

Sequential and Global Likelihood Ascent Search based Detection in Large MIMO Systems

Abhay Kumar Sah and A. K. Chaturvedi, *Senior Member, IEEE*

Abstract—Neighborhood search algorithms have been proposed for low complexity detection in large/massive multiple-input multiple-output (MIMO) systems. They iteratively search for the vector which minimizes the maximum likelihood (ML) cost in a fixed neighborhood. However, the ML solution may not lie in the searched space and also the search may go through a large number of intermediate vectors. Motivated by this, we first propose to cut down the size of the neighborhood so that the complexity of such algorithms can be reduced. Secondly, we also look for an update which is not restricted to be in a fixed neighborhood. This helps in improving the error performance. For the first purpose, we propose a metric and a few selection rules to decide whether or not to include a vector in the neighborhood. We use the indices of, say K , largest components of the metric for generating a reduced neighborhood set which is used to reduce the complexity of the existing algorithms while maintaining their error performance. Further, this reduced set facilitates the proposal of two new search algorithms. Simulation results show that the proposed algorithms have a much better error performance and also lower complexity compared to the existing algorithms.

Index Terms—Large MIMO, Neighborhood search algorithms, LAS algorithms.

I. INTRODUCTION

MULTIPLE input multiple output (MIMO) systems are an integral part of various wireless standards such as 4G, WiMax, 802.11n, HSPA+. They also seem indispensable for future wireless systems such as 5G etc [1]–[3]. Driven by the need for higher data rates, the number of transmit/receive antennas in future MIMO wireless systems may increase upto hundreds [4]–[6]. The realization of such large antenna systems requires several things, one of them being the availability of reliable and computationally efficient detectors. This is because the complexity of the maximum likelihood (ML) detector increases exponentially with the number of transmit antennas [7], [8]. The complexity of other near-optimal detectors such as [9]–[16] also increases rapidly with the number of transmit antennas while the performance of low complexity sub-optimal detectors based on MMSE or ZF criteria, or using VBLAST architecture, etc. [7] deteriorate with increasing number of antennas.

Various low-complexity detectors have been proposed in the literature for large/massive MIMO systems. These can be

broadly categorized as neighborhood search algorithms [17]–[23], lattice reduction (LR) aided algorithms [24]–[26], graph based algorithms [27]–[29], sparsity based algorithms [30]–[32], semi-definite relaxation (SDR) based approaches [33], [34], probabilistic data association (PDA) based algorithms [35], [36], and soft heuristics based algorithms [37], [38]. In this paper, we focus on neighborhood search algorithms.

The first neighborhood search algorithm reported is likelihood ascent search (LAS) algorithm [17]. This algorithm starts with some initial solution and searches for the best solution in the neighborhood (including the initial solution) until the improvement in ML cost saturates. The scheme suggested in [17] was designed for BPSK modulation only and later extended to complex constellations [18]. Two variants of LAS algorithm, namely, one symbol LAS (1-LAS) and multistage LAS (MLAS) were proposed in [18]. The 1-LAS algorithm looks for the best vector in a neighborhood which differs with the initial solution at one symbol only, leading to a poor error performance. While the MLAS improve its performance by also exploring the neighborhoods which differs with the initial solution at more than one symbols i.e. at two or three symbols, leading to significantly higher computations. Aiming to achieve better performance in terms of error performance as well complexity, [19] suggests some policies to avoid early terminations of the 1-LAS algorithm and named it as reactive tabu search (RTS) while [20] suggests initialization of the 1-LAS algorithm with multiple initial vectors. However, the improvement in performance comes at the price of extra computations as compared to 1-LAS and also the improvement is significant only for BPSK/QPSK schemes. To further improve the performance for higher order constellations, layered tabu search (LTS) was proposed in [22] which employs RTS algorithm multiple times requiring considerably higher computations.

All the above neighborhood search algorithms search every vector in the chosen neighborhood and therefore their complexity depends on the size of the neighborhood. However not every vector in the neighborhood causes reduction in ML cost, in fact, some of them may lead to a higher ML cost. Hence there is a need to carefully select the vectors which are likely to reduce the ML cost. Motivated by this idea, a metric for generating a one symbol update based reduced neighborhood is proposed in [39]. However, this reduced neighborhood approach can only reduce the complexity of the existing algorithms and not the error rate. This is because the search is restricted to be in a fixed neighborhood.

If the search is allowed to look for the best vector without any constraint on the neighborhood, the error rate can be further reduced. Using this idea, an unconstrained LAS (ULAS) algorithm was proposed in [40]. Subsequently, it has been extended to lattice reduction techniques in [41]. However,

Manuscript received December 27, 2016; revised May 16, 2017 and August 17, 2017; accepted September 29, 2017. This work was supported by Tata Consultancy Services Ltd. Project TCS/CS/20110191. The associate editor coordinating the review of this paper and approving it for publication was Taejoon Kim.

A. K. Sah is with the Department of Electrical Engineering, IIT Kanpur, Kanpur 208016, India (e-mail: abhaysah@iitk.ac.in).

A. K. Chaturvedi is with the Department of Electrical Engineering, IIT Kanpur, Kanpur 208016, India and also with the Department of Electronics and Communication Engineering, IIT Roorkee, Roorkee 247667, India (e-mail: akc@iitk.ac.in).

the ULAS algorithms in [40] and [41] can select only one neighboring vector per neighborhood leading to an inefficient search.

This paper is an attempt to arrive at an algorithm which can exploit the merits of both the above ideas i.e. an attempt to apply the reduced neighborhood approach (to reduce complexity) in a manner so that the search is not restricted to a fixed neighborhood (to reduce ML cost).

To achieve the above, we make the following two-fold contributions in this paper:

- *Reducing the complexity of existing neighborhood search algorithms:* We extend the idea of one symbol update based reduced neighborhood [39] to a multiple symbol update based reduced neighborhood. For this, we generalize the metric for one symbol update [39] to a multiple symbol update setting. This generalization requires usage of the properties of large channel matrices. Further, we identify the indices which need to be updated using a function of the largest (say K) components of the multiple symbol update based vector metric. We show that this approach is able to significantly reduce the complexity of the existing algorithms such as 1-LAS, MLAS, and RTS [19], and is also able to further reduce the complexity compared to the one symbol reduced neighborhood variants of 1-LAS, MLAS and RTS algorithms [39] without loss in their error performance.
- *Improving the error performance:* Next we improve the error performance by integrating the above approach with the idea of [40], [41] and propose two algorithms which are not restricted to search in a fixed neighborhood. The first one searches in a higher symbol neighborhood when there is no further improvement in the current neighborhood and continues this process till all the possible neighborhoods are exhausted. This helps in improving the error performance. The second one improves upon the first one by searching for the best vector in all the possible neighborhoods rather than searching sequentially in fixed neighborhoods of increasing sizes. This leads to a better exploration of the neighborhoods and also helps in faster convergence. It turns out the ULAS algorithm in [40], [41] is a special case of the second algorithm.

The rest of the paper is structured as follows: Section II describes the system model and the required initial background to develop the algorithm has been proposed in Section III. In III-A a novel metric is established first and then three selection rules are proposed in III-B to generate a reduced neighborhood. Using the reduced neighborhood, we propose two new search algorithms in Sections IV-A and IV-B respectively and their complexity has been analyzed in Section V. Section VI presents the simulations results and lastly, Section VII concludes the paper. In this paper, we use boldface capital letters ‘ \mathbf{A} ’ to denote matrices, boldface small letters ‘ \mathbf{a} ’ to denote vectors, small letters ‘ a ’ to denote elements of a vector/matrix, and calligraphic letters ‘ \mathcal{A} ’ to denote a set. The notations (\cdot) , $(\cdot)^T$, \mathbf{I}_N , and ‘ \setminus ’ stand for complex quantities, transpose, an $N \times N$ identity matrix, and the difference set operator, respectively. Similarly, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ represent the real and the imaginary part of a complex number, respectively.

II. PRELIMINARIES

A single user point to point MIMO system uses N_t number of transmit antennas for transmission and N_r number of receive antennas for reception. The input-output relationship can be mathematically modeled as

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\bar{\mathbf{x}} + \bar{\mathbf{n}}, \quad (1)$$

where $\bar{\mathbf{y}} = (\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{N_r})^T$ in which \bar{y}_i represents data received at the i th receive antenna and $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{N_t})^T$ is the transmitted signal vector where $\bar{x}_i \in \bar{\Omega}$ represents data transmitted through the i th transmit antenna. Here $\bar{\Omega}$ is a set of M complex symbols. The channel matrix $\bar{\mathbf{H}}$ is of dimension $(N_r \times N_t)$ with each coefficient $\bar{h}_{ij} \sim \mathcal{CN}(0, 1)$ and $\bar{\mathbf{n}} = (\bar{n}_1, \bar{n}_2, \dots, \bar{n}_{N_r})^T$ represents an $(N_r \times 1)$ i.i.d. additive white Gaussian noise (AWGN) vector with each $\bar{n}_i \sim \mathcal{CN}(0, \sigma^2)$.

It is worthwhile to mention that the system model in (1) is also valid for the uplink in multi-user massive MIMO systems [1]. The difference is that the vector $\bar{\mathbf{x}}$ will consist of the data of all users instead of a single user only.

For the MIMO systems employing real-valued or rectangular constellations, the complex system model in (1) can be decomposed into an equivalent real system model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2)$$

where $\mathbf{y} = (\Re\{\bar{\mathbf{y}}\}^T \Im\{\bar{\mathbf{y}}\}^T)^T$ is a $(2N_r \times 1)$ real equivalent received vector and $\mathbf{x} = (\Re\{\bar{\mathbf{x}}\}^T \Im\{\bar{\mathbf{x}}\}^T)^T$ represents a $(2N_t \times 1)$ real equivalent transmit vector, $\mathbf{n} = (\Re\{\bar{\mathbf{n}}\}^T \Im\{\bar{\mathbf{n}}\}^T)^T$ is a $(2N_r \times 1)$ equivalent noise vector and \mathbf{H} denotes the $(2N_r \times 2N_t)$ equivalent channel matrix given by

$$\mathbf{H} = \begin{bmatrix} \Re\{\bar{\mathbf{H}}\} & -\Im\{\bar{\mathbf{H}}\} \\ \Im\{\bar{\mathbf{H}}\} & \Re\{\bar{\mathbf{H}}\} \end{bmatrix}. \quad (3)$$

This real-equivalent decomposition (2) provides more flexibility in signal processing while delivering the same performance [42]. However, this equivalence does not always hold. For the detectors which are sensitive to channel ordering viz. VBLAST or for a tree search based detector, the real-valued system model can even outperform the complex-valued counterpart [43], [44], in terms of error performance, complexity, or both.

In this paper, for convenience we consider the real equivalent system model (2). Thus, we assume that the set of M complex symbols $\bar{\Omega}$ are taken from a square constellation, and therefore the elements of the real equivalent transmit vector belong to a set of \sqrt{M} real symbols Ω , where $\Omega = \{\pm 1, \pm 3 \dots \pm (\sqrt{M} - 1)\}$. However, one can apply the concepts which will be built in the subsequent sections to any complex constellation.

