Low-Complexity Lattice Reduction Aided Schnorr Euchner Sphere Decoder Detection Schemes with MMSE and SIC Pre-processing for MIMO Wireless Communication Systems

Goodwell Kapfunde, Oluyomi Simpson and Yichuang Sun
School of Physics, Engineering and Computer Science
University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, United Kingdom

Abstract - **The LRAD-MMSE-SIC-SE-SD (Lattice Reduction Aided Detection - Minimum Mean Squared Error-Successive Interference Cancellation - Schnorr Euchner - Sphere Decoder) detection scheme that introduces a trade-off between performance and computational complexity is proposed for Multiple-Input Multiple-Output (MIMO) in this paper. The Lenstra-Lenstra-Lovász (LLL) algorithm is employed to orthogonalise the channel matrix by transforming the signal space of the received signal into an equivalent reduced signal space. A novel Lattice Reduction aided SE-SD probing for the Closest Lattice Point in the transformed reduced signal space is hereby proposed. Correspondingly, the computational complexity of the proposed LRAD-MMSE-SIC-SE-SD detection scheme is independent of the constellation size while it is polynomial with reference to the number of antennas, and signal-to-noise-ratio (SNR). Performance results of the detection scheme indicate that SD complexity is significantly reduced at only marginal performance penalty.**

Keywords—MIMO, MMSE, SIC, Detection, Sphere Decoder

I. INTRODUCTION

A substantial number of researches have established that the Multiple-Input Multiple–Output (MIMO) antenna technology will continue to be of enormous significance to wireless communication systems, playing a very important role in improving the data throughput in modern wireless communication systems. [1-4]. It has been demonstrated in [5] that MIMO techniques, which have been widely studied due to their advantages over single antenna systems, can functionally improve link reliability without sacrificing bandwidth efficiency and transmit power. However, the major drawback of MIMO is the increased complexity of the detector scheme due to non-orthogonality of MIMO channels [6-7].

Linear detectors are popularly known for their significantly reduced complexity. However, the reduced complexity is achieved at strong performance penalty [8-13]. Non-linear detectors including sequential detectors yield better performance compared to linear detection schemes at the expense of one or more, or a combination of the following problems: error propagation, performance degradation, increased computational complexity or increased processing delay [13]. To overcome these problems, a trade-off between performance, and one or more of these issues has to be made.

Although the Maximum Likelihood (ML) detector yields an optimal solution to the MIMO detection problem, it cannot be implemented in practice as its computational complexity increases exponentially with the number of transmit and receive antennas, and with the constellation size. The Sphere Decoder (SD), a near–optimal method, can be used to solve this detection problem as it searches for lattice points confined in a hyper-sphere around the received point. SD based algorithms have been shown to be more efficient in estimating an ML solution [14-20]. However, superior performance of the SD is achieved at the detriment of its variable Non-deterministic Polynomial time (NP) hard complexity [17], when the initial radius is too large. On the other hand, a small initial radius leads to decoding failure.

Interestingly, the Depth-First-Search (DFS) SD-based algorithm is proposed as an optimal solution in [21, 22]. However, the problem associated with the proposed Depth-First-Search Schnorr Euchner - Sphere Decoder (DFS SE-SD) is linked with the equation:

$$\left\lceil \frac{-\boldsymbol{R_{SD}}'_{N_L-1} + \tilde{\mathbf{y}}_{N_L-1|N_L}}{\boldsymbol{r}_{N_L-1,N_L-1}} \right\rceil \le \boldsymbol{x}_{N_{L-1}} \le \left\lfloor \frac{\boldsymbol{R_{SD}}'_{N_L-1} + \tilde{\mathbf{y}}_{N_L-1|N_L}}{\boldsymbol{r}_{N_L-1,N_L-1}} \right\rfloor. \quad (1)$$

where $\tilde{\mathbf{y}}_{N_L-1|N_L}$ is the received signal conditioned to the already estimated symbol $\boldsymbol{x}_{N_L}$, $\boldsymbol{R_{SD}}'$is the new radius conditioned to initial sphere radius $\boldsymbol{R_{SD}}$. N_L is the number of rows in $\boldsymbol{R}$ which is a $N_t \times N_t$ upper triangular matrix, where represents the number of transmit antenna. $\lceil . \rceil$ and $\lfloor . \rfloor$ denote the rounding up and rounding down (quantisation) operations respectively. If $r_{N_L-1,N_L-1} \ll 1$, the initial radius conditioned on the initial radius increases drastically, thus, the number of the hypotheses $x_{N_{l-1}}$ rises exponentially with the increase in search radius leading to NP-hard complexity of the SD. In [14-16], the *K*-Best algorithm, also known as the Breadth-First Search (BFS), was also proposed to provide a sub-optimal solution to the MIMO detection problem. Here, *K* denotes the number of stored hypotheses $x_{N_{L-1}}$, also referred to as nodes or lattice points at each layer during the tree search detection process. Although this technique yields a near-optimal solution to the MIMO detection problem; the key issue is the reduction of the size K in order to achieve reasonably lower complexity.

The main contribution of this paper is to design an SD detection scheme that introduces a trade-off between performance, error propagation and complexity. Firstly, a novel Lattice Reduction Aided detection - Minimum Mean Squared Error-Successive Interference Cancellation (LRAD-MMSE-SIC) detection scheme is proposed. Secondly, the proposed scheme is then extended to the SD to reduce the search domain of the DFS based SD algorithm. The resulting reduced structure will be referred to as the LRAD-MMSE-SIC-SE-SD. Simulation and performance results prove that its computational complexity is independent of the constellation size while it is polynomial with respect to the number of antennas and the signal-to-noise ratio (SNR).

The rest of the paper is organised as follows: Section II describes the system model. Section III depicts the proposed novel SD-based MIMO detection algorithm. Section IV provides simulation results while Section V and VI analyses the computational complexity and, concludes the effects of the results, respectively.

