Massive BLAST: An Architecture for Realizing Ultra-High Data Rates for Large-Scale MIMO

Ori Shental, Member, IEEE, Sivarama Venkatesan, Alexei Ashikhmin, Fellow, IEEE, and Reinaldo A. Valenzuela, Fellow, IEEE

Abstract—A detection scheme for uplink massive MIMO, dubbed massive-BLAST or M-BLAST, is proposed. The derived algorithm is an enhancement of the well-celebrated soft parallel interference cancellation. Using computer simulations of both uncoded and coded streams in practical massive MIMO application scenarios, M-BLAST is shown to render a substantially better error performance, and consequently a noticeably higher throughput, than not only the benchmark alternative of a one-shot linear detector, but also than the original sequential V-BLAST, although the latter being computationally infeasible for massive MIMO.

I. INTRODUCTION

The introduction of V-BLAST (Vertical-Bell Laboratories Layered Space-Time, [1]) detection algorithm was one of the main enablers in the vast proliferation of multiple-input multiple-output (MIMO) systems in the last two decades. Massive MIMO (a.k.a. large-scale MIMO, originally introduced in [2], see also [3] and references therein) is essentially a scalable version of the point-to-point MIMO, or the multiuser MIMO (which is namely breaking up MIMO’s one end into multiple autonomous terminals), with many (tens, hundreds or even, in the era of millimeter wave wireless communications, thousands of) antennas on both link ends.

The current detection paradigm in massive MIMO mainly relies on simple (one-shot) linear signal processing schemes as the zero-forcing (ZF) and minimum-mean-square-error (MMSE) detectors. As a manifestation of successive interference cancellation (SIC), V-BLAST is not practically extendable for massive MIMO systems since the number of iterations required to peel off the various layers increases with the number of transmitting antennas.

As a potential remedy one may consider the utilization of parallel interference cancellation (PIC, or multistage detector [4]). A soft (hyperbolic tangent) decision version of PIC is known to be asymptotically optimal in the large-system limit [5], assuming it indeed converges. However the latter suffers from relatively slow convergence rate especially in a limit [5], assuming it indeed converges. However the latter is known to be asymptotically optimal in the large-system regime [4]). A soft (hyperbolic tangent) decision version of PIC parallel interference cancellation (SIC), V-BLAST is not practically extendable for massive MIMO systems since the number of transmitting antennas.

Thus, the number of receiving antennas.

The number of transmitting antennas.

The number of receiving antennas, \( M \), is taken to be much smaller than the number of receiving antennas, \( M \). Furthermore, hereinafter for derivation purposes we assume \( K, M \rightarrow \infty \) with a fixed ratio \( \beta \geq \frac{K}{M} \in \mathbb{R} < 1 \).

To simplify the M-BLAST derivation, we also make the naive postulate of perfect power control, yielding users with equal SNR. Later on, post derivation, the obtained M-BLAST algorithm is adopted to the more general case of non-equal SNR across users. An extension to the complex domain is straightforward.

The letter is organized as follows. The massive MIMO system model is described in Section II, Section III derives the M-BLAST algorithm and Section IV presents simulation results for the error performance and throughput gains. Finally, Section V ends this letter with some concluding remarks.

II. SYSTEM MODEL

Consider a basic massive MIMO [3] uplink channel setup with \( K \) transmitting users (single antenna each) and \( M \) receiving antennas at the base station. The MIMO channel adheres to

\[
y = \mathbf{H}x + \mathbf{n},
\]

where \( x \in \{\pm 1\}^K \) is an \( K \)-length hidden binary input vector, \( \mathbf{H} \in \mathbb{R}^{M \times K} \) is a fading channel matrix with entries \( H_{ij} \sim \mathcal{N}(0, \alpha^2) \), \( \mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_M) \) is the ambient \( M \)-length Gaussian noise vector and \( y \in \mathbb{R}^M \) is an \( M \)-length observed vector. The channel matrix, \( \mathbf{H} \), and noise variance, \( \sigma^2 \), are assumed to be known, either perfectly or to some extent, by the base station.

