This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

Massive MIMO Detection Techniques: A Survey

Mahmoud A. Albreem, *Senior Member, IEEE,* Markku Juntti, *Senior Member, IEEE,* and Shahriar Shahabuddin, *Member, IEEE*

***Abstract*—Massive multiple-input multiple-output (MIMO) is a key technology to meet the user demands in performance and quality of services (QoS) for next generation communication systems. Due to a large number of antennas and radio frequency (RF) chains, complexity of the symbol detectors increased rapidly in a massive MIMO uplink receiver. Thus, the research to find the perfect massive MIMO detection algorithm with optimal performance and low complexity has gained a lot of attention during the past decade. A plethora of massive MIMO detection algorithms has been proposed in the literature. The aim of this paper is to provide insights on such algorithms to a generalist of wireless communications. We garner the massive MIMO detection algorithms and classify them so that a reader can find a distinction between different algorithms from a wider range of solutions. We present optimal and near-optimal detection principles specifically designed for the massive MIMO system such as detectors based on a local search, belief propagation and box detection. In addition, we cover detectors based on approximate inversion, which has gained popularity among the VLSI signal processing community due to their deterministic dataflow and low complexity. We also briefly explore several nonlinear small-scale MIMO (2-4 antenna receivers) detectors and their applicability in the massive MIMO context. In addition, we present recent advances of detection algorithms which are mostly based on machine learning or sparsity based algorithms. In each section, we also mention the related implementations of the detectors. A discussion of the pros and cons of each detector is provided.**

***Index Terms*—5G, massive MIMO, detection, local search, belief propagation, approximate matrix inversion, lattice reduction, sparsity, machine learning, and compressive sensing.**

I. Introduction

THE number of mobile users is dramatically increasing every year. Users crave faster Internet access and instant access to the multimedia services. In addition, the implementation of smart cities has reached stages wherein a dense and heterogeneous set of devices positioned over the urban area generates Exabytes of data to be exchanged [1]. Figure 1 shows an exponential growth in the mobile data traffic in 2015–2021 [2][3]. This calls for higher data rates, larger network capacity, higher spectral efficiency, higher energy efficiency, and better mobility [4]. Therefore, researchers have proposed the 5G networks to handle the above mentioned issues resulted from billions of wireless devices. A combination of well-known

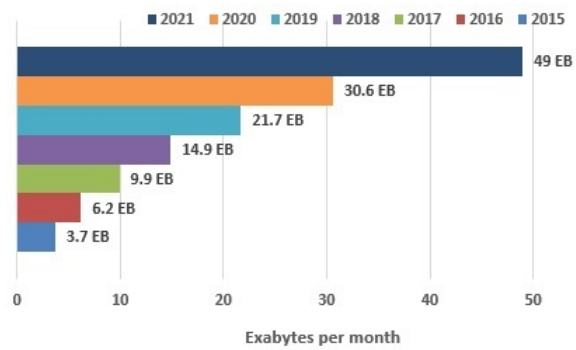

Figure 1. Global mobile data traffic forecast by 2021

and efficient technologies will be deployed in 5G networks such as the device-to-device (D2D) communication, the ultra dense networks (UDNs), the spectrum sharing, the centimeter wave (cmWave) or millimeter wave (mmWave), the internet of things (IoT), and the massive multiple-input multiple-output (MIMO) [5][6].

MIMO is a key technology that has been used since the third generation (3G) wireless networks to enhance performance of the wireless transceivers [7]. The idea is to use multiple antennas in the transmitter and the receiver to increase the spectral efficiency, the range and/or the link reliability. However, due to multiple interfering messages being transmitted from different antennas, the MIMO receiver is expected to use a detection mechanism to separate the symbols which are corrupted by interference and noise. The MIMO detector has been a topic of great interest during the past 50 years. Massive MIMO systems [8], [9] with a large number of antennas (up to hundreds) at the base station (BS) or access point are a natural extension of the conventional small-scale MIMO technology. The massive MIMO BS can serve a large number of user terminals with a single or few antennas in the same frequency band. The key feature of the classical massive MIMO system operating below 6 GHz carrier frequency is that the number of BS antennas is clearly larger than the total number of antennas in the user equipment within the cell or service area. Thereby the multiuser interference averages out to appear just as increased additive noise with the problems in channel estimation due to the pilot contamination [10].

The classical massive MIMO technology has been adopted for the fifth generation (5G) communication systems for below 6 GHz, wherein the scattering and multipath propagation in radio channels is rich. Thereby the interference averaging due to the large number of antenna elements makes the conventional matched filter (MF) based receivers often approx-

M. Albreem is with A'Sharqiyah University, Oman, and Centre for Wireless Communications, University of Oulu, Finland; e-mail: firstname.lastname@asu.edu.om and firstname.lastname@oulu.fi.

M. Juntti is with Centre for Wireless Communications, University of Oulu, Finland; e-mail: firstname.lastname@oulu.fi.

S. Shahabuddin is with Nokia, Finland and Centre for Wireless Communications, University of Oulu, Finland; e-mail: firstname.lastname@nokia.com and firstname.lastname@oulu.fi

This research has been financially supported by Academy of Finland 6Genesis Flagship (grant no. 318927), Nokia Foundation Centennial Grant, A'Sharqiyah University Research Visits Support Fund, and The Research Council of Oman.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

imately optimal. Very large antenna arrays are also needed at higher carrier frequencies, i.e., cmWave or mmWave bands and beyond toward the THz band. However, propagation channels are therein much more directive making the interference conditions rather different. Therefore, the term massive MIMO has not classically been used for those communications concepts, but this terminology varies from paper to paper. Large arrays are easier to implement and pack in the higher frequencies due to the smaller size of antennas. Therefore, the massive MIMO detection techniques may have a role in the cmWave or mmWave systems, although the propagation characteristics of the channels make the multiuser interference scenario quite different. We focus on the classical massive MIMO notion and detectors for systems operating below 6 GHz carrier frequency in this survey.

A. Relevant Prior Art

Massive MIMO has become a hot research topic in last few years, and, hence, several survey papers related to massive MIMO systems have been published [11], [12], [13], [14], [15], [16]. While these papers review a number of key topics of massive MIMO, none of them extensively discuss the detection techniques. A comprehensive review, comparison and discussion of the existing linear precoding mechanisms for massive MIMO according to different cell scenarios have been presented in [11]. It also discussed some standing challenges which related to the design of precoding mechanisms and practical implementations. Low complexity precoders suffered from a considerable performance loss, while a complicated precoder design is more difficult to implement practically. In [12], importance of the pilot contamination in massive MIMO is considered and hardware impairments are discussed. The article reviewed possible sources of pilot contamination, which include hardware impairments and non-reciprocal transceivers. Different mitigation techniques for pilot contamination have been categorized as pilot-based approach and subspace-based approach. In [13], challenges and benefits of the mmWave massive MIMO communication are reviewed. The paper discussed the enhancement in user throughput, spectral, energy efficiency and capacity. The design of mmWave massive MIMO communication system has to take into consideration the choice of the modulation technique, the signal waveform, the multiple access scheme, the user scheduling algorithm, the fronthaul design, the architecture of antenna array, precoding techniques, and health and safety issues. However, the holistic performance and evaluation of mmWave massive MIMO techniques in real-life scenarios and applications remains an open issue. In [14], an extensive investigation of massive MIMO propagation channels is presented and key differences from the conventional MIMO are discussed. It also reviewed the channel characteristics, measurements and models. Some future directions of channel models for massive MIMO are analyzed. It is concluded that the propagation channels will remain a vibrant research area in the coming few years. In [15], the combination of analog and digital beamforming structures using average channel state information (CSI) are reviewed. The hybrid beamforming structure keeps the number

Table I
ACRONYMS AND CORRESPONDING FULL MEANING

Acronyms	Full Form
ADMM	Alternating direction method of multipliers
ASIC	Application specific integrated circuit
BP	Belief propagation
BS	Base station
BER	Bit error rate
BB	Branch and bound
CD	Coordinate descent
CS	Compressive sensing
CG	Conjugate gradients
CDMA	Code division multiple access
CSI	Channel state information
CE	Constant envelop
D2D	Device-to-device
DFE	Decision feedback equalization
DBNIIR	Diagonal band Newton iteration with iterative refinement
EXIT	Extrinsic information transfer
ELR	Element-based lattice reduction
EVD	Eigenvalue decomposition
EP	Expectation propagation
GPU	Graphics processing unit
GMRES	Generalized minimal residual
GS	Gauss-Seidel
ISI	Inter-symbol interference
IoT	Internet of things
ICI	Inter-channel interference
IR	Iteration refinement
KSD	K-best sphere decoding
KL	Kullback-Leibler
LLL	Lenstra-Lenstra-Lovasz
LTE	Long term evolution
LOS	Line-of-sight
LLR	Log-likelihood ratios
LDPC	Low density parity check
LRA	Lattice reduction aided
LAS	Likelihood ascent search
MIMO	Multiple-input multiple-output
MIC	Multistage interference cancellation
MUD	Multiuser detection
mmWave	Millimeter wave
MMSE	Minimum mean square error
MF	Matched filter
MLSD	Maximum-likelihood sequence detection
MUD	Multiuser detector
mMTC	Massive machine-type communications
MSE	Mean-square error
MMP	Multipath matching pursuit
MRC	Maximum ratio combining
MEP	Multi-envelop precoding
MC	Multicell
NI	Newton iterations
NS	Neumann series
NIIR	Newton iteration with iterative refinement
OSD	Ordered sphere decoding
PAM	Pulse amplitude modulation
PIC	Parallel interference cancellation
PAPR	Peak-to-average-power ratio
QoS	Quality of service
QAM	Quadrature amplitude modulation
QP	Quadratic programming
RTS	Reactive tabu search
RF	Radio frequency
SOR	Successive over relaxation
SUMF	Single user matched filtering
SDM	Space division multiplexing
SD	Sphere decoding
SIMO	Single-input multiple-output
SNR	Signal-to-noise ratio
SIC	Successive interference cancellation
SR	Sequential Reduction
SLV	Shortest Longest Vector
STBC	Space-time block codes
SER	Symbol error rate
SC-FDMA	Single carrier frequency division multiple access
SC	Single-cell
UDNs	Ultra dense networks
VBLAST	Vertical Bell laboratories layered space-time
VCM	Virtual channel model
VLSI	Very-large-scale integration
WSSOR	Weighted symmetric successive over relaxation
ZF	Zero-forcing
3G	Third generation
4G	Fourth generation
5G	Fifth generation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

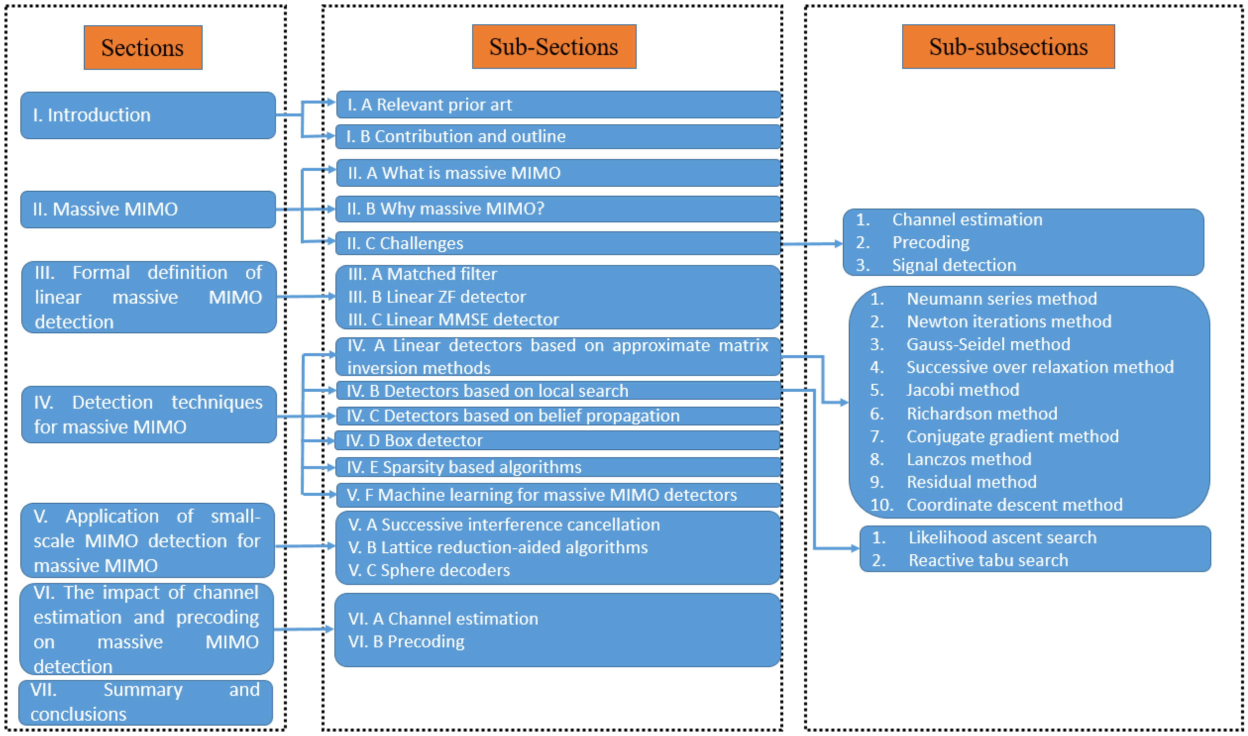

Figure 2. Outline of the article

of radio frequency (RF) chains within reasonable limits. It is shown that the hybrid beamforming techniques are promising for reducing the hardware cost and the training overhead. However, a trade-off between complexity and performance has to be considered in the design of different applications and channel characteristics.

A survey dated on 2015 [16] presented a detailed clarification of the MIMO detection fundamentals, and recited the half-a-century history of MIMO detection. The authors provided an extensive overview and milestones in the development of optimal, linear, interference cancellation MIMO detection. In addition, the tree-search detectors were reviewed extensively and milestones in development of the depth-first, the breadth-first and the best-first type were presented. The lattice reduction, probabilistic data association and semidefinite relaxation detectors were also reviewed extensively. In addition, relevant lessons were inferred from the rich heritage of small- and medium-scale MIMO detection for the sake of designing complexity-scalable MIMO detection algorithms that are potentially viable to large-scale MIMO systems. The authors briefly presented the recent advances of massive MIMO detectors where it is divided into two types. The first type corresponds to the case where the number of BS antennas is much larger than the number of active users and the second type corresponds to the case where the number of active users is comparable to the number of BS antennas. It is concluded that some conventional MIMO detectors might be infeasible with a certain type of massive MIMO systems. For instance, the family of tree-search based detectors will become less feasible in the first type and it might be invoked in the second type of massive MIMO systems.

B. Contribution and Outline

In this paper, an extensive survey on detection algorithms related to massive MIMO system is presented. Our particular focus is on performance and complexity trade-off as well as the practical implementation of detection algorithms. Although the survey in [16] is extensive, the primary focus of the article was not on massive MIMO systems. For instance, the effect of matrix inversion and approximate matrix inversion methods in the detection process are not presented in [16]. There is also a paucity of reviews on advanced detectors based on the local search, the belief propagation (BP), the BOX-detection, sparsity and machine learning approaches. To our best knowledge, this is the first survey to explore the detection mechanisms that pertains to only massive MIMO systems. A plethora of massive MIMO detection algorithms has been proposed in the literature. The aim of this paper is to provide insights on such algorithms to a generalist of wireless communications. In this paper, we garner the massive MIMO detection algorithms and present their performance-complexity profile so that a reader can find a distinction between different detection algorithms from a wider range of possible solutions. It starts off with a dive into the history of detectors for a small-scale MIMO. It then presents the concepts and benefits of massive MIMO system. After that, it discusses the signal detection challenge in massive MIMO system. Then, it surveys the corresponding detection solutions for massive MIMO systems starting with classical linear detectors with approximate matrix inversion methods such as the Neumann series (NS) method, the Newton iterations (NI) method, the Gauss-Seidel (GS) method, the successive over relaxation (SOR) method, the Jacobi method, the Richardson method, the

Lanczos method, the residual method, the coordinate descent (CD) method, and the conjugate gradients (CG) method. It also presents detectors based on nonlinear methods such as the successive interference cancellation (SIC), lattice reduction-aided algorithms, and the sphere decoding (SD). Sequentially this paper comprehensively reviews the detectors based on the local search, the belief propagation, the box detection, machine learning based detectors and sparsity based algorithms. One of the key features of this article is presenting the pros and cons of each detector based on the performance-complexity profile as well as the implementation stiffness.

Section II describes the concept of massive MIMO systems. Section IV presents the massive model and illustrates the detection techniques for massive MIMO. Finally, Section VII concludes the paper and presents the future directions in massive MIMO systems. For smooth readability, the outline of the article is depicted in Fig. 2 and the most used acronyms are presented in full form in Table I.

II. Massive MIMO

A. What is massive MIMO?

Massive MIMO is a scaled up version of the conventional small scale MIMO systems [8], [9]. As shown in Fig. 3, massive MIMO system is a multiuser communications solution that employs a large number (practically some dozens or hundreds, theoretically up to thousands) of antenna elements to serve simultaneously multiple users with a flexibility to opt what users to schedule for reception at any given time. The most common massive MIMO concept assumes that the user terminals have just a single antenna[1] and that the number of antennas at the BS is significantly larger than the number of served users. The introduction of massive MIMO had a tremendous impact on the research and development community during past decade. As a result, many next generation communication technologies, such as 5G below 6 GHz adopted massive MIMO as their key technology. Most of the massive MIMO literature focuses on mobile broadband type high rate problems with large data packets such that channel estimation and training makes clearly sense. The other application of interest is the massive machine-type communications (mMTC) wherein large number of connected devices are only sporadically active [17], [18].