II. System Model

The detection scheme for a symmetric MIMO system with N_t transmit and N_r receive antennas ($N_r = N_t$) is designed in this section. Fig.1 depicts the envisaged/designed MIMO system transmission model. At the transmitter, data is de-multiplexed into N_t data sub-streams also referred to as layers. These sub-streams are encoded by the Low Density Parity Check (LDPC) encoder and then interleaved bitwise by the inter-leaver before being mapped by the modulator $\mathbf{M}$ onto a $(N_t \times 1)$- complex valued transmit signal vector $\mathbf{x}$ of $\mathbf{M}$-QAM symbols. The symbols are transmitted by the N_t transmit antennas simultaneously in parallel over the flat fading channel, where $\mathbf{M}$ is the constellations size. For this design model, 4-QAM and 64-QAM modulation schemes is used.

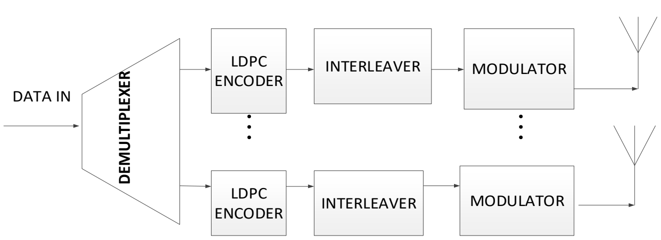


Fig. 1. MIMO system transmission model.

At the receiver, for simplicity, the received $N_r \times 1$ will be expressed by the following linear model

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} \quad (2)$$

where $\mathbf{n}$ denotes the AWGN noise of variance σ_n^2 observed at the receiver. The average transmit power of each antenna will be normalized to 1, i.e. $E\{\mathbf{xx}^H\} = \mathbf{I}_{N_t}$ and $E\{\mathbf{nn}^H\} = \sigma_n^2 \mathbf{I}_{N_r}$. $\mathbf{H}$ is the $N_r \times N_t$channel matrix whose elements are uncorrelated complex Gaussian fading gains with unit variance. In this design, a flat fading environment is assumed where the channel matrix $\mathbf{H}$ is constant over a time frame T and changes independently from frame to frame. It is also assumed that the channel matrix $\mathbf{H}$ is perfectly known at the receiver. The $N_r \times N_t$ complex valued system model in (1) can be decomposed into its $m \times n$ real-valued equivalent model which can be written as:

$$\begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} + \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix} \quad (3)$$

where $m = 2N_t$ and $n = 2N_r$ are the dimensions of the real valued channel matrix, $\begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix}$ and $\begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix}$ are $2N_r \times 1$ vectors, $\begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix}$ is a $2N_r \times 2N_t$ channel matrix and $\begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix}$ is a $2N_t \times 1$ vector. The corresponding dimensions of the received, noise and transmit vectors are given by $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{n} \in \mathbb{R}^n$ and $\mathbf{x} \in \chi^m$ respectively. χ denotes the finite set of real-valued transmit signals. The finite set χ, generated using an $\mathbf{M}$-QAM modulation scheme, is given by:

$$\chi = \left\{ \pm \frac{1}{2} x, \pm \frac{3}{2} x, \cdots \pm \frac{\sqrt{\mathbf{M}}-1}{2} x \right\} \quad (4)$$

where $\sqrt{\mathbf{M}}$ denotes the modulation index of the corresponding real-valued QAM modulator while the power normalization factor $x = (\sqrt{\frac{6}{\mathbf{M}-1}})$ is used for normalizing the power of the complex valued transmit signals. In this design, it is chosen in such a manner that the transmit power is normalized to 1. Each noiseless received signal is viewed as a point of infinite lattice spanned by $\mathbf{H}$ in the proposed design.

III. MIMO Detection Schemes

To recover the received signal, MIMO detection schemes are used to search for the received point located in the lattice generated by $\mathbf{H}$. The optimum ML detector performs an exhaustive search over the uncut set of transmit signals, $\mathbf{x} \in \chi^m$, and decides in favor of the transmit signal $\hat{\mathbf{x}}_{ML}$, that minimizes the Euclidian distance to the receive vector $\mathbf{y}$ and can be expressed as:

$$\hat{\mathbf{x}}_{ML} = \arg \min_{\mathbf{x} \in \chi} \|\boldsymbol{y} - \mathbf{Hx}\|^2 \quad (5)$$

However, the brute force ML detection is not feasible for larger number of transmit antennas or higher order modulation schemes as the computational effort is of order $\mathbf{M}^{N_t}$. A feasible alternative is the SD, which restricts the search space to a sphere of radius R_0. Nevertheless, the computational complexity is still high in comparison to simple, but suboptimal Successive Interference Cancellation (SIC) [21].

Notably, in this paper, a less complex SD detection scheme based on the hybrid LRAD-MMSE-SIC is proposed. It is well known that the MMSE yields an ML solution in perfectly orthogonal channels. Unfortunately, practical MIMO channels are non-orthogonal. In order to improve the performance of MMSE in practical non-orthogonal MIMO channels, the LRAD is proposed to transform the channel matrix $\mathbf{H}$ into a near-orthogonal channel matrix $\overline{\mathbf{H}}$.

However, $\overline{\mathbf{H}}$ is still not perfectly orthogonal. It has been shown in [13], [17] that the SIC detector is capable of achieving further improved performance compared to MMSE detector in non-orthogonal or near-orthogonal MIMO channels. The SIC detector achieves performance improvements by successively cancelling out the interference due to adjacent signal layers starting with the influence of the largest signal first, until the signal with the smallest power is detected. To yield improved performance, the ordered SIC detection is thus proposed in this design. First, an overview of the LRAD detection scheme is provided in the next sub-section, and then a detailed description of the proposed SD is explained.

