Reduced Neighborhood Search Algorithms for Low Complexity Detection in MIMO Systems

Abhay Kumar Sah† and A. K. Chaturvedi‡
Department of Electrical Engineering
Indian Institute of Technology Kanpur, Kanpur, India 208016
Email: {abhaysah†, akc‡}@iitk.ac.in

***Abstract*—Neighborhood search algorithms such as likelihood ascent search (LAS) and reactive tabu search (RTS) have been proposed for low complexity detection in multiple-input multiple-output (MIMO) systems having a large number of antennas. Both these algorithms are iterative and search for the vector which minimizes the maximum likelihood (ML) cost in the neighborhood. In this paper we propose a way to reduce the size of the neighborhood. For this, we propose a metric and a selection rule 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. This reduced set is used to evaluate the performance of the resulting LAS and RTS algorithms. Simulation results show that this reduces the complexity significantly while maintaining the error performance. We also show that the proposed reduced neighborhood algorithms can make MIMO systems with several hundred antenna pairs feasible.**

I. Introduction

Low complexity detection in multiple-input multiple-output (MIMO) systems with a large number of antennas is a key challenge [1], [2]. Several detection algorithms [3]–[11] have been proposed in the literature to address this problem. Except a few [10], [11] all others belong to a class of neighborhood search algorithms, primarily categorized into likelihood ascent search (LAS) [3] and reactive tabu search (RTS) [5] algorithms. The main idea of these algorithms is to begin with an initial guess, generate a neighborhood around this initial guess and replace it with the best solution in the neighborhood (in case of LAS the initial guess is also included). In LAS this process continues till there is no further reduction in the ML cost while in RTS some additional escape polices are used to avoid cycles and early terminations. Both these algorithms search the entire 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 lead to a higher ML cost. Hence there is a need to efficiently select the vectors which are likely to reduce the ML cost.

Motivated by this idea we suggest an error vector and channel coefficients based metric for generating a reduced neighborhood. The larger magnitude components of this metric will lead to inclusion of the corresponding vector in the neighborhood. Hence we select the indices of the K largest magnitude components and generate a reduced neighborhood with significantly lesser number of vectors than the full neighborhood. This reduced neighborhood is used in two versions [3], [4] of the LAS algorithm (reduced neighborhood 1-LAS and reduced neighborhood MLAS) and a reduced neighborhood RTS algorithm. We show that these proposed algorithms significantly reduce the computational complexity. We also show that this reduction in complexity makes MIMO systems with several hundreds of antennas [12] feasible.

The rest of the paper is structured as follows: Section II describes the system model and the required initial background to develop the algorithm which has been proposed in Section III. A novel metric is established in III-A, a selection rule is applied over it in III-B and lastly the reduced neighborhood algorithms have been proposed in III-C, III-D and III-E respectively. Section IV presents the simulations results and finally Section VI concludes the paper. In this paper bold face capital letters denote matrices, capital letters denote vectors and small letters are used to denote a scalar quantity. Bar is used to denote a complex quantity.

II. Preliminaries

A MIMO system uses N_t number of transmit antennas for transmission and N_r number of receive antennas for reception ($N_t \leq N_r$). The input-output relationship can be mathematically modeled as

$$\bar{Y} = \bar{\mathbf{H}}\bar{X} + \bar{N}, \tag{1}$$

where $\bar{Y} = (\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_{N_r})^T$ is $(N_r \times 1)$ received signal vector and $\bar{y}_i$ represents data received at ith receive antenna. $\bar{X} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_{N_t})^T$ is $(N_t \times 1)$ transmitted signal vector where $\bar{x}_i \in \bar{\Omega}$ represents data transmitted through ith transmit antenna. The $\bar{\Omega}$ is a set of M complex symbols such as square M-QAM constellation. $\bar{\mathbf{H}}$ denotes $(N_r \times N_t)$ channel matrix with each coefficient $\bar{h}_{ij} \sim \mathcal{CN}(0,1)$ and $\bar{N} = (\bar{n}_1, \bar{n}_2, \cdots, \bar{n}_{N_r})^T$ represents $(N_r \times 1)$ i.i.d. additive white Gaussian noise (AWGN) vector with each $\bar{n}_i \sim \mathcal{CN}(0, \sigma^2)$. This complex system model (1) can be formulated as equivalent real system model as follow

$$Y = \mathbf{H}X + N, \tag{2}$$

where $Y = (\Re\{\bar{Y}\}^T \ \Im\{\bar{Y}\}^T)^T$ is $(2N_r \times 1)$ real equivalent received vector. $\bar{X} = (\Re\{\bar{X}\}^T \ \Im\{\bar{X}\}^T)^T$ represents

This work was supported by Tata Consultancy Services Limited.