B. Why massive MIMO?

Massive MIMO technology relies on increasing the spatial multiplexing gain and the diversity gain by adding the number of antennas at the BS to serve users with relatively simple processing of signals from all the antennas [19][20]. The potential benefits of massive MIMO can be summarized as follows:

- Capacity and link reliability: Massive MIMO increases the diversity gain, and hence, provides link robustness as it resists fading [21]. It is approved that the capacity increases without a bound as the number of antenna increases, even under a pilot contamination, when multicell minimum mean square error (MMSE) precoding/combining and spatial channel correlation are used [22].
- Spectral efficiency: Massive MIMO improves the spectral efficiency (SE) of the cellular network by spatial multiplexing of a large number of user equipment's per cell [23]. Numerous antennas create more spatial data-streams, more throughput, more multiplexing gain, and hence achieve high spectral efficiency [24]. It is shown that the overall spectral efficiency in the massive MIMO can be ten times higher than in the conventional MIMO where tens of users will be served simultaneously in the same time-frequency resources [8].
- Energy efficiency: Due to coherent combining, the transmitted power is inversely proportionate to the number of transmit antennas [25]. As the number of transmit antennas increases, the transmit power will be significantly reduced. The power per antenna should be $\propto \frac{1}{n_t}$, where n_t is the number of antennas. Also, the throughput could be increased by increasing the number of transmit antennas and without increasing the transmit power [26]. Each antenna uses extremely low power, i.e., milliwatts [10]. Therefore, the energy efficiency increases and equivalently system reliability.
- Security enhancement and robustness improvement: Man-made interference and intentional jamming are serious concerns in modern wireless communication systems. Massive number of antenna terminals [8] leads to a large number of degrees of freedom which can be used to cancel the signals from intentional jammers [27]. In addition, massive MIMO systems are also inherently robust against passive eavesdropping attacks because of beamforming. However, the eavesdropper can take countermeasures by exploiting the high channel correlation in the vicinity of the user or the weakness of channel estimation [28].
- Cost efficiency: Massive MIMO eliminates the need for bulky items such as coaxial cables which used to connect the BS components, and hence reduces the system implementation cost [8]. In addition, massive MIMO uses cheap milliwatts amplifier instead of a multiple expensive high power amplifier [29]. Moreover, it has a potential to reduce the radiated power 1000 times and at the same time drastically maximize the data rates [8].
- Signal processing: A large number of antennas eliminates the interference effects, fast fading, uncorrelated noise and thermal noise, and hence simplifies the signal processing [1][30]. In addition, it is favorable propagation environment occurs when the channel responses from the base station to user terminals are different (mutually orthogonal, i.e., the inner products are zero). However, non-orthogonal channel vectors lead to advanced signal processing to suppress the interference.

One of the key properties of massive BS antenna arrays is so called channel hardening. It refers to the phenomenon where the massive MIMO channel matrix approaches their expected

[1]The massive MIMO is often in a way a misnomer, due to the single-antenna terminal. The system is actually a multiuser single-input multiple-output (SIMO) uplink or multiple-input single-output (MISO) downlink. As is customary in the literature, we refer in this paper both single and multiple antenna terminal case by massive MIMO system.

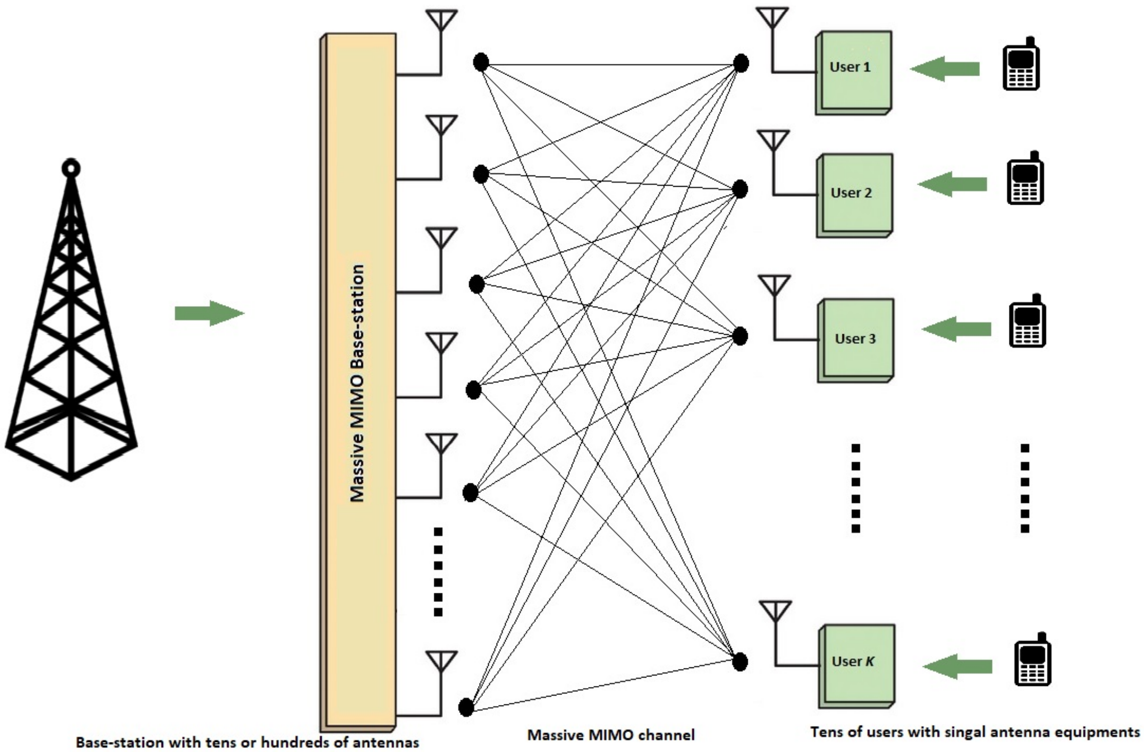

Figure 3. Massive MIMO architecture

values, when the number of antennas approaches infinity. In other words, the effective channel approaches deterministic and the off-diagonal components of the Gramian matrix become weaker compared to diagonal terms as the size of the channel gain matrix increases. This property can be exploited in the detection technique and the channel estimation. The simple matched filter (MF) approaches optimality in such a case [31]. However, this is only true with rich scattering and truly large antenna arrays. Therefore, advanced detection techniques are of interest in the practical propagation scenarios and correlated fading channels.

C. Challenges

Even though the numerous number of antennas benefits the communication system, massive MIMO imposes new challenges for the signal processing that can be categorized as:

1) Channel estimation: Channel estimation plays a major role in the overall performance of wireless systems. The base station requires an accurate estimation of the channel state information (CSI) to meet the potential advantages of the massive MIMO in practice. However, it is difficult to obtain the CSI for a large number of channels [32]. There is a need to exchange the CSI across the transmitters on a fast time scale and low-latency basis [33]. In addition, channel estimation is known to be hampered by the pilot contamination effect [34]. Channel information is obtained on the basis of finite-length pilot sequences in the presence of inter-cell interference. Hence, the pilot sequences from adjacent cells would contaminate each other. Therefore, channel estimation problems should be addressed in the massive MIMO to provide a substantial improvement in performance.

2) Precoding: It is an important signal processing scheme which uses the CSI at the transmitter to maximize the link performance. The BS has to precode the downlink data to focus the spatial data-streams at the users' location [35]. In other words, the transmit precoding can be used in the downlink to concentrate each signal at its intended receiver. In a non-LOS environment, the concept of focusing the antenna array toward a specific terminal becomes more complicated where the geographical point of multipath components have to be considered. In small-scale MIMO precoding schemes such as zerof-forcing (ZF) precoding, the symbols are modified in both amplitudes and phase at the baseband and supported by a dedicated radio frequency (RF) transceivers. Therefore, each antenna element requires a dedicated RF transceivers for assistance which is too expensive in the case of massive MIMO [36]. However, precoders with much less number of transceivers than total number of antennas will be more realistic and cost-effective to deploy [37]. In [38], constant envelop (CE) precoding techniques have been proposed to reduce the hardware costs. The transmitted signal is generated by varying only the phase of the constant amplitude baseband symbols, therefore, the peak-to-average-power ratio (PAPR) is significantly improved which makes utilizing cheap power amplifiers viable at the BS circuitry. Unfortunately, the hardware costs are reduced in an expenses of performance loss. This algorithm has been exploited in [39] where a multi-envelop precoding (MEP) technique has been proposed. The idea behind this technique is to utilize more than one but only few envelop levels, which reduces the needed additional power and increases the achievable data rates with respect to CE precoding. A plethora of literature can be found on precoding techniques such as the joint spatial division and multiplexing (JSDM) [40], [41], iterative algorithms for precoder optimization [42] and JSDM-finite alphabets [43].

3) Signal detection: Accurate and instantaneous CSI is needed at the BS to perform precoding in a forward link (downlink) and detection in a reverse link (uplink). Performance of the MF detector can be reasonably good in rich scattering channels with a small number of users. However, in spatially correlated channels and to increase the spectral efficiency, more advanced detectors are needed. Complexity of massive MIMO detection algorithm is affected by systems' size (number of antennas at both sides, transmitters and receiver terminals), the matrix-by-matrix multiplication, and the matrix inversion. However, a balance between performance and complexity should be considered [44]. We cover these aspects in this tutorial review.

III. Formal Definition of Linear massive MIMO Detection

We provide a formal definition of the massive MIMO system model in this section. The aim is to provide a relevant background to the readers for subsequent sections. There have been a reinvigorated interest in the traditional linear detectors since the introduction of massive MIMO systems. Therefore, we also present the linear detection mechanism in this section. We assume massive multi-user MIMO base-station (BS) is serving K single antenna users. The BS has a total N antennas where $K \leq N$. Assuming frequency-flat channel, the channel coefficients between K users and N BS antennas forms a matrix ($\mathbf{H}$) which can be expressed as

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1j} \\ h_{21} & h_{22} & \cdots & h_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ h_{i1} & h_{i2} & \cdots & h_{ij} \end{bmatrix}, \tag{1}$$

where h_{ij} is the channel gain or loss from jth transmit antenna to ith receive antenna. The channels are shown in lines between the BS and users in Fig. 3. The elements of the channel matrix $\mathbf{H} \in \mathbb{C}^{N\times K}$ are often assumed to be independent and identically distributed (i.i.d) Gaussian random variables with zero mean and unit variance. However, this is not always the case in truly directive channels. The K users transmit their symbols individually and we can form a symbol vector $\mathbf{x} = [x_1, x_2, ..., x_K]^T$ transmitted by all the users in the uplink or reverse direction. The BS receives a vector $\mathbf{y} = [y_1, y_2, ..., y_N]^T$ which is corrupted by channel effects and noise. The relationship between $\mathbf{x}$ and $\mathbf{y}$ can be characterized as

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n}, \tag{2}$$

where $\mathbf{n}$ is $N \times 1$ additive white Gaussian noise (AWGN) whose entries are i.i.d.. This model is generally adopted to derive a detection algorithm, where the CSI and the synchronization is assumed to be perfect at the BS.

The task of a MIMO detector is to determine the transmitted vector $\mathbf{x}$ based on the received vector $\mathbf{y}$. The maximum-likelihood sequence detection (MLSD) is an optimal algorithm to solve the MIMO detection problem. It performs an exhaustive search and examines all possible signals as illustrated by

$$\hat{\mathbf{x}}_{ML} = \arg \min_{\mathbf{x}\in O^K} \|\mathbf{y} - \mathbf{Hx}\|_2^2, \tag{3}$$

where $\hat{\mathbf{x}}$ is the estimated received signal. The ML problem is combinatorial in nature and the numerical standard algorithms for the convex optimization are not applicable. Therefore, the complexity of ML is exponential in the number of decision variables O^K [45]. The ML detector can be prohibitively complex even for a small-scale MIMO detection. For instance, a transmitter with four antennas supporting 64-QAM alphabet requires a 16.7×10^6 candidate comparisons for the ML detection. The linear detectors relax the discrete set O^K to a complex set so that the problem 3 can be solved with convex optimization methods and closed solution can be obtained. Linear detectors can be represented as multiplying the received signal $\mathbf{y}$ with the equalization matrix $\mathbf{A}^H$, $\hat{\mathbf{x}} = \mathcal{S}(\mathbf{A}^H\mathbf{y})$, followed by a slicer $\mathcal{S}(.)$, which quantizes each entry to the nearest neighbor in constellation [46]. The most conventional low complexity linear detectors such as the MF, the ZF algorithm and the MMSE algorithm are explained here.

A. MF detector

MF handles the interference from other sub-streams as purely noise by making $\mathbf{A} = \mathbf{H}$. The estimated received signal using MF is given by

$$\hat{\mathbf{x}}_{MF} = \mathcal{S}(\mathbf{H}^H\mathbf{y}), \tag{4}$$

which works properly when K is much smaller than N and it provides a worse performance compared to more complex detectors. MF, also called the maximum ratio combining (MRC), aims to maximize the received SNR of each stream by neglecting the effect of multiuser interference. If the channel is ill-conditioned, performance is severely degraded for a square MIMO system [47].

B. Linear ZF detector

ZF outperforms the MF detector and it aims to maximize the received signal-to-interference ration (SINR). The ZF mechanism is based on inverting the channel matrix $\mathbf{H}$ and thus, removing the effect of the channel. The equalization matrix of the ZF detector [48] is given by

$$\mathbf{A}_{ZF}^H = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H = \mathbf{H}^+, \tag{5}$$

where $\mathbf{H}^+$ is the Moore-Penrose pseudo-inverse of a matrix. The pseudo-inverse is used because $\mathbf{H}$ is not always a square matrix, i.e. the number of users is not equal to the number of antennas at BS. The estimated signal can be shown as

$$\hat{\mathbf{x}}_{ZF} = \mathcal{S}(\mathbf{A}_{ZF}^H\mathbf{y}). \tag{6}$$

It is clear that the ZF detector neglects the effect of noise and it works properly in interference-limited scenarios in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

expenses of higher computational complexity. However, the ZF detector and the MF may produce a noise enhancement in case of a small-valued coefficient channel. Therefore, MMSE detector is proposed to take the effect of noise in the equalization process.

C. Linear MMSE detector

The main idea of the MMSE detector is to minimize the mean-square error (MSE) between the transmitted $\mathbf{x}$ and the estimated signal $\mathbf{H}^H\mathbf{y}$ as given by

$$\mathbf{A}_{MMSE}^H = arg \min_{\mathbf{H}\in\mathbb{C}^{N\times K}} \mathbb{E}\|\mathbf{x}-\mathbf{H}^H\mathbf{y}\|^2, \tag{7}$$

where $\mathbb{E}$ is the expectation operator. MMSE detector takes the noise effect into consideration as

$$\mathbf{A}_{MMSE}^H = \left(\mathbf{H}^H\mathbf{H}+\frac{K}{SNR}\mathbf{I}\right)^{-1}\mathbf{H}^H, \tag{8}$$

where $\mathbf{I}$ is the identity matrix. The output of the MMSE detector can be obtained by

$$\hat{\mathbf{x}}_{MMSE} = \mathcal{S}(\mathbf{A}_{MMSE}^H\mathbf{y}). \tag{9}$$

Unlike the ZF detector in (6), the MMSE in (8) depends on a reduced noise enhancement and it requires a knowledge of the SNR. Therefore, the MMSE detector is capable of achieving a significantly better performance than the ZF detector when the noise power is large.

IV. Detection Techniques for Massive MIMO

The earliest massive MIMO detector dates back to 2008 when Vardhan *et al.* proposed a detection mechanism for massive MIMO based on likelihood ascent search [49]. The main issue with utilizing a large number of antennas is the high complexity involved and the proposed detector achieves a near-ML performance with low complexity. The next few years researchers proposed near-ML detectors using local search and belief propagation algorithms. As the number of antennas increases, complexity of the matrix inversion required for linear detectors also increase exponentially. In 2013, Wu *et al.* proposed an approximate inversion method based uplink detector in [50] and initiated a new direction for the massive MIMO detector research. This class of massive MIMO detectors are designed for specific massive MIMO configurations, i.e. when the number of antennas are high compared to the number of users. The effect of a channel hardening is higher for such massive MIMO configurations and thus, low complexity detectors can achieve high performance. The approximate inversion based detectors have been the most popular class of detectors since its introduction in 2013. In addition, symbol detectors utilizing the sparsity and the machine learning also gained attraction for massive MIMO configurations. In this section, we explore all the novel detectors which are proposed for the massive MIMO.

A. Linear Detectors Based on the Approximate Matrix Inversion

With a large number of transmit antennas, the channel hardening phenomenon can be exploited to cancel the characteristics of a small-scale fading [51] and it becomes dominant when the number of served users (K) is much lower than the number of receive antennas (N). This can be seen as a diagonlisation of the entries in the *Gram matrix* or *Gramian* $\mathbf{G} = \mathbf{H}^H\mathbf{H}$, where the non-diagonal components tend to zero and diagonal terms become closer to N [44][52]. As shown in (5), a matrix inversion of the *Gramian* matrix is required to equalize the received signal. It exhibits high computational complexity being one of the most complex operations in the linear and simple non-linear MIMO detectors. For the massive MIMO system, this problem becomes more severe as the dimension of the *Gramian* $\mathbf{G}$ increases [53]. Several methods have been proposed to reduce complexity by approximating the inverse of a matrix, rather than computing it [54]. Besides the cost of a matrix inversion, a challenge in matrix inversion lies on when the channel matrix is nearly singular and the system becomes ill-conditioned. In this case, the matrix inversion will not equalize the received signal [55], [56]. In order to overcome the inherent noise enhancement, modified detectors with approximate matrix inversion methods will be an essential. Therefore, detectors based on approximate matrix inversion will be presented and discussed below.

1) Neumann Series: The Neumann series (NS) is a popular method for approximating the matrix inversion which subsequently reduces complexity of the linear detector [54]. $\mathbf{G}$ can be decomposed into $\mathbf{G} = \mathbf{D}+\mathbf{E}$, where $\mathbf{D}$ is the main diagonal matrix and $\mathbf{E}$ is the non-diagonal matrix [57]. The NS expansion of $\mathbf{G}$ is given by

$$\mathbf{G}^{-1} = \sum_{n=0}^{\infty}\left(-\mathbf{D}^{-1}\mathbf{E}\right)^n\mathbf{D}^{-1}. \tag{10}$$

The polynomial expansion in (10) converges to the matrix inverse $\mathbf{G}^{-1}$ if

$$\lim_{n\to\infty}\left(-\mathbf{D}^{-1}\mathbf{E}\right)^n = 0. \tag{11}$$

In practice, a finite number of terms is utilized, and, thus, a fixed number of iterations of (10) is performed. As the number of iterations n increases, high precision of the matrix inverse will be achieved at the expense of extra complexity. The NS based algorithm reduces the computational complexity from $O\left(K^3\right)$ to $O\left(K^2\right)$ when the number of iterations $n\leq 2$ [58][59]. However, the NS method recursion is slow, therefore, high-order recursion method such as Schulz recursion can be used to accelerate the NS recursion in expenses of extra computational complexity [60].