Typical to massive MIMO uplink, the number of concurrently transmitting antennas, \( K \), is taken to be much smaller than the number of receiving antennas, \( M \). Furthermore, hereinafter for derivation purposes we assume \( K, M \rightarrow \infty \) with a fixed ratio \( \beta \geq \frac{K}{M} \in \mathbb{R} < 1 \).

To simplify the M-BLAST derivation, we also make the naive postulate of perfect power control, yielding users with equal SNR. Later on, post derivation, the obtained M-BLAST algorithm is adopted to the more general case of non-equal SNR across users. An extension to the complex domain is straightforward.
III. M-BLAST DERIVATION

The posterior probability associated with the channel model \( \text{Pr}(x|y, H, \sigma^2) \) can be written as
\[
\text{Pr}(x|y, H, \sigma^2) = \frac{1}{Z} \exp \left( \sum_{i<j} R_{ij} x_i x_j + \sum_i h_i x_i \right),
\]
(2)
where
\[
R \triangleq -\frac{H^T H}{\sigma^2}
\]
and
\[
h \triangleq \frac{H^T y}{\sigma^2}.
\]
(4)
Evidently, the normalizing partition function is defined as
\[
Z \triangleq \sum_x \exp \left( \sum_{i<j} R_{ij} x_i x_j + \sum_i h_i x_i \right).
\]
(5)
The symbols \( M_{ij} \) and \( v_i \) denote the \( ij \)th and \( i \)th scalar entries of the corresponding matrix \( M \) and vector \( v \), respectively, while \( \sum_{i<j} \) is a double summation over the upper-triangular entries of (the symmetric) matrix \( R \).

Denoting the partition function’s logarithm
\[
F \triangleq -\ln(Z),
\]
the desired vector of marginal posterior expectations is given, in the large-system limit, by
\[
\mathbb{E}(x|y, H, \sigma^2) = \hat{m} = \arg\min_{m} F(m),
\]
(7)
where \( m \) is the vector of expectations w.r.t. some \( \text{Pr}(x|y, H, \sigma^2) \), which is a tractable distribution approximating the actual intractable posterior distribution \( \text{Pr}(x|y, H, \sigma^2) \).

Leaving, for now, convergence issues aside, the partition function’s logarithm can be approximated via a Taylor expansion as \[ \text{[7]} \]
\[
F(m) = F_0(m) + \lambda F_1(m) + \frac{\lambda^2}{2!} F_2(m) + \ldots,
\]
(8)
with
\[
F_n \triangleq \frac{\partial^n F(m)}{\partial \lambda^n} \bigg|_{\lambda=0},
\]
(9)
to yield
\[
F_0(m) = \sum_i \left( \frac{1}{2} m_i \ln \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \ln \frac{1}{2} \right),
\]
(10)
\[
F_1(m) = -\sum_{i<j} R_{ij} m_i m_j - \sum_i h_i m_i,
\]
(11)
\[
F_2(m) = -\frac{1}{2} \sum_{i<j} R_{ij}^2 (1-m_i^2)(1-m_j^2).
\]
(12)

Hence, minimizing \[ \text{[8]} \] via differentiation with respect to \( m \) for \( \lambda = 1 \)
\[
0 = \frac{\partial F(m_i)}{\partial m_i} = \arctan(m_i) - \sum_j R_{ij} m_j - h_i + \frac{1}{\sigma^2} (1 - \beta \frac{m_i}{k} \sum_j (1-m_j^2)) m_i
\]
(13)

yields the following iterative equations
\[
m^{t+1} = \tanh \left( \frac{1}{\sigma^2} (m^t + H^T z^t) \right),
\]
(14)
\[
z^t = y - Hm^t + o^{t-1},
\]
(15)
\[
o^{t-1} \triangleq \beta z^{t-1} (1 - \left( \tanh^2 \left( \frac{1}{\sigma^2} (m^{t-1} + H^T z^{t-1}) \right) \right)),
\]
(16)
where \( \langle \cdot \rangle \) is an average of a vector, and the initial conditions are \( m^0 \leq 0 = z^0 \leq -1 = o^0 \leq -2 = 0 \).

