

# Signal Detection and Parameter Estimation in Massive MIMO Systems

Rodrigo C. de Lamare and Raimundo Sampaio-Neto

Centre for Telecommunications Studies (CETUC)

Pontifical Catholic University of Rio de Janeiro, Gávea - 22453-900, Rio de Janeiro, Brazil

Communications Research Group

Department of Electronics, University of York, York YO10 5DD, United Kingdom

Email: delamare@cetuc.puc-rio.br, raimundo@cetuc.puc-rio.br

**Abstract**—This book chapter reviews signal detection and parameter estimation techniques for multiuser multiple-antenna wireless systems with a very large number of antennas, known as massive multi-input multi-output (MIMO) systems. We consider both centralized antenna systems (CAS) and distributed antenna systems (DAS) architectures in which a large number of antenna elements are employed and focus on the uplink of a mobile cellular system. In particular, we focus on receive processing techniques that include signal detection and parameter estimation problems and discuss the specific needs of massive MIMO systems. Simulation results illustrate the performance of detection and estimation algorithms under several scenarios of interest. Key problems are discussed and future trends in massive MIMO systems are pointed out.

**Index Terms**—massive MIMO, signal detection, parameter estimation, algorithms,

## I. INTRODUCTION

Future wireless networks will have to deal with a substantial increase of data transmission due to a number of emerging applications that include machine-to-machine communications and video streaming [1]- [4]. This very large amount of data exchange is expected to continue and rise in the next decade or so, presenting a very significant challenge to designers of fifth-generation (5G) wireless communications systems [4]. Amongst the main problems are how to make the best use of the available spectrum and how to increase the energy efficiency in the transmission and reception of each information unit. 5G communications will have to rely on technologies that can offer a major increase in transmission capacity as measured in bits/Hz/area but do not require increased spectrum bandwidth or energy consumption.

Multiple-antenna or multi-input multi-output (MIMO) wireless communication devices that employ antenna arrays with a very large number of antenna elements which are known as massive MIMO systems have the potential to overcome those challenges and deliver the required data rates, representing a key enabling technology for 5G [5]- [8]. Among the devices of massive MIMO networks are user terminals, tablets, machines and base stations which could be equipped with a number of antenna elements with orders of magnitude higher than current devices. Massive MIMO networks will be structured by the following key elements: antennas, electronic components,

The work of the authors is funded by the Pontifical Catholic University of Rio de Janeiro and CNPq.

network architectures, protocols and signal processing. The network architecture, in particular, will evolve from homogeneous cellular layouts to heterogeneous architectures that include small cells and the use of coordination between cells [9]. Since massive MIMO will be incorporated into mobile cellular networks in the future, the network architecture will necessitate special attention on how to manage the interference created [10] and measurements campaigns will be of fundamental importance [11]- [13]. The coordination of adjacent cells will be necessary due to the current trend towards aggressive reuse factors for capacity reasons, which inevitably leads to increased levels of inter-cell interference and signalling. The need to accommodate multiple users while keeping the interference at an acceptable level will also require significant work in scheduling and medium-access protocols.

Another important aspect of massive MIMO networks lies in the signal processing, which must be significantly advanced for 5G. In particular, MIMO signal processing will play a crucial role in dealing with the impairments of the physical medium and in providing cost-effective tools for processing information. Current state-of-the-art in MIMO signal processing requires a computational cost for transmit and receive processing that grows as a cubic or super-cubic function of the number of antennas, which is clearly not scalable with a large number of antenna elements. We advocate the need for simpler solutions for both transmit and receive processing tasks, which will require significant research effort in the next years. Novel signal processing strategies will have to be developed to deal with the problems associated with massive MIMO networks like computational complexity and its scalability, pilot contamination effects, RF impairments, coupling effects, delay and calibration issues.

In this chapter, we focus on signal detection and parameter estimation aspects of massive MIMO systems. We consider both centralized antenna systems (CAS) and distributed antenna systems (DAS) architectures in which a large number of antenna elements are employed and focus on the uplink of a mobile cellular system. In particular, we focus on the uplink and receive processing techniques that include signal detection and parameter estimation problems and discuss specific needs of massive MIMO systems. We review the optimal maximum likelihood detector, nonlinear and linear suboptimal detectors and discuss potential contributions to the area. We also describe iterative detection and decoding algorithms, which

exchange soft information in the form of log likelihood ratios (LLRs) between detectors and channel decoders. Another important area of investigation includes parameter estimation techniques, which deal with methods to obtain the channel state information, compute the parameters of the receive filters and the hardware mismatch. Simulation results illustrate the performance of detection and estimation algorithms under scenarios of interest. Key problems are discussed and future trends in massive MIMO systems are pointed out.

This chapter is structured as follows. Section II reviews the signal models with CAS and DAS architectures and discusses the application scenarios. Section III is dedicated to detection techniques, whereas Section IV is devoted to parameter estimation methods. Section V discusses the results of some simulations and Section VI presents some open problems and suggestions for further work. The conclusions of this chapter are given in Section VII.

## II. SIGNAL MODELS AND APPLICATION SCENARIOS

In this section, we describe signal models for the uplink of multiuser massive MIMO systems in mobile cellular networks. In particular, we employ a linear algebra approach to describe the transmission and how the signals are collected at the base station or access point. We consider both CAS and DAS [15], [16] configurations. In the CAS configuration a very large array is employed at the rooftop or at the façade of a building or even at the top of a tower. In the DAS scheme, distributed radio heads are deployed over a given geographic area associated with a cell and these radio devices are linked to a base station equipped with an array through either fibre optics or dedicated radio links. These models are based on the assumption of a narrowband signal transmission over flat fading channels which can be easily generalized to broadband signal transmission with the use of multi-carrier systems.

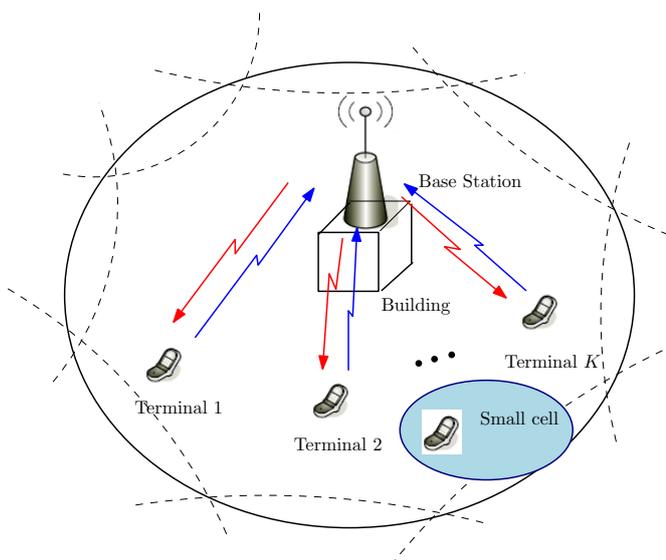


Fig. 1. Mobile cellular network with a CAS configuration.

The scenario we are interested in this work is that of mobile cellular networks beyond LTE-A [2] and 5G communications

[4], which is illustrated in Fig. 1. In such networks, massive MIMO will play a key role with the deployment of hundreds of antenna elements at the base station using CAS or using DAS over the cell of interest, coordination between cells and a more modest number of antenna elements at the user terminals. At the base station, very large antenna arrays could be deployed on the roof or on the façade of buildings. With further development in the area of compact antennas and techniques to mitigate mutual coupling effects, it is likely that the number of antenna elements at the user terminals (mobile phones, tablets and other gadgets) might also be significantly increased from 1–4 elements in current terminals to 10–20 in future devices. In these networks, it is preferable to employ time-division-duplexing (TDD) mode to perform uplink channel estimation and obtain downlink CSI by reciprocity for signal processing at the transmit side. This operation mode will require cost-effective calibration algorithms. Another critical requirement is the uplink channel estimation, which employs non-orthogonal pilots and, due to the existence of adjacent cells and the coherence time of the channel, needs to reuse the pilots [63]. Pilot contamination occurs when CSI at the base station in one cell is affected by users from other cells. In particular, the uplink (or multiple-access channel) will need CSI obtained by uplink channel estimation, efficient multiuser detection and decoding algorithms. The downlink (also known as the broadcast channel) will require CSI obtained by reciprocity for transmit processing and the development of cost-effective scheduling and precoding algorithms. A key challenge in the scenario of interest is how to deal with a very large number of antenna elements and develop cost-effective algorithms, resulting in excellent performance in terms of the metrics of interest, namely, bit error rate (BER), sum-rate and throughput. In what follows, signal models that can describe CAS and DAS schemes will be detailed.

### A. Centralized Antenna System Model

In this subsection, we consider a multiuser massive MIMO system with CAS using  $N_A$  antenna elements at the receiver, which is located at a base station of a cellular network installed at the rooftop of a building or a tower, as illustrated in Fig. 1. Following this description, we consider a multiuser massive MIMO system with  $K$  users that are equipped with  $N_U$  antenna elements and communicate with a receiver with  $N_A$  antenna elements, where  $N_A \geq KN_U$ . At each time instant, the  $K$  users transmit  $N_U$  symbols which are organized into a  $N_U \times 1$  vector  $\mathbf{s}_k[i] = [s_{k,1}[i], s_{k,2}[i], \dots, s_{k,N_U}[i]]^T$  taken from a modulation constellation  $\mathcal{A} = \{a_1, a_2, \dots, a_N\}$ . The data vectors  $\mathbf{s}_k[i]$  are then transmitted over flat fading channels. The received signal after demodulation, pulse-matched filtering and sampling is collected in an  $N_A \times 1$  vector  $\mathbf{r}[i] = [r_1[i], r_2[i], \dots, r_{N_A}[i]]^T$  with sufficient statistics for estimation and detection as described by

$$\begin{aligned} \mathbf{r}[i] &= \sum_{k=1}^K \gamma_k \mathbf{H}_k \mathbf{s}_k[i] + \mathbf{n}[i] \\ &= \sum_{k=1}^K \mathbf{G}_k \mathbf{s}_k[i] + \mathbf{n}[i], \end{aligned} \quad (1)$$

where the  $N_A \times 1$  vector  $\mathbf{n}[i]$  is a zero mean complex circular symmetric Gaussian noise with covariance matrix  $E[\mathbf{n}[i]\mathbf{n}^H[i]] = \sigma_n^2 \mathbf{I}$ . The data vectors  $\mathbf{s}_k[i]$  have zero mean and covariance matrices  $E[\mathbf{s}_k[i]\mathbf{s}_k^H[i]] = \sigma_{s_k}^2 \mathbf{I}$ , where  $\sigma_{s_k}^2$  is the user  $k$  transmit signal power. The elements  $h_{i,j}^k$  of the  $N_A \times N_U$  channel matrices  $\mathbf{H}_k$  are the complex channel gains from the  $j$ th transmit antenna of user  $k$  to the  $i$ th receive antenna. For a CAS architecture, the channel matrices  $\mathbf{H}_k$  can be modeled using the Kronecker channel model [14] as detailed by

$$\mathbf{H}_k = \Theta_R^{1/2} \mathbf{H}_k^o \Theta_T^{1/2}, \quad (2)$$

where  $\mathbf{H}_k^o$  has complex channel gains obtained from complex Gaussian random variables with zero mean and unit variance,  $\Theta_R$  and  $\Theta_T$  denote the receive and transmit correlation matrices, respectively. The components of correlation matrices  $\Theta_R$  and  $\Theta_T$  are of the form

$$\Theta_{R/T} = \begin{pmatrix} 1 & \rho & \rho^4 & \dots & \rho^{(N_a-1)^2} \\ \rho & 1 & \rho & \dots & \vdots \\ \rho^4 & \rho & 1 & \vdots & \rho^4 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{(N_a-1)^2} & \dots & \rho^4 & \rho & 1 \end{pmatrix} \quad (3)$$

where  $\rho$  is the correlation index of neighboring antennas and  $N_a$  is the number of antennas of the transmit or receive array. When  $\rho = 0$  we have an uncorrelated scenario and when  $\rho = 1$  we have a fully correlated scenario. The channels between the different users are assumed uncorrelated due to their geographical location.

