_{k,q}^{(u)}(t - \tau - qT_b) c_k^{(u)}(t; \tau) d\tau + w^{(u)}(t), \quad (1)$$

where $w^{(u)}(t)$ denotes the additive noise on the u -th subchannel, which is modelled as a complex-valued zero-mean white Gaussian process with a uniform two-sided PSD of $N_0/2$; $\dot{b}_k^{(u)}(q) \in \{\exp(j\frac{\pi}{4}), \exp(j\frac{3\pi}{4}), \exp(j\frac{5\pi}{4}), \exp(j\frac{7\pi}{4})\}$ represents the Differential Quadrature Phase Shift Keying (DQPSK) symbol sent by the k -th user on the u -th subcarrier during its q -th signalling interval of length T_b ; $\tilde{s}_{k,q}^{(u)}(t) = \sum_{n=0}^{N-1} \beta_{k,q}^{(u)}(n) \bar{\varphi}(t - nT_c)$ is the signature waveform employed by the k -th user to spread its q -th symbol on the u -th subcarrier, where N is the spreading gain, $T_c = T_b/N$ is the chip interval, and $\beta_{k,q}^{(u)} = [\beta_{k,q}^{(u)}(0) \beta_{k,q}^{(u)}(1) \cdots \beta_{k,q}^{(u)}(N-1)]^T$ is the long (aperiodic) spreading code that changes from symbol to symbol. Here we use $\bar{\varphi}(t)$ to denote the impulse response of the bandlimited Chip Pulse Shaping Filter (CPSF), which is typically a unit-energy Square Root Raised Cosine (SRRC) waveform having a roll-off factor of γ and time-supported over the interval of $[0, DT_c)$ with $D \in \mathbb{N}^+$ and $2D \leq N$ [19]. Finally, $c_k^{(u)}(t; \tau)$ stands for the low-pass, time-varying unit impulse response that characterises the UL fading channel experienced by the k -th user on its u -th subcarrier. The wireless channel under consideration is assumed as frequency-selective, where each subband signal suffers independent non-frequency-selective fading [20]. Hence we have $c_k^{(u)}(t; \tau) = \alpha_k^{(u)}(t) \delta(t - \tau_k)$, where $0 \leq \tau_k < T_b$ is the propagation delay of the k -th user, and the time-variant Complex Channel Gains (CCGs) $\{\alpha_k^{(u)}(t) | k = 1, \dots, K; u = 1, \dots, U\}$ are modelled as independent circularly-symmetric complex Gaussian random processes obeying the well-known Jakes auto-correlation function [21].

For each of the U frequency-multiplexed subchannel at the receiver there is a Chip Pulse Matched-Filter (CPMF) whose impulse response is $\varphi(t) = \bar{\varphi}(DT_c - t)$, followed by an Analog-to-Digital Converter (ADC) sampling at the rate of $\frac{1}{T_{\text{smp}}} = M/T_c$. At the design stage, let us regard $c_k^{(u)}(t; \tau)$ as being constant for about $(T_b + T_m + DT_c)$ seconds with $T_m \triangleq \max\{\tau_k | k = 1, 2, \dots, K\}$, and thereby define the so-called CCIR function that jointly accounts for the effects of the CPSF, the CPMF and the time-variant channel as follows [14], [22]:

$$g_k^{(u)}(t; q) = A_k \int_{-\infty}^{+\infty} \psi(t - \tau - qT_b) c_k^{(u)}(qT_b + \tau_k; \tau) d\tau = \alpha_k^{(u)}[q] \psi(t - \tau_k - qT_b). \quad (2)$$

In Eq. (2) we have $\alpha_k^{(u)}[q] \triangleq A_k \alpha_k^{(u)}(qT_b + \tau_k)$ and $\psi(t) \triangleq \int_{-\infty}^{+\infty} \bar{\varphi}(\tau) \varphi(\tau - t) d\tau$. Obviously, $\psi(t)$ a raised cosine waveform having a time-domain support range of $[0, 2DT_c)$. With the aid of Eq. (2), it can be readily shown that upon feeding $\tilde{r}^{(u)}(t)$ to the CPMF, the output signal $r^{(u)}(t)$ assumes the

following form:

$$\begin{aligned} r^{(u)}(t) &= \sum_{k=1}^K \sum_{q=-\infty}^{+\infty} \dot{b}_k^{(u)}(q) \underbrace{\sum_{n=0}^{N-1} \beta_{k,q}^{(u)}(n) g_k^{(u)}(t - nT_c; q)}_{h_k^{(u)}(t; q)} \\ &+ n^{(u)}(t) = \sum_{k=1}^K \sum_{q=-\infty}^{+\infty} \dot{b}_k^{(u)}(q) \underbrace{h_k^{(u)}(t; q)}_{y_k^{(u)}(t)} + n^{(u)}(t) \\ &= \sum_{k=1}^K y_k^{(u)}(t) + n^{(u)}(t), \end{aligned} \quad (3)$$

where $n^{(u)}(t) \triangleq \int_{-\infty}^{+\infty} w^{(u)}(\tau) \psi_{\text{rx}}(\tau - t) d\tau$ is the filtered (non-white) noise. Recalling $0 \leq T_m < T_b$, one may verify that $g_k^{(u)}(t; q)$ and $h_k^{(u)}(t; q)$ exhibit non-zero values only for $t \in [qT_b + \tau_k, qT_b + \tau_k + 2DT_c] \subseteq [qT_b, (q+1)T_b + 2DT_c]$ and $t \in [qT_b, (q+2)T_b + (2D-1)T_c]$, respectively, which in turn suggests that during the q -th symbol interval of $\mathcal{T}_q \triangleq [qT_b, (q+1)T_b]$, the term $y_k^{(u)}(t)$ in Eq. (3) is determined by at most three consecutive symbols, namely, $\dot{b}_k(q-2)$, $\dot{b}_k(q-1)$, and $\dot{b}_k(q)$. In order to accommodate bandlimited chip pulses, $r^{(u)}(t)$ has to be oversampled, i.e. the oversampling ratio $M \in \mathbb{N}^+$ must be higher than 1. By stacking the MN samples of $r^{(u)}(t)$ over \mathcal{T}_q , we obtain the following MN -dimensional data vector:

$$\begin{aligned} \mathbf{r}^{(u)}(q) &= \sum_{k=1}^K \sum_{p=-2}^0 \dot{b}_k^{(u)}(q+p) \mathbf{h}_k^{(u)}(q+p) + \mathbf{n}^{(u)}(q) \\ &= \sum_{k=1}^K \mathbf{y}_k^{(u)}(q) + \mathbf{n}^{(u)}(q), \end{aligned} \quad (4)$$

where $\mathbf{h}_k^{(u)}(q-2)$, $\mathbf{h}_k^{(u)}(q-1)$, $\mathbf{h}_k^{(u)}(q)$, $\mathbf{y}_k^{(u)}(q)$, and $\mathbf{n}^{(u)}(q)$ comprise the MN samples of $h_k^{(u)}(t; q-2)$, $h_k^{(u)}(t; q-1)$, $h_k^{(u)}(t; q)$, $y_k^{(u)}(t)$ and $n^{(u)}(t)$ within \mathcal{T}_q , respectively. Upon invoking the relationships documented in Equation (8) of [19], we have $\mathbf{h}_k^{(u)}(q-2) = \mathbf{C}_{k,-2}^{(u)}(q) \mathbf{g}_k^{(u)}(q-2)$, $\mathbf{h}_k^{(u)}(q-1) = \mathbf{C}_{k,-1}^{(u)}(q) \mathbf{g}_k^{(u)}(q-1)$, and $\mathbf{h}_k^{(u)}(q) = \mathbf{C}_{k,0}^{(u)}(q) \mathbf{g}_k^{(u)}(q)$, where $\mathbf{g}_k^{(u)}(q) \in \mathbb{C}^{(MN+2MD-1) \times 1}$ is the CCIR vector, which contains the samples $g_k^{(u)}(qT_b + \frac{nT_c}{M}; q)$ for $n = 1, 2, \dots, M(N+2D) - 1$, and the code-dependent matrices $\{\mathbf{C}_{k,p}^{(u)}(q) \in \mathbb{C}^{MN \times (MN+2MD-1)} | p = -2, -1, 0\}$ are composed by appropriately shifted versions of $\beta_{k,q}^{(u)}$ [13], [19]. Hence Eq. (4) can be rewritten in the form:

$$\mathbf{r}^{(u)}(q) = \sum_{k=1}^K \sum_{p=-2}^0 \dot{b}_k^{(u)}(q+p) \mathbf{C}_{k,p}^{(u)}(q) \mathbf{g}_k^{(u)}(q+p) + \mathbf{n}^{(u)}(q). \quad (5)$$

In the next section, Eq. (5) will be used as the starting point for developing the *structured* formulation of bandlimited long-code-based MC-DS-CDMA signals.

III. STRUCTURED FORMULATION OF THE RECEIVED SIGNAL

The proposed JTCE scheme operates in a block-by-block manner. More specifically, based on the data records $\{\mathbf{r}^{(u)}(q)\}$

$u = 1, \dots, U; q = 1, \dots, Q\}$ collected from the U subchannels within the observation window $\mathcal{T} \triangleq [T_b, (Q+1)T_b] = \cup_{q=1}^Q \mathcal{T}_q$, we aim for estimating the propagation delay as well as the amplitude and phase of the subcarrier-specific time-varying CCGs for all users without relying on any *a priori* knowledge about the transmitted symbols $\{\hat{b}_k^{(u)}(q) | k = 1, \dots, K; u = 1, 2, \dots, U; q = -1, \dots, Q\}$. For this purpose, instead of Eq. (5), we need an alternative *structured* formulation which characterises the dependence of the received signal on the channel parameters in a more explicit way.

To begin with, let us assume that the UL channel is solely occupied by user k . Then Eq. (5) reduces to:

$$\mathbf{r}^{(u)}(q) = \sum_{p=-2}^0 \hat{b}_k(q+p) \mathbf{C}_{k,p}^{(u)}(q) \mathbf{g}_k^{(u)}(q+p) + \mathbf{n}^{(u)}(q), \quad (6)$$

Recalling Eq. (2), we now have $\mathbf{g}_k^{(u)}(q+p) = \alpha_k^{(u)} [q+p] \psi(\tau_k)$, with $\psi(\tau_k) = [\psi(\frac{T_c}{M} - \tau_k) \dots \psi(T_b + \frac{(2MD-1)T_c}{M} - \tau_k)]^T \in \mathbb{R}^{(MN+2MD-1) \times 1}$. Hence Eq. (6) can be alternatively written as:

$$\mathbf{r}^{(u)}(q) = \mathbf{H}_{k,q}^{(u)}(\tau_k) \begin{bmatrix} \rho_k^{(u)}(q-2) & \rho_k^{(u)}(q-1) & \rho_k^{(u)}(q) \end{bmatrix}^T + \mathbf{n}^{(u)}(q), \quad (7)$$

where we have $\mathbf{H}_{k,q}^{(u)}(\tau_k) = [\mathbf{C}_{k,-2}^{(u)}(q)\psi(\tau_k) \ \mathbf{C}_{k,-1}^{(u)}(q)\psi(\tau_k) \ \mathbf{C}_{k,0}^{(u)}(q)\psi(\tau_k)] \in \mathbb{R}^{MN \times 3}$ and the Symbol Weighted CCGs (SW-CCGs) are defined as $\{\rho_k^{(u)}(q+p) = \hat{b}_k(q+p) \alpha_k^{(u)} [q+p] p = -2, -1, 0\}$. Stacking the Q consecutive data records collected from the entire observation interval, we arrive at a QMN -element vector $\mathbf{r}^{(u)} = [[\mathbf{r}^{(u)}(1)]^T \ [\mathbf{r}^{(u)}(2)]^T \ \dots \ [\mathbf{r}^{(u)}(Q)]^T]^T$ which, after some algebraic manipulation, can be formulated as:

$$\mathbf{r}^{(u)} = \underbrace{\text{diag} \left\{ \mathbf{H}_{k,1}^{(u)}(\tau_k), \mathbf{H}_{k,2}^{(u)}(\tau_k), \dots, \mathbf{H}_{k,Q}^{(u)}(\tau_k) \right\}}_{\mathbf{A}_k^{(u)}(\tau_k) \in \mathbb{R}^{QMN \times (Q+2)}} \mathbf{F} \boldsymbol{\rho}_k^{(u)} + \mathbf{n}^{(u)} = \mathbf{A}_k^{(u)}(\tau_k) \boldsymbol{\rho}_k^{(u)} + \mathbf{n}^{(u)}. \quad (8)$$

In Eq. (8), we have $\mathbf{F} = [\mathbf{F}_1^T \ \mathbf{F}_2^T \ \dots \ \mathbf{F}_Q^T]^T$ where $\mathbf{F}_q \in \mathbb{R}^{3 \times (Q+2)}$ is defined in the way that $\mathbf{F}_q(i, j) = 1$ if $j - i = q - 1$ and $\mathbf{F}_q(i, j) = 0$ otherwise, while the SW-CCG vector $\boldsymbol{\rho}_k^{(u)}$ takes the form $\boldsymbol{\rho}_k^{(u)} = [\rho_k^{(u)}(-1) \ \rho_k^{(u)}(0) \ \dots \ \rho_k^{(u)}(Q)]^T \in \mathbb{C}^{(Q+2) \times 1}$. Furthermore, the noise term $\mathbf{n}^{(u)} = [[\mathbf{n}^{(u)}(1)]^T \ [\mathbf{n}^{(u)}(2)]^T \ \dots \ [\mathbf{n}^{(u)}(Q)]^T]^T \in \mathbb{C}^{UQMN \times 1}$ is a complex-valued circularly-symmetric Gaussian vector with a zero mean and a covariance matrix of $\bar{\mathbf{\Lambda}} \in \mathbb{R}^{QMN \times QMN}$, which satisfies $\bar{\mathbf{\Lambda}}(i, j) = N_0 \psi(\frac{|i-j|T_c}{M} + 2DT_c)$.

