Large-Scale MIMO Detection for 3GPP LTE: Algorithms and FPGA Implementations

Michael Wu, Bei Yin, Guohui Wang, Chris Dick, Joseph R. Cavallaro, and Christoph Studer

Abstract—Large-scale (or massive) multiple-input multiple-output (MIMO) is expected to be one of the key technologies in next-generation multi-user cellular systems based on the upcoming 3GPP LTE Release 12 standard, for example. In this work, we propose—to the best of our knowledge—the first VLSI design enabling high-throughput data detection in single-carrier frequency-division multiple access (SC-FDMA)-based large-scale MIMO systems. We propose a new approximate matrix inversion algorithm relying on a Neumann series expansion, which substantially reduces the complexity of linear data detection. We analyze the associated error, and we compare its performance and complexity to those of an exact linear detector. We present corresponding VLSI architectures, which perform exact and approximate soft-output detection for large-scale MIMO systems with various antenna/user configurations. Reference implementation results for a Xilinx Virtex-7 XC7VX980T FPGA show that our designs are able to achieve more than 600 Mb/s for a 128 antenna, 8 user 3GPP LTE-based large-scale MIMO system. We finally provide a performance/complexity trade-off comparison using the presented FPGA designs, which reveals that the detector circuit of choice is determined by the ratio between BS antennas and users, as well as the desired error-rate performance.

Index Terms—Approximate matrix inversion, FPGA design, large-scale (or massive) MIMO, linear soft-output detection, minimum mean square error (MMSE), Neumann series, VLSI.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) in combination with spatial multiplexing [3] builds the foundation of most modern wireless communication standards, such as 3GPP LTE [4]–[6] or IEEE 802.11n [7]. MIMO technology offers significantly higher data rates over single-antenna systems by transmitting multiple data streams concurrently and in the same frequency band. Conventional MIMO wireless systems, however, already start to approach their throughput limits. Consequently, the deployment of novel transceiver technologies is of paramount importance in order to meet the ever-growing demand for higher data rates, better link reliability, and improved coverage, without further increasing the communication bandwidth [8]–[10].

M. Wu and B. Yin contributed equally to the paper.

Parts of this paper for a large-scale, point-to-point MIMO-OFDM system have been presented at the IEEE International Symposium on Circuit and Systems (ISCAS) [1] and the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) [2].

M. Wu, B. Yin, G. Wang, and J. Cavallaro are with the Dept. of ECE, Rice University, Houston, TX (e-mail: {mbw2, by2, wgh, cavallar}@rice.edu).

C. Dick is with Xilinx, Inc., San Jose, CA (e-mail: chris.dick@xilinx.com).

C. Studer is with the School of ECE, Cornell University, Ithaca, NY (e-mail: studer@cornell.edu)

This work was supported in part by Xilinx and by the US National Science Foundation under grants CNS-1265332, ECCS-1232274, EECS-0925942, and CNS-0923479.

A. Blessing and Curse of Massive MIMO

Large-scale (or massive) MIMO is an emerging technology, which postulates the use of antenna arrays having orders of magnitude more elements at the base station (BS) compared to conventional (small-scale) MIMO systems, while serving tens of users simultaneously and in the same frequency band [8]. This technology promises significant improvements in terms of spectral efficiency, link reliability, and coverage compared to conventional (small-scale) systems [9], [11], [12].

Unfortunately, the promised benefits of large-scale MIMO come at the cost of significantly increased computational complexity in the BS, as opposed to small-scale MIMO systems, which commonly deploy 2-to-4 antennas at both ends of the wireless link. In particular, data detection in the large-scale MIMO uplink is expected to be among the most critical tasks in terms of complexity and power consumption, as the presence of hundreds of antennas at the BS and a large number of users will increase the computational complexity by orders of magnitude. In addition, current cellular systems, such as 3GPP-LTE [4], [5] or LTE-Advanced (LTE-A) [6], rely on single-carrier frequency division multiple access (SC-FDMA), which further increases the dimensionality (and hence the complexity) of the underlying detection problem. As a consequence, optimal data detection methods, such as maximum-likelihood (ML) detection [13]–[15] or soft-output sphere decoding (SD) [16]–[18], whose (average) computational complexity scales exponentially in the number of transmitted data streams [19], [20], would simply result in prohibitive complexity. Hence, one has to resort to low-complexity (but sub-optimal) linear detection schemes [9] or stochastic detection algorithms [21] that deliver acceptable error-rate performance and scale favorably to the high-dimensional detection problems faced in SC-FDMA-based large-scale MIMO systems.