In [61], a MMSE parallel interference cancellation (MMSE-PIC) based algorithm is proposed to reduce the computational complexity by exploiting the NS expansion to replace the matrix-matrix multiplication of $\mathbf{G}$ with a matrix-vector multiplication. This method employed $n\leq 3$ for a MIMO size of 16×128. Compared to the ML detector, the computational complexity has been reduced to $O(nKN)$ with a marginal performance loss when $n = 3$ compared to the MMSE performance. Complexity can be reduced only when n is small.

In [62], an iterative detector has been proposed to manage complexity even for large n. The proposed algorithm exploits the NS method to compute the LLR for the channel decoder. The proposed algorithm has been tested at 12×70 MIMO system and it reduces the complexity from $O\left(K^3\right)$ to $O\left(K^2\right)$.

In [63], performance and complexity of the NS based linear detector has been investigated in the condition of Rician channel model. High throughput application specific integrated circuit (ASIC) is designed for the NS based detector in [64]. The ASIC achieves 3.8 Gbps for 128 antenna BS and 8 users for single carrier frequency division multiple access (SC-FDMA). The NS based detector is also implemented on a Xilinx Virtex-7 FPGA in [52]. The FPGA design achieves 600 Mbps for 128 antenna BS and 8 users.

It should be noted that the detection based on the NS method suffers from a considerable performance loss when the ratio between the number of BS antennas and user antennas, β, is large (close to 1) [65].

2) Newton Iteration method: The Newton iteration (NI) is also known as the Newton-Raphson method and it is an iterative method for finding the approximation of the matrix inverse [66]. For $\mathbf{G}$, estimation of the matrix inversion at *nth* iteration is given by

$$\mathbf{X}_n^{-1} = \mathbf{X}_{n-1}^{-1}\left(2\mathbf{I} - \mathbf{G}\mathbf{X}_{n-1}^{-1}\right), \tag{12}$$

which converges quadratically to the inverse matrix if

$$\left\|\mathbf{I} - \mathbf{G}\mathbf{X}_0^{-1}\right\| < 1. \tag{13}$$

High precision can be achieved using NI with quadratic convergence [67]. Like the NS method, it only requires a simple computation to accelerate the detection process. Although the NI requires one more matrix multiplication in each step, the approximation converges faster in comparison to the NS method. The NI requires only few iterations to approach a matrix inverse with impressive precision while the NS method requires more iterations for the same results [67]. In [68], the NI based linear detector has been used in 16×16 MIMO system. In [67][69], the NI with iterative refinement (NIIR) and diagonal band NIIR (DBNIIR) have been employed for a MIMO size of 16×128. The computational complexity reduced from $O\left(K^3\right)$ to $O\left(K^2\right)$.

In [70], NI and NS methods are exploited in the hybrid detector's design to achieve a fast convergence. The proposed hybrid detector has been employed in a 32×256 MIMO system where a faster convergence rate has been achieved compared to conventional NI and NS, with almost the same complexity when the iteration number is greater than or equal to 2.

3) Gauss-Seidel method: The Gauss-Seidel (GS) method is also known as the Liebmann method or the method of successive displacement. It is used to solve the linear system shown in (2). The *Gramian* matrix ($\mathbf{G}$) can be decomposed into $\mathbf{G} = \mathbf{D} + \mathbf{L} + \mathbf{U}$ [71][72], where $\mathbf{D}$, $\mathbf{L}$ and $\mathbf{U}$ are the diagonal component, the strictly lower triangular component, and the strictly upper triangular component, respectively. The GS method can be used to estimate the transmitted signal vector ($\hat{\mathbf{x}}$) [73] and its characterized by

$$\hat{\mathbf{x}}^{(n)} = (\mathbf{D} + \mathbf{L})^{-1}\left(\hat{\mathbf{x}}_{MF} - \mathbf{U}\hat{\mathbf{x}}^{(n-1)}\right), \qquad n = 1, 2, \cdots, \tag{14}$$

where n is the number of iterations and $\hat{\mathbf{x}}_{MF}$ is shown in (4). If there is no priori information about the initial solution $\hat{\mathbf{x}}^{(0)}$, it can be considered as zero [72]. According to [71], the GS iteration method outperforms the NS method with lower complexity. In [73], a detector based on the GS method has been proposed with initial solution based on the NS expansion with two terms. The proposed detector is implemented in the FPGA for 8×128 MIMO system. The parallel version of the GS method is implemented in [72]. It outperforms the implementation of [73] in terms of throughput for a 8×128 system.

It has also shown that detectors based the GS method can reduce the complexity to be $O\left(K^2\right)$ [59]. However, due to the GS internal sequential iterations structure, it is not well suited for parallel implementation [71], [74], [75].

4) Successive Over-Relaxation (SOR) method: The detected signal using the SOR iteration is described by

$$\hat{\mathbf{x}}^{(n)} = \left(\frac{1}{\omega}\mathbf{D} + \mathbf{L}\right)^{-1}\left(\hat{\mathbf{x}}_{MF} + \left(\left(\frac{1}{\omega} - 1\right)\mathbf{D} - \mathbf{U}\right)\hat{\mathbf{x}}^{(n-1)}\right), \tag{15}$$

where $n = 1, 2, \cdots$, and ω represents the relaxation parameter that plays a crucial role in the convergence rate. The GS method is a special case of the SOR method [76] where we can obtain (14) from (15) by setting $\omega = 1$. For uplink massive MIMO systems, the signal detection technique using the SOR method is convergent when the relaxation parameter ω satisfies $0 < \omega < 2$ [77].

The SOR method outperforms the NS approximation method in terms of performance and complexity reduction [77]. However, the detection algorithm using the GS method enjoys lower complexity than the SOR method [59]. In [78], a detector based on the SOR method has been proposed with iterative initial solution and an optimal relaxation parameter. The proposed detector achieved a significant improvement in detection performance when ratio of the number of BS antennas to the number of user terminal antennas, β, is small. The mathematical relationship between the antenna ratio β and the optimal relaxation parameter is defined based on Marchenko-Pastur law. The proposed detector employed in 16×80 MIMO system where the complexity has been reduced from $O\left(K^3\right)$ to $O\left(K^2\right)$. The SOR-based detector has been implemented on Xilinx Virtex-7 FPGA for a 8×128 system and achieved a throughput of 22 Mbps after two iterations.

5) Jacobi method: The Jacobi method is another simple iterative method for determining the solution of a diagonally dominant system where the estimated signal is given by

$$\hat{\mathbf{x}}^{(n)} = \mathbf{D}^{-1}\left(\hat{\mathbf{x}}_{MF} + (\mathbf{D} - \mathbf{A})\,\hat{\mathbf{x}}^{(n-1)}\right), \tag{16}$$

which holds if

$$\lim_{n\to\infty}\left(\mathbf{I} - \mathbf{D}^{-1}\mathbf{A}\right)^n = 0. \tag{17}$$

In the massive MIMO system, the condition in (17) can be met in a very high probability [79][80]. The initial estimation can be identified as

$$\hat{\mathbf{x}}^{(0)} = \mathbf{D}^{-1}\hat{\mathbf{x}}_{MF}. \tag{18}$$

The computational complexity of a detector using the Jacobi method is lower than the computational complexity obtained by the SOR method and the NS method [59]. In [79], a detector based on the Jacobi method has been designed to guarantee that the first iteration will be free of multiplication and hence, reduce the complexity. The detector has been tested in 8×64, 8×128, 16×128, and 16×256 MIMO systems. For a small number of iterations (up to 2), the computation complexity is $O\left(K^2\right)$ while it will be increased to $O\left(K^3\right)$ for high iterations.

6) Richardson method: The Richardson method concept is based on executing certain vector operations and multiplication by the channel matrix $\mathbf{H}$. It only utilizes symmetric matrices defined as positive at their execution and can be slowed as it approaches the exact solution over time. In order to achieve a fast convergence, a relaxation parameter ω has been introduced into iterative process and it satisfies $0<\omega<\frac{2}{\lambda}$ where λ is the largest eigenvalue of the symmetric positive definite matrix $\mathbf{H}$ [81]. The Richardson iterations can be described by

$$\mathbf{x}^{(n+1)}=\mathbf{x}^{(n)}+\omega\left(\mathbf{y}-\mathbf{H}\mathbf{x}^{(n)}\right)\quad n=0,1,2,\cdots. \tag{19}$$

The initial solution $\mathbf{x}^{(0)}$ can be identified as $2K\times 1$ zero vector without loss of generality as no a priori knowledge of the final solution is available [82]. The Richardson method achieves a near-ML performance but it requires plenty iterations [81][74].

Richardson method is a hardware friendly method and it can reduce the computational complexity from $O\left(K^3\right)$ to $O\left(K^2\right)$ for a wide range of n [58][59]. In [58], a reformulation of a Richardson method has been proposed for 16×128 MIMO hardware detector to reduce the computational complexity to $O(K)$. In [83], a graphics processing unit (GPU) friendly detector based a scalable Richardson method as been proposed. The proposed detector exploits the channel hardening gain to reduce the number of iterations n in the soft Richardson detector.

7) Conjugate Gradients method: Conjugate gradients (CG) method is another effective method to solve the linear equations through *nth* iterations [84][85]. The estimated signal ($\hat{\mathbf{x}}$) can be obtained using

$$\hat{\mathbf{x}}^{(n+1)}=\hat{\mathbf{x}}^{(n)}+\alpha^{(n)}\mathbf{p}^{(n)}, \tag{20}$$

where $\mathbf{p}^{(n)}$ is the conjugate direction with respect to $\mathbf{A}$, i.e.,

$$\left(\mathbf{p}^{(n)}\right)^H\mathbf{A}\mathbf{p}^{(j)}=0,\quad for\quad n\neq j, \tag{21}$$

where $\alpha^{(n)}$ is a scalar parameter.

The CG-based detection algorithm outperforms the NS-based detection scheme in terms of performance and complexity [85]. In [86], a detector based on the CG algorithm has been proposed and achieved a significant complexity reduction. In [75], a hybrid detector based on the CG algorithm and the Jacobi method has been proposed to speed up the convergence rate and improved performance. The CG-based detector is implemented in Xilinx Virtex-7 FPGA for a 128×8 in [87]. The CG-based detector is also implemented using a GPU platform in [88].

8) Lanczos method: Lanczos method is one of the Krylov subspace methods used to solve large sparse linear equations. It typically generates the orthogonal basis of the co-efficient matrix and finds the solution whose residual is orthogonal to Krylov subspace. The Lanczos method can be interpreted as a subspace approximation of the exact solution. This approximation converges quickly to the exact solution when the number of basis is large. However, the steps of Lanczos method can be divided into initialization and iterations. The iterative process can be concluded in a relationship between the estimated signal and received signal [89] as given by

$$\hat{\mathbf{x}}=\mathbf{Q}^{(n)}\mathbf{F}^{(n)-1}\mathbf{Q}^{(n)H}\mathbf{G}\mathbf{x}+\mathbf{Q}^{(n)}\mathbf{F}^{(n)-1}\mathbf{Q}^{(n)H}\mathbf{H}^H\mathbf{n}, \tag{22}$$

where $\mathbf{Q}$ and $\mathbf{F}$ are the matrix formed by orthogonal basis, and the tridiagonal matrix, respectively. For all iterative methods, i.e. the GS method, the SOR method, the Jacobi method, the Richardson method, and the Lanczos method, the initial solution $\hat{\mathbf{x}}^{(0)}$ plays a major role in its convergence. A possible selection of the initial solution [90] is given by

$$\hat{\mathbf{x}}^{(0)}=\mathbf{D}^{-1}\hat{\mathbf{y}}, \tag{23}$$

where $\mathbf{D}$ denotes the diagonal entries of the *Gram* matrix $\mathbf{G}$ and $\hat{\mathbf{y}}=\mathbf{H}^{-1}\mathbf{y}$.

Low complexity soft-output detection using the Lanczos method is proposed in [89]. It is capable of outperforming the existing NS approximation based detectors where the MMSE performance has been achieved within few iterations. In [91], the convergence speed of a Lanczos method based MU-MIMO detector has been analyzed by Kaniel-Paige-Saad theory. The storage requirement of the Lanczos method based detectors has been reduced in [90]. However, the Lanczos method either has low performance or requires high computational complexity under the time-varying channel [92].

9) Residual method: Residual method is an iterative method that concentrates on minimization of the residual norm rather than approximating the exact solution. In [93], the generalized minimal residual (GMRES) method has been used for the symbol detection to compute the MMSE filter without a matrix inversion. In the GMRES method, approximation of the exact solution $\mathbf{y}=\mathbf{H}\mathbf{x}$ can be considered by the vector $\mathbf{x}_s\in\tau_s$ that minimizes the norm of the residual vector $\mathbf{r}_s=\mathbf{H}\mathbf{x}_s-\mathbf{y}=$ where τ_s is given by

$$\tau_s=\text{span}\left\{\mathbf{y},\mathbf{H}\mathbf{y},\cdots,\mathbf{H}^{s-1}\mathbf{y}\right\}, \tag{24}$$

where $\mathbf{y},\mathbf{H}\mathbf{y},\cdots,\mathbf{H}^{s-1}\mathbf{y}$ are almost linearly independent vectors and span$\{\}$ denotes the set containing of all linear combinations of the vectors.

As mentioned earlier, the CG method has been utilized for the massive MIMO detection. A further improvement of the BER performance can be achieved by taking out a matrix-vector multiplication of the CG method, called the conjugate residual (CR) method. In addition, Cholesky factorization is utilized as a pre-condition algorithm to improve performance of the CR method. The proposed detectors have been employed for 32×128, 16×128, and 8×128 MIMO systems. The hardware architecture of the proposed methods is also proposed [94].

10) Coordinate Descent method: Coordinate descent (CD) is an iterative method that invert the high-dimensional linear system at low complexity. It obtains an approximate solution of a large number of the convex optimization using series of simple, coordinate-wise updates. The estimated solution can be concluded as

$$\hat{\mathbf{x}}_k = \left(\|\mathbf{h}_k\|^2 + N_0\right)^{-1} \mathbf{h}_k^H \left(\mathbf{y} - \sum_{j \neq k} \mathbf{h}_j \mathbf{x}_j\right), \qquad (25)$$

where N_o is the noise variance. (25) is computed sequentially for each user $k = 1, \cdots, K$, where the new result will be used immediately for the kth user in subsequent steps. These procedures will be repeated for a total number of K iterations. In [95], [96], low complexity optimized CD has been proposed with a corresponding high-throughput FPGA design for large-MIMO systems. The proposed FPGA reference design outperforms the existing approximate linear detectors in terms of hardware efficiency and BER performance. The proposed algorithm can support tens of users with hundreds of BS antennas.

As mentioned earlier, the above-mentioned methods can be utilized to avoid a direct matrix inversion, and hence, achieve complexity reduction. Meanwhile, complexity and performance will be influenced by the initial values, the number of iterations, and the relaxation parameter. By avoiding the cons shown in Table II, a balance between complexity and performance will be achieved.

B. Detectors Based on Local Search

Linear and nonlinear detectors become noncompetitive when used to serve massive MIMO systems because they require a matrix inverse calculation or QR-decomposition in which complexity is proportional to the number of antenna elements [97]. In order to achieve a satisfactory performance-complexity profile, detectors based on the local search have been proposed [98]. In the local search, the search focuses on a local region (neighborhood around it) one at a time and gradually approximates the best solution among the neighboring vectors. This process is continued for a certain number of iterations where the initial solution should be provided and it can be a random point from a suboptimal MIMO detector such as the ZF algorithm or the MMSE algorithm. After that the algorithm is terminated and the solution with a minimum cost function in the subset that called as the local optimum will be selected. The process can be iterated several times by changing the initial solution, the stopping criteria and escape strategies. The vector with the highest likelihood in the explored space will be identified as the solution [99].

This subsection reviews two local neighborhood search based algorithms, namely the likelihood ascent search (LAS) algorithm and the reactive tabu search (RTS) algorithm.

1) Likelihood Ascent Search: The concept of LAS is based on starting with an initial solution and keeps searching the neighborhood for a better solution. Usually, the initial solution vector is given by the linear ZF or the MMSE detector. An example of the neighborhood around the initial solution is the set of all vectors which differ from the initial solution in one coordinate [100]. The LAS algorithm includes several substages and each substage may consists more than one iteration. The iterations continue till the local optimum is reached in a substage. Afterwards, the second substage iterations start where two symbols update is applied. If the likelihood increases, the algorithm returns to the one symbol update stage. Otherwise, the algorithm moves on to a three symbol update and so on until the neighborhood fails to increase the likelihood. LAS can achieve a near-ML detection with a linear computational complexity [101]. The earliest near-optimal massive MIMO detector can be found in [49] where the LAS detector searched a sequence of bit vectors with a monotonic likelihood ascent.

LAS detector has been adopted to decode 16×16 and 32×32 space-time block codes (STBC) and reported interesting results that potentially enable the implementation of massive MIMO systems in [102]. In [103], a hardware implementation of a concatenated detector based on the LAS algorithm and turbo codes has been tested on 32×32 MIMO system after getting the initial solution using the ZF detector. It achieves more than 170 Mbps with 64 QAM with BER $10^{-1.5}$ - 10^{-2}. The work in [49] has been extended in [104] for several stages on 20×20 and 100×100 MIMO systems. The initial solution has been obtained from the output of the MF. Simulations show that the BER performance has been improved with complexity of $O\left(K^2\right)$. It is well known that not every vector in neighborhood search reduces the ML cost and some of them may cause an increase in ML cost. In order to reduce the size of the neighborhood search, a selection rule has been proposed in [105] which minimizes the ML cost. The simulation results on 32×32 MIMO system show a significant reduction in complexity while maintaining the BER performance. Motivated by the research in [106], researchers in [107] generalized the selection metric to reduce the size of the neighborhood, hence, complexity is reduced. Simulation results on 32×32 and 64×64 MIMO systems show a BER improvement and complexity reduction.

Simplicity is the key advantage of LAS detectors. In turn, the LAS algorithm suffers from the local minima that it first encounters and considers this minima to be the final solution vector. Performance of LAS detectors is also deteriorated when the modulation order increases. In addition, LAS detectors require a very large number of antennas to achieve the optimal performance. This number of antennas increases as the modulation order increases [108]. Moreover, the initial solution computation (which includes a matrix inversion in the ZF method and the MMSE method) increases the computational complexity.