A. Lattice Reduction Aided Detection Schemes

In this design, the columns $\boldsymbol{h}_l$ of the real-valued channel matrix $\mathbf{H}$ where $1 \leq l \leq 2N_t$ are regarded as the *basis* of a lattice spanned by the channel matrix $\mathbf{H}$ is adopted. It is also assumed that the possible transmit vectors are given the N_t-dimensional infinite integer space $\mathbb{Z}^{N_t}$. First, the estimate of the transmitted symbols are mapped to the appropriate QAM decision regions by performing scaling and shifting operation of the received signal in accordance to the LRAD principles as follows: $\tilde{\mathbf{x}} = \frac{\mathbf{x}}{d} + \mathbf{1}_N/2$, where d is the minimum distance between QAM constellation points and $\mathbf{1}_N$ denotes an all-ones $(N_t \times 1)$-dimensional vector. Next, the MIMO channel matrix $\mathbf{H}$ is transformed into an effective near-orthogonal channel matrix $\overline{\mathbf{H}}$ which yields an effective equivalent received signal model. This is achieved by using Lenstra-Lenstra-Lovasz (LLL) algorithm that decomposes $\mathbf{H}$ into $\mathbf{H} = \overline{\mathbf{H}}\mathbf{T}^{-1}$ where $\mathbf{T}$ is an $2N_t \times 2N_t$ *unimodular* matrix [18], i.e., $\mathbf{T}$ contains only integer entries and the determinant is det ($\mathbf{T}$) = 1 and $\mathbf{T}^{-1}$ is the inverse of the matrix $\mathbf{T}$. Mathematically, it is accepted that the inverse of a unimodular matrix always exist and contains only integer values, i.e.,$\mathbf{T}^{-1} \in \mathbb{Z}^m$ $(m = 2N_t)$. Thus, the effective transformed channel matrix $\overline{\mathbf{H}}$ which generates the *same* lattice as $\mathbf{H}$ is given by:

$$\overline{\mathbf{H}} = \mathbf{HT} \tag{6}$$

Finally, to further reduce the complexity of the proposed system, the channel matrix can be decomposed using $\mathbf{QR}$ decomposition [19], as $\mathbf{H} = \mathbf{QR}$, where the $N_r \times N_t$ matrix $\mathbf{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_{N_t}]$, consists of orthogonal columns of unit length ($\mathbf{Q}^T\mathbf{R} = \mathbf{I}_{N_t}$). $\mathbf{R}$ is the upper triangular matrix which consists of elements $r_{i,j}$ where $1 \leq i,j \leq N_t$. Thus, each column vector $\mathbf{h}_k$ of the channel matrix $\mathbf{H}$ is given by $\mathbf{h}_k = \sum_{l=1}^{k} r_{l,k}\boldsymbol{q}_l$ where k $(1 \leq k \leq N_t)$ is a counter. Here, the vector $\boldsymbol{q}_k$ denotes the direction of $\mathbf{h}_k$ perpendicular to the space spanned by $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_k$ and $r_{k,k}$ describes the corresponding length of $\mathbf{h}_k$. Additionally, $r_{l,k} = \boldsymbol{q}_l^T\mathbf{h}_k$ is the length of the projection of $\mathbf{h}_k$ onto $\boldsymbol{q}_l$ .The premise behind the LRAD technique is to transform a given basis $\mathbf{H}$ into a much better conditioned new basis $\overline{\mathbf{H}}$ with vectors of shortest length or, equivalently, into a basis consisting of near-orthogonal basis vectors. Likewise, the transformed channel matrix $\overline{\mathbf{H}}$ can be decomposed to $\overline{\mathbf{H}} = \overline{\mathbf{Q}}\overline{\mathbf{R}}$ in order to perform ordered SIC detection, where $\overline{\mathrm{QR}}$ is the transformed QR decomposition. The design description for the proposed MIMO detection scheme is provided in the next section.

B. LRAD-MMSE-SIC-SE-SD System Description

A detailed description of the proposed LRAD-MMSE-SIC-SE-SD MIMO detector is provided in this section. Fig. 2 shows the block diagram of the proposed detection scheme. This consists of five main blocks namely LLL Algorithm, LRAD Pre-processing, MMSE-SIC, the SE-SD and the Decision circuit. Each of the blocks addresses one or more of the issues mentioned in Section I. The LRAD linear detection

First, the LLL and LRAD Pre-processing blocks addresses the problems associated with the ill-conditioned and non-orthogonality of the channel matrix $\mathbf{H}$. The LLL algorithm in the LRAD pre-processor generates $\mathbf{T}$ and $\mathbf{T}^{-1}$ which are used to transform the channel matrix $\mathbf{H}$ and the transmit vector $\mathbf{x}$ into the $\mathbf{z}$ domain as:

$$\overline{\mathbf{H}} = \mathbf{HT} \;\;,\;\; \mathbf{z} = \mathbf{T}^{-1}\mathbf{x} \tag{7}$$

The transformed receive signal vector can then be rewritten as:

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} = \mathbf{HTT}^{-1}\mathbf{x} + \mathbf{n} = \overline{\mathbf{H}}\boldsymbol{z} + \mathbf{n} \tag{8}$$

Note that $\mathbf{Hx}$ and $\overline{\mathbf{H}}\boldsymbol{z}$ describe the same point in a lattice. The only difference is that the LLL-reduced channel matrix $\overline{\mathbf{H}}$ is much better conditioned and near-orthogonal than the original channel matrix $\mathbf{H}$. The condition of $\overline{\mathbf{H}}$ determines the noise amplification, hence the solution based on $\overline{\mathbf{H}}$ outperforms that based on $\mathbf{H}$. This solution does not only lead to performance gain of the overall system, but also reduces the computational complexity of the overall proposed LRAD-MMSE-SIC-SE-SD detector.