B. Contributions

This paper addresses the complexity issue of data detection in SC-FDMA-based large-scale MIMO systems in the uplink, i.e., where multiple users communicate with the BS. We focus on linear soft-output detection in combination with a new approximate matrix inversion method relying on a Neumann series expansion, which significantly reduces the computational complexity compared to that of an exact matrix inversion method. We analyze the implementation trade-offs associated with approximate and exact linear data detection in the large-scale MIMO uplink, and we show analytically that the approximation error caused by the proposed approximate inversion method depends on the ratio between BS antennas

and users. We show that the proposed approximation performs well for medium to large ratios between BS antennas and users, while exact linear detection is advantageous for small antenna ratios. We present reference FPGA designs for both, the approximate and exact matrix inversion, and for various antenna configurations, which enables us to characterize the associated hardware complexity vs. error-rate performance trade-offs. The resulting FPGA designs are—to the best of our knowledge—the first data detection engines for massive MIMO systems reported in the open literature that achieve a peak uplink throughput exceeding the $300\,\mathrm{Mb/s}$ specified in 3GPP LTE-Advanced operating at $20\,\mathrm{MHz}$ bandwidth [6].

C. Notation

Lowercase boldface letters stand for column vectors; uppercase boldface letters designate matrices. For a matrix $\mathbf{A}$, we denote its transpose and conjugate transpose $\mathbf{A}^T$ and $\mathbf{A}^H$, respectively. The entry in the kth row and ℓ^{th} column of a matrix $\mathbf{A}$ is denoted by $A_{k,\ell}$; the kth entry of a vector $\mathbf{a}$ is designated by a_k. The Frobenius norm and ℓ_2-norm of a matrix $\mathbf{A}$ and vector $\mathbf{a}$ are denoted by $\|\mathbf{A}\|_F$ and $\|\mathbf{a}\|_2$, respectively. The $M \times M$ identity matrix is denoted by $\mathbf{I}_M$, and $\mathbf{F}_M$ refers to the $M \times M$ discrete Fourier transform (DFT) matrix, normalized as $\mathbf{F}_M^H \mathbf{F}_M = \mathbf{I}_M$. In order to simplify notation, we make frequent use of the superscript $(\cdot)^{(i,j)}$ to indicate the ith base-station antenna and j^{th} user; the subscript $(\cdot)_w$ designates the SC-FDMA subcarrier index.

D. Paper Outline

The remainder of the paper is organized as follows. Section II introduces the uplink system model and outlines the basics of linear detection for SC-FDMA-based systems. The approximate matrix inversion approach, a corresponding error analysis, and an error-rate performance/complexity comparison are shown in Section III. Section IV details our VLSI architecture. Section V provides reference FPGA implementation results and a trade-off analysis. We conclude in Section VI. All proofs are relegated to the Appendices.

II. LARGE-SCALE MIMO IN LTE UPLINK

We next introduce the LTE uplink model and present a new and efficient method for linear soft-output minimum mean-square error (MMSE) detection in SC-FDMA-based systems.

A. LTE Uplink Model

We consider the large-scale multi-user (MU) MIMO uplink with B antennas at the base-station (BS) communicating with $U \le B$ single-antenna users.[1] To reduce the peak-to-average power ratio of the user equipment, LTE uplink employs SC-FDMA (short for single-carrier frequency division multiple access) [5]. The U users first encode their own transmit bits using channel encoders and then, map the coded bit stream

to time-domain constellation points in the finite alphabet $\mathcal{O}$ with cardinality $M = |\mathcal{O}|$ and average transmit power E_s per symbol. An L-point discrete Fourier transform (DFT) block[2] is used to perform modulation of these time-domain symbols onto orthogonal frequency bands. The L time-domain constellation points for the ith user are subsumed in the vector $\mathbf{x}^{(i)} = \left[x_1^{(i)}, \ldots, x_L^{(i)} \right]^T$. The output of the DFT block, namely the frequency-domain symbol, is defined as $\mathbf{s}^{(i)} = \left[\mathbf{s}_1^{(i)}, \ldots, \mathbf{s}_L^{(i)} \right]^T = \mathbf{F}_L \mathbf{x}^{(i)}$. Subsequent processing performed for each user corresponds to that of conventional orthogonal frequency-division multiplexing (OFDM) transmission [22]. Specifically, for each user, the frequency-domain symbols are first mapped onto data-carrying subcarriers and then, transformed back to the time domain with an inverse DFT (IDFT). After prepending the cyclic prefix to the time-domain symbols, all U users transmit their time-domain signals simultaneously over the wireless channel.