The parameters  $\gamma_k$  represent the large-scale propagation effects for user  $k$  such as path loss and shadowing which are represented by

$$\gamma_k = \alpha_k \beta_k, \quad (4)$$

where the path loss  $\alpha_k$  for each user is computed by

$$\alpha_k = \sqrt{\frac{L_k}{d_k^\tau}}, \quad (5)$$

where  $L_k$  is the power path loss of the link associated with user  $k$ ,  $d_k$  is the relative distance between the user and the base station,  $\tau$  is the path loss exponent chosen between 2 and 4 depending on the environment.

The log-normal shadowing  $\beta_k$  is given by

$$\beta_k = 10^{\frac{\sigma_k v_k}{10}}, \quad (6)$$

where  $\sigma_k$  is the shadowing spread in dB and  $v_k$  corresponds to a real-valued Gaussian random variable with zero mean and unit variance. The  $N_A \times N_U$  composite channel matrix that includes both large-scale and small-scale fading effects is denoted as  $\mathbf{G}_k$ .

### B. Distributed Antenna Systems Model

In this subsection, we consider a multiuser massive MIMO system with a DAS configuration using  $N_B$  antenna elements at the base station and  $L$  remote radio heads each with  $Q$  antenna elements, which are distributed over the cell and linked to the base station via wired links, as illustrated in

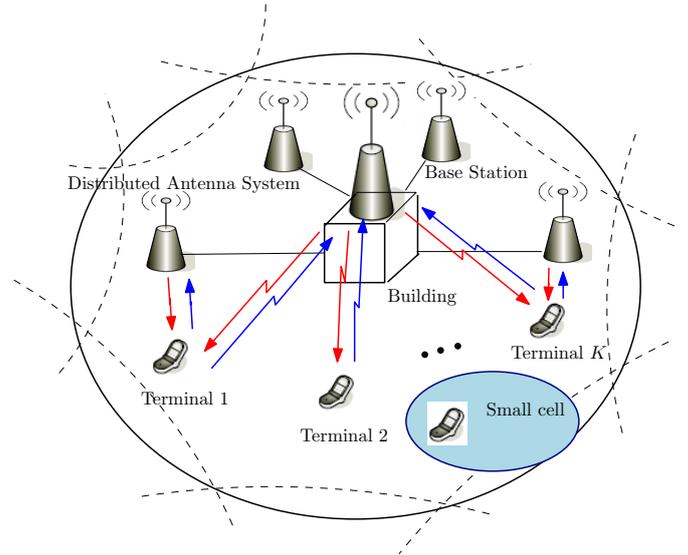


Fig. 2. Mobile cellular network with a DAS configuration.

Fig. 2. Following this description, we consider a multiuser massive MIMO system with  $K$  users that are equipped with  $N_U$  antenna elements and communicate with a receiver with a DAS architecture with a total of  $N_A = N_B + LQ$  antenna elements, where  $N_A \geq KN_U$ . In our exposition, when the number of remote radio heads is set to zero, i.e.,  $L = 0$ , the DAS architecture reduces to the CAS scheme with  $N_A = N_B$ .

At each time instant, the  $K$  users transmit  $N_U$  symbols which are organized into a  $N_U \times 1$  vector  $\mathbf{s}_k[i] = [s_{k,1}[i], s_{k,2}[i], \dots, s_{k,N_U}[i]]^T$  taken from a modulation constellation  $\mathcal{A} = \{a_1, a_2, \dots, a_N\}$ . The data vectors  $\mathbf{s}_k[i]$  are then transmitted over flat fading channels. The received signal after demodulation, pulse-matched filtering and sampling is collected in an  $N_A \times 1$  vector  $\mathbf{r}[i] = [r_1[i], r_2[i], \dots, r_{N_A}[i]]^T$  with sufficient statistics for estimation and detection as described by

$$\begin{aligned} \mathbf{r}[i] &= \sum_{k=1}^K \gamma_k \mathbf{H}_k \mathbf{s}_k[i] + \mathbf{n}[i] \\ &= \sum_{k=1}^K \mathbf{G}_k \mathbf{s}_k[i] + \mathbf{n}[i], \end{aligned} \quad (7)$$

where the  $N_A \times 1$  vector  $\mathbf{n}[i]$  is a zero mean complex circular symmetric Gaussian noise with covariance matrix  $E[\mathbf{n}[i]\mathbf{n}^H[i]] = \sigma_n^2 \mathbf{I}$ . The data vectors  $\mathbf{s}_k[i]$  have zero mean and covariance matrices  $E[\mathbf{s}_k[i]\mathbf{s}_k^H[i]] = \sigma_{s_k}^2 \mathbf{I}$ , where  $\sigma_{s_k}^2$  is the user  $k$  signal power. The elements  $h_{i,j}^k$  of the  $N_A \times N_U$  channel matrices  $\mathbf{H}_k$  are the complex channel gains from the  $j$ th transmit antenna to the  $i$ th receive antenna. Unlike the CAS architecture, in a DAS setting the channels between remote radio heads are less likely to suffer from correlation due to the fact that they are geographically separated. However, for the antenna elements located at the base station and at each remote radio head, the  $L + 1$  submatrices of  $\mathbf{H}_k$  can be modeled using the Kronecker channel model [14] as detailed in the previous subsection. The major difference between CAS

and DAS schemes lies in the large-scale propagation effects. Specifically, with DAS the links between the users and the distributed antennas experience in average lower path loss effects because of the reduced distance between their antennas. This helps to create better wireless links and coverage of the cell. Therefore, the large-scale propagation effects are modeled by an  $N_A \times N_A$  diagonal matrix given by

$$\gamma_k = \text{diag} \left( \underbrace{\gamma_{k,1} \cdots \gamma_{k,1}}_{N_B} \underbrace{\gamma_{k,2} \cdots \gamma_{k,2}}_Q \cdots \underbrace{\gamma_{k,L+1} \cdots \gamma_{k,L+1}}_Q \right) \quad (8)$$

where the parameters  $\gamma_{k,j}$  for  $j = 1, \dots, L+1$  denote the large-scale propagation effects like shadowing and pathloss from the  $k$ th user to the  $j$ th radio head. The parameters  $\gamma_{k,j}$  for user  $k$  and distributed antenna  $j$  are described by

$$\gamma_{k,j} = \alpha_{k,j} \beta_{k,j}, \quad j = 1, \dots, L+1 \quad (9)$$

where the path loss  $\alpha_{k,j}$  for each user and antenna is computed by

$$\alpha_{k,j} = \sqrt{\frac{L_{k,j}}{d_{k,j}^\tau}}, \quad (10)$$

where  $L_{k,j}$  is the power path loss of the link associated with user  $k$  and the  $j$ th radio head,  $d_{k,j}$  is the relative distance between the user and the radio head,  $\tau$  is the path loss exponent chosen between 2 and 4 depending on the environment. The log-normal shadowing  $\beta_{k,j}$  is given by

$$\beta_{k,j} = 10^{\frac{\sigma_k v_{k,j}}{10}}, \quad (11)$$

where  $\sigma_k$  is the shadowing spread in dB and  $v_{k,j}$  corresponds to a real-valued Gaussian random variable with zero mean and unit variance. The  $N_A \times N_U$  composite channel matrix that includes both large-scale and small-scale fading effects is denoted as  $\mathbf{G}_k$ .

### III. DETECTION TECHNIQUES

In this section, we examine signal detection algorithms for massive MIMO systems. In particular, we review various detection techniques and also describe iterative detection and decoding schemes that bring together detection algorithms and error control coding.

#### A. Detection Algorithms

In the uplink of the multiuser massive MIMO systems under consideration, the signals or data streams transmitted by the users to the receiver overlap and typically result in multiuser interference at the receiver. This means that the interfering signals cannot be easily demodulated at the receiver unless there is a method to separate them. In order to separate the data streams transmitted by the different users, a designer must resort to detection techniques, which are similar to multiuser detection methods [17].

The optimal maximum likelihood (ML) detector is described by

$$\hat{\mathbf{s}}_{\text{ML}}[i] = \arg \min_{\mathbf{s}[i] \in \mathcal{A}} \|\mathbf{r}[i] - \mathbf{G}\mathbf{s}[i]\|^2 \quad (12)$$

where the  $KN_U \times 1$  data vector  $\mathbf{s}[i]$  has the symbols of all users stacked and the  $KN_U \times N_A$  channel matrix  $\mathbf{G} = [\mathbf{G}_1 \dots \mathbf{G}_K]$  contains the channels of all users concatenated. The ML detector has a cost that is exponential in the number of data streams and the modulation order which is too costly for systems with a large number of antennas. Even though the ML solution can be alternatively computed using sphere decoder (SD) algorithms [19]- [22] that are very efficient for MIMO systems with a small number of antennas, the cost of SD algorithms depends on the noise variance, the number of data streams to be detected and the signal constellation, resulting in high computational costs for low SNR values, high-order constellations and a large number of data streams.

The high computational cost of the ML detector and the SD algorithms in scenarios with large arrays have motivated the development of numerous alternative strategies for MIMO detection, which are based on the computation of receive filters and interference cancellation strategies. The key advantage of these approaches with receive filters is that the cost is typically not dependent on the modulation, the receive filter is computed only once per data packet and performs detection with the aid of decision thresholds. Algorithms that can compute the parameters of receive filters with low cost are of central importance to massive MIMO systems. In what follows, we will briefly review some relevant suboptimal detectors, which include linear and decision-driven strategies.