Since $\mathbf{x}$ belongs to the set $\mathbb{Z}^p$, $\mathbf{z}$ also belong to $\mathbb{Z}^p$, where $p = N_t$. Therefore, $\mathbf{x}$ and $\boldsymbol{z}$ stem from the same set. The only difference here is that for $\mathbf{M}$-QAM the lattice is finite and the domain of $\mathbf{z}$ differs from $\boldsymbol{\chi}^m$. In other words, $\mathbf{z}$ now resides in a much reduced lattice. Further reduction in complexity is achieved by sorted $\overline{\mathbf{Q}}\,\overline{\mathbf{R}}$ decomposition of $\overline{\mathbf{H}}$. This leads to the generation of the estimate $\widetilde{\boldsymbol{y}}$ of the received signal $\boldsymbol{y}$ with an ordered upper triangular matrix $\overline{\mathbf{R}}$. The modified signal $\widetilde{\boldsymbol{y}}$ is then further processed by the SE-SD which utilises R_o computed by the MMSE-SIC processor. In the SIC block, signal detection starts with the most reliable signal, i.e., the signal with the largest amplitude.

Secondly, with the transformed received signal, the MMSE-SIC Pre-processing block generates a reliable initial radius R_{SE-SD} to be utilised in the SE-SD. The MMSE filter $\mathbf{G}_{MMSE} = (\overline{\mathbf{H}}^{-T}\overline{\mathbf{H}} + \sigma_n^2\mathbf{TT}^{-1})^{-1}\overline{\mathbf{H}}^T$ is applied to improve the accuracy of the estimate $\bar{\mathbf{z}}_{LR-MMSE}$, of $\boldsymbol{z}$. Applying the MMSE-filter $\mathbf{G}_{MMSE}$ to the lattice-reduced system yields $\bar{\mathbf{z}}_{LR-MMSE}$ as follows:

$$\bar{\mathbf{z}}_{LR-MMSE} = \mathbf{G}_{MMSE}\mathbf{y} = ((\overline{\mathbf{H}}^{-T}\overline{\mathbf{H}} + \sigma_n^2\mathbf{TT}^{-1})^{-1}\overline{\mathbf{H}}^T)\,\mathbf{y} \tag{9}$$

The output $\bar{\mathbf{z}}_{LR-MMSE-SIC}$ of the MMSE-SIC detector is equivalent to multiplying the original MMSE-SIC estimate $\bar{\mathbf{x}}_{MMSE-SIC}$ by$\mathbf{T}^{-1}$, the inverse of the unimodular matrix $\mathbf{T}$, that is:

$$\bar{\mathbf{z}}_{LR-MMSE-SIC} = \mathbf{T}^{-1}\bar{\mathbf{x}}_{MMSE-SIC} \tag{10}$$

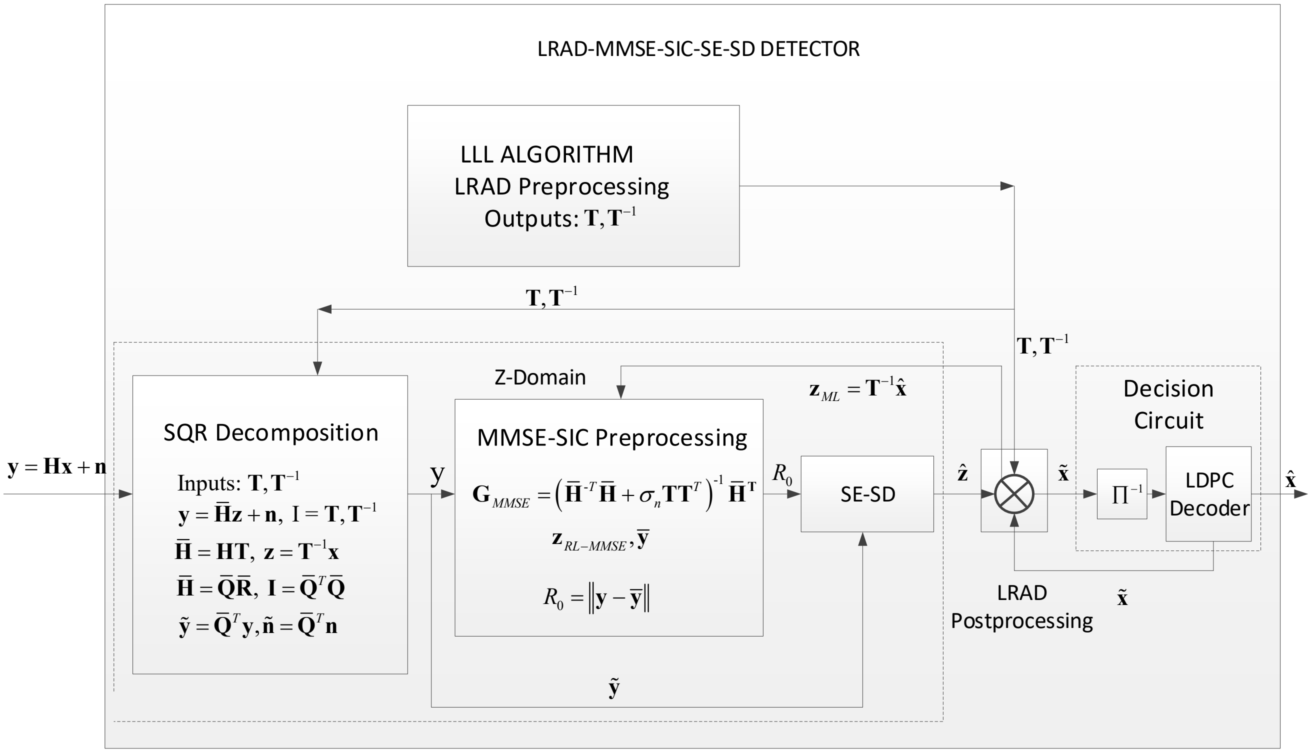


Fig. 2. Proposed LRAD-MMSE-SIC-SE-SD block diagram.