In order to jointly exploit the observations collected from all subcarriers for parameter estimation, let us proceed by stacking the components $\{\mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \dots, \mathbf{r}^{(U)}\}$ into a $UQMN$ -element vector \mathbf{r} . Upon invoking Eq. (8), the *structured* expression of \mathbf{r} may be obtained as follows:

$$\mathbf{r} = \mathbf{A}_k(\tau_k) \boldsymbol{\rho}_k + \mathbf{n}, \quad (9)$$

where $\mathbf{A}_k(\tau_k) = \text{diag}\{\mathbf{A}_k^{(1)}(\tau_k), \mathbf{A}_k^{(2)}(\tau_k), \dots, \mathbf{A}_k^{(U)}(\tau_k)\} \in \mathbb{R}^{UQMN \times U(Q+2)}$ and $\boldsymbol{\rho}_k = [[\boldsymbol{\rho}_k^{(1)}]^T \ [\boldsymbol{\rho}_k^{(2)}]^T \ \dots \ [\boldsymbol{\rho}_k^{(U)}]^T]^T \in \mathbb{C}^{U(Q+2) \times 1}$. The covariance matrix of the zero-mean Gaussian

noise vector $\mathbf{n} = [[\mathbf{n}^{(1)}]^T \ [\mathbf{n}^{(2)}]^T \ \dots \ [\mathbf{n}^{(U)}]^T]^T \in \mathbb{C}^{UQMN \times 1}$ takes the form $\mathbf{\Lambda} = \text{diag}\{\bar{\mathbf{\Lambda}}, \bar{\mathbf{\Lambda}}, \dots, \bar{\mathbf{\Lambda}}\} \in \mathbb{R}^{UQMN \times UQMN}$.

We now extend Eq. (9) to the multiuser system. For a K -user MC-DS-CDMA system, we readily have:

$$\mathbf{r} = \sum_{k=1}^K \mathbf{A}_k(\tau_k) \boldsymbol{\rho}_k + \mathbf{n} \quad (10a)$$

$$= \mathbf{A}(\boldsymbol{\tau}) \boldsymbol{\rho} + \mathbf{n}, \quad (10b)$$

where $\boldsymbol{\tau} = [\tau_1 \ \tau_2 \ \dots \ \tau_K]^T \in \mathbb{R}^{K \times 1}$, $\mathbf{A}(\boldsymbol{\tau}) = [\mathbf{A}_1(\tau_1) \ \mathbf{A}_2(\tau_2) \ \dots \ \mathbf{A}_K(\tau_K)] \in \mathbb{R}^{UQMN \times U(Q+2)K}$, and $\boldsymbol{\rho} = [\boldsymbol{\rho}_1^T \ \boldsymbol{\rho}_2^T \ \dots \ \boldsymbol{\rho}_K^T]^T \in \mathbb{C}^{U(Q+2)K \times 1}$. We note that instead of the CCGs, it is the SW-CCGs together with the user-specific propagation delays that are treated as the unknown parameters in Eq. (10). Incorporating the transmitted symbols into the SW-CCGs is necessary, since it allows us to perform JTCE without relying on the assistance of pilot sequences. Although the amplitudes of the SW-CCGs remain the same as those of the CCGs, their phases are determined by the random DQPSK symbols. Fortunately, this symbol-induced ambiguity will not pose a problem for the Multi-User Detector (MUD) provided differential encoding and decoding are adopted, as advocated in many existing blind channel estimators [6]–[8] as well as in this work. The relationship between the differentially coded symbols $\hat{b}_k^{(u)}(q)$ and the information-bearing Quadrature Phase Shift Keying (QPSK) symbols $b_k^{(u)}(q)$ is defined by the following formula:

$$b_k^{(u)}(q) = \hat{b}_k^{(u)}(q) \cdot (\hat{b}_k^{(u)}(q-1))^*, \quad (11)$$

where we assume furthermore that the four possible values of $b_k^{(u)}(q)$ (modulo- 2π) are mapped onto the quaternary alphabet $\{00, 01, 11, 10\}$ by Gray coding [21].

IV. NEAR-OPTIMAL JTCE BASED ON RWBS

A. Maximum Likelihood Joint Timing and Channel Estimates

Given Eq. (10b), the ML joint estimate of $\boldsymbol{\tau}$ and $\boldsymbol{\rho}$ can be determined from:

$$\{\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\rho}}\}^{\text{ML}} = \arg \min_{\{\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}\}} J^{\text{WLS}}(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}), \quad (12)$$

where $J^{\text{WLS}}(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}) = [\mathbf{r} - \mathbf{A}(\tilde{\boldsymbol{\tau}})\tilde{\boldsymbol{\rho}}]^H \boldsymbol{\Lambda}^{-1} [\mathbf{r} - \mathbf{A}(\tilde{\boldsymbol{\tau}})\tilde{\boldsymbol{\rho}}]$ is the (Weighted Least Squares) WLS cost function. Once $\tilde{\boldsymbol{\rho}}^{\text{ML}}$ is obtained, the corresponding amplitude and phase estimate can easily be extracted. At the first glance, Eq. (12) seems to pose a $(U(Q+2)+1)K$ -dimensional optimisation problem. However, a closer inspection reveals that for any given choice of $\tilde{\boldsymbol{\tau}} \in \mathbb{R}^{K \times 1}$, the best candidate of $\tilde{\boldsymbol{\rho}} \in \mathbb{C}^{U(Q+2)K \times 1}$ that minimises $J^{\text{WLS}}(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}})$ is given by

$$\tilde{\boldsymbol{\rho}}^{\text{WLS}}(\tilde{\boldsymbol{\tau}}) = (\mathbf{A}^H(\tilde{\boldsymbol{\tau}})\boldsymbol{\Lambda}^{-1}\mathbf{A}(\tilde{\boldsymbol{\tau}}))^{-1} \mathbf{A}^H(\tilde{\boldsymbol{\tau}})\boldsymbol{\Lambda}^{-1}\mathbf{r}. \quad (13)$$

That is, $\tilde{\boldsymbol{\rho}}$ is linked to $\tilde{\boldsymbol{\tau}}$ by Eq. (13), and the “independent” parameters are the K propagation delays. Hence Eq. (12) may be equivalently converted into the following reduced-scale K -dimensional optimisation problem

$$\begin{aligned} \hat{\boldsymbol{\tau}}^{\text{ML}} &= \arg \min_{\tilde{\boldsymbol{\tau}}} J^{\text{WLS}}(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}^{\text{WLS}}(\tilde{\boldsymbol{\tau}})) \\ &= \arg \min_{\tilde{\boldsymbol{\tau}}} [\mathbf{r} - \mathbf{A}(\tilde{\boldsymbol{\tau}})\tilde{\boldsymbol{\rho}}^{\text{WLS}}(\tilde{\boldsymbol{\tau}})]^H \boldsymbol{\Lambda}^{-1} [\mathbf{r} - \mathbf{A}(\tilde{\boldsymbol{\tau}})\tilde{\boldsymbol{\rho}}^{\text{WLS}}(\tilde{\boldsymbol{\tau}})]. \end{aligned} \quad (14)$$

Once $\hat{\tau}^{\text{ML}}$ is found, the corresponding ML estimate of ρ follows immediately as $\hat{\rho}^{\text{ML}} = \hat{\rho}^{\text{WLS}}(\hat{\tau}^{\text{ML}})$.

The optimal ML estimate $(\hat{\tau}^{\text{ML}}, \hat{\rho}^{\text{ML}})$ requires solving the optimisation problem (14), which is defined over the continuous K -dimensional space $[0, T_b)^K$ since $\tau_k \in [0, T_b)$ for $1 \leq k \leq K$. Finding the ML estimate $\hat{\tau}^{\text{ML}}$ by a brute-force exhaustive search over the space $[0, T_b)^K$ is computationally prohibitive. Even using the conventional grid search based scheme to find a near-ML solution is known to be NP-hard, or to impose an exponentially increasing complexity. Given that the symbol period is $T_b = NT_c$ with T_c being the chip period, we quantise the chip interval into the R equally-spaced bins. Then $[0, T_b)$ is quantised into the RN equally-spaced bins, and the number of candidates that will be visited by the K -dimensional exhaustive grid search is $(RN)^K$. Therefore, the complexity of this exhaustive grid based ML search is dictated by its number of CF evaluations, which is given by

$$N_{\text{CF-Eva}}^{\text{ML}} = (RN)^K. \quad (15)$$

The positive integer R denotes the resolution, which therefore determines how close the solution found by the exhaustive grid based ML search is to the true ML estimate. Observe the exponential complexity increase suggested by Eq. (15).

B. Motivation

Since direct solution of Eq. (14) via an exhaustive search seems unrealistic because of its formidable complexity, low-complexity suboptimal alternatives tend to have more appeal for practical applications. In our previous work [13], a PIC assisted suboptimal multiuser channel estimation scheme has been proposed, which may be readily invoked for the problem at hand. To be more specific, at the initial stage of the PIC we may just treat $\sum_{k'=1, k' \neq k}^K \mathbf{A}_{k'}(\tau_{k'}) \rho_{k'}$ in Eq. (10a) as part of the background noise and estimate $\{\tau_k, \rho_k\}$ by invoking the Structured Least Squares Search (SSLS) algorithm as suggested in [13]. From the second stage onwards, we will first regenerate the other $(K-1)$ users' contribution to the received signal along with their respective channel estimates obtained at the last stage, then subtract these restored components from \mathbf{r} and use the resultant decontaminated signal for the channel estimation of a specific user. Omitting the detailed analysis (c.f. [13]), we note that the above-mentioned PIC based approach entails a fairly modest complexity on the order of $\mathcal{O}(UQ^2MN^2KV)$, where V is the number of stages involved in the PIC. However, as it will be demonstrated by our simulation results, the performance of this PIC based approach remains poor even for a small number of users and high Signal-to-Noise-Ratios (SNRs). To interpret this limitation further, recall that in [13] we assumed a stationary channel state and pilot-assisted transmissions, hence there were only $(U+1)K$ parameters¹ to be estimated. By contrast, as blind estimation of time-varying channels is considered in this contribution, we now have as many as $(U(Q+2)+1)K$ unknowns in Eq. (9). Given the same amount of received data $\mathbf{r} \in \mathbb{C}^{UQMN \times 1}$, an increase in the number of unknown parameters renders the estimation process to become more

vulnerable to impairments, such as noise and interference. As a result, the channel estimates produced by the initial stage of the PIC are usually insufficiently accurate. The regeneration and cancellation based on such inaccurate (or even wrong) estimates tends to amplify, rather than mitigate the MAI [23], which corrupts the following stages.