Linear detectors [23] include approaches based on the receive matched filter (RMF), zero forcing (ZF) and minimum mean-square error (MMSE) designs that are described by

$$\hat{\mathbf{s}}[i] = Q(\mathbf{W}^H \mathbf{r}[i]), \quad (13)$$

where the receive filters are

$$\mathbf{W}_{\text{RMF}} = \mathbf{G}, \quad \text{for the RMF}, \quad (14)$$

$$\mathbf{W}_{\text{MMSE}} = \mathbf{G}(\mathbf{G}^H \mathbf{G} + \sigma_s^2 / \sigma_n^2 \mathbf{I})^{-1}, \quad \text{for the MMSE design}, \quad (15)$$

and

$$\mathbf{W}_{\text{ZF}} = \mathbf{G}(\mathbf{G}^H \mathbf{G})^{-1}, \quad \text{for the ZF design}, \quad (16)$$

and  $Q(\cdot)$  represents the slicer used for detection.

Decision-driven detection algorithms such as successive interference cancellation (SIC) approaches used in the Vertical-Bell Laboratories Layered Space-Time (VBLAST) systems [24]- [28] and decision feedback (DF) [29]- [46] detectors are techniques that can offer attractive trade-offs between performance and complexity. Prior work on SIC and DF schemes has been reported with DF detectors with SIC (S-DF) [24]- [28] and DF receivers with parallel interference cancellation (PIC) (P-DF) [39], [40], [44], combinations of these schemes and mechanisms to mitigate error propagation [43], [46], [47].

SIC detectors [24]- [28] apply linear receive filters to the received data followed by subtraction of the interference and subsequent processing of the remaining users. Ordering algorithms play an important role as they significantly affect the performance of SIC receivers. Amongst the existing criteria for ordering are those based on the channel norm, the SINR, the SNR and on exhaustive search strategies. The performance of exhaustive search strategies is the best followed by

SINR-based ordering, SNR-based ordering and channel norm-based ordering, whereas the computational complexity of an exhaustive search is by far the highest, followed by SINR-based ordering, SNR-based ordering and channel norm-based ordering. The data symbol of each user is detected according to:

$$\hat{s}_k[i] = Q(\mathbf{w}_k^H \mathbf{r}_k[i]), \quad (17)$$

where the successively cancelled received data vector that follows a chosen ordering in the  $k$ -th stage is given by

$$\mathbf{r}_k[i] = \mathbf{r}[i] - \sum_{j=1}^{k-1} \mathbf{g}_j \hat{s}_j[i], \quad (18)$$

where  $\mathbf{g}_j$  corresponds to the columns of the composite channel matrix  $\mathbf{G}_j$ . After subtracting the detected symbols from the received signal vector, the remaining signal vector is processed either by an MMSE or a ZF receive filter for the data estimation of the remaining users. The computational complexity of the SIC detector based on either the MMSE or the ZF criteria is similar and requires a cubic cost in  $N_A$  ( $O(N_A^3)$ ) although the performance of MMSE-based receive filters is superior to that of ZF-based detectors.

A generalization of SIC techniques, the multi-branch successive interference cancellation (MB-SIC) algorithm, employs multiple SIC algorithms in parallel branches. The MB-SIC algorithm relies on different ordering patterns and produces multiple candidates for detection, approaching the performance of the ML detector. The ordering of the first branch is identical to a standard SIC algorithm and could be based on the channel norm or the SINR, whereas the remaining branches are ordered by shifted orderings relative to the first branch. In the  $\ell$ -th branch, the MB-SIC detector successively detects the symbols given by the vector  $\hat{\mathbf{s}}_\ell[i] = [\hat{s}_{\ell,1}[i], \hat{s}_{\ell,2}[i], \dots, \hat{s}_{\ell,K}[i]]^T$ . The term  $\hat{\mathbf{s}}_\ell[i]$  represents the  $K \times 1$  ordered estimated symbol vector, which is detected according to the ordering pattern  $\mathbf{T}_\ell, \ell = 1, \dots, S$  for the  $\ell$ -th branch. The interference cancellation performed on the received vector  $\mathbf{r}[i]$  is described by:

$$\mathbf{r}_{\ell,k}[i] = \mathbf{r}[i] - \sum_{j=1}^{k-1} \mathbf{g}_{\ell,j} \hat{s}_{\ell,j}[i] \quad (19)$$

where the transformed channel column  $\mathbf{g}$  is obtained by  $\mathbf{g}_\ell = \mathbf{T}_\ell \mathbf{g}$ , the term  $\mathbf{g}'_k$  represents the  $k$ -th column of the ordered channel  $\mathbf{G}'$  and  $\hat{s}_{\ell,k}$  denotes the estimated symbol for each data stream obtained by the MB-SIC algorithm.

At the end of each branch we can transform  $\hat{\mathbf{s}}_\ell[i]$  back to the original order  $\tilde{\mathbf{s}}_\ell[i]$  by using  $\mathbf{T}_\ell$  as  $\tilde{\mathbf{s}}_\ell[i] = \mathbf{T}_\ell^T \hat{\mathbf{s}}_\ell[i]$ . At the end of the MB-SIC structure, the algorithm selects the branch with the minimum Euclidean distance according to

$$\ell_{opt} = \arg \min_{1 \leq \ell \leq S} \mathbf{C}(\ell) \quad (20)$$

where  $\mathbf{C}(\ell) = \|\mathbf{r}[i] - \mathbf{T}_\ell \mathbf{G} \tilde{\mathbf{s}}_\ell[i]\|$  is the Euclidean distance for the  $\ell$ -th branch. The final detected symbol vector is

$$\hat{s}_j[i] = Q(\mathbf{w}_{\ell_{opt},j}^H \mathbf{r}_{\ell_{opt},j}[i]), \quad j = 1, \dots, KN_U. \quad (21)$$

The MB-SIC algorithm can bring a close-to-optimal performance, however, the exhaustive search of  $S = K!$  branches is

not practical. Therefore, a reduced number of branches  $S$  must be employed. In terms of computational complexity, the MB-SIC algorithm requires  $S$  times the complexity of a standard SIC algorithm. However, it is possible to implement it using a multi-branch decision feedback structure [27], [42] that is equivalent in performance but which only requires a single matrix inversion as opposed to  $K$  matrix inversions required by the standard SIC algorithm and  $SK$  matrix inversions required by the MB-SIC algorithm.

DF detectors employ feedforward and feedback matrices that perform interference cancellation as described by

$$\hat{\mathbf{s}} = Q(\mathbf{W}^H \mathbf{r}[i] - \mathbf{F}^H \hat{\mathbf{s}}_o[i]), \quad (22)$$

where  $\hat{\mathbf{s}}_o$  corresponds to the initial decision vector that is usually performed by the linear section represented by  $\mathbf{W}$  of the DF receiver (e.g.,  $\hat{\mathbf{s}}_o = Q(\mathbf{W}^H \mathbf{r})$ ) prior to the application of the feedback section  $\mathbf{F}$ , which may have a strictly lower triangular structure for performing successive cancellation or zeros on the main diagonal when performing parallel cancellation. The receive filters  $\mathbf{W}$  and  $\mathbf{F}$  can be computed using various parameter estimation algorithms which will be discussed in the next section. Specifically, the receive filters can be based on the RMF, ZF and MMSE design criteria.

An often criticized aspect of these sub-optimal schemes is that they typically do not achieve the full receive-diversity order of the ML algorithm. This led to the investigation of detection strategies such as lattice-reduction (LR) schemes [30], [31], QR decomposition, M-algorithm (QRD-M) detectors [33], probabilistic data association (PDA) [34], [35], multi-branch [42], [45] detectors and likelihood ascent search techniques [48], [49], which can approach the ML performance at an acceptable cost for moderate to large systems. The development of cost-effective detection algorithms for massive MIMO systems is a challenging topic that calls for new approaches and ideas in this important research area.

## B. Iterative Detection and Decoding Techniques

Iterative detection and decoding (IDD) techniques have received considerable attention in the last years following the discovery of Turbo codes [50] and the application of the Turbo principle to interference mitigation [29], [50]–[58]. More recently, work on IDD schemes has been extended to low-density parity-check codes (LDPC) [54] and [57] and their extensions which compete with Turbo codes. The goal of an IDD system is to combine an efficient soft-input soft-output (SISO) detection algorithm and a SISO decoding technique as illustrated in Fig. 3. Specifically, the detector produces log-likelihood ratios (LLRs) associated with the encoded bits and these LLRs serve as input to the decoder. Then, in the second phase of the detection/decoding iteration, the decoder generates a posteriori probabilities (APPs) after a number of (inner) decoding iterations for encoded bits of each data stream. These APPs are fed to the detector to help in the next iterations between the detector and the decoder, which are called outer iterations. The joint process of detection/decoding is then repeated in an iterative manner until the maximum number of (inner and outer) iterations is reached. In mobile cellular networks, a designer can employ convolutional, Turbo

or LDPC codes in IDD schemes for interference mitigation. LDPC codes exhibit some advantages over Turbo codes that include simpler decoding and implementation issues. However, LDPC codes often require a higher number of decoding iterations which translate into delays or increased complexity. The development of IDD schemes and decoding algorithms that perform message passing with reduced delays are of fundamental importance in massive MIMO systems because they will be able to cope with audio and 3D video which are delay sensitive.

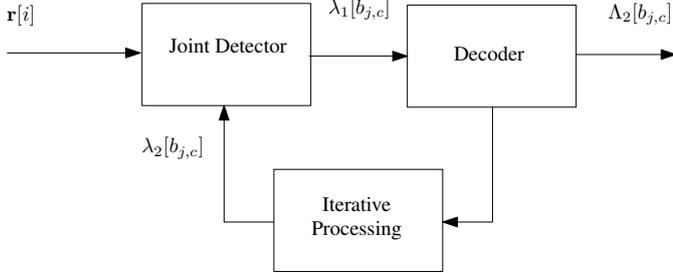


Fig. 3. Block diagram of an IDD scheme.

