Efficient Soft-Input Soft-Output MIMO Chase Detectors for arbitrary number of streams

Ahmad Gomaa and Louay M.A. Jalloul, Senior Member, IEEE

Abstract—We present novel soft-input soft-output (SISO) multiple-input multiple-output (MIMO) detectors based on the Chase detection principle in the context of iterative and decoding (IDD). The proposed detector complexity is linear in the signal modulation constellation size and the number of spatial streams. Two variants of the SISO detector are developed, referred to as SISO B-Chase and SISO L-Chase. An efficient method is presented that uses the decoder output to modulate the signal constellation decision boundaries inside the detector leading to the SISO detector architecture. The performance of these detectors significantly improves with just a few number of IDD iterations. The effect of transmit and receive antenna correlation is simulated. For the high-correlation case, the superiority of SISO B-Chase over the SISO L-Chase is demonstrated.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) antenna communication systems are known to achieve large spatial multiplexing and diversity gains in multi-path rich fading channels. All communication systems require some sort of error correction coding for reliable reception, e.g. turbo codes or low-density parity-check codes. The turbo principle as described in has been successfully applied and extended to coded MIMO systems with iterative detection and decoding (IDD) [3, 4]. The ”outer” code is the turbo or LDPC code and the ”inner” code is spatial multiplexing and transmission of the symbols over the multiple antennas. With IDD there is an iterative exchange of information between the MIMO detector and the channel decoder which has been shown to achieve near channel capacity [5]. In IDD architectures, the log-likelihood ratios (LLRs) of the code bits are generated by the MIMO detector and passed to the channel decoder, which computes the extrinsic LLRs and feeds them back to the detector. The detector exploits the a priori LLRs from the decoder to generate more accurate LLRs for the channel decoder to start a new iteration.

There is a rich body of literature on MIMO (or symbol) detectors ranging in complexity from linear minimum mean-squared error (MMSE) [3, 6] to sphere detection [7]. These detectors also differ in terms of their relative performance where a sphere detector typically performs close to maximum likelihood (ML) whereas the linear MMSE is inferior to ML. The core receiver complexity remains to be dominated by the MIMO detector especially when the number of spatial streams and the size of the signal modulation constellation on each stream are large, e.g. 4-layers and 64-QAM.

The drawback of MIMO sphere detectors has been their hardware implementation complexity [6]. In spite of recent advances in complexity reduction of MIMO sphere detectors, as reported in [9]-[11], their nondeterministic processing throughput remains to present a challenge for hardware implementation [2].

In this paper, we develop efficient soft-input soft-output MIMO detectors in the context of IDD. Our proposed MIMO detectors have linear complexity. Furthermore, their structures are generalizable to an arbitrary number of spatial streams and signal modulation constellation size and, moreover, have fixed detection times. Our proposed MIMO detectors are generalization of the Chase family of detection algorithms (referred to in the sequel as SISO B-Chase and SISO L-Chase). Our contribution over the approach in [1] is two-fold: First, in our proposed SISO B-Chase detector, we account for the residual error variance in the symbol detection, making the detector capable of generating soft output information, unlike the B-Chase detector in [1] where only hard output is available. Second, we extend both of the L-Chase and B-Chase detection algorithm to use soft-input through an efficient method for processing the a priori information obtained from the decoder. We compare our detection method to that of [13] that uses soft feedback detection. We show large performance gains with the new detection method over iterative SIOF.

The rest of the paper is organized as follows. Section II introduces the system model. Our soft-input soft-output MIMO detection methods are described in Sections III and IV and their performances are shown in Section V. Finally, Section VI has the concluding remarks.

Notations: Unless otherwise stated, lower case and upper case bold letters denote vectors and matrices, respectively, and 0 denotes the all-zero column vector. $I_m$ denotes the identity matrix of size $m$. The $k$-th element of the vector $x$ is denoted by $x(k)$. Furthermore, $\| x \|_l$, $l=1, 2$, and $E[.]$ denote the absolute value, $l_2$-norm, and statistical expectation, respectively, while $(.)^H$ denotes the complex conjugate transpose operation.

