A Hybrid Probabilistic Data Association-Sphere Decoding Detector for Multiple-Input–Multiple-Output Systems

Georgios Latsoudas and Nicholas D. Sidiropoulos, *Senior Member, IEEE*

***Abstract*—A hybrid probabilistic data association (PDA)-sphere decoding (SD) algorithm is proposed for signal detection in multiple-input–multiple-output (MIMO) systems. The key idea is to reduce the dimension of the problem solved via SD by first running a single stage of the PDA to fix symbols that can be decoded with high reliability. Simulations under a multiple antenna Rayleigh fading scenario show that this two-step algorithm attains a considerably better performance-complexity tradeoff than SD and PDA for low to moderate signal-to-noise ratio (SNR) or higher problem dimensions.**

***Index Terms*—Integer least squares, multiple-input–multiple-output (MIMO) decoding, probabilistic data association (PDA), sphere decoder.**

I. Introduction

MULTIPLE antenna systems have attracted great interest in recent years, due to the capacity improvement that they afford. Vertical Bell Laboratories Layered Space Time (V-BLAST) [3] is a widely known multiple antenna spatial multiplexing system targeting high spectral efficiencies. Unfortunately, the associated maximum-likelihood (ML) detector amounts to a constrained integer least-squares problem, whose exact solution entails exhaustive search. Thus, following the so-called *nulling and cancelling* detector [3], several computationally efficient detection algorithms have been developed for or adapted to V-BLAST.

Sphere Decoding (SD) [11], Probabilistic Data Association (PDA) [8], [10], and Semi-Definite Relaxation (SDR) [9] are three multiple-input–multiple-output (MIMO) detectors that can provide near-optimal performance at relatively low complexity in certain scenarios. Among them, SD appears to be prevalent in the recent literature. Numerous variants and improvements of SD have recently been developed, e.g., [1], [2], [12], and [13], incorporating more sophisticated schemes for increasing the associated search radius and organizing the computations in a more efficient manner, e.g., the Schnorr–Euchner (SE) SD, which uses an improved search strategy [1], [2]. A drawback of the SD family of detectors is that, for close-to-ML performance, complexity remains high in the low signal-to-noise ratio (SNR) regime or when the number of symbols to be jointly detected is large [5], [6].

The PDA is a simpler detection method, which, however, generally provides worse performance than SD. SD, PDA, SDR, and several other algorithms have recently been compared in the context of code division multiple access (CDMA) multiuser detection [4]. A corresponding comparison for the multiple antenna Rayleigh fading scenario (as in V-BLAST) has not been undertaken, to the best of our knowledge. Thorough comparisons are nontrivial, because complexity and performance of SD and SDR depend on a number of parameters. Our experience in [7] indicates that SDR is inferior to SD at high SNR.

In this letter, we propose a hybrid PDA-SD algorithm that attains a better performance-complexity tradeoff than either of its constituent components. At each stage of the decoding process, the PDA produces a set of soft decision metrics that can be used to assess how reliable associated hard decisions would be at that point. The basic idea, then, is to execute *a single stage* of the PDA algorithm and fix those symbols that can be detected with high reliability. After cancelling the effect of those symbols, a reduced-dimensionality problem is passed to SD for decoding. This reduces the complexity of SD and improves the performance of PDA. Our simulations show that the proposed algorithm enjoys an error performance close to that of SD over a wide range of SNR, at a significantly reduced computational cost.

We use the SD algorithm in Viterbo–Boutros (VB-SD) [11], with an initial radius chosen according to [5], and the SE-SD in [1] and [2], with a search radius set to infinity. Note, however, that the initial PDA stage can also be combined with other variants of SD or SDR. The key here is that *dimensionality reduction* via single-stage PDA preprocessing can provide significant computational relief at a small performance cost.

II. System Model

The aforementioned techniques are applicable to a broad range of MIMO communication systems. Herein, we focus on V-BLAST for concreteness. V-BLAST is a symbol synchronized multiple antenna system with n_T transmit and n_R receive antennas, with $n_T \leq n_R$. The input stream of bits is mapped to a particular constellation, and the resulting symbol stream is demultiplexed into n_T substreams. The transmissions are organized into bursts of L symbol periods. It is assumed that

Manuscript received July 28, 2004; revised October 28, 2004. This work was supported in part by the U.S. ARO under ERO Contract N62558-03-C-0012 and by the EU under U-BROAD STREP 506790S. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yiteng (Arden) Huang.

