# Two-Tier Channel Estimation Aided Near-Capacity MIMO Transceivers Relying on Norm-Based Joint Transmit and Receive Antenna Selection

Peichang Zhang, Sheng Chen, *Fellow, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—We propose a norm-based joint transmit and receive antenna selection (NBJTRAS) aided near-capacity multiple-input–multiple-output (MIMO) system relying on the assistance of a novel two-tier channel estimation scheme. Specifically, a rough estimate of the full MIMO channel is first generated using a low-complexity, low-training-overhead minimum mean square error based channel estimator, which relies on reusing a modest number of radio frequency (RF) chains. NBJTRAS is then carried out based on this initial full MIMO channel estimate. The NBJTRAS aided MIMO system is capable of significantly outperforming conventional MIMO systems equipped with the same modest number of RF chains while dispensing with the idealized simplifying assumption of having perfectly known channel state information (CSI). Moreover, the initial subset channel estimate associated with the selected subset MIMO channel matrix is then used for activating a powerful semi-blind joint channel estimation and turbo detector–decoder, in which the channel estimate is refined by a novel block-of-bits selection based soft-decision aided channel estimator (BBSB-SDACE) embedded in the iterative detection and decoding process. The joint channel estimation and turbo detection–decoding scheme operating with the aid of the proposed BBSB-SDACE channel estimator is capable of approaching the performance of the near-capacity maximum-likelihood (ML) turbo transceiver associated with perfect CSI. This is achieved without increasing the complexity of the ML turbo detection and decoding process.**

***Index Terms*—Multiple-input–multiple-output (MIMO), near-capacity system, joint transmit and receive antenna selection, joint channel estimation and three-stage turbo detection–decoding.**

## I. Introduction

In recent years, multiple-input–multiple-output (MIMO) systems have attracted significant attention owing to their capability of increasing the reliability and/or bandwidth efficiency of communication systems [1], [2]. However, since MIMO systems utilize multiple radio frequency (RF) chains, their power consumption and hardware costs are substantial. Moreover, for massive MIMO systems and particularly for millimeter-wave based MIMO systems [2]–[4], the number of available antenna array elements increases massively [5], [6], while in practice the number of available RF chains is typically limited. As a remedy, antenna selection (AS) offers a low-cost, low-complexity technique of reducing the number of RF chains utilized at the transmitter and/or receiver, while retaining the significant advantages of MIMO systems. In AS aided MIMO systems, an additional number of antenna elements are employed to allow AS techniques to select a subset of antennas associated with the optimal or near-optimal channel condition (e.g., a subset associated with the highest equivalent signal to noise ratio (SNR)) from the whole antenna set to form the actual MIMO communication system, which therefore provides significant performance gain to MIMO systems. Generally, AS can be classified into three categories, namely transmit AS (TxAS), receive AS (RxAS), and joint transmit/receive AS (JTRAS) [7].

### A. Review of Antenna Selection Techniques

TxAS schemes conceived for MIMO systems have been studied in [8] and [9]. In particular, three AS criteria (ASCs) were proposed for space shift keying (SSK) systems in [8], which were the max-norm based AS (ASC1), maximum norm difference based AS (ASC2) and the hybrid scheme combining ASC1 and ASC2. The simulation results of [8] showed that AS techniques are capable of improving the performance of SSK aided MIMO systems and that interestingly, ASC1 outperformed both ASC2 and the hybrid design. Additionally, a pair of TxAS techniques were proposed in [9] for spatial modulation systems, where it was shown that the proposed capacity-optimized AS scheme outperformed the Euclidean distance optimized AS scheme. Similarly, RxAS schemes designed for MIMO systems have been investigated in [10], [11]. More specifically, the authors of [10] proposed an optimal RxAS scheme for space–time trellis codes, which selected the receive antennas (RAs) having the highest instantaneous SNR. Additionally, a RxAS scheme was proposed for vertical Bell Laboratories layered space–time MIMO systems in [11], where it was demonstrated that the system's performance was improved with the aid of AS in terms of the achievable block error rate.

As a hybrid version of TxAS and RxAS, JTRAS schemes were investigated in [12]–[15], where it was observed that MIMO systems employing JTRAS were capable of improving

Manuscript received December 21, 2013; revised April 23, 2014 and June 24, 2014; accepted June 25, 2014. Date of publication July 11, 2014; date of current version January 7, 2015. This work was supported in part by RC-UK under the auspices of the India–U.K. Advanced Technology Centre, by EU's Concerto project, and by the European Research Council's Advanced Fellow Grant. The associate editor coordinating the review of this paper and approving it for publication was M. McKay.

The authors are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: pz3g09@ecs.soton.ac.uk; sqc@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2014.2334325

the achievable system performance, while maintaining a low hardware complexity compared to the conventional MIMO systems employing the same number of RF chains and operating without JTRAS. Generally, there are two major types of AS algorithms—capacity based AS (CBAS) and norm-based AS (NBAS) [12]. The main idea of CBAS is to select those antennas that can maximize the system's MIMO channel capacity [1]. However, it is well-known that the optimal CBAS requires exhaustive search over all the possible subsets of the full channel matrix, which becomes impractical for systems having a large number of transmit antennas (TAs) and/or RAs [13]. Diverse sub-optimal CBAS techniques were developed in [12], [13], which were capable of reducing the AS complexity at the cost of a certain performance loss. As another efficient yet low-complexity category of AS algorithms, NBAS techniques are studied in [14], [15], which aim for selecting specific antennas that are capable of maximizing the system's SNR. It was shown that NBAS algorithms were capable of approaching the performance of CBAS techniques, while imposing a lower AS complexity. More specifically, a near-optimal JTRAS method was presented in [14], which first selects the RA having the maximum receive SNR and then selects some of the TAs that are related to the selected RA. A major limitation of this near-optimal AS method is that it is restricted to select a single RA. A JTRAS algorithm was proposed for two-hop amplify-and-forward relaying systems in [15], where only a single TA/RA pair is selected in each phase of relay communication.

It has been widely recognized [8]–[15] that AS techniques are capable of significantly improving the performance of coherently detected MIMO systems based on the assumption of perfectly known channel state information (CSI), in comparison to the conventional MIMO systems equipped with the same number of RF chains. However, in practice, CSI has to be acquired. A standard channel estimation (CE) technique is the training based CE (TBCE), where pilot symbols are used for acquiring an estimated CSI prior to actual data transmission. An analytical framework that enables the evaluation of the performance of multiple-branch diversity systems with the aid of TBCE was developed in [16], where the TBCE scheme was shown to be capable of preserving the diversity order of a MIMO system at the cost of a SNR penalty. The conventional training-based minimum mean square error (MMSE) channel estimator was employed in [17] for RxAS aided space–time coded MIMO systems communicating over Rayleigh flat fading channels, which however only considered selecting a single RA. The performance of the training-based MMSE channel estimator was investigated in [18] for employment in orthogonal frequency-division multiplexing based MIMO systems using RxAS, where AS was simply performed based on the received signal power quantified prior to CE. However, the conventional TBCE schemes adopted in [17], [18] are capable of generating accurate MIMO CSI only at the cost of imposing a potentially excessive pilot-overhead, which not only significantly erodes the system's throughput but also results in an excessive CE complexity. Additionally, it has been shown in [19], [20] that for AS aided MIMO systems, AS requires a less accurate CSI, while data detection must rely on a very accurate channel estimate. According to this observation, an efficient CE method was proposed for RxAS in [19], where just-sufficient training pilots are firstly sent for RxAS and then extra pilots are sent for further refining the channel estimate associated with the selected antennas. Similarly, a dual pilot-based training scheme was proposed in [20] for an AS aided multi-user orthogonal frequency-division multiple access system, where an uplink user firstly transmits a reference signal to the base station for acquiring the CSI for AS as well as for frequency-domain transmission scheduling. Then the uplink user sends a second reference signal for further refining the channel estimate for supporting data detection. Both these two novel schemes are capable of improving the attainable system performance at the cost of transmitting extra pilots. Therefore, the challenge here is also the acquisition of accurate MIMO CSI without imposing an excessive training overhead.

### *B. Review of Near-Capacity MIMO Systems*

Under the idealized conditions of perfectly known CSI, the three-stage turbo detection and decoding structure of [1] is capable of attaining a near-capacity MIMO performance, while imposing a low detection–decoding latency. The challenge of approaching the optimal near-capacity MIMO performance again is the acquisition of accurate MIMO CSI without imposing an excessive pilot-overhead and an excessive channel estimation complexity [21]. The existing state-of-the-art solutions [22]–[31] combine the decision-directed (DD) CE (DDCE) solutions with powerful iterative detection and decoding schemes in order to form semi-blind joint CE and turbo detection–decoding, where only a modest training overhead is required for generating an initial MMSE CE or least squares CE. The most effective schemes [26]–[31] employ soft-decision aided CEs in the semi-blind joint CE and turbo detection–decoding process, which are more robust against error propagation than the hard-decision aided CE schemes. Consequently, these joint soft-decision based CE and turbo detection–decoding schemes are capable of achieving a better overall system performance than their hard-decision based counterparts.

However, even the best existing semi-blind iterative soft-decision aided CE and turbo detection–decoding structures fail to approach the optimal maximum-likelihood (ML) turbo detection–decoding performance bound associated with perfect CSI. The reason is simply because given a small training overhead and a low SNR, a large percentage of the detected bits is likely to be erroneous and hence the error propagation is typically serious even for soft-decision aided CE schemes. Moreover, in order to benefit from the iterative gain of turbo detection–decoding, the DDCE in these structures takes place after the convergence of the turbo detection–decoding process, and this introduces an extra iterative loop between the DDCE scheme and the turbo detector–decoder, which significantly increases the complexity imposed. Furthermore, these existing schemes rely on the entire frame of the detected soft or hard bits for CE, and the complexity of the associated DDCE may become unacceptably high. This is because the number of bits in a single interleaved frame of a turbo code is very large, and typically thousands of bits are contained in a turbo coded frame.

Therefore, in order for these existing semi-blind joint CE and turbo detection–decoding arrangements to approach the ML performance bound associated with perfect CSI, it is necessary to employ a substantial training-overhead, which dramatically erodes the system's effective throughput.

The lack of accurate and efficient CE for MIMO CSI has long been the stumbling-block of near-capacity operation. Recently, however, an effective solution has been found for near-capacity three-stage-concatenated MIMO turbo-transceivers operating without AS [32]. More specifically, a novel block-of-bits selection based soft-decision aided channel estimator (BBSB-SDACE) was proposed in [32], which selects the high-quality reliable detected symbols or blocks of bits based on the *a posteriori* information produced by the MIMO soft-demapper within the original inner turbo loop of the unity-rate-code (URC) decoder and MIMO detector. Since this BBSB-SDACE only utilizes a "just-sufficient-number" of the high-quality detected symbols for CE, it does not suffer from the usual performance degradation imposed by erroneous decisions. Furthermore, this measure dramatically reduces the complexity of the DDCE. Additionally, as a benefit of selecting only the reliably detected symbols for CE, it is no longer necessary to wait for the convergence of the three-stage turbo detection–decoding process, before the DDCE scheme can be activated. Consequently, the CE is naturally embedded in the original three-stage turbo detection–decoding process. Therefore no extra iterative loop is required between the CE scheme and the three-stage MIMO detector–decoder. Hence, the complexity of this joint BBSB-SDACE and three-stage turbo detector–decoder remains similar to that of the idealized three-stage turbo receiver relying on perfect CSI. Most importantly, as demonstrated in [32], this semi-blind joint BBSB-SDACE and three-stage turbo detection–decoding structure is capable of approaching the optimal ML performance bound of the idealized three-stage turbo detector–decoder associated with perfect CSI, while only imposing the same number of turbo iterations as the latter.

### C. Our Novel Contributions

Against the above background, we propose a two-tier CE (TTCE) assisted and norm-based JTRAS (NBJTRAS) aided near-capacity MIMO system in this paper. Our novel contributions are as follows.

1) Firstly, we propose a new NBJTRAS aided near-capacity MIMO system, which achieves low system complexity by employing only a small number of RF chains. The system is capable of significantly outperforming the non-AS aided conventional MIMO systems utilizing the same number of RF chains, both in terms of its bit error ratio (BER) and throughput. Additionally, the proposed NBJTRAS aided MIMO system is capable of achieving extra diversity gains over that of the conventional MIMO system relying on the same number of RF chains and operating without AS, albeit this gain is achieved at the cost of employing more antenna elements than the latter.
2) Secondly, we propose a novel TTCE scheme relying on a low training overhead for assisting the NBJTRAS aided MIMO system to approach the capacity-optimal MIMO performance bound associated with perfect CSI, which maintains a high system throughput, while imposing a low computational complexity.

   To be more explicit, in tier one of the proposed TTCE scheme, a low-complexity low-pilot-overhead MMSE based CE scheme relying on RF chain reuse generates a coarse initial estimate of the full MIMO channel matrix using only a low number of training symbol blocks. Then NBJTRAS is carried out based on this rough CE, and the selected TAs and RAs are activated for actual data transmission.