From these fixed-point equations an approximation to the desired posterior expectation can be inferred after a predetermined number of iterations.

Note that arbitrarily setting \( o^t = 0 \) (i.e., removing \[ \text{[10]} \] from the set of equations), the fixed-point equations \[ \text{[14]-[15]} \] boil down to the well-known (soft) PIC. Therefore in this sense, per a given number of iterations, the derived M-BLAST scheme can be viewed as simply an improvement (i.e., a more accurate Taylor expansion to the partition function logarithm) over the conventional PIC. Furthermore, complexity-wise the computation of the additional term (w.r.t. ordinary soft PIC) in M-BLAST, \( o^t \) \[ \text{[15]} \] (originating from what is known as the Onsager correction term \[ \text{[12], [6]} \]), requires only a straightforward and simple processing of already obtained information from previous iterations (namely \[ \text{[14]-[15]} \]).

In order to:
1) account for non-equal SNR across users, let \( A \) be a diagonal matrix with the corresponding user magnitudes, \( A_{ii} = \sqrt{P_i} \), where \( P_i \) is the transmit power of the \( i \)th user;
2) better facilitate compatibility with ordinary (soft) PIC also for \( finite-size \) massive MIMO systems, for which \( D = \text{diag}(H^T H)A \) can be only approximately replaced by the magnitude matrix \( A \):

the self-consistency equations are rewritten as
\[
m^{t+1} = \tanh \left( \frac{1}{\sigma^2} A(m^t + H^T z^t) \right),
\]
(17)
\[
z^t = y - HAm^t + o^{t-1},
\]
(18)
\[
o^{t-1} \triangleq \beta z^{t-1} (1 - \left( \tanh^2 \left( \frac{1}{\sigma^2} (A(m^{t-1} + H^T z^{t-1}) \right) \right)),
\]
(19)

Also note that convergence is guaranteed as long as
\[
\beta \left( 1 - \left( \tanh^2 \left( \frac{1}{\sigma^2} A(m^t + H^T z^t) \right) \right) \right) < 1.
\]
(20)

Hence the proposed scheme is typically suitable for uplink (underloaded) massive MIMO scenarios with \( \beta < 1 \). Finally, note that the obtained self-consistency iterative equations \[ \text{[17]-[19]} \] may provide a rationalization to recent literature on \( damped \) interference cancellation schemes (e.g., \[ \text{[8]} \]) originally established mainly on heuristics, thus heavily rely on simulation-based optimization of the damping factor, rather than firm theoretical justification.

IV. SIMULATION RESULTS

The proposed M-BLAST scheme \[ \text{[17]-[19]} \] with \( t = 10 \) iterations is simulated in an uplink flat-fading massive MIMO channel with \( K = 500 \) BPSK transmitting users and \( M = 1000 \) receiving antennas at the base station (thus the load \( \beta = \ldots \)
The error performance of the M-BLAST, in bit-error-rate (BER), is compared to the non-fading single-input single-output (SISO) AWGN lower bound and to several conventional detectors: a one-shot linear MMSE, ZF- and MMSE-based SIC (V-BLAST) and ordinary soft PIC, with also $t = 10$ iterations. We first assume uncoded streams and users transmitting with equal SNR.

Fig. 1 plots the BER vs. $E_b/N_0$ for the different detectors assuming perfect channel state information (CSI) and ideal SNR estimation at the base station. As shall be explained in the following, a typical operational SNR range is chosen in compliance with 3GPP standardization and recent massive MIMO studies.

M-BLAST is observed to yield a reduced BER behaviour across the entire examined $E_b/N_0$ range compared to the common MMSE and V-BLAST (being impractical as inherently requires $K = 500$ extraction iterations) error performance. For the intermediate SNR levels M-BLAST exhibits non-negligible gains also over the soft PIC.

Quantitatively, for the working point example of 1% BER, M-BLAST yields gains of approximately 2dB (vs. MMSE), 1dB (V-BLAST), 0.3dB (PIC) and is only 0.5dB away from the SISO-AWGN bound. The inset in Fig. 1 displays the BER for a (currently) more commercially implementable setup of $K = 32$ users and $M = 96$ receiving antennas (thus total of 128 antennas with load $\beta = 1/3$). Similar trend and gains in error performance of M-BLAST over the linear MMSE are observed.