At the receiver our objective is to find the actual transmitted symbol vector \mathbf{x} among all the possible \sqrt{M}^{2N_t} transmit vectors which is nearest to the received signal vector \mathbf{y} for the given channel matrix \mathbf{H} . Mathematically, this is stated as

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \Omega^{2N_t}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2, \quad (4)$$

and is well known as ML detection. Here $\|\cdot\|$ denotes L_2 norm. We define the Euclidean cost function as

$$\phi(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 = \sum_{i=1}^{2N_r} \left| y_i - \sum_{j=1}^{2N_t} h_{ij} x_j \right|^2, \quad (5)$$

and we call it ML cost or simply cost in the rest of the paper. The computational complexity of ML detection is exponential in nature and cannot be applied practically even for small number of antenna pairs.

In the literature, for MIMO systems up to hundred antenna pairs, several neighborhood search algorithms such as [17]–[19] have been suggested. These algorithms proceed iteratively, initialize with an initial solution vector and then search for the best solution in the neighborhood of the initial solution vector. The L symbol neighborhood $\mathcal{N}_L(\mathbf{x})$ of a vector \mathbf{x} is defined as follows: Consider a vector $\tilde{\mathbf{x}}$ which differs from \mathbf{x} at exactly L symbols then $\tilde{\mathbf{x}} \in \mathcal{N}_L(\mathbf{x})$. The set which contains these L indices is defined as \mathcal{I}_k . It can be seen that for an L symbol neighborhood there will be a total of $\binom{2N_t}{L}$ \mathcal{I}_k 's. Mathematically, a neighboring vector $\tilde{\mathbf{x}} \in \mathcal{N}_L(\mathbf{x})$ can be expressed as

$$\tilde{x}_i = \begin{cases} \omega_j, & i \in \mathcal{I}_k \ \& \ \omega_j \in \{\Omega \setminus x_i\} \\ x_i, & i \notin \mathcal{I}_k, \end{cases} \quad (6)$$

where $i = 1, 2, \dots, 2N_t$, $j = 1, 2, \dots, \sqrt{M} - 1$ and $k = 1, 2, \dots, \binom{2N_t}{L}$. Thus, there are $(\sqrt{M} - 1)^L \binom{2N_t}{L}$ number of vectors in $\mathcal{N}_L(\mathbf{x})$. For example let us consider the set $\Omega = \{-3, -1, 1, 3\}$ and a vector $\mathbf{x} = [1 \ 3 \ -1]$ of length 3. The 1-symbol neighborhood of \mathbf{x} can be constructed by flipping the elements of this vector one by one with the other elements of the set Ω . Thus corresponding to the first element of this vector, there are three neighboring vectors: $[-3 \ 3 \ -1]$, $[-1 \ 3 \ -1]$ and $[3 \ 3 \ -1]$ in the 1-symbol neighborhood of $[1 \ 3 \ -1]$. Similarly for the other two locations we have six more vectors.

Based on the above neighborhood definition there exist different versions of neighborhood search algorithms such as 1-LAS algorithm [17], MLAS algorithm [18], multiple output selection LAS [20], RTS [19], random restart reactive tabu search [21], layered tabu search [22] etc. One can refer to these cited papers for more detail.

III. REDUCED SIZE NEIGHBORHOOD

In neighborhood search algorithms the aim is to find the minimum Euclidean cost vector (best vector) in the neighborhood. Even for a moderate choice of L , M , and N_t , the size of the neighborhood can be very large. This is because for a fixed value of L , there are $\binom{2N_t}{L}$ different sets of indices i.e. \mathcal{I}_k 's and corresponding to each \mathcal{I}_k there are $(\sqrt{M} - 1)^L$ possible neighbors. Since the neighborhood search algorithms are iterative in nature, neighborhood choice is to be made at each iteration but we should include only those neighbors in the neighborhood which are likely to reduce the ML cost. Therefore, we propose to break this problem into two parts. First is to identify a few sets of indices \mathcal{I}_k 's which are likely to cause reduction in the ML cost and the second is to find the best vector for every such \mathcal{I}_k . The first part has been addressed in [39] while the second part has been discussed in [18], but

only for $L = 1$. However, complexity of the first part alone is of the order of N_t^L because of $\binom{2N_t}{L}$ number of \mathcal{I}_k 's, which makes it infeasible even for moderate values of L . We address both these issues a fresh by deriving a novel multiple symbol update based metric in the next sub-section.

A. Multiple symbol update based metric to reduce the size of the neighborhood

Let \mathbf{e}^r denote the error vector corresponding to the r th iteration, defined as

$$\mathbf{e}^r = \mathbf{y} - \mathbf{H}\mathbf{x}^r, \quad (7)$$

where \mathbf{x}^r is the initial solution i.e. the solution obtained at the $(r - 1)$ th iteration. The L_2 norm of this vector \mathbf{e}^r will give the ML cost corresponding to the vector \mathbf{x}^r . Our objective is to minimize $\|\mathbf{e}^r\|^2$. The j th element of \mathbf{e}^r can be expressed as

$$e_j^r = y_j - \sum_{i=1}^{2N_t} h_{ji} x_i^r, \quad (8)$$

where $j = 1, 2, \dots, 2N_r$. For a fixed set of indices say $\mathcal{I} = \{i_1, i_2, \dots, i_L\}$ consider replacing the x_i^r in \mathbf{x}^r with a neighboring point in the constellation to generate $\tilde{\mathbf{x}}^r$. Mathematically this can be captured by

$$x_i^{r+1} = x_i^r + \eta_i d_{min} \quad \forall i \in \mathcal{I}, \quad (9)$$

where η_i is a non-zero integer and d_{min} is the minimum distance between the symbols in the real equivalent system. Consequent change in e_j^r can be expressed as

$$e_j^{r+1} = e_j^r - d_{min} \sum_{s=1}^L \eta_{i_s} h_{ji_s}. \quad (10)$$

where $i_s \in \mathcal{I}$ in η_{i_s} and h_{ji_s} corresponds to the j th row and i_s th column of \mathbf{H} . We propose to find the set of indices \mathcal{I} and the corresponding values of $\boldsymbol{\eta}$ which minimize $\sum_{j=1}^{2N_r} |e_j^{r+1}|^2$. This can be expressed as

$$\begin{aligned} (\boldsymbol{\eta}^\circ, \mathcal{I}^\circ) &= \underset{\boldsymbol{\eta}, \mathcal{I}}{\operatorname{argmin}} \sum_{j=1}^{2N_r} \left(e_j^r - d_{min} \sum_{s=1}^L \eta_{i_s} h_{ji_s} \right)^2 \\ &= \underset{\boldsymbol{\eta}, \mathcal{I}}{\operatorname{argmin}} \|\mathbf{e}^r - d_{min} \mathbf{H}_{\mathcal{I}} \boldsymbol{\eta}\|^2. \end{aligned} \quad (11)$$

where $\mathbf{H}_{\mathcal{I}}$ contains the L columns of the channel matrix \mathbf{H} , the indices of which are included in \mathcal{I} and $\boldsymbol{\eta} = [\eta_{i_1}, \eta_{i_2}, \dots, \eta_{i_L}]^T$. We denote the values of $\boldsymbol{\eta}$ and \mathcal{I} which minimize (11) as $\boldsymbol{\eta}^\circ$ and \mathcal{I}° respectively. However, finding an optimal solution is computationally expensive. Therefore we first solve it for a fixed \mathcal{I} using the pseudo inverse approach [45] and denote the solution as $\boldsymbol{\eta}^*$ as given below

$$\boldsymbol{\eta}^* = \left\lceil \frac{(\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}})^{-1} \mathbf{H}_{\mathcal{I}}^T \mathbf{e}^r}{d_{min}} \right\rceil = \frac{(\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}})^{-1} \mathbf{H}_{\mathcal{I}}^T \mathbf{e}^r}{d_{min}} + \boldsymbol{\delta}, \quad (12)$$

where $\boldsymbol{\delta} = [\delta_{i_1}, \delta_{i_2}, \dots, \delta_{i_L}]^T$ is the rounding error and $\lceil \cdot \rceil$ denotes the rounding operation required to ensure that the elements of $\boldsymbol{\eta}^*$ are integers. It may be noted that here we have relaxed the non-zero integer constraint to achieve a better

update. However, η^* is, in general, different from η^o . Using (12) in (11) we get the following optimization problem

$$\begin{aligned} & \min_{\mathcal{I}} \left\| \mathbf{e}^r - \mathbf{H}_{\mathcal{I}} \left((\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}})^{-1} \mathbf{H}_{\mathcal{I}}^T \mathbf{e}^r + d_{min} \boldsymbol{\delta} \right) \right\|^2 \\ & = \min_{\mathcal{I}} \left\{ -\mathbf{e}^{rT} \mathbf{H}_{\mathcal{I}} (\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}})^{-1} \mathbf{H}_{\mathcal{I}}^T \mathbf{e}^r + d_{min}^2 \|\mathbf{H}_{\mathcal{I}} \boldsymbol{\delta}\|^2 \right\}. \end{aligned} \quad (13)$$

Given the channel-hardening behavior [46] in large/massive MIMO systems, the off-diagonal elements of $\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}}$ can be assumed to be small compared to the diagonal elements. Hence, we have used a relation¹ from [47] to approximate the inverse of $\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}}$ as follows

$$\left[(\mathbf{H}_{\mathcal{I}}^T \mathbf{H}_{\mathcal{I}})^{-1} \right]_{ij} \approx \begin{cases} \frac{1}{\|\mathbf{h}_i\|^2}, & i = j \\ -\frac{\mathbf{h}_i^T \mathbf{h}_j}{\|\mathbf{h}_i\|^2 \|\mathbf{h}_j\|^2}, & i \neq j. \end{cases} \quad (14)$$

where $[\cdot]_{ij}$ represents the (i, j) th element. Using the above relation, (13) can be approximated as

$$\begin{aligned} & \min_{\mathcal{I}} \left[\sum_{s=1}^L \left\{ \underbrace{-\frac{(\mathbf{e}^{rT} \mathbf{h}_{i_s})^2}{\|\mathbf{h}_{i_s}\|^2}}_A + \underbrace{\delta_{i_s}^2 d_{min}^2 \|\mathbf{h}_{i_s}\|^2}_B \right\} \right. \\ & \left. + \sum_{\substack{s=1 \\ s \neq t}}^L \sum_{t=1}^L \left\{ \underbrace{\frac{(\mathbf{e}^{rT} \mathbf{h}_{i_s}) (\mathbf{e}^{rT} \mathbf{h}_{i_t})}{\|\mathbf{h}_{i_s}\|^2 \|\mathbf{h}_{i_t}\|^2}}_C + \underbrace{\delta_{i_s} \delta_{i_t} d_{min}^2}_D \right\} \underbrace{\mathbf{h}_{i_s}^T \mathbf{h}_{i_t}}_E \right]. \end{aligned} \quad (15)$$

The cross term $C \times E$ is the ratio of off-diagonal elements of $\mathbf{H}^H \mathbf{H}$ to the leading diagonal elements of $\mathbf{H}^H \mathbf{H}$. Given the context of large/massive MIMO systems, this ratio can be taken to be negligible compared to the term A. Similarly, term $D \times E$ can also be ignored compared to the term B. Hence (15) can be simplified to

$$\max_{\mathcal{I}} \sum_{s=1}^L \left\{ \frac{(\mathbf{e}^{rT} \mathbf{h}_{i_s})^2}{\|\mathbf{h}_{i_s}\|^2} - \delta_{i_s}^2 d_{min}^2 \|\mathbf{h}_{i_s}\|^2 \right\}. \quad (16)$$

In the above maximization problem, the second term is negligible compared to the first term (justified below) and hence can be ignored. We denote $\frac{\mathbf{e}^{rT} \mathbf{h}_k}{\|\mathbf{h}_k\|}$ by f_i for $i = 1, 2, \dots, 2N_t$ and use it to compute $\sum_{s=1}^L f_{i_s}^2$ for all possible $\binom{2N_t}{L}$ \mathcal{I}_k 's to generate a multiple symbol reduced neighborhood. If we compute this metric for $L = 1$, it turns out to be same as the metric for one symbol update as proposed in [39].