At the BS, each receive antenna obtains a mixture of the time-domain signals from all users. For data detection, the time-domain signals received at each antenna are first transformed back into the frequency domain using a DFT. The data-carrying symbols are then extracted from the DFT's output. Assuming a sufficiently long cyclic-prefix (i.e., longer than the delay spread of the channel's impulse response), the received frequency-domain symbols can be modeled using the standard input-output relation $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}$, with the following definitions:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^{(1)} \\ \vdots \\ \mathbf{y}^{(B)} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}^{(1,1)} & \cdots & \mathbf{H}^{(1,U)} \\ \vdots & \ddots & \vdots \\ \mathbf{H}^{(B,1)} & \cdots & \mathbf{H}^{(B,U)} \end{bmatrix},$$

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}^{(1)} \\ \vdots \\ \mathbf{s}^{(U)} \end{bmatrix}, \quad \text{and } \mathbf{n} = \begin{bmatrix} \mathbf{n}^{(1)} \\ \vdots \\ \mathbf{n}^{(B)} \end{bmatrix}.$$

Here, the vector $\mathbf{y}^{(i)} = \left[y_1^{(i)}, \ldots, y_L^{(i)} \right]^T$ contains the received symbols on the ith antenna in the frequency domain, where $y_w^{(i)}$ is the symbol received on the wth subcarrier of the ith antenna. The $L \times L$ diagonal matrix $\mathbf{H}^{(i,j)} = \mathrm{diag}\left(h_1^{(i,j)}, \ldots, h_L^{(i,j)} \right)$ contains the channel's frequency response of length L between the ith receive antenna and j^{th} transmit antenna on its main diagonal, and $\mathbf{n}^{(i)} = \left[n_1^{(i)}, \ldots, n_L^{(i)} \right]^T$ models thermal noise at the ith receive antenna in the frequency domain. The entries of the vector $\mathbf{n}^{(i)}$ are assumed to be i.i.d. zero-mean Gaussian with variance N_0 per complex entry.

B. Linear MMSE Detection

The task of a data detector for MIMO systems is to compute soft-estimates in the form of log-likelihood ratio (LLR) values for each coded bit, given the channel matrix[3] $\mathbf{H}$ and receive vector $\mathbf{y}$. In order to arrive at low computational complexity for data detection in SC-FDMA-based large-scale

---

[1]More generally, instead of having U single-antenna users, the proposed system may equivalently support U spatial streams, which can, for example, be shared among a smaller number of user terminals that are equipped with more than one antenna.

[2]In practice, the DFT and inverse DFT are carried out by fast (inverse) Fourier transform (I/FFT) units.

[3]In practice, channel-state information is acquired using pilot sequences specified by the standard [23]. For the sake of simplicity, we assume perfect channel state information (CSI) throughout the paper. An investigation of the impact of imperfect CSI on the error-rate performance is left for future work.

MIMO systems, we focus exclusively on linear soft-output detection [24]. Linear detection for SC-FDMA mainly consists of the following two steps: (i) *channel equalization* to generate estimates of the frequency domain symbols, and (ii) *soft-output computation* to generate LLRs from the equalized frequency domain symbols. Both of these steps are detailed next.

(i) Channel equalization: The most common approach to linear MIMO detection is the minimum-mean square error (MMSE) equalizer, which computes equalized frequency-domain symbols as $\hat{\mathbf{s}} = \mathbf{W}\mathbf{y}$ with the MMSE equalization matrix defined as follows [3]:

$$\mathbf{W} = \left(\mathbf{H}^H\mathbf{H} + N_0 E_\mathrm{s}^{-1}\mathbf{I}_{LU}\right)^{-1}\mathbf{H}^H.$$

Since the effective channel matrix $\mathbf{H}$ is built from diagonal $L \times L$ submatrices, we can apply MMSE equalization on a per-subcarrier basis. Specifically, the received frequency symbols on the wth subcarrier in the frequency domain can be modeled as $\mathbf{y}_w = \mathbf{H}_w\mathbf{s}_w + \mathbf{n}_w$, where

$$\mathbf{y}_w = \begin{bmatrix} y_w^{(1)} \\ \vdots \\ y_w^{(B)} \end{bmatrix}, \quad \mathbf{H}_w = \begin{bmatrix} h_w^{(1,1)} & \cdots & h_w^{(1,U)} \\ \vdots & \ddots & \vdots \\ h_w^{(B,1)} & \cdots & h_w^{(B,U)} \end{bmatrix},$$

$$\mathbf{s}_w = \begin{bmatrix} s_w^{(1)}, \ldots, s_w^{(U)} \end{bmatrix}^T, \quad \text{and} \quad \mathbf{n}_w = \begin{bmatrix} n_w^{(1)}, \ldots, n_w^{(B)} \end{bmatrix}^T.$$

Here, $y_w^{(i)}$ is the frequency symbol received on the wth subcarrier for the ith antenna, and $h_w^{(i,j)}$ is the frequency gain (or attenuation) on the wth subcarrier between the ith receive antenna and j^th transmit antenna. The scalar $s_w^{(j)}$ denotes the symbol transmitted by the jth user on the wth subcarrier; the scalar $n_w^{(i)}$ models thermal noise at the ith receive antenna on the wth subcarrier. With this reformulation, the equalized symbols on the wth subcarrier are given by $\hat{\mathbf{s}}_w = \mathbf{W}_w\mathbf{y}_w$, with the per-subcarrier MMSE equalization matrix defined as

$$\mathbf{W}_w = \left(\mathbf{H}_w^H\mathbf{H}_w + N_0 E_\mathrm{s}^{-1}\mathbf{I}_U\right)^{-1}\mathbf{H}_w^H. \quad (1)$$