Clearly, the above-mentioned suboptimal PIC based approach is unsatisfactory. It is therefore highly desirable to find the optimal or near-optimal solution of the JTCE problem at an affordable complexity. For this objective, we turn to a population based Guided Random Search (GRS) procedure, known as the Repeated Weighted Boosting Search (RWBS) [15], [16], for solving the optimisation problem in Eq. (14). The RWBS algorithm belongs to the class of so-called Evolutionary Algorithms (EAs), which are further exemplified by the family of Genetic Algorithms (GAs) [24], [25], Ant Colony Optimisation (ACO) [26], [27], Particle Swarm Optimisation (PSO) [28], [29] and Differential Evolution Algorithm (DEA) [30], [31], just to name a few. The research on EA has long been among the hottest topics in the computational intelligence community. Vast amounts of empirical results exist in the literature, which have demonstrated that appropriately configured EAs are capable of finding the global optimum or a near-optimum solutions even for the most challenging optimisation problems at an affordable computational complexity. Moreover, the theoretical analysis of EAs has made significant progresses in the past few years (see [32] and the references therein). More specifically, we now know that many NP-hard problems can be turned into the so-called EA-easy class [32] and then solved by a well-configured EA algorithm at a complexity which increases at most polynomially instead of exponentially with the problem size. Besides, both empirical experience and theoretical analysis have suggested that whether a particular NP-hard problem can be turned into an EA-easy one by a given EA critically depends on whether the algorithmic parameter settings of this EA are appropriately matched to the given problem.

The RWBS algorithm [15], [16] was proposed as a population based GRS global optimisation algorithm. A number of successful applications have since been reported, covering the diverse fields of machine learning, chaotic system stabilisation, image and signal processing as well as wireless communication designs [17], [33]–[46]. The RWBS is essentially a multi-start search technique [47], constituted by a stochastic search component and a local search component, like most EAs. The local optimisation mechanism within the RWBS is based on an iterative, adaptive, weighted convex combination of the population of individuals. In conjunction with a reflection operator, the convex combination generates new solutions in a manner similar to the simplex method, but the adaptive weight update process, which is based on a modified boosting technique [48], significantly boosts the power of this local optimiser. The stochastic or global search is constituted by repeating the local search a number of times throughout the generations, where each generation starts from a randomly initialised population of individuals combined with an elitism mechanism. Interestingly, the RWBS shares a number of similarities with the GA that employs a convex crossover operator, single offspring selection, and elitism [15],

¹More specifically, each of the K active users has 1 propagation delay and U subcarrier-specific CCGs that had to be estimated.

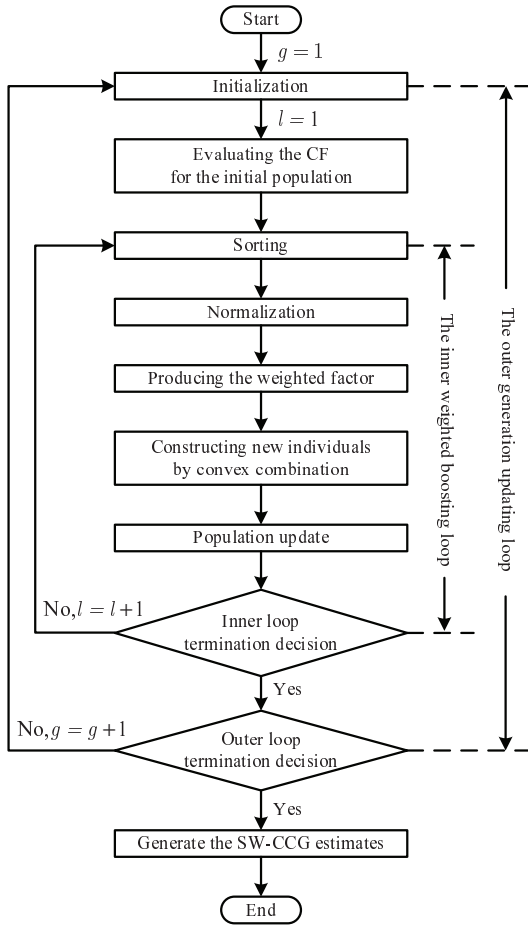


Fig. 2. The flowchart of the proposed RWBS-based JTCE scheme.

[16]. Both algorithms are population-based techniques which combine a local search based on current members of the population, i.e. convex combination in RWBS and crossover in a GA, as well as a stochastic search, i.e. randomly initialised population members in RWBS and mutation in a GA, for the sake of preventing the algorithm from converging towards local optima. The appeal of RWBS lies in its two prominent merits: 1) it has an extremely simple structure that may be implemented at modest complexity, and 2) it has only a small number of tunable parameters that have to be tailored according to the specific problem of interest. In the following discourse, we will demonstrate that the RWBS-based JTCE estimator offers ML-estimation quality, while only imposing a fraction of the full ML-based estimator's complexity.

C. Algorithm

Fig. 2 depicts the flowchart of our RWBS-based JTCE estimator. In particular, the RWBS procedure adopted to solve the optimisation given in Eq. (14) consists of an inner Weighted Boosting Search (WBS) loop nested in an outer generation updating loop, which searches for the optimal or near-optimal delay estimates and subsequently extracts the ML SW-CCG estimates. Referring to Fig. 2, we now detail every component of the RWBS-based estimator.

Outer Loop. Initialisation: Initialisation is performed at every generation of the RWBS process by creating a population of P_S potential solutions or *individuals*, where P_S is

referred to as the population size. The individuals represent legitimate candidates for the multi-user propagation delay vector τ , and they are denoted by $\{\tilde{\tau}^{(g,i)} | g = 1, 2, \dots; i = 1, 2, \dots, P_S\}$, with g being the generation index and i being the index of the individual in the population. For the first generation, $\{\tilde{\tau}^{(1,i)} | i = 1, 2, \dots, P_S\}$ are uniformly and randomly initialised in the K -dimensional parameter space $[0, T_b]^K$. For the g -th generation ($g > 1$), an elitism strategy is adopted by setting $\tilde{\tau}^{(g,1)} = \hat{\tau}^{(g-1)}$, where $\hat{\tau}^{(g-1)}$ denotes the best solution found in the $(g-1)$ -th generation, while the remaining $(P_S - 1)$ individuals in the population are generated using a certain *filling algorithm*². Following the initialisation of the population, the WBS loop index is set to $l = 1$.

CF Evaluation: The quality of an individual is reflected by its CF value. The lower the CF value of an individual, the higher its “fitness”, and vice versa. The CF values associated with $\{\tilde{\tau}^{(g,i)} | i = 1, 2, \dots, P_S\}$ are evaluated:

$$J_i = J^{\text{WLS}}(\tilde{\tau}^{(g,i)}, \hat{\rho}^{\text{WLS}}(\tilde{\tau}^{(g,i)})), \quad 1 \leq i \leq P_S.$$

Inner Loop. Sorting: The best and worst individuals in the g -th generation population are determined:

$$\begin{cases} \tilde{\tau}^{(g,\text{best})} = \tilde{\tau}^{(g,i_\Delta)}, & \text{where } i_\Delta = \arg \min_{1 \leq i \leq P_S} \{J_i\}, \\ \tilde{\tau}^{(g,\text{worst})} = \tilde{\tau}^{(g,i_\nabla)}, & \text{where } i_\nabla = \arg \max_{1 \leq i \leq P_S} \{J_i\}. \end{cases}$$

For convenience, let us define:

$$J_{\text{best}} \triangleq J^{\text{WLS}}(\tilde{\tau}^{(g,\text{best})}, \hat{\rho}^{\text{WLS}}(\tilde{\tau}^{(g,\text{best})})), \quad (16a)$$

$$J_{\text{worst}} \triangleq J^{\text{WLS}}(\tilde{\tau}^{(g,\text{worst})}, \hat{\rho}^{\text{WLS}}(\tilde{\tau}^{(g,\text{worst})})). \quad (16b)$$

Normalisation: The normalised CF values corresponding to $\{\tilde{\tau}^{(g,i)} | i = 1, 2, \dots, P_S\}$ are simply given as

$$\bar{J}_i = \frac{J_i}{\sum_{i'=1}^{P_S} J_{i'}}, \quad 1 \leq i \leq P_S.$$

Unlike J_i , The normalised CF value \bar{J}_i provides the *relative merit* associated with $\tilde{\tau}^{(g,i)}$, which serves as a reasonable metric for guiding the search toward the optimal solution.

Combination Weight Generation: The algorithm generates a new individual as the convex combination of all the individuals in the population. Intuitively, the individuals with low CF values should have high weights so that the new individual generated by the convex combination may have an even lower CF value. Therefore, the weight-generation procedure should augment the influence of meritorious individuals (having low CF values) with the aid of larger weights, while suppressing the deficient individuals (having high CF values) using smaller weights. The weight $\omega_i(l)$ assigned to the i -th individual $\tilde{\tau}^{(g,i)}$ (where $1 \leq i \leq P_S$) in the l -th iteration of the inner WBS loop is self-adapted in the following way [15]. First, a weighting factor β_l is computed according to

$$\beta_l = \frac{\eta_l}{1 - \eta_l} \quad \text{with} \quad \eta_l = \sum_{i=1}^{P_S} \omega_i(l-1) \bar{J}_i. \quad (17)$$

²It is worth mentioning that the efficiency of the RWBS procedure conceived for the problem at hand substantially depends on the specific choice of filling algorithm. For this reason, we will elaborate on this issue later in a dedicated paragraph.

Then the weights $\omega_i(l)$ are updated from their previous values $\omega_i(l-1)$ according to the formulae:

$$\begin{cases} \tilde{\omega}_i(l) = \omega_i(l-1)\beta_l^{\bar{J}_i}, & \text{if } 0 \leq \beta_l \leq 1, \\ \tilde{\omega}_i(l) = \omega_i(l-1)\beta_l^{(1-\bar{J}_i)}, & \text{if } \beta_l > 1, \end{cases} \quad (18)$$

$$\omega_i(l) = \frac{\tilde{\omega}_i(l)}{\sum_{i'=1}^{P_S} \tilde{\omega}_{i'}(l)}. \quad (19)$$

The reason for the two different scenarios of weight updating in Eq. (18) according to $\beta_l \lesseqgtr 1$ is that the exponential function β_l^x is monotonically increasing when $\beta_l > 1$ and it is monotonically decreasing for $\beta_l < 1$, where $0 < x < 1$. Since we would like to increase the weights of individuals having lower CF values and to decrease the weights of individuals having higher CF values, the weights are therefore updated using the two rules of Eq. (18). The weights computed by the formulae (18)-(19) reflect the fitness ratios of their corresponding individuals, as is desired. Moreover, for $1 \leq i \leq P_S$, we have $0 \leq \omega_i(l) \leq 1$ and $\sum_{i=1}^{P_S} \omega_i(l) = 1$, which meet the requirements for a convex combination. Naturally, at the initial stage we have no reason to favour any of the individuals. Therefore, before starting the WBS inner loop, all the initial weights can be set to $\omega_i(0) = 1/P_S$, $1 \leq i \leq P_S$.

Constructing New Individuals: A new individual is constructed as the *convex combination* of the existing P_S individuals based on their associated weights:

$$\tilde{\tau}^{(g, P_S+1)} = \sum_{i=1}^{P_S} \omega_i(l) \tilde{\tau}^{(g, i)}.$$

A ‘‘mirror image’’ of $\tilde{\tau}^{(g, P_S+1)}$ with respect to $\tilde{\tau}^{(g, \text{best})}$ along the direction of $\tilde{\tau}^{(g, \text{best})} - \tilde{\tau}^{(g, P_S+1)}$ is also generated

$$\tilde{\tau}^{(g, P_S+2)} = \tilde{\tau}^{(g, \text{best})} + \left(\tilde{\tau}^{(g, \text{best})} - \tilde{\tau}^{(g, P_S+1)} \right).$$

Since the influence of the meritorious individuals having low CF values have been augmented by higher weights, typically $\tilde{\tau}^{(g, P_S+1)}$ or $\tilde{\tau}^{(g, P_S+2)}$ will exhibit a CF value even lower than J_{best} .