The massive MIMO systems described at the beginning of this chapter are considered here with convolutional codes and an iterative receiver structure consists of the following stages: A soft-input-soft-output (SISO) detector and a maximum *a posteriori* (MAP) decoder. Extensions to other channel codes are straightforward. These stages are separated by interleavers and deinterleavers. The soft outputs from the detector are used to estimate LLRs which are interleaved and serve as input to the MAP decoder for the convolutional code. The MAP decoder [52] computes *a posteriori* probabilities (APPs) for each stream's encoded symbols, which are used to generate soft estimates. These soft estimates are subsequently used to update the receive filters of the detector, de-interleaved and fed back through the feedback filter. The detector computes the *a posteriori* log-likelihood ratio (LLR) of a symbol (+1 or -1) for every code bit of each data stream in a packet with  $P$  symbols as given by

$$\Lambda_1[b_{j,c}[i]] = \log \frac{P[b_{j,c}[i] = +1 | \mathbf{r}[i]]}{P[b_{j,c}[i] = -1 | \mathbf{r}[i]]}, \quad (23)$$

$$j = 1, \dots, KN_U, \quad c = 1, \dots, C,$$

where  $C$  is the number of bits used to map the constellation. Using Bayes' rule, the above equation can be written as

$$\begin{aligned} \Lambda_1[b_{j,c}[i]] &= \log \frac{P[\mathbf{r}[i] | b_{j,c}[i] = +1]}{P[\mathbf{r}[i] | b_{j,c}[i] = -1]} + \log \frac{P[b_{j,c}[i] = +1]}{P[b_{j,c}[i] = -1]} \\ &= \lambda_1[b_{j,c}[i]] + \lambda_2^p[b_{j,c}[i]], \end{aligned} \quad (24)$$

where  $\lambda_2^p[b_{j,c}[i]] = \log \frac{P[b_{j,c}[i] = +1]}{P[b_{j,c}[i] = -1]}$  is the *a priori* LLR of the code bit  $b_{j,c}[i]$ , which is computed by the MAP decoder processing the  $j$ th data/user stream in the previous iteration, interleaved and then fed back to the detector. The superscript  $p$  denotes the quantity obtained in the previous iteration. Assuming equally likely bits, we have  $\lambda_2^p[b_{j,c}[i]] = 0$  in the first iteration for all streams/users. The quantity  $\lambda_1[b_{j,c}[i]] =$

$\log \frac{P[\mathbf{r}[i] | b_{j,c}[i] = +1]}{P[\mathbf{r}[i] | b_{j,c}[i] = -1]}$  represents the *extrinsic* information computed by the SISO detector based on the received data  $\mathbf{r}[i]$ , and the prior information about the code bits  $\lambda_2^p[b_{j,c}[i]]$ ,  $j = 1, \dots, KN_U$ ,  $c = 1, \dots, C$  and the  $i$ th data symbol. The extrinsic information  $\lambda_1[b_{j,c}[i]]$  is obtained from the detector and the prior information provided by the MAP decoder, which is de-interleaved and fed back into the MAP decoder of the  $j$ th data/user stream as the *a priori* information in the next iteration.

For the MAP decoding, we assume that the interference plus noise at the output  $z_j[i]$  of the receive filters is Gaussian. This assumption has been reported in previous works and provides an efficient and accurate way of computing the extrinsic information. Thus, for the  $j$ th stream/user and the  $q$ th iteration the soft output of the detector is

$$z_j^{(q)}[i] = V_j^{(q)} s_j[i] + \xi_j^{(q)}[i], \quad (25)$$

where  $V_j^{(q)}[i]$  is a scalar variable equivalent to the magnitude of the channel corresponding to the  $j$ th data stream and  $\xi_j^{(q)}[i]$  is a Gaussian random variable with variance  $\sigma_{\xi_j^{(q)}}^2[i]$ . Since we have

$$V_j^{(q)} = E[s_j^* z_j^{(q)}] \quad (26)$$

and

$$\sigma_{\xi_j^{(q)}}^2 = E[|z_j^{(q)}[i] - V_j^{(q)}[i] s_j[i]|^2], \quad (27)$$

the receiver can obtain the estimates  $\hat{V}_j^{(q)}$  and  $\hat{\sigma}_{\xi_j^{(q)}}^2$  via corresponding sample averages over the received symbols. These estimates are used to compute the *a posteriori* probabilities  $P[b_{j,c}[i] = \pm 1 | z_j^{(q)}[i]]$  which are de-interleaved and used as input to the MAP decoder. In what follows, it is assumed that the MAP decoder generates APPs  $P[b_{j,c}[i] = \pm 1]$ , which are used to compute the input to the receiver. From (25) the extrinsic information generated by the iterative receiver is given by

$$\begin{aligned} \lambda_1[b_{j,c}[i]] &= \log \frac{P[z_j^{(q)} | b_{j,c}[i] = +1]}{P[z_j^{(q)} | b_{j,c}[i] = -1]}, \quad i = 1, \dots, P, \\ &= \log \frac{\sum_{\mathbf{s} \in \mathbf{S}_c^{+1}} \exp\left(-\frac{|z_j^{(q)} - V_j^{(q)} \mathbf{S}|^2}{2\sigma_{\xi_j^{(q)}}^2}\right)}{\sum_{\mathbf{s} \in \mathbf{S}_c^{-1}} \exp\left(-\frac{|z_j^{(q)} - V_j^{(q)} \mathbf{S}|^2}{2\sigma_{\xi_j^{(q)}}^2}\right)}, \end{aligned} \quad (28)$$

where  $\mathbf{S}_c^{+1}$  and  $\mathbf{S}_c^{-1}$  are the sets of all possible constellations that a symbol can take on such that the  $c$ th bit is 1 and -1, respectively. Based on the trellis structure of the code, the MAP decoder processing the  $j$ th data stream computes the *a posteriori* LLR of each coded bit as described by

$$\begin{aligned} \Lambda_2[b_{j,c}[i]] &= \log \frac{P[b_{j,c}[i] = +1 | \lambda_1^p[b_{j,c}[i]; \text{decoding}]]}{P[b_{j,c}[i] = -1 | \lambda_1^p[b_{j,c}[i]; \text{decoding}]]} \\ &= \lambda_2[b_{j,c}[i]] + \lambda_1^p[b_{j,c}[i]], \end{aligned} \quad (29)$$

$$\text{for } j = 1, \dots, KN_U, \quad c = 1, \dots, C.$$

The computational burden can be significantly reduced using the max-log approximation. From the above, it can be seen that the output of the MAP decoder is the sum of the prior

information  $\lambda_1^p[b_{j,c}[i]]$  and the extrinsic information  $\lambda_2[b_{j,c}[i]]$  produced by the MAP decoder. This extrinsic information is the information about the coded bit  $b_{j,c}[i]$  obtained from the selected prior information about the other coded bits  $\lambda_1^p[b_{j,c}[i]]$ ,  $j \neq k$ . The MAP decoder also computes the *a posteriori* LLR of every information bit, which is used to make a decision on the decoded bit at the last iteration. After interleaving, the extrinsic information obtained by the MAP decoder  $\lambda_2[b_{j,c}[i]]$  for  $j = 1, \dots, KN_U$ ,  $c = 1, \dots, C$  is fed back to the detector, as prior information about the coded bits of all streams in the subsequent iteration. For the first iteration,  $\lambda_1[b_{j,c}[i]]$  and  $\lambda_2[b_{j,c}[i]]$  are statistically independent and as the iterations are computed they become more correlated and the improvement due to each iteration is gradually reduced. It is well known in the field of IDD schemes that there is no performance gain when using more than 5 – 8 iterations.

The choice of channel coding scheme is fundamental for the performance of iterative joint detection schemes. More sophisticated schemes than convolutional codes such as Turbo or LDPC codes can be considered in IDD schemes for the mitigation of multi-beam and other sources of interference. LDPC codes exhibit some advantages over Turbo codes that include simpler decoding and implementation issues. However, LDPC codes often require a higher number of decoding iterations which translate into delays or increased complexity. The development of IDD schemes and decoding algorithms that perform message passing with reduced delays [60]–[62] are of great importance in massive MIMO systems.

#### IV. PARAMETER ESTIMATION TECHNIQUES

Amongst the key problems in the uplink of multiuser massive MIMO systems are the estimation of parameters such as channels gains and receive filter coefficients of each user as described by the signal models in Section II. The parameter estimation task usually relies on pilot (or training) sequences, the structure of the data for blind estimation and signal processing algorithms. In multiuser massive MIMO networks, non-orthogonal training sequences are likely to be used in most application scenarios and the estimation algorithms must be able to provide the most accurate estimates and to track the variations due to mobility within a reduced training period. Standard MIMO linear MMSE and least-squares (LS) channel estimation algorithms [67] can be used for obtaining CSI. However, the cost associated with these algorithms is often cubic in the number of antenna elements at the receiver, i.e.,  $N_A$  in the uplink. Moreover, in scenarios with mobility the receiver will need to employ adaptive algorithms [66] which can track the channel variations. Interestingly, massive MIMO systems may have an excess of degrees of freedom that translates into a reduced-rank structure to perform parameter estimation. This is an excellent opportunity that massive MIMO offers to apply reduced-rank algorithms [72]– [82] and further develop these techniques. In this section, we review several parameter estimation algorithms and discuss several aspects that are specific for massive MIMO systems such as TDD operation, pilot contamination and the need for scalable estimation algorithms.

##### A. TDD operation

One of the key problems in modern wireless systems is the acquisition of CSI in a timely way. In time-varying channels, TDD offers the most suitable alternative to obtain CSI because the training requirements in a TDD system are independent of the number of antennas at the base station (or access point) [6], [63], [64] and there is no need for CSI feedback. In particular, TDD systems rely on reciprocity by which the uplink channel estimate is used as an estimate of the downlink channel. An issue in this operation mode is the difference in the transfer characteristics of the amplifiers and the filters in the two directions. This can be addressed through measurements and appropriate calibration [65]. In contrast, in a frequency division duplexing (FDD) system the training requirements is proportional to the number of antennas and CSI feedback is essential. For this reason, massive MIMO systems will most likely operate in TDD mode and will require further investigation in calibration methods.

##### B. Pilot contamination

The adoption of TDD mode and uplink training in massive MIMO systems with multiple cells results in a phenomenon called pilot contamination. In multi-cell scenarios, it is difficult to employ orthogonal pilot sequences because the duration of the pilot sequences depends on the number of cells and this duration is severely limited by the channel coherence time due to mobility. Therefore, non-orthogonal pilot sequences must be employed and this affects the CSI employed at the transmitter. Specifically, the channel estimate is contaminated by a linear combination of channels of other users that share the same pilot [63], [64]. Consequently, the detectors, precoders and resource allocation algorithms will be highly affected by the contaminated CSI. Strategies to control or mitigate pilot contamination and its effects are very important for massive MIMO networks.