II. SYSTEM MODEL

We consider MIMO systems, where $N_L$ QAM symbols are linearly precoded using the precoding matrix $W$ of size $N_t \times N_L$ and then transmitted over $N_t \geq N_L$ antennas. The receiver detects the transmitted symbols (streams) using $N_r \geq N_L$ receive antennas. The input-output relation is given by

$$ y = \mathbf{H} W \mathbf{s} + \mathbf{n} = \sum_{i=1}^{N_L} \mathbf{h}_i s_i + \mathbf{n} $$

(1)

where $y$, $s$, $n$ and $\mathbf{H}$ denote the $N_r \times 1$ received signal, $N_{t} \times 1$ transmitted symbols, $N_r \times 1$ background noise plus intercell interference, and $N_r \times N_t$ channel matrix, respectively. Furthermore, $h_i$ is the $i$-th column vector of the equivalent
channel matrix $H = HW$, and $s_i$ is the $i$-th transmitted $M$-QAM symbol chosen from constellation $\chi$. The symbol, $s_i$, represents $q = \log_2(M)$ code bits $c_{i1} = [c_{i1} \ c_{i2} \ \cdots \ c_{iq}]$. The above relationship models both single-carrier systems over flat fading channels and orthogonal frequency division multiplexing (OFDM) systems over frequency-selective channels where Equation (1) applies to each subcarrier. Assuming a known channel at the receiver and zero-mean circularly symmetric complex Gaussian noise $n$ of covariance matrix $C_{nn} = B^{-1}$, we write the max-log maximum-a-posteriori (MAP) detector LLR of the bit $c_{ik}$ as follows [4], [5]:

$$L(c_{ik}) = \max_{s_i \in \chi_{k,1}} \eta_{\text{max-log}}(s) - \max_{s_i \in \chi_{k,0}} \eta_{\text{max-log}}(s)$$

(2)

$$\eta_{\text{max-log}}(s) = \sum_{m=1}^{N_L} b_{mn} L_a(c_{mn}) - \|y - \sum_{m=1}^{N_L} h_m s_m\|_B^2$$

(3)

where $s = [s_1 \ s_2 \ \ldots \ s_{N_L}]^T$, $\|x\|_B^2 = x^H B x$, and $\chi_{k,1}$ and $\chi_{k,1}$ denote the constellation sets where the $k$-th bit is ‘1’ and ‘0’, respectively. Furthermore, $L_a(c_{mn})$ denotes the a priori LLR (computed by the decoder) of the bit $c_{mn}$, while $\{b_{mn}\}_{n=1}^{N_L} \in \{0, 1\}$ denote the bit vector representation of $s_m$. The exact brute-force solution of (2) requires the computation of $M^{N_L \times N_L}$ metrics, which is quite complex for large signal modulation constellation sizes and large number of spatial streams. However, in this work we show how we approximate the max-log MAP solution in (2) using much lower complexity. By whitening the noise and multiply $y$ by $\sqrt{B}$, we arrive at an equivalent system model given by

$$\bar{y} = \sqrt{B} y = \bar{H} s + \bar{n}$$

(4)

where $\bar{H} = \sqrt{B} H$ and $\bar{n} = \sqrt{B} n \sim N(0, I_{N_r})$.

### III. SOFT-INPUT SOFT-OUTPUT L-CHASE DETECTORS

In order to generate the LLRs of $s_i$, we first reorganize the columns of $H$ such that its $i$-th and last columns are swapped to get $H_i = \bar{H} P_i$, where $P_i$ is the corresponding $N_L \times N_L$ permutation matrix with $P_i^2 = I_{N_L}$. Next, we obtain the QR decomposition of $\bar{H}_i \equiv \bar{Q}_i \bar{R}_i$ and rotate $y$ by $Q_i^H$ to get

$$y_i = Q_i^H \bar{y} = \bar{R}_i P_i s + \bar{n}_i$$

(5)

where $Q_i^H \bar{Q}_i = I_{N_L}$, $\bar{n}_i = Q_i^H \bar{n} \sim N(0, I_{N_r})$ and $\bar{R}_i$ is an $N_L \times N_L$ upper triangular matrix. Furthermore, $P_i s = [s_{j_1} \ s_{j_2} \ \ldots \ s_{j_N}]^T$ where $j_N = i$ and $j_1 = N_L$. The matrix $\bar{R}_i$ can be partitioned as follows:

$$\bar{R}_i = \begin{bmatrix} \bar{R}_{i1} & \bar{r}_i \\ 0^T & d_i \end{bmatrix}$$