Justification for Ignoring the Second Term in (16): The maximization problem in (16) involves two terms in which the second term occurs due to the rounding operation in (12). Let us quantify the second term. For simplicity, we begin by assuming a special case of $L = 1$ i.e. the initial solution vector (say \mathbf{x}_2) differs from the transmitted vector (say \mathbf{x}_1) at one index (say u th) only. Then the u th index of \mathbf{x}_2 can be written in terms of the u th index of \mathbf{x}_1 as $x_{2u} = x_{1u} + md_{min}$ where

¹The inverse of a symmetric matrix \mathbf{A} , in which the off-diagonal elements are small compared to the diagonal elements can be approximated as $\mathbf{A}^{-1} \approx \mathbf{A}_d^{-1} - \mathbf{A}_d^{-1} \mathbf{A}_{od} \mathbf{A}_d^{-1}$ where, \mathbf{A}_d and \mathbf{A}_{od} denote the matrices formed by taking the diagonal and off-diagonal elements of \mathbf{A} such that $\mathbf{A} = \mathbf{A}_d + \mathbf{A}_{od}$.

m is an integer. Now by evaluating (12) we get the following expression

$$\begin{aligned} \eta_i^* & = \frac{\mathbf{h}_i^T (\mathbf{y} - \mathbf{H} \mathbf{x}_1 - md_{min} \mathbf{h}_u)}{d_{min} \|\mathbf{h}_i\|^2} + \delta_i \\ & = \frac{\mathbf{h}_i^T (\mathbf{n} - md_{min} \mathbf{h}_u)}{d_{min} \|\mathbf{h}_i\|^2} + \delta_i \\ & = -\frac{m \mathbf{h}_i^T \mathbf{h}_u}{\|\mathbf{h}_i\|^2} + \frac{\mathbf{h}_i^T \mathbf{n}}{d_{min} \|\mathbf{h}_i\|^2} + \delta_i. \end{aligned} \quad (17)$$

Since from the construction $x_{2u} = x_{1u} + md_{min}$, we know that the true value of $\eta_i^* = -m$ for $i = u$ and 0 for $i \neq u$, the value of δ_i will be $\frac{-\mathbf{h}_i^T \mathbf{n}}{d_{min} \|\mathbf{h}_i\|^2}$ for $i = u$. It may be noted that the rounding errors at the indices $i \neq u$ are not of interest because in the maximization problem (16) an index is considered only if the corresponding η_i^* is non-zero.

Now coming to the general case i.e. an initial vector that differs from the transmitted vector at L indices. We denote such indices by u_1, u_2, \dots, u_L . Following the same argument, the second term in (16) can be expressed as

$$\sum_{i=1}^L \delta_{u_i}^2 d_{min}^2 \|\mathbf{h}_{u_i}\|^2 = \sum_{i=1}^L \left(\frac{\mathbf{h}_{u_i}^T \mathbf{n}}{\|\mathbf{h}_{u_i}\|} \right)^2. \quad (18)$$

On the right hand side of this equation the numerator is a summation of N_r products of two independent zero mean Gaussian distributed random variables while the denominator is a squared sum of N_r independent zero mean unit variance Gaussian distributed random variables. For a large value of N_r , as is true for large MIMO systems, the overall term $\frac{\mathbf{h}_{u_i}^T \mathbf{n}}{\|\mathbf{h}_{u_i}\|}$ is likely to be very small with high probability. For example, taking $N_r = 50$ and SNR = 10 dB, this term will lie between -0.09 to 0.09 with probability 0.997. Its square will be even smaller. As a result, the second term in (16) (right hand side of (18)) can be assumed to be negligible compared to the first term, and can be ignored. This is specially true for small values of L , which is likely to be the case because the symbol error rate for the initial solution vector obtained through ZF or MMSE is of the order of 10^{-1} , for a wide range of SNR.

B. Selection rules

For generating an L symbol reduced neighborhood we propose to use $\sum_{s=1}^L f_{i_s}^2$ as a metric. Let us define it as

$$u_k = \sum_{s=1}^L f_{i_s}^2 \quad \forall \mathcal{I}_k, \quad (19)$$

where $k = 1, 2, \dots, \binom{2N_t}{L}$. Using these u_k 's we propose to use either of the following three selection strategies for the selection of the set of indices \mathcal{I}_k , which need to be updated to construct the reduced neighborhood.

1) *K most likely sets selection:* We choose the K sets of indices i.e. \mathcal{I}_k (where K is a significantly smaller number compared to $2N_t$) having the largest u_k values and generate a reduced set of neighbors. This means we are limiting the possible neighbors as well as the set \mathcal{I}_k from $\binom{2N_t}{L}$ to K . To avoid degradation in error performance, a suitable choice for the value of K has to be made. A very small value can lead to

a loss in error performance whereas a higher value will reduce the gain in complexity. We have observed that a selection of a maximum of 10% of all possible neighborhoods is sufficient for this purpose.

2) *Normalization based selection*: We propose to select the most likely set i.e. the \mathcal{I}_k for which u_k is maximum and the sets for which u_k is close to the maximum value. For convenience, let us normalize u_k 's and denote the normalized value \tilde{u}_k as

$$\tilde{u}_k = \frac{u_k}{u_{max}} \quad \forall k = 1, 2, \dots, \binom{2N_t}{L}, \quad (20)$$

where u_{max} is the maximum value taken by u_k 's. It may be noted that \tilde{u}_k 's will take values in the range 0 to 1. We select all the sets of \mathcal{I}_k 's for which \tilde{u}_k 's have a value close to 1 (say 0.8).

3) *Sum square normalization based selection*: Algorithms like RTS improve their error performance by allowing the search to proceed even if the neighboring vector has a higher ML cost. In terms of the normalization based selection above, this would mean lowering the threshold. However, lowering the threshold will not be a good strategy because we observed that \tilde{u}_k values are clustered at either high values or at low values and hence setting the threshold to a low value may allow a very large number of neighbors to be a part of search. Instead, a better way will be to select the \tilde{u}_k 's for which $\sum_k \tilde{u}_k$ is greater than some threshold. However finding this sum and then estimating a reasonable threshold, both need considerable effort. Hence, we propose a normalization such that $\sum_k \tilde{u}_k^2 = 1$ i.e.

$$\tilde{u}_k = \frac{u_k}{\sqrt{\sum_k |u_k|^2}} \quad \forall k = 1, 2, \dots, \binom{2N_t}{L}. \quad (21)$$

It may be noted that $\sum_k \tilde{u}_k^2$ will take values in the range 0 to 1 and by choosing an appropriate threshold we can get the right mix of complexity and performance.

It is worthwhile to mention here that since we need to select only a few sets of \mathcal{I}_k 's for which u_k values are large, it is not required to compute all u_k 's. One can determine the necessary u_k values obtained by sorting f_i 's in descending order of their magnitude and taking the first few combinations.

Using (6), we generate a reduced neighborhood (\mathcal{RN}) which differs with \mathbf{x} only at the indices contained in \mathcal{I}_k . Let us denote an element of the reduced neighborhood as $\tilde{\mathbf{x}}$. The i th element of $\tilde{\mathbf{x}}$ can be expressed as

$$\tilde{x}_i^r = \begin{cases} [x_i^r + \eta_i^* d_{min}], & i \in \mathcal{I}_k \\ x_i^r, & i \notin \mathcal{I}_k, \end{cases} \quad (22)$$

where $i = 1, 2, \dots, 2N_t$. Here the operator $[\cdot]$ is defined as

$$[A] = \begin{cases} \sqrt{M} - 1, & A > \sqrt{M} - 1 \\ -\sqrt{M} + 1, & A < -\sqrt{M} + 1 \\ A, & otherwise. \end{cases} \quad (23)$$

We will update the indices in the set \mathcal{I}_k only. This reduced neighborhood definition (22) can be used to reduce the complexity of an existing neighborhood algorithms by

Algorithm 1: Sequential LAS Algorithm

Input : $\mathbf{y}, \mathbf{H}, \Omega, max_{stage}$
Output: $\hat{\mathbf{x}}$

```

1 Initialization  $r = 0$  &  $L = 1$ ;
2  $\mathbf{x}^r \leftarrow$  output of MF/ZF/MMSE detector;
3  $Cost_{next} \leftarrow \|\mathbf{y} - \mathbf{H}\mathbf{x}^r\|^2$  &  $Cost_{pre} \leftarrow \infty$ ;
4 while  $L \leq max_{stage}$  do
    Determining the sets of indices to be updated
5  $f_i = \frac{(\mathbf{y} - \mathbf{H}\mathbf{x}^r)^T \mathbf{h}_i}{\|\mathbf{h}_i\|} \quad \forall i = 1, 2, \dots, 2N_t$ ;
6 Generate all possible combinations of  $L$  indices and save it
  as  $\mathcal{I}_k$  where  $k = 1, 2, \dots, \binom{2N_t}{L}$ ;
7  $u_k = \sum_{i \in \mathcal{I}_k} f_i^2 \quad \forall k = 1, 2, \dots, \binom{2N_t}{L}$ ;
8 Apply the chosen selection rule on  $u_k$ 's and select few
   $\mathcal{I}_k$ 's;
9  $Cost_{pre} \leftarrow Cost_{next}$ ;
10  $\mathbf{x}_{temp} \leftarrow \mathbf{x}^r$ ;
    Updating the selected sets of indices
11 for  $j \leftarrow 1$  to number of selected  $\mathcal{I}_k$ 's do
12     Select one set of indices  $\mathcal{I}_k$  and proceed;
13      $\eta_{\mathcal{I}_k}^* = \left\lceil \frac{(\mathbf{H}_{\mathcal{I}_k}^T \mathbf{H}_{\mathcal{I}_k})^{-1} \mathbf{H}_{\mathcal{I}_k}^T \mathbf{e}^r}{d_{min}} \right\rceil$ ;
14     for  $i \leftarrow 1$  to  $2N_t$  do
15         if  $i \in \mathcal{I}_k$  then
16              $\tilde{x}_i = [x_i^r + \eta_i^* d_{min}]$ ;
17         else
18              $\tilde{x}_i = x_i^r$ ;
19         end
20     end
    Updating the best available vector
21  $Cost_{temp} = \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}\|^2$ ;
22 if  $(Cost_{temp} < Cost_{next})$  then
23      $\mathbf{x}_{temp} \leftarrow \tilde{\mathbf{x}}$ ;
24      $Cost_{next} = Cost_{temp}$ ;
25 end
26 end
    Determining the neighborhood search size
27 if  $(Cost_{next} < Cost_{pre})$  then
28      $L = 1$ ;
29 else
30      $L = L + 1$ ;
31 end
32  $\mathbf{x}^{r+1} \leftarrow \mathbf{x}_{temp}$ ;
33  $r = r + 1$ ;
34 end
35 return  $\hat{\mathbf{x}} \leftarrow \mathbf{x}^{r+1}$ .