##### C. Estimation of Channel Parameters

Let us now consider channel estimation techniques for multiuser Massive MIMO systems and employ the signal models of Section II. The channel estimation problem corresponds to solving the following least-squares (LS) optimization problem:

$$\hat{\mathbf{G}}[i] = \arg \min_{\mathbf{G}[i]} \sum_{l=1}^i \lambda^{i-l} \|\mathbf{r}[l] - \mathbf{G}[i]\mathbf{s}[l]\|^2, \quad (30)$$

where the  $N_A \times KN_U$  matrix  $\mathbf{G} = [\mathbf{G}_1 \dots \mathbf{G}_K]$  contains the channel parameters of the  $K$  users, the  $KN_U \times 1$  vector contains the symbols of the  $K$  users stacked and  $\lambda$  is a forgetting factor chosen between 0 and 1. In particular, it is common to use known pilot symbols  $\mathbf{s}[i]$  in the beginning of the transmission for estimation of the channels and the other receive parameters. This problem can be solved by computing the gradient terms of (30), equating them to a zero matrix and manipulating the terms which yields the LS estimate

$$\hat{\mathbf{G}}[i] = \mathbf{Q}[i]\mathbf{R}^{-1}[i], \quad (31)$$

where  $\mathbf{Q}[i] = \sum_{l=1}^i \lambda^{i-l} \mathbf{r}[l]\mathbf{s}^H[l]$  is a  $N_A \times KNU$  matrix with estimates of cross-correlations between the pilots and

the received data  $\mathbf{r}[i]$  and  $\mathbf{R}[i] = \sum_{l=1}^i \lambda^{i-l} \mathbf{s}[l] \mathbf{s}^H[l]$  is an estimate of the auto-correlation matrix of the pilots. When the channel is static over the duration of the transmission, it is common to set the forgetting factor  $\lambda$  to one. In contrast, when the channel is time-varying one needs to set  $\lambda$  to a value that corresponds to the coherence time of the channel in order to track the channel variations.

The LS estimate of the channel can also be computed recursively by using the matrix inversion lemma [66], [67], which yields the recursive LS (RLS) channel estimation algorithm [27] described by

$$\mathbf{P}[i] = \lambda^{-1} \mathbf{P}[i-1] - \frac{\lambda^{-2} \mathbf{P}[i-1] \mathbf{s}[i] \mathbf{s}^H[i] \mathbf{P}[i-1]}{1 + \lambda^{-1} \mathbf{s}^H[i] \mathbf{P}[i-1] \mathbf{s}[i]} \mathbf{s}[i], \quad (32)$$

$$\mathbf{T}[i] = \lambda \mathbf{T}[i-1] + \mathbf{r}[i] \mathbf{s}^H[i], \quad (33)$$

$$\hat{\mathbf{G}}[i] = \mathbf{T}[i] \mathbf{P}[i], \quad (34)$$

where the computational complexity of the RLS channel estimation algorithm is  $N_A(KN_U)^2 + 4(KN_U)^2 + 2N_A(KN_U) + 2KN_U + 2$  multiplications and  $N_A(KN_U)^2 + 4(KN_U)^2 - KN_U$  additions [27].

An alternative to using LS-based algorithms is to employ least-mean square (LMS) techniques [68], which can reduce the computational cost. Consider the mean-square error (MSE)-based optimization problem:

$$\hat{\mathbf{G}}[i] = \arg \min_{\mathbf{G}[i]} E \|\mathbf{r}[i] - \mathbf{G}[i] \mathbf{s}[i]\|^2, \quad (35)$$

where  $E[\cdot]$  stands for expected value. This problem can be solved by computing the instantaneous gradient terms of (35), using a gradient descent rule and manipulating the terms which results in the LMS channel estimation algorithm given by

$$\hat{\mathbf{G}}[i+1] = \hat{\mathbf{G}}[i] + \mu \mathbf{e}[i] \mathbf{s}^H[i], \quad (36)$$

where the error vector signal is  $\mathbf{e}[i] = \mathbf{r}[i] - \hat{\mathbf{G}}[i] \mathbf{s}[i]$  and the step size  $\mu$  should be chosen between 0 and  $2/\text{tr}[\mathbf{R}]$  [66]. The cost of the LMS channel estimation algorithm in this scheme is  $N_A(KN_U)^2 + N_A(KN_U) + KN_U$  multiplications and  $N_A(KN_U)^2 + N_A KN_U + N_A - KN_U$  additions. The LMS approach has a cost that is one order of magnitude lower than the RLS but the performance in terms of training speed is worse. The channel estimates obtained can be used in the ML rule for ML detectors and SD algorithms, and also to design the receive filters of ZF and MMSE type detectors outlined in the previous section.

#### D. Estimation of Receive Filter Parameters

An alternative to channel estimation techniques is the direct computation of the receive filters using LS techniques or adaptive algorithms. In this subsection, we consider the estimation of the receive filters for multiuser Massive MIMO systems and employ again the signal models of Section II. The receive filter estimation problem corresponds to solving the LS optimization problem described by

$$\mathbf{w}_{k,o}[i] = \arg \min_{\mathbf{w}_k[i]} \sum_{l=1}^i \lambda^{i-l} |s_k[l] - \mathbf{w}_k^H[i] \mathbf{r}[l]|^2, \quad (37)$$

where the  $N_A \times 1$  vector  $\mathbf{w}_k$  contains the parameters of the receive filters for the  $k$ th data stream, the symbol  $s_k[i]$  contains the symbols of the  $k$ th data stream. Similarly to channel estimation, it is common to use known pilot symbols  $s_k[i]$  in the beginning of the transmission for estimation of the receiver filters. This problem can be solved by computing the gradient terms of (37), equating them to a null vector and manipulating the terms which yields the LS estimate

$$\mathbf{w}_{k,o}[i] = \mathbf{R}_r^{-1}[i] \mathbf{p}_k[i], \quad (38)$$

where  $\mathbf{R}_r[i] = \sum_{l=1}^i \lambda^{i-l} \mathbf{r}[l] \mathbf{r}^H[l]$  is the auto-correlation matrix of the received data and  $\mathbf{p}_k[i] = \sum_{l=1}^i \lambda^{i-l} \mathbf{r}[l] s_k^H[l]$  is a  $N_A \times 1$  vector with cross-correlations between the pilots and the received data  $\mathbf{r}[i]$ . When the channel is static over the duration of the transmission, it is common to set the forgetting factor  $\lambda$  to one. Conversely, when the channel is time-varying one needs to set  $\lambda$  to a value that corresponds to the coherence time of the channel in order to track the channel variations. In these situations, a designer can also compute the parameters recursively, thereby taking advantage of the previously computed LS estimates and leading to the RLS algorithm [66] given by

$$\mathbf{k}[i] = \frac{\lambda^{-1} \mathbf{P}[i-1] \mathbf{r}[i]}{1 + \lambda^{-1} \mathbf{r}^H[i] \mathbf{P}[i-1] \mathbf{r}[i]}, \quad (39)$$

$$\mathbf{P}[i] = \lambda^{-1} \mathbf{P}[i-1] - \lambda^{-1} \mathbf{k}[i] \mathbf{r}^H[i] \mathbf{P}[i-1], \quad (40)$$

$$\mathbf{w}_k[i] = \mathbf{w}_k[i-1] - \mathbf{k}[i] e_{k,a}^*[i], \quad (41)$$

where  $e_{k,a}[i] = s_k[i] - \mathbf{w}_k^H[i-1] \mathbf{r}[i]$  is the *a priori* error signal for the  $k$ th data stream. Several other variants of the RLS algorithm could be used to compute the parameters of the receive filters [69]. The computational cost of this RLS algorithm for all data streams corresponds to  $KN_U(3N_A^2 + 4N_A + 1)$  multiplications and  $KN_U(3N_A^2 + 2N_A - 1) + 2N_A KN_U$  additions

A reduced complexity alternative to the RLS algorithms is to employ the LMS algorithm to estimate the parameters of the receive filters. Consider the mean-square error (MSE)-based optimization problem:

$$\mathbf{w}_{k,o}[i] = \arg \min_{\mathbf{w}_k[i]} E \|\mathbf{s}_k[i] - \mathbf{w}_k^H[i] \mathbf{r}[i]\|^2, \quad (42)$$

Similarly to the case of channel estimation, this problem can be solved by computing the instantaneous gradient terms of (42), using a gradient descent rule and manipulating the terms which results in the LMS estimation algorithm given by

$$\hat{\mathbf{w}}_k[i+1] = \hat{\mathbf{w}}_k[i] + \mu e_k^*[i] \mathbf{r}[i], \quad (43)$$

where the error signal for the  $k$ th data stream is  $e_k[i] = s_k[i] - \hat{\mathbf{w}}_k^H[i] \mathbf{r}[i]$  and the step size  $\mu$  should be chosen between 0 and  $2/\text{tr}[\mathbf{R}]$  [66]. The cost of the LMS estimation algorithm in this scheme is  $KN_U(N_A + 1)$  multiplications and  $KN_U N_A$  additions.

In parameter estimation problems with a large number of parameters such as those found in massive MIMO systems, an effective technique is to employ reduced-rank algorithms which perform dimensionality reduction followed by parameter estimation with a reduced number of parameters. Consider

the mean-square error (MSE)-based optimization problem:

$$[\bar{\mathbf{w}}_{k,o}[i], \mathbf{T}_{D,k,o}[i]] = \arg \min_{\bar{\mathbf{w}}_k[i], \mathbf{T}_{D,k}} E[|s_k[i] - \bar{\mathbf{w}}_k^H[i] \mathbf{T}_{D,k}^H[i] \mathbf{r}[i]|^2] \quad (44)$$

where  $\mathbf{T}_{D,k}[i]$  is an  $N_A \times D$  matrix that performs dimensionality reduction and  $\bar{\mathbf{w}}_k[i]$  is a  $D \times 1$  parameter vector. Given  $\mathbf{T}_{D,k}[i]$ , a generic reduced-rank RLS algorithm [82] with  $D$ -dimensional quantities can be obtained from (39)-(41) by substituting the  $N_A \times 1$  received vector  $\mathbf{r}[i]$  by the reduced-dimension  $D \times 1$  vector  $\bar{\mathbf{r}}[i] = \mathbf{T}_{D,k}^H[i] \mathbf{r}[i]$ .

A central design problem is how to compute the dimensionality reduction matrix  $\mathbf{T}_{D,k}[i]$  and several techniques have been considered in the literature, namely:

- Principal components (PC):  $\mathbf{T}_{D,k}[i] = \phi_D[i]$ , where  $\phi_D[i]$  corresponds to a unitary matrix whose columns are the  $D$  eigenvectors corresponding to the  $D$  largest eigenvectors of an estimate of the covariance matrix  $\hat{\mathbf{R}}[i]$ .
- Krylov subspace techniques:  $\mathbf{T}_{D,k}[i] = [\mathbf{t}_k[i] \hat{\mathbf{R}}[i] \mathbf{t}_k[i] \dots \hat{\mathbf{R}}^{D-1}[i] \mathbf{t}_k[i]]$ , where  $\mathbf{t}_k[i] = \frac{\mathbf{t}_k[i]}{\|\mathbf{t}_k[i]\|}$ , for  $k = 1, 2, \dots, D$  correspond to the bases of the Krylov subspace [70]- [74].
- Joint iterative optimization methods:  $\mathbf{T}_{D,k}[i]$  is estimated along with  $\bar{\mathbf{w}}_k[i]$  using an alternating optimization strategy and adaptive algorithms [75]- [83].