The inset in Fig. 1 shows the BER for total of 128 antennas case.

Fig. 2 plots the BER vs. $E_b/N_0$ for the different detectors under imperfect CSI (assuming 6dB pilot boost) and non-ideal SNR estimation at the base station. For the latter, the estimated noise variance at the receiver, $\hat{\sigma}^2$, is randomly taken from a uniform distribution within the range $(1 \pm X)\sigma^2$ ($X = 1\%$ for the below described LTE use-case, which is a typical value for static users, and $X = 5\%$ in the massive MIMO use-case for pedestrian users).

Even under such non-ideal, yet pragmatic, conditions M-BLAST is shown to yield an improved error performance across the examined $E_b/N_0$ range. Again, for the intermediate SNR levels M-BLAST presents substantial gains also over the conventional soft PIC. Particularly, for the 1% BER reference point M-BLAST yields gains of about 2dB (vs. MMSE), 1dB (V-BLAST), 0.7dB (PIC) and is now approximately 1dB apart from the SISO-AWGN lower bound.

We now simulate the case of users with non-equal SNR (accounting for, e.g., path loss and non-perfect power control) and coded streams. We repeat the same setup as in the aforementioned model (of Fig. 1) but with user streams encoded by rate-1/2 convolutional codes and user SNRs taken from 3GPP-compliant cumulative distribution function (CDF). In plotting the coded BER vs. user’s average $E_b/N_0$ in Fig. 3 an individual detection (using either the benchmark linear MMSE or the proposed M-BLAST with $t = 5$ iterations) and decoding (using soft Viterbi algorithm) scheme is adopted. Again, the inset presents the corresponding coded BER curves for the $K = 32$ users and $M = 96$ receiving antennas scenario. Looking at the coded BER, non-negligible gains of about 1dB for the $K = 500$ setup and about 0.5dB for the $K = 32$ configuration, for M-BLAST over the linear MMSE, are observed. Similar relative gains of M-BLAST over MMSE were also observed under imperfect CSI and non-ideal SNR estimation.

Next, the achievable user’s uplink throughput is evaluated for the various detectors in the following manner. First, the simulated post-detection signal-to-interference-and-noise-ratio (SINR) for the different detectors are plugged into Shannon’s SISO-AWGN capacity, serving for our purposes as an upper bound on the user throughput. Second, the CDF of a user’s SINR in a multi-cell network is first learned from two mentioned model (of Fig. 1) but with user streams encoded by rate-1/2 convolutional codes and user SNRs taken from 3GPP-compliant cumulative distribution function (CDF). In plotting the coded BER vs. user’s average $E_b/N_0$ in Fig. 3 an individual detection (using either the benchmark linear MMSE or the proposed M-BLAST with $t = 5$ iterations) and decoding (using soft Viterbi algorithm) scheme is adopted. Again, the inset presents the corresponding coded BER curves for the $K = 32$ users and $M = 96$ receiving antennas scenario. Looking at the coded BER, non-negligible gains of about 1dB for the $K = 500$ setup and about 0.5dB for the $K = 32$ configuration, for M-BLAST over the linear MMSE, are observed. Similar relative gains of M-BLAST over MMSE were also observed under imperfect CSI and non-ideal SNR estimation.

1) 3GPP’s specifications for the uplink in a LTE’s multi-cell 3D-UMi (urban micro-cell) channel (e.g., [9]);

2) Furthermore, a straightforward enumeration of ‘multiply & accumulate’ (MAC) operations shows that for even a much more moderate setup of $K = 32$ and $M = 96$, MMSE-based V-BLAST requires more than 500 times the MAC operations than linear MMSE, while the proposed M-BLAST requires only as low as the order of $3t/5$ times the operations than the one-shot linear MMSE.

3) As may be envisioned for Internet of Things (IoT) or massive Machine-Type Communications (mMTC) use-cases.
2) A massive MIMO uplink system.