A key method to arrive at low-complexity linear MMSE detection was put forward in [25]. This approach first computes the matched-filter (MF) output as $\mathbf{y}_w^\mathrm{MF} = \mathbf{H}_w^H\mathbf{y}_w$ and the Gram matrix $\mathbf{G}_w = \mathbf{H}_w^H\mathbf{H}_w$ for each subcarrier w, followed by forming the regularized Gram matrix $\mathbf{A}_w = \mathbf{G}_w + N_0 E_\mathrm{s}^{-1}\mathbf{I}_U$. The equalized symbols per subcarrier are then computed as $\hat{\mathbf{s}}_w = \mathbf{A}_w^{-1}\mathbf{y}_w^\mathrm{MF}$, which requires the *explicit* computation of a $U \times U$-dimensional matrix inverse.[4]

(ii) LLR computation: To obtain symbol estimates in the time domain, the MMSE detector performs an IDFT on the equalized frequency domain symbols for each user. The time-domain symbol estimates for the ith user are given by $\hat{\mathbf{x}}^{(i)} = \mathbf{F}_L^H\hat{\mathbf{s}}^{(i)}$, where $\mathbf{F}_L^H$ is the IDFT matrix and $\hat{\mathbf{x}}^{(i)} = \begin{bmatrix} \hat{x}_1^{(i)}, \ldots, \hat{x}_L^{(i)} \end{bmatrix}^T$ contains the time-domain symbol estimates of the symbols transmitted by the ith user. To extract LLRs from the time-domain symbol estimates, we approximate each estimate as an independent Gaussian random variable. In

particular, the estimated t^th symbol transmitted from the ith user is modeled as $\hat{x}_t^{(i)} = \mu^{(i)}x_t^{(i)} + e_t^{(i)}$, where $\mu^{(i)}$ is the effective channel gain and $e_t^{(i)}$ is the post-equalization noise-plus-interference (NPI) variance. Let ν_i^2 be the variance of $e_t^{(i)}$ and b be the bit index of the LLR associated with the tth symbol transmitted from the ith user. With this model, the max-log LLRs can be computed as [25], [27]

$$L_t^{(i)}(b) = \rho_i^2 \left(\min_{a \in \mathcal{O}_b^0} \left| \frac{\hat{x}_t^{(i)}}{\mu^{(i)}} - a \right|^2 - \min_{a' \in \mathcal{O}_b^1} \left| \frac{\hat{x}_t^{(i)}}{\mu^{(i)}} - a' \right|^2 \right), \quad (2)$$

where $\rho_i^2 = \left(\mu^{(i)}\right)^2 / \nu_i^2$ is the post-equalization signal-to-noise-plus-interference ratio (SINR), and $\mathcal{O}_b^0$ and $\mathcal{O}_b^1$ correspond to the sets of constellation symbols for which the bth bit equals to 0 and 1, respectively.

In order to obtain an explicit formulation of the effective channel gain $\mu^{(i)}$ as well as the NPI variance ν_i^2, we can write the tth symbol estimate of the ith user as follows:

$$\hat{x}_t^{(i)} = \mathbf{f}_t^H\hat{\mathbf{s}}^{(i)} = \mathbf{f}_t^H\mathbf{W}^{(i,:)}\mathbf{y}.$$

Here, $\mathbf{W}^{(i,:)} = \begin{bmatrix} \mathbf{W}^{(i,1)}, \ldots, \mathbf{W}^{(i,B)} \end{bmatrix}$ is a horizontal concatenation of the ith block row of (diagonal) submatrices of $\mathbf{W}$. The row vector $\mathbf{f}_t^H$ corresponds to the tth row of the IDFT matrix $\mathbf{F}_L^H$. Let $\mathbf{H}^{(:,j)} = \begin{bmatrix} \mathbf{H}^{(1,j)}, \ldots, \mathbf{H}^{(B,j)} \end{bmatrix}^T$ be the horizontal concatenation of the jth block column of (diagonal) submatrices of $\mathbf{H}$, consisting of the frequency-domain channel responses between the receive antennas and the transmit antenna associated with the jth user. We first compute the effective channel gain:

$$\mu^{(i)}x_t^{(i)} = \mathsf{E}\left[\mathbf{f}_t^H\mathbf{W}^{(i,:)}\mathbf{y} \,|\, x_t^{(i)}\right] = L^{-1}\mathrm{tr}(\mathbf{W}^{(i,:)}\mathbf{H}^{(:,i)})x_t^{(i)}.$$