## V. SIMULATION RESULTS

In this section, we illustrate some of the techniques outlined in this article using massive MIMO configurations, namely, a very large antenna array, an excess of degrees of freedom provided by the array and a large number of users with multiple antennas. We consider QPSK modulation, data packets of 1500 symbols and channels that are fixed during each data packet and that are modeled by complex Gaussian random variables with zero mean and variance equal to unity. For coded systems and iterative detection and decoding, a non-recursive convolutional code with rate  $R = 1/2$ , constraint length 3, generator polynomial  $g = [7 \ 5]_{\text{oct}}$  and 4 decoding iterations is adopted. The numerical results are averaged over  $10^6$  runs. For the CAS configuration, we employ  $L_k = 0.7$ ,  $\tau = 2$ , the distance  $d_k$  to the BS is obtained from a uniform discrete random variable between 0.1 and 0.95, the shadowing spread is  $\sigma_k = 3$  dB and the transmit and receive correlation coefficients are equal to  $\rho = 0.2$ . The signal-to-noise ratio (SNR) in dB per receive antenna is given by  $\text{SNR} = 10 \log_{10} \frac{KN_U \sigma_{s_r}^2}{RC \sigma_n^2}$ , where  $\sigma_{s_r}^2 = \sigma_s^2 E[|\gamma_k|^2]$  is the variance of the received symbols,  $\sigma_n^2$  is the noise variance,  $R < 1$  is the rate of the channel code and  $C$  is the number of bits used to represent the constellation. For the DAS configuration, we use  $L_{k,j}$  taken from a uniform random variable between 0.7 and 1,  $\tau = 2$ , the distance  $d_{k,j}$  for each link to an antenna is obtained from a uniform discrete random variable between 0.1 and 0.5, the shadowing spread is  $\sigma_{k,j} = 3$  dB and the transmit and receive correlation coefficients for the antennas that are co-located are equal to  $\rho = 0.2$ . The signal-to-noise ratio (SNR) in dB per receive antenna for the DAS configuration is given by  $\text{SNR} = 10 \log_{10} \frac{KN_U \sigma_{s_r}^2}{RC \sigma_n^2}$ , where  $\sigma_{s_r}^2 = \sigma_s^2 E[|\gamma_{k,j}|^2]$  is the variance of the received symbols.

In the first example, we compare the BER performance against the SNR of several detection algorithms, namely, the RMF with multiple users and with a single user denoted as single user bound, the linear MMSE detector [17], the SIC-MMSE detector using a successive interference cancellation [26] and the multi-branch SIC-MMSE (MB-SIC-MMSE) detector [27], [42], [45]. We assume perfect channel state information and synchronization. In particular, a scenario with  $N_A = 64$  antenna elements at the receiver,  $K = 32$  users and  $N_U = 2$  antenna elements at the user devices is considered, which corresponds to a scenario without an excess of degrees of freedom with  $N_A \approx KN_U$ . The results shown in Fig. 4 indicate that the RMF with a single user has the best performance, followed by the MB-SIC-MMSE, the SIC-MMSE, the linear MMSE and the RMF detectors. Unlike previous works [6] that advocate the use of the RMF, it is clear that the BER performance loss experienced by the RMF should be avoided and more advanced receivers should be considered. However, the cost of linear and SIC receivers is dictated by the matrix inversion of  $N_A \times N_A$  matrices which must be reduced for large systems. Moreover, it is clear that a DAS configuration is able to offer a superior BER performance due to a reduction of the average distance from the users to the receive antennas and a reduced correlation amongst the set of  $N_a$  receive antennas, resulting in improved links.

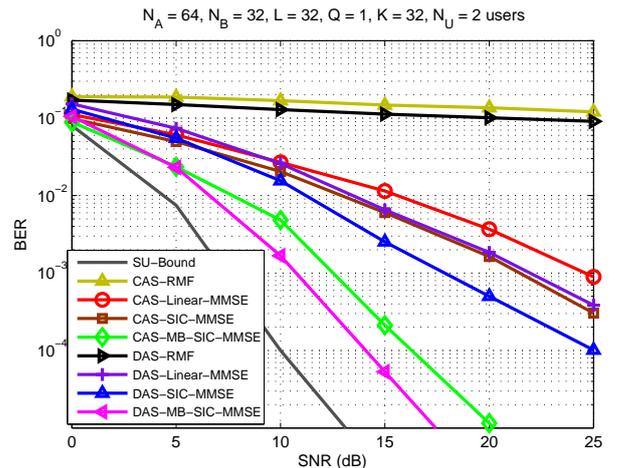


Fig. 4. BER performance against SNR of detection algorithms in a scenario with  $N_A = 64$ ,  $N_B = 32$ ,  $L = 32$ ,  $Q = 1$ ,  $K = 32$  users and  $N_U = 2$  antenna elements.

In the second example, we consider the coded BER performance against the SNR of several detection algorithms with a DAS configuration using perfect channel state information, as illustrated in Fig. 5. The results show that the BER is much lower than that obtained for an uncoded systems as indicated in Fig. 4. Specifically, the MB-SIC-MMSE algorithm obtains the best performance followed by the SIC-MMSE, the linear MMSE and the RMF techniques. Techniques like the MB-SIC-MMSE and SIC-MMSE are more promising for systems with a large number of antennas and users as they can operate with lower SNR values and are therefore more energy efficient. Interestingly, the RMF can offer a BER performance that is acceptable when operating with a high SNR that is not energy

efficient and has the advantage that it does not require a matrix inversion. If a designer chooses stronger channel codes like Turbo and LDPC techniques, this choice might allow the operation of the system at lower SNR values.

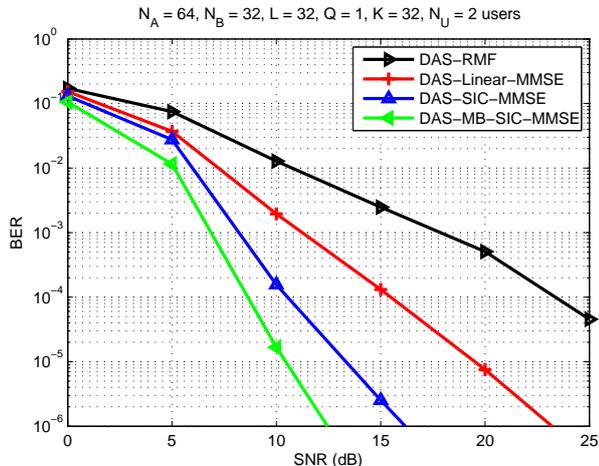


Fig. 5. Coded BER performance against SNR of detection algorithms with DAS in a scenario with  $N_A = 64$ ,  $N_B = 32$ ,  $L = 32$ ,  $Q = 1$ ,  $K = 32$  users,  $N_U = 2$  antenna elements and 4 iterations.

In the third example, we assess the estimation algorithms when applied to the analyzed detectors. In particular, we compare the BER performance against the SNR of several detection algorithms with a DAS configuration using perfect channel state information and estimated channels with the RLS and the LMS algorithms. The channels are estimated with 250 pilot symbols which are sent at the beginning of packets with 1500 symbols. The results shown in Fig. 6 indicate that the performance loss caused by the use of the estimated channels is not significant as it remains within 1-2 dB for the same BER performance. The main problems of the use of the standard RLS and LMS is that they require a reasonably large number of pilot symbols to obtain accurate estimates of the channels, resulting in reduced transmission efficiency.

In the fourth example, we evaluate the more sophisticated reduced-rank estimation algorithms to reduce the number of pilot symbols for the training of the receiver filters. In particular, we compare the BER performance against the number of received symbols for a SIC type receiver using a DAS configuration and the standard RLS [66], the Krylov-RLS [?] and JIO-RLS [82] and the JIDF-RLS [80] algorithms. We provide the algorithms pilots for the adjustment of the receive filters and assess the BER convergence performance. The results shown in Fig. 7 illustrate that the performance of the reduced-rank algorithms is significantly better than the standard RLS algorithm, indicating that the use of reduced-rank algorithms can reduce the need for pilot symbols. Specifically, the best performance is obtained by the JIDF-RLS algorithm, followed by the JIO-RLS, the Krylov-RLS and the standard RLS techniques. In particular, the reduced-rank algorithms can obtain a performance comparable to the standard RLS algorithm with a fraction of the number of pilot symbols required by the RLS algorithm. It should be remarked that for larger training periods the standard RLS algorithm will

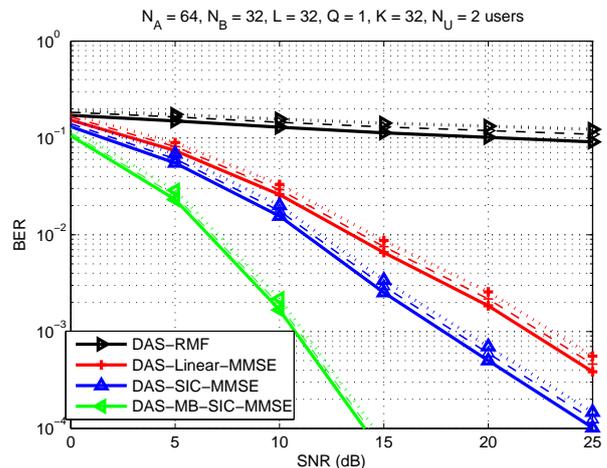


Fig. 6. BER performance against SNR of detection algorithms in a scenario with channel estimation,  $N_A = 64$ ,  $N_B = 32$ ,  $L = 32$ ,  $Q = 1$ ,  $K = 32$  users and  $N_U = 2$  antenna elements. Parameters:  $\lambda = 0.999$  and  $\mu = 0.05$ . The solid lines correspond to perfect channel state information, the dashed lines correspond to channel estimation with the RLS algorithm and the dotted lines correspond to channel estimation with the LMS algorithm.

converge to the MMSE bound and the reduced-rank algorithms might converge to the MMSE bound or to higher MSE values depending on the structure of the covariance matrix  $\mathbf{R}$  and the choice of the rank  $D$ .