For the latter scenario the user’s SINR CDF, as recently been evaluated ([10], and specifically Section VII and Fig. 4 therein), is adopted here to benchmark the typical user SINRs and the corresponding achievable uplink throughput gains of M-BLAST in the uplink of massive MIMO cellular network.

The observed gains in maximal throughput for a certain fraction of the users (in % out of the total number of users in a cell) with the lowest achievable information rates (i.e., the maximal rate of % of the ‘rate-deprived’ users) of M-BLAST w.r.t. the legacy detectors are summarized in the following tables for perfect and imperfect CSI scenarios. Looking first at the 3GPP uplink use-case (Table I upper row), M-BLAST introduces significant throughput gains, compared to the legacy detectors (and especially vs. the state-of-the-art MMSE in the shaded column), across the throughput CDF, under ideal CSI conditions, thus may be beneficial for all users within a cell.

Under the implementation penalty (Table II upper row), the trend in gains remains (especially for the 10% cell-edge and 90% center-cell users), although no gain emerges for the median (50%) throughput compared to MMSE and (impractical) MMSE-SIC. Note that M-BLAST gains vs. the conventional soft PIC intensify, alluding to the improved robustness of the suggested interference cancellation architecture under such pragmatic CSI conditions. For the massive MIMO use-case (lower row on Tables I and II), similar attractive behavior to the LTE use-case of M-BLAST is demonstrated except for zero gain, compared to MMSE, appearing this time for 10% cell-edge users under imperfect CSI conditions.

V. Conclusion

This letter advocates for the utilization of an improved interference cancellation scheme, which is based upon parallel rather than successive detection architecture (as the legacy V-BLAST), thus being more suitable for large-scale MIMO applications. Note that in addition to the massive MIMO realm, M-BLAST may also be extremely beneficial for IoT or 5G’s mMTC for which a lot of users (devices) transmit simultaneously. The derived M-BLAST algorithm also calls for a new look at PIC-like design as an alternative to the widespread conception of SIC as a detector for non-orthogonal multiple-access (NOMA, [11]) in the uplink. Study of M-BLAST architecture for higher constellations is currently underway.

REFERENCES

[1] P. W. Wolniansky, G. J. Foschini, G. Golden, and R. A. Valenzuela, “V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel,” in Signals, Systems, and Electronics, 1998, ISEE 98. 1998 URSI International Symposium on. IEEE, 1998, pp. 295–300.

[2] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590–3600, 2010.

[3] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, Fundamentals of Massive MIMO. Cambridge University Press, 2016.

[4] M. K. Varanasi and B. Aazhang, “Multistage detection in asynchronous code-division multiple-access communications,” IEEE Trans. Commun., vol. 38, no. 4, pp. 509–519, 1990.

[5] T. Tanaka and M. Okada, “Approximate belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection,” IEEE Trans. Inf. Theory, vol. 51, no. 2, pp. 700–706, 2005.

[6] J. Yedidia, “An idiosyncratic journey beyond mean field theory,” Advanced mean field methods: Theory and practice, pp. 21–36, 2001.

[7] T. Pfefla, “Convergence condition of the TAP equation for the infinite-ranged Ising spin glass model,” Journal of Physics A: Mathematical and general, vol. 15, no. 6, p. 1971, 1982.

[8] P. Som, T. Datta, A. Chockalingam, and B. S. Rajan, “Improved large-MIMO detection based on damped belief propagation,” in Information Theory (ITW 2010, Cairo), 2010 IEEE Information Theory Workshop on. IEEE, 2010, pp. 1–5.

[9] “3GPP self-evaluation methodology and results,” 3GPP, Tech. Rep., Dec 2009.

[10] A. Adhikary, A. Ashikhmin, and T. L. Marzetta, “Uplink interference reduction in large-scale antenna systems,” IEEE Trans. Commun., vol. 65, no. 5, pp. 2194–2206, 2017.

[11] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, “Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends,” IEEE Commun. Mag., vol. 53, no. 9, pp. 74–81, Sep. 2015.