The estimate $\bar{\mathbf{z}}_{LR-MMSE-SIC}$ can be remapped to a lattice point $\bar{\mathbf{H}}\bar{\mathbf{z}}_{LR-MMSE-SIC}$ to yield:

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\bar{\mathbf{z}}_{LR-MMSE-SIC} \tag{11}$$

The initial SE-SD radius can thus be modelled by:

$$R^2{}_{SE-SD} = R_0^2 = \|\mathbf{y} - \bar{\mathbf{y}}\|^2 \tag{12}$$

The initial radius R_o is fed into the input of the SE-SD. This is then processed using the SE-SD algorithm to yield $\hat{\mathbf{z}}$. The SE-SD estimate can be recovered by multiplying $\hat{\mathbf{z}}$ by the unimodular matrix **T** as follows:

$$\widetilde{\mathbf{x}} = \mathbf{T}\hat{\mathbf{z}} \tag{13}$$

This is fed into the decision circuit which finally generates the estimate $\hat{\mathbf{x}}$ of the received signal $\mathbf{x}$. To prevent error propagation in the MMSE-SIC pre-processing unit, the LDPC error corrected estimate $\hat{\mathbf{x}}$ is feedback via the LRAD post-processing unit where it is transformed into the z-domain as follows:

$$\mathbf{z}_{ML} = \mathbf{T}^{-1}\hat{\mathbf{x}} \tag{14}$$

The effect of this feedback does not only ensure that error propagation is effectively arrested, but it also ensures that the overall performance of the system improves significantly.

IV. Simulation, Performance Results and Discussion

The simulations and the discussions of performance results for the proposed LRAD-MMSE-SIC-SE-SD detection strategy are presented in this section.

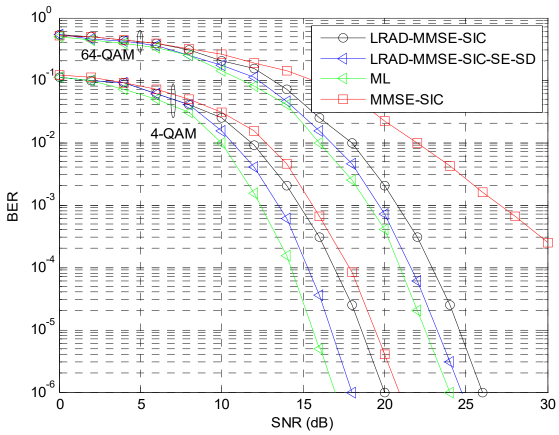
A. System Model for Simulation.

In this Section, two MIMO configurations were considered: 4x4 MIMO and 4x6 MIMO setups. Signals were transmitted in blocks of data over uncorrelated flat fading MIMO channels. A random information source was used to generate a stream of independent and identically distributed (i.i.d) information bits which were subsequently encoded using Low Density Parity Check codes of a code rate $R_C = {}^{K}/_{N} = 1/2$, where K is the number of information bits and N (=2K) is the sum of the information bits and parity check (redundant) bits. Note that the redundant bits constitute twice the information bits. The bit streams were then interleaved and divided into $N_{\mathbb{C}}$ blocks of N_t. The bit streams were further mapped into L bits, where L denotes the number of bits per modulated symbol, resulting in $M = 2^L$ different constellation points. 4-QAM and 64-QAM modulation schemes were applied on each sub stream as representatives for the respective low and high spectral efficiency regimes. The constellation points were finally mapped onto a vector transmit symbol $x \in \mathbb{C}^{[N_t \times 1]}$ whose components x_t were taken from some complex signal $\mathbb{C}$. It was further assumed that the transmitter and receiver were perfectly synchronized in time and frequency.

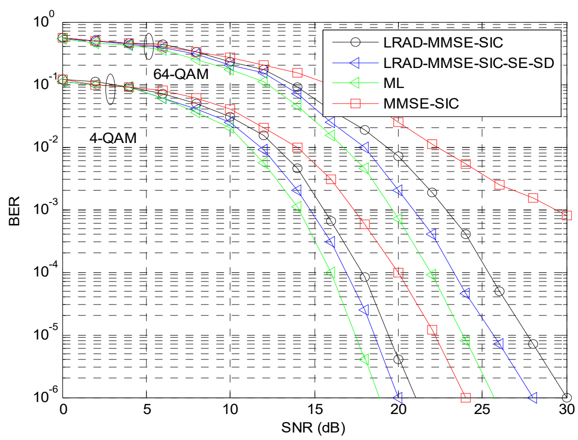
At the receiver, the proposed SD was used to detect the most likely vector x that was transmitted based on prior knowledge of $\boldsymbol{y}$, **H** and the statistics of noise **n** where $\boldsymbol{y} \in \mathbb{C}^{[N_r \times 1]}$ is the received signal, $\mathbf{H} \in \mathbb{C}\ (N_r \mathrm{X}\ N_t)$ is the channel matrix, and $\mathbf{n} \in \mathbb{C}^{[N_r \times 1]}$ is the receiver noise. The entries of $\boldsymbol{y}$, **H** and **n** are zero-mean, circularly symmetric complex Gaussian random variables, and the entries h{i, j} of **H** are normalized to have unit variance. The average energy per transmit symbol is denoted by E_s. The corresponding average energy per bit is denoted by E_b while the double-sided power spectral density of the complex noise (N_o/2 per real dimension) is denoted by N_o. The signal-to-noise ratio is thus given by SNR = E_s / N_o.

B. Performance Results and Discussion.

The Bit Error Rate (BER) performance results are presented for different modulation orders and different number of transmit and receive antenna configurations. The performance results for the proposed LRAD-MMSE- SIC-SE-SD are evaluated by comparing them with the ML,



(a) ½ rate LDPC Coded 4x4 MIMO System Setup.