The authors are with the Department of Electronic and Computer Engineering, Technical University of Crete, University Campus-Kounoupidiana, 731 00 Chania-Crete, Greece (e-mail: latsoud@telecom.tuc.gr; nikos@telecom.tuc.gr).

Digital Object Identifier 10.1109/LSP.2005.843779

the channel is frequency flat and block fading (i.e., its variation is negligible over the L symbol periods comprising a burst and random from one burst to the next). The channel is assumed to be known to the receiver but not to the transmitter. From the discrete-time baseband-equivalent viewpoint, the system can be represented as

$$\tilde{\mathbf{r}} = \sqrt{\frac{\rho}{n_T}}\tilde{\mathbf{A}}\tilde{\mathbf{s}} + \tilde{\mathbf{n}} = \tilde{\mathbf{H}}\tilde{\mathbf{s}} + \tilde{\mathbf{n}} \quad (1)$$

where $\tilde{\mathbf{r}} = [\tilde{r}_1, \tilde{r}_2, \ldots, \tilde{r}_{n_R}]^T, \tilde{\mathbf{s}} = [\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_{n_T}]^T$ are the receive and the transmit vector, respectively, $\tilde{\mathbf{A}}$ is a generally complex $n_R \times n_T$ channel matrix with entries $\tilde{a}_{ij}$, and $\tilde{\mathbf{n}}$ is a white Gaussian circularly symmetric $n_R \times 1$ noise vector with covariance matrix $2\sigma^2\mathbf{I}$. The normalized amplitude $\sqrt{(\rho/n_T)}$ ensures that the SNR is constant for a given noise variance, irrespective of n_T. Assuming rich scattering, the elements of $\tilde{\mathbf{A}}$ are modeled as independent and identically distributed (i.i.d.) circularly symmetric Gaussian variables with zero mean and unit variance of the real and imaginary parts. For simplicity, we assume that the transmitted symbols are taken from a 4-QAM constellation, but the ideas generalize to higher order constellations. In order to transform the above model to a real-valued one, define

$$\mathbf{s} := [\Re(\tilde{\mathbf{s}}^T) \quad \Im(\tilde{\mathbf{s}}^T)]^T \quad (2)$$

$$\mathbf{r} := [\Re\{\tilde{\mathbf{r}}^T\} \quad \Im\{\tilde{\mathbf{r}}^T\}]^T \quad (3)$$

$$\mathbf{A} := \begin{bmatrix} \Re\{\tilde{\mathbf{A}}\} & -\Im\{\tilde{\mathbf{A}}\} \\ \Im\{\tilde{\mathbf{A}}\} & \Re\{\tilde{\mathbf{A}}\} \end{bmatrix} \quad (4)$$

$$\mathbf{n} := [\Re\{\tilde{\mathbf{n}}^T\} \quad \Im\{\tilde{\mathbf{n}}^T\}]^T \quad (5)$$

where $\Re, \Im$ denote the real and the imaginary part, respectively. Using the above vectors and matrices, we obtain the real-valued vector equation

$$\mathbf{r} = \sqrt{\frac{\rho}{n_T}}\mathbf{A}\mathbf{s} + \mathbf{n} = \mathbf{H}\mathbf{s} + \mathbf{n}. \quad (6)$$