Population Update: The CF values for the two new individuals, $\tilde{\tau}^{(g, P_S+1)}$ and $\tilde{\tau}^{(g, P_S+2)}$, are computed, and they are denoted by J_{P_S+1} and J_{P_S+2} , respectively. Define

$$i_* \triangleq \arg \min_{i=P_S+1, P_S+2} \{J_i\}.$$

Then the current worst individual $\{J_{\text{worst}}, \tilde{\tau}^{(g, \text{worst})}\}$ is replaced by the newly generated individual $\{J_{i_*}, \tilde{\tau}^{(g, i_*)}\}$. It can readily be seen that at each iteration of the WBS inner loop, the population is guaranteed to improve, because the average CF value of the population is reduced.

Inner Loop Termination. The convergence of the WBS inner loop is detected by the following *Test*:

If $\|\tilde{\tau}^{(g, P_S+1)} - \tilde{\tau}^{(g, P_S+2)}\| \leq E_B$, where E_B specifies the accuracy of the WBS, or alternatively, if $l \geq N_B$, where N_B denotes the maximum allowed WBS inner loop iterations, then the WBS inner loop is terminated with the output $\hat{\tau}^{(g)} = \tilde{\tau}^{(g, \text{best})}$.

Otherwise, set $l = l + 1$ and the inner loop continues by returning to the point **Inner Loop**.

Outer Loop Termination. The convergence of the generation outer loop is detected by the following *Test*:

If $\|\hat{\tau}^{(g)} - \hat{\tau}^{(g-1)}\| \leq E_G$, where E_G specifies the accuracy of the RWBS, or alternatively, if $g \geq N_G$, where N_G denotes the maximum allowed generation outer loop iterations, then the generation outer loop is terminated with the output $\hat{\tau}^{\text{RWBS}} = \hat{\tau}^{(g)}$.

Otherwise, set $g = g + 1$ and the outer loop continues by returning to the point **Outer Loop**.

SW-CCG estimation: Perform the SW-CCG estimation as $\hat{\rho}^{\text{RWBS}} = \hat{\rho}^{\text{WLS}}(\hat{\tau}^{\text{RWBS}})$, and if necessary, extract the associated amplitude and phase estimates. ■

For a convenient implementation, the termination tests for the inner and outer loops are usually simplified as $l \geq N_B$ and $g \geq N_G$, respectively. It can be readily shown that the computational complexity of this RWBS-based JTCE estimator is determined by the total number of its CF evaluations, which is given by [15]

$$N_{\text{CF-Eva}}^{\text{RWBS}} = (P_S - 1 + 2N_B) \cdot N_G. \quad (20)$$

In Section VII, we will demonstrate that $N_{\text{CF-Eva}}^{\text{RWBS}}$ is only a fraction of $N_{\text{CF-Eva}}^{\text{ML}}$ given in Eq. (15). Two comments on the above RWBS procedure are in order.

Comment 1. The filling algorithm or initialisation scheme: As discussed in Subsection IV-B, given a particular NP-hard optimisation problem and a specific EA, whether this EA can turn the given problem into an EA-easy one critically depends on whether this EA is appropriately configured to match the given problem [32]. An important algorithmic parameter of the RWBS is the choice of its initialisation scheme or filling algorithm invoked for creating the initial population of each generation. We introduce an ad-hoc filling algorithm that will be referred to as the *Parallel One-Dimensional Randomisation* (PODR), which proves to match well the optimisation problem at hand.

To augment the role played by the filling algorithm in handling local minima of the CF, let us consider a simple $K = 3$ scenario, where the true multiuser propagation delay vector is assumed to be $\tau = [\tau_1 \ \tau_2 \ \tau_3]^T$. In our simulations, it is frequently observed that at the end of a certain generation, say generation g , the above RWBS procedure may yield a temporary estimate such as $\hat{\tau}^{(g)} = [\hat{\tau}_1^{(g)} \ \hat{\tau}_2^{(g)} \ \hat{\tau}_3^{(g)}]^T$, where $\hat{\tau}_k^{(g)} \approx \tau_k$ for example holds for $k = 1$ and 2 but not for $k = 3$. Hence, $\hat{\tau}^{(g)}$ is a local minimum of the CF, and due to the elitist strategy mentioned above, it will be retained in the $(g+1)$ -th generation as $\tilde{\tau}^{(g+1, 1)}$. As suggested by [15], a convenient way of ‘‘filling’’ the $(g+1)$ -th generation is the *Pure Randomisation* (PR), i.e. we simply choose the other (P_S-1) individuals $\{\tilde{\tau}^{(g+1, 2)}, \tilde{\tau}^{(g+1, 3)}, \dots, \tilde{\tau}^{(g+1, P_S)}\}$ on the parameter space in a totally random basis. This PR initialisation scheme has been shown to work well for many practical applications, as it is an effective means to explore other areas of the parameter space. However, for our current problem, as $\tilde{\tau}^{(g+1, 1)}$ is already a local minimum, unless we are lucky enough to come across a new candidate $\tilde{\tau}^{(g+1, i)}$ with $i \neq 1$ satisfying $\{\tilde{\tau}_k^{(g+1, i)} \approx \tau_k | k = 1, 2, 3\}$, $\tilde{\tau}^{(g+1, 1)}$ will perpetually receive the highest weight in the boosting process or convex combination, and $\hat{\tau}^{(g+1)}$ shall therefore

be (at least approximately) equal to $\tilde{\tau}^{(g+1,1)}$, namely, $\hat{\tau}^{(g)}$. In other words, the RWBS procedure might get stuck around the local minimum $\hat{\tau}^{(g)}$. Another popular filling algorithm for RWBS is based on the *Mutation* (Mut) [17], which generates $\{\tilde{\tau}^{(g+1,i)} | i \neq 1\}$ according to $\tilde{\tau}^{(g+1,i)} = \hat{\tau}^{(g)} + \xi$, where ξ is a real-valued Gaussian random vector with a zero mean and a covariance matrix of $\sigma^2 \mathbf{I}_K$. In order to maintain a good convergence performance, σ has to be quite small, e.g. in [17] $\sigma = 0.04$ is chosen. As the $(P_S - 1)$ new individuals are all generated within close proximity of $\hat{\tau}^{(g)}$, The Mut scheme may also be inefficient, especially when for example $\hat{\tau}_3^{(g)}$ is distant from τ_3 .

In order to circumvent the difficulty associated with the two existing initialisation schemes described above, we propose the PODR filling algorithm. Let us define an operator \mathfrak{X} which takes a K -dimensional vector $\tau \in \mathbb{R}^{K \times 1}$ as well as a positive integer $1 \leq k \leq K$ as its input, and produces a K -dimensional output vector $\bar{\tau} = \mathfrak{X}(\tau, k) \in \mathbb{R}^{K \times 1}$, so that $\bar{\tau}\langle k \rangle$ is a random variable uniformly distributed over $[0, T_b)$, while the other $(K - 1)$ entries of $\bar{\tau}$ remain the same as their corresponding counterparts in the input vector τ . Then, as the first step of PODR, we divide the $(P_S - 1)$ individuals that have to be generated into $(K + 1)$ groups $\{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{K+1}\}$, each group having the same size of $\bar{P}_S = (P_S - 1)/(K + 1)$.³ The \bar{P}_S individuals belonging to the last group, namely, \mathcal{G}_{K+1} , will be generated by the PR. By contrast, for $1 \leq k \leq K$, the \bar{P}_S individuals belonging to \mathcal{G}_k will be generated by independently invoking $\mathfrak{X}(\hat{\tau}^{(g)}, k)$ \bar{P}_S times.

If the PODR is used as our filling algorithm for the above $K = 3$ example, for instance, the $(g + 1)$ -th generation of the RWBS shall comprise $\tilde{\tau}^{(g+1,1)}$, \mathcal{G}_1 , \mathcal{G}_2 , \mathcal{G}_3 , and \mathcal{G}_4 . Note that all the \bar{P}_S individuals in \mathcal{G}_3 inherit the first two entries of $\hat{\tau}^{(g)}$, which have already been accurately estimated, and it is only their third entries that are randomised. In this way, there is a far better chance to come up with a new individual in \mathcal{G}_3 that approaches the global minimum, than by the PR which re-guesses all the entries, or by the Mut which only searches around $\hat{\tau}^{(g)}$. We will give a quantitative comparison among the PODR, PR and Mut initialisation schemes using our numerical results in Section VII.

Comment 2. Possible rank-deficiency of $\mathbf{A}_k(\tilde{\tau}_k^{(g,i)})$ and $\mathbf{A}(\tilde{\tau}^{(g,i)})$: Let us assume that $\tilde{\tau}^{(g,i)}$ is the i -th individual in the g -th generation and $\tilde{\tau}_k^{(g,i)}$ its k -th entry. Then special attention should be paid to the issue that when $\tilde{\tau}_k^{(g,i)} \in [0, T_b - (2D - 1)T_c]$, we will have $\mathbf{A}_k(\tilde{\tau}_k^{(g,i)})\langle :, m \rangle = \mathbf{0}_{UQMN}$ for $m = 1, ((Q + 2) + 1), \dots, ((U - 1)(Q + 2) + 1)$, because the received signal \mathbf{r} in this case is completely independent of $b_k(-1)$ or $\{\rho_k^{(1)}(-1), \rho_k^{(2)}(-1), \dots, \rho_k^{(U)}(-1)\}$. To avoid $\mathbf{A}(\tilde{\tau}^{(g,i)}) = [\mathbf{A}_1(\tilde{\tau}_1^{(g,i)}) \mathbf{A}_2(\tilde{\tau}_2^{(g,i)}) \dots \mathbf{A}_K(\tilde{\tau}_K^{(g,i)})]$ becoming rank-deficient, we have to delete these all-zero columns and thereby shrink $\mathbf{A}(\tilde{\tau}^{(g,i)})$ to a $UQMN \times U((Q + 2)K - 1)$ -dimensional matrix. By applying the reduced-size $\mathbf{A}(\tilde{\tau}^{(g,i)})$ to Eq. (13), we arrive at $\tilde{\rho}^{\text{WLS}}(\tilde{\tau}^{(g,i)}) = [[\tilde{\rho}_1^{\text{WLS}}(\tilde{\tau}^{(g,i)})]^T \quad [\tilde{\rho}_2^{\text{WLS}}(\tilde{\tau}^{(g,i)})]^T \quad \dots \quad [\tilde{\rho}_K^{\text{WLS}}(\tilde{\tau}^{(g,i)})]^T]^T \in \mathbb{C}^{U((Q+2)K-1) \times 1}$, where $\tilde{\rho}_k^{\text{WLS}}(\tilde{\tau}^{(g,i)}) \in \mathbb{C}^{U(Q+1) \times 1}$

is the estimate for $[\rho_k^{(1)}(0) \rho_k^{(1)}(1) \dots \rho_k^{(1)}(Q) \rho_k^{(2)}(0) \dots \rho_k^{(U)}(0) \dots \rho_k^{(U)}(Q)]^T$. Note that sometimes there might be more than one entries of $\tilde{\tau}^{(g,i)}$ falling into $[0, T_b - (2D - 1)T_c]$. If that is the case, all the associated all-zero columns should be deleted from $\mathbf{A}(\tilde{\tau}^{(g,i)})$, which implies that the number of columns contained in the reduced-size $\mathbf{A}(\tilde{\tau}^{(g,i)})$ (and therefore the dimension of $\tilde{\rho}^{\text{WLS}}(\tilde{\tau}^{(g,i)})$) deviates from $U(Q + 1)K$ to $U(Q + 2)K$, depending on how many entries of $\tilde{\tau}^{(g,i)}$ are located within $[0, T_b - (2D - 1)T_c]$.