2) Reactive Tabu Search: Reactive tabu search (RTS) is a more competitive local neighborhood search which adds additional restrictions to avoid an early termination. The RTS algorithm also starts with an initial solution vector. It imposes the search to visit several neighborhood solutions to achieve a satisfactory performance. In defining the neighborhood in a given iteration, the RTS algorithm seeks to avoid cycling by making moves to solution vectors of past few iterations as ”tabu”. In other words, the RTS algorithm is banning certain vectors from being included in the neighborhood list. Due to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

the escaping strategy, RTS can find better minimas. Unlike the LAS algorithm, the RTS algorithm involves several parameters such as the stopping criteria parameters, initial tabu period, maximum number of iterations [100]. Therefore, the RTS algorithm is capable of outperforming the LAS detector at complexity overhead. In [109], the RTS based decoding of 12×12 STBC has been proposed. In [110], the RTS detector depends on running multiple tabu searches where each search starts with a random initial vector and selects the best vector among solution vectors. The RTS algorithm based on multiple random restarts and threshold stopping criterion for under-determined MIMO systems has been proposed [111]. In [105], the RTS algorithm with a selection metric has been proposed for generating a reduced neighborhood set. Compared to the ML detector, simulation results for 32×32 MIMO systems show a significant complexity reduction with a satisfactory BER performance.

Although the RTS based detector achieves a near-ML performance and lower complexity than the ML technique, and the LAS algorithm, it suffers high computational complexity and performance degradation when the modulation order increases.

C. Detectors Based on Belief Propagation

Detectors based on the local search is usually need to compute an initial solution vector which increases complexity which is not required in belief propagation (BP). In addition, the LAS algorithm gets trapped in the local minima problem. In turn, detectors based on the BP algorithm are less likely suffer from the local minimum problem and achieve a better performance in general [98].

The BP algorithm is a tree search based algorithm which recursively searches for the optimal solution in a reduced search space. In signal processing, large variety of algorithms can be viewed as examples of the probability / belief propagation, which work by passing the message in a graphical model. For instance, the Baysian belief networks and Markov random fields, the turbo codes, the low density parity check (LDPC) codes [112], [113] and the message passing are popular examples of utilizing BP in CDMA [114] and MIMO detection [115], [116].

In the BP algorithm, the channel between the transmit and receive antennas are presented by channel response (h_{ij}) as illustrated in Fig. 4 where j and i denote the transmit and receive antennas, respectively. Each transmit antenna transmits individually an independent symbol. The transmitted symbols are summed at the receive antenna with different weights according to gains of the channel. It is clear that the transmitted symbols and received signals are mutually dependent. This property can be exploited to model the MIMO system by a factor graph (Tanner graph) [117] as presented in Fig. 4(b). Factor graph can be utilized to remove the interference between the transmitted symbols. At the transmitter side, symbol nodes have information on transmitted symbols independently. Sequentially, the signal observed at the receiver will be stored in the observation node. Therefore, reliability messages δ_{ij} and β_{ij} are iteratively exchanged and transferred between both symbol node and observation node as shown in Fig. 4(b). The channel response will be utilized to determine the coupling strength and the number of major connections. The nodes in Tanner graph are connected to each other, thus, the graph contains many loops where many of them could cause performance degradation in the message passing. The number of major connections will not become so large with high number of antenna elements. Therefore, the effective number of loops with high impact is a few which suits the detection in massive MIMO systems to achieve low complexity [118]. The transmitted signals are detected at each observation node and the result is passed as a message (extrinsic information) to each symbol node. The extrinsic information (priori values for the *jth* node are summed, and a posteriori log-likelihood ratios (LLR) of each bit is calculated and utilized for a decision output after a certain number of iterations. In other words, the MIMO channel can be illustrated as a certain graphical model, while detection of the channel input is equivalent to performing inference in corresponding graph [119]. The posteriori probability of each transmitted symbol is approximated by passing messages that marginalize over other symbols in a factor graph. This process will be repeated until achieving the convergence. The BP based detectors achieved a near-optimal performance when the number of antennas is large and the channel correlation is reasonably low [98]. On the other hand, the convergence degrades in a bad conditioned factor graph.

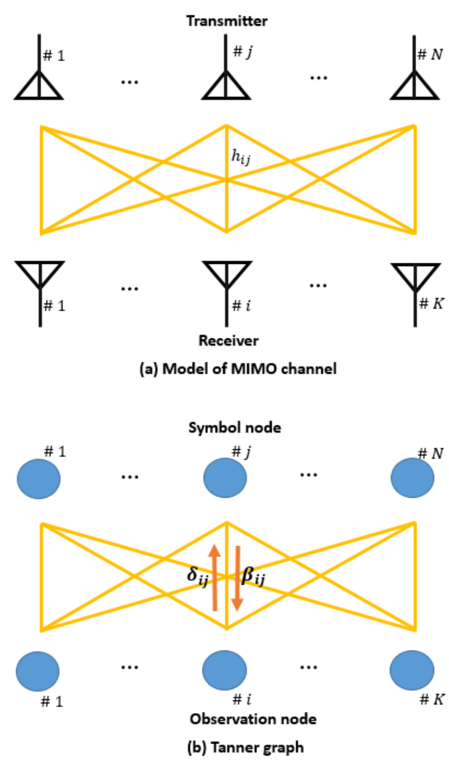

Figure 4. Example of a Tanner graph for a MIMO channel [98]

In [120], an analysis of the message passing detector based on BP is provided with several parameters, such as the number of users, the number of antennas, and the damping factor. The proposed detector achieved performance gain over a factor graph BP algorithm without extra expenses of computational complexity.

In [121], a detector based on the BP algorithm and message passing on Markov random field presentation has been proposed for non-orthogonal STBC 16×16 and 24×24 systems. This work has been extended in [122] using factor graphs. A Gaussian approximation based on BP has employed in 16×16 MIMO system to significantly reduce complexity [123], [124]. In [125], the Markov random field, the message damping and the Gaussian approximation of interference methods have been utilized with a local neighborhood search algorithm to enhance the performance of 16×16, 24×24, and 32×32 MIMO systems. The proposed detector has achieved a significant complexity reduction. In [126], the extrinsic information transfer (EXIT) of a factor graph based on a message passing algorithm has been utilized for a detection in 16×16, 64×64 and 256×256 MIMO systems. A detector based on the EXIT and the BP has been implemented in [118] using 100×100 MIMO system. Another detector based on the BP has been implemented using an antenna array of 100 elements with second-order calculations [98]. In [127], a hardware architecture with a parallel pipline and a simple logic structure has been presented for the 32×32 MIMO system with 4-QAM.

Minimum Kullback-Leibler (KL) divergence criteria has been exploited to approximate the original discrete messages with continuous messages [128], [115]. This detector achieved a near-optimal performance for a 64×64 MIMO system. Channel hardening has been exploited for the detection purposes using the MF, the MMSE, the message passing, the low density parity check (LDPC) and the exit chart matching for 128×128 and 128×32 MIMO systems [31], [129], [130]. Also, BP and LDPC based detectors have been addressed in [131] for large MIMO systems. In addition, a generalized approximate message passing detection with EXIT has been proposed in [132].

A hybrid RTS and BP has been proposed for 16×16 and 32×32 MIMO systems [133]. The output of the BP algorithm is fed back to the RTS algorithm for the next iteration. The hybrid algorithm performed better than the RTS algorithm at the expense of an extra complexity.

D. BOX Detection

Infinity norm or a BOX-constrained detection is a convex relaxation of the ML decoder where the signal can be recovered through the efficient convex optimization followed by a hard thresholding [134], [135]. As shown in Fig. 5, it relaxes the constraint $x\in C^K$ to a convex hull around the constellation set [134], [136].

In [96], a CD based BOX equalization has been implemented for a 128 BS antenna and 8 users system. It shows a satisfactory hardware efficiency with low hardware complexity. In [136], a detector based on a BOX-constrained equalization and alternating direction method of multipliers (ADMM) has been proposed. In addition, an implementation of the VLSI architecture for 16×16 massive MU-MIMO system is provided. In order to outperform performance of linear detectors, the ratio between the number of user terminals and the number of BS terminals is small (less than two). A derivation of the symbol error rate (SER) for the BOX-decoder in the large system is provided in [135]. The proposed BOX decoder significantly outperforms the linear ZF and MMSE decoder.

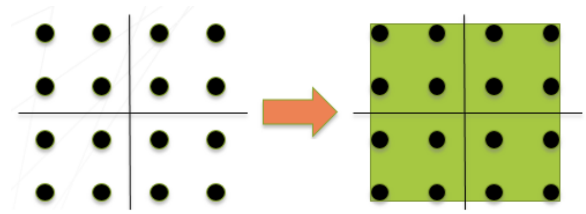

Figure 5. Idea of BOX-constrained detection

Table II presents the pros and cons of several detection algorithms in massive MIMO systems.

E. Sparsity based Algorithms

In the massive MIMO model, a large number of antennas increases the degree of freedom and hence, the magnitude of some of channel coefficients becomes zero or a small enough to be neglected [142]. The linear transformation of $\mathbf{H}$ to be a sparse is given by

$$\mathbf{H}=\mathbf{T}\,\mathbf{H}_s\,\mathbf{B}, \tag{26}$$

where $\mathbf{T}$ and $\mathbf{B}$ are unitary matrices to perform a linear transformation of $\mathbf{H}$ to get a sparse matrix $\mathbf{H}_s$. Using the compressive sensing (CS), a channel matrix $\mathbf{H}$ can be estimated by exploiting the sparse structure of $\mathbf{H}_s$ [143]. The main idea behind the CS relies on the fact that the sparse signals $\mathbf{z}$ can be reconstructed from compressed measurements $\mathbf{y}=\Phi\mathbf{z}$ through a convex programming provided that the signal to be recovered is a sparse (i.e., the number of zero elements in the vector is large) [144]. It can be used even when the length of measurement vectors ($\mathbf{y}$) is less than the length of $\mathbf{z}$ [145] where the model is shown as

$$\mathbf{y}=\Phi\,\mathbf{z}+\mathbf{n}, \tag{27}$$

where Φ is the measurement process and it should be a full rank matrix and $\mathbf{n}$ is a noise vector.

In [146], a multipath matching pursuit (MMP) detector using the SD technique has been proposed. It depends on identifying the location of errors in the low complexity initial solution and then a residual update strategy is applied to improve the localization accuracy by using a thresholding function. Finally, the SD-MMP algorithm will be applied to reduce the variation in the number of errors. The detector has been examined in 32×32, 64×64 and 128×128 4-QAM and 16-QAM MIMO systems where the complexity is $O\left(NK^2\right)$. In [137], a hidden sparsity resulting from the decision feedback equalization has been exploited to iteratively boost the detection with the computational complexity equals to $O\left(K^3\right)$. The CS technique has been correcting the symbol error from the output of linear detectors [147]. Figure 6

Table II
PROS AND CONS OF DETECTION ALGORITHMS IN MASSIVE MIMO SYSTEMS

Algorithm	Pros	Cons
MF	• Low complexity. • It Works properly if columns of the propagation matrix are nearly orthogonal.	• It achieves low performance in the ill-conditioned environment.
ZF	• Low complexity. • It has a satisfactory performance in interference-limited environments and impose noise enhancement. • Perform better than the MF.	• Low performance in the ill-conditioned channel matrix. • In a square massive MIMO, it does not improve either the diversity gain or the computational complexity.
MMSE	• Low complexity. • It reduces the noise enhancement. • In a medium and high SNR, it outperforms the ZF algorithm.	• Low performance in the ill-conditioned channel matrix. • In a square massive MIMO, it does not improve either the diversity gain or the computational complexity.
SIC	• It outperforms the ZF method and the MMSE method.	• Performance is influenced by the first detected signal. • Compared with the ZF method and the MMSE method, it has high computational complexity.
LRA	• It modifies the ill-conditioned channel matrix to be more orthogonal. • Good performance.	• High computational complexity.
SD	• ML performance can be achieved. The *K*-best variant of the list size LSD can be configured to maintain a balance between complexity and performance.	• High computational complexity when the list size is not fixed.
BP	• The ML performance can be achieved when the channel correlation is low.	• It is difficult to find the optimal damping factor. • Performance degrades in a bad conditioned factor graph. • In general, the convergence is not always guaranteed in this iterative method.
Local search	• It minimizes the ML cost in a fixed neighborhood.	• Complexity depends on the size of the neighborhood. • Not every vector in the neighborhood causes reduction in complexity.
BOX Detection	• A satisfactory hardware efficiency with low hardware complexity [96].	• Poor performance when the ratio of the number of user terminals to the number of BS antennas is close to one.
Sparsity based algorithms	• The ML performance can be achieved. • Complexity is lower than local search algorithms [137].	• In the CS, the number of local minima which produces convergence errors would be increased by highly sparse constraints [138]. • Methods such as a sparse Baysian learning (SBL) are well suited for handling local minima at the expense of high complexity [138].
The NS method	• Low complexity.	• It suffers from a considerable performance loss when the ratio between BS antennas and user antennas is large (close to 1). • Approximation converges slower in comparison with the Newton iteration [52][86][78].
NI	• A fast convergence can be achieved if the condition in (13) satisfied.	• It requires more calculations to obtain the initial estimation [67], [139].
The GS method	• It achieves a near-optimal performance even when the ratio between BS antennas and user antennas is close to one.	• Due to an internal sequential iterations structure, the GS method is hard for a parallel implementation [71], [74], [75].
The SOR method	• It achieves a near-optimal performance even when the ratio between BS antennas and user antennas is large. • On a Xilinx Virtex-7 FPGA, the SOR based detector can achieve more than 3x throughput per slice compared with the NS based detector and 2x throughput per slice with the CG-based detectors [76].	• The *Gram* matrix should be pre-computed and provided as an input which increases the computational complexity [61]. • It has an uncertain relaxation parameter $0 < \omega < 2$.
The Jacobi method	• It achieves a near-optimal performance when the ratio between BS antennas and user antennas is small. • It can be implemented in a parallel manner [79].	• It converges slowly, and hence, implying higher latency [75]. • Performance is not improved over iterations when the ratio of user terminals to BS antennas is close to one.
The Richardson method	• It has a hardware friendly approach with a reasonable performance [83].	• It requires a large number of iterations to converge [81], [140]. • It needs a stability to ensure convergence and the spectral radius of the matrix should be less than 1 [83]. • It has an uncertain relaxation parameter $0 < \omega < \frac{2}{\lambda}$ where λ is the largest eigenvalue of the $\mathbf{H}$ [70], [83].
The CG method	• It achieves a near-optimal performance when the ratio between BS antennas and user antennas is large [85].	• It requires a large number of iterations [85][70]. • It includes many division operations [85][70].
The Lanczos method	• It converges to performance of the MMSE method within few iterations [141]. • It suits a parallelism-optimized hardware architecture [141].	• It has low BER performance or require high computational complexity under time-varying channels [92].
The residual method	• It achieve a satisfactory performance even when the ratio between BS antennas and user antennas is large.	• It requires a pre-conditioning algorithm for a satisfactory BER performance [94].
The CD method	• It achieves a satisfactory performance even when the ratio between BS antennas and user antennas is large.	• As shown in (25), the estimated solution has an inverse component that increases complexity [95], [96].
The Lanczos method	• It converges to performance of the MMSE within few iterations [141]. • It suits a parallelism-optimized hardware architecture [141].	• It has low BER performance or require high computational complexity under time-varying channels [92].

shows an example of the post detection sparse error recovery technique. It includes a conventional detector followed by a slicer as well as a sparse error recovery algorithm such as the orthogonal matching pursuit or MMP.

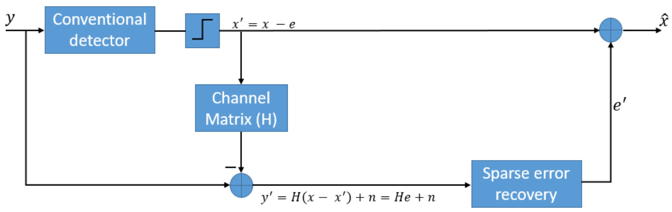

Figure 6. Example of overall structure of the post detection sparse error recovery technique [147]

F. Machine learning for massive MIMO detection

Machine learning uses algorithms to build analytical models, assisting computers learn from data. In last decade, machine learning has success stories in the big data, the speech recognition, the natural language processing, the computer vision, and so on. There are interesting applications of machine learning in communication system such as the detector design for the MIMO transmission [148], [149], [150], [151], the channel estimation based massive MIMO [152], [153], and the pilot allocation for the massive MIMO [154]. A survey of machine learning techniques utilized in communication systems can be found in [155].

In signal detection, the classical detection theory tries to choose the best estimate of the unknown symbols while machine learning tries to select the best algorithm to be applied. In machine learning, detection rules are very crucial in learning process whereas the key hypotheses in classical detection are the unknown symbols [148]. In machine learning, "learning" the algorithm is the most computationally expensive stage and it can be completed off line. Once the optimal rule algorithm is found, it can be implemented in real time systems with low computational complexity. However, the computational complexity will be increased once we get a new observation. In last few years, the "deep" revolution has been witnessed which is associated with the use of complicated and significant classes of algorithms (also known as architectures) such as the neural networks with many non-linear operations and layers. A comprehensive overview of machine learning and deep learning can be found in [156]. In a standard neural network (NN), many simple processors (neurons) are connected to produce a sequence of real-valued activations. Input neurons get activated through sensors recognize the environment while other neurons get activated through weighted connections from formally active neurons. Learning is about finding weights that make the NN achieve the desired behavior. It depends on the number of neurons, therefore, the desired behavior may require long chains of computational stages, where each stage transforms the aggregate activation of the network. Deep learning is about precisely specifying credit across many such stages [156].

Although, deep learning is a promising approach, it is not yet well-investigated in the massive MIMO detector design. However, there is a limited work in this field. In [157], an investigation on how techniques from deep learning can be utilized to train a detection algorithm from samples of transmitted and received signals is conducted. In [148], a deep learning network for 30×60 MIMO detection is proposed. The results show that the proposed deep networks can achieve high accuracy and low complexity even in the ill-conditioned channels. In [158], a deep learning detector integrated with the BP algorithm has been proposed for 8×16 MIMO system to further improve the detection performance of BP algorithms. Compared with the BP detector, the proposed detector has achieved performance improvement with low complexity.

Table III presents the chronology of detection algorithms in massive MIMO systems. To our best knowledge, the earliest detector in the context of massive MIMO has been proposed in 2008.