   In tier two of the proposed scheme, a powerful semi-blind BBSB-SDACE and three-stage turbo detection–decoding structure is employed, which is similar to the one proposed in [32]. This scheme relies on the selected subset channel matrix, obtained in the tier-one NBJTRAS stage, as the initial MIMO CE for activating the turbo detection and decoding process invoked for detecting the data as well as for refining the CE.

   Our simulation results demonstrate that the optimal ML performance bound associated with perfect CSI can be approached with the aid of the proposed TTCE scheme, without increasing the complexity of the ML turbo detection and decoding process, while maintaining a high system throughput. Additionally, we show that the proposed AS scheme is capable of improving the accuracy of CE.

The rest of this contribution is organized as follows. Section II describes the three-stage turbo MIMO transceiver system model employed and presents the proposed NBJTRAS aided MIMO scheme, while our TTCE scheme conceived for the NBJTRAS aided MIMO system is detailed in Section III. The achievable performance of the proposed TTCE assisted and NBJTRAS aided MIMO system is investigated in Section IV, while our conclusions are provided in Section V. Finally, the Appendix defines the abbreviations adopted.

## II. System Overview

The following notational conventions are adopted throughout our discussions. Boldface capital and lower-case letters stand for matrices and column vectors, respectively, while $\mathbb{C}$ denotes the field of complex numbers. The inverse operation is denoted by $(\ )^{-1}$, while $(\ )^{\mathrm{T}}$ and $(\ )^{\mathrm{H}}$ represent the transpose and conjugate transpose operators. Furthermore, $\|\ \|$ and $|\ |$ denote the norm and magnitude operators, respectively. Additionally, $\mathrm{E}\{\ \}$ denotes the expectation operator and $\det\{\ \}$ is the matrix determinant operator, while $\Pr\{\ \}$ represents the probability. Finally, the $(M \times M)$-element identity matrix is denoted by $\boldsymbol{I}_M$, and $\boldsymbol{H}\langle i, j\rangle$ is the $i$th-row and $j$th-column element of $\boldsymbol{H}$.

### A. Three-Stage Turbo MIMO Transceiver

The structure of the three-stage turbo encoder employed at the transmitter is shown in Fig. 1(a), where the two-stage inner encoder is formed by an $L$-level quadrature amplitude modulation (QAM) MIMO transmitter combined with a URC

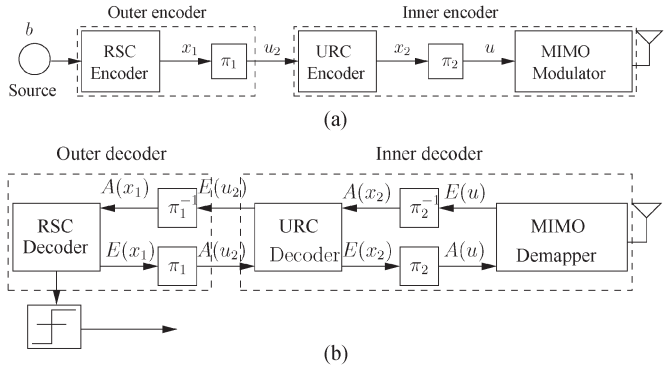


Fig. 1. Three-stage serial-concatenated turbo transmitter and receiver. (a) Three-stage turbo encoder. (b) Three-stage turbo decoder.

encoder, while a half-rate recursive systematic code (RSC) is employed as the outer encoder. In this way, a three-stage serial-concatenated RSC-URC-modulator scheme is created as the transmitter. The explicit benefit of incorporating a low-complexity memory-1 URC is that it has an infinite-duration impulse response (IIR), which allows the system to beneficially spread the *extrinsic* information across the iterative decoder components without increasing its delay [1], [33]–[35]. As a result, a vanishingly low BER is attainable [1], [34], [35].

The schematic of the corresponding three-stage turbo decoder adopted at the MIMO receiver is portrayed in Fig. 1(b), which consists of an $L$-QAM MIMO detector, a URC decoder and a RSC decoder. More explicitly, the composite inner decoder is formed by the combined MIMO detector and the URC decoder, where the associated *a priori* information and *extrinsic* information are firstly interleaved and exchanged $I_{\rm in}$ times. The outer decoder is constituted by the RSC decoder, where the information gleaned from the inner decoder is iteratively exchanged $I_{\rm out}$ times. As a benefit of the URC decoder's IIR structure, the extrinsic information transfer (EXIT) curve is capable of reaching the (1.0, 1.0) point of perfect convergence in the EXIT chart, implying that no error floor occurs, which is a necessary condition for near-capacity operation and for achieving a vanishingly low BER [1], [34], [35].

### B. Full MIMO System

Consider the MIMO system employing $N_T$ TAs and $N_R$ RAs as well as $L_T$ transmit and $L_R$ receive RF chains, whilst adopting $L$-QAM signaling. If the hardware resources are affordable and we have $L_T = N_T$ and $L_R = N_R$, a full $(N_R \times N_T)$-element MIMO system can be realized as:

$$\boldsymbol{y}(i) = \boldsymbol{H}\boldsymbol{s}(i) + \boldsymbol{v}(i), \tag{1}$$

where $i$ denotes the symbol index, $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$ is the full MIMO channel matrix whose elements obey the complex-valued zero-mean Gaussian distribution $\mathcal{CN}(0,1)$ with a variance of 1/2 per dimension, $\boldsymbol{s}(i) \in \mathbb{C}^{N_T}$ is the transmitted $L$-QAM symbol vector, and $\boldsymbol{y}(i) \in \mathbb{C}^{N_R}$ is the received signal vector, while $\boldsymbol{v}(i) \in \mathbb{C}^{N_R}$ is the noise vector whose elements obey $\mathcal{CN}(0, N_{\rm o})$ with a variance of $N_{\rm o}/2$ per dimension. The number of bits per $L$-QAM symbol (BPS) is $\text{BPS} = \log_2(L)$. The system's SNR is defined as $\text{SNR} = E_{\rm s}/N_{\rm o}$, where $E_{\rm s}$ is the average symbol energy. By adopting the three-stage turbo transceiver structure of Section II-A, this MIMO scheme is capable of achieving a near-capacity performance [1].

Let us further define the number of bits per block (BPB) as $\text{BPB} = N_T \cdot \text{BPS}$. At the receiver we extract the *a priori* log-likelihood ratios (LLRs) $\{L_a(u_k)\}_{k=1}^{\rm BPB}$ from the channel decoder, where $\{u_k\}_{k=1}^{\rm BPB}$ indicates the corresponding bits that are mapped to the symbol vector $\boldsymbol{s}(i)$. Then the *a posteriori* LLRs produced by the ML MIMO soft-demapper are expressed as [36]

$$L_p(u_k) = L_p(k) = \ln \frac{\sum\limits_{\boldsymbol{s}^n \in \{\boldsymbol{s}_{u_k=1}\}} \exp(p_n)}{\sum\limits_{\boldsymbol{s}^n \in \{\boldsymbol{s}_{u_k=0}\}} \exp(p_n)}, \tag{2}$$

where $\{\boldsymbol{s}_{u_k=1}\}$ and $\{\boldsymbol{s}_{u_k=0}\}$ represent the $L$-QAM symbol vector sets with the corresponding bits being $u_k = 1$ and $u_k = 0$, respectively. The probability metrics $\{p_n\}_{n=1}^{L^{N_T}}$ of the legitimate $L$-QAM symbol vectors $\{\boldsymbol{s}^n\}_{n=1}^{L^{N_T}}$ are given as

$$p_n = -\frac{\|\boldsymbol{y}(i) - \boldsymbol{H}\boldsymbol{s}^n\|^2}{N_0} + \sum_{k=1}^{\rm BPB} \tilde{u}_k L_a(u_k), \tag{3}$$

where $\{\tilde{u}_k\}_{k=1}^{\rm BPB}$ indicates the corresponding bits that are mapped to the specific symbol vector $\boldsymbol{s}^n$. For large-scale MIMO systems, we may opt for using reduced-complexity near-optimum detection schemes, such as the K-best sphere detector [37], [38], in order to avoid the exponentially increasing complexity imposed by the ML detector.

### C. Proposed NBJTRAS Aided MIMO System

The implementation of the above full MIMO system requires that the number of RF chains $L_T$ employed at the transmitter is equal to the number of TAs $N_T$ and the number of RF chains $L_R$ used at the receiver is equal to the number of RAs $N_R$. In practice, however, the number of affordable RF chains is often limited, and we have $L_T < N_T$ and $L_R < N_R$, particularly for large-scale MIMO systems. For a MIMO system of $L_T < N_T$ and $L_R < N_R$, the full MIMO system (1) is "virtual," i.e., the full channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$ is "virtual," since the communications only occur over an $(L_R \times L_T)$-element subset channel matrix $\boldsymbol{H}_{sub} \in \mathbb{C}^{L_R \times L_T}$. The conventional MIMO system operating without the aid of AS refers to the MIMO system that only employs $L_T$ TAs and $L_R$ RAs. In the generic case of $L_T < N_T$ and $L_R < N_R$, in order to efficiently utilize the available hardware resources, it is desirable to choose the most appropriate $L_T$ TAs from the full set of $N_T$ TAs and the most appropriate $L_R$ RAs from the full set of $N_R$ RAs to form a desired $(L_R \times L_T)$-element MIMO channel for actual data communications. Our proposed NBJTRAS aided MIMO system is depicted in Fig. 2, where we assume for the time being that the "virtual" full channel matrix $\boldsymbol{H}$ is perfectly known. In this NBJTRAS aided MIMO system, our proposed NBJTRAS algorithm constructs a MIMO system where each block of $L_T$

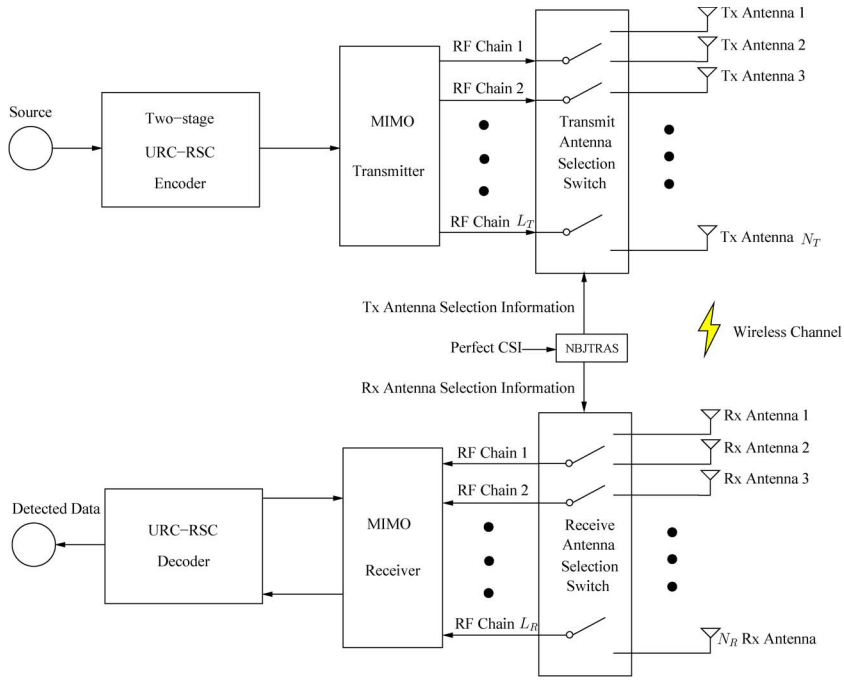


Fig. 2. Proposed NBJTRAS aided three-stage serial-concatenated turbo coded near-capacity MIMO systems.

symbols is transmitted over the activated subset channel matrix $\boldsymbol{H}_{sub} \in \mathbb{C}^{L_R \times L_T}$ with $\boldsymbol{H}_{sub} \subset \boldsymbol{H}$.

Generally speaking, increasing the channel gain is equivalent to reducing the effects of noise, which yields an improved system performance. This motivates our NBAS approach that selects the specific TAs and RAs related to the subset channel matrix having the highest channel norm. Let $\widetilde{\boldsymbol{H}}_{sub} \in \mathbb{C}^{L_R \times L_T}$ denote the subset candidates of the full channel matrix $\boldsymbol{H}$. The selected subset $\boldsymbol{H}_{sub}$ based on the NBAS criterion is found by solving the following optimization problem:

$$\begin{aligned} \boldsymbol{H}_{sub} &= \arg \max_{\widetilde{\boldsymbol{H}}_{sub} \subset \boldsymbol{H}} \|\widetilde{\boldsymbol{H}}_{sub}\|^2 \\ &= \arg \max_{\widetilde{\boldsymbol{H}}_{sub} \subset \boldsymbol{H}} \sum_{n_t=1}^{L_T} \sum_{n_r=1}^{L_R} \left| \widetilde{\boldsymbol{H}}_{sub}\langle n_r, n_t \rangle \right|^2. \end{aligned} \quad (4)$$

Solving the above optimization problem by exhaustive search requires us to evaluate the norms of the $\mathsf{C}_{N_R}^{L_R} \times \mathsf{C}_{N_T}^{L_T}$ candidate subset matrices, where $\mathsf{C}_k^n = k!/n!(k-n)!$, $\mathsf{C}_{N_R}^{L_R}$ and $\mathsf{C}_{N_T}^{L_T}$ are the row-dimension and column-dimension combinations of $\boldsymbol{H}_{sub}$, respectively. This will impose an extremely high computational complexity, particularly for large-scale MIMO systems associated with high $N_T$ and/or $N_R$. We propose a novel NBJTRAS scheme for solving the optimization problem (4) at a substantially reduced complexity. We now detail our NBJTRAS scheme.