V. CRAMÉR-RAO LOWER BOUND

In order to evaluate the performance of the proposed RWBS-based estimator, in this section we derive the CRLB [18] for the blind JTCE problem at hand, following the approach proposed in [19]. By treating each SW-CCG as a function of its amplitude and phase, Eq. (10b) can be rewritten as⁴:

$$\mathbf{r} = \mathbf{A}(\tau)\boldsymbol{\rho}(\mathbf{a}, \boldsymbol{\theta}) + \mathbf{n}, \quad (21)$$

where \mathbf{a} and $\boldsymbol{\theta}$ are two real-valued vectors having the same size of $\boldsymbol{\rho}$ so that for $1 \leq m \leq U(Q + 2)K$, it holds that $\rho\langle m \rangle = \mathbf{a}\langle m \rangle \exp(\mathbf{j}\boldsymbol{\theta}\langle m \rangle)$. Given Eq. (21), the Log-Likelihood Function (LLR) of the received signal \mathbf{r} conditioned on the unknown channel parameters $\boldsymbol{\eta} = [\tau^T \mathbf{a}^T \boldsymbol{\theta}^T]^T$ can be shown to obey:

$$\ln L(\mathbf{r}|\boldsymbol{\eta}) = -[\mathbf{r} - \mathbf{A}(\tau)\boldsymbol{\rho}(\mathbf{a}, \boldsymbol{\theta})]^H \boldsymbol{\Lambda}^{-1} [\mathbf{r} - \mathbf{A}(\tau)\boldsymbol{\rho}(\mathbf{a}, \boldsymbol{\theta})] - UQMN \ln(\pi \det(\boldsymbol{\Lambda})). \quad (22)$$

For space economy, we have to report the Fisher Information Matrix (FIM) [18] as follows without providing its detailed derivation:

$$\begin{aligned} \bar{\mathbf{F}} &= E \left\{ \left[\frac{\partial \ln L(\mathbf{r}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \right] \left[\frac{\partial \ln L(\mathbf{r}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \right]^T \right\} \\ &= \begin{bmatrix} \bar{\mathbf{F}}_{\tau\tau} & \bar{\mathbf{F}}_{\tau\mathbf{a}} & \bar{\mathbf{F}}_{\tau\boldsymbol{\theta}} \\ \bar{\mathbf{F}}_{\tau\mathbf{a}}^T & \bar{\mathbf{F}}_{\mathbf{a}\mathbf{a}} & \bar{\mathbf{F}}_{\mathbf{a}\boldsymbol{\theta}} \\ \bar{\mathbf{F}}_{\tau\boldsymbol{\theta}}^T & \bar{\mathbf{F}}_{\mathbf{a}\boldsymbol{\theta}}^T & \bar{\mathbf{F}}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{bmatrix}, \end{aligned} \quad (23)$$

In the above expression, $\bar{\mathbf{F}}_{\tau\tau} \in \mathbb{R}^{K \times K}$, $\bar{\mathbf{F}}_{\tau\mathbf{a}} \in \mathbb{R}^{K \times U(Q+2)K}$, $\bar{\mathbf{F}}_{\tau\boldsymbol{\theta}} \in \mathbb{R}^{K \times U(Q+2)K}$, $\bar{\mathbf{F}}_{\mathbf{a}\mathbf{a}} \in \mathbb{R}^{U(Q+2)K \times U(Q+2)K}$, $\bar{\mathbf{F}}_{\mathbf{a}\boldsymbol{\theta}} \in \mathbb{R}^{U(Q+2)K \times U(Q+2)K}$,

⁴Note that the considerations raised in *Comment 2.* of Section IV also apply to Eq. (21), i.e. if some entries of τ fall into $[0, T_b - (2D - 1)T_c]$, then the corresponding all-zero columns should be deleted from $\mathbf{A}(\tau)$. Therefore, we need to remove the corresponding entries from $\boldsymbol{\rho}$, \mathbf{a} and $\boldsymbol{\theta}$, which in this situation have no influence on the received signal \mathbf{r} and therefore cannot be estimated by any means. With the reduced-sized $\mathbf{A}(\tau)$, $\boldsymbol{\rho}$, \mathbf{a} and $\boldsymbol{\theta}$, the following derivation of the CRLB bears little difference from the regular case where no entry of τ falls into $[0, T_b - (2D - 1)T_c]$, as is assumed in this section for simplicity.

³Here we assume $\bar{P}_S \in \mathbb{N}^+$.

and $\overline{\mathbf{F}}_{\theta\theta} \in \mathbb{R}^{U(Q+2)K \times U(Q+2)K}$ are defined as:

$$\overline{\mathbf{F}}_{\tau\tau}\langle k, k' \rangle = 2\Re \left\{ \boldsymbol{\rho}^H \frac{\partial \mathbf{A}^H}{\partial \tau\langle k \rangle} \boldsymbol{\Lambda}^{-1} \frac{\partial \mathbf{A}}{\partial \tau\langle k' \rangle} \boldsymbol{\rho} \right\}, \quad (24a)$$

$$\overline{\mathbf{F}}_{\tau\mathbf{a}}\langle k, m \rangle = 2\Re \left\{ \boldsymbol{\rho}^H \frac{\partial \mathbf{A}^H}{\partial \tau\langle k \rangle} \boldsymbol{\Lambda}^{-1} \mathbf{A} \frac{\partial \boldsymbol{\rho}}{\partial \mathbf{a}\langle m \rangle} \right\}, \quad (24b)$$

$$\overline{\mathbf{F}}_{\tau\boldsymbol{\theta}}\langle k, m \rangle = 2\Re \left\{ \boldsymbol{\rho}^H \frac{\partial \mathbf{A}^H}{\partial \tau\langle k \rangle} \boldsymbol{\Lambda}^{-1} \mathbf{A} \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}\langle m \rangle} \right\}, \quad (24c)$$

$$\overline{\mathbf{F}}_{\mathbf{a}\mathbf{a}}\langle m, n \rangle = 2\Re \left\{ \frac{\partial \boldsymbol{\rho}^H}{\partial \mathbf{a}\langle m \rangle} \mathbf{A}^H \boldsymbol{\Lambda}^{-1} \mathbf{A} \frac{\partial \boldsymbol{\rho}}{\partial \mathbf{a}\langle n \rangle} \right\}, \quad (24d)$$

$$\overline{\mathbf{F}}_{\mathbf{a}\boldsymbol{\theta}}\langle m, n \rangle = 2\Re \left\{ \frac{\partial \boldsymbol{\rho}^H}{\partial \mathbf{a}\langle m \rangle} \mathbf{A}^H \boldsymbol{\Lambda}^{-1} \mathbf{A} \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}\langle n \rangle} \right\}, \quad (24e)$$

$$\overline{\mathbf{F}}_{\boldsymbol{\theta}\boldsymbol{\theta}}\langle m, n \rangle = 2\Re \left\{ \frac{\partial \boldsymbol{\rho}^H}{\partial \boldsymbol{\theta}\langle m \rangle} \mathbf{A}^H \boldsymbol{\Lambda}^{-1} \mathbf{A} \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}\langle n \rangle} \right\}, \quad (24f)$$

where $\frac{\partial \boldsymbol{\rho}}{\partial \mathbf{a}\langle m \rangle} = [0 \cdots 0 \exp(\mathbf{j}\boldsymbol{\theta}\langle m \rangle) 0 \cdots 0]^T$, $\frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}\langle m \rangle} = [0 \cdots 0 \mathbf{j}\mathbf{a}\langle m \rangle \exp(\mathbf{j}\boldsymbol{\theta}\langle m \rangle) 0 \cdots 0]^T$, and $\frac{\partial \mathbf{A}}{\partial \tau\langle k \rangle}$ can be computed by the following ‘‘formula chain’’ (c.f. Eqs.(8)-(10)):

$$\frac{\partial \mathbf{A}}{\partial \tau\langle k \rangle} = \begin{bmatrix} \mathbf{O} & \cdots & \mathbf{O} & \frac{\partial \mathbf{A}_k(\tau_k)}{\partial \tau_k} & \mathbf{O} & \cdots & \mathbf{O} \end{bmatrix}, \quad (25a)$$

$$\frac{\partial \mathbf{A}_k(\tau_k)}{\partial \tau_k} = \text{diag} \left\{ \frac{\partial \mathbf{A}_k^{(1)}(\tau_k)}{\partial \tau_k}, \frac{\partial \mathbf{A}_k^{(2)}(\tau_k)}{\partial \tau_k}, \dots, \frac{\partial \mathbf{A}_k^{(U)}(\tau_k)}{\partial \tau_k} \right\}, \quad (25b)$$

$$\frac{\partial \mathbf{A}_k^{(u)}(\tau_k)}{\partial \tau_k} = \text{diag} \left\{ \frac{\partial \mathbf{H}_{k,1}^{(u)}(\tau_k)}{\partial \tau_k}, \frac{\partial \mathbf{H}_{k,2}^{(u)}(\tau_k)}{\partial \tau_k}, \dots, \frac{\partial \mathbf{H}_{k,Q}^{(u)}(\tau_k)}{\partial \tau_k} \right\} \mathbf{F}, \quad (25c)$$

$$\frac{\partial \mathbf{H}_{k,q}^{(u)}(\tau_k)}{\partial \tau_k} = \begin{bmatrix} \mathbf{C}_{k,-2}^{(u)}(q) \frac{\partial \psi(\tau_k)}{\partial \tau_k} & \mathbf{C}_{k,-1}^{(u)}(q) \frac{\partial \psi(\tau_k)}{\partial \tau_k} \\ \mathbf{C}_{k,0}^{(u)}(q) \frac{\partial \psi(\tau_k)}{\partial \tau_k} \end{bmatrix}. \quad (25d)$$

In Eq. (25d), we have $\frac{\partial \psi(\tau_k)}{\partial \tau_k} = \left[\dot{\psi} \left(\frac{T_c}{M} - \tau_k \right) \cdots \dot{\psi} \left(T_b + \frac{(2MD-1)T_c}{M} - \tau_k \right) \right]^T \in \mathbb{R}^{(MN+2MD-1) \times 1}$. Here $\dot{\psi}(t)$ represents the 1st-order derivative of $\psi(t)$, for which a closed-form formula has been provided in [49]. Finally, given $\overline{\mathbf{F}}$, the CRLB may be obtained as the diagonal entries of $(\overline{\mathbf{F}})^{-1}$.

VI. DIFFERENTIAL MUD BASED ON BLIND JTCE

Let us now briefly portray our differential (non-coherent) MUD with the aid of the proposed blind JTCE scheme. Given the estimated SW-CCG vector:

$$\begin{aligned} \hat{\boldsymbol{\rho}} &= \begin{bmatrix} \hat{\rho}_1^{(1)}(-1) & \cdots & \hat{\rho}_1^{(1)}(Q) & \cdots & \hat{\rho}_1^{(U)}(-1) & \cdots & \hat{\rho}_1^{(U)}(Q) \\ \hat{\rho}_2^{(1)}(-1) & \cdots & \hat{\rho}_2^{(1)}(Q) & \cdots & \hat{\rho}_2^{(U)}(-1) & \cdots & \hat{\rho}_2^{(U)}(Q) \\ \vdots & & \vdots & & \vdots & & \vdots \\ \hat{\rho}_K^{(1)}(-1) & \cdots & \hat{\rho}_K^{(1)}(Q) & \cdots & \hat{\rho}_K^{(U)}(-1) & \cdots & \hat{\rho}_K^{(U)}(Q) \end{bmatrix}^T, \end{aligned} \quad (26)$$

the decision statistics for $b_k^{(u)}(q)$ can be formed as:

$$\hat{b}_k^{(u)}(q) = \hat{\rho}_k^{(u)}(q) \left(\hat{\rho}_k^{(u)}(q-1) \right)^*. \quad (27)$$

To see the validity of Eq. (27), let us assume that the SW-CCG vector has been estimated with satisfactory reliability, i.e. the relationship $\hat{\rho}_k^{(u)}(q) \approx \rho_k^{(u)}(q)$ is true for all the possible values of u, k and q involved in Eq. (26). In this case, Eq. (27) can be further developed as:

$$\begin{aligned} \hat{b}_k^{(u)}(q) &\approx \rho_k^{(u)}(q) \left(\rho_k^{(u)}(q-1) \right)^* \\ &= \dot{b}_k^{(u)}(q) \left(\dot{b}_k^{(u)}(q-1) \right)^* \alpha_k^{(u)}(q) \left(\alpha_k^{(u)}(q-1) \right)^* \\ &\approx b_k^{(u)}(q) \left| \alpha_k^{(u)}(q) \right|^2. \end{aligned} \quad (28)$$

The last approximation in the above expression follows from the assumption that the time-variant CCGs remain approximately constant within at least two consecutive symbol intervals [7], [8]. As shown by Eq. (28), $\hat{b}_k^{(u)}(q)$ is an estimate of $b_k^{(u)}(q)$ up to an amplitude of $|\alpha_k^{(u)}(q)|^2$. Hence by decoding the quadrant where $\arctan(\hat{b}_k^{(u)}(q))$ falls, we can obtain an estimate of the two information bits carried by $b_k^{(u)}(q)$.