$(2N_t \times 1)$ real equivalent transmit vector and each $x_i \in \Omega$. $\Omega = \{\pm 1, \pm 3 \cdots \pm (\sqrt{M}-1)\}$ is a set of $\sqrt{M}$ real symbols drawn from one dimensional constellation i.e. $\sqrt{M}$-PAM. $\mathbf{H}$ denotes the $(2N_r \times 2N_t)$ equivalent channel matrix given by

$$\mathbf{H} = \begin{bmatrix} \Re\{\bar{\mathbf{H}}\} & -\Im\{\bar{\mathbf{H}}\} \\ \Im\{\bar{\mathbf{H}}\} & \Re\{\bar{\mathbf{H}}\} \end{bmatrix} \tag{3}$$

and $N = (\Re\{\bar{N}\}^T \; \Im\{\bar{N}\}^T)^T$ is $(2N_r \times 1)$ equivalent noise vector. We will use this equivalent real system model (2) throughout the paper. At receiver, our objective is to find the actual transmitted symbol vector X among the all possible $\sqrt{M}^{2N_t}$ transmit vectors which is nearest to the received signal vector Y for the given channel matrix $\mathbf{H}$. Mathematically, this is stated as

$$\hat{X} = \underset{X \in \Omega^{2N_t}}{\operatorname{argmin}} \|Y - \mathbf{H}X\|^2 \tag{4}$$

and well known as ML detection. The computational complexity of ML detection is exponential in nature and practically it can not be applied even for small number of antenna pairs. In the literature, MIMO systems for up to hundred of antenna pairs, LAS algorithm [3], [4] and RTS [5] algorithm have been suggested. These algorithms proceed iteratively, initialize with a initial solution vector and search for the best solution in the neighborhood of the initial solution vector until there is no further reduction in the ML cost (4). The neighborhood of a vector $X = (x_1, x_2, \cdots, x_{2N_t})^T$ where $x_i \in \Omega \, \forall i = 1, \cdots, 2N_t$ is defined as follows: Let us denote it by $\mathcal{N}_{\mathcal{L}}(X)$ and consider a vector $\tilde{X} \in \mathcal{N}_{\mathcal{L}}(X)$ which differs from the original vector at exactly $\mathcal{L}$-symbols. We define a set $\mathcal{I}_k$ which contains the indices at which $\tilde{X}$ differs from X and for a $\mathcal{L}$ symbol update there are $\binom{2N_t}{\mathcal{L}}$ such $\mathcal{I}_k$'s. Now, the neighboring vector $\tilde{X} \in \mathcal{N}_{\mathcal{L}}(X)$ can be defined mathematically as

$$\tilde{x}_i(j,k) = \begin{cases} \omega_j, & \forall \omega_j \neq x_i, \, \omega_j \in \Omega \;\&\; i \in \mathcal{I}_k \\ x_i, & i \notin \mathcal{I}_k, \end{cases} \tag{5}$$

where $i = 1, 2, \ldots, 2N_t$, $j = 1, 2, \cdots, \sqrt{M} - 1$ and $k = 1, 2, \cdots, \binom{2N_t}{\mathcal{L}}$. The $\tilde{x}_i(j,k)$ represents the ith element of (j,k)th neighbor of X. This means there are $(\sqrt{M} - 1)^{\mathcal{L}}\binom{2N_t}{\mathcal{L}}$ number of vectors in the $\mathcal{L}$-Symbol neighborhood of X or $\mathcal{N}_{\mathcal{L}}(X)$. For example if we consider the set $\Omega = \{-3, -1, 1, 3\}$ and a vector $X = $ [1 3 -1] of length 3. Then the 1-symbol neighborhood of this vector X can be constructed by flipping the elements of this vector one by one with the other elements of the set Ω. Corresponding to the first element of this vector we have three vectors in 1-symbol neighborhood [-3 3 -1], [-1 3 -1] and [3 3 -1]. Similarly for other two locations we have six more vectors. Based on the above neighborhood definition there exist different versions of LAS and RTS algorithms such as 1-LAS algorithm [3], MLAS algorithm [4], multiple output selection LAS [9], random restart reactive tabu search [7], layered tabu search [8] etc. One can refer to these cited papers for more detail.