III. Hybrid Algorithm

The hybrid algorithm consists of the following steps. As in [10], we premultiply (6) with $\mathbf{H}^T$, which yields

$$\mathbf{z} = \mathbf{H}^T\mathbf{r} = \mathbf{G}\mathbf{s} + \mathbf{v} \quad (7)$$

where $\mathbf{G} := \mathbf{H}^T\mathbf{H}$ is a symmetric positive definite[1] matrix, and $\mathbf{v} = \mathbf{H}^T\mathbf{n}$ is a noise vector with covariance matrix $\sigma^2\mathbf{G}$. We then apply one stage of the PDA detector (steps 1–5 in [8]) to the system in (7) and, thus, obtain a vector $\mathbf{p}$ that contains the associated probabilities for the elements of $\mathbf{s}$. Let D denote the subset of bits that satisfy

$$\mathbf{p}(i) \in [0, \tau] \cup [1 - \tau, 1] \quad (8)$$

with τ to be suitably chosen. $\bar{D}$ will henceforth denote the complement of D. We then make hard decisions for the bits in D, that is, set $\hat{s}_i = \text{sign}(\mathbf{p}(i) - 0.5), \forall i \in D$ and collect these decisions in a vector $\hat{\mathbf{s}}_D$. Now, expand (6) as

$$\mathbf{r} = [\mathbf{H}_D \quad \mathbf{H}_{\bar{D}}] \begin{bmatrix} \mathbf{s}_D \\ \mathbf{s}_{\bar{D}} \end{bmatrix} + \mathbf{n}$$

with obvious notation. Assuming perfect decisions for the bits in D (that is, $\hat{\mathbf{s}}_D = \mathbf{s}_D$), the residual subsystem after cancellation is

$$\mathbf{y}_{\bar{D}} := \mathbf{r} - \mathbf{H}_D\mathbf{s}_D = \mathbf{H}_{\bar{D}}\mathbf{s}_{\bar{D}} + \mathbf{n}.$$

After compacting

$$\mathbf{y}_c := \mathbf{H}_{\bar{D}}^T\mathbf{y}_{\bar{D}} = \mathbf{H}_{\bar{D}}^T\mathbf{H}_{\bar{D}}\mathbf{s}_{\bar{D}} + \mathbf{H}_{\bar{D}}^T\mathbf{n} = \mathbf{G}_{\overline{DD}}\mathbf{s}_{\bar{D}} + \mathbf{v}_{\bar{D}}$$

the noise vector $\mathbf{v}_{\bar{D}}$ is colored Gaussian with zero mean and covariance matrix $\sigma^2\mathbf{G}_{\overline{DD}}$. Introduce the Cholesky factorization

$$\mathbf{G}_{\overline{DD}} = \mathbf{L}_{\overline{DD}}^T\mathbf{L}_{\overline{DD}} \quad (9)$$

and premultiply the system with $\mathbf{L}_{\overline{DD}}^{-T}$ to obtain

$$\mathbf{x} := \mathbf{L}_{\overline{DD}}^{-T}\mathbf{y}_\mathbf{c} = \mathbf{L}_{\overline{DD}}\mathbf{s}_{\bar{D}} + \mathbf{w} \quad (10)$$

where the noise vector $\mathbf{w}$ is white Gaussian with covariance matrix $\sigma^2\mathbf{I}$. We now apply SD to (10). Let K be the number of elements in $\bar{D}$. As suggested in [5], the initial radius for VB-SD is set to $C = aK\sigma^2$, with a such that

$$\int_0^{aK/2} \frac{x^{(K/2-1)}}{\Gamma(K/2)} e^{-x} dx = 0.99. \quad (11)$$

Alternatively, SE-SD can be used in the second stage of the hybrid algorithm. We try both VB-SD and SE-SD in our simulations.

Threshold Parameter

The threshold parameter τ should be small enough to ensure that the PDA stage makes reliable decisions. On the other hand, τ should not be too small, for otherwise, the inclusion of the PDA stage will yield little if any dimensionality reduction benefit.

While it is clear that τ should be made smaller with increasing SNR, choosing it based on analytical considerations appears intractable. Our experience is that the following choice is reasonable: $\tau = 10^{-p}$ [hard-limited within $(0, 0.45]$], with $p := 3.5((8\sigma^2)/(\rho))^{-1.55}$. This setting is well supported by our simulation results, which are reported next.

IV. Simulation Results

In our simulations, each burst comprises $L = 100$ symbol intervals. Over each symbol interval, n_T 4-QAM symbols $(\pm(1/\sqrt{2}) \pm \jmath(1/\sqrt{2}))$ are simultaneously transmitted. For each burst, a new realization of the Rayleigh channel matrix is generated. For the bit-error rate (BER) plots, we use a dynamic Monte Carlo simulation: For each SNR, the simulation stops when both the number of errors has reached 150, and the

[1]With probability 1, under the i.i.d. Rayleigh assumption.