Clearly, the estimation of the SW-CCG vector $\boldsymbol{\rho}$ plays a central part in the demodulation process, and to a great extent it predetermines the overall achievable BER performance. Recall the last component of the proposed RWBS-based estimator in Subsection IV-C, **SW-CCG estimation**, where the SW-CCG vector estimate $\hat{\boldsymbol{\rho}}^{\text{RWBS}} = \tilde{\boldsymbol{\rho}}^{\text{WLS}}(\hat{\boldsymbol{\tau}}^{\text{RWBS}})$ is obtained by setting $\hat{\boldsymbol{\tau}} = \hat{\boldsymbol{\tau}}^{\text{RWBS}}$ in Eq. (13). From a MUD point of view, Eq. (13) amounts to recovering the SW-CCGs with the aid of a De-Correlation (DC) receiver⁵, and it is in general inferior to its Minimum Mean Squared Error (MMSE) counterpart, which takes the form:

$$\hat{\boldsymbol{\rho}}^{\text{MMSE}} = \mathbf{A}^H(\boldsymbol{\tau}) \mathbf{C}_\rho^H (\mathbf{A}(\boldsymbol{\tau}) \mathbf{C}_\rho \mathbf{A}^H(\boldsymbol{\tau}) + \boldsymbol{\Lambda}^H)^{-1} \mathbf{r}, \quad (29)$$

where \mathbf{C}_ρ denotes the auto-correlation matrix of the SW-CCG vector $\boldsymbol{\rho}$, i.e. $\boldsymbol{\Sigma}_\rho = E(\boldsymbol{\rho}\boldsymbol{\rho}^H)$. As the SW-CCGs associated with different users, subcarriers, or symbols (i.e. different k, u or q) may be assumed to be independent of each other, $\mathbf{C}_\rho \in \mathbb{R}^{U(Q+2)K \times 1}$ is a diagonal matrix with its diagonal elements given by $\boldsymbol{\rho} \odot \boldsymbol{\rho}^*$. This implies that \mathbf{C}_ρ is dependent only on the squared amplitude of the fading gains. In practice, Eq. (29) can be implemented by replacing the matrices $\mathbf{A}(\boldsymbol{\tau})$ and \mathbf{C}_ρ with their RWBS-based estimates $\mathbf{A}(\hat{\boldsymbol{\tau}}^{\text{RWBS}})$ and $\mathbf{C}_{\hat{\boldsymbol{\rho}}^{\text{RWBS}}}$, respectively.

The above discussion has assumed that independent information streams are transmitted on each subcarrier to achieve maximum uplink bit rate. In the cases where some users prefer to send the same information stream on several subcarriers simultaneously to achieve frequency diversity, the decision statistic formulated in Eq. (27) has to be modified to:

$$\hat{b}_k^l(q) = \sum_{u \in \mathcal{U}_k^l} \lambda_k^{(u)}(q) \hat{\rho}_k^{(u)}(q) \left(\hat{\rho}_k^{(u)}(q-1) \right)^*, \quad (30)$$

⁵In contrast to the classic form of DC receiver (e.g. Eq. (6.28) of [3]), Eq. (13) in the current paper has been adapted to accommodate bandlimited noise.

where \mathcal{U}_k^l represents the set of subcarriers employed for transmitting the l -th information stream of user k , and $\lambda_k^{(u)}(q)$ is a combining weight, which may be optimised for example according to the Maximum signal-to-noise Ratio Combining (MRC) rule, as discussed in [7].

VII. PERFORMANCE EVALUATION

A. Simulation System Set Up and Performance Metrics

The system parameters adopted in our simulations are $M = 2$, $N = 16$, $D = 4$, $U = 2$, $\gamma = 0.22$ and $f_d = 0.2$, where f_d is the Doppler frequency normalised by the symbol rate $1/T_b$. The long spreading codes are generated as random bipolar sequences, where every chip has the same probability of being $+1$ or -1 . We assume $A_2 = A_3 = \dots = A_K$ and refer to $20 \log \left(\frac{A_2}{A_1} \right)$ as the Near-Far Ratio (NFR). Unless otherwise stated, the asynchronous MC-DS-CDMA uplink is assumed to be shared by $K = 10$ users with the NFR tied to 10 dB, and we will always report the results associated with the lowest-power user, i.e. user 1, for characterising the near-far resistance of the proposed channel estimator. As in all other applications [15]–[17], the algorithmic parameters of RWBS have to be determined empirically. For our PODR assisted RWBS procedure (RWBS-PODR), we have found $P_S = (20(K+1)+1)$ (i.e. $\bar{P}_S = 20$), $N_G = 40$, $N_B = P_S$, and $E_G = E_B = 0.000001$ to be appropriate.

The primary performance criteria of our interest include the probability of correct acquisition (denoted as P_{ca}), the Mean Square Error (MSE) of the delay, amplitude and phase estimates, as well as the Bit Error Rate (BER) of the differential multiuser detectors with the aid of the proposed blind JTCE scheme. Note that the acquisition of a specific user is deemed as “correct” only when the corresponding delay estimation error has an absolute value less than $T_c/2$ [13], [14], [19], [22], [49]. Moreover, to highlight how the system’s performance is impacted by the tradeoff between the achievable frequency diversity and the bit rate, in conducting our simulations, we have considered the following two specific scenarios: 1) a single information stream is transmitted via all available subcarriers for achieving MAXimum frequency DIVersity (MAXDIV), and 2) an independent information stream is assigned to each subcarrier for attaining MAXimum BiT Rate (MAXBTR).

B. Evaluation of RWBS-Based JTCE Estimator

In Fig. 3-A, the P_{ca} values of the RWBS-based channel estimators which adopt different filling algorithms or initialisation schemes are plotted against the number of active users K , in comparison with those of the PIC-based counterparts relying on different number of stages V . The results appearing in Fig. 3 are generated for a fairly benign environment of $E_b/N_0 = 15$ dB and NFR = 0 dB, where E_b is the energy per bit. It is observed from Fig. 3-A that the acquisition capabilities of the RWBS procedures assisted by the PR and Mut initialisation schemes (abbreviated hereafter as RWBS-PR and RWBS-Mut, respectively) suffer a substantial degradation when K increases from 2 to 3. By contrast, the proposed RWBS-PODR estimator does not exhibit such deficiency. To obtain further understandings into this remarkable distinction,

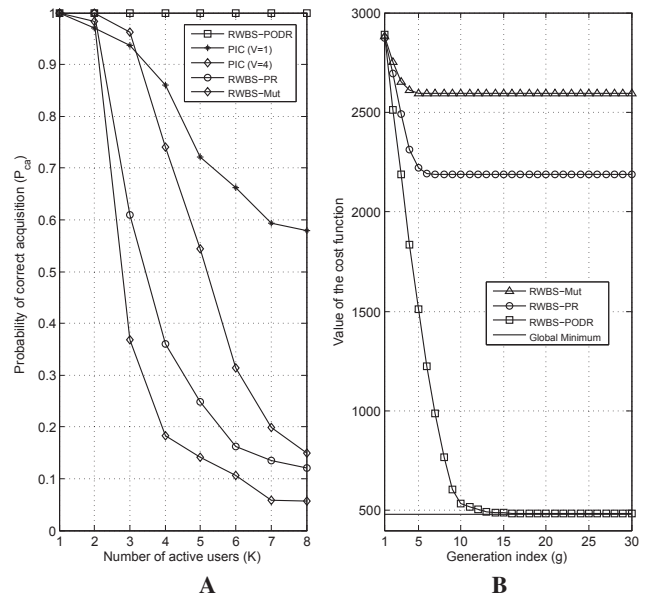


Fig. 3. **A:** P_{ca} v.s. number of active users K , and **B:** value of the cost function v.s. generation index for the case of $K = 5$, given $E_b/N_0 = 15$ dB and NFR = 0 dB. The results presented in the figure are generated for the MAXDIV scenario and are averaged over 1000 independent runs. Simulation results for the MAXBTR scenario under the same conditions demonstrate a quite similar trend to their MAXDIV counterparts shown in this figure, and they are hence not shown here for space reason.

we plot the CF value against the generation index for the RWBS-PR, RWBS-Mut and RWBS-PODR in Fig. 3-B, with K set to 5. It is observed that when the PODR is adopted as the filling algorithm, the CF value can reach its global minimum as the population evolves, which shows that the RWBS-PODR is capable of capturing the ML estimate with the number of generation iterations $g \geq 15$. On the other hand, the CF values yielded by the RWBS-PR and RWBS-Mut upon their convergence are well above the global minimum. This implies that they tend to frequently get trapped at a local minimum of the CF, and therefore cannot guarantee reliable parameter estimation for all users. Furthermore, as indicated in Fig. 3-A, the PIC-based schemes are also incompetent for reliable blind JTCE. The reason is that, owing to the “error-propagation” effects mentioned in Subsection IV-B, some of the cancellation stages are in fact detrimental rather than beneficial as the number of users K increases.

Figs. 4 and 5 are dedicated to the performance evaluation of the proposed RWBS-PODR aided channel estimator. Indeed, it would be meaningful to provide a comparison between the RWBS-PODR estimator and its optimal ML counterpart based on exhaustive search. However, the latter is excessively complex and its simulation turns out to be unacceptably time-consuming even for a moderate number of users. Instead, we compare the MSE of the RWBS-PODR estimator to the CRLB, which serves as a lower bound for the attainable MSE of all unbiased estimators [18], including the ML and the RWBS-PODR based ones. To quantify the dependence of the JTCE quality on the length of observation window, all the results in Figs. 4 and 5 have been generated for the cases of $Q = 10$ and $Q = 20$, respectively. Figs. 4-A plots the P_{ca} values achieved against the E_b/N_0 values for both the MAXDIV and MAXBTR scenarios, where it can

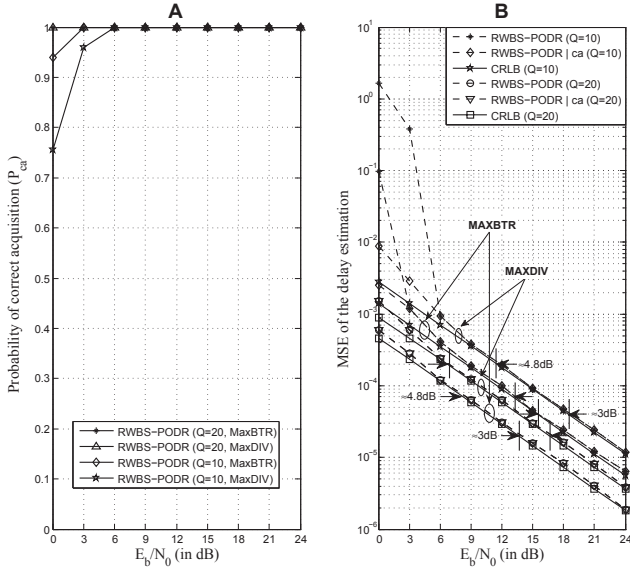


Fig. 4. **A:** P_{ca} v.s. E_b/N_0 , and **B:** the MSE of delay estimation (normalized by T_c^2) v.s. E_b/N_0 . $K = 10$ users, NFR = 10 dB, and all results are averaged over 1000 independent runs.

be seen that for the observation window of $Q = 20$, P_{ca} reaches the value of 1 for all the E_b/N_0 values tested. For the observation window of $Q = 10$, however, P_{ca} approaches 1 only when $E_b/N_0 \geq 3$ dB for the MAXBTR case, or when $E_b/N_0 \geq 6$ dB for the MAXDIV case. In Fig. 4-B, the “unconditioned” MSE of the delay estimation and its counterpart conditioned on correct acquisition (ca) are plotted along with their corresponding CRLB. We observe that the MSE within the high E_b/E_0 range and the corresponding CRLB are 3 dB better in the MAXBTR scenario than in the MAXDIV scenario. This can be explained by the fact that a MAXBTR receiver actually enjoys a $10 \log_{10}(U)$ dB SNR gain compared to a MAXDIV one, provided that they both operate under the same E_b/N_0 , since the latter has a spreading gain (in the time and frequency domains) which is twice ($U = 2$) as much as that of the former. As to the effect of the observation window length, considering the MAXDIV scenario as an example, we note that in the case of $Q = 10$, the unconditioned and conditioned MSEs coincide with each other when E_b/N_0 exceeds 6 dB, for the proposed channel estimator achieves $P_{ca} \approx 1$ under this SNR condition as shown in Fig. 4-A. If the observation window lasts for $20T_b$, the same phenomenon occurs for $E_b/N_0 \geq 0$ dB. More importantly, the MSE of delay estimation approaches its CRLB in the case of $Q = 10$ for $E_b/N_0 \geq 9$ dB, while for $Q = 20$, only $E_b/N_0 \geq 6$ dB is required. The advantage associated with a longer observation window has also been confirmed in the MAXBTR scenario.