*NBJTRAS Algorithm:* Given the full channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$, without loss of generality, we assume $\mathsf{C}_{N_R}^{L_R} \leq \mathsf{C}_{N_T}^{L_T}$. The algorithm accomplishes the optimization (4) in the following two steps.

*Step 1)—Row Dimension Operations:* Let $i_r \in \{1, 2, \ldots, \mathsf{C}_{N_R}^{L_R}\}$ be the row combination index and let us denote the row indices corresponding to the $i_r$th sub-matrix $\boldsymbol{H}_{i_r} \in \mathbb{C}^{L_R \times N_T}$ by $\boldsymbol{l}_{i_r} = [l_{i_r}^1 l_{i_r}^2 \ldots l_{i_r}^{L_R}]^{\mathrm{T}}$. Then we have

$$\boldsymbol{H}_{i_r} = \begin{bmatrix} \boldsymbol{h}_{l_{i_r}^1}^{\mathrm{T}} \\ \boldsymbol{h}_{l_{i_r}^2}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{h}_{l_{i_r}^{L_R}}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_{i_r}\langle 1,1 \rangle & \cdots & \boldsymbol{H}_{i_r}\langle 1, N_T \rangle \\ \boldsymbol{H}_{i_r}\langle 2,1 \rangle & \cdots & \boldsymbol{H}_{i_r}\langle 2, N_T \rangle \\ \vdots & \cdots & \vdots \\ \boldsymbol{H}_{i_r}\langle L_R,1 \rangle & \cdots & \boldsymbol{H}_{i_r}\langle L_R, N_T \rangle \end{bmatrix}, \quad (5)$$

where $\boldsymbol{h}_x^{\mathrm{T}}$ is the $x$th row of $\boldsymbol{H}$. The evaluation of

$$m_{i_r}^x = \sum_{j=1}^{L_R} |\boldsymbol{H}_{i_r}\langle j, x \rangle|^2, \quad 1 \leq x \leq N_T, \quad (6)$$

where $m_{i_r}^x$ represents the magnitude of the $x$th column in $\boldsymbol{H}_{i_r}$, yields the norm metric vector

$$\boldsymbol{m}_{i_r}^{\mathrm{T}} = \left[ m_{i_r}^1 \; m_{i_r}^2 \; \ldots \; m_{i_r}^{N_T} \right]. \quad (7)$$

Applying the operations of (6) and (7) to all the $\mathsf{C}_{N_R}^{L_R}$ possible combinations leads to the norm metric matrix $\boldsymbol{M} \in \mathbb{C}^{\mathsf{C}_{N_R}^{L_R} \times N_T}$ given by

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{m}_1^{\mathrm{T}} \\ \boldsymbol{m}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{m}_{\mathsf{C}_{N_R}^{L_R}}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} m_1^1 & m_1^2 & \cdots & m_1^{N_T} \\ m_2^1 & m_2^2 & \cdots & m_2^{N_T} \\ \vdots & \vdots & \cdots & \vdots \\ m_{\mathsf{C}_{N_R}^{L_R}}^1 & m_{\mathsf{C}_{N_R}^{L_R}}^2 & \cdots & m_{\mathsf{C}_{N_R}^{L_R}}^{N_T} \end{bmatrix}. \quad (8)$$

*Step 2)—Column Dimension Operations:* Find the largest $L_T$ elements in the $i_r$th row of $\boldsymbol{M}$ and sum them up, which is denoted as $m_{\max}^{i_r}$, as well as record the column indices of these $L_T$ elements in the index vector $\boldsymbol{l}_{i_c}(i_r) =$

$[l_{i_c}^1(i_r)\ l_{i_c}^2(i_r)\ \ldots\ l_{i_c}^{L_T}(i_r)]^{\mathrm{T}}$. This produces the max-norm metric vector

$$\boldsymbol{m}_{\max}^{\mathrm{T}} = \left[m_{\max}^1\ m_{\max}^2\ \ldots\ m_{\max}^{\mathsf{C}_{N_R}^{L_R}}\right]. \tag{9}$$

Next find

$$\bar{i}_r = \arg \max_{1 \le i_r \le \mathsf{C}_{N_R}^{L_R}} m_{\max}^{i_r}. \tag{10}$$

Then the selected TA and RA indices are specified by $\boldsymbol{l}_{i_c}(\bar{i}_r)$ and $\boldsymbol{l}_{\bar{i}_r}$, respectively, and the corresponding subset channel matrix $\boldsymbol{H}_{sub}$ is the optimal solution of (4).

*Complexity Analysis:* The complexity of the above NBJTRAS algorithm can be shown to be on the order of $C_{\mathrm{NBJTRAS}} \approx \mathcal{O}((N_T \cdot (L_R+1)+1) \cdot \mathsf{C}_{N_R}^{L_R})$, while that of the exhaustive search is given by $C_{\mathrm{ES}} \approx \mathcal{O}((L_R \cdot L_T) \cdot (\mathsf{C}_{N_T}^{L_T} \cdot \mathsf{C}_{N_R}^{L_R}))$. It can be clearly seen that $C_{\mathrm{NBJTRAS}}$ is significantly lower than $C_{\mathrm{ES}}$. If $\mathsf{C}_{N_R}^{L_R} > \mathsf{C}_{N_T}^{L_T}$, the NBJTRAS starts with **Step 1**) of the Column Dimension Operations followed by **Step 2**) of Row Dimension Operations, and the complexity of this algorithm is $C_{\mathrm{NBJTRAS}} \approx \mathcal{O}((N_R \cdot (L_T+1)+1) \cdot \mathsf{C}_{N_T}^{L_T})$.

*Additional Diversity Order Attained:* Given $L_R$ and $L_T$, the achievable multiplexing gain of the MIMO system is determined. We define the loading factor of AS as

$$f_{AS}(N_T, N_R) = \frac{N_T + N_R}{L_T + L_R}, \tag{11}$$

which determines the additional diversity order attainable by JTRAS schemes, such as our NBJTRAS algorithm, over the conventional MIMO system formed by employing $L_T$ TAs and $L_R$ RAs as well as operating without AS.

## III. Two-Tier Channel Estimation

In the previous section, we assumed having a perfect knowledge of the CSI for the NBJTRAS aided MIMO system. Let us now focus on eliminating this assumption. The TBCE schemes of [17], [18] may be preferred for estimating the MIMO CSI owing to their algorithmic simplicity, albeit they impose a substantial pilot overhead. Fortunately, it has been shown in [39] that the achievable diversity order obtained under the assumption of perfect CSI is not reduced, when an imperfect CE is used for AS. In other words, AS is relatively insensitive to the CE error. Consequently, a simple TBCE scheme relying on a small to modest pilot overhead may be sufficient for assisting our proposed NBJTRAS aided MIMO system. However, there are two critical issues which must be resolved first.

Note that the estimate of the full MIMO channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$ is required for AS, but we can only configure an $(L_R \times L_T)$-element MIMO physically. Therefore, a way must be found to estimate the "virtual" full MIMO channel matrix based on the limited affordable hardware resources. An attractive solution is to reuse the available RF chains for estimating all the $(L_R \times L_T)$-element subset MIMO channel matrices and consequently to form the estimate of this full MIMO channel matrix based on these estimated subset channel matrices. The resultant full MIMO estimate can then be adopted for NBJTRAS. Moreover, the selected subset MIMO CE can be employed for data detection. However, unlike the AS operation, data detection is sensitive to CE errors, and the coarse CE obtained by the TBCE scheme based on a small pilot overhead is far too inaccurate for the system to attain near-capacity operation. To obtain an accurate estimate of the MIMO CSI based on a TBCE scheme would impose a substantial pilot overhead and hence would erode the system's effective throughput quite considerably. A design alternative is to employ the powerful semi-blind BBSB-SDACE scheme of [32], which is capable of approaching the optimal MIMO performance bound associated with perfect CSI, without increasing the training overhead and the associated computational complexity. These two considerations motivate the design of our novel TTCE scheme for assisting the NBJTRAS aided MIMO system, which is illustrated in Fig. 3.

### A. Tier One: TBCE for Full Channel Matrix

In tier one CE, a low-complexity training based MMSE channel estimator is adopted for obtaining an initial CE. Since AS is relatively insensitive to the CE errors [39] and we may not need a high-accuracy CE, only a small number of training blocks is utilized for this TBCE scheme. In this way, a high effective system throughput is maintained. However, the available RF chains must be reused for the estimation of the "virtual" full channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$. For the sake of simplicity and without loss of generality, we assume that the ratios $N_T/L_T$ and $N_R/L_R$ are both integers. Then the number of the subset channel matrices that have to be estimated is $(N_T/L_T) \times (N_R/L_R)$. More specifically, we have to estimate the subset channel matrices $\boldsymbol{H}^{(i,j)} \in \mathbb{C}^{L_R \times L_T}$ for $i \in \{1, 2, \ldots, N_R/L_R\}$ and $j \in \{1, 2, \ldots, N_T/L_T\}$ in order to form the full MIMO channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$.

We assume that the number of the training blocks available is $M_T$ and the training data for estimating $\boldsymbol{H}^{(i,j)}$ are arranged as

$$\boldsymbol{Y}_{tM_T}^{(i,j)} = \left[\boldsymbol{y}^{(i,j)}(1) \quad \boldsymbol{y}^{(i,j)}(2) \quad \ldots \quad \boldsymbol{y}^{(i,j)}(M_T)\right], \tag{12}$$

$$\boldsymbol{S}_{tM_T}^{(i,j)} = \left[\boldsymbol{s}^{(i,j)}(1) \quad \boldsymbol{s}^{(i,j)}(2) \quad \ldots \quad \boldsymbol{s}^{(i,j)}(M_T)\right], \tag{13}$$

where $\boldsymbol{y}^{(i,j)}(q) \in \mathbb{C}^{L_R}$ is the received signal vector corresponding to the transmitted symbol vector $\boldsymbol{s}^{(i,j)}(q) \in \mathbb{C}^{L_T}$ for $1 \le q \le M_T$. The MMSE estimate of $\boldsymbol{H}^{(i,j)}$ based on the training data (12) and (13) is readily obtained as

$$\widehat{\boldsymbol{H}}^{(i,j)} = \boldsymbol{Y}_{tM_T}^{(i,j)} \left( \left(\boldsymbol{S}_{tM_T}^{(i,j)}\right)^{\mathrm{H}} \boldsymbol{S}_{tM_T}^{(i,j)} + N_{\mathrm{o}} \cdot \boldsymbol{I}_{M_T} \right)^{-1} \left(\boldsymbol{S}_{tM_T}^{(i,j)}\right)^{\mathrm{H}}, \tag{14}$$

and the estimate of the full channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$ can be formed according to

$$\widehat{\boldsymbol{H}} = \begin{bmatrix} \widehat{\boldsymbol{H}}^{(1,1)} & \widehat{\boldsymbol{H}}^{(1,2)} & \cdots & \widehat{\boldsymbol{H}}^{\left(1,\frac{N_T}{L_T}\right)} \\ \widehat{\boldsymbol{H}}^{(2,1)} & \widehat{\boldsymbol{H}}^{(2,2)} & \cdots & \widehat{\boldsymbol{H}}^{\left(2,\frac{N_T}{L_T}\right)} \\ \vdots & \vdots & \cdots & \vdots \\ \widehat{\boldsymbol{H}}^{\left(\frac{N_R}{L_R},1\right)} & \widehat{\boldsymbol{H}}^{\left(\frac{N_R}{L_R},2\right)} & \cdots & \widehat{\boldsymbol{H}}^{\left(\frac{N_R}{L_R},\frac{N_T}{L_T}\right)} \end{bmatrix}. \tag{15}$$

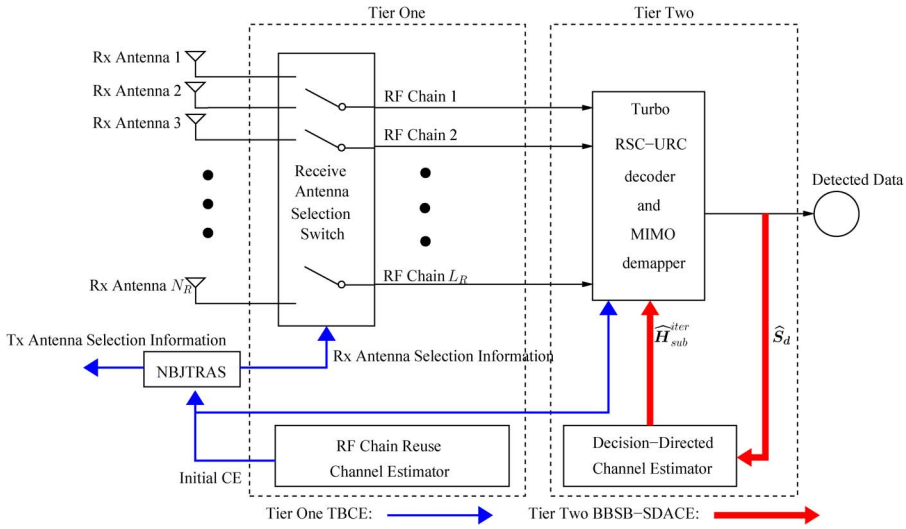


Fig. 3. Proposed two-tier channel estimation scheme for assisting the NBJTRAS aided MIMO system.

