EFFICIENT ML DETECTION FOR MIMO CHANNELS: ORDERED SPHERE DECODING

Karen Su and Ian J. Wassell, Cambridge University Engineering Department

Key words to describe the work: Sphere decoding, Maximum Likelihood (ML) detection, nearest lattice point search, Multiple Input Multiple Output (MIMO) channels, pre-processing, ordering.

Key results: The Ordered Sphere Decoder (OSD) performs ML detection for MIMO channels at a greatly reduced computational cost, while also providing a competitive sub-optimal solution in fixed time.

How does the work advance the state-of-the-art?: The OSD is evaluated against state-of-the-art works, demonstrating a dramatic complexity reduction, particularly in the important higher modulation order regimes.

Motivation (problems addressed): The OSD is a new ML MIMO channel decoding proposal that addresses the issues of decoding complexity and delay over a wide range of operating regions.

INTRODUCTION

The ML detection of signals transmitted over MIMO channels is an important communications problem that is well-known to be NP-complete. A current state-of-the-art solution is the Sphere Decoder (SD) [3, 7], whose computational cost is polynomial in the average case [4]. So bright are its prospects for industrial application that VLSI implementations have already been reported in the literature (e.g., [1]). Even so, existing SDs exhibit two major weaknesses: Their complexity coefficients can become large when the problem dimension is high, e.g., at the spectral efficiencies demanded by future services, and the variance of their computation times can also be large, leading to undesirable highly variable decoding delays. In this paper, we present the OSD, a solution that successfully tackles these critical challenges.

SPHERE DECODING

A MIMO channel detector produces a set of symbols $\mathbf{s} \in \mathcal{X}^M$ given a set of signals $\mathbf{v} \in \mathbb{R}^N$ observed at the output of the communication channel, typically modelled as a linear system $\mathbf{H} \in \mathbb{R}^{N \times M}$ combined with an additive noise vector $\mathbf{n} \in \mathbb{R}^N$.[1] We assume that $M \leq N$ and that $\mathbf{H}$ is of full rank M, i.e., there are at least as many observations as symbols to be detected. Since the transmitted symbols are drawn from a known finite alphabet $\mathcal{X}$ of size B, the detector's goal is to choose one of the B^M possible transmitted symbol vectors based on the data.

The SD is an ML detector, i.e., it returns a solution

$$\mathbf{s}_* = \underset{\mathbf{s} \in \mathcal{X}^M}{\operatorname{argmax}}\, P(\mathbf{v} \text{ is observed} \,|\, \mathbf{s} \text{ was sent}) \tag{1}$$

$$= \underset{\mathbf{s} \in \mathcal{X}^M}{\operatorname{argmin}}\, |\mathbf{v} - \mathbf{H}\mathbf{s}|^2, \tag{2}$$

where we make the further assumption that the additive noise $\mathbf{n}$ is white and Gaussian.

Solving (2) via brute force requires computing the distances from $\mathbf{v}$ to an exponential number of points. The SD on the other hand, is based on efficiently enumerating only those points located within some distance of the observed $\mathbf{v}$. Assuming that at least one such point is found, it then follows that the nearest point to $\mathbf{v}$ in the exponential search set must also be the nearest one in the smaller set. Thus an optimal solution can be declared after computing, on average, only a polynomial number of distances.

Underlying all known SDs is a weighted $(M+1)$-level B-ary tree, as depicted in Fig. 1 for the case where $M = B = 2$ and $\mathcal{X} = \{-1, 1\}$.

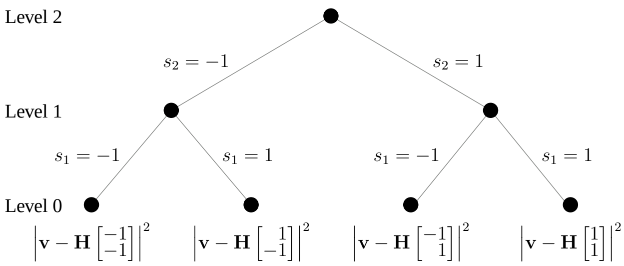

Figure 1: Weighted 3-level binary tree, $\mathcal{X} = \{-1, 1\}$.

Its nodes are numbered from the *root* at level M to the B^M *leaves* at level 0. Traversing a branch from level i to $i-1$ assigns a value to symbol s_i. Thus descending from the root to a leaf assigns values to all M symbols. To each branch can be assigned a non-negative weight, and to each node a weight equal to the sum of the branches along its path from the root, such that the leaf node weights are precisely the squared distances in (2). Within this context, an optimal solution is given by the symbol vector assigned to *the smallest weight leaf node* in the tree of Fig. 1.

THE ORDERED SPHERE DECODER

A block diagram of the proposed decoder is shown in Fig. 2. It is formed by cascading an ordering module with an Automatic Sphere Decoder (ASD) [5].

The ordering module takes as its inputs the vector of observed signals $\mathbf{v}$ and the channel matrix $\mathbf{H}$. It is a fixed time algorithm with a polynomial complex-

[1] In this work, all signals and coefficients are real numbers; we recall that complex signal detection can be written as an equivalent problem in twice the number of real dimensions.