III. Proposed Algorithm

In neighborhood based algorithms the aim is to find the minimum Euclidean distance vector (best vector) in the neighborhood. This problem can be broken into two parts. First is to identify a small set of indices $\mathcal{I}_k$'s which will cause reduction in the ML cost and the second is to find the best vector for the every $\mathcal{I}_k$. For an $\mathcal{L}$ symbol neighborhood there are $\binom{2N_t}{\mathcal{L}}$ such $\mathcal{I}_k$'s and corresponding to each $\mathcal{I}_k$ there are $(\sqrt{M}-1)^{\mathcal{L}}$ possible neighbors. The second part has been addressed in [4] for $\mathcal{L} = 1$ and 2 while the first part is yet to be addressed in the literature. Complexity of the first part alone is of the order of $N_t^{\mathcal{L}}$ which makes it infeasible for a huge antenna setting. Now we derive a simple metric to reduce the complexity of first part significantly.

A. A simple metric to reduce the size of the neighborhood

Let us define the cost vector for a solution vector X^r at rth iteration as

$$E^r = Y - \mathbf{H}X^r. \tag{6}$$

The L_2 norm of this vector E^r will give the ML cost of the vector X^r. Our objective is to minimize $\|E^r\|^2$. This can be done by minimizing $|e_j^r| \; \forall j = 1, 2, \cdots 2N_r$. Here the e_j^r's can be expressed as

$$e_j^r = y_j - \sum_{i=1}^{2N_t} h_{ji} x_i^r. \tag{7}$$

For a particular value of i consider replacing the x_i^r in X^r with the neighboring point in the constellation to generate $\tilde{X}^r$. The neighboring point can be smaller or larger than the x_i^r. Mathematically this can be captured by $\tilde{x}_i^r = x_i^r + \eta \, d_{min}$ where η is an integer and d_{min} is the minimum distance between symbols in the real equivalent system. Consequence change in e_j^r can be expressed as $\tilde{e}_j^r = e_j^r - \eta \, d_{min} \, h_{ji}$. We propose to find the η and the index i which minimize $\sum_{j=1}^{2N_r} |\tilde{e}_j|^2$. This can be expressed as

$$\min_{\eta, i} \sum_{j=1}^{2Nr} (e_j^r - \eta \, d_{min} \, h_{ji})^2 = \min_{\eta, i} \|E^r - \eta \, d_{min} \, H_i\|^2 \tag{8}$$

where H_i represents the ith column of the channel matrix $\mathbf{H}$. We denote the optimal value of η and i which minimize (8) as η^* and i^*. Taking partial derivative of the objective function in (8) with respect to η for a fixed i and equating it to zero we get η as a function of H_i. Let us denote this as η_i^* as given below

$$\eta_i^* = \left\lceil \frac{{E^r}^T H_i}{d_{min} \|H_i\|^2} \right\rfloor = \frac{{E^r}^T H_i}{d_{min} \|H_i\|^2} + \delta_i \tag{9}$$

where $\lceil \cdot \rfloor$ denotes the rounding operation and δ_i is the rounding error. As we have considered the change at only one index (i.e. i), (9) is similar to [4] for $\mathcal{L} = 1$. Even though η_i^* is not necessarily optimal for $\mathcal{L} > 1$, we show below that it is helpful in identifying i's at which the change will lead to maximum reduction in the cost function (4). We do this by

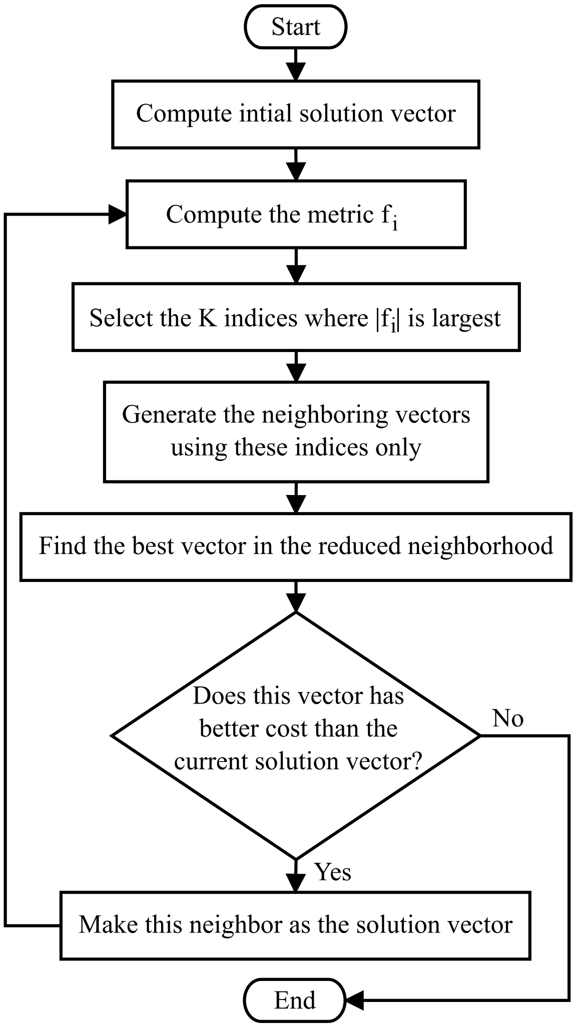

Fig. 1. Reduced Neighborhood (RN) LAS Algorithm.

