

A Fast Sphere Decoding Framework for Space-Frequency Block Codes

Zoltan Safar
Department of Innovation
IT University of Copenhagen
Copenhagen, Denmark
E-mail: safar@itu.dk

Weifeng Su, and K. J. Ray Liu
Department of Electrical and Computer Engineering
University of Maryland
College Park, MD, USA
E-mail: {weifeng, kjrlui}@eng.umd.edu

Abstract—In order to make multi-antenna OFDM systems an attractive choice for practical applications, implementation issues such as decoding complexity must be addressed successfully. In this paper, we propose a computationally efficient decoding algorithm for space-frequency block codes. The central part of the algorithm is a complex-domain, modulation-independent sphere decoding framework. The simulation results demonstrate that the proposed algorithm can significantly reduce the decoding complexity. For 64 QAM modulation, we observe up to about 73% reduction in the required FLOP count per code block compared to previously proposed methods without noticeable performance degradation.

I. INTRODUCTION

The recently proposed space-time (ST) and space-frequency (SF) coded multiple-input-multiple-output (MIMO) systems have promised considerable performance improvement over the single-antenna systems. However, the computational complexity of the maximum likelihood (ML) decoding algorithm may hamper the widespread use of such systems, so the development of low complexity ST and SF decoding algorithms is a problem of paramount importance.

Computationally efficient decoding algorithms have only been proposed for decoding ST block codes in quasi-static, flat fading environment [1], [2]. For ST block codes transmitted over temporally evolving channels and for SF block-coded MIMO-OFDM systems, where the channel changes along the frequency axis, low complexity decoding algorithms still do not exist in the literature.

The sphere decoding algorithm was introduced in [3] assuming a single-antenna, real-valued fading channel model. Later results [4], [5] generalized the algorithm to complex-valued MIMO channels. A reduced complexity algorithm was proposed in [6], where the signal coordinates were sorted according to their partial metrics and explored in this order. The authors of [7] achieved complexity reduction by exploring the signal coordinates in a zig-zag order. This approach was further refined and improved in [8]. In [9], the sphere decoding algorithm was applied to equalize frequency selective MIMO channels. All of these works considered uncoded MIMO systems and assumed quasi-static, flat fading channels. Moreover, they formulated the sphere decoding problem in the real domain, so the resulting algorithms can only be used with

modulation methods that can be decomposed into the product of two real constellations (for example, square QAM).

A complex-domain sphere decoding algorithm was described in [10]. This work considered iterative (turbo) decoding in a MIMO system where linear ST mapping was combined with an outer channel code, and a sphere detector was used to approximate the log-likelihood ratio in a computationally efficient way. This approach was specific to modulation methods that can be decomposed into PSK constellations, and its objective was to identify a set of candidate solutions, as opposed to finding the best candidate solution with the maximum possible efficiency.

In this work, we propose a general framework for decoding SF block codes constructed from orthogonal and quasi-orthogonal designs based on the idea of sphere decoding. We formulate the decoding problem in the complex domain, which allows us to exploit the distance structure of complex signal constellations. We propose a modulation-independent sphere decoding framework by interpreting the sphere decoding problem as a greedy, constrained depth-first search algorithm. Due to the modular structure of the framework, it can be used to construct a decoding algorithm that can be used with any memoryless modulation, and it can also be tailored to a particular modulation method.

II. SYSTEM MODEL AND NOTATION

Consider a SF-coded MIMO-OFDM system having K transmit antennas, L receive antennas and M sub-carriers, with M being a multiple of K . Suppose that the frequency selective fading channels between each pair of transmit and receive antennas have P independent delay paths and the same power delay profile. The MIMO channel is assumed to be constant over each OFDM block period. The channel impulse response from transmit antenna k to receive antenna l at time τ is modeled as $h_{k,l}(\tau) = \sum_{p=0}^{P-1} \beta_{k,l}(p)\delta(\tau - \tau_p)$, where τ_p is the delay and $\beta_{k,l}(p)$ is the complex amplitude of the p -th path between transmit antenna k and receive antenna l . The $\beta_{k,l}(p)$'s are modeled as zero-mean, complex Gaussian random variables with variances $E[|\beta_{k,l}(p)|^2] = \delta_p^2$. The powers of the P paths are normalized such that $\sum_{p=0}^{P-1} \delta_p^2 = 1$. The frequency response of the channel is given by $H_{k,l}(f) = \sum_{p=0}^{P-1} \beta_{k,l}(p)e^{-j2\pi f\tau_p}$. We assume that the MIMO channel

is spatially uncorrelated, i.e. the $\beta_{k,l}(p)$'s are independent for different indices (k, l) .