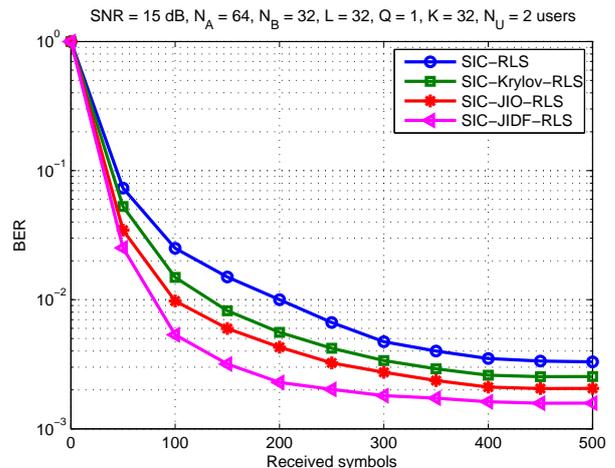


Fig. 7. BER performance against the number of received symbols for SIC receivers with a DAS architecture operating at SNR = 15 dB in a scenario with the estimation of the receive filters,  $N_A = 64$ ,  $N_B = 32$ ,  $L = 32$ ,  $Q = 1$ ,  $K = 32$  users and  $N_U = 2$  antenna elements. Parameters:  $\lambda = 0.999$ ,  $D = 5$  for all reduced-rank methods, and interpolators with  $I = 3$  parameters and 12 branches for the JIDF scheme.

## VI. FUTURE TRENDS AND EMERGING TOPICS

In this section, we discuss some future signal detection and estimation trends in the area of massive MIMO systems and point out some topics that might attract the interest of researchers. The topics are structured as:

- Signal detection:

→ Cost-effective detection algorithms: Techniques to perform dimensionality reduction [71]- [83] for detection problems will play an important role in massive MIMO devices. By reducing the number of effective processing elements, detection algorithms could be applied. In addition, the development of schemes based on RMF with non-linear interference cancellation capabilities might be a promising option that can close the complexity gap between RMF and more costly detectors.

→ Decoding strategies with low delay: The development of decoding strategies for DAS configurations with reduced delay will play a key role in applications such as audio and video streaming because of their delay sensitivity. Therefore, we novel message passing algorithms with smarter strategies to exchange information should be investigated along with their application to IDD schemes [60]–[62].

→ Mitigation of impairments: The identification of impairments originated in the RF chains of massive MIMO systems, delays caused by DAS schemes will need mitigation by smart signal processing algorithms. For example, I/Q imbalance might be dealt with using widely-linear signal processing algorithms [74], [87], [88].

→ Detection techniques for multicell scenarios: The development of detection algorithms for scenarios with multiple and small cells requires approaches which minimize the need for channel state information from adjacent cells and the decoding delay [84]- [86].

- Parameter estimation:

→ Blind algorithms: The development of blind estimation algorithms for the channel and receive filter parameters is important for mitigating the problem of pilot contamination [89]- [94].

→ Reduced-rank and sparsity-aware algorithms: the development of reduced-rank and sparsity-aware algorithms that exploit the mathematical structure of massive MIMO channels is an important topic for the future along with features that lend themselves to implementation [70]- [83].

## VII. CONCLUDING REMARKS

This chapter has presented signal detection and estimation techniques for multiuser massive MIMO systems. We consider the application to cellular networks with massive MIMO along with CAS and DAS configurations. Recent signal detection algorithms have been discussed and their use with iteration detection and decoding schemes has been considered. Parameter estimation algorithms have also been reviewed and studied in several scenarios of interest. Numerical results have illustrated some of the discussions on signal detection and estimation techniques along with future trends in the field.

## REFERENCES

- [1] Cisco and/or its affiliates, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017", Tech. Rep., Cisco Systems, Inc., Jan. 2013.
- [2] Requirements for Further Advancements for E-UTRA (LTE-Advanced), 3GPP TR 36.913 Standard, 2011.
- [3] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Enhancements for Very High Throughput for Operation in Bands Below 6GHz, IEEE P802.11ac/D1.0 Standard., Jan. 2011.
- [4] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong and J. Yao, "5G on the Horizon", IEEE Vehicular Technology Magazine, September 2013.
- [5] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [6] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors and F. Tufvesson, "Scaling up MIMO: Opportunities, and challenges with very large arrays", IEEE Signal Processing Mag., vol. 30, no. 1, pp. 40-60, Jan. 2013.
- [7] R. C. de Lamare, "Massive MIMO Systems: Signal Processing Challenges and Future Trends", URSI Radio Science Bulletin, December 2013.
- [8] J. Nam, J.-Y. Ahn, A. Adhikary, and G. Caire, "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information," in 46th Annual Conference on Information Sciences and Systems (CISS), 2012.
- [9] R. Combes, Z. Altman and E. Altman, "Interference coordination in wireless networks: A flow-level perspective," Proceedings IEEE INFOCOM 2013, vol., no., pp.2841,2849, 14-19 April 2013.
- [10] R. Aggarwal, C. E. Koksal, and P. Schniter, "On the design of large scale wireless systems", IEEE J. Sel. Areas Commun, vol. 31, no. 2, pp. 215-225, Feb. 2013.
- [11] C. Shepard, H. Yu, N. Anand, L. E. Li, T. L. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in ACM Int. Conf. Mobile Computing and Networking (MobiCom), Istanbul, Turkey, Aug. 2012.
- [12] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, "Channel measurements for large antenna arrays," in IEEE International Symposium on Wireless Communication Systems (ISWCS), Paris, France, Aug. 2012.
- [13] X. Gao, F. Tufvesson, O. Edfors, and F. Rusek, "Measured propagation characteristics for very-large MIMO at 2.6 GHz," in Proc. of the 46th Annual Asilomar Conference on Signals, Systems, and Computers., Pacific Grove, California, USA, Nov. 2012.
- [14] J.P. Kermoal, L. Schumacher, and K.I. Pedersen et al., "A stochastic MIMO radio channel model with experimental validation," IEEE J. Sel. Areas Commun., vol. 20, no. 6, pp. 1211 – 1226, Aug. 2002.
- [15] W. Choi, J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," IEEE Transactions on Wireless Communications, vol.6, no.1, pp.69,73, Jan. 2007.
- [16] L. Dai, "An Uplink Capacity Analysis of the Distributed Antenna System (DAS): From Cellular DAS to DAS with Virtual Cells," IEEE Transactions on Wireless Communications, vol.13, no.5, pp.2717,2731, May 2014.
- [17] S. Verdú, *Multiuser Detection*, Cambridge, 1998.
- [18] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels", IEEE Trans. on Inf. Theory, vol. 45, no. 5, pp.1639–1642, July 1999.
- [19] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum likelihood detection and the search for the closest lattice point," IEEE Trans. Inform. Theory, vol. 49, pp. 2389–2402, Oct. 2003.
- [20] Z. Guo and P. Nilsson, "Algorithm and Implementation of the K-Best Sphere Decoding for MIMO Detection," IEEE Journal on Selected Areas in Communications, vol. 24, no. 3, pp. 491–503, March 2006.
- [21] C. Studer, A. Burg, and H. Bolcskei, Soft-output sphere decoding: algorithms and VLSI implementation," IEEE J. Sel. Areas Commun., vol. 26, pp. 290-300, Feb. 2008.
- [22] B. Shim and I. Kang, "On further reduction of complexity in tree pruning based sphere search," IEEE Trans. Commun., vol. 58, no. 2, pp. 417–422, Feb. 2010.
- [23] A. Duel-Hallen, "Equalizers for Multiple Input Multiple Output Channels and PAM Systems with Cyclostationary Input Sequences," IEEE J. Select. Areas Commun., vol. 10, pp. 630-639, April, 1992.
- [24] G. D. Golden, C. J. Foschini, R. A. Valenzuela and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture", Electronics Letters, vol. 35, No.1, January 1999.