```

replacing the full neighborhood with the proposed reduced neighborhood.

IV. PROPOSED SEQUENTIAL AND GLOBAL LIKELIHOOD ASCENT SEARCH ALGORITHMS

The error performance of neighborhood search algorithms can be improved if the search is not restricted to a fixed symbol neighborhood. In this section, we use the reduced neighborhood definition given in (22) and propose two search algorithms which are not restricted to a fixed symbol neighborhood. In the first one, we search in a higher symbol neighborhood when the reduction in ML cost saturates in the current neighborhood while in the second one, we allow the algorithm to search in all possible neighborhoods. The goal is to improve the error performance while retaining the low complexity feature.

A. Sequential LAS algorithm

The algorithm begins with an initial vector which can be generated through any one of the linear detectors such as matched filter (MF), zero forcing (ZF) or minimum mean square error (MMSE). We denote this initial vector as $\mathbf{x}^{(0)}$. The expressions for $\mathbf{x}^{(0)}$ for these receivers are

$$\begin{aligned} \mathbf{x}_{MF}^{(0)} &= \lceil \mathbf{H}^T \mathbf{y} \rceil_{\Omega}, \\ \mathbf{x}_{ZF}^{(0)} &= \lceil (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} \rceil_{\Omega}, \\ \mathbf{x}_{MMSE}^{(0)} &= \lceil (\mathbf{H}^T \mathbf{H} + \sigma^2 \mathbf{I})^{-1} \mathbf{H}^T \mathbf{y} \rceil_{\Omega}, \end{aligned} \quad (24)$$

where $\lceil \cdot \rceil_{\Omega}$ represents element wise rounding operation to the set Ω . Corresponding to the initial vector, we compute the metric u_k (19) where $k = 1, 2, \dots, \binom{2N_t}{L}$ for one symbol neighborhood i.e. $L = 1$. Next we apply any one of the selection rules discussed in Section III-B to select the indices at which $\mathbf{x}^{(0)}$ will be updated using (12) and (22). If K (out of $2N_t$) indices are selected, there are only K updated vectors. These updated vectors will constitute the one symbol reduced neighborhood \mathcal{RN}_1 .

From all these updated vectors, we select the vector that minimizes the cost (5). If there is a reduction in the cost (5) compared to the initial solution vector, we replace the initial vector with the selected vector otherwise use the initial solution vector to go to the next step i.e. generate the reduced 2-symbol neighborhood \mathcal{RN}_2 followed by a reduced 3-symbol neighborhood \mathcal{RN}_3 and so on, up to some fixed \mathcal{RN}_L . We refer to this sequential neighborhood search algorithm as sequential LAS (SLAS) algorithm.

For generating an $L (> 1)$ -symbol reduced neighborhood, the process is identical except that instead of a single index, update is carried out at a set of L indices denoted as \mathcal{I}_k for $k = 1, 2, \dots, \binom{2N_t}{L}$ to arrive at an updated vector. Similar to one symbol reduced neighborhood only K (out of $\binom{2N_t}{L}$) such sets of \mathcal{I}_k 's will be selected for update. It may be noted that for $L = 1$ this reduces to what has been stated above for one-symbol neighborhood.

For a given \mathcal{I}_k , there are $(\sqrt{M} - 1)^L$ possible updates. To find a reasonable update among all these, we use the result in (12) as follows

$$\boldsymbol{\eta}_{\mathcal{I}_k}^* = \left\lceil \frac{(\mathbf{H}_{\mathcal{I}_k}^T \mathbf{H}_{\mathcal{I}_k})^{-1} \mathbf{H}_{\mathcal{I}_k}^T \mathbf{e}^r}{d_{min}} \right\rceil, \quad (25)$$

where $\mathbf{H}_{\mathcal{I}_k}$ denotes the matrix obtained by deleting the columns of \mathbf{H} whose indices are not part of \mathcal{I}_k .

If at any stage either \mathcal{RN}_2 , \mathcal{RN}_3 or any of the higher reduced neighborhoods finds a better solution than the solution obtained through \mathcal{RN}_1 , we repeat the above process considering that solution as the initial vector. The algorithm terminates when there is no reduction in the cost (4) at all the stages. The complete SLAS algorithm is provided in tabular form in Algorithm 1 in which max_{stage} denotes the maximum number of stages up to which the search proceeds.

It may be noted that SLAS is similar to MLAS but only in the sense that both search sequentially. Given its high complexity MLAS was conceived for only 2-symbol or, at most 3-symbol neighborhood searches [18] and is not feasible for higher neighborhood searches. Although SLAS operates

Algorithm 2: Global LAS Algorithm

Input : $\mathbf{y}, \mathbf{H}, \Omega, max_{stage}$

Output: $\hat{\mathbf{x}}$

```

1 Initialization  $r = 0$  &  $L = 1$ ;
2  $\mathbf{x}^r \leftarrow$  output of MF/ZF/MMSE detector;
3  $Cost_{next} \leftarrow \|\mathbf{y} - \mathbf{H}\mathbf{x}^r\|^2$  &  $Cost_{pre} \leftarrow \infty$ ;
4 while ( $Cost_{next} < Cost_{pre}$ ) do
   Computing the likelihood metric
5    $f_i = \frac{(\mathbf{y} - \mathbf{H}\mathbf{x}^r)^T \mathbf{h}_i}{\|\mathbf{h}_i\|} \forall i = 1, 2, \dots, 2N_t$ ;
6   for  $L \leftarrow 1$  to  $2N_t$  do
     Determining the sets of indices to be updated
7     Generate all possible combinations of  $L$  indices and
8     save it as  $\mathcal{I}_k$  where  $k = 1, 2, \dots, \binom{2N_t}{L}$ ;
9      $u_k = \sum_{i \in \mathcal{I}_k} f_i^2 \forall k = 1, 2, \dots, \binom{2N_t}{L}$ ;
10    Apply the chosen selection rule on  $u_k$ 's and select few
11     $\mathcal{I}_k$ 's;
12     $Cost_{pre} \leftarrow Cost_{next}$ ;
13     $\mathbf{x}_{temp} \leftarrow \mathbf{x}^r$ ;
14    Updating the selected sets of indices
15    for  $j \leftarrow 1$  to number of selected  $\mathcal{I}_k$ 's do
16      Select one set of indices  $\mathcal{I}_k$  and proceed;
17       $\boldsymbol{\eta}_{\mathcal{I}_k}^* = \left\lceil \frac{(\mathbf{H}_{\mathcal{I}_k}^T \mathbf{H}_{\mathcal{I}_k})^{-1} \mathbf{H}_{\mathcal{I}_k}^T \mathbf{e}^r}{d_{min}} \right\rceil$ ;
18      for  $i \leftarrow 1$  to  $2N_t$  do
19        if  $i \in \mathcal{I}_k$  then
20           $\tilde{x}_i = [x_i^r + \eta_i^* d_{min}]$ ;
21        else
22           $\tilde{x}_i = x_i^r$ ;
23        end
24      end
25      Updating the best available vector
26       $Cost_{temp} = \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}\|^2$ ;
27      if ( $Cost_{temp} < Cost_{next}$ ) then
28         $\mathbf{x}_{temp} \leftarrow \tilde{\mathbf{x}}$ ;
29         $Cost_{next} = Cost_{temp}$ ;
30      end
31    end
32  end
33   $\mathbf{x}^{r+1} \leftarrow \mathbf{x}_{temp}$ ;
34   $r = r + 1$ ;
35 end
36 return  $\hat{\mathbf{x}} \leftarrow \mathbf{x}^{r+1}$ .

```

on a reduced neighborhood, in a way it can be viewed as a generalization of the MLAS algorithm. It is worthwhile to point out that this generalization would not have been possible without the metric (19) for a multiple symbol update and the value of $\boldsymbol{\eta}_{\mathcal{I}_k}^*$ as provided in (25).

B. Global LAS algorithm

In the above algorithm, we search in a higher symbol neighborhood only if the search in the current neighborhood is unable to improve the existing solution. On the other hand, if in the current neighborhood there exists a solution with a lower cost, the algorithm is reset to search from one symbol neighborhood again, with this solution as an initial solution vector. Thus, at a given iteration, SLAS does not look for the best solution vector in all possible symbol neighborhoods. A better way will be to look for a global solution i.e. amongst all the possible symbol neighborhoods. However, this approach will explode the complexity. Hence, we strike a balance by

using reduced neighborhoods (22). To summarize, we begin by generating all possible L -symbol reduced neighborhoods (22) of a ZF/MMSE/MF based initial solution vector (24) i.e. \mathcal{RN}_L 's $\forall L = 1$ to $2N_t$. We determine the vector among these reduced neighborhoods which minimizes the cost (5). Considering this vector as an initial solution vector, we repeat the above search process. It is worthwhile to mention here that for generating an L -symbol reduced neighborhood, one can use any one of the selection rules proposed in Section III-B. The complete algorithm is provided in Algorithm 2. We refer to this neighborhood search algorithm as global LAS (GLAS) algorithm.

C. Discussion

It may be noted that for $max_{stage} = 1$, Algorithm 1 can be viewed as different reduced neighborhood versions of the 1-LAS algorithm [17], one for each selection rule. We refer to it as RN-1-LAS algorithm. If the K most likely sets selection rule is used, this RN-1-LAS algorithm is identical to the RN-1-LAS algorithm in [39]. Similarly, for the special case $max_{stage} = 3$, Algorithm 1 corresponds to reduced neighborhood versions of the MLAS algorithm [18], one for each selection rule. We refer to it as RN-MLAS algorithm. It may be pointed out that these RN-MLAS versions are better than the RN-MLAS in [39]. This will be discussed further in Section VI-A. Lastly, the unconstrained LAS (ULAS) algorithm proposed in [40] is a special case of Algorithm 2 using the K most likely sets selection rule with $K = 1$.

The reduced neighborhood definition in (22) can also be applied to other neighborhood search algorithms. For example, the RTS algorithm [19] is similar to the LAS algorithm except that it moves to the best vector in the neighborhood even if the best vector in the neighborhood is worse than the current solution vector. This strategy allows the algorithm to escape from early terminations. The process is continued for a fixed number of iterations and avoids cycles using a tabu matrix. Finally, the best solution vector across all the iterations is declared as the solution vector. In the reduced neighborhood variant, instead of the full neighborhood, we use the reduced neighborhood definition given in (22). Rest of the steps are same as in RTS. We call this variant as the RN-RTS algorithm.