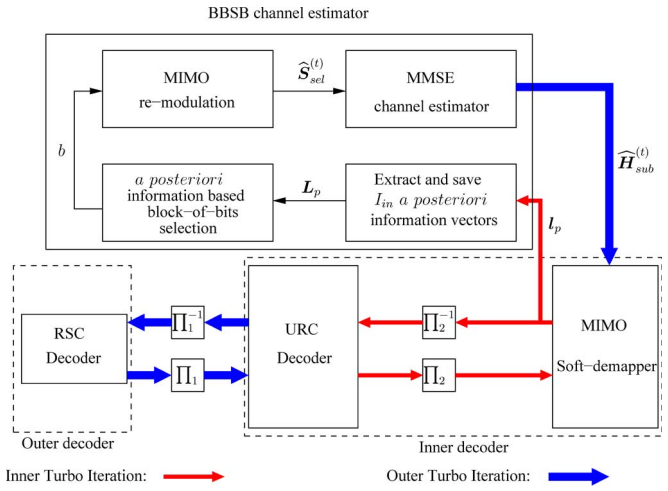


Fig. 4. Joint BBSB-SDACE and three-stage turbo detection–decoding structure, where The BBSB-SDACE is naturally embedded within the three-stage turbo detection–decoding process.

Then the NBJTRAS described in Section II-C is carried out based on this estimated full channel matrix $\widehat{\boldsymbol{H}} \in \mathbb{C}^{N_R \times N_T}$, which also yields the coarse estimate $\widehat{\boldsymbol{H}}_{sub} \in \mathbb{C}^{L_R \times L_T}$ of the subset channel matrix $\boldsymbol{H}_{sub} \in \mathbb{C}^{L_R \times L_T}$ for the specifically selected subset MIMO system over which the actual data transmission will take place.

### B. Tier Two: BBSB-SDACE

When using a short training length $M_T$, the accuracy of the MMSE estimate $\widehat{\boldsymbol{H}}_{sub}$ is poor. Recall that data detection is more sensitive to CE errors than the NBJTRAS. In tier two, we use a powerful iterative BBSB-SDACE scheme, similar to the one proposed in [32], for refining the initial TBCE $\widehat{\boldsymbol{H}}_{sub}$. This novel iterative BBSB-SDACE scheme is embedded naturally within the original three-stage turbo detection–decoding structure, as illustrated in Fig. 4. Note that there is no additional iterative loop involving the BBSB-SDACE block and the three-stage turbo detector–decoder. In other words, our soft decision aided CE is embedded in the original outer loop of the three-stage turbo detector–decoder process, and the CE update occurs concurrently with the original outer turbo decoding iteration. Moreover, our CE does not use the entire frame of the detected bits. Rather, it only selects the high-quality or reliable decisions. Specifically, the *a posteriori* information (2) output by the MIMO soft-demapper provides the confidence levels of binary 1 s and 0 s [1]. Therefore, based on this confidence level, we opt for using only the reliable decisions from the MIMO soft-demapper's output sequence for CE. Removing most of the erroneous decisions from the DDCE leads to a much more accurate CE, which in turn enhances the performance of the three-stage turbo detection–decoding process. Consequently, the joint BBSB-SDACE and three-stage turbo detector–decoder of Fig. 4 is capable of approaching the performance bound of the idealized three-stage turbo ML detector–decoder associated with perfect CSI [32], despite imposing only the same detection–decoding complexity as the idealized three-stage turbo detector–decoder associated with perfect CSI.

*BBSB-SDACE Scheme:* Let $M_F$ be the length of the observation data output by the MIMO demodulator, which is expressed as

$$\boldsymbol{Y}_{dM_F} = [\boldsymbol{y}(1) \quad \boldsymbol{y}(2) \quad \ldots \quad \boldsymbol{y}(M_F)], \tag{16}$$

where $\boldsymbol{y}(i) \in \mathbb{C}^{L_R}$ represents the received signal vector corresponding to the transmitted symbol vector $\boldsymbol{s}(i) \in \mathbb{C}^{L_T}$. Let us now detail our BBSB-SDACE scheme of Fig. 4.

Step 1) Set the outer turbo iteration index to $t = 0$ and the initial channel estimate to $\widehat{\boldsymbol{H}}_{sub}^{(t)} = \widehat{\boldsymbol{H}}_{sub}$.

Step 2) Given $\widehat{\boldsymbol{H}}_{sub}^{(t)}$, perform ML soft-demapping for the observation data $\boldsymbol{Y}_{\mathrm{d}M_F}$ of (16). The MIMO soft-demapper exchanges its soft information with the URC inner decoder for $I_{\mathrm{in}}$ iterations, yielding the $I_{\mathrm{in}}$ vectors of the *a posteriori* information as defined in (2), which can be arranged as seen in the following *a posteriori* information matrix

$$\boldsymbol{L}_p = [\boldsymbol{l}_p^1 \, \boldsymbol{l}_p^2 \, \ldots \, \boldsymbol{l}_p^{I_{\mathrm{in}}}]^{\mathrm{T}} \in \mathbb{C}^{I_{\mathrm{in}} \times L_F}, \tag{17}$$

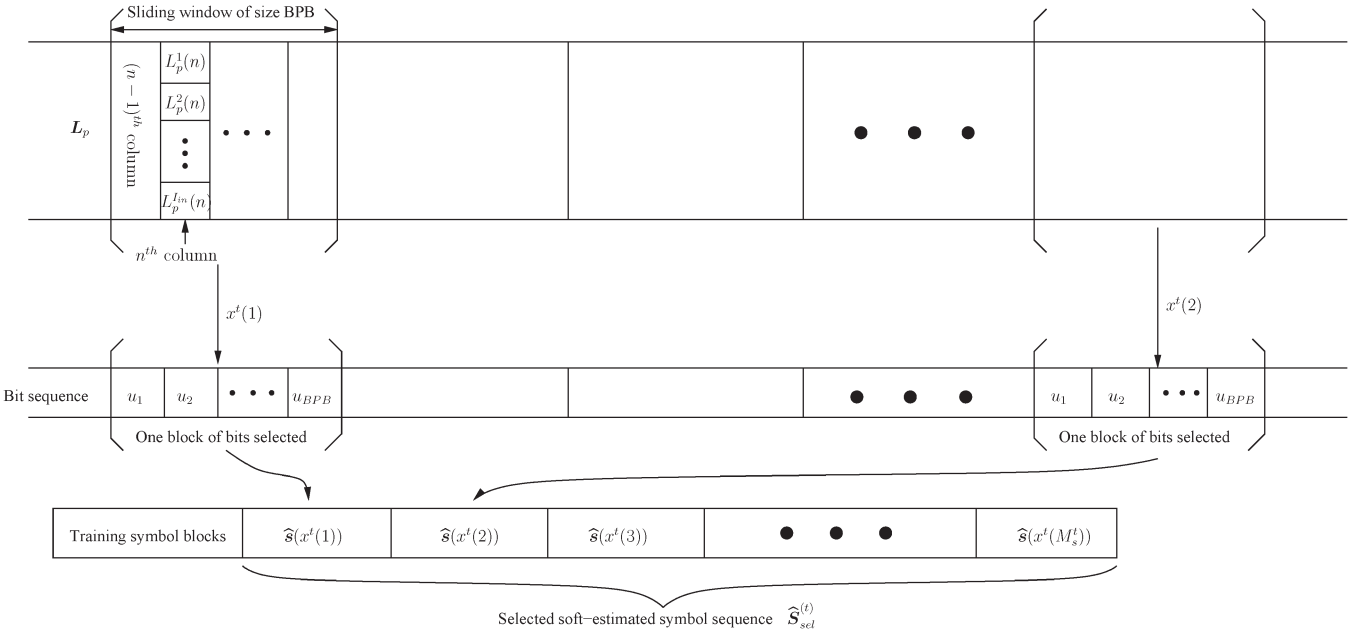


Fig. 5. Illustration of the sliding-window process using a window-size of BPB bits to select reliable detected symbol vectors.

where $L_F = \text{BPB} \cdot M_F$ is the total number of bits in a frame, $\boldsymbol{l}_p^i = [L_p^i(1)\ L_p^i(2)\ \ldots\ L_p^i(L_F)]^{\mathrm{T}} \in \mathbb{C}^{L_F}$ for $1 \le i \le I_{\text{in}}$ is the *a posteriori* information vector obtained during the $i$th inner iteration. The $n$th column of $\boldsymbol{L}_p$ contains the $I_{\text{in}}$ soft decisions $\{L_p^1(n), L_p^2(n), \ldots, L_p^{I_{\text{in}}}(n)\}$ for the $n$th information bit obtained in the $I_{\text{in}}$ inner decoder iterations, which can be exploited to judge whether the $n$th detected bit is reliable or not. Specifically, the $n$th detected bit is judged to be high quality when either of the following two criteria is met:

*Criterion 1:* If the soft decisions in the $n$th column of $\boldsymbol{L}_p$ share similar values, these soft decisions may result in a stable and reliable bit decision, which is hence considered to be correct. Specifically, the criterion for the $n$th detected bit to be judged as a correct one is

$$\frac{\left|L_p^1(n) - L_p^2(n)\right| + \cdots + \left|L_p^{I_{\text{in}}-1}(n) - L_p^{I_{\text{in}}}(n)\right|}{|\mu|} \in (0, T_h), \tag{18}$$

where $\mu$ is the mean of the soft decisions in the $n$th column of $\boldsymbol{L}_p$, which is employed to normalize the sum of the absolute differences between adjacent soft decisions (i.e., the differences between soft decisions of $L_p^1(n)$ and $L_p^2(n)$, $L_p^2(n)$ and $L_p^3(n)$, and so on) and therefore to reduce the effects of the varying soft decisions $\{L_p^1(n), L_p^2(n), \ldots, L_p^{I_{\text{in}}}(n)\}$ associated with different outer iterations, while $T_h$ denotes the pre-defined block-of-bits selection threshold. Note that $\{L_p^1(n), L_p^2(n), \ldots, L_p^{I_{\text{in}}}(n)\}$ are different for different outer iterations. Therefore, without the normalization factor $|\mu|$, a different $T_h$ value might be required for each outer iteration, which would complicate the selection process.

*Criterion 2:* If the absolute values of the soft decisions in the $n$th column of $\boldsymbol{L}_p$ are in monotonically ascending order and these soft decisions share the same sign, namely, we have,

$$\left|L_p^1(n)\right| < \left|L_p^2(n)\right| < \cdots < \left|L_p^{I_{\text{in}}}(n)\right| \text{ and}$$

$$\text{sign}\left\{L_p^1(n)\right\} = \text{sign}\left\{L_p^2(n)\right\} = \cdots = \text{sign}\left\{L_p^{I_{\text{in}}}(n)\right\},$$

then the $n$th detected bit may be regarded as a correct one.

By checking through the columns of $\boldsymbol{L}_p$, only high-confidence decision blocks are selected and the corresponding symbol block indices can be determined by a sliding-window based method using a window-size of BPB bits. More explicitly, the corresponding symbol vector is selected for CE only when the BPB consecutive detected bits of a block are all regarded as correct. This process yields an integer-index vector, denoted as $\boldsymbol{x}^t = [x^t(1)\ x^t(2)\ \ldots\ x^t(M_s^t)]^{\mathrm{T}}$ in which $x^t(i)$ is the position or index of the $i$th selected symbol vector in the transmitted symbol vector sequence. The number of the selected symbol vectors $M_s^t$ varies within $\{1, 2, \ldots, M_{\text{sel}}\}$, where $M_{\text{sel}} \ll M_F$ is the maximum number of blocks invoked for DDCE. Specifically, whenever the number of selected reliable symbol vectors $M_s^t$ reaches the limit $M_{\text{sel}}$, the sliding-window process ends; otherwise, the sliding-window process examines all the possible bit blocks and outputs the $M_s^t$ selected symbol vectors. Thus, $M_s^t$ varies at each outer turbo iteration $t$, and $M_s^t \le M_{\text{sel}}$. By using this index vector, the corresponding observation data can be selected from (16), and rearranged as

$$\boldsymbol{Y}_{\text{sel}}^{(t)} = \left[\boldsymbol{y}\left(x^t(1)\right)\ \boldsymbol{y}\left(x^t(2)\right)\ \ldots\ \boldsymbol{y}\left(x^t\left(M_s^t\right)\right)\right]. \tag{19}$$

This sliding-window process used for selecting reliable detected symbol vectors is illustrated in Fig. 5.