Since $\mathbf{W}^{(i,j)}$ and $\mathbf{H}^{(i,j)}$ are both diagonal matrices, we can write $\mu^{(i)}$ as a sum of per-subcarrier operations. In particular, let $\mathbf{w}_{i,w}^H$ be the ith row of $\mathbf{W}_w$ and $\mathbf{h}_{i,w}$ be the ith column of $\mathbf{H}_w$. Then, we obtain the effective channel gain as

$$\mu^{(i)} = L^{-1}\sum_{w=1}^{L}\mathbf{w}_{i,w}^H\mathbf{h}_{i,w}. \quad (3)$$

We next compute the post-equalization NPI variance ν_i^2 of the residual noise plus interference as

$$\nu_i^2 = \mathsf{E}\left[\left|\hat{x}_t^{(i)}\right|^2\right] - \mathsf{E}\left[\left|\mu^{(i)}x_t^{(i)}\right|^2\right]$$

$$= \mathsf{E}\left[\mathbf{f}_t^H\mathbf{W}^{(i,:)}(\mathbf{Hs} + \mathbf{n})(\mathbf{Hs} + \mathbf{n})^H(\mathbf{W}^{(i,:)})^H\mathbf{f}_t\right] - E_s\left|\mu^{(i)}\right|^2.$$

The MMSE equalization matrix can be written in two ways [25], i.e., either

$$\mathbf{W} = \left(\mathbf{H}^H\mathbf{H} + N_0 E_\mathrm{s}^{-1}\mathbf{I}_{LU \times LU}\right)^{-1}\mathbf{H}^H \text{ or}$$

$$\mathbf{W} = \mathbf{H}^H\left(\mathbf{HH}^H + N_0 E_\mathrm{s}^{-1}\mathbf{I}_{LB \times LB}\right)^{-1}.$$

Hence, we have $\mathbf{W}\left(E_s\mathbf{HH}^H + N_0\mathbf{I}_{LB \times LB}\right) = E_s\mathbf{H}^H$; this allows us to rewrite the post-equalization NPI in compact form as follows [25]:

$$\nu_i^2 = E_s\mu^{(i)} - E_s\left|\mu^{(i)}\right|^2. \quad (4)$$

---

[4]We are aware of the fact that the estimate $\hat{\mathbf{s}}_w$ could be computed without forming the explicit inverse $\mathbf{A}_w^{-1}$, e.g., via the Cholesky decomposition combined with forward/backward substitution [26]. However, soft-output detection as performed here requires the explicit inverse $\mathbf{A}_w^{-1}$ to compute the post-equalization SINR (see Section II-B).

We emphasize that both parameters $\mu^{(i)}$ and ν_i^2 are functions of $\mathbf{A}_w^{-1}$, $\forall w$. Consequently, an explicit computation of the inverses $\mathbf{A}_w^{-1}$, $\forall w$, is necessary for the computation of LLR values using the approach detailed above.

III. APPROXIMATE MMSE DETECTION VIA NEUMANN SERIES EXPANSION

The computation of all per-subcarrier inverses $\mathbf{A}_w^{-1}$, $\forall w$, in (1) is responsible for the main computational complexity of linear MMSE detection in SC-FDMA-based large-scale MIMO systems. For a conventional small-scale LTE uplink scenario, i.e., where the number of receive antennas B and users U is small (on the order of $U, B \leq 6$), existing VLSI designs for linear detection, such as [28]–[30], compute the exact inverse explicitly. For large-scale MIMO systems with a large number of users U however, the computation of the inverse $\mathbf{A}_w^{-1}$ can quickly result in excessive complexity. Hence, practical solutions for large-scale MIMO detection in LTE necessitate low-complexity matrix inversion methods—a corresponding approximate solution is proposed next.

A. Neumann Series Approximation

For large-scale MIMO systems, where the number of receive antennas is larger than the number of single-antenna users, i.e., for $U \ll B$, the Gram matrices $\mathbf{G}_w$, and, consequently $\mathbf{A}_w$, become diagonally dominant [9]. In fact, for i.i.d. Gaussian channel matrices $\mathbf{H}_w$ (with properly normalized entries) and in the large antenna limit, [8] shows that $\mathbf{G}_w \to \mathbf{I}_U$. Inspired by this central property of large-scale MIMO, one can derive a low-complexity approximation of the inverse. In particular, let $\mathbf{A}_w \approx \mathbf{D}_w$, where $\mathbf{D}_w$ is the main diagonal of $\mathbf{A}_w$. As a result, the inverse $\mathbf{A}_w^{-1}$ can be approximated by $\mathbf{D}_w^{-1}$, which requires evidently much lower complexity than that of the exact inverse. Unfortunately, for realistic antenna/user configurations, such a crude approximation would cause a significant performance loss. Hence, to arrive at an accurate approximation of the inverse at low computational complexity, we propose to use a Neumann series expansion.