V. COMPLEXITY ANALYSIS

Let us first discuss the complexity of the three reduced neighborhood variants. For this we begin with the RN-MLAS algorithm. The main steps in the RN-MLAS algorithm are: *i*) initialization, *ii*) computation of $\mathbf{e}^{r^T} \mathbf{h}_i$, $\|\mathbf{h}_i\|^2$ for $i = 1, 2, \dots, 2N_t$ and $\langle \mathbf{h}_i, \mathbf{h}_j \rangle$ for $i \neq j$, *iii*) selecting the indices, *iv*) generating the neighborhood and *v*) the search operation. The first two steps are also required by the MLAS algorithm. The RN-MLAS algorithm differs from step *iii*) onwards which is an additional step not required by MLAS. For RN-MLAS, step *iii*) has $O(\log N_t)$ complexity while for both the algorithms the overall complexity depends on the complexity of steps *iv*) and *v*) which, in turn, depends on the size of the neighborhood. This leads to $O(K)$ and $O(N_t^3)$ complexity per iteration for RN-MLAS and MLAS respectively. For the case of RN-1-LAS the complexity is still $O(K)$ in comparison

TABLE I
COMPLEXITY ORDER OF THE KEY STEPS OF SLAS AND GLAS ALGORITHMS

Key Steps	Operations	Complexity Order	Frequency of Computation
Initialization	ZF/MMSE	$O(N_r N_t^2)$	Once
Likelihood metric	$\mathbf{e}^{r^T} \mathbf{h}_i \forall i$	$O(N_r N_t)$	Once per iteration
	$\ \mathbf{h}_i\ ^2 \forall i$	$O(N_r N_t)$	
Selecting the indices	sorting of f_i 's	$O(\log N_t)$	
Generating the neighborhood	$u_k \forall k = 1, \dots, K$	$O(K)$	2N _t times per iteration
	$\mathbf{W}_{\mathcal{I}}^{-1} = (\mathbf{H}_{\mathcal{I}}^H \mathbf{H}_{\mathcal{I}})^{-1}$	$O(L^2)$	
	$\boldsymbol{\eta}_{\mathcal{I}}^*$	$O(L^2)$	
Determining the best update	Euclidean cost	$O(N_r N_t)$	
	decision	$O(1)$	

to $O(N_t)$ of 1-LAS. Thus, depending on the value of K , the overall complexity of reduced neighborhood variants of LAS algorithms can be significantly lower than that of LAS algorithms. Similar conclusion can be drawn for the RN-RTS algorithm also.

Since the reduced neighborhood variants save computational complexity, we have allowed search in higher symbol neighborhoods so that the error performance can be enhanced. For $max_{stage} = 2N_t$ the overall size of all $2N_t$ neighborhoods is of the order of $O(KN_t)$ which is still less than $O(N_t^3)$ of MLAS. However with the increasing value of L the computations required to find $\boldsymbol{\eta}^*$ may become significant because of the need to compute matrix inverse and one can argue that searching in a higher neighborhood will lead to increase in complexity. This is indeed true. However if we use the fact that for a selected set of indices \mathcal{I}' in an $L + 1$ symbol neighborhood there exist a subset of indices \mathcal{I} for which the inverse has already been computed at the L symbol neighborhood stage, then we can compute the inverse efficiently. Let us denote $\mathbf{W}_{\mathcal{I}'} \text{ as } \mathbf{H}_{\mathcal{I}'}^H \mathbf{H}_{\mathcal{I}'}$ then we can express $\mathbf{W}_{\mathcal{I}'}$ as

$$\mathbf{W}_{\mathcal{I}'} = \begin{bmatrix} \mathbf{W}_{\mathcal{I}} & \mathbf{W}_{\mathcal{I},(\mathcal{I}' \setminus \mathcal{I})} \\ \mathbf{W}_{(\mathcal{I}' \setminus \mathcal{I}),\mathcal{I}} & \mathbf{W}_{(\mathcal{I}' \setminus \mathcal{I}),(\mathcal{I}' \setminus \mathcal{I})} \end{bmatrix}, \quad (26)$$

where $\mathcal{I}' \setminus \mathcal{I}$ denotes the set difference between the sets \mathcal{I}' and \mathcal{I} . Since $\mathcal{I}' \supset \mathcal{I}$ and the size of \mathcal{I}' and \mathcal{I} differ by one, the size of $\mathcal{I}' \setminus \mathcal{I}$ is exactly one. Therefore, if $\mathbf{W}_{\mathcal{I}}^{-1}$ exists then $\mathbf{W}_{\mathcal{I}'}^{-1}$ can be expressed in terms of $\mathbf{W}_{\mathcal{I}}^{-1}$ as follows

$$\mathbf{W}_{\mathcal{I}'}^{-1} = \begin{bmatrix} \mathbf{W}_{\mathcal{I}}^{-1} + \frac{1}{\rho} \mathbf{W}_{\mathcal{I}}^{-1} \mathbf{b} \mathbf{b}^T \mathbf{W}_{\mathcal{I}}^{-1} & -\frac{1}{\rho} \mathbf{W}_{\mathcal{I}}^{-1} \mathbf{b} \\ -\frac{1}{\rho} \mathbf{b}^T \mathbf{W}_{\mathcal{I}}^{-1} & \frac{1}{\rho} \end{bmatrix}, \quad (27)$$

where $\rho = \mathbf{W}_{(\mathcal{I}' \setminus \mathcal{I}),(\mathcal{I}' \setminus \mathcal{I})} - \mathbf{b}^T \mathbf{W}_{\mathcal{I}}^{-1} \mathbf{b}$ and $\mathbf{b} = \mathbf{W}_{\mathcal{I},(\mathcal{I}' \setminus \mathcal{I})} = \mathbf{W}_{(\mathcal{I}' \setminus \mathcal{I}),\mathcal{I}}^T$. The computational complexity of finding this matrix inverse is $O(L^2)$. Hence the computations required for finding $\mathbf{W}_{\mathcal{I}'}^{-1} \forall L = 1, 2, \dots, 2N_t$ is $O(N_t^3)$. The complexity of all the key steps involved in SLAS and GLAS algorithms including their frequency of computation have been summarized in Table I. The overall per iteration complexity of both the proposed algorithms is $O(N_r N_t^2)$.

VI. SIMULATION RESULTS

There are two important aspects that need to be investigated: *i*) As discussed in Section V the per iteration complexity of reduced neighborhood variants such as RN-1-LAS, RN-MLAS,

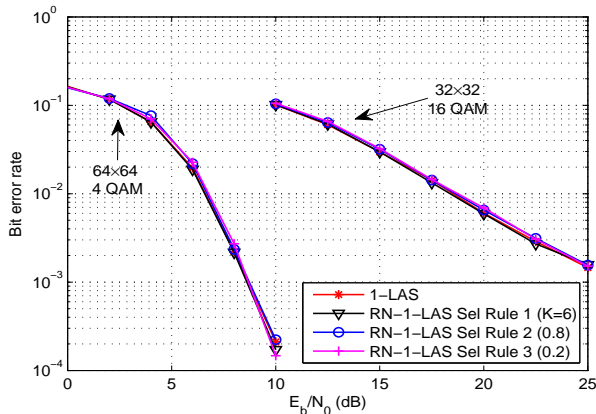


Fig. 1. Bit error rate performance of 1-LAS and its reduced neighborhood variants for a 64×64 system with 4-QAM and for a 32×32 system with 16-QAM.

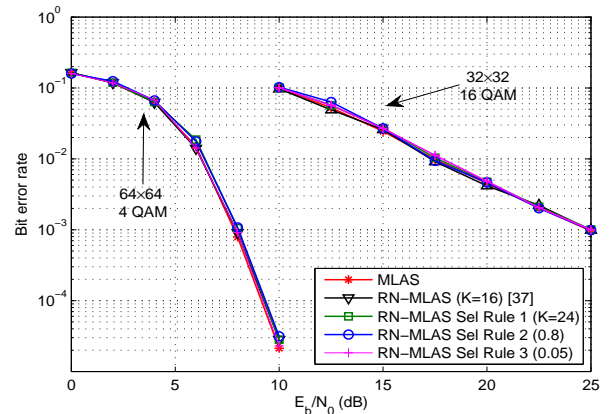


Fig. 3. Bit error rate performance of MLAS and its reduced neighborhood variants for a 64×64 system with 4-QAM and for a 32×32 system with 16-QAM.

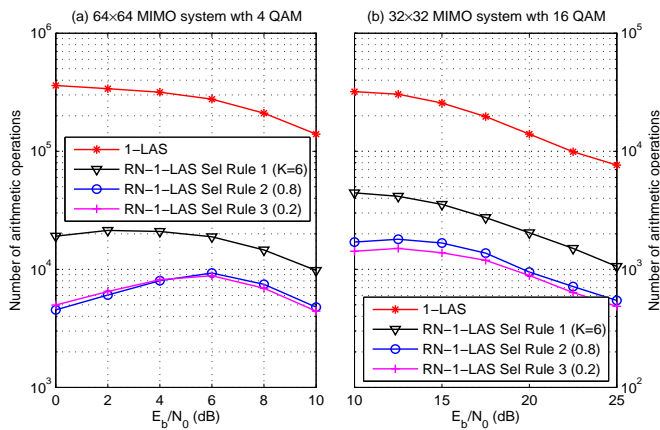


Fig. 2. Number of arithmetic operations required by 1-LAS and its reduced neighborhood variants for a 64×64 system with 4-QAM and for a 32×32 system with 16-QAM.

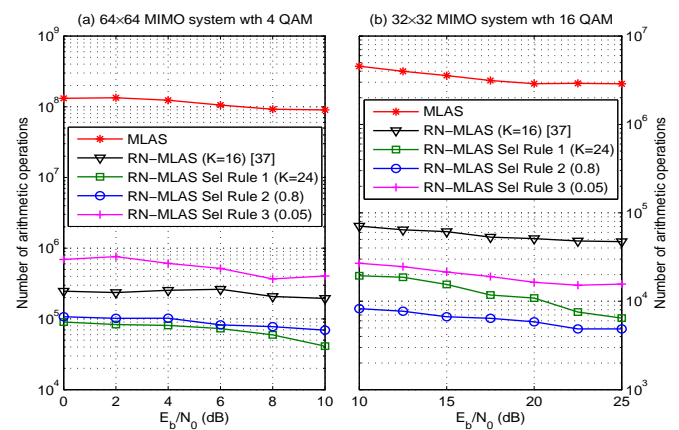


Fig. 4. Number of arithmetic operations required by MLAS and its reduced neighborhood variants for a 64×64 system with 4-QAM and for a 32×32 system with 16-QAM.

and RN-RTS can be significantly lower than their original counterparts. Will they be able to retain the bit error rate (BER) of LAS/RTS despite this reduction in complexity? *ii*) Whether the reduced neighborhood based approaches i.e. SLAS and GLAS can be utilized to improve the error performance, if complexity is not the focus? We will find answers to both these questions in this section.