- [25] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Trans. Signal Processing*, vol. 51, pp. 1722-1730, July 2003.
- [26] A. Rontogiannis, V. Kekatos, and K. Berberidis, "A Square-Root Adaptive V-BLAST Algorithm for Fast Time-Varying MIMO Channels," *IEEE Signal Processing Letters*, Vol. 13, No. 5, pp. 265-268, May 2006.
- [27] R. Fa, R. C. de Lamare, "Multi-Branch Successive Interference Cancellation for MIMO Spatial Multiplexing Systems", *IET Communications*, vol. 5, no. 4, pp. 484 - 494, March 2011.
- [28] P. Li, R. C. de Lamare and R. Fa, "Multiple Feedback Successive Interference Cancellation Detection for Multiuser MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2434 - 2439, August 2011.
- [29] J. H. Choi, H. Y. Yu, Y. H. Lee, "Adaptive MIMO decision feedback equalization for receivers with time-varying channels", *IEEE Trans. Signal Proc.*, 2005, 53, no. 11, pp. 4295-4303.
- [30] C. Windpassinger, L. Lampe, R.F.H. Fischer, T.A Hehn, "A performance study of MIMO detectors," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, August 2006, pp. 2004-2008.
- [31] Y. H. Gan, C. Ling, and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Trans. Signal Processing*, vol. 56, no. 7, July 2009.
- [32] Q. Zhou and X. Ma, Element-Based Lattice Reduction Algorithms for Large MIMO Detection, *IEEE J. Sel. Areas Commun*, vol. 31, no. 2, pp. 274-286, Feb. 2013.
- [33] K. J. Kim, J. Yue, R. A. Iltis, and J. D. Gibson, "A QRD-M/Kalman filter-based detection and channel estimation algorithm for MIMO-OFDM systems", *IEEE Trans. Wireless Communications*, vol. 4, pp. 710-721, March 2005.
- [34] Y. Jia, C. M. Vithanage, C. Andrieu, and R. J. Piechocki, "Probabilistic data association for symbol detection in MIMO systems," *Electron. Lett.*, vol. 42, no. 1, pp. 38-40, Jan. 2006.
- [35] S. Yang, T. Lv, R. Maunder, and L. Hanzo, "Unified Bit-Based Probabilistic Data Association Aided MIMO Detection for High-Order QAM Constellations", *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 981-991, 2011.
- [36] M. K. Varanasi, "Decision feedback multiuser detection: A systematic approach," *IEEE Trans. on Inf. Theory*, vol. 45, pp. 219-240, January 1999.
- [37] J. F. Rößler and J. B. Huber, "Iterative soft decision interference cancellation receivers for DS-CDMA downlink employing 4QAM and 16QAM," in *Proc. 36th Asilomar Conf. Signal, Systems and Computers*, Pacific Grove, CA, Nov. 2002.
- [38] J. Luo, K. R. Pattipati, P. K. Willet and F. Hasegawa, "Optimal User Ordering and Time Labeling for Ideal Decision Feedback Detection in Asynchronous CDMA", *IEEE Trans. on Communications*, vol. 51, no. 11, November, 2003.
- [39] G. Woodward, R. Ratasuk, M. L. Honig and P. Rapajic, "Minimum Mean-Squared Error Multiuser Decision-Feedback Detectors for DS-CDMA," *IEEE Trans. on Communications*, vol. 50, no. 12, December, 2002.
- [40] R.C. de Lamare, R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels", *IEEE Communications Letters*, vol. 7, no. 2, Feb. 2003, pp. 73 - 75.
- [41] F. Cao, J. Li, and J. Yang, "On the relation between PDA and MMSE-ISDIC," *IEEE Signal Processing Letters*, vol. 14, no. 9, Sep. 2007.
- [42] R.C. de Lamare, R. Sampaio-Neto, "Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA systems", *IEEE Trans. Commun.*, vol. 56, no. 5, May 2008, pp. 778-789.
- [43] Y. Cai and R. C. de Lamare, "Adaptive Space-Time Decision Feedback Detectors with Multiple Feedback Cancellation", *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, October 2009, pp. 4129 - 4140.
- [44] P. Li and R. C. de Lamare, "Adaptive Decision-Feedback Detection With Constellation Constraints for MIMO Systems", *IEEE Transactions on Vehicular Technology*, vol. 61, no. 2, 853-859, 2012.
- [45] R. C. de Lamare, "Adaptive and Iterative Multi-Branch MMSE Decision Feedback Detection Algorithms for Multi-Antenna Systems", *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, February 2013.
- [46] M. Reuter, J.C. Allen, J. R. Zeidler, R. C. North, "Mitigating error propagation effects in a decision feedback equalizer", *IEEE Transactions on Communications*, vol. 49, no. 11, November 2001, pp. 2028 - 2041.
- [47] R.C. de Lamare, R. Sampaio-Neto, A. Hjørungnes, "Joint iterative interference cancellation and parameter estimation for cdma systems", *IEEE Communications Letters*, vol. 11, no. 12, December 2007, pp. 916 - 918.
- [48] K. Vardhan, S. Mohammed, A. Chockalingam and B. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems", *IEEE J. Sel. Commun.*, vol. 26, no. 3, pp. 473-485, Apr. 2008.
- [49] P. Li and R. D. Murch, Multiple Output Selection-LAS Algorithm in Large MIMO Systems, *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 399-401, May 2010.
- [50] C. Berrou and A. Glavieux, "Near optimum error-correcting coding and decoding: Turbo codes," *IEEE Trans. Commun.*, vol. 44, Oct. 1996.
- [51] C. Douillard et al., "Iterative correction of intersymbol interference: Turbo equalization," *European Trans. Telecommun.*, vol. 6, no. 5, pp. 507-511, Sept.-Oct. 1995.
- [52] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046-1061, July 1999.
- [53] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean square error equalization using a priori information," *IEEE Trans. Signal Processing*, vol. 50, pp. 673-683, Mar. 2002.
- [54] B. Hochwald and S. ten Brink, Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389-399, Mar. 2003.
- [55] J. Hou, P. H. Siegel, L. B. Milstein, "Design of multi-input multi-output systems based on low-density Parity-check codes," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 601- 611, April 2005.
- [56] H. Lee, B. Lee, and I. Lee, "Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM Systems," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 504-513, Mar. 2006.
- [57] J. Wu, H.-N. Lee, "Performance Analysis for LDPC-Coded Modulation in MIMO Multiple-Access Systems," *IEEE Transactions on Communications*, vol. 55, no. 7, pp. 1417-1426, July 2007
- [58] X. Yuan, Q. Guo, X. Wang, and Li Ping, "Evolution analysis of low-cost iterative equalization in coded linear systems with cyclic prefixes," *IEEE J. Select. Areas Commun. (JSAC)*, vol. 26, no. 2, pp. 301-310, Feb. 2008.
- [59] J. W. Choi, A. C. Singer, J. Lee, N. I. Cho, "Improved linear soft-input soft-output detection via soft feedback successive interference cancellation," *IEEE Trans. Commun.*, vol.58, no.3, pp.986-996, March 2010.
- [60] M. J. Wainwright, T. S. Jaakkola, and A.S. Willsky, "A new class of upper bounds on the log partition function," *IEEE Trans. Information Theory*, vol. 51, no. 7, pp. 2313 - 2335, July 2005.
- [61] H. Wymeersch, F. Penna and V. Savic, "Uniformly Reweighted Belief Propagation for Estimation and Detection in Wireless Networks," *IEEE Trans. Wireless Communications*, vol. PP, No. 99, pp. 1-9, Feb. 2012.
- [62] J. Liu, R. C. de Lamare, "Low-Latency Reweighted Belief Propagation Decoding for LDPC Codes," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1660-1663, October 2012.
- [63] J. Jose, A. Ashikhmin, T. L. Marzetta, S. Vishwanath, "Pilot Contamination and Precoding in Multi-Cell TDD Systems," *IEEE Transactions on Wireless Communications*, vol.10, no.8, pp. 2640-2651, August 2011.
- [64] A. Ashikhmin and T. L. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, Jul. 2012.
- [65] R. Rogalin, O. Y. Bursalioğlu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, K. Psounis, "Scalable Synchronization and Reciprocity Calibration for Distributed Multiuser MIMO," *IEEE Transactions on Wireless Communications*, vol.13, no.4, pp.1815-1831, April 2014.
- [66] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice- Hall, 2002.
- [67] M. Biguesh, A.B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals", *IEEE Transactions on Signal Processing*, vol. 54 no. 3, March 2006.
- [68] T. Wang, R. C. de Lamare, P. D. Mitchell, "Low-Complexity Set-Membership Channel Estimation for Cooperative Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol.60, no.6, pp. 2594,2607, July 2011.
- [69] R. C. de Lamare, P. S. R. Diniz, "Set-Membership Adaptive Algorithms Based on Time-Varying Error Bounds for CDMA Interference Suppression", *IEEE Transactions on Vehicular Technology*, vol.58, no.2, pp.644-654, Feb. 2009.
- [70] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1700-1708, Oct. 2003.

- [71] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. on Communications*, vol. 50, no. 6, June 2002.
- [72] Y. Sun, V. Tripathi, and M. L. Honig, "Adaptive, iterative, reduced-rank (turbo) equalization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2789–2800, Nov. 2005.
- [73] R. C. de Lamare, M. Haardt and R. Sampaio-Neto, "Blind Adaptive Constrained Reduced-Rank Parameter Estimation based on Constant Modulus Design for CDMA Interference Suppression," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, June 2008.
- [74] N. Song, R. C. de Lamare, M. Haardt, and M. Wolf, "Adaptive Widely Linear Reduced-Rank Interference Suppression based on the Multi-Stage Wiener Filter", *IEEE Transactions on Signal Processing*, vol. 60, no. 8, 2012.
- [75] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank MMSE filtering with interpolated FIR filters and adaptive interpolators", *IEEE Signal Processing Letters*, vol. 12, no. 3, March, 2005.
- [76] R. C. de Lamare and Raimundo Sampaio-Neto, "Reduced-rank Interference Suppression for DS-CDMA based on Interpolated FIR Filters", *IEEE Communications Letters*, vol. 9, no. 3, March 2005.
- [77] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Interference Suppression for DS-CDMA Systems based on Interpolated FIR Filters with Adaptive Interpolators in Multipath Channels", *IEEE Trans. Vehicular Technology*, Vol. 56, no. 6, September 2007.
- [78] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank MMSE Parameter Estimation based on an Adaptive Diversity Combined Decimation and Interpolation Scheme," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, April 15-20, 2007, vol. 3, pp. III-1317-III-1320.
- [79] R. C. de Lamare and R. Sampaio-Neto, "Reduced-Rank Adaptive Filtering Based on Joint Iterative Optimization of Adaptive Filters", *IEEE Signal Processing Letters*, vol. 14, no. 12, December 2007.
- [80] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering," *IEEE Trans. Signal Process.*, vol. 57, no. 7, July 2009, pp. 2503-2514.
- [81] R. C. de Lamare and R. Sampaio-Neto, "Reduced-Rank Space-Time Adaptive Interference Suppression With Joint Iterative Least Squares Algorithms for Spread-Spectrum Systems," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, March 2010, pp.1217-1228.
- [82] R.C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2482-2494, July 2011.
- [83] R. C. de Lamare, R. Sampaio-Neto and M. Haardt, "Blind adaptive constrained constant-modulus reduced-rank interference suppression algorithms based on interpolation and switched decimation", *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 681-695, February 2011.
- [84] H. Dai, A. F. Molisch and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection", *IEEE Trans. Wireless Communications*, vol. 3, no. 2, pp. 442–453, Mar. 2004.
- [85] P. Marsch and G. Fettweis, "Uplink CoMP under a constrained backhaul and imperfect channel knowledge", *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1730–1742, Jun. 2011.
- [86] P. Li and R. C. de Lamare, "Distributed Iterative Detection With Reduced Message 2 Passing for Networked MIMO Cellular Systems", *IEEE Transactions on Vehicular Technology*, vol.63, no.6, pp. 2947-2954, July 2014.
- [87] P. Chevalier and A. Blin, "Widely linear MVDR beamformers for the reception of an unknown signal corrupted by noncircular interferences," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5323–5336, 2007.
- [88] N. Song, W. U. Alokozai, R. C. de Lamare and M. Haardt, " Adaptive Widely Linear Reduced-Rank Beamforming Based on Joint Iterative Optimization", *IEEE Signal Processing Letters*, vol. 21, no. 3, March 2014.
- [89] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 944–960, Jul. 1995.
- [90] R. C. de Lamare and R. Sampaio-Neto, "Low-complexity variable step-size mechanisms for stochastic gradient algorithms in minimum variance CDMA receivers" *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2302-2317, June 2006.
- [91] C. Xu, G. Feng, and K. S. Kwak, "A modified constrained constant modulus approach to blind adaptive multiuser detection," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1642–1648, Sep. 2001.
- [92] R. C. de Lamare and R. Sampaio-Neto, "Blind adaptive code-constrained constant modulus algorithms for CDMA interference suppression in multipath channels", *IEEE Communications Letters*, vol. 9, no. 4, pp. 334-336, April 2005.
- [93] R. C. de Lamare and R. Sampaio-Neto, "Blind adaptive MIMO receivers for space-time block-coded DS-CDMA systems in multipath channels using the constant modulus criterion", *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 21-27, January 2010.
- [94] L. Wang and R. C. de Lamare, "Adaptive constrained constant modulus algorithm based on auxiliary vector filtering for beamforming", *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5408-5413, October 2010.