V. Application of Small-Scale MIMO Detectors for Massive MIMO

A plethora of small-scale MIMO detectors exists in the literature and many of their application for massive MIMO has not been explored. In this section, we explore few nonlinear small-scale MIMO detectors briefly which have been used for massive MIMO systems.

A. Successive interference cancellation

It is a nonlinear detector based on the linear detector such as ZF and MMSE [211]. It cancels the interference caused by multiple antennas. The detection and canceling are performed in a serial fashion. First, a signal will be selected and detected using the linear ZF or MMSE detector. The interference of the detected signal is canceled. Then, the second signal is detected and canceled from the remaining signals set, and so on. This process will be repeated until all signals are detected [212]. Figure 7 presents an example of a SIC detector.

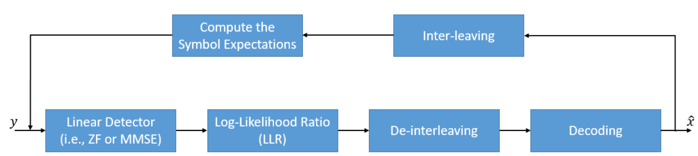

Figure 7. Example of a SIC detector

Performance of the SIC detector will be influenced by the first detected signal, thereby, the signal with the highest signal-to-noise-plus-interference (SINR) is detected first to achieve the best possible error rate performance [213]. Then, the second strongest signal will be detected and canceled from the remaining signals set. The process is iterated until all signals are detected. This method is often called ordered-SIC (OSIC) or V-BLAST which improves the diversity gain in low and moderate SNR [214] where (28) shows the bias removal in MMSE filter.

$$\mathbf{W}_{SIC} = \mathbf{B}\,\mathbf{A}_{MMSE}^{H}. \tag{28}$$

where $\mathbf{B}$ is a diagonal matrix and $\mathbf{A}_{MMSE}^{H}$ is shown in (8). The OSIC algorithm outperforms the ZF method and

Table III
CHRONOLOGY OF DETECTION TECHNIQUES FOR MASSIVE MIMO

Year	Summary of work performed	Reference
2008	Proposed a multistage LAS rooted in Hopefield neural networks.	[159]
2008	The LAS detector has been proposed for large-scale MIMO systems.	[49], [102]
2008	A detector based on LAS and turbo codes has been implemented.	[103]
2009	Updated-based LAS algorithm has been proposed by utilizing the multistage multi-symbol update based strategy.	[160]
2009	A detector based on the MCMC Gibbs Sampling method has been proposed.	[161]
2009	A detector based on the RTS algorithm has been proposed for non-orthogonal space-time block codes with large MIMO.	[109]
2009	A detector based on the BP algorithm for non-orthogonal STBCs with large dimensions has been proposed.	[121]
2010	A detector based on the LAS algorithm has been proposed.	[162]
2010	A detector based on the RTS algorithm has been proposed.	[110]
2010	A detector based on the RTS algorithm has been proposed.	[107]
2010	A detector based on the BP algorithm, Markov random fields, and factor graphs has been proposed.	[122]
2011	A hybrid detector based on the RTS and the BP has been proposed.	[133]
2011	Proposed a layered local neighborhood search based on lower bound on the ML bit performance.	[163]
2011	A class of the BP based detector has been proposed using message passing on graphical models.	[125]
2011	A particle swarm optimization (CPSO) and factor-graph have been utilized in the detection scheme.	[164]
2011	A detector based on tree approximation of the Gaussian distribution has been proposed.	[165]
2012	A detector based on the ML detection and genetic optimization has been proposed	[166]
2012	A detector based on the RTC algorithm for underdetermined massive MIMO has been proposed.	[111]
2013	Monte-Carlo-Sampling based massive MIMO detection algorithm has been proposed.	[167]
2013	Proposed an element-based lattice reduction algorithms based on minimizing the diagonal elements in the noise covariance matrix.	[168]
2013	Decision feedback (DF) detection algorithm has been proposed.	[169]
2013	An improved lattice reduction aided algorithm has been proposed.	[170]
2013	A detector based on the ML algorithm and a heuristic programming method has been proposed.	[171]
2013	A detector based on extrinsic information transfer (EXIT) and message passing algorithm has been proposed.	[126]
2013	A detector based on extrinsic information transfer (EXIT) has been proposed.	[118]
2014	Exploit the sparsity produced by the output of the conventional linear detectors and employ the compressed sensing techniques to correct the errors.	[147]
2014	Message passing algorithm through factor graph has been proposed.	[128]
2014	Proposed a receiver based on message passing detector by exploiting a Gaussian approximation on the off-diagonal terms of the channel matrix.	[129], [31]
2014	A detector based on the MMSE method and the SOR algorithm has been proposed.	[77]
2014	A detector based on the MMSE method and the Richardson method has been proposed.	[172]
2014	A sequential decoder based on the Fano algorithm has been proposed.	[173]
2014	A detector based on convex optimization has been proposed.	[174]
2014	A detector based on non-binary the BP algorithm and Gaussian approximation has been proposed.	[131]
2015	The proposed algorithm depends on leveraging the hidden sparsity produced from the decision feedback equalization to iteratively boost the detection.	[137]
2015	Proposed iterative neighbourhood search algorithms such as the LAS algorithm and RTS algorithm.	[105]
2015	Proposed a half sparse decomposition of the data signal vector to relax the ML problem into another minimization problem.	[175]
2015	Proposed an iterative decoding strategy by exploiting the fact that the transmit constellation is discrete, and hence, re-model the channel with a sparse input belonging to the binary set (0,1).	[176]
2015	Used the CG method to transform the MIMO detection into minimizing the quadratic function.	[177]
2015	The proposed algorithm exploits advantages of the MMSE method property and the relaxation iteration (RI) method to avoid a matrix inversion.	[178]
2015	Proposed an iterative NS expansion algorithm for the MMSE method to avoid the direction computation of the matrix inversion.	[62]
2015	Proposed a two stages multi-branch linear minimum output energy (MOE) receiver to collect symbols from different paths and then, selects the less samples.	[179]
2015	Detection scheme based on adaptive reduced rank receive processing has been proposed.	[180]
2015	Analyzed the effects of the ratio of the number of massive MIMO antennas to the number of users based on the approximation of the ZF-Matrix inversion method.	[181]
2015	MMSE with the GS method have been proposed to perform a detection process.	[71]
2015	A two stage quadratic programming (QP) detector has been proposed	[182]
2015	A detector based on quadratic minimization and convex constraints has been proposed	[183]
2015	A detector based on the MF and the LAS algorithm has been proposed.	[104]
2015	A detector based on the Gaussian approximate BP (GABP) has been proposed.	[123]
2016	Proposed an approximate matrix inverse suffices for finding linear and nonlinear detector solutions such as the ZF method, the MMSE method and the SD algorithm.	[68]
2016	A detector using Markov chain Monte CArlo (MCMC) strategy has been proposed for the massive MIMO system.	[184]
2016	Proposed a concatenation based improved error localization (CBIEL) detector which exploits the sparsity characteristics offered by the MMSE output vector.	[185]
2016	Quadratic programming (QP) and branch and bound (BB) are developed to achieve low-complexity and high performance detectors.	[108]
2016	A scalable soft detection method is proposed based on the Richardson method.	[83]
2016	Proposed a sorted-decision-feedback differential detection (DFDD) in combination with noncoherent decision-feedback (nDFE).	[186]
2016	This article discussed the feasibility of an online failure detection algorithm for massive MIMO applications.	[187]
2016	Proposed ML detection scheme with the assistance of a two stage ranking mechanism for massive MIMO systems.	[188]
2016	Proposed an adjustable SD partitioning to the transmission channel with low latency overhead.	[189]
2016	The NS expansion method based on the MMSE detection algorithm is proposed. A direct matrix inversion method has been replaced by matrix-vector multiplications.	[61]
2016	The MAP estimation-based error recovery method is proposed. It is based on the fact that the error vector is not only sparse but also discrete.	[190]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

2016	Exploits the Jacobi method in the linear detection to reduce the intensive matrix inversion and also proposed a multiplication-free initial estimate for the Jacobi method to reduce complexity.	[79]
2016	Proposed a turbo detection scheme and used an outer forward error correcting (FEC) code feeded by sparse vector.	[191]
2016	Proposed a block iterative support detection (block-ISD) that exploits the block sparsity inherent in the block-sparse equivalent channel response (CIR) generated by considering the spatial correlations of the MIMO channels.	[192]
2016	Proposed a joint channel estimation and data detection algorithm to achieve optimality in generalized likelihood ratio test (GLRT).	[193]
2016	Proposed a detection method based on GS method. The NS method is also employed for initialization.	[73]
2016	Proposed a polynomial expansion (PE) method for matrix inversion based on a diagonal band NI.	[67]
2016	A detector based on Element-based lattice reduction and the K-Best algorithm has been proposed.	[194]
2016	A detector based on the MMSE method and the CD algorithm has been proposed.	[95]
2016	A detector based on the BP algorithm has been proposed.	[127]
2016	A detector based on semidefinite relaxations, convex and concave quadratic has been proposed.	[195]
2016	A detector based on the local search has been proposed.	[99]
2016	A detector based on multiple feedback SIC and shadow area constraint (SAC) has been proposed.	[195]
2016	A detector based on the normalized MF belief in Gaussian BP has been proposed.	[124]
2016	A detector based on Gaussian message passing iterative technique has been proposed.	[196]
2017	Proposed an iterative sequential detection algorithm. In every iteration, symbol transmitted from each user is detected and updated sequentially while nulling the interference from all other users.	[197]
2017	Proposed a SD multipath matching pursuit (SD-MMP) algorithm.	[146]
2017	Proposed a parallel detection with QR decomposition and a M-Algorithm (QRM-MLD).	[198]
2017	Proposed a hardware architectures of the Richardson method based massive MIMO detectors.	[58]
2017	Reduced the number of required quadratic programs (QPs) by using a likelihood based branching criteria and a node selection strategy.	[199]
2017	Proposed a kernel Hilber space (RKHS) based block symbol detector that works on decomposed blocks of the observations, and selectively decides the use of an incoming observation.	[200]
2017	In order to avoid more redundant visited nodes in the tree search, ordered sphere decoding (OSD) is divided into multiple OSDs.	[201]
2017	Exploits the channel sparsity to directly factorizing the received signal matrix.	[202]
2017	Investigated performance of linear detectors for uplink scenario over correlated Rician fading.	[63]
2017	Two methods are proposed. In the first method, channel matrices are pre-processed by the BP prior to the normal BP algorithms. In the second method, channel matrices are pre-processed by a de-correlation matrix, then followed by normal BP algorithms.	[115]
2017	Proposed an approximate detection method based on the NI, and also proposed upgrade methods named the NI with iterative refinement (NIIR) and diagonal band NIIR (DBNIIR) which combine the NI and the DBNI method with the iterative refinement (IR).	[69]
2017	Proposed an improved convex semidefinite relaxation detector (RFRD) based on the LAS algorithm for detecting high-order modulation signals in the massive MIMO system.	[203]
2017	A detector based on the MMSE algorithm and the symmetric SOR has been proposed.	[204]
2017	A detector based on the MMSE method and the GS method has been proposed.	[205]
2017	A detector based on multiple feedback and ordered SIC has been proposed.	[206]
2017	A detector based on the MCMC layered Gipps sampling (GS) algorithm has been proposed.	[139]
2017	A detector based on the K-best SIC has been proposed.	[207]
2017	A detector based on a generalized approximate message passing (GAMP) has been proposed.	[132]
2017	A detector based on the MCMC and Gibbs sampling algorithm has been proposed.	[139]
2017	A detector based on message passing and BP has been analyzed.	[120]
2017	A detector based on deep learning network has been proposed and examined in ill-conditioned channels.	[148]
2018	Utilized the stair matrix instead of the diagonal matrix in the detection scheme.	[208]
2018	Proposed a new iterative method using the stair matrix to achieve the symbol estimation close to the linear MMSE estimation.	[209]
2018	Message passing detection algorithm with no division or exponential operations is proposed.	[130]
2018	A feasibility study on various linear and non linear detectors including lattice reduction techniques have been conducted.	[210]
2018	Deep learning has been utilized to train the detection algorithm from samples of transmitted and received signals.	[157]
2018	A deep learning detector integrated with the BP algorithm has been proposed.	[158]

the MMSE method but also it suffers from a considerable computational complexity [108] [188]. In [108], two stages of the quadratic programming (QP) detector with a SIC algorithm are proposed and examined in different square massive MIMO configurations.

B. Lattice reduction-aided algorithms

Performance of linear detectors can be improved by modifying the ill-conditioned channel matrix to be more orthogonal using LRA methods [215], [216]. Figure 8 shows the decision regions before and after the LRA. The decision regions generated by non-orthogonal basis vectors are less immune than the decision regions generated by orthogonal basis vectors. The new channel matrix ($\widetilde{\mathbf{H}}$) is given by

$$\widetilde{\mathbf{H}} = \mathbf{H}\mathbf{T}, \tag{29}$$

where $\mathbf{T}$ is a unimodular matrix [217]. Hence, (2) can be written as

$$\mathbf{y} = \widetilde{\mathbf{H}}\mathbf{z} + \mathbf{w} \quad where \quad \mathbf{z} = \mathbf{T}^{-1}\mathbf{x}. \tag{30}$$

The aim of LRA algorithms is to find the vector of shortest length, which results in lower orthogonality deficiency and hence, it achieves a considerable performance. Many LRA algorithms have been proposed such as the Minkowski reduction, the Hermite-Korkin-Zolotarev (HKZ) reduction, the Lenstra-Lenstra-Lovasz (LLL) and its complex counterparts [218], the Seysen's algorithms (SA) [219] and an element-based lattice reduction (ELR) [220]. Linear or SIC detectors can be preprocessed using lattice reduction algorithms to achieve a significant performance gain and the computational complexity will be increased [215][217][221]. Complexity of LRA based detectors is independent of the employed signal constellation, hence, LRA detectors are hardware-friendly [222].

In order to enhance the asymptotic performance of linear detectors, two element-based lattice reduction (ELR) algorithms have been proposed to reduce the diagonal elements

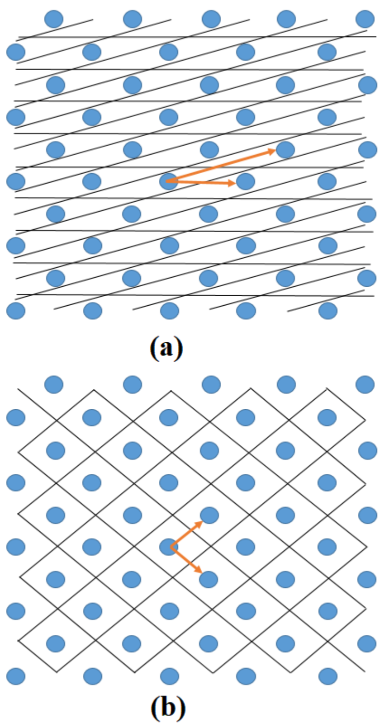

Figure 8. Idea of lattice reduction aided algorithms (a) before LRA (b) after LRA

of the noise covariance matrix of linear detectors [168]. The first algorithm is the shortest longest vector (SLV) reduction and the second algorithm is the shortest longest basis (SLB) reduction. Due to high complexity of the SLV algorithm and the SLB algorithm, two hybrid methods are proposed to find the sub-optimal solutions to the SLV and SLB reductions, namely, ELR-SLV and ELR-SLB reductions. The proposed detectors employed in 64×64 MIMO system with 256-QAMELR-SLB achieved better error performance than ELR-SLV at the cost of higher complexity [168].

In [223], a sequential reduction (SR) scheme has been proposed to emphasize reducing one lattice vector with another "close" vector generated from a subspace spanned by other basis vectors, and there is a freedom to choose this "close" vector. The proposed algorithm employed in a 50×50 MIMO system with 64QAM modulations. In [210], the LRA algorithm has been utilized with linear and non linear detectors for 128×128 MIMO system with 64QAM modulation and a near-ML performance has been achieved without an expense of higher complexity.

C. Sphere decoder

The main idea behind the sphere decoder (SD) algorithm is to search only through the constellation points that are confined within a sphere with a predetermined radius "d" [224][225]. Figure 9 shows the geometrical representation of the SD algorithm where the small blue nodes represent all possible transmitted symbols. By visiting all these nodes, the ML solution is achieved. In order to reduce complexity, the SD restricts the search within the sphere with a predetermined radius "d". SD algorithm is also considered as a tree research where the visited branches are dependent on the channel characteristics and the noise variance. Therefore, ML performance can be achieved an complexity can be reduced by eliminating the lattice points that lay inside the sphere as long as d is properly selected [226]. The channel matrix ($\mathbf{H}$) can be factorized by the QR decomposition to a unitary matrix ($\mathbf{Q}$) and an upper triangular matrix ($\mathbf{R}$). Therefore, the mathematical representation of the SD search [227] is given by

$$\hat{\mathbf{x}}_{SD} = arg \min_{x \in \mathbb{C}^K} \left\{ \|\hat{\mathbf{y}} - \mathbf{R}\mathbf{x}\|^2 \leq d^2 \right\}. \tag{31}$$

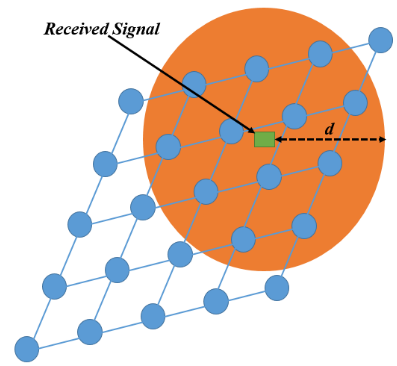

Figure 9. Idea of Sphere Decoder [228]

Several SD detectors have been proposed to improve the BER performance and to reduce the complexity [229], [230], [231], [232], [233]. Although SD avoids the exponential complexity of the ML detection [234], its average complexity is exponential in the number of transmit antennas. It is also not very hardware-friendly due to a variable complexity with different signals and channels. This leads to a non-fixed detection throughput which is not desirable in real time applications. In turn, the K-best SD algorithm is well appreciated in hardware implementation due to its fixed throughput and complexity. The K-best algorithm is also known as the M-algorithm or a beam search in the artificial intelligence literature. Unlike the conventional SD algorithm, which recursively explores all candidates within the initial radius, the K-best SD algorithm maintains only the best K branches for the next research level and other nodes are neglected [235], [236], [237]. This reduces the computational complexity at each level. The K-best is a sub-optimal algorithm and it does not pledge the ML performance because not all branches that satisfy the radius constraint are kept. It needs a large K value to achieve the ML performance, which increases the complexity and the power consumption. It also has performance degradation in general [238]. Therefore, a modified branch and bound (BB) algorithms have been proposed. In [108], another BB search tree algorithm is proposed to improve performance of an iterative QP detection. The proposed detector examined in 16×16, 32×32 and 64×64 MIMO systems with the complexity of $O\left(nNK^3\right)$. In this algorithm, reduced and controlled BB algorithm is proposed where the width and the depth of the BB tree are reduced. In [199], a likelihood based branching criteria to reduce the number of QPs required is proposed and tested in 32×32 and 64×64 16-QAM MIMO systems. In general, in large-scale MIMO, the SD algorithm and its variants, the K-best SD algorithm, and the BB demonstrate a poor performance.