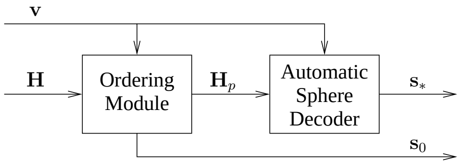

Figure 2: Block diagram of the OSD.

ity that is roughly comparable to a handful of matrix inversions or QR factorizations. It produces at its output an ordered channel matrix $\mathbf{H}_p$ and a sub-optimal initial guess for the detected symbol vector $\mathbf{s}_0$. The construction of $\mathbf{H}_p$ is based on a heuristic rule designed to reduce the number of nodes (at all levels) whose weights are less than that of the smallest weight leaf node. Please see [6] for more details.

The ASD is a new sphere decoding algorithm that takes as its inputs the observed vector $\mathbf{v}$ and a channel matrix $\mathbf{H}$. Its complexity is characterized by the product of the number of nodes expanded ν and the per-node processing time τ. We show in [5] that all known SDs expand at least as many nodes as the ASD, and also that for all known SDs, including the ASD, τ is linear in M. Therefore we propose to compare the computational performance of different SDs by considering their respective average values of ν.

SIMULATION RESULTS

Fig. 3 shows the average number of nodes expanded by the OSD and by a radius-adaptive SD based on the Schnorr-Euchner enumeration (SEA) [2,3] over a range of Signal-to-Noise Ratio (SNR)s. As expected, $\nu_{OSD} \leq \nu_{SEA}$ regardless of the SNR. Also, for all of the decoders, there is a lower bound of $\nu \geq 2M$, where M is the number of transmit antennas; the factor of two arises because there are two real dimensions per complex dimension.

A vast improvement is realized by the OSD, especially at low SNRs, where the complexity of existing sphere decoders is not widely considered to be competitive. Even more noteworthy is that this improvement remains significant even as the modulation order is increased, and that the computational profile of the OSD is consistently almost an order of magnitude more efficient than current proposals.

CONCLUSIONS

In this paper we present a new ML MIMO decoding proposal that is both effective in addressing critical weaknesses of current proposals and timely, as the sphere decoder vies for a central role in current 4G standardization efforts. Simulation results demonstrate that the OSD offers an immense reduction in the computational cost of sphere decoding. We also

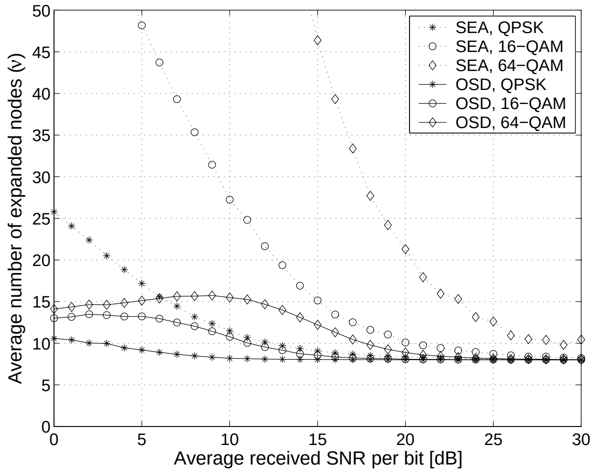

Figure 3: Average number of expanded nodes (ν) vs. average received SNR per bit for the SEA and OSD decoders over a 4:4 MIMO flat fading channel using QPSK, 16-QAM and 64-QAM modulations.

highlight its effectiveness in the higher modulation order regimes that have traditionally been prohibitively complex for SDs. Finally, an advantageous by-product of the OSD is the competitive sub-optimal solution $\mathbf{s}_0$, which is computed in fixed time and can be returned early to delay-sensitive applications that are unable to wait for the ML solution. The bit error rate of $\mathbf{s}_0$ taken as a sub-optimal solution is comparable to that of the well-known V-BLAST system with MMSE nulling, cancellation and ordering.

REFERENCES

[1] A. Burg, M. Borgmann, C. Simon, M. Wenk, M. Zellweger, and W. Fichtner. Performance tradeoffs in the VLSI implementation of the sphere decoding algorithm. In *IEE 3G Mobile Communication Technologies Conference*, October 2004.

[2] A. M. Chan and I. Lee. A new reduced-complexity sphere decoder for multiple antenna systems. In *IEEE International Conference on Communications*, volume 1, pages 460–464, April 2002.

[3] M. O. Damen, H. E. Gamal, and G. Caire. On maximum-likelihood detection and the search for the closest lattice point. *IEEE Transactions on Information Theory*, 49(10):2389–2402, October 2003.

[4] B. Hassibi and H. Vikalo. On the sphere decoding algorithm: Part I, The expected complexity. *To appear in IEEE Transactions on Signal Processing*, 2004.

[5] K. Su, C. N. Jones, and I. J. Wassell. An automatic sphere decoder. *Submitted to IEEE Transactions on Information Theory*, 2004.

[6] K. Su and I. J. Wassell. An enhanced ordering for efficient sphere decoding. In *IEEE International Conference on Communications (to appear)*, May 2005.

[7] E. Viterbo and J. Boutros. A universal lattice code decoder for fading channels. *IEEE Transactions on Information Theory*, 45(5):1639–1642, July 1999.