The input bit stream is divided into b bit long segments, creating B -ary ($B = 2^b$) source symbols. The encoder forms M/K source symbol blocks, each containing N source symbols. Source symbol $s_i \in \{0, 1, \dots, B-1\}$, $i = 0, 1, \dots, N-1$, is mapped onto a complex channel symbol (or constellation point) x_i according to $x_i = \Omega(s_i)$, where the function $\Omega(\cdot)$ represents the modulation operation. The average energy of the constellation will be denoted by E_{avg} .

Then, the SF encoder forms two-dimensional, square code-words from the channel symbols. The SF codeword corresponding to the t -th ($t = 0, 1, \dots, M/K - 1$) source symbol block can be expressed as a K by K matrix $\mathbf{C} = \{c_k[Kt+m]\}$ ($k, m = 0, 1, \dots, K-1$), where $c_k[i]$ denotes the channel symbol transmitted over the i -th sub-carrier by transmit antenna k . We assume that each $c_k[i]$ is either zero, or a channel symbol or a negative and/or complex conjugate of a channel symbol corresponding to a source symbol in the appropriate source symbol block.

At the receiver, after matched filtering, removing the cyclic prefix, and applying FFT, the received signal corresponding to the t -th source symbol block at sub-carrier $Kt + m$ ($m = 0, 1, \dots, K-1$) and receive antenna l is given by

$$y_l[Kt+m] = \sum_{k=0}^{K-1} H_{k,l}[Kt+m]c_k[Kt+m] + n_l[Kt+m], \quad (1)$$

where $H_{k,l}[i] = H_{k,l}(i\Delta f)$ is the channel frequency response at the i -th sub-carrier between transmit antenna k and receive antenna l , $\Delta f = 1/T$ is the sub-carrier separation in the frequency domain, and T is the OFDM symbol period. We assume that the channel state information $H_{k,l}[i]$ is known at the receiver, but not at the transmitter. In (1), $n_l[i]$ denotes the complex, zero-mean, additive white Gaussian noise component at the i -th sub-carrier at receive antenna l . The variance of the noise samples is assumed to be $1/(\rho\gamma)$, where the scaling factor γ is defined as $\gamma = b/(K^2 E_{avg})$, so ρ is the signal to noise ratio (SNR) per bit at each sub-carrier at each receive antenna. In the sequel, we will focus our attention on decoding a single code block, so a simplified notation will be used by dropping the block index t :

$$y_l[m] = \sum_{k=0}^{K-1} H_{k,l}[m]c_k[m] + n_l[m], \quad (2)$$

for $m = 0, 1, \dots, K-1$.

III. EQUIVALENT REPRESENTATION

In general, SF coding introduces spatial and frequency-domain dependence among the code symbols $c_k[i]$ within a code block \mathbf{C} . For example, in case of the 2×2 orthogonal design ($K = 2, N = 2$) [1]

$$\mathbf{C} = \begin{bmatrix} x_0 & x_1 \\ -x_1^* & x_0^* \end{bmatrix}, \quad (3)$$

the channel symbols transmitted from different transmit antennas and through different sub-carriers are clearly related.

However, to be able to use a sphere decoder (to be able to make sequential decisions on the sent signal coordinates), it is necessary to transform the received signal to an equivalent signal representation, where the coordinates of the sent signal vector are independent. Due to space limitations, we will only consider SF codes constructed from the 2×2 orthogonal design (3) in this paper, but our approach can be extended to more transmit antennas and quasi-orthogonal designs.