substituting the value of η_i^* from (9) to (8) leading to the following optimization problem

$$\begin{aligned}
&\min_i \left\| E^r - \left(\frac{E^{rT} H_i}{d_{min}\|H_i\|^2} + \delta_i \right) d_{min} H_i \right\|^2 \\
&= \min_i \left\{ -\frac{\left(E^{rT} H_i\right)^2}{\|H_i\|^2} + \delta_i^2 d_{min}^2 \|H_i\|^2 \right\} \\
&= \max_i \left\{ \frac{\left(E^{rT} H_i\right)^2}{\|H_i\|^2} - \delta_i^2 d_{min}^2 \|H_i\|^2 \right\}. \qquad (10)
\end{aligned}$$

For one symbol neighborhood the ith column of the channel matrix $\mathbf{H}$ which maximizes the objective function in (10) is the optimal i^*.

In the above maximization problem we observe that the second term in (10) is negligible (discussed later in Section V) compared to the first term and hence can be ignored. We denote $\frac{E^{rT} H_i}{\|H_i\|}$ by f_i for $i = 1, 2, \cdots, 2N_t$ and use it to generate an $\mathcal{L}$ symbol reduced neighborhood in the next subsection.

B. Selection Rule

We propose to update only those indices which have large $|f_i|$ values. For this we choose the indices (say K which is significantly smaller compared to $2N_t$) having the largest $|f_i|$ 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 $\binom{2N_t}{\mathcal{L}}$ to $\binom{K}{\mathcal{L}}$ and at the same time not likely to compromise on the error performance.

We use the neighborhood definition (5) to generate a reduced neighborhood ($\mathcal{RN}$) which differs with X only at the indices contained in $\mathcal{I}_k$. Let us denote an element of the reduced neighborhood as $\tilde{X}$. The ith element of $\tilde{X}$ can be expressed as

$$\tilde{x}_i^r = \begin{cases} [x_i^r + \eta_i^* d_{min}], & i \in \mathcal{I}_k \\ x_i^r, & i \notin \mathcal{I}_k \end{cases} \qquad (11)$$

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

$$[A] = \begin{cases} \sqrt{M} - 1, & A > \sqrt{M} - 1 \\ -\sqrt{M} + 1, & A < -\sqrt{M} + 1 \\ A, & otherwise. \end{cases} \qquad (12)$$

We will update the indices in the set $\mathcal{I}_k$ only. The overall process is summarized in the flow chart in Fig. 1.

Now we use the above reduced neighborhood definition (11) to propose two reduced neighborhood based algorithms.

C. Reduced Neighborhood 1-LAS (RN-1-LAS)

In this algorithm we start with an initial vector which can be generated through either a matched filter (MF), zero forcing (ZF) or a minimum mean square error (MMSE) receiver. We denote this initial vector as X^0. The expressions for X^0 for these receivers are

$$\begin{aligned}
X_{MF}^{(0)} &= \lceil \mathbf{H}^T Y \rfloor, \\
X_{ZF}^{(0)} &= \lceil (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T Y \rfloor, \\
X_{MMSE}^{(0)} &= \lceil (\mathbf{H}^T \mathbf{H} + \sigma^2 \mathbf{I})^{-1} \mathbf{H}^T Y \rfloor, \qquad (13)
\end{aligned}$$

where the subscripts denote the type of receiver and $\lceil \cdot \rfloor$ represents element wise rounding operation to the set Ω. One can begin with any one of the above initial vector. Corresponding to the initial vector we compute the metric f_i for $i = 1, 2, \cdots, 2N_t$ and apply the selection rule discussed in III-B on f_i to choose indices i from $\{1, 2, \cdots, 2N_t\}$. Now based on the selected indices we generate a one symbol reduced neighborhood $\mathcal{RN}_1$ by updating the symbols of X^0 at these indices only. It means, if K number of indices are selected out of $2N_t$ then there are only K vectors in the reduced neighborhood.

We use the optimal η_i^* expressed in (9) for generating the reduced neighborhood and select the vector that minimizes the cost function (4). If there is an improvement in the cost compared to initial vector we replace the initial vector with the selected vector otherwise terminate the algorithm. The complete algorithm is given in tabular form in Algorithm 1.