We start by rewriting the inverse $\mathbf{A}_w^{-1}$ with the following Neumann series expansion [31]:

$$\mathbf{A}_w^{-1} = \sum_{n=0}^{\infty} \left(\mathbf{X}^{-1} \left(\mathbf{X} - \mathbf{A}_w \right) \right)^n \mathbf{X}^{-1}, \qquad (5)$$

which holds if $\lim_{n\to\infty}(\mathbf{I} - \mathbf{X}^{-1}\mathbf{A}_w)^n = \mathbf{0}_{U \times U}$ is satisfied. By decomposing the regularized Gram matrix $\mathbf{A}_w$ such that $\mathbf{A}_w = \mathbf{D}_w + \mathbf{E}_w$, where $\mathbf{D}_w$ is the main diagonal of $\mathbf{A}_w$ and $\mathbf{E}_w$ is the hollow, regularized Gram matrix, we can rewrite the Neumann series in (5) as

$$\mathbf{A}_w^{-1} = \sum_{n=0}^{\infty} (-\mathbf{D}_w^{-1}\mathbf{E}_w)^n \mathbf{D}_w^{-1}, \qquad (6)$$

where we substitute $\mathbf{X}$ in (5) by $\mathbf{D}_w$. Note that if $\lim_{n\to\infty}(-\mathbf{D}_w^{-1}\mathbf{E}_w)^n = \mathbf{0}_{U \times U}$, then the series expansion in (6) is guaranteed to converge.

The key idea of the proposed approximate inversion method is to keep only the first K terms of the Neumann series (6). Concretely, we compute a K-term approximation as follows:

$$\widetilde{\mathbf{A}}_{w\,|\,K}^{-1} = \sum_{n=0}^{K-1} (-\mathbf{D}_w^{-1}\mathbf{E}_w)^n \mathbf{D}_w^{-1}, \qquad (7)$$

which can be computed at low computational complexity for approximations consisting of only a few Neumann series terms, i.e., for small values of K. With this approximation, the resulting *approximate* MMSE equalization matrix is given by $\widetilde{\mathbf{W}}_{w\,|\,K} = \widetilde{\mathbf{A}}_{w\,|\,K}^{-1}\mathbf{H}_w^H$. For $K = 1$, we obtain $\widetilde{\mathbf{A}}_{w\,|\,1}^{-1} = \mathbf{D}_w^{-1}$, which is simply a scaled version of the MF detector, as $\widetilde{\mathbf{W}}_{w\,|\,1}^{-1} = \mathbf{D}_w^{-1}\mathbf{H}_w^H$. We emphasize that the row-wise scaling induced by $\mathbf{D}_w^{-1}$ does not affect the detection process, as long as $\mathbf{D}_w^{-1}$ exists. Hence, the proposed approximation (7) simply coincides with the MF detector for $K = 1$. For $K = 2$, we obtain $\widetilde{\mathbf{A}}_{w\,|\,2}^{-1} = \mathbf{D}_w^{-1} - \mathbf{D}_w^{-1}\mathbf{E}_w\mathbf{D}_w^{-1}$, whose computational complexity only scales with $O(U^2)$ operations; this is in contrast to the $O(U^3)$ complexity scaling required by computing an exact inverse. Hence, a second-order Neumann series approximation can be obtained at lower computational complexity. For $K = 3$, we obtain

$$\widetilde{\mathbf{A}}_{w\,|\,3}^{-1} = \mathbf{D}_w^{-1} - \mathbf{D}_w^{-1}\mathbf{E}_w\mathbf{D}_w^{-1} + \mathbf{D}_w^{-1}\mathbf{E}_w\mathbf{D}^{-1}\mathbf{E}_w\mathbf{D}_w^{-1}, \quad (8)$$

whose complexity scales with $O(U^3)$, which is equivalent to that of an exact inverse. Nevertheless, evaluating (8) requires fewer arithmetic operations than an explicit evaluation of $\mathbf{A}^{-1}$. Note that for $K \geq 4$, computing the exact inverse can be of lower complexity than the proposed approximation, e.g., when using a Cholesky factorization (see Section III-D).[5]

B. Analysis of the Approximation Error

We next analytically characterize the error induced by the approximate inverse (7) for MMSE estimation. To this end, we define the approximation error as $\boldsymbol{\Delta}_{w\,|\,K} = \mathbf{A}_w^{-1} - \widetilde{\mathbf{A}}_{w\,|\,K}^{-1}$, which is equivalent to

$$\boldsymbol{\Delta}_{w\,|\,K} = \sum_{n=K}^{\infty} (-\mathbf{D}_w^{-1}\mathbf{E}_w)^n \mathbf{D}_w^{-1}$$
$$= \left(-\mathbf{D}_w^{-1}\mathbf{E}_w\right)^K \sum_{n=0}^{\infty} (-\mathbf{D}_w^{-1}\mathbf{E}_w)^n \mathbf{D}_w^{-1}$$
$$= \left(-\mathbf{D}_w^{-1}\mathbf{E}_w\right)^K \mathbf{A}_w^{-1}.$$