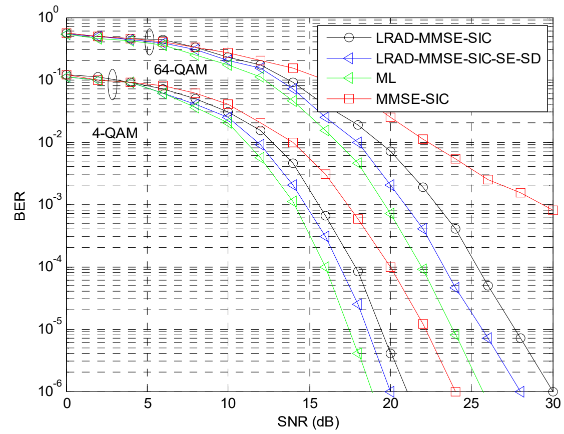


(b) Uncoded 4x4 MIMO System Setup.

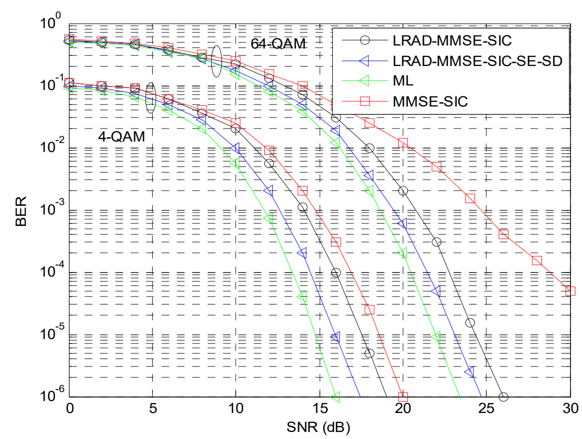
Fig. 3. Performance results for (a) coded and, (b) uncoded 4x4 MIMO System Setup.

LRAD-MMSE-SIC, and the MMSE-SIC detection schemes. The assumption made is that MIMO signals are transmitted through uncorrelated Rayleigh flat fading channels with the transmit power normalized to unit. Fig. 3 shows the BER results for ½ rate LDPC coded and uncoded 4x4 MIMO setups for the proposed LRAD-MMSE-SIC-SE-SD. 4-QAM and 64-QAM modulation schemes were applied on each sub stream as representatives for the respective low and high spectral efficiency regimes.

It can be clearly seen that the proposed LRAD-MMSE-SIC-SE-SD achieves substantial performance improvements in comparison with the LRAD-MMSE-SIC and MMSE-SIC detection schemes. The results in Fig. 3(a) show that the proposed LRAD-MMSE-SIC-SE-SD scheme achieves performance improvement of about 2dB and 3dB at a BER of 10^{-3} for the cases of both coded and uncoded 4-QAM transmissions compared to the performance of the LRAD-MMSE-SIC and MMSE-SIC detection schemes respectively. The performance improvement arises from the combination of the reduced search space introduced by the LRAD schemes and the optimal ordering due to the MMSE-SIC. However, there is a marginal overall performance loss of the proposed LRAD-MMSE-SIC-SE-SD compared to the ML although the performance is 1dB within that of the ML throughout the SNR regimes.



(a) Uncoded 4x4 MIMO System Setup.



(b) Uncoded 4x6 MIMO System Setup.

Fig. 4. Performance results for proposed LRAD-MMSE-SIC-SE-SD

The performance loss comes with a benefit of significantly reduced computational complexity. A similar trend in BER performance is also observed for the cases of both coded and uncoded 64-QAM although there is a modulation efficiency improvement penalty of about 6dB decrease in performance at a BER of 10^{-3} compared to the case of the 4-QAM. The proposed scheme is about 2dB and 8dB better than the LRAD-MMSE-SIC and MMSE-SIC schemes respectively at a BER of 10^{-3} for the uncoded case shown in Fig. 3(b).

The benefit of equipping both the transmitter and the receiver is demonstrated in Fig. 4. This is demonstrated by comparing the BER results for uncoded 4x4 MIMO setup and 4x6 MIMO setup where 4-QAM and 64-QAM modulation schemes are applied on each sub stream. Again, the proposed LRAD-MMSE-SIC-SE-SD achieves substantial performance improvements in comparison to the LRAD-MMSE-SIC and MMSE-SIC detection schemes.

As can be clearly seen in Fig. 4, the proposed LRAD-MMSE-SIC-SE-SD benefits substantially from reduced error probability on the first layer by equipping the receiver with more antennas (6 antennas in this case compared to 4 antennas in Fig. 3). It can also be clearly seen that low order modulation schemes perform better than higher order modulation schemes at the cost of bandwidth inefficiency. These results demonstrate that equipping the transmitter and the receiver with more antennas in conjunction with higher order modulation schemes can be attractive where high data rates are the main target of the wireless communication system.

V. Computational Complexity Analysis

The goal of the MIMO detector is to solve the Closest Lattice Point Search (CLPS) problem as efficiently as possible with minimum computational complexity. However, the ML performs an exhaustive search which is impractical for real-time systems. Fortunately, the emergence of the SD has restored the lost hope for high data rate MIMO systems. This section investigates the complexity aspects of the proposed LRAD-MMSE-SIC-SE-SD for MIMO systems. The complexity analysis focuses on uncoded MIMO transmission, i.e., hard output detection.

So far, most of the available results on the complexity for sphere detection have focused on the average behavior [20-23], and complexity exponents at moderately high SNR [21]. It is stated in [24], that the SD has a considerably higher worst case, but lower average complexity than other tree search based detection algorithms.

In this paper, the average complexity of the proposed LRAD-MMSE-SIC-SE-SD is investigated by estimating the number of arithmetic operations conducted to yield an estimate of the ML solution. This is then compared to the enumeration techniques in [25]. For a brute force search (ML detection), the maximum number of arithmetic operations A_o required to compute an exhaustive ML solution is $A_o^{ML} = \sum_{l=1}^{N_L} \chi^l$. To make a fair comparison, the number of each addition, multiplication, and extraction of a square-root will be counted as one operation as proposed in [25]. The upper limit for the original Fincke-Pohst (FP) is used as a figure of merit or yardstick against which the complexity of the proposed LRAD-MMSE-SIC-SE-SD is measured and is given by [24]:

$$\frac{1}{6}(2m^3 + 3m^2 - 5m) + \frac{1}{2}(m^2 + 12m - 7) \times \left(\left(2\left\lfloor\sqrt{d^2 t}\right\rfloor + 1\right)\binom{\lfloor 4d^2 t\rfloor + m - 1}{\lfloor 4d^2 t\rfloor} + 1\right) \quad (15)$$

where $m = N_t$, d is the sphere radius and $t = \max\left(r_{1,1}^2, \cdots, r_{m,m}^2\right)$ while that in [26] require cubic $O\ (m^3)$ computations.