Step 3) Based on the selected high-confidence detected blocks of bits having the symbol vector indices $\boldsymbol{x}^t$,

generate the soft-estimate of each symbol element as [29]

$$\widehat{s}^{\,m}\left(x^t(n)\right) = \sum_{l=1}^{L} s^l \Pr\left\{s^m\left(x^t(n)\right) = s^l\right\} = \sum_{l=1}^{L} s^l \cdot \frac{\exp\left(\sum_{j=1}^{\mathrm{BPS}} \widetilde{u}_j L_a(u_j)\right)}{\prod_{j=1}^{\mathrm{BPS}} \left(1+\exp\left(L_a(u_j)\right)\right)}, \quad (20)$$

for $1 \le n \le M_s^t$, where $\{s^l\}_{l=1}^{L}$ denotes the $L$-QAM symbol set, $m \in \{1, 2, \ldots, L_T\}$ indicates the symbol index in the soft-estimated symbol vector $\widehat{\boldsymbol{s}}(x^t(n))$, and $\{\widetilde{u}_j\}_{j=1}^{\mathrm{BPS}}$ represents the bit mapping corresponding to $\{s^l\}_{l=1}^{L}$. By arranging the soft-estimated symbol vectors as

$$\widehat{\boldsymbol{S}}_{\mathrm{sel}}^{(t)} = \left[\widehat{\boldsymbol{s}}\left(x^t(1)\right)\ \widehat{\boldsymbol{s}}\left(x^t(2)\right)\ \ldots\ \widehat{\boldsymbol{s}}\left(x^t\left(M_s^t\right)\right)\right], \quad (21)$$

as well as defining $\boldsymbol{Y}_{\mathrm{t+sel}}^{(t)} = [\boldsymbol{Y}_{tM_T}\ \boldsymbol{Y}_{\mathrm{sel}}^{(t)}]$ and $\widehat{\boldsymbol{S}}_{\mathrm{t+sel}}^{(t)} = [\boldsymbol{S}_{tM_T}\ \widehat{\boldsymbol{S}}_{\mathrm{sel}}^{(t)}]$, the resultant soft decision-directed MMSE estimate is readily given by

$$\widehat{\boldsymbol{H}}^{(t+1)} = \boldsymbol{Y}_{\mathrm{t+sel}}^{(t)} \left(\left(\widehat{\boldsymbol{S}}_{\mathrm{t+sel}}^{(t)}\right)^{\mathrm{H}} \widehat{\boldsymbol{S}}_{\mathrm{t+sel}}^{(t)} + N_{\mathrm{o}} \boldsymbol{I}_{M_T+M_s^t}\right)^{-1} \times \left(\widehat{\boldsymbol{S}}_{\mathrm{t+sel}}^{(t)}\right)^{\mathrm{H}}, \quad (22)$$

where $\boldsymbol{Y}_{tM_T} \in \mathbb{C}^{L_R \times M_T}$ and $\boldsymbol{S}_{tM_T} \in \mathbb{C}^{L_T \times M_T}$ denote the tier-one training data used for generating the initial MMSE estimate $\widehat{\boldsymbol{H}}_{sub}$. This update occurs as the soft information is exchanged between the two-stage inner decoder and the outer RSC decoder, as indicated in Fig. 4.

Step 4) Set $t = t + 1$. If $t < I_{\mathrm{out}}$, repeat Steps 2) and 3); otherwise, stop.

*Remarks:* We first elaborate on the two criteria used for selecting high-quality bits. The idea behind *Criterion 1* is that if the decisions for the $n$th bit are relatively similar during the inner turbo iterations, the $n$th bit decision may be regarded as reliable. This makes sense because following a number of outer iterations, a stable state may be reached by the turbo decoder and hence the stable decisions of the inner decoder are likely to be the correct ones. Our experience detailed in [32] suggests that most of the chosen bit blocks or symbols are selected according to *Criterion 1*. As for *Criterion 2*, we note that if the absolute values of the decisions for a specific bit are in monotonically ascending order and these decisions share the same polarity, the corresponding bit decision is likely to be correct. This makes sense, because the correct decisions are likely to benefit from an iteration gain and this will lead to increasing the absolute values of the soft-decisions as the number of inner iterations increases. This type of reliable decisions may not always be spotted according to *Criterion 1* and hence *Criterion 2* allows us to select these high-quality decisions, when they do occur.

An important point to note is that our scheme fully exploits the information provided by the entire inner iterative turbo process, as manifested in the $n$th column of the *a posteriori* information matrix $\boldsymbol{L}_p$ in (17) which records the $I_{\mathrm{in}}$ soft decisions $\{L_p^1(n), L_p^2(n), \ldots, L_p^{I_{\mathrm{in}}}(n)\}$ for the $n$th information bit obtained in the $I_{\mathrm{in}}$ inner decoder iterations. The decision as to whether to select the $n$th bit or not is based on examining the entire inner turbo decoding-detecting process, rather than on an individual LLR value. Therefore, our scheme is capable of making a high-confidence decision regarding whether the $n$th detected bit is reliable or not.

Next, we discuss the role of the block-of-bits selection threshold $T_h$. In our proposed BBSB-SDACE algorithm, $T_h$ is invoked for carefully controlling the selection of reliable bits based on *Criterion 1*, which has a significant effect on the overall performance of the proposed scheme. Therefore, again the value of $T_h$ employed in **step 2**) for *Criterion 1* should be carefully chosen. To be more explicit, if too small a value of $T_h$ is chosen, the similarity between the adjacent soft decisions of a specific detected bit has to be very high to ensure that the left-hand-side value of the criterion (18) falls within the selection range of $(0, T_h)$. As a result, the number of the selected decision blocks may be insufficient for DDCE even after examining the entire sequence of $L_F$ bit decisions. By contrast, too large a value may result in the number of the selected blocks reaching the maximum limit value of $M_{\mathrm{sel}}$ after only examining a small initial portion of the $L_F$ bit decisions and hence the selected blocks may contain many "low confidence" decisions. Both of these two situations will result in a performance degradation. However, apart from these relatively extreme cases, our experience documented in [32] suggests that the performance of our semi-blind scheme is insensitive to the value of $T_h$. Specifically, there exists a relatively wide range of values for $T_h$, which allows our scheme to approach its optimal performance without increasing the number of turbo iterations. This range of optimal values for $T_h$ depends on both the modulation scheme and on the MIMO channel.

### C. Complexity of Proposed TTCE Assisted and NBJTRAS Aided MIMO

The total complexity $C_{\mathrm{total}}$ of the proposed TTCE assisted and NBJTRAS aided near-capacity MIMO scheme consists of three major parts, namely, the complexity $C_{\mathrm{t1ce}}$ of the tier-one TBCE, the complexity $C_{\mathrm{turbo}}$ of the three-stage turbo detection–decoding process, and finally the complexity $C_{\mathrm{t2ce}}$ of the tier-two soft DDCE that is embedded in the turbo detection–decoding process. Thus, we have

$$C_{\mathrm{total}} = C_{\mathrm{t1ce}} + C_{\mathrm{turbo}} + C_{\mathrm{t2ce}}. \quad (23)$$

In the tier-one TBCE, we have to carry out the MMSE CE process, as specified in (14) $((N_T/L_T) \times (N_R/L_R))$ times, and each MMSE CE imposes a complexity of the order of $\mathcal{O}(M_T^3)$. Therefore, we have

$$C_{\mathrm{t1ce}} = \left(\frac{N_T}{L_T} \times \frac{N_R}{L_R}\right) \cdot \mathcal{O}\left(M_T^3\right). \quad (24)$$

The complexity of the tier-one TBCE is kept low by using only a small number of training blocks, i.e., using a small $M_T$,

albeit this results in a reduced-accuracy full channel matrix estimate (15). Fortunately, this does not really matter, as AS is insensitive to CE errors. Moreover, in the tier-two stage, the BBSB-SDACE scheme will update the coarse CE $\widehat{\boldsymbol{H}}_{sub}$ selected by the NBJTRAS algorithm for the corresponding activated MIMO sub-system into an accurate CE in order to guarantee that the three-stage turbo detection–decoding process approaches the optimal performance bound associated with perfect CSI.

Let $C_{\mathrm{RSC}}$, $C_{\mathrm{URC}}$ and $C_{\mathrm{ML}}$ denote the complexity of the RSC decoder, the URC decoder, and the ML MIMO soft-demapper, respectively. Since the two-stage inner turbo loop requires $I_{\mathrm{in}}$ iterations and the outer turbo loop requires $I_{\mathrm{out}}$ iterations, the computational complexity of the idealized three-stage turbo receiver provided with perfect CSI can readily be expressed by

$$C_{\mathrm{ideal-turbo}} = I_{\mathrm{out}} \left( C_{\mathrm{RSC}} + I_{\mathrm{in}} (C_{\mathrm{ML}} + C_{\mathrm{URC}}) \right). \quad (25)$$

Since our TTCE assisted and NBJTRAS aided three-stage turbo MIMO receiver does not increase the number of turbo iterations required, its data detection–decoding process imposes the same complexity as the idealized three-stage turbo receiver. Hence, we have

$$C_{\mathrm{turbo}} = C_{\mathrm{ideal-turbo}}. \quad (26)$$

On the other hand, in the tier-two stage, the soft decision-directed MMSE estimate (22), which has a complexity lower than $\mathcal{O}((M_T + M_{\mathrm{sel}})^3)$, is calculated $I_{\mathrm{out}}$ times. Therefore, we have

$$C_{\mathrm{t2ce}} < I_{\mathrm{out}} \cdot \mathcal{O}\left( (M_T + M_{\mathrm{sel}})^3 \right). \quad (27)$$

It can be seen that although a turbo coded data frame is typically very long and hence $L_F$ is in the thousands, the complexity $C_{\mathrm{t2ce}}$ can be maintained at a reasonable level by setting a not-too-large limit $M_{\mathrm{sel}}$, say in the hundreds. Such a modest limit value is capable of ensuring for the BBSB-SDACE assisted three-stage turbo detector–decoder to approach the optimal performance bound of the idealized three-stage turbo detector–decoder associated with perfect CSI [32].

Since $C_{\mathrm{t2ce}}$ is low while $C_{\mathrm{t1ce}}$ is negligible compared to $C_{\mathrm{turbo}}$, the complexity of our TTCE assisted and NBJTRAS aided near-capacity MIMO scheme is only slightly higher than that of the idealized three-stage turbo detector–decoder, which has a perfect CSI.

## IV. Simulation Results and Discussions

A quasi-static independent Rayleigh fading environment was considered in our simulations. The generic MIMO system having $N_T$ TAs and $N_R$ RAs as well as employing $L_T$ transmit RF chains and $L_R$ receive RF chains, while adopting $L$-QAM signaling, was denoted by MIMO$(N_T, N_R; L_T, L_R; L-\mathrm{QAM})$. In order to keep both the hardware complexity and the power consumption at the same level, the number of RF chains employed, namely $L_T$ and $L_R$, were equal in both the NBJTRAS aided MIMO system and in the conventional MIMO system operating without AS, implying that the MIMO channel matrix activated for data communication had the same dimension of $\boldsymbol{H}_{sub} \in \mathbb{C}^{L_R \times L_T}$ for the both systems. More specifically, the conventional MIMO system operating without AS had the MIMO(4, 2; 4, 2; 4-QAM) arrangement, while the NBJTRAS aided MIMO system employed the MIMO$(N_T, N_R; 4, 2; 4-\mathrm{QAM})$ structure along with the AS loading factor of $f_{AS}(N_T, N_R)$. All the results were averaged over 100 channel realizations.

The generator polynomials of the half-rate RSC encoder were expressed in binary format as $G_{RSC} = [1, 0, 1]_2$ and $G^r_{RSC} = [1, 1, 1]_2$, while those of the URC encoder were $G_{URC} = [1, 0]_2$ and $G^r_{URC} = [1, 1]_2$, where $G^r_{RSC}$ and $G^r_{URC}$ denoted the feedback polynomials of the RSC and URC encoders, respectively. The number of inner iterations and outer iterations were set to $I_{\mathrm{in}} = 3$ and $I_{\mathrm{out}} = 5$. An interleaver length of 160,000 bits was used by the three-stage serial-concatenated turbo encoder-decoder of Fig. 1, which corresponded to $L_F = 20,000$ symbol blocks or vectors of $\boldsymbol{s}(i) \in \mathbb{C}^{L_T}$ for $1 \leq i \leq L_F$. The transmitted signal power was normalized to unity, and therefore the SNR was given as $1/N_{\mathrm{o}}$.

Three metrics were used for assessing the achievable performance, namely, the BER, the mean CE error (MCE) of the channel estimator, and the MIMO channel's capacity. The MCE is defined by

$$J_{\mathrm{MCE}}(\widehat{\boldsymbol{H}}_{sub}) = \frac{\|\boldsymbol{H}_{sub} - \widehat{\boldsymbol{H}}_{sub}\|^2}{\|\boldsymbol{H}_{sub}\|^2}, \quad (28)$$

where $\boldsymbol{H}_{sub} \in \mathbb{C}^{L_R \times L_T}$ denotes the true channel matrix of the activated MIMO system and $\widehat{\boldsymbol{H}}_{sub}$ its estimate. The MIMO channel's capacity is given by [1]

$$C_{\mathrm{MIMO}}(N_{\mathrm{o}}, L_T, L_R) = E\left\{ \log\left( \det\left\{ \boldsymbol{I}_{L_R} + \frac{N_{\mathrm{o}}}{L_T} \boldsymbol{H}_{sub} \boldsymbol{H}^{\mathrm{H}}_{sub} \right\} \right) \right\}, \quad (29)$$

where $N_{\mathrm{o}}$ is the channel's noise power and $\boldsymbol{H}_{sub} \in \mathbb{C}^{L_R \times L_T}$ denotes the corresponding MIMO channel matrix used for data communication.