(6)

where $\bar{R}_{i1}$ is the upper triangular $(N_L - 1) \times (N_L - 1)$ matrix constructed by the first $N_L - 1$ rows and columns of $\bar{R}_i$. Also, $\bar{r}_i$ is the last column of $\bar{R}_i$, excluding its $(N_L, N_L)$ element denoted by $d_i$. Next, we get

$$\bar{y}_i = \begin{bmatrix} \bar{R}_{i1}^{-1} \\ 0 \end{bmatrix} y_i = \begin{bmatrix} I_{N_L-1} & c_i \\ 0^T & d_i \end{bmatrix} P_i s + \bar{n}_i$$

(7)

where $c_i = \bar{R}_{i1}^{-1} \bar{r}_i$ is obtained using the back substitution technique [14] where no matrix inversion is needed. Using the new input-output relation in (7), we approximate the LLR in (2) as follows:

$$L(c_{ik}) \approx \max_{s_i \in \chi_{k,1}} \eta_{\text{L-Chase}}(s) - \max_{s_i \in \chi_{k,0}} \eta_{\text{L-Chase}}(s)$$

(8)

$$\eta_{\text{L-Chase}}(s) = \sum_{m=1}^{N_L} b_{mn} L_a(c_{mn}) - \left| \bar{y}_i - \begin{bmatrix} \bar{I}_{N_L-1} & c_i \\ 0^T & d_i \end{bmatrix} P_i s \right|^2$$

$$= \sum_{m=1}^{N_L} b_{mn} L_a(c_{mn}) - \left| \bar{y}_i (N_L) - d_i s_i \right|^2$$

(9)

$$- \sum_{l=1}^{N_L-1} \left( \bar{y}_i(l) - c_i(l) s_i - s_j l \right)^2 E[|\bar{n}_i(l)|^2]$$

where $E[|\bar{n}_i(l)|^2] = 1$ and $E[|\bar{n}_i(l)|^2]$ is the $l_2$-norm of the $l$-th row of $\bar{R}_i^{-1}$. Note that the correlations between the elements of $\bar{n}_i$ are not accounted for in (9) in order to reduce the algorithm complexity. We rewrite the maximization problems in (8) as:

$$\max_{s_i \in \chi_{k,1}(0)} \eta_{\text{L-Chase}}(s) = \max_{s_i \in \chi_{k,1}(0)} \left( \sum_{n=1}^{N_L} b_{n} L_a(c_{n}) - \left| \bar{y}_i (N_L) - d_i s_i \right|^2 E[|\bar{n}_i(l)|^2] \right)$$

$$+ \sum_{l=1}^{N_L-1} \max_{s_{j} \in \chi_{k,1}(l)} \left( \sum_{n=1}^{N_L} b_{n} L_a(c_{n}) - \left| \bar{y}_i(l) - c_i(l) s_i - s_j l \right|^2 E[|\bar{n}_i(l)|^2] \right)$$

(10)

where we enumerate over the stream of interest $s_i$, and compute $\alpha_i + \sum_{l=1}^{N_L-1} \alpha_{i,l}$ for every possible instance of $s_i$. Then, we run the maximization over $\chi_{k,1}$ and $\chi_{k,0}$ to get the LLR. Although (10) reveals that we need to compute $(N_L - 1)^2$ metrics, we show in the next subsection how we exactly and efficiently solve for the sub-maximization problem $\alpha_{i,l}$ using much fewer metric computations.