The transformed equivalent received signal vector $\mathbf{y}_l = [y_l[0], y_l^*[1]]^T$ for receive antenna l can be rewritten in matrix-vector form as

$$\mathbf{y}_l = \mathbf{H}_l \mathbf{x} + \mathbf{n}_l,$$

where $\mathbf{x} = [x_0, x_1]^T$ is the $N \times 1$ transmitted channel symbol vector, $\mathbf{n}_l = [n_l[0], n_l^*[1]]^T$ is the equivalent noise component, and \mathbf{H}_l is defined as

$$\mathbf{H}_l = \begin{bmatrix} H_{0,l}[0] & H_{1,l}[0] \\ H_{1,l}^*[1] & -H_{0,l}^*[1] \end{bmatrix}.$$

By collecting the received signal and noise components corresponding to different receive antennas in $KL \times 1$ vectors as $\mathbf{y} = [\mathbf{y}_0^T, \dots, \mathbf{y}_{L-1}^T]^T$, and $\mathbf{n} = [\mathbf{n}_0^T, \dots, \mathbf{n}_{L-1}^T]^T$, the equivalent received signal can be expressed as

$$\mathbf{y} = \mathbf{H} \mathbf{x} + \mathbf{n}, \quad (4)$$

where the $KL \times N$ matrix \mathbf{H} is the equivalent channel matrix, defined as $\mathbf{H} = [\mathbf{H}_0^T, \dots, \mathbf{H}_{L-1}^T]^T$. Note that the above described equivalent representation has the following properties that are important from the viewpoint of the sphere decoding algorithm. First, the coordinates of the noise vector \mathbf{n} are independent, zero mean, complex Gaussian random variables with variance $1/(\rho\gamma)$. Second, the coordinates of the \mathbf{x} vector are independent. Third, the matrix \mathbf{H} has at least as many rows as columns, independently of the number of receive antennas. Fourth, since the entries in the matrix \mathbf{H} are complex, zero mean, Gaussian random variables and we have assumed that the MIMO channel is spatially independent, the matrix \mathbf{H} has full (column) rank with high probability.

IV. THE PROPOSED ALGORITHM

For systems described by (4), to decode the sent signal vector \mathbf{x} with the maximum likelihood (ML) algorithm, the task is to find a valid signal vector \mathbf{x} that minimizes the metric $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$. Unfortunately, in some cases this can only be performed by exhaustive search over all valid signal vectors. To alleviate this computational burden, sphere decoding can be used, where the decoder searches over only a subset of \mathbf{x} vectors that lie within a hyper-sphere of radius r centered around the received signal vector, i.e. $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \leq r^2$. We have divided the proposed algorithm in two parts: the preprocessing stage and the searching stage.

A. Preprocessing Stage

The purpose of this stage is to transform the expression $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$ in such a form that the decisions on the coordinates of \mathbf{x} can be made sequentially. First, the complex Cholesky factorization of the matrix $\mathbf{H}^H \mathbf{H}$ is calculated, obtaining an

$N \times N$ complex, upper triangular matrix $\mathbf{R} = \{R_{k,l}\}$ with positive and real diagonal elements. We used a simple extension of the ‘‘Gaxpy’’ version of the real Cholesky decomposition algorithm [11]. Then, the elements of the $N \times N$ matrix $\mathbf{Q} = \{Q_{k,l}\}$ are calculated as $Q_{k,k} = R_{k,k}^2$ and $Q_{k,l} = R_{k,l}/R_{k,k}$ for $k < l$. After that, the zero forcing solution $\mathbf{z} = \mathbf{H}^+\mathbf{y}$ is determined, where \mathbf{H}^+ denotes the pseudo-inverse of \mathbf{H} . This can be done most efficiently by solving the lower triangular system $\mathbf{R}^H\mathbf{w} = \mathbf{H}^H\mathbf{y}$ for \mathbf{w} and solving the upper triangular system $\mathbf{R}\mathbf{z} = \mathbf{w}$ for \mathbf{z} . The last step is to obtain the modified radius $r'^2 = r^2 - \|\mathbf{y}\|^2 + \|\mathbf{H}\mathbf{z}\|^2$. Using these quantities, the sphere decoding problem can be expressed in the following way: find only those \mathbf{x} signal vectors that satisfy $\|\mathbf{R}(\mathbf{z} - \mathbf{x})\|^2 \leq r'^2$.