Algorithm 1: Reduced Neighborhood 1-LAS Algorithm

Input : $Y, \mathbf{H}, \Omega, K$
Output: $\hat{X}$

Initialization $r = 0$;
$X^r \leftarrow$ output of MF/ZF/MMSE detetor;
$Cost_{next} \leftarrow \|Y - \mathbf{H}X^r\|^2$ & $Cost_{pre} \leftarrow \infty$;
while $Cost_{next} < Cost_{pre}$ **do**
  $f_i = \frac{(Y-\mathbf{H}X^r)^T H_i}{\|H_i\|}$ $\forall i = 1, 2, \cdots 2N_t$;
  $Ind \leftarrow$ select K indices where $|f_i|$ is largest;
  $Cost_{pre} \leftarrow Cost_{next}$;
  $X_{temp} \leftarrow X^r$;
  for $j \leftarrow 1$ **to** *length* (Ind) **do**
    $i \leftarrow j$th element of Ind;
    $\eta_i^* = \left\lceil \frac{f_i}{d_{min}\|H_i\|} \right\rfloor$;
    for $k \leftarrow 1$ **to** $2N_t$ **do**
      if $k = i$ **then**
        $\tilde{x}_k = [x_i + \eta_i^* d_{min}]$;
      else
        $\tilde{x}_k = x_i$;
      end
    end
    $Cost_{temp} = \|Y - \mathbf{H}\tilde{X}\|^2$;
    if $(Cost_{temp} < Cost_{next})$ **then**
      $X_{temp} \leftarrow \tilde{X}$;
      $Cost_{next} = Cost_{temp}$;
    end
  end
  $X^{r+1} \leftarrow X_{temp}$;
  $r = r + 1$;
end
return $\hat{X} \leftarrow X^{r+1}$.

D. Reduced Neighborhood MLAS (RN-MLAS)

The reduced neighborhood MLAS algorithm starts by obtaining a solution with RN-1-LAS algorithm. We use the solution to generate the reduced 2-symbol neighborhood set $\mathcal{RN}_2$ followed by reduced 3-symbol neighborhood set $\mathcal{RN}_3$. If at any stage either $\mathcal{RN}_2$ or $\mathcal{RN}_3$ finds a better solution than the solution obtained through 1-LAS, we repeat the above process considering this solution as the initial vector. The algorithm terminates when there is no improvement in the cost function (4) at all the three stages.

For generating an $\mathcal{L}$-symbol reduced neighborhood of a vector X, we compute the metric f_i for $i = 1, 2, \cdots, 2N_t$ and then choose the indices using the selection rule. Now we construct the set $\mathcal{I}_k$ using these selected indices only and apply the neighborhood definition (11). For a particular $\mathcal{I}_k$, $\mathcal{L}$ symbols leads to $(\sqrt{M}-1)^{\mathcal{L}}$ possible updates. For a 2-symbol update considering $\mathcal{I}_k = \{i_1,\, i_2\}$, the optimal value of $\eta_{i_1}^*$ and $\eta_{i_2}^*$ is given by [4]

$$\begin{bmatrix} \eta_{i_1}^* \\ \eta_{i_2}^* \end{bmatrix} = \left\lceil \frac{1}{d_{min}} \begin{bmatrix} \langle H_{i_1}, H_{i_1} \rangle & \langle H_{i_1}, H_{i_2} \rangle \\ \langle H_{i_2}, H_{i_1} \rangle & \langle H_{i_2}, H_{i_2} \rangle \end{bmatrix}^{-1} \begin{bmatrix} E^{r^T} H_{i_1} \\ E^{r^T} H_{i_2} \end{bmatrix} \right\rfloor, \tag{14}$$

which is an extension of the result in (9). Similarly this result can be extended and used for a 3-symbol update.

E. Reduced Neighborhood RTS (RN-RTS)

RTS algorithm is similar to LAS algorithm, except that it moves to the best vector in the neighborhood even if the best vector in the neighborhood is worse, in terms of ML cost, than the current solution vector. This strategy allows the algorithm to escape from early terminations. The process is repeated for a fixed number of iterations and to avoid cycles a tabu matrix is used. Finally, the best solution vector across all the iterations is declared as the solution vector. The detailed algorithm can be found in [5].

In RN-RTS, instead of the full neighborhood, we use the reduced neighborhood definition given in (11). Rest of the steps are same as in RTS.