A. Comparison of reduced neighborhood algorithms

We begin with examining the performance of the reduced neighborhood variants of 1-LAS, MLAS and RTS algorithms and compare them with their original versions taking MMSE solution as the initial vector. For this we consider a 64×64 MIMO system with 4-QAM modulation and a 32×32 MIMO system with 16-QAM modulation. The corresponding BER results are shown in Fig. 1, Fig. 3 and Fig. 5 while their complexity comparisons are shown in terms of average number of arithmetic operations in Fig. 2, Fig. 4 and Fig. 6 respectively. It may be noted that throughout this paper, we do not count the arithmetic operations involved in obtaining the initial MMSE solution. This is because all the algorithms under consideration use the same initial solution.

From Fig. 1 we can see that the error performance of RN-1-LAS matches the performance of 1-LAS for all the three

selection rules. In the first rule, the value of K was taken to be 6 which lead to a selection of only 6 neighbors from a total of 64 neighbors, while for the second and third selection rules the threshold was set to be 0.8 and 0.2 respectively. From Fig. 2 we can see that RN-1-LAS provides at least 90% reduction in complexity for all the selection rules. Among them, the third selection rule outperforms the first and second in terms of complexity. From Fig. 5 and Fig. 6, a similar conclusion can be drawn for RN-RTS as well. In this case, the value of K was chosen to be same as RN-1-LAS while the threshold values were 0.3 and 0.7 for the second and third rules respectively. Here we require a higher threshold value because the RTS algorithm requires some redundancy to continue the iterations even if the cost does not reduce.

Next we discuss the results for RN-MLAS algorithm. We observe from Fig. 3 and Fig. 4 that the error performance of RN-MLAS (for all the selection rules) matches with MLAS while there is a drastic reduction in complexity (more than 99%). Further, the RN-MLAS has a significantly reduced complexity compared to the one symbol update based RN-MLAS in [39]. To achieve near MLAS performance, we have taken the value of K in RN-MLAS to be 24. This lead to a selection of only 24 neighbors which is significantly lower compared to the 560 neighbors selected by the one symbol

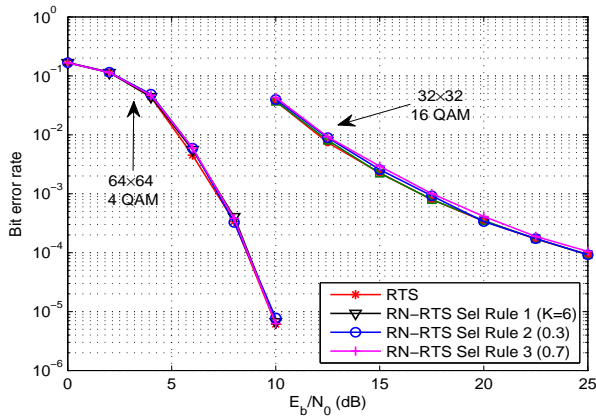


Fig. 5. Bit error rate performance of RTS and its reduced neighborhood variants for a 64×64 system with 4-QAM and for a 32×32 system with 16-QAM.

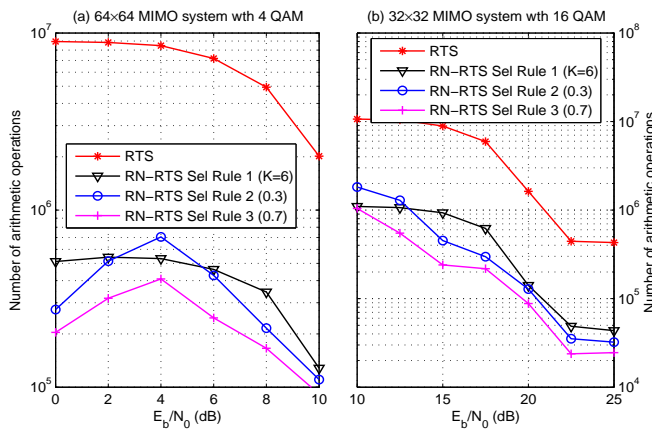


Fig. 6. Number of arithmetic operations required by RTS and its reduced neighborhood variants for a 64×64 system with 4-QAM and for a 32×32 system with 16-QAM.

update based RN-MLAS in [39] or the 41664 neighbors selected by MLAS. We chose the threshold values for the second and third selection rules to be 0.8 and 0.05 respectively. From Fig. 4 one can notice that the second rule outperforms the others for the 32×32 MIMO system with 16-QAM modulation while the first rule is slightly superior than the second rule for the 64×64 MIMO system with 4-QAM modulation.

We can conclude that the third rule is a clear winner for the algorithms which use one symbol neighborhood definitions i.e. RN-1-LAS and RN-RTS while for the algorithms based on multiple symbol neighborhoods i.e. RN-MLAS, first and second selection rules are better. Therefore, we will use the first two rules for evaluating the performance of the multiple symbol update based algorithms like SLAS and GLAS.

B. Comparison of SLAS and GLAS algorithms with the existing LAS algorithms

Let us investigate whether the idea of using a reduced neighborhood accompanied by searching in all the possible neighborhoods i.e. for $max_{stage} = 2N_t$ will yield improvement in error performance as compared to 1-LAS and MLAS algorithms. For this we have considered a 32×32 system

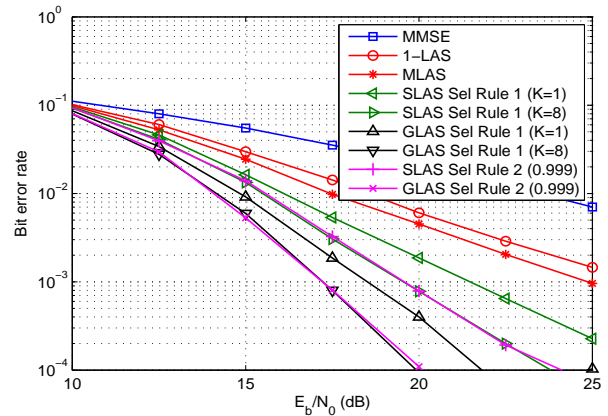


Fig. 7. Bit error rate performance taking MMSE solution as the initial vector for a 32×32 system with 16-QAM.

with 16-QAM modulation. The corresponding BER curves have been shown in Fig. 7. Whereas the expected complexity in terms of number of arithmetic operations along with the worst-case complexity and standard deviation is shown in Table II. The worst-case complexity is the maximum number of arithmetic operations required over 10^5 number of sample points. From the figure one can observe that the proposed SLAS and GLAS algorithms outperform the 1-LAS and MLAS algorithms in terms of error performance as well as complexity. For example, a BER of 10^{-3} can be achieved by GLAS with the first selection rule for $K = 8$ around 17.2 dB which is considerably lower than the 25 dB required by MLAS to achieve the same performance while the expected and worst-case complexities both are at least 25 and 9 times lower, respectively than the expected complexity of 1-LAS. Even if one compares the error performance of SLAS with the first selection rule for $K = 1$ (the worst possible option among SLAS and GLAS), it can be observed that there is a gain of around 3 dB compared to MLAS and the expected and worst-case complexities are much lower than the expected complexity of 1-LAS. This implies that the proposed schemes are much better in terms of both the error performance as well as complexity. One can also observe that the complexities of SLAS and GLAS have much lower standard deviations compared to 1-LAS and MLAS.

Further, we compare SLAS and GLAS for the first and second selection rules. We choose the threshold values for the second rule such that the respective BER performances of SLAS and GLAS match with the first selection rule for $K = 8$. Thus, we consider the threshold values as 0.999. Comparing

TABLE II
EXPECTED COMPLEXITY, WORST-CASE COMPLEXITY AND THE STANDARD DEVIATION IN TERMS OF NUMBER OF ARITHMETIC OPERATIONS ($\times 10^4$) PER BIT FOR A 32×32 SYSTEM WITH 16-QAM.

Detection Algorithm	$E_b/N_0 = 15$ dB			$E_b/N_0 = 20$ dB		
	Expected Complexity	Worst-Case Complexity	Standard Deviation	Expected Complexity	Worst-Case Complexity	Standard Deviation
1-LAS	2.704	517.2	36.566	1.523	291.2	20.587
MLAS	347.2	69400	4896.8	297.9	59550	4211.1
SLAS Sel Rule 1 ($K = 1$)	0.874	1.454	0.1024	0.8118	1.404	0.0628
SLAS Sel Rule 1 ($K = 8$)	4.807	9.344	0.777	4.591	9.25	0.501
SLAS Sel Rule 2 (0.999)	4.067	9.269	0.867	3.867	9.145	0.741
GLAS Sel Rule 1 ($K = 1$)	3.726	9.798	1.269	2.791	8.428	0.663
GLAS Sel Rule 1 ($K = 8$)	12.46	40.78	4.334	9.69	33.71	2.16
GLAS Sel Rule 2 (0.999)	13.22	43.27	4.45	10.03	32.35	2.83

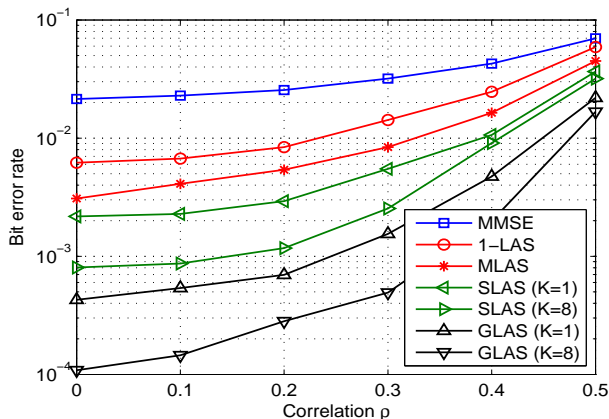


Fig. 8. Effect of correlation on bit error rate performance for a 32×32 system with 16-QAM at $E_b/N_0 = 20$ dB.

their complexity, one can see from Table II that the numbers of arithmetic operations required for SLAS and GLAS with the second selection rule are nearly of the same magnitude as with the first selection rule (for SLAS it is slightly lower whereas for GLAS it is slightly higher). Therefore, in the sequel we will use only the first selection rule for evaluating the performance of SLAS and GLAS.

It may be noted that the BER performance of SLAS and GLAS can be improved by choosing different selection parameters such as choosing a higher value of K . This means that by allowing additional complexity, these algorithms can improve upon their BER performance which is not possible in the case of 1-LAS or MLAS.