B. Searching Stage

The searching stage generates the sent signal vectors \mathbf{x} satisfying the sphere constraint and selects the decoded signal vector. To be able to make the decisions on the signal vector \mathbf{x} coordinate by coordinate, the sphere constraint is expressed in a recursive manner by defining the quantities α_k and T_k as follows.

$$\alpha_k = z_k + \sum_{l=k+1}^{N-1} Q_{k,l}(z_l - x_l) = z_k + \sum_{l=k+1}^{N-1} Q_{k,l}\Delta_l$$

for $k = N-1, N-2, \dots, 0$, where x_k and z_k denote the k -th coordinates of \mathbf{x} and \mathbf{z} , respectively, and $\Delta_k = z_k - x_k$ is defined for computational convenience. $T_{N-1} = r'^2$, and

$$T_k = T_{k+1} - Q_{k+1,k+1}|\alpha_{k+1} - x_{k+1}|^2$$

for $k = N-2, N-3, \dots, 0$. The quantity T_k can be thought of as the remaining squared distance between the partial solution $x_{N-1}, x_{N-2}, \dots, x_{k+1}$ and the surface of the sphere, and α_k can be interpreted as the k -th received signal component given $x_{N-1}, x_{N-2}, \dots, x_{k+1}$. The sphere decoding problem can be solved by going over the coordinates of \mathbf{x} starting from $k = N-1$ and continuing down to $k = 0$. For each k value, we need to calculate α_k and T_k , and determine those valid x_k values that satisfy

$$|\alpha_k - x_k|^2 \leq T_k/Q_{k,k}. \quad (5)$$

Then, we go over each x_k value, decrease the value of k by one, and repeat the procedure. If the $k = 0$ level is completed, we have found an \mathbf{x} vector inside the sphere.

The original sphere decoding algorithm [3] was based on an algorithm developed for finding the shortest vector in a lattice. Our objective is to devise a decoding framework for any memoryless modulation method, so the signal coordinates x_k may take on any complex value, and the set of vectors $\{\mathbf{H}\mathbf{x}\}$ does not form a lattice. To extend the sphere decoding to arbitrary constellations, we provide an alternative interpretation. We abandon the lattice concept and look at the sphere decoding problem from a different angle. For each k value, all x_k values satisfying the partial constraint (5) are enumerated and explored, so the search space can be represented by a

tree. The sphere constraint limits the number of branches emanating from each node of the tree. Finally, as suggested in [6], the valid x_k values are sorted according to the metric $d_k = |\alpha_k - x_k|^2$ in increasing order, and they are explored in this order, which corresponds to a ‘‘greedy’’ strategy. As a consequence, the proposed sphere decoding framework can be interpreted as a greedy, constrained depth-first tree search algorithm.

The flowchart of the proposed algorithm can be observed in Figure 1. The main differences between this algorithm and the one described in [6] are that we formulated the decoding problem in the complex domain and not in the real domain, and that our algorithm can be used with any constellation, not only with square QAM.

The variable d indicates whether a node is reached from above (‘‘DOWN’’), i.e. it is visited for the first time, or from below (‘‘UP’’), i.e. it has been visited before. The variable n_c counts the number of valid candidate solutions that have been generated. The index k is checked, and if it is non-negative, the bottom level of the tree has not been reached yet. If the current tree node is visited for the first time, the value of α_k is calculated, and the function SYMLIST generates the list of possible source symbols $s_{k,i}$, whose corresponding channel symbols $x_{k,i} = \Omega(s_{k,i})$ satisfy the normalized partial constraint $|\alpha_k - x_{k,i}|^2 \leq T_k/Q_{k,k}$. The SYMLIST function takes α_k and the normalized partial constraint $T_k/Q_{k,k}$ as inputs and produces 3 outputs. The first output, n_k ($0 \leq n_k \leq B$), is the number of symbols satisfying the current partial constraint, so there will be n_k branches emanating from the current node. The second output is the symbol list $\mathbf{s}_k = [s_{k,0}, s_{k,1}, \dots, s_{k,n_k-1}]^T$, and the vector $\mathbf{d}_k = [d_{k,0}, d_{k,1}, \dots, d_{k,n_k-1}]^T$ whose coordinates are the metrics $d_{k,i} = |\alpha_k - x_{k,i}|^2$. The symbols in \mathbf{s}_k are ordered according to increasing $d_{k,i}$ values. The symbol pointer p_k points to the current element of \mathbf{s}_k . The next source symbol is taken from the list, the new partial constraint value is calculated, and the algorithm moves one level down the tree.