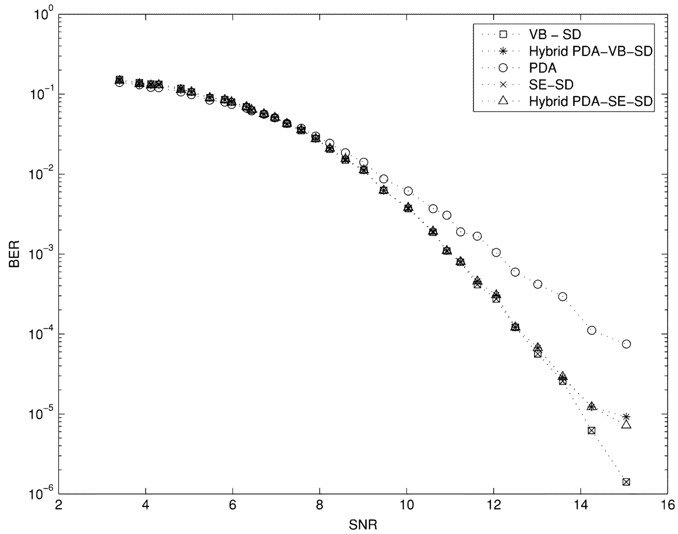

Fig. 1. Probability of error comparison for 4-QAM with $n_T = n_R = 16$. Dynamic Monte Carlo simulation.

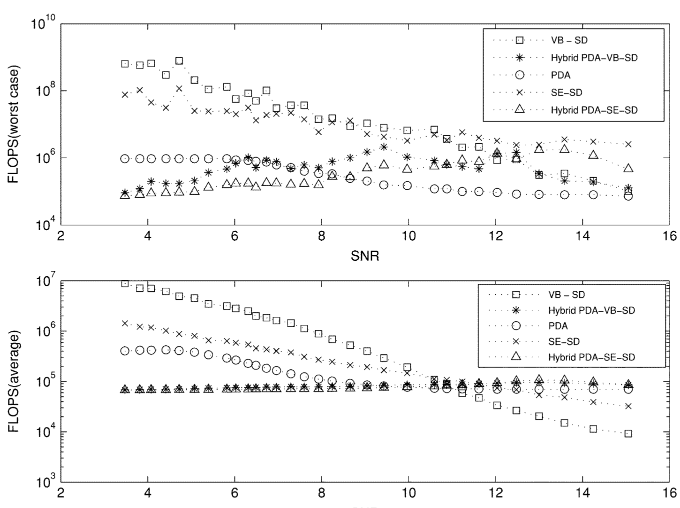

Fig. 2. Computational cost versus SNR, $n_T = n_R = 16$, 4-QAM, 10^4 Monte Carlo runs.

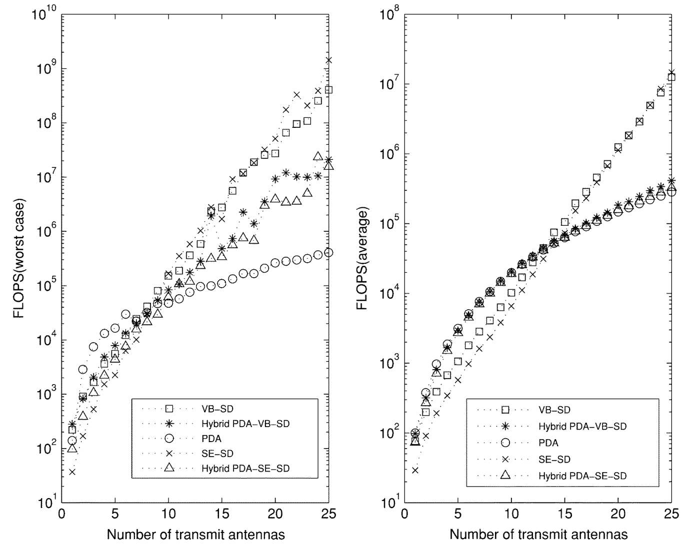

Fig. 3. Computational cost versus n_T, $n_T = n_R$, 4-QAM, SNR $= 10$ dB, 10^4 Monte Carlo runs.

number of bursts has reached five. This ensures sufficient averaging in the low error rate regime while reducing unnecessarily long runs in the high error rate regime. For the computational complexity plots, we use 10^4 (100 bursts of 100 symbol vectors each) Monte Carlo runs per datum reported.