Similar trends to those described above can also be found in Fig. 5-A and Fig. 5-B, where the amplitude and phase estimation performance are investigated, respectively. However, there are two distinctions worth mentioning. Firstly, the gap between the conditioned and unconditioned MSE is less evident for the amplitude or phase offset estimation than for the delay estimation. This behaviour has also been witnessed in [13], where pilot-assisted timing and channel estimation was invoked for bandlimited long-code-based MC-DS-CDMA

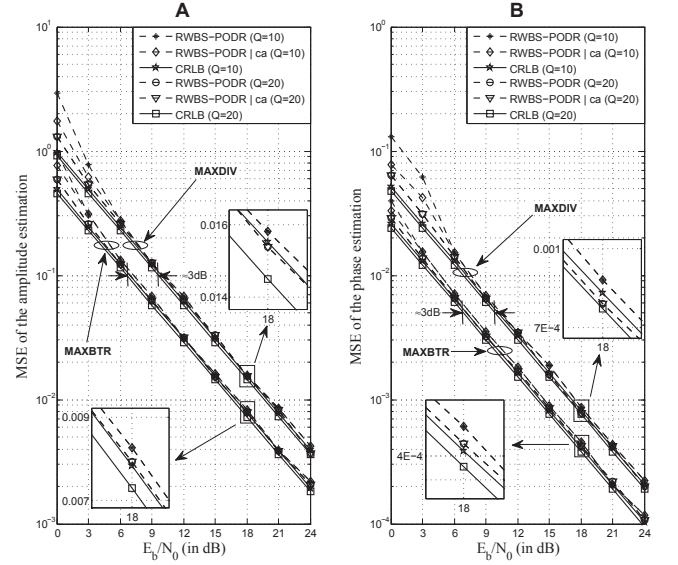


Fig. 5. **A:** the MSE of amplitude estimation (normalized by the mean value of the true squared amplitude) v.s. E_b/N_0 , and **B:** the MSE of phase estimation (normalized by $4\pi^2$) v.s. E_b/N_0 . $K = 10$ users, NFR = 10 dB, and all results are averaged over 1000 independent runs.

systems transmitting over quasi-static fading channels. Secondly, from Fig. 4-B we observe that when Q is increased from 10 to 20, the performance of the delay estimation exhibits a 4.8 dB improvement in the high E_b/N_0 range for both the MAXDIV and MAXBTR scenarios. By contrast, the amplitude and phase estimations hardly benefit from an extended observation window, as can be seen in Fig. 5. To gain a deeper insight into this interesting discrepancy, let us consider a $20T_b$ -long observation window, which can be conceptually divided into two parts, i.e. the *original* part and the *extended* part, each having a length of $10T_b$. Since the propagation delay is deemed to be constant over the entire observation window, both the *original* and *extended* parts may be exploited for the delay estimation. Thus the delay estimator relying on the entire observation window tends to be better than those relying on the *original* part or the *extended* part alone. On the other hand, the CCGs are assumed to be time-variant and their values can only be viewed as stationary on the symbol duration level. Therefore, the *extended* part of the observation window cannot improve the estimation of the CCGs (and hence their amplitude and phase) associated with the *original* part, or vice versa. It follows the inference that an extended observation window does not necessarily guarantee to improve the accuracy of the amplitude or phase estimation.

C. Evaluation of Differential MUD Based on Proposed Blind JTCE

In Fig. 6, we report the BERs of the differential DC and MMSE detectors based on the proposed blind RWBS-PODR JTCE estimator, in comparison to that of the idealised differential MMSE detector relying on the perfect channel state information. From Fig. 4, we note that a $20T_b$ -long observation window brings about 4.8 dB advantage for the delay estimation against its $10T_b$ -long counterpart, while from Fig. 5, we seen that an extended observation window does not guarantee an improvement in the accuracy of the CCG estimate because

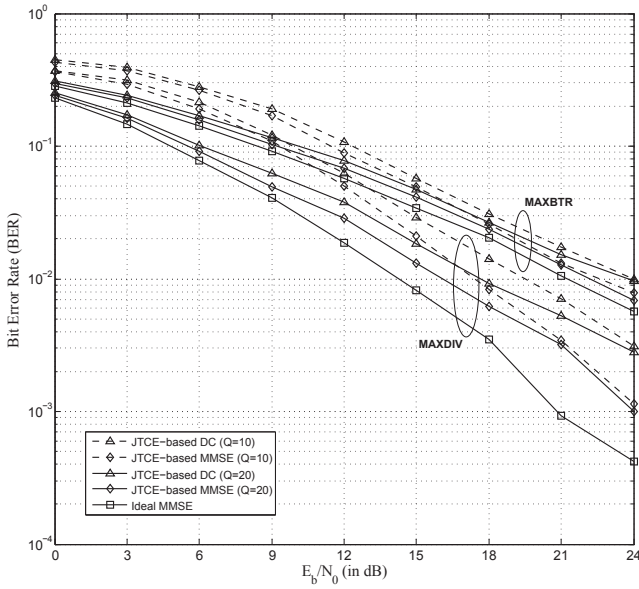


Fig. 6. System's achievable BER versus E_b/N_0 . $K = 10$ and NFR = 10dB. Every result is obtained by repeating the simulation until at least 100 bit errors are found.

of the time-varying nature of the CCGs. Interestingly, Fig. 6 also demonstrates that the BER improvement due to increasing the observation window from $Q = 10$ to $Q = 20$ becomes less evident and may even be imperceptible as the E_b/N_0 grows, and this conclusion holds for both the MAXDIV and MAXBTR setups. Another interesting observation is that although the MAXBTR setting offers the 3 dB superiority over the MAXDIV counterpart as far as the channel estimation is concerned, it is generally inferior to the latter in terms of BER. Obviously, the MAXDIV receiver can effectively exploit frequency diversity and thereby enhances the reliability of the decision statistics when independent fading is present on each subcarrier. Overall, the results shown in Fig. 6 confirm that the proposed blind JTCE algorithm is effective, since the multiuser detectors using the delay and channel estimates provided by this blind estimator achieve satisfactory BER performance.

D. Complexity Analysis

Let us now compare the complexity of our RWBS-PODR estimator with that of its ML counterpart. Recalling Eq. (14), we note that both the RWBS-PODR and ML based estimators rely on searching for the minimum of the CF within the K -dimensional parameter space $[0, T_b)^K$. Therefore, the complexity of these two estimators may be evaluated by a common formula

$$C = C_{\text{CF-Eva}} \cdot N_{\text{CF-Eva}}, \quad (31)$$

where $N_{\text{CF-Eva}}$ is the total number of CF evaluations involved in the search process, and $C_{\text{CF-Eva}}$ is the complexity per CF evaluation. Since $C_{\text{CF-Eva}}$ is independent of the estimation scheme, the complexity of the RWBS-PODR and ML based estimators are mainly dictated by their corresponding $N_{\text{CF-Eva}}$ values, denoted as $N_{\text{CF-Eva}}^{\text{RWBS-PODR}}$ and $N_{\text{CF-Eva}}^{\text{ML}}$, respectively. Furthermore, $N_{\text{CF-Eva}}^{\text{RWBS-PODR}}$ is specified in Eq. (20), while $N_{\text{CF-Eva}}^{\text{ML}}$ is specified in Eq. (15) for the exhaustive grid-based ML search with R being the number of equally-spaced bins for quantising the chip interval.

To gain a more tangible impression on the efficiency of the proposed RWBS-PODR estimator with respect to its ML counterpart, given $K = 10$ and $N = 16$ in the simulation system, let us now calculate $N_{\text{CF-Eva}}^{\text{ML}}$ assuming $R = 2$ as the quantisation resolution⁶ and $N_{\text{CF-Eva}}^{\text{RWBS-PODR}}$ using the RWBS algorithmic parameters adopted in our simulation, namely, $P_S = (20(K+1)+1) = 221$, $N_G = 40$ and $N_B = P_S = 221$. It is straightforward to arrive at:

$$N_{\text{CF-Eva}}^{\text{ML}} = (2 \cdot 16)^{10} \approx 1.126 \times 10^{15}, \quad (32a)$$

$$N_{\text{CF-Eva}}^{\text{RWBS-PODR}} = (221 - 1 + 2 \cdot 221) \cdot 40 = 26\,480, \quad (32b)$$

$$N_{\text{CF-Eva}}^{\text{ML}} / N_{\text{CF-Eva}}^{\text{RWBS-PODR}} = 1.126 \times 10^{15} / 26\,480 \approx 4.25 \times 10^{10}. \quad (32c)$$

To conclude our complexity analysis, we note that $C_{\text{CF-Eva}}$ is principally determined by the computational overhead required for calculating $\tilde{\rho}^{\text{WLS}}(\tilde{\tau})$, which is formulated in Eq. (13). Skipping the detailed analysis, it may be worked out that the calculation of Eq. (13) entails a complexity of

$$\begin{aligned} C_{\text{CF-Eva}} \approx & U^3(Q^3 + 2Q^2)KM^2N^2 + U^2(Q^2 + 2Q)KMN \\ & + U^2(Q+2)^2K^2 + U^3(Q+2)^2QK^2MN \\ & + \eta U^3(Q+2)^3K^3, \end{aligned} \quad (33)$$

where the term $\eta U^3(Q+2)^3K^3$ is contributed by $[\mathbf{A}^H(\tilde{\tau})\mathbf{\Lambda}^{-1}\mathbf{A}(\tilde{\tau})]^{-1}$, and η is a constant relying on the specific approach invoked for this matrix inversion.

VIII. CONCLUSIONS

In this paper, we have conceived a JTCE scheme based on the RWBS optimisation algorithm for bandlimited long-code-aided MC-DS-CDMA systems, which requires neither the transmission of known pilots nor the assumption that the channel state remains constant within the entire observation window. An ad-hoc filling algorithm named the PODR is proposed to work in conjunction with the RWBS procedure, which is capable of effectively avoiding local minima of the CF, and is shown to be more efficient than existing generation initialisation alternatives for the problem at hand. The CRLB corresponding to the JTCE task of interest has been derived as a performance benchmark. Quantitatively, for the case of $K = 10$ users, $E_b/N_0 \geq 3$ dB and a 10 dB NFR, the RWBS-PODR estimator using an observation window of $20T_b$ is shown to approach the CRLB at a complexity which is 10 orders of magnitude lower in comparison to its full ML-based counterpart relying on exhaustive grid-based search. Hence our solution strikes a very efficient and practical near-optimal low-complexity alternative to the ML-based estimator.

As a closing remark, we note that albeit the RWBS has a concise structure which allows easy programming and tuning, a whole host of other EAs may be applied to the multi-variate optimisation problem under consideration (c.f. Eq. (14)). For future work, a comparative study in the context of JTCE may be pursued to compare the RWBS, in terms of both

⁶In fact, $R = 2$ is the minimum possible value for the resolution of the grid search and it is often insufficient for achieving the necessary timing accuracy in practice. Here by choosing $R = 2$ we expect to obtain a *lower bound* for the complexity of the ML estimator.

performance and efficiency, with other EA alternatives, such as the GA, ACO, PSO and DEA.

ACKNOWLEDGMENTS

Mr. Tiexing Wang is gratefully acknowledged for his assistance with some the diagrams in this paper.