If the value of k becomes negative, we have reached the bottom level, so a valid candidate solution \mathbf{x} has been found. At this point, the radius of the sphere is reduced to further decrease the decoding complexity by adjusting all partial constraint values such that the last solution satisfies the constraints with equality (the ‘‘surplus’’ partial constraint T_{-1} is subtracted from each partial constraint). The last \mathbf{x} vector is the best solution so far, so the corresponding source symbols are saved in the $\{s_i^{best}\}$ variables by overwriting the previous solution. The source symbol lists are also modified by the UPDATE function, which keeps only those source symbols on the list whose corresponding d_k metric values satisfy the new partial constraints. Since the symbols are ordered according to the corresponding d_k values, this can be done simply by changing the value of n_k at each level.

V. SIMULATION RESULTS

To illustrate the performance of the proposed sphere decoding algorithm, we provide some simulation results. The

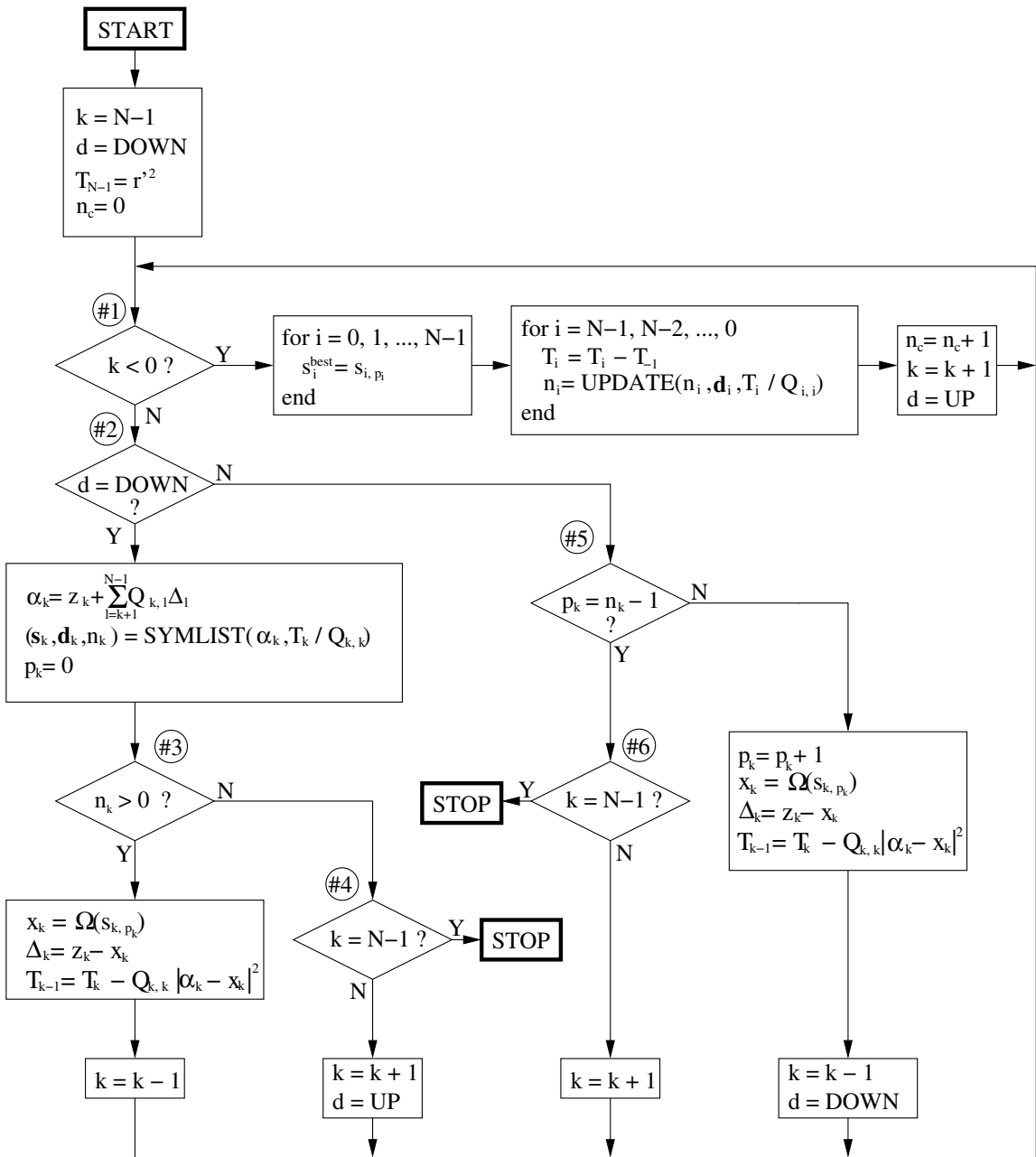


Fig. 1. The flowchart of the decoding algorithm

simulated communication system had 2 transmit antennas ($K = 2$), 2 receive antennas ($L = 2$), and the 2×2 orthogonal design (3) with 64 QAM modulation was used as the SF code ($N = 2$). The OFDM modulation had 128 sub-carriers with an OFDM symbol period of $128 \mu s$. The frequency selective MIMO channel was modeled by the COST 207 Typical Urban 6-ray power delay profile. The initial radius was set to $r = 10$ to ensure ML performance.

We compared the total decoding complexity, including the preprocessing and searching stages, of five different decoding algorithms: the ML decoding algorithm (performing exhaustive search), the real-domain sphere decoding algorithm described in [6] (preprocessing stage: real Cholesky

decomposition), the real-domain sphere decoding algorithm described in [8] (preprocessing stage: real QR decomposition), and the proposed decoding framework (preprocessing stage: complex Cholesky decomposition) with two symbol list generating method: the modulation-independent search and the fast QAM-specific search, corresponding to two different implementations of the SYMLIST function. The first method, the modulation-independent approach, assumes that there is no apriory information available on the used constellation, so it works with any memoryless modulation. The second method, the fast QAM-specific search approach, performs a nearest neighbor signal point search around the α_k values. The details of these algorithms can be found in [12].