VI. The Impact of Channel Estimation and Precoding on Massive MIMO detection

A. Channel Estimation

The detectors presented in Section IV assumes a perfect CSI is available at the receiver. In practical systems, it is not possible to obtain a perfect CSI due to imperfections of a channel estimation and quantization errors [239]. This imperfection is more severe in case of massive MIMO detection due to a pilot contamination. In [240], [241], [239], several symbol detection techniques whereby CSI is estimated from orthogonal pilot sequences. As the MIMO size becomes large, acquiring CSI through pilot sequences transmission is impractical because of a large overhead and the time required for a pilot transmission may exceed the channel coherence time. The CSI acquisition overhead can be compensated by the potential gain due to the sparsity of the massive MIMO channel in a transformed domain [202]. In [242], a channel estimation based on the eigenvalue decomposition (EVD) of the correlation matrix of received vectors has been proposed. This technique is sensitive to the accuracy of the sample correlation matrix as well as the size of the antenna array. This technique has been exploited in [243] where a blind channel estimation has been performed followed by a symbol detection based on the expectation propagation (EP) algorithm. Briefly, EP aims to find the closest approximation for the conditional marginal distribution of a required variable in an iterative refinement procedure. The link performance will be maximized using reliable CSI and precoding.

ML decoder is formulated under imperfect CSI and recursive tree search algorithms [244]. The authors proposed a recursive search algorithm which is similar to the sphere decoding. The proposed decoder results in a near-ML performance with a significant complexity reduction. In [239], performance of the EP detector in practical situations in which imperfect CSI is available has been investigated. Results show that the lack of perfect CSI produces a significant performance loss of the EP detector. In addition, the EP detector shows high sensitivity to the channel estimation error at high SNR. In order to avoid the mentioned problem, a modified EP detector with utilization of a correlation matrix of the channel estimation error has been proposed. The modified EP detector produced a significant improve in performance. A massive MMSE detector is proposed for a 1-bit quantization and a channel estimation error in [245]. A closed form expression is presented for the uplink achievable rate and the total system throughput is compared against conventional MIMO configurations with higher order modulation.

B. Precoding

The precoding techniques can be applied to simplify the receivers in multiuser MIMO. The interference of the transmission can be removed in the transmitter if the CSI is available at the transmitter side. This process is sometimes referred to as pre-equalization. Due to their similar operations, the linear MIMO pre-equalization methods can be viewed as a dual of MIMO detection. A comprehensive overview and comparison between different linear precoding techniques under both single-cell (SC) and multicell (MC) can be found in [11]. The survey presents linear methods such as MF, ZF and MMSE precoding which are counterpart of linear detection methods. In addition, advanced precoding schemes such as the H-infinity, the max-SINR and the multilayer precoding are also presented in [11].

In [246], approximate algorithms such as the polynomial expansion method, the conjugate gradient (CG), the Gauss-Seidel (GS) method, the Jacobi method and the Newton iterations have been utilized in detection and precoding purposes. The algorithms supposed to obtain enough precision within few iterations (1 or 2). In order to improve the precision with little complexity cost, the approximate algorithms have been combined with the iteration refinement (IR). In [85], a precoder using the CG method is proposed to be utilized with realistic antenna configurations. The SOR based precoding scheme has been proposed in [247] and it shows faster convergence rate than the Neumann-based precoding. The SOR based precoding scheme is very sensitive to the relaxation parameter selection. Therefore, a precoder based on weighted symmetric successive over relaxation (WSSOR) has been proposed to reduce the complexity of the matrix inversion [248]. It also has a simple method to select the relaxation parameter and weighting factor based on the configuration parameters. Similar to the detection problems, the precoding methods are typically reliant CSI at the transmitter side.

VII. Summary and Conclusions

Massive MIMO is destined to provide great and improved user experience, delivery of new revenue generated exciting mobile services. Consequently, massive MIMO would remain a strong competitor in the next decade for both developed and emerging markets. A significant research dedicated to the receiver's design has been proposed. In this paper, a review of various detection techniques for massive MIMO systems is provided. Although linear detectors suffer from mediocre performance, the ZF method and the MMSE method are found to play a crucial role in the receiver design due to their relative simplicity. They are also used in the initialization and preprocessing for other detectors. Local search and BP based detectors may achieve a promising balance between performance and complexity. In addition, non linear detectors achieved a near-optimal ML performance but with high computational complexity.

A sparse representation of the physical channels, e.g., via the virtual channel model (VCM), is a potential direction of new innovation for decoding in massive MIMO. By employing the VCM for a uniform linear antenna arrays and under a flat fading produces, a large number of zero components will be obtained. Precoding for massive MIMO systems using VCM sparsity has been presented in [43]. Additionally, machine learning based massive MIMO detectors are at an early stage

[148]. Instead of the classical detection theory which tries to find the best estimate of the unknowns, machine learning could choose the best algorithm to be applied. The learning stage is computationally expensive but it can be performed off-line to find the optimal algorithm. In addition, the K-best SD (KSD) and the ordered SD (OSD) are not yet investigated in the context of massive MIMO. Although the complex-valued modulation constellations are often employed in digital communications, performance of several detectors have not been explored yet for different constellations such as performance of the BOX detector for 256-QAM. Besides that, detectors are using the diagonal matrix. However, utilization of the stair matrix[2] in NS expansion has been proposed in [208] and it shows a promising complexity-performance trade-off in the context of massive MIMO. By using the diagonal matrix, the normalized MSE is always higher than that of utilizing the stair matrix in the same system configuration. In addition, a detector based on the stair matrix requires less iterations to achieve the same level of the MSE in using the diagonal matrix. As a result, the detector with a stair matrix requires less computational complexity in implementation. Therefore, the work in [208] can be extended to test the efficiency of the stair matrix in all existing detectors. Concurrently, proposed detectors can be extended to the frequency selective channel considering their possible use in mmWave massive MIMO systems. The possible extended detectors may also exploit channel statistics such as temporal and spatial correlation to reduce complexity overhead.

References

[1] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, February 2016.

[2] "Cisco visual networking index: global mobile data traffic forecast update, 2015-2020," *White Paper*, pp. 1–39, February 2016.

[3] "Cisco visual networking index: global mobile data traffic forecast update, 2016-2021," *White Paper*, pp. 1–35, February 2017.

[4] M. A. M. Albreem, "5G wireless communication systems: Vision and challenges," in *Proc. Int. Conf. on Computers, Commun., and Contr. Technol.*, April 2015, pp. 493–497.

[5] M. A. M. Albreem, A. El-Saleh, M. Isa, W. Salah, and M. Juso, "Green internet of things: an overivew," in *Proc. IEEE Int. Conf. on Smart Instrumentation, Measurement and Applications*, November 2017, pp. 1–6.

[6] P. Gandotra, R. K. Jha, and S. Jain, "Green communication in next generation cellular networks: A survey," *IEEE Access*, vol. 5, pp. 11 727–11 758, June 2017.

[7] A. Adjoudani, E. C. Beck, A. P. Burg, G. M. Djuknic, T. G. Gvoth, D. Haessig, S. Manji, M. A. Milbrodt, M. Rupp, D. Samardzija, A. B. Siegel, T. Sizer, C. Tran, S. Walker, S. A. Wilkus, and P. W. Wolniansky, "Prototype experience for MIMO BLAST over third-generation wireless system," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 440–451, April 2003.

[8] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, February 2014.

[9] M. Pappa, C. Ramesh, and M. N. Kumar, "Performance comparison of massive MIMO and conventional MIMO using channel parameters," in *Proc. Int. Conf. on Wireless Commun., Signal Process. and Networking*, March 2017, pp. 1808–1812.

[10] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[2]Stair matrix is a particular tridiagonal matrix where the off-diagonal elements in either the even or the odd row are zeros.

[11] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive MIMO linear precoding: A survey," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3920–3931, Dec 2018.

[12] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO—5G system," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, Secondquarter 2016.

[13] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, Secondquarter 2018.

[14] P. Zhang, J. Chen, X. Yang, N. Ma, and Z. Zhang, "Recent research on massive MIMO propagation channels: A survey," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 22–29, December 2018.

[15] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[16] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale mimos," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1941–1988, 2015.

[17] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Trans. Commun.*, Nov. 2018.

[18] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, "Toward massive machine type cellular communications," *IEEE Trans. Wireless Commun.*, vol. 24, no. 1, pp. 120–128, February 2017.

[19] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 168–175, July 2016.

[20] X. Luo, "Multiuser massive MIMO performance with calibration errors," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4521–4534, July 2016.

[21] D. A. Basnayaka and H. Haas, "Spatial modulation for massive MIMO," in *Proc. IEEE Int. Conf. Commun.*, June 2015, pp. 1945–1950.

[22] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan 2018.

[23] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[24] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293–1308, Feb 2016.

[25] J. Chen, H. Chen, H. Zhang, and F. Zhao, "Spectral-energy efficiency tradeoff in relay-aided massive MIMO cellular networks with pilot contamination," *IEEE Access*, vol. 4, pp. 5234–5242, July 2016.

[26] Q. He, L. Xiao, X. Zhong, and S. Zhou, "Increasing the sum-throughput of cells with a sectorization method for massive MIMO," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1827–1830, Oct 2014.

[27] T. T. Do, E. Björnson, E. G. Larsson, and S. M. Razavizadeh, "Jamming-resistant receivers for the massive MIMO uplink," *IEEE Trans. Inf. Forensics Security*, vol. PP, no. 99, pp. 210–223, March 2017.

[28] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 21–27, June 2015.

[29] H. Alshamary, "Coherent and non-coherent data detection algorithms in massive MIMO," Master's thesis, University of Iowa, 2017.

[30] S. Jin, X. Wang, Z. Li, K. K. Wong, Y. Huang, and X. Tang, "On massive MIMO zero-forcing transceiver using time-shifted pilots," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 59–74, Jan 2016.

[31] T. L. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 847–860, Oct 2014.

[32] T. L. Marzetta, "How much training is required for multiuserMIMO?" in *Proc. Annual Asilomar Conf. Signals, Syst., Comp.*, Oct 2006, pp. 359–363.

[33] A. Zaib, M. Masood, A. Ali, W. Xu, and T. Y. Al-Naffouri, "Distributed channel estimation and pilot contamination analysis for massive MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4607–4621, Nov 2016.

[34] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, February 2013.

[35] Y. Liang, H. Li, F. Li, R. Song, and L. Yang, "Channel compensation for reciprocal TDD massive MIMO-OFDM with IQ imbalance," *IEEE Wireless Commun. Lett.*, vol. PP, no. 99, pp. 778–781, August 2017.

[36] J. Cai, B. Rong, and S. Sun, "A low complexity hybrid precoding scheme for massive MIMO system," in *Int. Symp. on Commun. and Inf. Technol.*, Sep. 2016, pp. 638–641.

[37] Y. Ren, Y. Wang, and G. Xu, "Two-stage hybrid precoding for massive MIMO systems," in *Proc. Int. Conf. on Computers, Commun., and Contr. Technol.*, April 2015, pp. 294–297.

[38] M. Kazemi, H. Aghaeinia, and T. M. Duman, "Discrete-phase constant envelope precoding for massive MIMO systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2011–2021, May 2017.

[39] M. Gümüş and T. M. Duman, "Multi-envelope precoding for massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 720–723, Oct 2018.

[40] Y. Jeon, C. Song, S. R. Lee, S. Maeng, J. Jung, and I. Lee, "New beamforming designs for joint spatial division and multiplexing in large-scale MISO multi-user systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3029–3041, May 2017.

[41] J. Nam, A. Adhikary, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct 2014.

[42] M. A. Girnyk, M. Vehkapera, and L. K. Rasmussen, "Large-system analysis of correlated MIMO multiple access channels with arbitrary signaling in the presence of interference," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2060–2073, April 2014.

[43] T. Ketseoglou and E. Ayanoglu, "Downlink precoding for massive MIMO systems exploiting virtual channel model sparsity," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, 2018.

[44] A. A. Lu, X. Gao, Y. R. Zheng, and C. Xiao, "Low complexity polynomial expansion detector with deterministic equivalents of the moments of channel Gram matrix for massive MIMO uplink," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 586–600, Feb 2016.

[45] J. C. Chen, "A low complexity data detection algorithm for uplink multiuser massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1701–1714, Aug 2017.

[46] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE Global Telecommun. Conf.*, vol. 1, Nov 2002, pp. 424–428 vol.1.

[47] Y. G. Lim, C. B. Chae, and G. Caire, "Performance analysis of massive MIMO for cell-boundary users," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6827–6842, Dec 2015.

[48] D. J. Costello, "Fundamentals of wireless communication," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 919–920, Feb 2009.

[49] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 473–485, April 2008.

[50] M. Wu, B. Yin, A. Vosoughi, C. Studer, J. R. Cavallaro, and C. Dick, "Approximate matrix inversion for high-throughput data detection in the large-scale MIMO uplink," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2013, pp. 2155–2158.

[51] A. A. Esswie, M. El-Absi, O. A. Dobre, S. Ikki, and T. Kaiser, "A novel FDD massive MIMO system based on downlink spatial channel estimation without CSIT," in *Proc. IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.

[52] M. Wu, B. Yin, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, "Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 916–929, Oct 2014.

[53] H. Prabhu, O. Edfors, J. Rodrigues, L. Liu, and F. Rusek, "Hardware efficient approximative matrix inversion for linear pre-coding in massive MIMO," in *Proc. IEEE Int. Symp. on Circuits and Systems*, June 2014, pp. 1700–1703.

[54] H. Prabhu, J. Rodrigues, O. Edfors, and F. Rusek, "Approximative matrix inverse computations for very-large MIMO and applications to linear pre-coding systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, April 2013, pp. 2710–2715.

[55] I. Al-Nahhal, M. Alghoniemy, A. B. A. El-Rahman, and Z. Kawasaki, "Modified zero forcing decoder for ill-conditioned channels," in *2013 IFIP Wireless Days (WD)*, Nov 2013, pp. 1–3.

[56] I. Al-Nahhal, M. Alghoniemy, O. Muta, and A. B. A. El-Rahman, "Reduced complexity k-best sphere decoding algorithms for ill-conditioned mimo channels," in *IEEE Annual Cons. Commun. Netw. Conf.*, Jan 2016, pp. 183–187.

[57] B. Kang, J. H. Yoon, and J. Park, "Low complexity massive MIMO detection architecture based on Neumann method," in *Proc. Int. SoC Design Conf.*, Nov 2015, pp. 293–294.

[58] B. Kang, J. Yoon, and J. Park, "Low-complexity massive MIMO detectors based on Richardson method," in *ETRI J.*, vol. 39, no. 3, Nov 2017, pp. 326–335.

[59] Z. Zhang, J. Wu, X. Ma, Y. Dong, Y. Wang, S. Chen, and X. Dai, "Reviews of recent progress on low-complexity linear detection via iterative algorithms for massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, July 2016, pp. 1–6.

[60] M. Wu, B. Yin, K. Li, C. Dick, J. R. Cavallaro, and C. Studer, "Implicit vs. explicit approximate matrix inversion for wideband massive MU-MIMO data detection," *J. Signal Process., Springer*, Dec 2017. [Online]. Available: https://doi.org/10.1007/s11265-017-1313-z

[61] L. Fang, L. Xu, and D. D. Huang, "Low complexity iterative MMSE-PIC detection for medium-size massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 108–111, Feb 2016.

[62] F. Wang, C. Zhang, X. Liang, Z. Wu, S. Xu, and X. You, "Efficient iterative soft detection based on polynomial approximation for massive MIMO," in *Proc. Int. Conf. on Wireless Commun. Signl Process.*, Oct 2015, pp. 1–5.

[63] S. Ghacham, M. Benjillali, and Z. Guennoun, "Low-complexity detection for massive MIMO systems over correlated Rician fading," in *Proc. Int. Wireless Commun. and Mobile Comput. Conf.*, June 2017, pp. 1677–1682.

[64] B. Yin, M. Wu, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, "A 3.8Gb/s large-scale MIMO detector for 3GPP LTE-advanced," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, May 2014, pp. 3879–3883.

[65] M. Albreem and M. Juntti, "On approximate matrix inversion methods for massive MIMO detectors," in *Proc. IEEE Wireless Commun. and Networking Conf.*, April 2019, pp. 1–6.

[66] M. Ylinen, A. Burian, and J. Takala, "Direct versus iterative methods for fixed-point implementation of matrix inversion," in *Proc. IEEE Int. Symp. on Circuits and Systems*, vol. 3, May 2004, pp. III–225–8 Vol.3.

[67] C. Tang, C. Liu, L. Yuan, and Z. Xing, "High precision low complexity matrix inversion based on Newton iteration for data detection in the massive MIMO," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 490–493, March 2016.

[68] V. Gupta, A. K. Sah, and A. K. Chaturvedi, "Iterative matrix inversion based low complexity detection in large/massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 712–717.

[69] C. Tang, C. Li, L. Yuan, and Z. Xing, "Approximate iteration detection with iterative refinement in massive MIMO systems," *IET Commun.*, vol. 11, no. 7, pp. 1152–1157, 2017.

[70] L. Shao and Y. Zu, "Joint newton iteration and neumann series method of convegence-accelerating matrix inversion approximation in linear precoding for massive MIMO systems," *Mathematical Problems in Engineering, Hindawi*, vol. 2016, 2016.

[71] L. Dai, X. Gao, X. Su, S. Han, C. L. I, and Z. Wang, "Low-complexity soft-output signal detection based on Gauss Seidel method for uplink multiuser large-scale MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4839–4845, Oct 2015.