Complexity Analysis: The main steps in the RN-MLAS algorithm are: $i)$ initialization, $ii)$ computation of $E^{r^T} H_i$, $\|H_i\|^2$ for $i = 1, 2, \cdots, 2N_t$ and $\langle H_i, H_j \rangle$ 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 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 $O(K^3)$ and $O({N_t}^3)$ complexity per transmit vector for RN-MLAS and MLAS respectively. For the case of RN-1-LAS the complexity is $O(K)$ in comparison to $O(N_t)$ of 1-LAS. Thus, depending on the value of K, the overall complexity of RN-LAS algorithms can be significantly lower than that of LAS algorithms. Similar conclusions can be drawn for RN-RTS algorithm too.

IV. Simulation Results

There are two aspects that need to be investigated: first whether this potentially significant reduction in complexity of RN-1-LAS, RN-MLAS and RN-RTS can be accompanied by near LAS/RTS performance in terms of bit error rate (BER). Secondly, whether this reduction in complexity can make a huge antenna setting, say of the order of several hundreds of antennas [12] feasible. However, the answer of the second part is linked to the answer of the first part.

Let us consider a 32×32 MIMO system for 4-QAM and 16-QAM modulations. The corresponding BER results are shown in Fig. 2 and Fig. 4. From the figure we can see that the error performance of RN-1-LAS matches the performance of 1-LAS. In this case the value of K was taken to be 6 which lead to a selection of only 6 neighbors from a total of 64 neighbors. From Fig. 3 and Fig. 5 we can see that this leads to more than 90% reduction in complexity. From both these

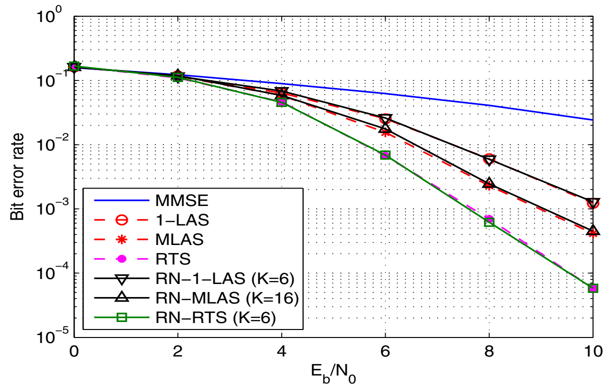

Fig. 2. Bit error rate performance of reduced neighborhood algorithms taking MMSE solution as the initial vector for a 32×32 MIMO system with 4-QAM modulation.

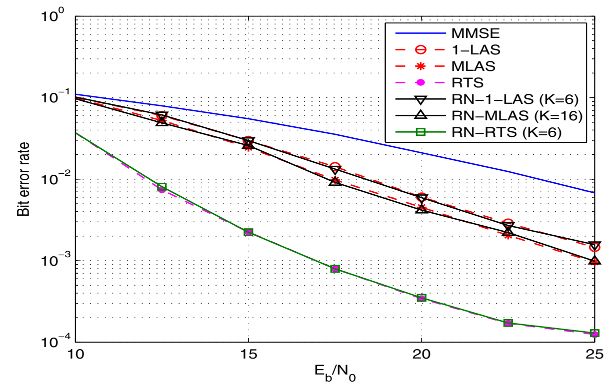

Fig. 4. Bit error rate performance of reduced neighborhood algorithms taking MMSE solution as the initial vector for a 32×32 MIMO system with 16-QAM modulation.

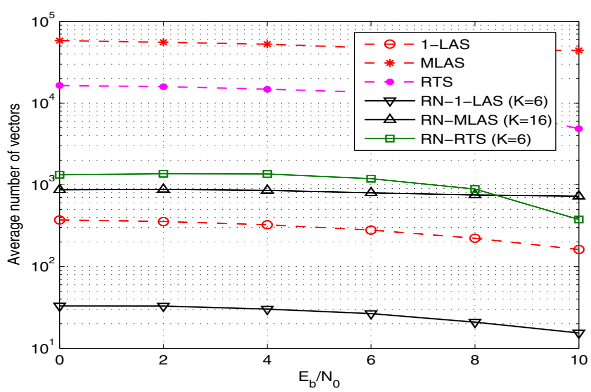

Fig. 3. Average number of vectors searched by the reduced neighborhood algorithms taking MMSE solution as the initial vector for a 32×32 MIMO system with 4-QAM modulation.