For a 4x4 MIMO with a 64-QAM constellation, the upper bound number of arithmetic operations used in simulations is $A_o = 10000$. Fig. 5 shows the average value of the arithmetic operations A_o for the proposed LRAD-MMSE-SIC-SE-SD plotted against SNR. According to the results obtained to date in the literature, the expected complexity of FP-SD is only polynomial in the problem size for a wide range of SNRs [21, 26-28]. However, it was proven in [28, 29] that there exists a lower bound exponent on the complexity of the FP-SD.

While this implies that the complexity of the FP-SD will always grow exponentially with the problem size, the rate of exponential growth depends strongly on the SNR as is depicted in Fig. 5. Here, the number of required arithmetic computations decreases substantially as the SNR is increased and the detection complexity eventually approaches that of linear suboptimal detectors, i.e., the detection complexity eventually approaches a constant value. Conversely, the number of arithmetic computations increases substantially as the SNR is decreased and the detection complexity eventually approaches that of a brute force search. The FP-SD complexity is also extremely sensitive to the choice of the search radius.

The goal of a wireless communication system is delivering high data rates at minimum transmit power and at a much reduced detector complexity.

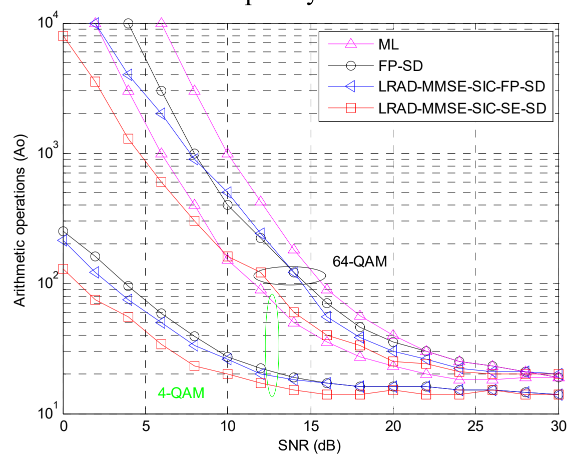


Fig. 5. Average arithmetic operations without statistical pruning.

As can be seen in Fig. 5, the proposed LRAD-MMSE-SIC-SE-SD can achieve the desired BER performance (see Fig. 3 & Fig. 4) at much reduced complexity within the whole range of the SNR regime (0-30dB) compared to the ML and original FP-SD for both cases of 4-QAM and 64-QAM transmissions.

The reduction in complexity of the proposed LRAD-MMSE-SIC-SE-SD is partly attributed to the MMSE-SIC pre-processing with layer ordering and partly due to the reduced search domain introduced by the LRAD scheme.

The complexity of the proposed LRAD-MMSE-SIC-SE-SD can be further reduced by applying statistical tree pruning, particularly in the low SNR regime. Fig. 6 illustrates that the proposed LRAD-MMSE-SIC-SE-SD algorithm solves the CLPS problem far more efficiently than the original FP-SD with the application of statistical tree pruning, particularly in the low to medium SNR regime, and for the case of higher order modulation. The complexity of the proposed LRAD-MMSE-SIC-SE-SD can be reduced by several orders of magnitude for 64-QAM transmission at SNRs below 10dB. The average complexity of the proposed LRAD-MMSE-SIC-SE-SD can be reduced by a factor of about 30 in the low SNR regime for the case of 64-QAM transmission as can be clearly seen in Fig. 6. However, the reduction in complexity is not noticeable for the case of 4-QAM transmission.

The average complexity of the original FP-SD becomes largely independent of the operating SNR by employing LRAD-MMSE-SIC based pre-processing. A reduction of 80-90% compared to a FP-SD without tree pruning can be achieved. However, the complexity reduction is achieved at the expense of performance loss over the whole BER range arising from the sub optimality due to statistical tree pruning. Overall, the complexity for both the LRAD-MMSE-SIC-SE-SD and the LRAD-MMSE-SIC-FP-SD is much lower than the ML for both cases of 4-QAM and 64-QAM transmissions. Based on these results in Fig. 5 and Fig. 6, it can be concluded that the proposed SE-SD with LRAD-MMSE-SIC pre-processing is the most attractive option for solving the CLPS problem, i.e., the best option for solving the ML detection problem.

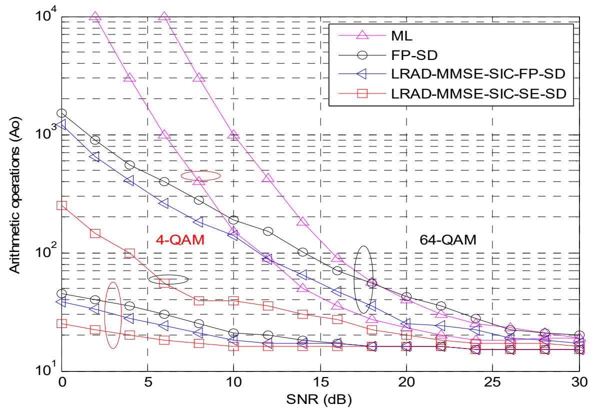


Fig. 6. Average arithmetic operations with statistical pruning.

VI. Conclusion

The LRAD-MMSE-SIC-SE-SD detection scheme which introduces a trade-off between performance and complexity was designed and presented in this paper. The reduction in complexity mainly results from the transformation of the channel matrix into a near-orthogonal channel and Sorted QR Decomposition (SQRD). As a result, low computational complexity is generated independent of constellation size and the number of antennas. Hence, this novel SD detection scheme is more efficient and applicable for MIMO in 5G mobile communication systems.