[72] Z. Wu, Y. Xue, X. You, and C. Zhang, "Hardware efficient detection for massive MIMO uplink with parallel Gauss-Seidel method," in *Proc. Int. Conf. on Digital Signal Process.*, Aug 2017, pp. 1–5.

[73] Z. Wu, C. Zhang, Y. Xue, S. Xu, and X. You, "Efficient architecture for soft-output massive MIMO detection with Gauss-Seidel method," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2016, pp. 1886–1889.

[74] X. Qin, Z. Yan, and G. He, "A near-optimal detection scheme based on joint steepest descent and Jacobi method for uplink massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 276–279, Feb 2016.

[75] W. Song, X. Chen, L. Wang, and X. Lu, "Joint conjugate gradient and Jacobi iteration based low complexity precoding for massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, July 2016, pp. 1–5.

[76] P. Zhang, L. Liu, G. Peng, and S. Wei, "Large-scale mimo detection design and FPGA implementations using SOR method," in *Proc. IEEE Int. Conf. Commun. Software and Net.*, June 2016, pp. 206–210.

[77] X. Gao, L. Dai, Y. Hu, Z. Wang, and Z. Wang, "Matrix inversion-less signal detection using SOR method for uplink large-scale MIMO systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2014, pp. 3291–3295.

[78] Q. Deng, L. Guo, C. Dong, J. Lin, D. Meng, and X. Chen, "High-throughput signal detection based on fast matrix inversion updates for uplink massive multiuser multiple-input multi-output systems," *IET Commun.*, vol. 11, no. 14, pp. 2228–2235, 2017.

[79] B. Y. Kong and I. C. Park, "Low-complexity symbol detection for massive MIMO uplink based on Jacobi method," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Sept 2016, pp. 1–5.

[80] Y. Lee, "Decision-aided Jacobi iteration for signal detection in massive MIMO systems," *IEE Electron. Lett.*, vol. 53, no. 23, pp. 1552–1554, 2017.

[81] X. Gao, L. Dai, Y. Ma, and Z. Wang, "Low-complexity near-optimal signal detection for uplink large-scale MIMO systems," *IEE Electron. Lett.*, vol. 50, no. 18, pp. 1326–1328, August 2014.

[82] A. Bjorck, *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, 1996. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/1.9781611971484

[83] H. Costa and V. Roda, "A scalable soft Richardson method for detection in a massive MIMO system," *Przeglad Elektrotechniczny*, vol. 92, no. 5, pp. 199–203, August 2016.

[84] J. Jin, Y. Xue, Y. L. Ueng, X. You, and C. Zhang, "A split pre-conditioned conjugate gradient method for massive MIMO detection," in *Proc. IEEE Workshop on Signal Proess. Syst.*, Oct 2017, pp. 1–6.

[85] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "Conjugate gradient-based soft-output detection and precoding in massive MIMO systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2014, pp. 3696–3701.

[86] J. Zhou, Y. Ye, and J. Hu, "Biased MMSE soft-output detection based on Jacobi method in massive mimo," in *Proc. IEEE Int. Conf. on Commun. Problem-solving*, Dec 2014, pp. 442–445.

[87] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "VLSI design of large-scale soft-output MIMO detection using conjugate gradients," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2015, pp. 1498–1501.

[88] K. Li, B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "Accelerating massive MIMO uplink detection on GPU for SDR systems," in *Proc. IEEE Dallas Circuits and Syst. Conf.*, Oct 2015, pp. 1–4.

[89] C. Xiao, X. Su, J. Zeng, L. Rong, X. Xu, and J. Wang, "Low-complexity soft-output detection for massive MIMO using SCBiCG and Lanczos methods," *China Commun.*, vol. 12, no. Supplement, pp. 9–17, December 2015.

[90] X. Jing, A. Li, and H. Liu, "A low-complexity Lanczos-algorithm-based detector with soft-output for multiuser massive MIMO systems," *Digital Signal Process., Elsevier*, vol. 69, no. Supplement, pp. 41–49, October 2017.

[91] H. Zhang, G. Peng, and L. Liu, "Low complexity signal detector based on Lanczos method for large-scale MIMO systems," in *Proc. Int. Conf. on Electron. Info. and Emergency Commun.*, June 2016, pp. 6–9.

[92] J. Chen and V. K. N. Lau, "Multi-stream iterative SVD for massive MIMO communication systems under time varying channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, May 2014, pp. 3152–3156.

[93] A. Abdaoui, M. Berbineau, and H. Snoussi, "GMRES interference canceler for doubly iterative MIMO system with a large number of antennas," in *Proc. IEEE Int. Symp. on Signal Process. and Info. Technol.*, Dec 2007, pp. 449–453.

[94] Y. Yang, Y. Xue, X. You, and C. Zhang, "An efficient conjugate residual detector for massive MIMO systems," in *Proc. IEEE Workshop on Signal Proess. Syst.*, Oct 2017, pp. 1–6.

[95] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer, "FPGA design of a coordinate descent data detector for large-scale MU-MIMO," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2016, pp. 1894–1897.

[96] M. Wu, C. Dick, J. Cavallaro, and C. Studer, "High-throughput data detection for massive MU-MIMO-OFDM using coordinate descent," *IEEE Trans. Circuits Syst. I*, vol. 63, no. 12, pp. 2357–2367, Dec 2016.

[97] S. Rahaman, S. Shahabuddin, M. B. Hossain, and S. Shahabuddin, "Complexity analysis of matrix decomposition algorithms for linear MIMO detection," in *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, May 2016, pp. 927–932.

[98] W. Fukuda, T. Abiko, T. Nishimura, T. Ohgane, Y. Ogawa, Y. Ohwatari, and Y. Kishiyama, "Low-complexity detection based on belief propagation in a massive MIMO system," in *Proc. IEEE Veh. Technol. Conf.*, June 2013, pp. 1–5.

[99] M. Chaudhary, N. K. Meena, and R. S. Kshetrimayum, "Local search based near optimal low complexity detection for large mimo system," in *Proc. Int. Conf. on adv. Networks and Telecommun. Syst.*, Nov 2016, pp. 1–5.

[100] A. Chockalingam, "Low-complexity algorithms for large-MIMO detection," in *Int. Symp. on Commun., Control and Signal Process.*, March 2010, pp. 1–6.

[101] Y. Sun, "A family of linear complexity likelihood ascent search multiuser detectors for CDMA communications," in *Proc. Annual Asilomar Conf. Signals, Syst., Comp.*, vol. 2, Oct 2000, pp. 1163–1167 vol.2.

[102] S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "High-rate space-time coded large MIMO systems: Low-complexity detection and performance," in *Proc. IEEE Global Telecommun. Conf.*, Nov 2008, pp. 1–5.

[103] B. Cerato and E. Viterbo, "Hardware implementation of a low-complexity detector for large MIMO," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2009, pp. 593–596.

[104] A. A. J. Pereira and R. Sampaio-Neto, "A random-list based LAS algorithm for near-optimal detection in large-scale uplink multiuser MIMO systems," in *Proc. ITG Workshop Smart Antennas*, March 2015, pp. 1–5.

[105] A. K. Sah and A. K. Chaturvedi, "Reduced neighborhood search algorithms for low complexity detection in MIMO systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2015, pp. 1–6.

[106] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[107] N. Srinidhi, T. Datta, A. Chockalingam, and B. S. Rajan, "Layered tabu search algorithm for large-MIMO detection and a lower bound on ML performance," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2010, pp. 1–5.

[108] A. Elghariani and M. Zoltowski, "Low complexity detection algorithms in large-scale MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1689–1702, March 2016.

[109] N. Srinidhi, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive tabu search," in *Proc. IEEE Int. Symp. Inform. Theory*, June 2009, pp. 1993–1997.

[110] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-restart reactive tabu search algorithm for detection in large-MIMO systems," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1107–1109, December 2010.

[111] ——, "Low-complexity near-optimal signal detection in underdetermined large-MIMO systems," in *Proc. Nat. Conf. on Commun.*, Feb 2012, pp. 1–5.

[112] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, "Turbo decoding as an instance of Pearl's (belief propagation) algorithm," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb 1998.

[113] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar 1999.

[114] K. Takeuchi, T. Tanaka, and T. Kawabata, "Performance improvement of iterative multiuser detection for large sparsely spread CDMA systems by spatial coupling," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1768–1794, April 2015.

[115] Y. Gao, H. Niu, and T. Kaiser, "Massive MIMO detection based on belief propagation in spatially correlated channels," in *Proc. ITG Workshop Smart Antennas*, Feb 2017, pp. 1–6.

[116] A. Mezghani and J. A. Nossek, "Belief propagation based MIMO detection operating on quantized channel output," in *Proc. IEEE Int. Symp. Inform. Theory*, June 2010, pp. 2113–2117.

[117] H. . Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan 2004.

[118] T. Abiko, W. Fukuda, T. Nishimura, T. Ohgane, Y. Ogawa, Y. Ohwatari, and Y. Kishiyama, "An EXIT chart analysis for belief-propagation based detection in a large-scale MIMO system," in *Proc. IEEE Veh. Technol. Conf.*, June 2013, pp. 1–5.

[119] D. Bickson and D. Dolev, "Linear detection via belief propagation," in *Proc. Annual Allerton Conf. Commun., Contr., Computing*, p. 07.

[120] A. M. Mussi and T. Abrão, "Message passing detection for large-scale MIMO systems: damping factor analysis," *IET Signal Process.*, vol. 11, no. 8, pp. 923–935, 2017.

[121] M. Suneel, P. Som, A. Chockalingam, and B. S. Rajan, "Belief propagation based decoding of large non-orthogonal STBCs," in *Proc. IEEE Int. Symp. Inform. Theory*, June 2009, pp. 2003–2007.

[122] P. Som, T. Datta, A. Chockalingam, and B. S. Rajan, "Improved large-MIMO detection based on damped belief propagation," in *Proc. IEEE Inform. Theory Workshop*, Jan 2010, pp. 1–5.

[123] Y. Zhang, L. Huang, J. Song, J. Li, and W. Liu, "A low-complexity detector for uplink massive MIMO systems based on Gaussian approximate belief propagation," in *Proc. Int. Conf. on Wireless Commun. Signl Process.*, Oct 2015, pp. 1–5.

[124] T. Takahashi, S. Ibi, and S. Sampei, "On normalization of matched filter belief in GaBP for large MIMO detection," in *Proc. IEEE Veh. Technol. Conf.*, Sept 2016, pp. 1–6.

[125] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Low-complexity detection in large-dimension MIMO-ISI channels using graphical models," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1497–1511, Dec 2011.

[126] T. L. Narasimhan and A. Chockalingam, "EXIT chart based design of irregular LDPC codes for large-MIMO systems," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 115–118, January 2013.

[127] J. Yang, C. Zhang, S. Xu, and X. You, "Efficient stochastic detector for large-scale MIMO," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, March 2016, pp. 6550–6554.

[128] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 902–915, Oct 2014.

[129] T. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large MIMO systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, April 2014, pp. 815–820.

[130] J. Zeng, J. Lin, and Z. Wang, "Low complexity message passing detection algorithm for large-scale MIMO systems," *IEEE Wireless Commun. Lett.*, vol. PP, no. 99, pp. 1–1, 2018.

[131] T. Narasimhan and A. Chockalingam, "Detection and decoding in large-scale MIMO systems: A non-binary belief propagation approach," in *Proc. IEEE Veh. Technol. Conf.*, May 2014, pp. 1–5.

[132] Y. Xiong, N. Wei, and Z. Zhang, "A low-complexity iterative GAMP-based detection for massive MIMO with low-resolution ADCs," in *Proc. IEEE Wireless Commun. and Networking Conf.*, March 2017, pp. 1–6.

[133] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "A hybrid RTS-BP algorithm for improved detection of large-MIMO M-QAM signals," in *Proc. Nat. Conf. on Commun.*, Jan 2011, pp. 1–5.

[134] C. Thrampoulidis, E. Abbasi, W. Xu, and B. Hassibi, "Ber analysis of the BOX relaxation for BPSK signal recovery," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, March 2016, pp. 3776–3780.

[135] C. Thrampoulidis, W. Xu, and B. Hassibi, "Symbol error rate performance of Box-relaxation decoders in massive MIMO," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3377–3392, July 2018.

[136] S. Shahabuddin, M. Juntti, and C. Studer, "ADMM-based infinity norm detection for large MU-MIMO: Algorithm and VLSI architecture," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 2017, pp. 1–4.

[137] X. Peng, W. Wu, J. Sun, and Y. Liu, "Sparsity-boosted detection for large mimo systems," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 191–194, Feb 2015.

[138] O. Tanchuk and B. Rao, "Exploiting sparsity during the detection of high-order QAM signals in large dimension MIMO systems," in *Proc. Annual Asilomar Conf. Signals, Syst., Comp.*, Nov 2014, pp. 101–105.

[139] M. Mandloi and V. Bhatia, "Layered Gibbs sampling algorithm for near-optimal detection in large-MIMO systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, March 2017, pp. 1–6.

[140] J. Minango and A. C. Flores, "Low-complexity MMSE detector based on refinement Jacobi method for massive mimo uplink," *Physical Communication*, vol. 26, pp. 128 – 133, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1874490717302550

[141] Y. Xue, L. Liu, G. Peng, H. Zhang, S. Yin, J. Wei, and S. Wei, "Hardware efficient signal detector based on Lanczos method for massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. Software and Net.*, May 2017, pp. 523–527.

[142] G. Wunder, H. Boche, T. Strohmer, and P. Jung, "Sparse signal processing concepts for efficient 5G system design," *IEEE Access*, vol. 3, pp. 195–208, 2015.

[143] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104–118, Jan 2015.

[144] S. Kwon, J. Wang, and B. Shim, "Multipath matching pursuit," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2986–3001, May 2014.

[145] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, July 2007.

[146] A. K. Sah and A. K. Chaturvedi, "An MMP-based approach for detection in large MIMO systems using sphere decoding," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 158–161, April 2017.

[147] J. W. Choi and B. Shim, "New approach for massive MIMO detection using sparse error recovery," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2014, pp. 3754–3759.

[148] N. Samuel, T. Diskin, and A. Wiesel, "Deep MIMO detection," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, July 2017, pp. 1–5.

[149] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," *CoRR*, vol. abs/1707.07980, 2017. [Online]. Available: http://arxiv.org/abs/1707.07980

[150] S. S. Mosleh, L. Liu, C. Sahin, Y. R. Zheng, and Y. Yi, "Brain-inspired wireless communications: Where reservoir computing meets MIMO-OFDM," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4694–4708, Oct 2018.

[151] X. Yan, F. Long, J. Wang, N. Fu, W. Ou, and B. Liu, "Signal detection of MIMO-OFDM system based on auto encoder and extreme learning machine," in *Int. Joint Conf. on Neur. Net.*, May 2017, pp. 1602–1606.

[152] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Veh. Technol. Mag.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.

[153] C. Wen, W. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, Oct 2018.

[154] K. Kim, J. Lee, and J. Choi, "Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 828–831, April 2018.

[155] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec 2017.

[156] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85 – 117, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0893608014002135

[157] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5663–5678, Nov 2018.

[158] X. Liu and Y. Li, "Deep MIMO detection based on belief propagation," in *Proc. IEEE Inform. Theory Workshop*, Nov 2018, pp. 1–5.

[159] S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity near-ML performance achieving algorithm for large MIMO detection," in *Proc. IEEE Int. Symp. Inform. Theory*, July 2008, pp. 2012–2016.

[160] S. K. Mohammed, A. Zaki, A. Chockalingam, and B. S. Rajan, "High-rate space-time coded large-MIMO systems: Low-complexity detection and channel estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 958–974, Dec 2009.

[161] M. Hansen, B. Hassibi, A. G. Dimakis, and W. Xu, "Near-optimal detection in MIMO systems using Gibbs sampling," in *Proc. IEEE Global Telecommun. Conf.*, Nov 2009, pp. 1–6.

[162] 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.

[163] N. Srinidhi, T. Datta, A. Chockalingam, and B. Rajan, "Layered tabu search algorithm for large-MIMO detection and a lower bound on ML performance," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2955–2963, November 2011.

[164] C. Knievel, M. Noemm, and P. A. Hoeher, "Low-complexity receiver for large-MIMO space-time coded systems," in *Proc. IEEE Veh. Technol. Conf.*, Sept 2011, pp. 1–5.

[165] J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4973–4982, Aug 2011.

[166] P. Svac, F. Meyer, E. Riegler, and F. Hlawatsch, "Low-complexity detection for large MIMO systems using partial ML detection and genetic programming," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, June 2012, pp. 585–589.

[167] T. Datta, N. A. Kumar, A. Chockalingam, and B. S. Rajan, "A novel Monte-Carlo-sampling-based receiver for large-scale uplink multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3019–3038, Sept 2013.

[168] 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, February 2013.

[169] R. C. de Lamare, "Adaptive and iterative multi-branch MMSE decision feedback detection algorithms for multi-antenna systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5294–5308, October 2013.

[170] K. A. Singhal, T. Datta, and A. Chockalingam, "Lattice reduction aided detection in large-MIMO systems," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, June 2013, pp. 594–598.

[171] P. Švač, F. Meyer, E. Riegler, and F. Hlawatsch, "Soft-heuristic detectors for large MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4573–4586, Sept 2013.

[172] X. Gao, L. Dai, C. Yuen, and Y. Zhang, "Low-complexity MMSE signal detection based on Richardson method for large-scale MIMO systems," in *Proc. IEEE Veh. Technol. Conf.*, Sept 2014, pp. 1–5.

[173] K. S. Ali, W. Abediseid, and M. S. Alouini, "Sequential decoders for large MIMO systems," in *Proc. Int. Symp. on Modeling and Opt. in Mobile, Ad Hoc, and Wireless Net.*, May 2014, pp. 709–716.

[174] S. Wang, Y. Li, and J. Wang, "Convex optimization based multiuser detection for uplink large-scale MIMO under low-resolution quantization," in *Proc. IEEE Int. Conf. Commun.*, June 2014, pp. 4789–4794.

[175] Z. Hajji, K. Amis, A. Aïssa-El-Bey, and F. Abdelkefi, "Low-complexity half-sparse decomposition-based detection for massive MIMO transmission," in *Proc. Int. Conf. on Commun. and Networking*, Nov 2015, pp. 1–6.