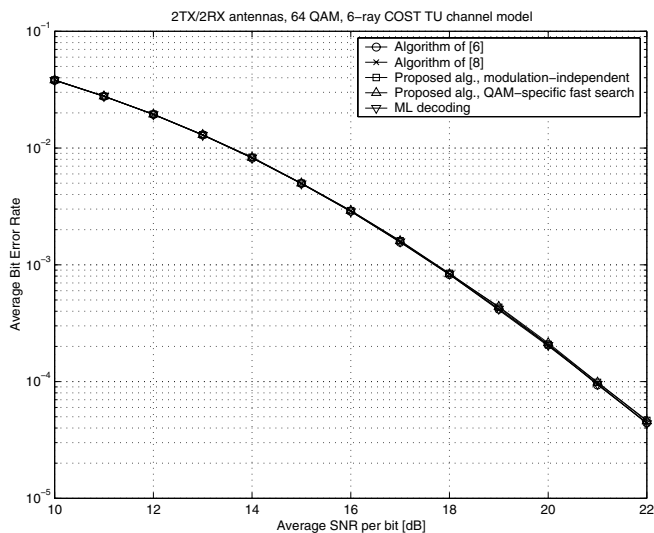


Fig. 2. Average bit error rate

Figure 2 depicts the bit error rate (BER) curves as functions of the average SNR. The total number of FLOPs per decoded code block, including both stages, as a function of the SNR is depicted in Figure 3. The FLOP count for the ML decoding algorithm was 331776. Based on the above figures, we can make several observations. First, using sphere decoding, the computational complexity of the SF decoder can be reduced by orders of magnitude without perceptible performance loss in the meaningful BER range. Second, the complexity of the decoding algorithms do not change considerably as the SNR increases. The reason is that the initial radius was kept constant, and most of the time, the first solution found by the greedy tree search was actually the ML solution. Finally, the proposed QAM-specific fast search algorithm is the computationally most efficient: the total number of FLOPs per code block was reduced to about 64% of the FLOP count of the algorithm in [6], and to about 27% of the FLOP count of the algorithm in [8]. The performance gain compared to [6] is due to the higher efficiency of the complex-domain processing, and the avoidance of the unnecessary enumeration and sorting operations. The higher complexity of [8] is mainly the consequence of the complexity of the real QR decomposition of an 8×4 matrix (as opposed to the complex Cholesky decomposition of a 2×2 matrix).

We also performed simulations with 16 QAM modulation. The observed tendencies were similar, and the total FLOP count per code block was reduced by the proposed algorithm to about 71% and 26% of that of the methods in [6] and [8], respectively.

VI. CONCLUSION

We proposed a sphere decoding approach for decoding SF block codes. We formulated the sphere decoding problem in the complex domain and developed a modulation independent decoding framework by interpreting the decoding problem as a greedy, constrained depth-first tree search algorithm. The simulation results demonstrated significant complexity

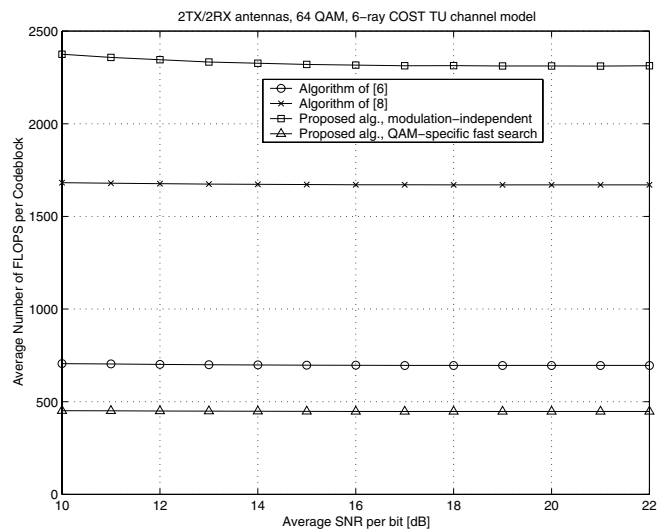


Fig. 3. Average number of FLOPs per code block

reduction compared to previously existing algorithms. For problems of larger size (larger K , L and B values), even more pronounced complexity reduction is expected.

REFERENCES

- [1] S. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE JSAC*, vol. 16, no. 8, pp. 1451-1458, 1998.
- [2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, 1999.
- [3] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels", *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1639-1642, 1997.
- [4] M. Damen, A. Chkeif and J. Belfiore, "Lattice code decoder for space-time codes", *IEEE Communication Letters*, vol. 4, no. 5, pp. 161-163, 2000.
- [5] M. Damen, K. Abed-Meraim and J. Belfiore, "Generalized lattice decoder for asymmetrical space-time communication architecture", *Proc. of IEEE ICASSP*, vol. 5, pp. 2581-2584, 2000.
- [6] A. Chan and I. Lee, "A new reduced complexity sphere decoder for multiple antenna systems", *Proc. of IEEE ICC*, vol. 1, pp. 460-464, 2002.
- [7] E. Agrell, T. Eriksson, A. Vardy and K. Zeger, "Closest point search in lattices", *IEEE Transactions on Information Theory*, vol. 48, no. 8, pp. 2201-2214, August 2002.
- [8] M. Damen, H. El Gamal and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point", *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2389-2402, October 2003.
- [9] H. Vikalo and B. Hassibi, "Maximum likelihood sequence detection of multiple antenna systems over dispersive channels via sphere decoding", *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 5, pp. 525-531, 2002.
- [10] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel", *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389-399, March 2003.
- [11] G. Golub and C. Van Loan, *Matrix computations*, Johns Hopkins University Press, 1996.
- [12] Z. Safar, W. Su and K. J. R. Liu, "Fast sphere decoding of space-frequency block codes via nearest neighbor signal point search", *European Wireless Conference*, February 2004.