### A. Efficient computation of the sub-maximization problem

Separating out the real and imaginary parts of the sub-maximization problem $\alpha_{i,l}$, we write [15]:

$$\alpha_{i,l} = \max_{s_{r,j}} \left( \sum_{n \in I_r} b_{n} L_a(c_{n}) - \left| z_{r,l} - s_{r,j} \right|^2 E[|\bar{n}_i(l)|^2] \right)$$

$$+ \max_{s_{l,j} \in \psi} \left( \sum_{n \in I_l} b_{n} L_a(c_{n}) - \left| z_{l,l} - s_{l,j} \right|^2 E[|\bar{n}_i(l)|^2] \right)$$

(11)

where $z_{r,l}(s_{r,j})$ and $z_{l,l}(s_{l,j})$ denote the real and imaginary parts of $z_i(s_{j_l})$, respectively, where $z_i = \bar{y}_i(l) - c_i(l) s_i$. Furthermore, $I_r$ and $I_l$ are the bit indices corresponding to $s_{r,j}$ and $s_{l,j}$, respectively, while $\psi$ denotes the $\sqrt{M}$-PAM one-dimensional constellation shown in Fig. 1, which corresponds to the real or imaginary parts of $s_j$. We use the a priori LLRs to modulate [15] the decision thresholds between

$x_{\pi} \ldots x_3 \ x_2 \ x_1$
the $\sqrt{M}\text{-PAM}$ constellation symbols in Fig. 1. Based on the modulated thresholds, we apply simple slicers on $z_{r,l}$ and $z_{l,l}$ to obtain the solutions of the real and imaginary maximization problems in (11) denoted by $s_{r,j}^*$ and $s_{l,j}^*$, respectively, as:

$$s_{r,j}^* = \arg \max_{s_{r,j}\in \mathbb{C}} \left( \sum_{n\in I_r} b_{jn} L_a(c_{jn}) - \frac{(z_{r,l} - s_{r,j})^2}{E[|\tilde{n}_l(l)|^2]} \right)$$

$$s_{l,j}^* = \arg \max_{s_{l,j}\in \mathbb{C}} \left( \sum_{n\in I_l} b_{mn} L_a(c_{jn}) - \frac{(z_{l,l} - s_{l,j})^2}{E[|\tilde{n}_l(l)|^2]} \right)$$

is the modulated boundary between the one-dimensional constellation symbols $x_m$ and $x_u$ in Fig. 1 where $\{b_{mn}\}_{n=1}^N$ is the bit vector representation of any $M\text{-QAM}$ complex symbol whose real part equals $x_m$. Similarly, we obtain $s_{l,j}$. Despite that the solution of (10) requires enumeration over $s_i$, the $M - \sqrt{M}$ boundaries, $\{D_{mu}\}$, are computed only once and not for each instance of $s_i$. Hence, the total number of metric computations required to detect $s_i$ is $N_L M - (N_L - 1)\sqrt{M} \ll (N_L - 1)M^2$ for large $N_L$ and $M$.

IV. SOFT-INPUT SOFT-OUTPUT B-CHASE DETECTORS

To detect the $i$-th stream, we reorganize the columns of $\mathbf{H}$ such that its $i$-th column is moved to the last column position. The remaining $N_L - 1$ columns are BLAST-sorted and placed at the first $N_L - 1$ column positions. The reorganized channel matrix is $\mathbf{H}_i = \mathbf{H} \mathbf{A}_i$, where $\mathbf{A}_i$ is the $N_L \times N_L$ permutation matrix with $\mathbf{A}_{i,j} = 1$. Next, we obtain the QR decomposition of $\mathbf{H}_i \equiv \mathbf{Q}_i \mathbf{R}_i$ and rotate $y$ by $\mathbf{Q}_i^H$ to get

$$y_i = \mathbf{Q}_i^H \mathbf{y} = \mathbf{R}_i \mathbf{A}_i \mathbf{s} + \mathbf{n}_i$$

where $\mathbf{Q}_i^H \mathbf{Q}_i = \mathbf{I}_{N_L}$, $\mathbf{n}_i = \sim N(0,I_{N_L})$ and $\mathbf{R}_i$ is an $N_L \times N_L$ upper triangle matrix whose $(k,l)$ element is denoted by $r_{kl}$. Also, $\mathbf{A}_i \mathbf{s} = [s_{j_1} s_{j_2} \ldots s_{j_N_L}]^T$ with $j_{N_L} = i$ and $j_i = N_L$. We approximate the LLR in (2) as follows:

$$L(c_{i,k}) \approx \max_{s_{i,k}\in \mathbb{C}} \eta_{\text{B-Chase}}(s_{i,k}) = \max_{s_{i,k}\in \mathbb{C}} \eta_{\text{B-Chase}}(s_{i,k})$$

where $\eta_{\text{B-Chase}}$ is the same as $\eta_{\text{max-log}}$ in (2) except that the symbols $\{s_{j_l}\}_{l=1}^{N_L-1}$ are replaced by their estimates,

$$\begin{cases} \hat{s}_j \mid f = l+1, \{ \hat{s}_j \}_{f=1}^{N_L-1}, \text{ instead of being nullled out as in (7). Furthermore,} \end{cases}$$

$$\text{the residual error after subtracting the symbols estimates is}$$

$$\text{being accounted for using its variance } \var_i \text{ given by}$$

$$\var_i = 1 + \sum_{f=1}^{N_L-1} |r_{lf}|^2 \var(s_{j_l})$$

$$\text{where the symbols estimates and variances are}$$

$$\hat{s}_j = E \left[ \left| s_j \right| \text{Ld}(c_{jn}) \right]_{n=1}^{q}$$

$$\var(s_{j_l}) = E \left[ \left| s_{j_l} - \hat{s}_{j_l} \right| \text{Ld}(c_{jn}) \right]_{n=1}^{q}$$

$$\text{are the post-detection LLRs of the bits representing the symbol }$$

$$\text{s}_{j_l} \text{ obtained as follows:}$$

$$\hat{s}_{j} = \sum_{s_m \in \mathbb{C}} \eta_{\text{B-Chase}}^q \left( \frac{1}{2} \sum_{s_m \in \mathbb{C}} \left| s_{m} \right|^2 (1 + (2b_{mn} - 1)\tanh(L_{p(c_{jn})}/2) \right)$$

$$\text{where } \tanh(.) \text{ is the hyperbolic tangent function.}$$

Next, we rewrite the maximization problem in (15) as:

$$\max_{s_i \in \mathbb{C}} \eta_{\text{B-Chase}}(s_i) = \max_{s_i \in \mathbb{C}} \left( \sum_{n=1}^{q} b_{jn} L_a(c_{jn}) - |y_i(N_L) - d_is_i|^2 \right)$$

$$+ \sum_{l=1}^{N_L-1} \max_{s_l \in \mathbb{C}} \left( \sum_{n=1}^{q} b_{jn} L_a(c_{jn}) - |z_l - s_l|^2 \right) \var(r_{l,l})$$

$$z_l = y_i(l) - r_{l,N_L} s_i - \sum_{f=1}^{N_L-1} r_{lf} \hat{s}_{j_l}$$

Then, we use our algorithm in Section III-A to solve for the maximization problem $\alpha_{i,l}$ in (26), with $E[|\tilde{n}_l(l)|^2]$ replaced by $\var_i/r_{l,l}^2$ while $z_{r,l}$ and $z_{l,l}$ are replaced by the real and imaginary parts, respectively, of $z_i$ in (25).
code rates, namely, 0.5 and 0.83, respectively. Furthermore, two standard multi-path channel models [18] are simulated; namely, the pedestrian-B (PEDB) channel with no antenna correlation, and the extended pedestrian-A (EPA) channel with high antenna correlations, where both transmit and receive correlation coefficients are 0.9. Perfect channel knowledge is assumed at the receiver. Figs. [2] and [3] show that both of the proposed algorithms (SISO L-Chase and B-Chase) outperform SIOF in [13] by several dBs. The waterfall phenomenon is not observed in these figures since we simulate frequency selective channels rather than frequency flat channels. The benefit from the iterations in SIOF receiver is minimal due to the lack of coding gain when using code rate of 0.83 in the case of Fig. [2].

High channel correlations also diminish the benefit of the SIOF method as seen in Fig. [4] due its inability of separating out the correlated streams. Comparing Figs. [2] and [3] we observe the superiority of SISO B-Chase over SISO L-Chase algorithm for channels with high antenna correlation.