Next, we study the performance of SLAS and GLAS in a correlated MIMO channel. We consider the popular Kronecker model [48] according to which the complex MIMO channel can be expressed as

$$\bar{\mathbf{H}}_c = \Phi_t^{1/2} \bar{\mathbf{H}} \Phi_r^{1/2}, \quad (28)$$

where $\bar{\mathbf{H}}_c$, $\Phi_t = E\{\bar{\mathbf{H}}_c^H \bar{\mathbf{H}}_c\}$, and $\Phi_r = E\{\bar{\mathbf{H}}_c \bar{\mathbf{H}}_c^H\}$, represent the correlated channel matrix, transmit spatial correlation, and receive spatial correlation, respectively. Further for equal transmit and receive spatial correlation, the (i, j) th coefficients of Φ_t and Φ_r can be modeled as $\rho^{|i-j|}$ [49], where ρ is a number between 0 to 1, zero represents an uncorrelated case while one represents a fully correlated scenario. The BER results for $\rho = 0$ to 0.5 at $E_b/N_0 = 20$ dB are shown in Fig. 8. From the figure, we can observe that GLAS still outperforms the others. However with increasing correlation, the performance of all the algorithms deteriorates, and as the correlation increases, the difference between the algorithms decreases. It may be noted that the overall deterioration in the performance is because of the inferior quality of the initial solution (MMSE/ZF) due to correlation, and the increasing invalidity of the approximation in (16).

C. Comparison of SLAS and GLAS algorithms with other existing algorithms

Now we compare the SLAS and GLAS algorithms beyond the LAS family of algorithms i.e. with the best from the

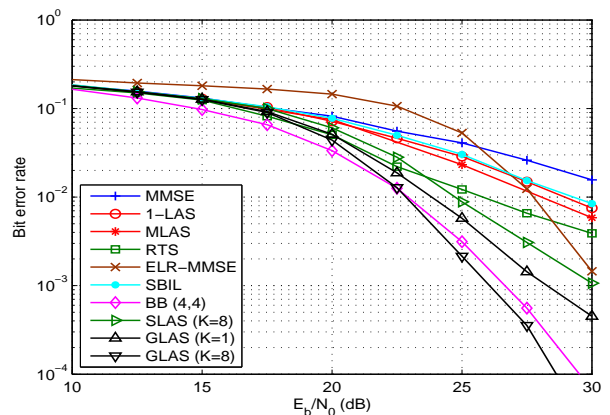


Fig. 9. Bit error rate performance for a 64×64 MIMO system with 64-QAM.

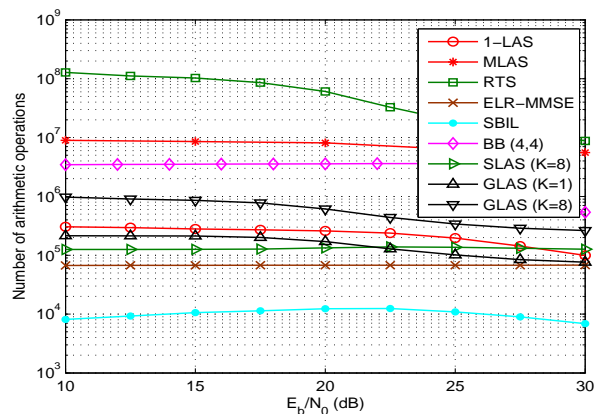


Fig. 10. Number of arithmetic operations required for a 64×64 MIMO system with 64-QAM.

available algorithms for large MIMO systems such as [19]–[22], [24], [25], [30], [32], [33], [37]. These algorithms can be grouped into four categories i.e. neighborhood search algorithms [19]–[22], algorithms based on LR techniques [24], sparse error recovery based algorithms [30], [32], and SDR based detectors [33]. We have chosen the most recent ones from each category i.e. RTS [19], [22], ELR-MMSE [24], sparsity boosted iterative linear detector (SBIL) [30], and branch and bound algorithm (BB) [33]. This is because these are the best known algorithms in their respective categories. However, due to its very high complexity we have skipped comparison with LTS [22] though it is true that LTS can provide a better error performance compared to RTS.

We compare the performance of the proposed SLAS and GLAS algorithms with RTS, ELR-MMSE, SBIL, and BB detectors for a 64×64 MIMO system with 64-QAM modulation. The bit error performance has been shown in Fig. 9 while the total number of arithmetic operations required is shown in Fig. 10. In the figure the results for SLAS and GLAS are provided for three settings of K i.e. SLAS for $K = 8$ and GLAS for $K = 1$ and 8. The simulation parameters for RTS, SBIL and BB are taken to be the same as in [22], [30], and [33], respectively. For the sake of completeness we have also showed the performance of 1-LAS, MLAS and the MMSE receiver. It may be noted that in the legends corresponding to BB, the numbers inside the braces i.e. BB (4,4), represent the

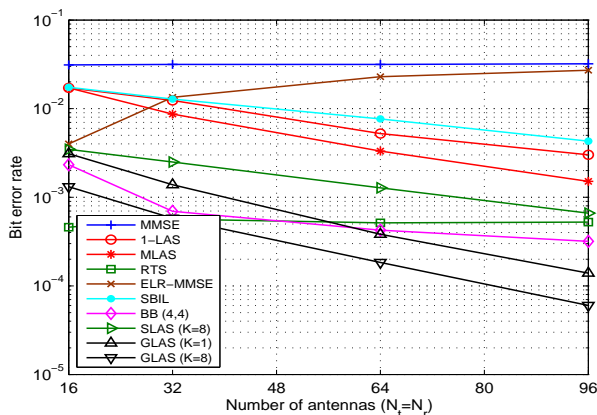


Fig. 11. Bit error rate performance for a 16-QAM at $E_b/N_0 = 18$ dB.

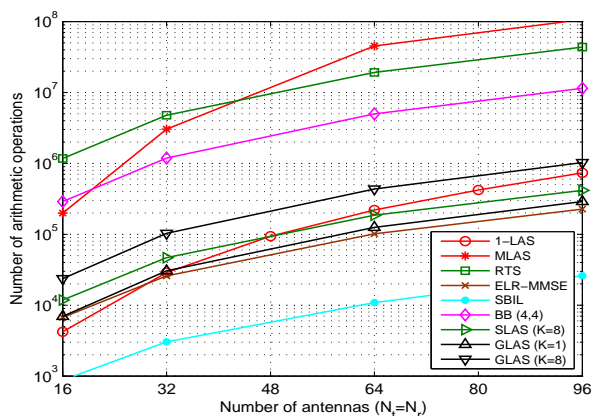


Fig. 12. Number of arithmetic operations required for a 16-QAM at $E_b/N_0 = 18$ dB.

depth and breadth of the algorithm.

From Fig. 9 one can see that among the considered variants of SLAS and GLAS, GLAS for $K = 8$ has the best error performance while SLAS for $K = 8$ has the most inferior performance. However, in terms of error performance, even SLAS ($K = 8$) is found to be superior compared to the other algorithms except BB, while in terms of complexity it loses out only to ELR-MMSE and SBIL. The others not only have an inferior error performance but also have a higher number of arithmetic operations. Only in the case of ELR-MMSE [24] and SBIL [30] we have a lower number of arithmetic operations but that comes at the cost of a worse error performance (see Fig. 9). Among the different categories of algorithms with which we are comparing SLAS and GLAS, we know that BB [33] is the best in terms of error performance. Comparing BB with the best of the three variants of GLAS/SLAS being considered i.e. GLAS ($K = 8$), it can be seen that GLAS ($K = 8$) has a gain of 1 dB at a BER of 10^{-4} while the savings in the number of arithmetic operations are more than 80%.

Further, we compare the performance for increasing number of antennas. We have shown the BER performance in Fig. 11 while the corresponding number of arithmetic operations required are shown in Fig. 12. Comparing with BB and RTS, we observe that the BER performance of RTS and BB saturate while the three variants of GLAS/SLAS continue to

TABLE III
ROI PER DECADE OF E_b/N_0 FOR A 64×64 SYSTEM WITH 16-QAM MODULATION.

Detection Algorithm	ROI for the E_b/N_0 range 20 dB to 30 dB
MMSE	0.72
SBIL	0.96
1-LAS	0.98
MLAS	1.11
RTS	1.11
SLAS ($K = 8$)	1.75
ELR-MMSE	2.01
GLAS ($K = 1$)	2.05
BB (4,4)	2.53
GLAS ($K = 8$)	3.31

improve with increasing number of antennas. Thus, after 32 and 64 antennas, GLAS ($K = 8$) and GLAS ($K = 1$) respectively, are superior compared to both RTS and BB. From the trend, this appears to be true for SLAS ($K = 1$) as well, but only after 96 antennas. Equally important to note, the complexities of the three variants of GLAS/SLAS are much lower than both RTS and BB, throughout the range of antennas considered. From the two figures it can also be seen that the three variants of GLAS/SLAS have superior error performance compared to the remaining algorithms. The trends with respect to complexity are same as those in the earlier discussed Fig. 10. It may be noted that unlike the other algorithms, BB does not require an initial MMSE solution. However, it has several orders of magnitude higher complexity than SLAS/GLAS algorithms and addition of the complexity of obtaining the initial MMSE solution in SLAS/GLAS does not make any material difference to this comparison.

D. Getting a sense of diversity order

In this sub-section we introduce a rate of improvement (ROI) metric to measure the improvement in error performance with SNR, of the different algorithms under consideration. We define it as

$$ROI = -\frac{\log BER_y - \log BER_x}{(\log y - \log x)} \quad (29)$$

where $x, y (> x)$ (not in dB) denote any two E_b/N_0 values and BER_x, BER_y are the corresponding BERs. In other words, if x and y differ by 10 dB, 10^{ROI} will provide the factor by which the BER in the BER vs E_b/N_0 curve will fall per decade of E_b/N_0 .

ROI metric in (29) is not the same as diversity order but can be viewed as an attempt to get a sense of the diversity orders achieved by the various detection algorithms under consideration.

We have compared the ROI's of the algorithms being considered for a 64×64 system with 16-QAM modulation. The algorithms have been arranged in increasing order of ROI in Table III. The E_b/N_0 range has been taken from 20 dB to 30 dB. From the table it can be seen that the entries corresponding to GLAS have a better ROI compared to the other algorithms. It can also be observed that the ROI values for GLAS improve with increasing value of K . Further, since the performance of ELR-MMSE and BB deteriorates as the number of antennas

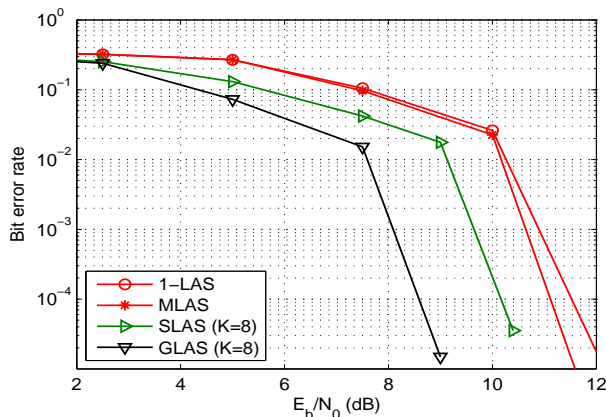


Fig. 13. Bit error rate performance for a Turbo coded 32×32 MIMO system with 16-QAM.

increase (can be inferred from Fig. 11) their ROI will come down as the number of antennas increase.