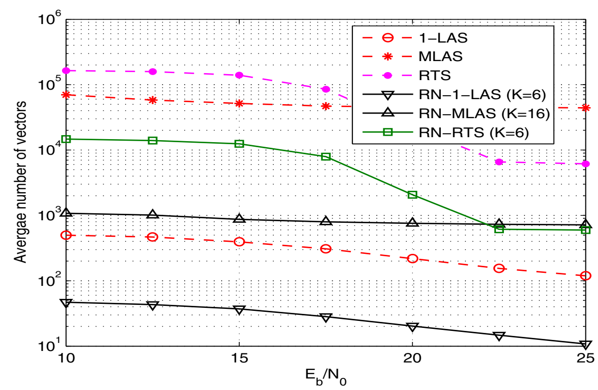

Fig. 5. Average number of vectors searched by the reduced neighborhood algorithms taking MMSE solution as the initial vector for a 32×32 MIMO system with 16-QAM modulation.

figures, 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 and the other simulation parameters are same as in [5]. For RN-MLAS, the value of K was taken to be 16 which lead to a selection of only 560 neighbors which is significantly lower compared to the 41664 neighbors selected by MLAS. It leads to a near MLAS error performance while providing more than 98% reduction in complexity.

This significantly reduced complexity makes the proposed algorithms candidates for realizing much larger systems. Therefore, we now study the performance of RN-1-LAS, RN-MLAS and RN-RTS for massive MIMO systems. We consider 4-QAM modulation at an E_b/N_0 value of 10 dB. For RN-1-LAS/RN-RTS and RN-MLAS, we consider antenna pairs up to 1000 and 250 respectively while for 1-LAS/RTS and MLAS, because of their higher complexity, we could consider antenna pairs only up to 400 and 150 respectively. From Fig. 6 it is clear that RN-1-LAS, RN-MLAS and RN-RTS have the same error performance as 1-LAS, MLAS and RTS respectively (their respective curves overlap). Further, at this value of E_b/N_0 it can be seen that for the RTS family the BER starts increasing. Fig. 7 shows that RN-1-LAS and RN-RTS provide approximately 95% reduction in complexity while RN-MLAS provides more than 95% savings in complexity. For the proposed algorithms, RN-1-LAS/RN-RTS and RN-MLAS the value of K was chosen to be 5% and 10% respectively of $2N_t$.

From Fig. 7 we observe that the average number of vectors are increasing with the number of transmit antennas but it may be noted that this does not convey the per symbol complexity. This is because the number of transmitted symbols also increase with the number of transmit antennas. In fact, the rate of increase in average computations per symbol reduces as the number of antennas increase.

V. Justification for Ignoring the Term Involving δ

The maximization problem in (10) involves two terms in which the second term occurs due to the rounding error δ_i (9). Let us quantify this rounding error. We begin by assuming that the initial solution vector (say X_2) differs from the transmitted vector (say X_1) at one index (say uth) only. Then the uth index of X_2 can be written in terms of the uth index of X_1 as $x_{2u} = x_{1u} + m d_{min}$ where m is an integer. Now by evaluating

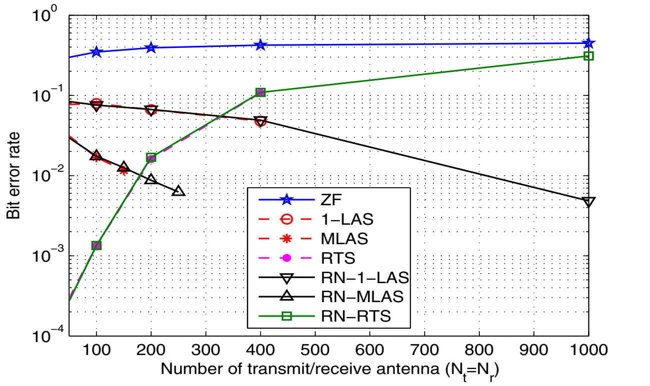

Fig. 6. Bit error rate performance of reduced neighborhood algorithms taking ZF solution as the initial vector for a 4-QAM modulation at $E_b/N_0 = 10$dB.

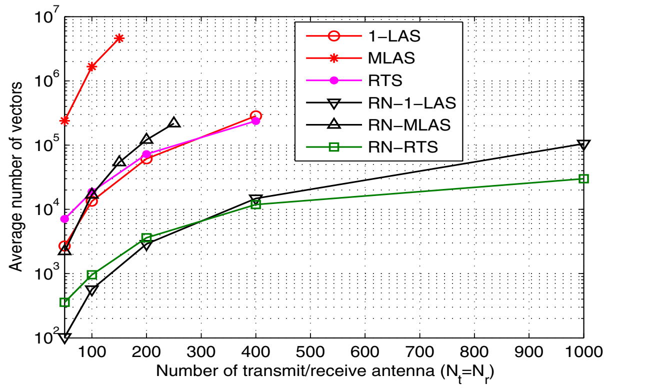

Fig. 7. Average number of vectors searched by the reduced neighborhood algorithms taking ZF solution as the initial vector for 4-QAM modulation at $E_b/N_0 = 10$dB.