Now, consider the situation of using the approximate $\widetilde{\mathbf{A}}_{w\,|\,K}^{-1}$ in place of $\mathbf{A}_w^{-1}$ to compute the equalized frequency-domain symbols, i.e.,

$$\hat{\mathbf{s}}_{w\,|\,K} = \widetilde{\mathbf{A}}_{w\,|\,K}^{-1}\mathbf{H}_w^H\mathbf{y}_w = \mathbf{A}_w^{-1}\mathbf{y}_w^{\text{MF}} - \boldsymbol{\Delta}_{w\,|\,K}\mathbf{y}_w^{\text{MF}}$$

with $\mathbf{y}_w^{\text{MF}} = \mathbf{H}_w^H\mathbf{y}_w$ and $\hat{\mathbf{s}}_w = \mathbf{A}_w^{-1}\mathbf{y}_w^{\text{MF}}$ being the exact estimate. We can bound the ℓ_2-norm of the *residual estimation*

---

[5]For approximations with $K = 2^n$ terms and $n \geq 2$, efficient ways of evaluating (7) exist. In particular, a clever re-arrangement and factorization of terms yields solutions which only require $2(n - 1)$ matrix multiplications.

error resulting from this approximate equalization by

$$
\begin{aligned}
\|\mathbf{\Delta}_{w\,|\,K}\mathbf{y}_w^{\mathrm{MF}}\|_2 &= \|(-\mathbf{D}_w^{-1}\mathbf{E}_w)^K\mathbf{A}_w^{-1}\mathbf{y}_w^{\mathrm{MF}}\|_2 \\
&\le \|(-\mathbf{D}_w^{-1}\mathbf{E}_w)^K\|_F\|\mathbf{A}_w^{-1}\mathbf{y}_w^{\mathrm{MF}}\|_2 \\
&\le \|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^K\|\hat{\mathbf{s}}_w\|_2.
\end{aligned}
\tag{9}
$$

From (9), we see that if the condition

$$
\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F < 1
\tag{10}
$$

is satisfied, then the approximation error approaches zero exponentially fast as $K \to \infty$. Moreover, one can show that (10) is a sufficient condition for (6) to converge.

We now show that the condition $\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F < 1$ is satisfied with high probability for large-scale MIMO systems with a larger number of BS antennas B than users U, and if the entries of $\mathbf{H}_w \in \mathbb{C}^{B\times U}$ are assumed to be i.i.d. circularly symmetric complex Gaussian with unit variance. More specifically, we arrive at a condition that only depends on U and B for (i) the proposed Neumann series to converge and (ii) the residual approximation error (9) to be small. The following theorem, which is proven in Appendix A, makes this behavior explicit.

Theorem 1. *Let $B > 4$ and the entries of $\mathbf{H}_w \in \mathbb{C}^{B\times U}$ be i.i.d. circularly symmetric complex Gaussian with unit variance. Then, we have*

$$
\begin{aligned}
&\Pr\big\{\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^K < \alpha\big\} \\
&\ge 1 - \frac{(U^2-U)}{\alpha^{\frac{2}{K}}}\sqrt{\frac{2B(B+1)}{(B-1)(B-2)(B-3)(B-4)}}.
\end{aligned}
\tag{11}
$$

We emphasize that this theorem provides conditions[6] for which the Neumann series converges with a certain probability; this can be accomplished by setting $\alpha = 1$ and $K = 1$ and by inspecting the convergence condition (10). Furthermore, Theorem 1 provides conditions for which the residual estimation error (9) is small. In both cases, we can see from Theorem 1 that increasing the ratio between the number of BS antennas B and the number of users U increases the probability of convergence. Moreover, for $\alpha < 1$, increasing K also increases the probability that the residual estimation error caused by a K-term approximation in (9) is smaller than α.

We note that Theorem 1 also provides insight into the behavior in the large-antenna limit, i.e., for $B \to \infty$ while U is held constant. In this case, we have $\Pr\big\{\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^K < \alpha\big\} \to 1$ for $\alpha \in (0,1]$, which implies (i) that the Neumann series converges with probability 1 and (ii) that the approximation error for any K-term approximation is arbitrary small, which includes the MF detector (corresponding to $K = 1$). We note that this behavior is in accordance with existing results for MF detection in large-scale MIMO systems [8], [32].

---

[6]The result in (11) also holds for the case where the regularization term $N_0 E_s^{-1}$ vanishes, which coincides to ZF detection. As a consequence, the condition (11) is rather pessimistic and is likely to be sub-optimal, especially for $N_0 E_s^{-1} > 0$. The derivation of a tighter condition is left for future work.