### A. NBJTRAS Aided MIMO Systems With Perfectly Known CSI

Our investigations commenced with the EXIT chart analysis of the proposed NBJTRAS aided three-stage turbo MIMO$(N_T, N_R; 4, 2; 4-\mathrm{QAM})$ with the AS loading factor of $f_{AS}(8, 4) = 2$, in comparison to that of the conventional three-stage turbo MIMO(4, 2; 4, 2; 4-QAM) without AS, assuming for the time being that the CSI was perfectly known. It can be seen from the EXIT charts shown in Fig. 6 that for the proposed NBJTRAS aided MIMO system having an AS loading factor of 2, an open EXIT-tunnel exists between the EXIT curve of the amalgamated inner MIMO soft-demapper-URC decoder and the outer RSC decoder based on perfect CSI at SNR $=$ 1.4 dB. The actual Monte-Carlo simulation based stair-case shaped decoding trajectory, which closely matches the EXIT curves, is also shown at SNR $= 1.4$ dB for this NBJTRAS aided MIMO system. The trajectory shows that the point of perfect convergence at (1.0, 1.0) can be reached with the aid

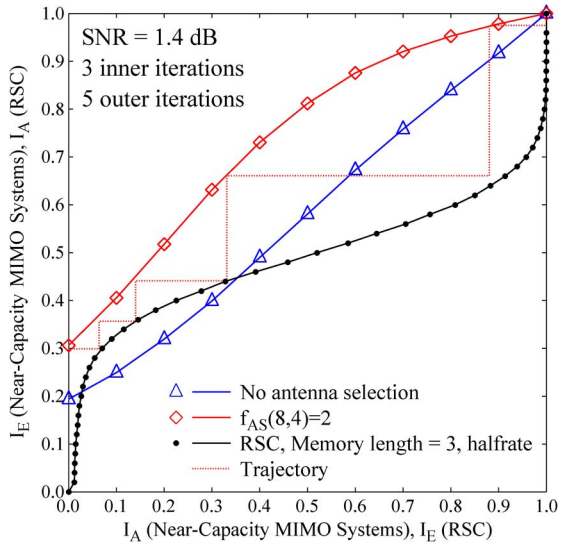


Fig. 6. EXIT chart analysis of our NBJTRAS aided MIMO$(N_T, N_R; 4, 2; 4\text{-QAM})$ with the AS loading factor of $f_{AS}(8, 4) = 2$, in comparison to that of the conventional MIMO$(4, 2; 4, 2; 4\text{-QAM})$ without AS.

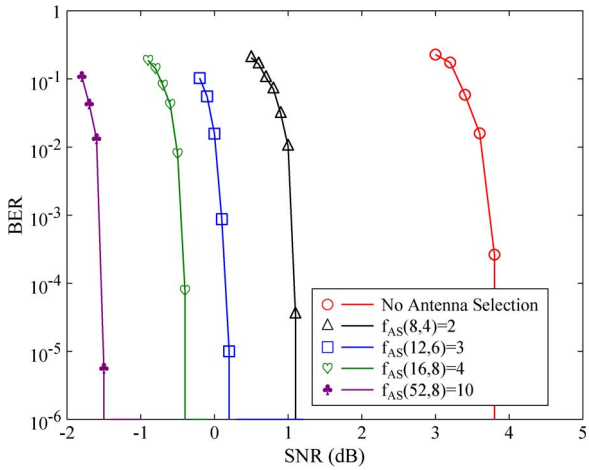


Fig. 7. BER performance of our NBJTRAS aided MIMO$(N_T, N_R; 4, 2; 4\text{-QAM})$ with various AS loading factors $f_{AS}(N_T, N_R)$, in comparison to that of the conventional MIMO$(4, 2; 4, 2; 4\text{-QAM})$ without AS.

of $I_{\text{out}} = 5$ iterations, implying that the proposed NBJTRAS aided MIMO scheme is capable of achieving a vanishingly low BER at SNR $= 1.4$ dB. This is confirmed by the BER performance shown in Fig. 7, where it can be seen that for the case of $f_{AS}(8, 4) = 2$, the "turbo-cliff" of the BER curve is observed just before the point of SNR $= 1.4$ dB. Fig. 6 also shows the EXIT curve of the conventional MIMO system operating without AS. Unlike the proposed NBJTRAS aided MIMO system, the conventional MIMO system dispensing with AS fails to achieve an open tunnel between the EXIT curve of the amalgamated inner MIMO soft-demapper-URC decoder and the outer RSC decoder. This implies that the conventional MIMO system operating without AS cannot attain a vanishingly low BER at SNR $= 1.4$ dB, which is confirmed by its BER performance shown in Fig. 7, where the actual convergence point of this conventional MIMO system using no AS is near SNR $= 4$ dB.

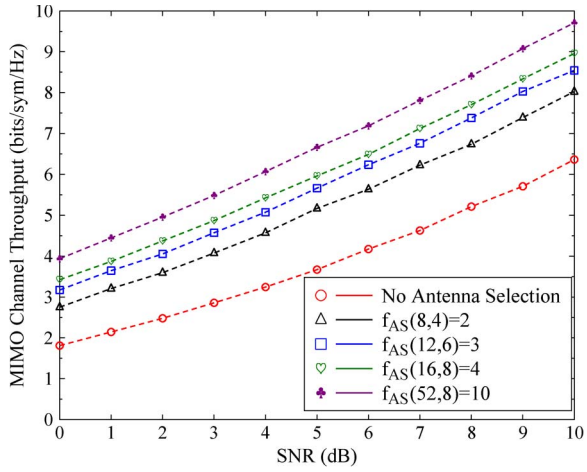


Fig. 8. MIMO channel throughput performance of our NBJTRAS aided MIMO$(N_T, N_R; 4, 2; 4 - \text{QAM})$ with various AS loading factors $f_{AS}(N_T, N_R)$, in comparison to that of the conventional MIMO$(4, 2; 4, 2; 4\text{-QAM})$ without AS.

The BER performance of our NBJTRAS aided MIMO$(N_T, N_R; 4, 2; 4 - \text{QAM})$ system is depicted in Fig. 7 for various AS loading factors, in comparison to the performance of the conventional MIMO$(4, 2; 4, 2; 4\text{-QAM})$ system operating without AS. It can be seen from Fig. 7 that given perfect CSI, the BER performance of the NBJTRAS aided MIMO system is significantly better than that of the conventional MIMO system using no AS. More specifically, the BER of the conventional MIMO system operating without AS achieves an infinitesimally low BER at about SNR $= 4$ dB, while the BER curve of the NBJTRAS aided MIMO system with $f_{AS}(8, 4) = 2$ reaches the same BER level at about SNR $= 1.4$ dB, yielding a significant SNR gain of about 2.6 dB. The BER curves of our NBJTRAS aided MIMO system associated with the AS loading factors of $f_{AS}(12, 6) = 3$, $f_{AS}(16, 8) = 4$ and $f_{AS}(52, 8) = 10$ are also shown in Fig. 7, which converge to a vanishingly low BER at about SNR $= 0.4$ dB, $-0.4$ dB and $-1.4$ dB, respectively, achieving SNR gains of about 3.6 dB, 4.4 dB and 5.4 dB, respectively, compared to the conventional MIMO system operating without AS. It can be seen that for the NBJTRAS aided MIMO system, higher performance gains are achieved by increasing the AS loading factor, at the cost of requiring more antennas. Most interestingly, although the rate of the gain improvement does appear to slow down as the AS loading factor increases, further significant gains are achieved, as the AS loading factor tends to large values. This is dissimilar to the standard diversity order trends, where the achieved gain tends to saturate upon increasing the diversity order to large values.

The achievable throughputs of the NBJTRAS aided MIMO$(N_T, N_R; 4, 2; 4 - \text{QAM})$ recorded for four different AS loading factors are compared to that of the conventional MIMO$(4, 2; 4, 2; 4\text{-QAM})$ operating without AS in Fig. 8. It can be seen from Fig. 8 that the higher the AS loading factor, the larger the achievable throughput gain of the proposed NBJTRAS scheme. This trend is similar to the BER performance enhancement attained by the NBJTRAS scheme depicted in Fig. 7. Specifically, given the AS loading factor of $f_{AS}(8, 4) = 2$ and the SNR value of 10 dB, the throughput

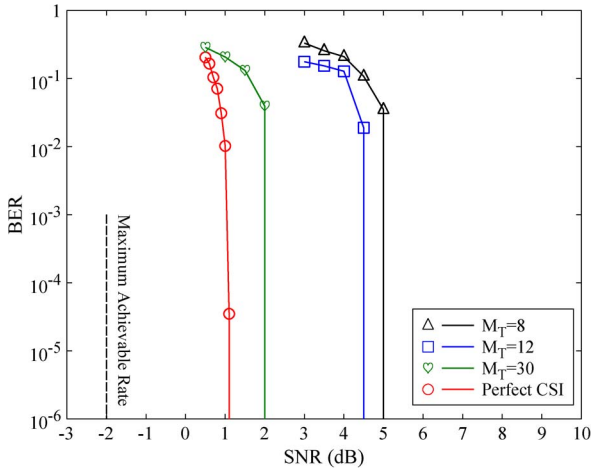


Fig. 9. Achievable BER performance of the NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM) with the AS loading factor of $f_{AS}(8,4) = 2$ assisted by the standard training based channel estimator.

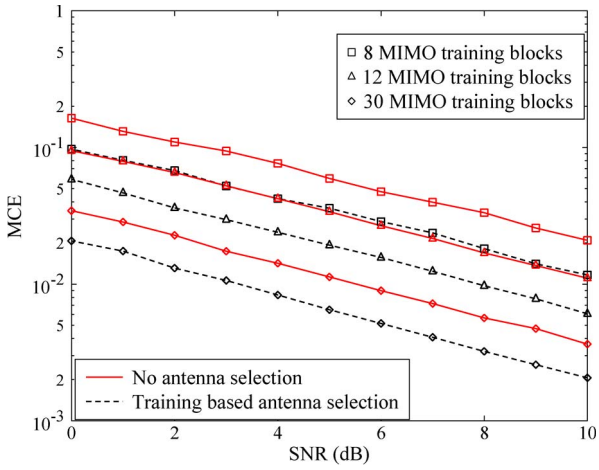


Fig. 10. Achievable MCE performance of the NBJTRAS aided MIMO(8, 4; 4, 2;-QAM) with the AS loading factor of $f_{AS}(8,4) = 2$ and the conventional MIMO(4, 2; 4, 2; 4-QAM) without AS, both assisted by the standard training based channel estimator.

obtained by the NBJTRAS aided MIMO system is approximately 1.5 [bits/symbol/Hz] higher than that of the conventional MIMO system operating without AS, while this becomes about 2 [bits/symbol/Hz] for the AS loading factor of $f_{AS}(12,6) = 3$.

### *B. Standard TBCE for NBJTRAS Aided MIMO*

Next we removed the assumption of having perfectly known CSI and investigated the achievable performance of the NBJTRAS aided three-stage turbo MIMO$(N_T, N_R; 4, 2; 4\text{-QAM})$ arrangement assisted by the standard TBCE scheme. For this purpose, we set the AS loading factor to $f_{AS}(8,4) = 2$. The corresponding BER performance is characterized in Fig. 9, using the optimal performance bound associated with perfect CSI as the benchmark. It can be seen from Fig. 9 that when the $M_T = 8$ MIMO training blocks are employed for the CE, the system's BER converges to a vanishingly low value at the SNR of 5 dB, while an infinitesimally low BER is attained at the SNR of approximately 4.5 dB, when the training data length increases to $M_T = 12$. When the CE utilizes the training data length of $M_T = 30$, it becomes capable of assisting the NBJTRAS aided MIMO(4, 2, 4-QAM) scheme to achieve a vanishingly low BER at the SNR value of 2 dB, but there still exists a performance gap of approximately 1 dB with respect to the benchmark associated with perfect CSI. Evidently, the NBJTRAS aided MIMO system assisted by the standard TBCE scheme having a training data length up to $M_T = 30$ is incapable of approaching the performance bound of the idealized NBJTRAS aided MIMO system associated with perfect CSI.