The implementation of PDA does not incorporate the bit-flip stage [8]. The internal threshold parameter of PDA is set to $\epsilon = 10^{-2}/(4\text{SNR})$ as in [8] (note that this is different from our hard decoding threshold τ). The initial radius of SD is set as in Section III; if SD fails to find a point inside the sphere, the radius is increased by one, up to five times (six searches at most). For the SE-SD algorithm, we set the search radius to infinity, which ensures that the ML solution will be found.

Fig. 1 shows the BER performance of PDA, SD, and the hybrid PDA-SD algorithm as a function of $\text{SNR} := 10\log_{10}(\rho/\sigma^2)$, for $n_T = n_R = 16$. Fig. 2 shows the associated average and worst-case computational costs per symbol vector, measured in Floating Point Operations (FLOPS). Finally, Fig. 3 shows FLOPS versus n_T, with $n_T = n_R$, for SNR $= 10$ dB.

V. Conclusion

We have presented a two-stage hybrid PDA-SD algorithm for signal detection in MIMO systems. The basic idea is *dimensionality reduction* via hard decoding and cancellation of those symbols that can be quickly and reliably detected via a single PDA stage. In the V-BLAST scenario considered, simulations show that the proposed hybrid algorithm attains performance close to SD, at a complexity close to PDA. The dimensionality reduction idea can also be applied in conjunction with other variants of SD or SDR.

Acknowledgment

The authors would like to thank W. Zhao and G. Giannakis (Univ. Minnesota) for sharing their MATLAB implementation of SE-SD. They also would like to thank the reviewers for their constructive comments.

References

[1] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[2] A. Chan and I. Lee, "A new reduced-complexity sphere decoder for multiple antenna systems," in *Proc. ICC*, vol. 1, New York, Apr. 28–May 2 2002, pp. 460–464.

[3] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, pp. 14–15, Jan. 1999.

[4] F. Hasegawa, J. Luo, K. Pattipati, P. Willet, and D. Pham, "Speed and accuracy comparison of techniques for multiuser detection in synchronous CDMA," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 540–545, Apr. 2004.

[5] B. Hassibi and H. Vikalo, "On sphere decoding algorithm, part I: Expected complexity," *IEEE Trans. Signal Process.*, to be published.

[6] J. Jaldén and B. Ottersten, "An exponential lower bound on the expected complexity of sphere decoding," in *Proc. ICASSP*, Montreal, QC, Canada, May 17–21, 2004.

[7] G. Latsoudas, "Comparison and Speedup of Near-Optimal Decoding Algorithms for MIMO Systems," diploma thesis, Dept. Elect. Comput. Eng., Tech. Univ. Crete, May 2004.

[8] J. Luo, K. Pattipati, P. Willett, and F. Hasegawa, "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 8, pp. 361–363, Sep. 2001.

[9] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, "Quasi-ML multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 912–922, Apr. 2002.

[10] D. Pham, K. R. Pattipati, P. K. Willett, and J. Luo, "A generalized probabilistic data association detector for multiple antenna systems," *IEEE Commun. Lett.*, vol. 8, no. 4, pp. 205–207, Apr. 2004.

[11] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 1639–1642, Jul. 1999.

[12] R. Wang and G. B. Giannakis, "Approaching MIMO capacity with reduced-complexity soft sphere-decoding," in *Proc. WCNC*, Atlanta, GA, Mar. 21–25, 2004.

[13] W. Zhao and G. B. Giannakis, "Sphere decoding algorithms with improved radius search," in *Proc. WCNC*, Atlanta, GA, Mar. 21–25, 2004.