References

[1] K. Pratik, B. D. Rao and M. Welling, "RE-MIMO: Recurrent and Permutation Equivariant Neural MIMO Detection," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 459-473, 2021.

[2] J. Liao, J. Zhao, F. Gao and G. Y. Li, "A Model-Driven Deep Learning Method for Massive MIMO Detection," in *IEEE Communications Letters*, vol. 24, no. 8, pp. 1724-1728, Aug. 2020.

[3] M. Khani, M. Alizadeh, J. Hoydis and P. Fleming, "Adaptive Neural Signal Detection for Massive MIMO," in *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5635-5648, Aug. 2020.

[4] Y. Ding, N. Li, Y. Wang, S. Feng and H. Chen, "Widely Linear Sphere Decoder in MIMO Systems by Exploiting the Conjugate Symmetry of Linearly Modulated Signals," in *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6428-6442, 15 Dec.15, 2016.

[5] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal,* vol. 1, pp. 41-59, 1996.

[6] H. Wang, A. Kosasih, C. Wen, S. Jin and W. Hardjawana, "Expectation Propagation Detector for Extra-Large Scale Massive MIMO," in IEEE Transactions on Wireless Communications, vol. 19, no. 3, pp. 2036-2051, March 2020.

[7] M. M. Mansour, "Low-Complexity Soft-Output MIMO Detectors Based on Optimal Channel Puncturing," in IEEE Transactions on Wireless Communications, vol. 20, no. 4, pp. 2729-2745, April 2021.

[8] H. -Y. Lu, L. -P. Chang and H. -S. Hung, "Partial Tree Search Assisted Symbol Detection for Massive MIMO Systems," in IEEE Transactions on Vehicular Technology, vol. 69, no. 11, pp. 13319-13327, Nov. 2020.

[9] E. V. Goncharov, "Comparison of PIC and SIC with Lattice Reduction cancellation schemes for V-BLAST MIMO system," 2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2019, pp. 320-324.

[10] J. Park and J. Chun, "Improved lattice reduction-aided MIMO successive interference cancellation under imperfect channel estimation," *IEEE Transactions on Signal Processing,* vol. 60, pp. 3346-3351, 2012.

[11] B. Gestner, X. Ma, and D. V. Anderson, "Incremental lattice reduction: motivation, theory, and practical implementation," *IEEE Transactions on Wireless Communications,* vol. 11, pp. 188-198, 2012.

[12] S. Aubert, Y. Nasser, and F. Nouvel, "Lattice reduction-aided minimum mean square error k-best detection for mimo systems," in *Computing, Networking and Communications (ICNC), 2012 International Conference on*, pp. 1066-1070, 2012.

[13] D. Wübben and K. D. Kammeyer, "Low complexity successive interference cancellation for MIMO–OFDM systems," *European Transactions on Telecommunications,* vol. 18, pp. 457-466, 2007.

[14] S. Han, T. Cui, and C. Tellambura, "Improved K-best sphere detection for uncoded and coded MIMO systems," *IEEE Wireless Communications Letters,* vol. 1, pp. 472-475, 2012.

[15] R. El Chall, F. Nouvel, M. Hélard, and M. Liu, "Low complexity K-Best based iterative receiver for MIMO systems," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2014 6th International Congress on*, 2014, pp. 451-455.

[16] A. K. Sah and A. K. Chaturvedi, "Stopping Rule based Iterative Tree Search for Low-Complexity Detection in MIMO Systems," *IEEE Transactions on Wireless Communications,* 2016.

[17] L. Li, R. C. de Lamare, and A. G. Burr, "Complex sphere decoding with a modified tree pruning and successive interference cancellation," in *Wireless Communication Systems (ISWCS), 2012 International Symposium on*, pp. 226-230, 2012.

[18] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen,* vol. 261, pp. 515-534, 1982.

[19] A. Chauhan and R. Mehra, "Analysis of QR Decomposition for MIMO Systems," in *Electronic Systems, Signal Processing and Computing Technologies (ICESC), 2014 International Conference on*, pp. 69-73, 2014.

[20] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Transactions on Information Theory,* vol. 49, pp. 2389-2402, 2003.

[21] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *IEEE Transactions on Signal Processing,* vol. 53, pp. 2806-2818, 2005.

[22] C.-F. Liao and Y.-H. Huang, "Reduced-complexity LLL algorithm for lattice-reduction-aided MIMO detection," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*, pp. 1451-1455, 2009.

[23] A. Mejri and G. R.-B. Othman, "Reduced-Complexity Stack Decoder for MIMO Systems," in *Vehicular Technology Conference (VTC Spring)*, pp. 1-5, 2015.

[24] Y. Dai, S. Sun, and Z. Lei, "A comparative study of QRD-M detection and sphere decoding for MIMO-OFDM systems," in *Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on*, 2005, pp. 186-190.

[25] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of Computation,* vol. 44, pp. 463-471, 1985.

[26] H. Vikalo and B. Hassibi, "On the sphere-decoding algorithm II. Generalizations, second-order statistics, and applications to communications," *IEEE Transactions on Signal Processing,* vol. 53, pp. 2819-2834, 2005.

[27] O. Hayat, M. Qaisrani, and M. Akbar, "Sphere decoding in the presence of channel uncertainty," in *Emerging Technologies (ICET), 2010 6th International Conference on*, 2010, pp. 216-220.

[28] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Transactions on Signal Processing,* vol. 53, pp. 1474-1484, 2005.

[29] G. Papa, D. Ciuonzo, G. Romano, and P. S. Rossi, "A Dominance-Based Soft-Input Soft-Output MIMO Detector with Near-Optimal Performance," *IEEE Transactions on Communications,* vol. 62, pp. 4320-4335, Dec 2014.