We further investigated the achievable MCE $J_{\text{MCE}}(\widehat{\boldsymbol{H}}_{sub})$ performance of the standard TBCE scheme, when assisting the NBJTRAS aided MIMO system as well as when assisting the conventional MIMO system operating without AS, where $\widehat{\boldsymbol{H}}_{sub} \in \mathbb{C}^{L_R \times L_T}$ denoted the channel matrix estimate of the MIMO system activated for data communication. The MCE results obtained for both systems are compared in Fig. 10, where the number of training blocks was set to $M_T = 8, 12$ and 30, respectively, for both systems. It can readily be seen from Fig. 10 that the MCE of the training based channel estimator for the NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM) with $f_{AS}(8,4) = 2$ is approximately 3 dB lower than the MCE of the same training based channel estimator for the conventional MIMO(4, 2; 4, 2; 4-QAM) operating without AS. The significance of the results shown in Fig. 10 warrants further discussion. For the conventional MIMO(4, 2; 4, 2; 4-QAM) using no AS, the channel estimator estimates the (2 $\times$ 4)-element MIMO channel matrix based on the $M_T$ training symbol blocks. The MCE between this (2 $\times$ 4)-element MIMO estimate and the true (2 $\times$ 4)-element MIMO channel invoked for data communication is then calculated and depicted in Fig. 10. For the NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM) using $f_{AS}(8,4) = 2$—given the same number of $M_T$ training symbol blocks—the channel estimator relying on RF chain reuse estimates the four (2 $\times$ 4)-element MIMO channel matrices to form the estimate of the "virtual" (4 $\times$ 8)-element full MIMO channel matrix. The NBJTRAS algorithm then selects a (2 $\times$ 4)-element estimated subset MIMO channel matrix from this estimated full MIMO channel matrix. The MCE between this selected (2 $\times$ 4)-element MIMO estimate and the corresponding true (2 $\times$ 4)-element MIMO channel activated for data communication is then calculated and depicted in Fig. 10. The results of Fig. 10 are remarkable—the same MMSE channel estimator relying on the same training length of $M_T$ is used for both systems, and yet the MCE obtained by the NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM) with $f_{AS}(8,4) = 2$ is approximately 3 dB lower than that achieved by the conventional MIMO(4, 2; 4, 2; 4-QAM) using no AS. This clearly demonstrates that the proposed NBJTRAS scheme is capable of improving the TBCE accuracy, and this also proves that AS is generally beneficial in terms of enhancing the accuracy of TBCE.

### *C. Proposed TTCE for NBJTRAS Aided MIMO*

We are now ready to investigate the overall performance of our proposed TTCE assisted and NBJTRAS aided three-stage

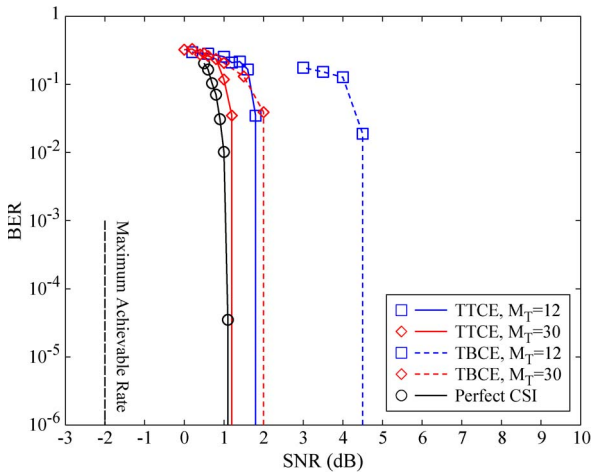


Fig. 11. Achievable BER performance of the proposed TTCE assisted and NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM), in comparison to that of the standard TBCE scheme assisted NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM). Both the systems employ an AS loading factor of $f_{AS}(8,4) = 2$.

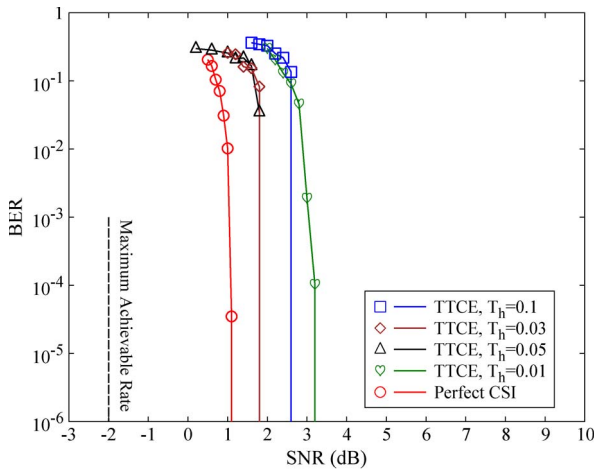


Fig. 12. Effects of threshold $T_h$ to the achievable performance of the TTCE assisted and NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM) with $f_{AS}(8,4) = 2$ and $M_T = 12$ initial training data.

turbo MIMO system. The maximum number of selected symbol blocks invoked for the BBSB-SDACE scheme was set to $M_{\text{sel}} = 500$, and we tested the effects of $M_T = 12$ and 30 MIMO training blocks. The achievable BER performance of the TTCE assisted and NBJTRAS aided MIMO(8, 4; 4, 2; 4-QAM) scheme is shown in Fig. 11 as solid curves, in comparison to those of the standard TBCE assisted and NBJTRAS aided three-stage turbo MIMO(8, 4; 4, 2; 4-QAM) arrangement, which are depicted as dashed curves in Fig. 11. The both systems adopted the AS loading factor of $f_{AS}(8,4) = 2$. The optimal performance bound of the idealized NBJTRAS aided three-stage turbo MIMO(8, 4; 4, 2; 4-QAM) with $f_{AS}(8,4) = 2$ and associated with perfect CSI is also included in Fig. 11 as a benchmark. It can be seen from Fig. 11 that the proposed TTCE assisted and NBJTRAS aided MIMO system is capable of achieving an infinitesimally low BER at SNR $\approx 1.8$ dB for $M_T = 12$. Hence it outperforms the standard TBCE assisted and NBJTRAS aided MIMO system by about 2.7 dB. This clearly demonstrates the power of the tier-two BBSB-SDACE scheme in improving the accuracy of the MIMO CE. Additionally, we note that there exists a small performance gap of approximately 0.7 dB between the TTCE aided MIMO system using $M_T = 12$ initial training data blocks and the perfect CSI performance bound. The reason for this phenomenon can be explained as follows. The CE error of the tier-one TBCE scheme has two effects. Firstly, the full channel matrix estimate $\widehat{\boldsymbol{H}}$ of (15) contains the CE error which in turn will impose AS errors. Secondly, the estimate $\widehat{\boldsymbol{H}}_{sub}$ of the selected subset MIMO system used for actual communication also contains the CE error. Even though the effects imposed by the CE error in $\widehat{\boldsymbol{H}}_{sub}$ on the system's achievable performance will be completely eliminated by the tier-two BBSB-SDACE scheme [32], the effects of the AS error cannot be dealt with by the tier-two DDCE scheme. Given a low number of initial training data, such as in the case of $M_T = 12$, the AS error will be noticeable, and this leads to a slight degradation of the overall performance observed in Fig. 11. By increasing the number of initial training data, we will be able to reduce the AS error and consequently to mitigate this performance loss. Observe in Fig. 11 that the TTCE assisted and NBJTRAS aided MIMO system is capable of closely approaching the optimal performance bound associated with perfect CSI with the aid of $M_T = 30$ initial training data blocks.

The effects of the block-of-bits selection threshold $T_h$ on the achievable BER performance were also investigated by varying the value of $T_h$ in the set $\{0.01, 0.03, 0.05, 0.1\}$ under the same system configuration. The corresponding results are shown in Fig. 12, where it can be seen that for $T_h = 0.03$ and 0.05, the same system performance is attained with the aid of the BBSB-SDACE scheme. However, for a threshold value of $T_h = 0.01$, a performance degradation occurs, since the number of decision blocks selected for CE is probably insufficient, for such a low threshold. On the other hand, given a high value of $T_h = 0.1$, some unreliable decision blocks may have been selected for CE and this may lead to a performance degradation. The results of Fig. 12 clearly confirm that as long as the threshold value is not chosen to be too high or too low, the performance of the BBSB-SDACE scheme remains insensitive to the actual value of $T_h$. Indeed, there exists a range of values for $T_h$, which allow the BBSB-SDACE scheme to attain its full performance potential. For this system, values in the interval $T_h \in [0.03, 0.05]$ are all appropriate.

Fig. 13 characterizes the MCE convergence performance of the proposed TTCE assisted and NBJTRAS aided three-stage turbo MIMO(8, 4; 4, 2; 4-QAM) associated with $M_T = 12$ initial training blocks and an AS loading factor of $f_{AS}(8,4) = 2$. Additionally, the MCE performance of the standard TBCE assisted and NBJTRAS aided three-stage turbo MIMO(8, 4; 4, 2; 4-QAM) scheme associated with $M_T = 12$ and 500 training blocks as well as the same $f_{AS}(8,4) = 2$ are also presented in Fig. 13 as comparison. From the results shown in Fig. 13, it can be seen that the tier-two BBSB-SDACE scheme is capable of substantially improving the accuracy of the CE by approximately 15 dB, and its MCE converges in 5 iterations from the initial MCE of the TBCE scheme with the aid of $M_T = 12$ training data to that of the TBCE scheme using $M_T = 500$ training

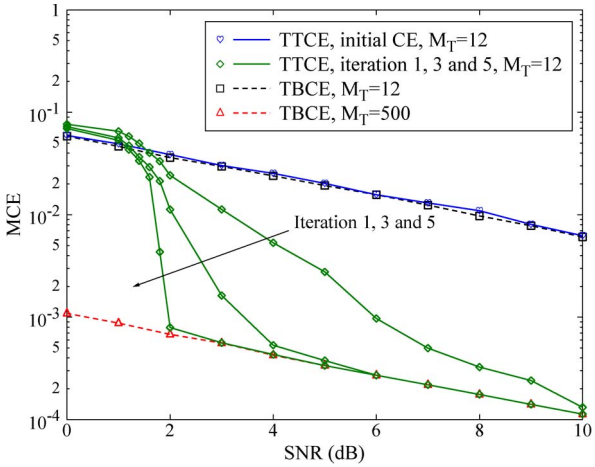


Fig. 13. MCE performance of the proposed TTCE assisted and NBJTRAS aided MIMO$(8, 4; 4, 2;$ 4-QAM) with $M_T = 12$, in comparison to those of the standard TBCE scheme assisted and NBJTRAS aided MIMO$(8, 4; 4, 2;$ 4-QAM) with $M_T = 12$ and 500. Both the systems employ $f_{AS}(8, 4) = 2$.

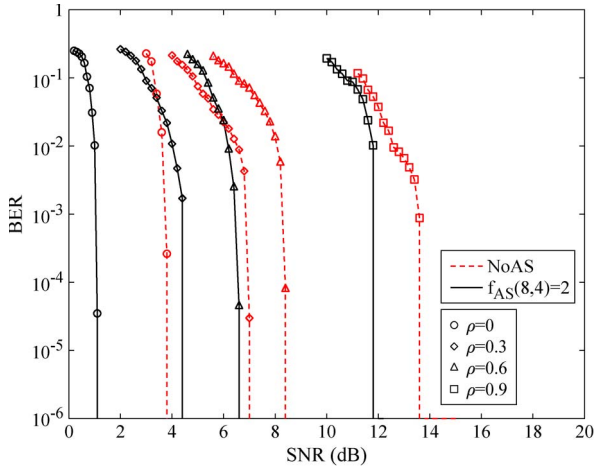


Fig. 14. Achievable BER performance of the NBJTRAS aided MIMO$(8, 4;$ $4, 2;$ 4-QAM) with $f_{AS}(8, 4) = 2$, in comparison to those of the conventional MIMO$(4, 2; 4, 2;$ 4-QAM)without AS, under various spatially correlated channel environments. Both the systems assume perfectly known CSI.

data for SNR $> 1.8$ dB. This is because in our simulations the BBSB-SDACE scheme selects no more than $M_{\text{sel}} = 500$ high-quality data blocks per frame for the DDCE and, furthermore, with the aid of the tier-two BBSB-SDACE scheme, our TTCE assisted and NBJTRAS aided MIMO$(8, 4; 4, 2;$ 4-QAM) arrangement attains a vanishingly low BER at the SNR value of 1.8 dB, as shown in Fig. 11. Therefore, under the operational conditions of SNR $> 1.8$ dB the selected data symbols are all correct and they are as "good" as the training data symbols.

### *D. NBJTRAS Aided MIMO Systems in a Spatially Correlated Environment*

The investigations carried out so far assumed an independent fading channel environment, an assumption that is commonly made in the literature of AS techniques. However, in practice, the MIMO channels are often spatially correlated, because the antenna spacing may not be sufficiently high to experience independently fading MIMO channels. Let us hence investigate the impact of spatial correlation on the BER performance of the proposed NBJTRAS algorithm. The NBJTRAS aided MIMO$(8, 4; 4, 2;$ 4-QAM) system using $f_{AS}(8, 4) = 2$ was adopted. The channel's spatial correlation factor was set to $\rho = 0, 0.3, 0.6$ and $0.9$, where $\rho = 0$ implied independent fading and $\rho = 1$ indicated fully correlated fading. Furthermore, a perfectly known full MIMO CSI was assumed and again all the results were averaged over 100 channel realizations. The simulation results obtained are shown in Fig. 14 in comparison to those of the conventional MIMO$(4, 2; 4, 2;$ 4-QAM) system using no AS. As expected, the BER performance of both the NBJTRAS aided MIMO system and of the conventional MIMO degrade, as the correlation between the MIMO channels increases, because increasing the channel's spatial correlation is expected to reduce the diversity gain of the MIMO system. It can also be observed that the NBJTRAS aided MIMO system is capable of outperforming the conventional MIMO system operating without AS in a spatially correlated channel environment. To be more explicit, in the independent fading environment, a performance gain of about SNR $= 2.6$ dB is achieved by the NBJTRAS aided MIMO system over the conventional MIMO system operating without AS. At the spatial correlation value of $\rho = 0.3$, the NBJTRAS aided MIMO still outperforms the conventional MIMO by approximately 2.6 dB. As the spatial correlation value is increased to $\rho = 0.6$ and $\rho = 0.9$, the performance gain is reduced to approximately 1.8 dB. Based on these results, we may conclude that at a low spatial correlation level of say $\rho < 0.3$, the NBJTRAS aided MIMO system is capable of achieving the same performance gain over the conventional MIMO system using no AS as in the independent fading environment, while in the highly correlated channel environment of say $\rho > 0.6$, the NBJTRAS aided MIMO is still capable of outperforming the conventional MIMO, but provides a smaller performance gain.