(9) we get the following expression

$$\eta_i^* = \left\lceil \frac{(Y - \mathbf{H}X_1 - md_{min}\, H_u)^T H_i}{d_{min}\|H_i\|^2} \right\rfloor = \left\lceil \frac{(N - md_{min}\, H_u)^T H_i}{d_{min}\|H_i\|^2} \right\rfloor = -\frac{m{H_u}^T H_i}{\|H_i\|^2} + \frac{N^T H_i}{d_{min}\|H_i\|^2}. \tag{15}$$

Comparing (9) with (15) we get $\eta_i^* = -m$ for $i = u$ and 0 for $i \neq u$ and $\delta_i = \frac{N^T H_i}{d_{min}\|H_i\|^2}$.

Let us generalize this for an initial vector that differs from the transmitted vector at $\mathcal{L}$ symbols and denote such indices by $u_1, u_2, \cdots, u_{\mathcal{L}}$. Following the same argument we get

$$\delta_i = \frac{1}{d_{min}} \sum_{i=1}^{\mathcal{L}} \frac{N^T H_{u_i}}{\|H_{u_i}\|^2} \tag{16}$$

Corresponding to this δ_i the second term in (10) can be expressed as

$$\delta_i^2 d_{min}^2 \|H_i\|^2 = \left(\sum_{i=1}^{\mathcal{L}} \frac{N^T H_{u_i}}{\|H_{u_i}\|^2} \right)^2 \|H_i\|^2 \tag{17}$$

In this equation, the term $\frac{N^T H_{u_i}}{\|H_{u_i}\|^2}$ can be seen as a sample average of the product of two independent and identically distributed, zero mean Gaussian random variables. For large values of N_r this is likely to be close to zero. As a result the right hand side of (17) is likely to be of a small magnitude, in particular, for small values of $\mathcal{L}$ which is likely to be the case. This is 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. In view of this, the second term in (10) is negligible compared to the first term and can be ignored.

VI. Conclusion

Existing neighborhood search algorithms, like LAS and RTS, look for the best vector in the entire neighborhood which leads to a high computational complexity. In this paper we 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 achieve a significant reduction in complexity. We have also evaluated the performance of the proposed algorithms for massive MIMO systems employing up to a thousand antennas.

References

[1] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 742–758, Oct 2014.

[2] M. Wu, B. Yin, G. Wang, C. Dick, J. Cavallaro, and C. Studer, "Large-Scale MIMO Detection for 3GPP LTE: Algorithms and FPGA Implementations," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 916–929, Oct 2014.

[3] K. Vardhan, S. Mohammed, A. Chockalingam, and B. Rajan, "A Low-Complexity Detector for Large MIMO Systems and Multicarrier CDMA Systems," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 3, pp. 473–485, April 2008.

[4] S. Mohammed, A. Chockalingam, and B. Sundar Rajan, "A Low-complexity near-ML performance achieving algorithm for large MIMO detection," in *Information Theory, ISIT 2008. IEEE International Symposium on*, July 2008, pp. 2012–2016.

[5] B. Rajan, S. Mohammed, A. Chockalingam, and N. Srinidhi, "Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive tabu search," in *Information Theory, ISIT 2009. IEEE International Symposium on*, June 2009, pp. 1993–1997.

[6] A. Chockalingam, "Low-complexity algorithms for large-MIMO detection," in *Communications, Control and Signal Processing, ISCCSP 2010. 4th International Symposium on*, March 2010, pp. 1–6.

[7] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-Restart Reactive Tabu Search Algorithm for Detection in Large-MIMO Systems," *Communications Letters, IEEE*, vol. 14, no. 12, pp. 1107–1109, Dec. 2010.

[8] N. Srinidhi, T. Datta, A. Chockalingam, and B. Rajan, "Layered Tabu Search Algorithm for Large-MIMO Detection and a Lower Bound on ML Performance," *Communications, IEEE Transactions on*, vol. 59, no. 11, pp. 2955–2963, November 2011.

[9] P. Li and R. Murch, "Multiple output selection-LAS algorithm in large MIMO systems," *Communications Letters, IEEE*, vol. 14, no. 5, pp. 399–401, May 2010.

[10] Q. Zhou and X. Ma, "Element-Based Lattice Reduction Algorithms for Large MIMO Detection," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 274–286, February 2013.

[11] P. Svac, F. Meyer, E. Riegler, and F. Hlawatsch, "Soft-Heuristic Detectors for Large MIMO Systems," *Signal Processing, IEEE Transactions on*, vol. 61, no. 18, pp. 4573–4586, Sept 2013.

[12] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?" *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 160–171, February 2013.