[176] Y. Fadlallah, A. Aïssa-El-Bey, K. Amis, and D. Pastor, "Low-complexity detector for very large and massive MIMO transmission," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, June 2015, pp. 251–255.

[177] J. Zhou, J. Hu, J. Chen, and S. He, "Biased MMSE soft-output detection based on conjugate gradient in massive MIMO," in *Proc. IEEE Int. Conf. on ASIC*, Nov 2015, pp. 1–4.

[178] R. Guo, X. Li, W. Fu, and Y. Hei, "Low-complexity signal detection based on relaxation iteration method in massive MIMO systems," *China Commun.*, vol. 12, no. Supplement, pp. 1–8, December 2015.

[179] T. Li, S. Patole, and M. Torlak, "Low complexity detection for massive MIMO under multipath fading with limited storage resources," in *Proc. IEEE Int. Conf. Commun. Workshop*, June 2015, pp. 1316–1321.

[180] Y. Cai, R. C. de Lamare, B. Champagne, B. Qin, and M. Zhao, "Adaptive reduced-rank receive processing based on minimum symbol-error-rate criterion for large-scale multiple-antenna systems," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4185–4201, Nov 2015.

[181] D. Zhu, B. Li, and P. Liang, "On the matrix inversion approximation based on Neumann series in massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, June 2015, pp. 1763–1769.

[182] A. Elghariani and M. Zoltowski, "A quadratic programming-based detector for large-scale MIMO systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, March 2015, pp. 387–392.

[183] Y. Fadlallah, A. Aïssa-El-Bey, K. Amis, D. Pastor, and R. Pyndiah, "New iterative detector of MIMO transmission using sparse decomposition," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3458–3464, Aug 2015.

[184] L. Bai, T. Li, J. Liu, Q. Yu, and J. Choi, "Large-scale MIMO detection using MCMC approach with blockwise sampling," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3697–3707, Sept 2016.

[185] Abhishek, A. K. Sah, and A. K. Chaturvedi, "Improved sparsity behaviour and error localization in detectors for large MIMO systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2016, pp. 1–6.

[186] G. Yammine and R. F. H. Fischer, "Soft-decision decoding in noncoherent massive MIMO systems," in *Proc. ITG Workshop Smart Antennas*, March 2016, pp. 1–7.

[187] D. Finchera, M. D. Migliore, M. Lucido, F. Schettino, and G. Panariello, "Online failure detection in large massive MIMO linear arrays," in *Proc. Int. Applied Computational Electrom. Soc. Symp.*, March 2017, pp. 1–2.

[188] H. Y. Lu and X. He, "Two-stage-ranking assisted symbol detection for massive MIMO systems," in *Proc. IEEE Int. Conf. on Sigal and Image Process.*, Aug 2016, pp. 519–523.

[189] K. Nikitopoulos, D. Chatzipanagiotis, C. Jayawardena, and R. Tafazolli, "Multisphere: Massively parallel tree search for large sphere decoders," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2016, pp. 1–6.

[190] R. Hayakawa and K. Hayashi, "Error recovery with relaxed MAP estimation for massive MIMO signal detection," in *Proc. IEEE Int. Symp. Inform. Theory and its Applications*, Oct 2016, pp. 478–482.

[191] Z. Hajji, K. Amis, and A. Aïssa-El-Bey, "Turbo detection based on sparse decomposition for massive MIMO transmission," in *Proc. Int. Symp. on Turbo Codes and Iterative Info. Process.*, Sept 2016, pp. 290–294.

[192] W. Shen, L. Dai, Y. Shi, Z. Gao, and Z. Wang, "Massive MIMO channel estimation based on block iterative support detection," in *Proc. IEEE Wireless Commun. and Networking Conf.*, April 2016, pp. 1–6.

[193] H. A. J. Alshamary and W. Xu, "Efficient optimal joint channel estimation and data detection for massive MIMO systems," in *Proc. IEEE Int. Symp. Inform. Theory*, July 2016, pp. 875–879.

[194] O. H. Toma and M. El-Hajjar, "Element-based lattice reduction aided K-best detector for large-scale MIMO systems," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, July 2016, pp. 1–5.

[195] N. Souto, M. Ribeiro, and P. Sebastião, "Semidefinite relaxations for MIMO transmissions with high-order QAM constellations," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 984–988, July 2016.

[196] L. Liu, C. Yuen, Y. L. Guan, Y. Li, and Y. Su, "Convergence analysis and assurance for Gaussian message passing iterative detector in massive MU-MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6487–6501, Sept 2016.

[197] M. Mandloi and V. Bhatia, "Low-complexity near-optimal iterative sequential detection for uplink massive MIMO systems," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 568–571, March 2017.

[198] S. Higuchi, C. J. Ahn, and K. y. Hashimoto, "A reduced complexity and latency for massive MIMO using parallel detection algorithm," in *Proc. IEEE Reg. 10 Conf.*, Nov 2016, pp. 1996–1999.

[199] S. Agarwal, A. K. Sah, and A. K. Chaturvedi, "Likelihood-based tree search for low complexity detection in large MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 450–453, Aug 2017.

[200] R. Mitra and V. Bhatia, "Kernel-based parallel multi-user detector for massive-MIMO," *J. Comput. & Elect. Eng., Elsevier*, vol. 2, 2017.

[201] J. C. Chi, Y. C. Yeh, I. W. Lai, and Y. H. Huang, "Sphere decoding for spatial permutation modulation MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, May 2017, pp. 1–8.

[202] J. Zhang, X. Yuan, and Y. J. A. Zhang, "Blind signal detection in massive MIMO: Exploiting the channel sparsity," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, 2017.

[203] L. Li, W. Meng, and C. Li, "Semidefinite further relaxation on likelihood ascent search detection algorithm for high-order modulation in massive MIMO system," *IET Commun.*, vol. 11, no. 6, pp. 801–808, 2017.

[204] Y. Sun, Z. Li, C. Zhang, R. Zhang, F. Yan, and L. Shen, "Low complexity signal detector based on SSOR iteration for large-scale MIMO systems," in *Proc. Int. Conf. on Wireless Commun. Signl Process.*, Oct 2017, pp. 1–6.

[205] Z. Wu, L. Ge, X. You, and C. Zhang, "Efficient near-MMSE detector for large-scale MIMO systems," in *Proc. IEEE Workshop on Signal Proess. Syst.*, Oct 2017, pp. 1–6.

[206] M. Mandloi, M. A. Hussain, and V. Bhatia, "Improved multiple feedback successive interference cancellation algorithms for near-optimal MIMO detection," *IET Commun.*, vol. 11, no. 1, pp. 150–159, 2017.

[207] M. Manish, H. Azahar, and B. Vimal, "Adaptive multiple stage K-best successive interference cancellation algorithm for MIMO detection," *J. Telecommun. Syst., Springer*, vol. 66, no. 1, pp. 1–16, Sep 2017. [Online]. Available: https://doi.org/10.1007/s11235-016-0270-3

[208] F. Jiang, C. Li, Z. Gong, and R. Su, "Stair matrix and its applications to massive MIMO uplink data detection," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, 2018.

[209] F. Jing, C. Li, Z. Gong, and R. Su, "Extrinsic information analysis of a new iterative method using the stair matrix for massive MIMO uplink signal detection," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2018.

[210] M. M. Lwin and M. F. M. Salleh, "Feasibility of lattice reduction aided detection techniques in massive MIMO," in *Proc. Int. Conf. on Signals and Syst.*, May 2018, pp. 131–135.

[211] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, vol. 2, 2000, pp. II737–II740 vol.2.

[212] S. Loyka and F. Gagnon, "Analytical framework for outage and BER analysis of the V-BLAST algorithm," in *Proc. Int. Zurich Seminar Broadband Commun.*, 2004, pp. 120–123.

[213] Y. Jiang, M. K. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: An in-depth study of the high SNR regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2008–2026, April 2011.

[214] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI/IEEE Convention Radio Science*, Sep 1998, pp. 295–300.

[215] D. Wubben, R. Bohnke, V. Kuhn, and K. D. Kammeyer, "Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice- reduction," in *Proc. IEEE Int. Conf. Commun.*, vol. 2, June 2004, pp. 798–802 Vol.2.

[216] S. Shahabuddin, J. Janhunen, Z. Khan, M. Juntti, and A. Ghazi, "A customized lattice reduction multiprocessor for MIMO detection," in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2015, pp. 2976–2979.

[217] D. Wubben, D. Seethaler, J. Jalden, and G. Matz, "Lattice reduction," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 70–91, May 2011.

[218] M. Taherzadeh, A. Mobasher, and A. K. Khandani, "LLL reduction achieves the receive diversity in MIMO decoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4801–4805, Dec 2007.

[219] W. Zhang, X. Ma, and A. Swami, "Designing low-complexity detectors based on Seysen's algorithm," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3301–3311, October 2010.

[220] Q. Zhou and X. Ma, "Designing low-complexity detectors for generalized SC-FDMA systems," in *Proc. Conf. Inform. nces Syst. (CISS)*, March 2011, pp. 1–6.

[221] C. Ling, "On the proximity factors of lattice reduction-aided decoding," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2795–2808, June 2011.

[222] B. Gestner, W. Zhang, X. Ma, and D. V. Anderson, "Lattice reduction for MIMO detection: From theoretical analysis to hardware realization," *IEEE Trans. Circuits Syst. I*, vol. 58, no. 4, pp. 813–826, April 2011.

[223] S. Lyu and C. Ling, "Sequential lattice reduction," in *Proc. Int. Conf. on Wireless Commun. Signl Process.*, Oct 2016, pp. 1–5.

[224] M. A. M. Albreem and M. F. M. Salleh, "Near-A_n-lattice sphere decoding technique assisted optimum detection for block data transmission systems," *IEICE Trans. Commun.*, vol. E96-B, no. 1, pp. 365–359, Jan 2013.

[225] M. A. Albreem and M. F. M. Salleh, "Radius selection for lattice sphere decoder-based block data transmission systems," *Wireless Net., Springer*, vol. 22, no. 2, pp. 655–662, June 2016.

[226] M. El-Khamy, H. Vikalo, B. Hassibi, and R. J. McEliece, "On the performance of sphere decoding of block codes," in *Proc. IEEE Int. Symp. Inform. Theory*, July 2006, pp. 1964–1968.

[227] M. A. M. Albreem and N. A. H. B. Ismail, "A review: detection techniques for LTE system," *J. Telecommun. Syst., Springer*, vol. 63, no. 2, pp. 153–168, Oct 2016. [Online]. Available: https://doi.org/10.1007/s11235-015-0112-8

[228] M. A. M. Albreem and M. F. M. Salleh, "Regularized lattice sphere decoding for block data transmission systems," *Wireless Pers. Commun., Kluwer*, vol. 82, no. 3, pp. 1833–1850, Jun 2015. [Online]. Available: https://doi.org/10.1007/s11277-015-2317-2

[229] C. A. Shen and A. M. Eltawil, "A radius adaptive K-best decoder with early termination: Algorithm and VLSI architecture," *IEEE Trans. Circuits Syst. I*, vol. 57, no. 9, pp. 2476–2486, Sept 2010.

[230] T. H. Kim and I. C. Park, "Small-area and low-energy K-best MIMO detector using relaxed tree expansion and early forwarding," in *Proc. IEEE Int. Symp. on Low-Power Electron. and Design*, Aug 2010, pp. 231–236.

[231] Y. H. Wu, Y. T. Liu, H. C. Chang, and Y. C. Liao, "Early-pruned K-best sphere decoding algorithm based on radius constraints," in *Proc. IEEE Int. Conf. Commun.*, May 2008, pp. 4496–4500.

[232] Q. Li and Z. Wang, "Early-pruning K-best sphere decoder for MIMO systems," in *Proc. IEEE Workshop on Signal Proess. Syst.*, Oct 2007, pp. 40–44.

[233] K. C. Lai, C. C. Huang, and J. J. Jia, "Variation of the fixed-complexity sphere decoder," *IEEE Commun. Lett.*, vol. 15, no. 9, pp. 1001–1003, September 2011.

[234] T. Cui, S. Han, and C. Tellambura, "Probability-distribution-based node pruning for sphere decoding," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1586–1596, May 2013.

[235] S. Shahabuddin, O. Silvén, and M. Juntti, "Programmable ASIPs for Multimode MIMO Transceiver," *Journal of Signal Processing Systems*, vol. 90, no. 10, pp. 1369–1381, Oct. 2018.

[236] S. Shahabuddin, J. Janhunen, E. Suikkanen, H. Steendam, and M. Juntti, "An adaptive detector implementation for MIMO-OFDM downlink," in *2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, June 2014, pp. 305–310.

[237] E. Suikkanen, J. Janhunen, S. Shahabuddin, and M. Juntti, "Study of adaptive detection for MIMO-OFDM systems," in *2013 International Symposium on System on Chip (SoC)*, Oct 2013, pp. 1–4.

[238] Q. Li and Z. Wang, "Reduced complexity K-best sphere decoder design for MIMO systems," *Circuits, Systems & Signal Processing*, vol. 27, no. 4, pp. 491–505, Aug 2008. [Online]. Available: https://doi.org/10.1007/s00034-008-9039-6

[239] K. Ghavami and M. Naraghi-Pour, "MIMO detection with imperfect channel state information using expectation propagation," *IEEE Trans. Veh. Commun.*, vol. 66, no. 9, pp. 8129–8138, Sep. 2017.

[240] C. Wen, S. Jin, K. Wong, J. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, March 2015.

[241] J. Céspedes, P. M. Olmos, M. Sánchez-Fernández, and F. Perez-Cruz, "Expectation propagation detection for high-order high-dimensional MIMO systems," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2840–2849, Aug 2014.

[242] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, March 2012, pp. 3249–3252.

[243] K. Ghavami and M. Naraghi-Pour, "Blind channel estimation and symbol detection for multi-cell massive MIMO systems by expectation propagation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 943–954, Feb 2018.

[244] B. S. Thian and A. Goldsmith, "Decoding for MIMO systems with imperfect channel state information," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2010, pp. 1–6.

[245] A. Abdallah, M. M. Mansour, A. Chehab, and L. M. A. Jalloul, "MMSE detection for 1-bit quantized massive mimo with imperfect channel estimation," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.*, June 2018, pp. 1–5.

[246] C. Tang, Y. Tao, Y. Chen, C. Liu, L. Yuan, and Z. Xing, "Approximate iteration detection and precoding in massive MIMO," *China Commun.*, vol. 15, no. 5, pp. 183–196, May 2018.

[247] T. Xie, Q. Han, H. Xu, Z. Qi, and W. Shen, "A low-complexity linear precoding scheme based on SOR method for massive MIMO systems," in *Proc. IEEE Veh. Technol. Conf.*, May 2015, pp. 1–5.

[248] L. Zhang and Y. Hu, "Low complexity wssor-based linear precoding for massive MIMO systems," in *Int. Conf. on Cloud Comp. and Big Data*, Nov 2016, pp. 122–126.

Mahmoud A. Albreem (SM'16) received the BEng. degree in electrical engineering from Islamic University of Gaza (IUG), Palestine, in 2008. He received the MSc. and PhD. degrees from the University Sains Malaysia (USM), Malaysia, in 2010 and 2013, respectively. In 2014-2016, Dr. Albreem appointed as a Senior Lecturer with the University Malaysia Perlis (UniMAP). Dr. Albreem is currently an Assistant Professor and head of electronics and communications engineering department at A'Sharqiyah University (ASU), Oman. He is also a visiting Assistant Professor to the Centre for Wireless Communications (CWC), University of Oulu, Finland. He was a recipient of the Nokia Foundation Centennial Grant, 2018. His research interest includes signal processing for communication systems and information theory.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2935810, IEEE Communications Surveys & Tutorials

Markku Juntti (S'93-M'98-SM'04) received his M.Sc. (EE) and Dr.Sc. (EE) degrees from University of Oulu, Oulu, Finland in 1993 and 1997, respectively.

Dr. Juntti was with University of Oulu in 1992–98. In academic year 1994–95, he was a Visiting Scholar at Rice University, Houston, Texas. In 1999–2000, he was a Senior Specialist with Nokia Networks. Dr. Juntti has been a professor of communications engineering since 2000 at University of Oulu, Centre for Wireless Communications (CWC), where he leads the Communications Signal Processing (CSP) Research Group. He also serves as Head of CWC – Radio Technologies (RT) Research Unit. His research interests include signal processing for wireless networks as well as communication and information theory. He is an author or co-author in almost 500 papers published in international journals and conference records as well as in books *Wideband CDMA for UMTS* in 2000–2010, *Handbook of Signal Processing Systems* in 2013 and 2018 and *5G Wireless Technologies* in 2017. Dr. Juntti is also an Adjunct Professor at Department of Electrical and Computer Engineering, Rice University, Houston, Texas, USA.

Dr. Juntti is an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS and was an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2002–2008. He was Secretary of IEEE Communication Society Finland Chapter in 1996–97 and the Chairman for years 2000–01. He has been Secretary of the Technical Program Committee (TPC) of the 2001 IEEE International Conference on Communications (ICC 2001), and the Co-Chair of the Technical Program Committee of 2004 Nordic Radio Symposium and 2006 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006), and the General Chair of 2011 IEEE Communication Theory Workshop (CTW 2011). He has served as Co-Chair of the Signal Processing for Communications Symposium of Globecom 2014 Signal Processing for Communications Symposium, IEEE GlobalSIP 2016 Symposium on Transceivers and Signal Processing for 5G Wireless and mm-Wave Systems, ACM NanoCom 2018, and ISWCS 2019.

Shahriar Shahabuddin Dr. Shahriar Shahabuddin received his MSc and PhD from Centre for Wireless Communications, University of Oulu, Finland in 2012 and 2019 respectively under the supervision of professor Markku Juntti. During Spring 2015, he worked at Computer Systems Laboratory of Cornell University, USA in Professor Christoph Studer's group. Shahriar received distinction in MSc and the best masters thesis award of the department of communications engineering, University of Oulu in 2012. In addition, he received several scholarships and grants such as Nokia Foundation Scholarship, University of Oulu Scholarship Foundation Grant, Taunu Tonningen Foundation Grant during his PhD. Shahriar's reserach interest includes VLSI signal processing, MIMO detection and precoding, 5G security and machine learning applications for wireless communications. He is currently working at Nokia, Finland as a SoC Specialist.