## V. Conclusion

In this paper, we have proposed a novel TTCE assisted and NBJTRAS aided three-stage turbo coded MIMO system, and our contribution has been twofold. Firstly, we have developed a low-complexity yet effective NBJTRAS aided near-capacity three-stage turbo coded MIMO system, which significantly outperforms the conventional MIMO system having the same number of RF chains and operating without AS, in terms of the achievable BER performance and throughput. Secondly, we have proposed a new TTCE scheme relying on a low training overhead for assisting the NBJTRAS aided MIMO system to approach the optimal MIMO performance bound associated with perfect CSI, which maintains a high system effective throughput while imposing a low computational complexity. More specifically, in tier one of the proposed TTCE scheme, we use a low-complexity low-pilot-overhead channel estimator relying on RF chain reuse to obtain an initial coarse estimate of the full MIMO channel matrix. This allows the NBJTRAS to be carried out based on this coarse TBCE. In tier two of the proposed scheme, we adopt a powerful semi-blind BBSB-SDACE scheme for assisting the three-stage turbo data detection–decoding process. Our simulation results have

demonstrated that our novel TTCE assisted and NBJTRAS aided MIMO system is indeed capable of approaching the optimal ML performance bound associated with perfect CSI, with the aid of a modest training overhead and without a significant increase in computational complexity. Our investigations have also revealed that the proposed NBJTRAS scheme is capable of significantly improving the accuracy of the TBCE.

## APPENDIX

List of Abbreviations

| | |
|---|---|
| AS | Antenna selection. |
| ASC | Antenna selection criterion. |
| BBSB-SDACE | Block-of-bits selection based soft-decision aided channel estimator. |
| BER | Bit error ratio. |
| BPB | Bits per block. |
| BPS | Bits per symbol. |
| CBAS | Capacity based antenna selection. |
| CE | Channel estimation. |
| CSI | Channel state information. |
| DDCE | Decision-directed channel estimation. |
| EXIT | EXtrinsic Information Transfer. |
| IIR | Infinite-duration impulse response. |
| JTRAS | Joint transmit and receive antenna selection. |
| LLR | Log-likelihood ratio. |
| MCE | Mean channel error. |
| MIMO | Multiple-input multiple-output. |
| ML | Maximum-likelihood. |
| MMSE | Minimum mean square error. |
| NBAS | Norm based antenna selection. |
| NBJTRAS | Norm based joint transmit and receive antenna selection. |
| QAM | Quadrature amplitude transmitter. |
| RA | Receive antenna. |
| RF | Radio frequency. |
| RSC | Recursive systematic code. |
| RxAS | Receive antenna selection. |
| SNR | Signal-to-noise ratio. |
| SSK | Space shift keying. |
| TA | Transmit antenna. |
| TxAS | Transmit antenna selection. |
| TBCE | Training based channel estimation. |
| TTCE | Two-tier channel estimation. |
| URC | Unity rate code. |

## REFERENCES


[1] L. Hanzo, O. R. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation*. Chichester, U.K.: Wiley, 2009.

[2] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[3] K.-C. Huang and Z. Wang, "Terahertz terabit wireless communication," *IEEE Microw. Mag.*, vol. 12, no. 4, pp. 108–116, Jun. 2011.

[4] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[5] K.-C. Huang and Z. Wang, "Millimeter-wave circular polarized beam-steering antenna array for Gigabit wireless communications," *IEEE Trans. Antennas Propag.*, vol. 54, no. 2, pp. 743–746, Feb. 2006.

[6] F. Khan and Z. Pi, "mmWave Mobile Broadband (MMB): Unleashing the 3–300 GHz spectrum," in *Proc. 34th IEEE Sarnoff Symp.*, Princeton, NJ, USA, May 3/4, 2011, pp. 1–6.

[7] S. Sanayei and A. Norsratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[8] W.-H. Chung and C.-Y. Hung, "Multi-antenna selection using space shift keying in MIMO systems," in *Proc. IEEE VTC-Spring*, Yokohama, Japan, May 6–9, 2012, pp. 1–5.

[9] R. Rajashekar, K. S. V. Hari, and L. Hanzo, "Antenna selection in spatial modulation systems," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 521–524, Mar. 2013.

[10] A. S. Hiwale, A. A. Ghatol, and S. D. Bhad, "Performance analysis of space–time trellis codes with receive antenna selection," in *Proc. 4th Int. Conf. Wireless Commun. Sensor Netw.*, Allahabad, India, Dec. 27–29, 2008, pp. 148–152.

[11] D. Liu and D. K. C. So, "Performance based receive antenna selection for V-BLAST systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 214–225, Jan. 2009.

[12] S. Sanayei and A. Nosratinia, "Capacity maximizing algorithms for joint transmit–receive antenna selection," in *Conf. Rec. 38th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 7–10, 2004, pp. 1773–1776.

[13] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 339–347, Feb. 2004.

[14] T. Gucluoglu and T. M. Duman, "Space–time coded systems with joint transmit and receive antenna selection," in *Proc. IEEE ICC*, Glasgow, U.K., Jun. 24–28, 2007, pp. 5305–5310.

[15] A. Yilmaz and O. Kucur, "Error performance of joint transmit and receive antenna selection in two hop amplify-and-forward relay system over Nakagami-$m$ fading channels," in *Proc. IEEE 21st PIMRC*, Istanbul, Turkey, Sep. 26–29, 2010, pp. 2198–2203.

[16] W. M. Gifford, M. Z. Win, and M. Chiani, "Diversity with practical channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1935–1947, Jul. 2005.

[17] Q. Ma and C. Tepedelenlioglu, "Antenna selection for space–time coded systems with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 710–719, Feb. 2007.

[18] A. B. Narasimhamurthy and C. Tepedelenlioglu, "Antenna selection for MIMO-OFDM systems with channel estimation error," *IEEE Trans. Veh. Technol.*, vol. 58, no. 5, pp. 2269–2278, Jun. 2009.

[19] V. Kristem, N. B. Mehta, and A. F. Molisch, "A novel, balanced, energy-efficient training method for receive antenna selection," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2742–2752, Sep. 2010.

[20] S. Kashyap and N. B. Mehta, "Joint antenna selection and frequency-domain scheduling in OFDMA systems with imperfect estimates from dual pilot training scheme," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3473–3483, Jul. 2013.

[21] P. Zhang, S. Chen, and L. Hanzo, "Reduced-complexity near-capacity joint channel estimation and three-stage turbo detection for coherent space–time shift keying," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1902–1912, May 2013.

[22] M. Loncar, R. Muller, J. Wehinger, C. F. Mecklenbrauker, and T. Abe, "Iterative channel estimation and data detection in frequency-selective fading MIMO channels," *Eur. Trans. Telecommun.*, vol. 15, no. 5, pp. 459–470, Sep./Oct. 2004.

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

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

[25] J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Turbo multi-user detection for OFDM/SDMA systems relying on differential evolution aided iterative channel estimation," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1621–1633, Jun. 2012.

[26] M. Sandell, C. Luschi, P. Strauch, and R. Yan, "Iterative CE using soft decision feedback," in *Proc. IEEE GLOBECOM*, Sydney, N.S.W., Australia, Nov. 8–12, 1998, pp. 3728–3733.

[27] S. Song, A. C. Singer, and K. Sung, "Soft input CE for turbo equalization," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2885–2894, Oct. 2004.

[28] B. Hu *et al.*, "Iterative joint CE and successive interference cancellation using a SISO-SAGE algorithm for coded CDMA," in *Conf. Rec. 38th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 7–10, 2004, pp. 622–626.

[29] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: Theory and experiments," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.
[30] Y. Wu, X. Zhu, and A. K. Nandi, "Soft-input turbo CE for single-carrier multiple-input–multiple-output systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3867–3873, Sep. 2009.
[31] W. Haselmayr, D. Schellander, and A. Springer, "Iterative CE and turbo equalization for time-varying channels in a coded OFDM-LTE system for 16-QAM and 64-QAM," in *Proc. 21st IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, Istanbul, Turkey, Sep. 26–29, 2010, pp. 614–618.
[32] P. Zhang, S. Chen, and L. Hanzo, "Embedded iterative semi-blind channel estimation for three-stage-concatenated MIMO-aided QAM turbo-transceivers," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 439–446, Jan. 2014.
[33] I. Land, P. A. Hoeher, and S. Gligorević, "Computation of symbol-wise mutual information in transmission systems with logAPP decoders and application to EXIT charts," in *Proc. 5th Int. ITG Conf. SCC*, Erlangen, Germany, Jan. 2004, pp. 195–202.
[34] J. Wang *et al.*, "Near-capacity three-stage MMSE turbo equalization using irregular convolutional codes," in *Proc. TURBOCODING*, Munich, Germany, Apr. 3–7, 2006, pp. 1–6.
[35] R. G. Maunder and L. Hanzo, "Iterative decoding convergence and termination of serially concatenated codes," *IEEE Trans. Veh. Technol.*, vol. 90, no. 1, pp. 216–224, Jan. 2010.
[36] C. Xu, S. Sugiura, S. X. Ng, and L. Hanzo, "Spatial modulation and space–time shift keying: Optimal performance at a reduced detection complexity," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 206–216, Jan. 2013.
[37] Z. Guo and P. Nilsson, "Algorithm and implementation of the $K$-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.
[38] L. Wang, L. Xu, S. Chen, and L. Hanzo, "*A priori*-LLR-threshold-assisted $K$-best sphere detection for MIMO channels," in *Proc. IEEE VTC-Spring*, Singapore, May 11–14, 2008, pp. 867–871.
[39] T. Gucluoglu and E. Panayirci, "Performance of transmit and receive antenna selection in the presence of channel estimation errors," *IEEE Commun. Lett.*, vol. 12, no. 5, pp. 371–373, May 2008.

**Peichang Zhang** received the B.Eng. degree (with first-class honor) in electronic engineering from the University of Central Lancashire, Preston, U.K., in 2009 and the M.Sc. degree (with distinction) in wireless communications from the University of Southampton, Southampton, U.K., in 2010. He is currently working toward the Ph.D. degree with the Communications, Signal Processing and Control Research Group, School of Electronics and Computer Science, University of Southampton.

His research interests include cooperative communications, coherent and non-coherent detection, iterative detection, as well as channel estimation.

**Sheng Chen** (M'90–SM'97–F'08) received the B.Eng. degree in control engineering from the East China Petroleum Institute, Dongying, China, in 1982, the Ph.D. degree in control engineering from the City University, London, U.K., in 1986, and the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2005.

From 1986 to 1999, he held research and academic appointments at the Universities of Sheffield, Edinburgh, and Portsmouth, all in U.K. Since 1999, he has been with the School of Electronics and Computer Science, University of Southampton, where he currently holds the post of Professor in Intelligent Systems and Signal Processing. He is a Distinguished Adjunct Professor at King Abdulaziz University, Jeddah, Saudi Arabia. He has published over 460 research papers. His recent research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, and evolutionary computation methods and optimization.

Prof. Chen is a Chartered Engineer and a Fellow of IET. He is an ISI highly cited researcher in the engineering category (March 2004).

**Lajos Hanzo** (M'91–SM'92–F'04) received the M.S. degree in electronics and the Ph.D. degree from the Technical University of Budapest, Budapest, Hungary, in 1976 and 1983, respectively; the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2004; and the "Doctor Honoris Causa" degree from the Technical University of Budapest in 2009.

During his 35-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. He is currently 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) U.K., the European IST Programme, and the Mobile Virtual Centre of Excellence (VCE), U.K. During 2008–2012, he was a Chaired Professor with Tsinghua University, Beijing China. He is an enthusiastic supporter of industrial and academic liaison, and he offers a range of industrial courses. He has successfully supervised more than 80 Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications totaling in excess of 10 000 pages, and published more than 1400 research entries at IEEE Xplore. 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.

Prof. Hanzo is a Fellow of the Royal Academy of Engineering, the Institution of Engineering and Technology, and the European Association for Signal Processing. He is also a Governor of the IEEE Vehicular Technology Society. During 2008–2012, he was the Editor-in-Chief of the IEEE Press. He has served as the Technical Program Committee Chair and the General Chair of IEEE conferences, has presented keynote lectures, and has been awarded a number of distinctions.