E. Extension to coded systems

Lastly, we explain the extension of SLAS and GLAS algorithms to coded systems. While hard decoding can be done using the detected symbol vector, soft decoding in such large systems is a challenge. This is because soft decoding requires the log likelihood ratio (LLR) values to be computed over a candidate set [50], [51]. As suggested in [52], we propose to generate the candidate sets in Algorithm 1 and Algorithm 2 by considering \mathbf{x}_{temp} - the intermediate solution vectors. This set is updated at every iteration until the algorithm terminates.

For evaluating the performance of coded systems, a Turbo code with constraint length 4 and a random interleaver with a length which is an integer multiple of 320 bits has been considered, for a 32×32 MIMO system with 16-QAM modulation. The convolutional code feedforward generator polynomials are $1 + D^2 + D^3$, $1 + D + D^2 + D^3$ and the feedback generator polynomial is $1 + D^2 + D^3$. The internal interleaver of the Turbo encoder was taken to be of length 64, arranged in the descending order starting from 64. For decoding we have used the iterative decoder in [52] with four iterations. The results obtained are shown in Fig. 13 for SLAS and GLAS with $K = 8$ and compared with 1-LAS and MLAS. From the figure it can be seen that the proposed SLAS and GLAS are better for coded systems too. Since the LLR values are computed over the set of intermediate solution vectors, the trend of complexities will also remain the same as in Fig. 10.

It may be noted that in contrast to the candidate sets used above, the size of the candidate sets in LAS and RTS in [18] and [22] respectively, are very large. This makes the exact LLR computations infeasible and therefore, [18] and [22] used an approximation to compute the LLR values. However, due to the significantly smaller number of candidates in SLAS and GLAS, the exact LLR computations are feasible now.

VII. CONCLUSION

Existing neighborhood search algorithms like LAS and RTS look for the best vector in a fixed neighborhood and search the entire fixed neighborhood. In this paper we first propose a method to obtain a reduced neighborhood. For this, we

derive a metric to select the vectors in the neighborhood which reduce the ML cost. This has helped us to achieve a significant reduction in complexity without loss in error performance. Then we use this reduced neighborhood definition to propose two new search algorithms which are not restricted to a fixed neighborhood, namely SLAS and GLAS. Simulation results show that the proposed algorithms have much improved error performance as well as lesser computational complexity compared to not only LAS and RTS algorithms but other existing algorithms as well. Lastly, we point out that the approach used may provide additional coding gain in coded systems using soft decoding.

REFERENCES

- [1] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [2] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [3] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.
- [4] J. Hoydis, S. ten Brink, and M. Debbah, "Massive mimo in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [5] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [6] J. Mietzner, R. Schober, L. Lampe, W. Gerstacker, and P. Hoeher, "Multiple-antenna techniques for wireless communications - a comprehensive literature survey," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 2, pp. 87–105, Second Quarter 2009.
- [7] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO wireless communications*. Cambridge University press, 2007.
- [8] E. Larsson, "MIMO detection methods: How they work [lecture notes]," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 91–95, May 2009.
- [9] S. W. Kim and K. P. Kim, "Log-likelihood-ratio-based detection ordering in V-BLAST," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 302–307, Feb. 2006.
- [10] S.-R. Lee, S.-H. Park, S. W. Kim, and I. Lee, "Enhanced detection with new ordering schemes for V-BLAST systems," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1648–1651, Jun. 2009.
- [11] B. S. Thian and A. Goldsmith, "A reduced-complexity MIMO receiver via channel ordering," in *Proc. 2009 IEEE Global Commun. Conf. (GLOBECOM)*, no. 1, pp. 1–6.
- [12] C.-Y. Hung, R. Chang, and W.-H. Chung, "A hybrid MMSE and K-best detection scheme for MIMO systems," in *Proc. 2012 IEEE Veh. Technol. Conf. (VTC Fall)*, pp. 1–5.
- [13] S. Han and C. Tellambura, "A complexity-efficient sphere decoder for MIMO systems," in *Proc. 2011 IEEE Int. Conf. Commun. (ICC)*, pp. 1–5.
- [14] 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.
- [15] T.-H. Kim, "Low-complexity sorted QR decomposition for MIMO systems based on pairwise column symmetrization," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1388–1396, Mar. 2014.
- [16] G. Romano, D. Ciuonzo, P. Salvo Rossi, and F. Palmieri, "Low-complexity dominance-based sphere decoder for MIMO systems," *Signal Process.*, vol. 93, no. 9, pp. 2500–2509, 2013.
- [17] K. Vardhan, S. Mohammed, A. Chockalingam, and B. 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, Apr. 2008.
- [18] S. Mohammed, A. Chockalingam, and B. Sundar Rajan, "A low-complexity near-ML performance achieving algorithm for large MIMO detection," in *Proc. 2008 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2012–2016.
- [19] B. Rajan, S. Mohammed, A. Chockalingam, and N. Srinidhi, "Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive tabu search," in *Proc. 2009 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1993–1997.

- [20] P. Li and R. Murch, "Multiple output selection-LAS algorithm in large MIMO systems," *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 399–401, May 2010.
- [21] 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, Dec. 2010.
- [22] 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, Nov. 2011.
- [23] A. A. Pereira and R. Sampaio-Neto, "A random-list based LAS algorithm for near-optimal detection in large-scale uplink multiuser MIMO systems," in *Proc. 2015 19th Int. ITG Workshop on Smart Antennas (WSA)*, pp. 1–5.
- [24] Q. Zhou and X. Ma, "Element-based lattice reduction algorithms for large MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 274–286, Feb. 2013.
- [25] K. Ali, W. Abediseid, and M.-S. Alouini, "Sequential decoders for large MIMO systems," in *Proc. 2014 12th Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2014, pp. 709–716.
- [26] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," *IEEE Trans. Commun.*, vol. 56, no. 2, pp. 309–318, Feb. 2008.
- [27] P. Som, T. Datta, A. Chockalingam, and B. Rajan, "Improved large-MIMO detection based on damped belief propagation," in *Proc. 2010 IEEE Workshop Inf. Theory (ITW) 2010*, pp. 1–5.
- [28] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. 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.
- [29] T. 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, Jan. 2013.
- [30] 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.
- [31] Y. Fadlallah, A. Aissa-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.
- [32] J. W. Choi and B. Shim, "New approach for massive MIMO detection using sparse error recovery," in *Proc. 2014 IEEE Global Commun. Conf. (GLOBECOM)*, pp. 3754–3759.
- [33] 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, Mar. 2016.
- [34] N. Souto, M. Ribeiro, and P. Sebastio, "Semidefinite relaxations for MIMO transmissions with high-order QAM constellations," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 984–988, Jul. 2016.
- [35] A. Yellepeddi, K. J. Kim, C. Duan, and P. Orlik, "On probabilistic data association for achieving near-exponential diversity over fading channels," in *Proc. 2013 IEEE Int. Conf. Commun. (ICC)*, pp. 5409–5414.
- [36] S. Yang, L. Wang, T. Lv, and L. Hanzo, "Approximate bayesian probabilistic-data-association-aided iterative detection for MIMO systems using arbitrary M-ary modulation," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1228–1240, Mar. 2013.
- [37] P. Svac, 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, Sep. 2013.
- [38] 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.
- [39] A. K. Sah and A. K. Chaturvedi, "Reduced neighborhood search algorithms for low complexity detection in MIMO systems," in *Proc. 2015 IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1–6.
- [40] —, "Beyond fixed neighborhood search in the likelihood ascent algorithm for MIMO systems," in *Proc. 2016 IEEE Int. Conf. Commun. (ICC)*, pp. 1–6.
- [41] —, "An unconstrained likelihood ascent based detection algorithm for large MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 2262–2273, Apr. 2017.
- [42] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.
- [43] R. F. H. Fischer and C. Windpassinger, "Real versus complex-valued equalisation in V-BLAST systems," *Electron. Lett.*, vol. 39, no. 5, pp. 470–471, Mar. 2003.
- [44] T.-H. Liu and C.-N. Chiu, "On fast preprocessing schemes for the real-valued spatially multiplexed MIMO detectors," *Int. J. Commun. Syst.*, vol. 27, no. 2, pp. 357–376, Feb. 2014.
- [45] Å. Björck, *Numerical methods for least squares problems*. SIAM, 1996.
- [46] 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, Fourth-quarter 2015.
- [47] P. Sprent, "Inversion of nearly diagonal matrices," *The Math. Gazette*, vol. 49, no. 368, pp. 184–89, 1965.
- [48] J. P. Kermaol, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.
- [49] D. Wubben, V. Kuhn, and K. D. Kammeyer, "On the robustness of lattice-reduction aided detectors in correlated MIMO systems," in *Proc. 2004 IEEE 60th Veh. Technol. Conf. (VTC Fall)*, vol. 5, pp. 3639–3643.
- [50] C. Studer and H. Bolcskei, "Soft-input soft-output single tree-search sphere decoding," *IEEE Trans. on Inf. Theory*, vol. 56, no. 10, pp. 4827–4842, Oct. 2010.
- [51] G. Papa, D. Ciuonzo, G. Romano, and P. Salvo Rossi, "A dominance-based soft-input soft-output MIMO detector with near-optimal performance," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4320–4335, Dec. 2014.
- [52] B. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.



Abhay Kumar Sah (S'15) received the B.Tech. degree in Electronics and Communications Engineering from the Jaypee University of Information Technology, Waknaghat (H.P.), India, in 2010. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, IIT Kanpur, Kanpur, India. His current research interests are in wireless communication and multi-antenna systems. He was a recipient of the TCS Research Scholarship and the Vice-Chancellors Gold Medal for outstanding performance in his bachelors degree.



A. K. Chaturvedi (S'91-M'96-SM'03) received the B.Tech., M.Tech., and Ph.D. degrees in Electrical Engineering from Indian Institute of Technology Kanpur in 1986, 1988, and 1995, respectively. He served the Department of Electronics Engineering at Indian Institute of Technology, Banaras Hindu University, Varanasi from 1994 to 1996. Subsequently, he joined the faculty of the Department of Electronics and Computer Engineering at Indian Institute of Technology Roorkee. In 1999, he moved to Indian Institute of Technology Kanpur where he also held the positions of Head of the Department of Electrical Engineering, Dean of Research & Development and Deputy Director. He is now the Director of IIT Roorkee. Prof. Chaturvedi was the Coordinator of the BSNL-IITK Telecom Centre of Excellence which has done a large number of projects for the Indian telecom sector. He is a recipient of the Distinguished Teacher award of IIT Kanpur and Tan Chin Tuan Fellowship of Nanyang Technical University, Singapore. Prof. Chaturvedi is a member of the Nomination and Election Committee of IEEE Communication Society. He has also served on the Teaching Awards Committee of IEEE. He is a founding member of the Telecom Standards Development Society of India (TSDSI) and the current Chair of the Joint Telematics Group which organizes the annual National Conference on Communications. Prof. Chaturvedi was a member of the DoT committee which recommended criteria for spectrum allocation to telecom operators, in 2008. His research interests are in Communication Theory and Wireless Communications. He has supervised a large number of graduate students and publishes regularly in IEEE journals in the area of Communications and flagship conferences of IEEE Communication Society. He is known for his contributions to waveform shaping and sequence design. Nowadays, he is working on the detection problem in massive MIMO systems.