REFERENCES

- [1] S. Kondo and L. B. Milstein, "Performance of multicarrier DS CDMA systems," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 238–246, Feb. 1996.
- [2] L. Hanzo, L.-L. Yang, E. L. Kuan, and K. Yen, *Single- and Multi-Carrier DS-CDMA: Multi-User Detection, Space-Time Spreading, Synchronization, Standards and Networking*. John Wiley & Sons, 2003.
- [3] L.-L. Yang, *Multicarrier Communications*. John Wiley & Sons, 2009.
- [4] L.-L. Yang and L. Hanzo, "Multicarrier DS-CDMA: a multiple access scheme for ubiquitous broadband wireless communications," *IEEE Commun. Mag.*, vol. 41, no. 10, pp. 116–124, Oct. 2003.
- [5] S. L. Miller and B. J. Rainbolt, "MMSE detection of multicarrier CDMA," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2356–2362, Nov. 2000.
- [6] J. Namgoong, T. F. Wong, and J. S. Lehnert, "Subspace multiuser detection for multicarrier DS-CDMA," *IEEE Trans. Commun.*, vol. 48, no. 11, pp. 1897–1908, Nov. 2000.
- [7] U. Madhow, K. Bruvold, and L. J. Zhu, "Differential MMSE: a framework for robust adaptive interference suppression for DS-CDMA over fading channels," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1377–1390, Aug. 2005.
- [8] D. J. Sadler and A. Manikas, "MMSE multiuser detection for array multicarrier DS-CDMA in fading channels," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2348–2358, July 2005.
- [9] B. Hu, L.-L. Yang, and L. Hanzo, "Time- and frequency-domain-spread generalized multicarrier DS-CDMA using subspace-based blind and group-blind space-time multiuser detection," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 3235–3241, Sept. 2008.
- [10] Z. Xu and M. K. Tsatsanis, "Blind channel estimation for long-code multiuser CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 4, pp. 988–1001, Apr. 2000.
- [11] S. Buzzi and H. V. Poor, "On parameter estimation in long-code DS/CDMA systems: Cramér-Rao bounds and least-squares algorithms," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 545–559, Feb. 2003.
- [12] S. Buzzi, L. Venturino, A. Zappone, and A. De Maio, "Blind user detection in doubly dispersive DS/CDMA fading channels," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1446–1451, Mar. 2010.
- [13] S. Wang, J. Lu, J. An, and L. Hanzo, "Uplink channel estimation for bandlimited MC-DS-CDMA systems relying on long spreading codes," in *Proc. 2011 IEEE VTC – Fall*, pp. 1–5.
- [14] S. Wang, A. Wang, J. An, and L. Hanzo, "Timing acquisition for bandlimited long-code DS-CDMA in doubly-selective fading channels," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 671–674, Nov. 2011.
- [15] S. Chen, X. Wang and C. J. Harris, "Experiments with repeating weighted boosting search for optimization in signal processing applications," *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 35, no. 4, pp. 682–693, Aug. 2005.
- [16] S. F. Page, S. Chen, C. J. Harris, and N. M. White, "Repeated weighted boosting search for discrete or mixed search space and multiple-objective optimisation," *Applied Soft Computing*, vol. 12, no. 9, pp. 2740–2755, Sept. 2012.
- [17] J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Joint channel estimation and multiuser detection for SDMA/OFDM based on dual repeated weighted boosting search," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3265–3275, Sept. 2011.
- [18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.
- [19] S. Buzzi and V. Massaro, "Parameter estimation and multiuser detection for bandlimited long-code CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2307–2317, June 2008.
- [20] B. Smida, L. Hanzo, and S. Affes, "Exact BER performance of asynchronous MC-DS-CDMA over fading channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1249–1254, Apr. 2010.
- [21] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [22] S. Buzzi, L. Venturino, and A. Zappone, "Multipath delay acquisition in asynchronous doubly-selective DS/CDMA fading channels," *IEEE Commun. Lett.*, vol. 14, no. 4, pp. 276–278, Apr. 2010.
- [23] D. Divsalar, M. K. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 258–268, Feb. 1998.
- [24] J. H. Holland, *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975.
- [25] L. D. Davis, *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, 1991.
- [26] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 26, no. 1, pp. 29–41, Feb. 1996.
- [27] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Mag.*, vol. 1, no. 4, pp. 28–39, Nov. 2006.
- [28] J. Kennedy and R. Eberhart, "Particle swarm optimization," in: *Proc. 1995 IEEE Int. Conf. Neural Networks*, vol. 4, pp. 1942–1948.
- [29] J. Kennedy and R. Eberhart, *Swarm Intelligence*. Morgan Kaufmann, 2001.
- [30] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Springer Verlag, 2005.
- [31] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evolutionary Computation*, vol. 13, no. 2, pp. 398–417, Apr. 2009.
- [32] X. Yao, "Unpacking and understanding evolutionary algorithms," in *Advances in Computational Intelligence*, J. Liu, C. Alippi, B. Bouchon-Meunier, G. W. Greenwood, and H. A. Abbass, editors. Springer-Verlag, 2012, pp. 60–76.
- [33] S. Chen, X. X. Wang and D. J. Brown, "Orthogonal least squares regression with tunable kernels," *Electron. Lett.*, vol. 41, no. 8, pp. 484–486, 2005.
- [34] X. P. Zong, Y. Xu, L. Hao, and X. L. Huai, "Camera calibration based on the RBF neural network with tunable nodes for visual servoing in robotics," in *Proc. 2006 IEEE/RSS Int. Conf. Intelligent Robots and Systems*, pp. 5708–5712.
- [35] S. Chen, X. X. Wang, X. Hong and C. J. Harris, "Kernel classifier construction using orthogonal forward selection and boosting with Fisher ratio class separability measure," *IEEE Trans. Neural Networks*, vol. 17, no. 6, pp. 1652–1656, June 2006.
- [36] X. X. Wang, S. Chen, D. Lowe, and C. J. Harris, "Sparse support vector regression based on orthogonal forward selection for the generalised kernel model," *Neurocomputing*, vol. 70, nos. 1-3, pp. 462–474, 2006.
- [37] J. X. Yu, H. Q. Cao, and Y. B. He, "A new tree structure code for equivalent circuit and evolutionary estimation of parameters," *Chemometrics and Intelligent Laboratory Systems*, vol. 85, no. 1, pp. 27–39, 2007.
- [38] M. Zhang, J. G. Zhou, L. H. Fu, and T. T. He, "Hybrid wavelet model construction using orthogonal forward selection with boosting search," in *Proc. 2007 Int. Conf. Fuzzy Systems and Knowledge Discovery*, pp. 341–345.
- [39] S. Chen, X. X. Wang, and C. J. Harris, "NARX-based nonlinear system identification using orthogonal least squares basis hunting," *IEEE Trans. Control Systems Technol.*, vol. 16, no. 1, pp. 78–84, 2008.
- [40] M. Abuthinien, S. Chen, and L. Hanzo, "Semi-blind joint maximum likelihood channel estimation and data detection for MIMO systems," *IEEE Signal Process. Lett.*, vol. 15, pp. 202–205, 2008.
- [41] M. Zhang, L. Fu, G. Wang and T. He, "Improved orthogonal least-squares regression with tunable kernels using a tree structure search algorithm," *IEEE Signal Process. Lett.*, vol. 15, pp. 653–656, 2008.
- [42] S. A. Hoseini, R. P. Torghabeh, M. Kaveh, and H. Khaloozadeh, "Designing stabilizing regulators for chaotic systems using repeated weighted boosting search method," in: *Proc. 2009 Int. Conf. Computer, Control and Communication*, pp. 1–6.
- [43] M. Zhang, J. Zhou, L. Fu, and T. He, "Hybrid wavelet model construction using orthogonal forward selection with boosting search," *Int. J. Business Intelligence and Data Mining*, vol. 3, no. 4, pp. 437–450, 2009.
- [44] S. Chen, X. Hong, B. L. Luk, and C. J. Harris, "Construction of tunable radial basis function networks using orthogonal forward selection," *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 39, no. 2, pp. 457–466, 2009.
- [45] J. Jiang and X.-P. Zhang, "A new player-enabled rapid video navigation method using temporal quantization and repeated weighted boosting search," in *Proc. 2010 IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops*, pp. 64–71.
- [46] L. H. Fu, M. Zhang, and H. W. Li, "Sparse RBF networks with multi-kernels," *Neural Process. Lett.*, vol. 32, no. 3, pp. 235–247, 2010.
- [47] F. Schoen, "Stochastic techniques for global optimization: a survey of recent advances," *J. Global Optimization*, vol. 1, no. 3, pp. 207–228, 1991.

- [48] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
- [49] H. Li, R. Wang, and K. Amleh, "Blind code-timing estimation for CDMA systems with bandlimited chip waveforms in multipath fading channels," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 141–149, Jan. 2006.



Shuai Wang received his B.Eng degree and Ph.D. degree from Zhengzhou University and Beijing Institute of Technology, P. R. China, in 2005 and 2012 respectively, both in Communications Engineering. From 2010 to 2011 he was a visiting Ph.D. student in the School of Electronics and Computer Science, University of Southampton, UK, under the financial support of China Scholarship Council. He has been with the School of Information Science and Electronics, Beijing Institute of Technology since July 2012, where he now serves as a lecturer. His research

interest lies in the application of general signal processing theory to wireless communications systems with special emphasis on synchronization, channel estimation, and anti-jammer techniques.



Sheng Chen (M'90-SM'97-F'08) received his B.Eng degree from the East China Petroleum Institute, China, in January 1982, and his Ph.D degree from the City University, London, in September 1986, both in control engineering. In 2005, he was awarded the higher doctorate degree, Doctor of Sciences (DSc), 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, where he currently holds the post of Professor in Intelligent Systems and Signal Processing. Dr Chen's 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 optimisation. He has published over 460 research papers. Dr. Chen is a Fellow of IET. He is a Distinguished Adjunct Professor at the King Abdulaziz University, Jeddah, Saudi Arabia. He is an ISI highly cited researcher in the engineering category (March 2004).



Aihua Wang (M'08) received her B.Eng degree from Hebei University, Baoding, P. R. China in 1986, M.S.c degree from Xidian University, Xi'an, P. R. China in 1993, and Ph.D. degree from Beijing Institute of Technology, P. R. China in 2001, respectively, all in Communications and Information Systems. She now holds the post as an Associate Professor in the School of Information Science and Electronics, Beijing Institute of Technology. Dr. Wang has published more than 40 journal and conference papers. She also holds (or co-holds) 16

patents. Her main research interest lies in channel estimation synchronization techniques in high-dynamic environments.



Jianping An (M'08) received his B.Eng degree from Information Engineering University of the People's Liberation Army, Zhengzhou, P. R. China in 1987, M.S.c degree from Beijing Institute of Technology, P. R. China in 1992 and Ph.D degree from Beijing Institute of Technology in 1996, all in Communications Engineering. Since 1996 He has been with the School of Information Science and Electronics, Beijing Institute of Technology, where he now holds the post of Full Professor. He also serves as the Vice Dean of the School of Information

Science and Electronics as well as the Chairman of the Communications and Information Systems Department. From 2010 to 2011 He was an Visiting Professor at University of California, San Diego. The research activities of his group in Beijing Institute of Technology are supported by the General Programs and Special Funds for Scientific Equipments of National Scientific Foundation of China, the National High Technology Research and Development Programs of China (A.K.A. "863" Programs), and the Chinese National Programs for Fundamental Research and Development (A.K.A. "973" Programs). He has published more than 150 journal and conference articles and holds (or co-holds) 37 patents. He has received various awards for his academic achievements and the resultant industrial influences, including the National Award for Scientific and Technological Progress of China (1997) and the Excellent Young Teacher Award by the China's Ministry of Education (2000). Since 2010 he has been serving as a Chief Reviewing Expert for the Information Technology Division, National Scientific Foundation of China. Prof. An's current research interest is focused on digital signal processing theory and algorithms for communication systems.



Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded the honorary doctorate "Doctor Honoris 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 successfully supervised 80 PhD students, co-authored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1300 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 a 100-strong 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. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. His research is funded by the European Research Council's Senior Research Fellow Grant. For further information on research in progress and associated publications please refer to <http://www-mobile.ecs.soton.ac.uk>