VI. CONCLUSION

Based on the Chase detection principle, previously known for producing hard decisions, we developed soft-input soft-output MIMO detectors for two classes of algorithms: SISO B-Chase and SISO L-Chase. An efficient method was described for using the decoder output LLRs to modulate the QAM signal constellation decision boundaries in the detector. The performances of the SISO B-Chase and SISO L-Chase were compared with the SIOF detector in [13] under different coding rates and channel conditions. Simulation results showed that the new proposed detectors have 3 to 4 dB advantage over SIOF at 1% BLER after 3 iterations. Simulations also showed that SISO B-Chase has superior performance over SISO L-Chase for channels with high correlations.

REFERENCES

[1] D. Waters and J. Barry, “The Chase family of detection algorithms for multiple-input multiple-output channels,” IEEE Transactions on Signal Processing, vol. 56, no. 2, pp. 739–747, Feb 2008.
[2] C. Douillard and et al, “Iterative correction of intersymbol interference: turbo equalization,” Eue Trans. on Telecommun., vol. 6, September 1995.
[3] M. Tuchler, A. Singer, and R. Koetter, “Minimum mean squared error equalization using a priori information,” IEEE Trans. on Sig. Proc., vol. 50, no. 3, pp. 673–683, 2002.
[4] H. Vikalo, B. Hassibi, and T. Kailath, “Iterative decoding for MIMO channels via modified sphere decoding,” IEEE Transactions on Wireless Communications, vol. 3, no. 6, pp. 2299–2311, 2004.
[5] B. Hochwald and S. ten Brink, “Achieving near-capacity on a multiple-antenna channel,” IEEE Trans. on Commun., vol. 51, no. 3, pp. 389–399, 2003.
[6] D. Wübben, R. Böhnke, V. Kühn, and K. Kammeyer, “MMSE extension of V-BLAST based on sorted QR decomposition,” in Proc. IEEE Vehicular Technol. Conf. (VTC), Orlando, Florida, Oct. 2003, pp. 508–512.
[7] 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.
[8] D. Garrett, L. Davis, S. ten Brink, B. Hochwald, and G. Knagge, “Silicon complexity for maximum likelihood MIMO detection using spherical decoding,” IEEE J. Solid-State Circuits, vol. 39, no. 9, pp. 1544–1552, Sep. 2004.
[9] R. Wang and G. B. Giannakis, “Approaching MIMO channel capacity with reduced-complexity soft sphere decoding,” in Proc. IEEE Wireless Commun. and Netwo. Conf. (WCNC), vol. 3, Atlanta, GA, Mar. 2004, pp. 1620–1625.
[10] M. S. Yee, “Max-log-MAP sphere decoder,” in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process. (ICASSP), vol. 3, Philadelphia, PA, Mar. 2005, pp. 1013–1016.

[11] C. Studer, A. Burg, and H. Bölcskei, “Soft-output sphere decoder: Algorithms and VLSI implementation,” IEEE J. Sel. Areas Commun., vol. 26, no. 2, pp. 290–300, Feb. 2008.

[12] C. Studer, M. Wenk, and A. Burg, “VLSI implementation of hard- and soft-output sphere decoding for wide-band MIMO systems,” in VLSI-Soc: Forward-Looking Trends in IC and Systems Design, ser. IFIP Advances in Information and Communication Technology, J. L. Ayala, D. Atienza, and R. Reis, Eds. Springer Berlin Heidelberg, 2012, ch. 6, pp. 128–154.

[13] J. Choi, A. Singer, L. Jungwoo, and N. Cho, “Improved linear soft-input soft-output detection via soft feedback successive interference cancellation,” IEEE Trans. on Comm., vol. 58, no. 3, pp. 986–996, 2010.

[14] G. H. Golub and C. F. V. Loan, Matrix Computations-Third Edition. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[15] A. Gomaa and L. Jalloul, “Efficient soft-input soft-output detection of dual-layer MIMO systems,” appeared online, IEEE Wireless Comm. Letters, 2014.

[16] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, “V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel,” in International Symposium on Signals, Systems, and Electronics, 1998.

[17] Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding, 3GPP Std. TS 36.212. [Online]. Available: http://www.3gpp.org

[18] “High speed downlink packet access: UE radio transmission and reception (FDD),” 3GPP TR 25.890, 2002–2005.