C. Channel Gain and NPI Variance Computation

Computation of the max-log LLRs via the proposed Neumann series approximation is carried out by simply replacing the exact inverse $\mathbf{A}_w^{-1}$ by the approximation $\tilde{\mathbf{A}}_{w\,|\,K}^{-1}$ to perform MMSE equalization (2). For this approximation, the effective channel gain $\tilde{\mu}_K^{(i)}$ and the variance of the residual post-equalization NPI variance $\tilde{\nu}_{i\,|\,K}^2$ now depend on the number of Neumann series terms.

In order to compute the effective channel gain $\tilde{\mu}_K^{(i)}$, we first construct the $LU \times LU$ matrix $\widetilde{\mathbf{W}}_{\cdot\,|\,K}^{-1}$ from the sub-carrier equalization matrices $\widetilde{\mathbf{W}}_{w\,|\,K}^{-1}, \forall w$, as explained in Section II-B. With this, we have $\tilde{\mu}_K^{(i)} x_t^{(i)} = \mathsf{E}\big[\mathbf{f}_t^H\widetilde{\mathbf{W}}_K^{(i,:)}\mathbf{y}\,|\,x_t^{(i)}\big]$, which can be rewritten as in (3) by replacing $\mathbf{W}^{(i,:)}$ with $\widetilde{\mathbf{W}}_{\cdot\,|\,K}^{(i,:)}$. Consequently, the effective channel gain is given by $\tilde{\mu}_K^{(i)} = L^{-1}\sum_{w=1}^L \widetilde{\mathbf{w}}_{i,w\,|\,K}^H\mathbf{h}_{i,w}$, where $\widetilde{\mathbf{w}}_{i,w\,|\,K}^H$ is the ith row of $\widetilde{\mathbf{W}}_{w\,|\,K}^{(i,i)}$ and $\mathbf{h}_{i,w\,|\,K}$ the ith column of $\mathbf{H}_w^{(i,i)}$.

In order to compute the post-equalization NPI variance $\tilde{\nu}_{i|K}^2$ one might assume that it simply corresponds to $E_s\tilde{\mu}_K^{(i)} - E_s|\tilde{\mu}_K^{(i)}|^2$ as in (4). Unfortunately, this expression no longer holds, because of the following fact:

$$
\widetilde{\mathbf{W}}_{\cdot\,|\,K}(E_s\mathbf{H}\mathbf{H}^H + N_0\mathbf{I}_{LB}) \ne \mathbf{H}^H.
$$

Furthermore, the above NPI variance expression is not guaranteed to be non-negative and hence, using it to compute LLR values inevitably results in poor error-rate performance. As a consequence, an alternative expression for $\tilde{\nu}_{i|K}^2$ is required when using the approximate matrix inverse for data detection. Following the steps of the derivation of $\tilde{\nu}_i^2$ in Section II-B and by replacing $\mathbf{W}^{(i,:)}$ with $\widetilde{\mathbf{W}}_{\cdot\,|\,K}^{(i,:)}$, the exact post-equalization NPI variance can be expressed as:

$$
\begin{aligned}
\tilde{\nu}_{i|K}^2 =&\,\mathbf{f}_t^H\widetilde{\mathbf{W}}_{\cdot\,|\,K}^{(i,:)}(E_s\mathbf{H}\mathbf{H}^H + \dots \\
&N_0\mathbf{I}_{LB})(\widetilde{\mathbf{W}}_{\cdot\,|\,K}^{(i,:)})^H\mathbf{f}_t - E_s\Big|\tilde{\mu}_K^{(i)}\Big|^2.
\end{aligned}
\tag{12}
$$

Since $\mathbf{H}^H(E_s\mathbf{H}\mathbf{H}^H + N_0\mathbf{I}_{LB}) = (E_s\mathbf{H}^H\mathbf{H} + N_0\mathbf{I}_{LU})\mathbf{H}^H$, we have:

$$
\tilde{\nu}_{i|K}^2 = E_s\mathbf{f}_t^H(\tilde{\mathbf{A}}_{\cdot\,|\,K}^{-1})^{(i,:)}\mathbf{A}\mathbf{G}(\tilde{\mathbf{A}}_{\cdot\,|\,K}^{-1})^{(i,:)}\mathbf{f}_t - E_s\Big|\tilde{\mu}_K^{(i)}\Big|^2.
$$

As $(\tilde{\mathbf{A}}_{\cdot\,|\,K}^{-1})^{(i,i)}$ is diagonal, we can decompose the computation as the sum of per-subcarrier operations. To this end, let $\tilde{\mathbf{a}}_{i,w|K}^H$ be the ith row of $\widetilde{\mathbf{A}}_{w|K}^{-1}$, then:

$$
\tilde{\nu}_{i|K}^2 = E_s\sum_{w=1}^L \tilde{\mathbf{a}}_{i,w|K}^H\mathbf{A}_w\mathbf{G}_w\tilde{\mathbf{a}}_{i,w|K} - E_s\Big|\tilde{\mu}_K^{(i)}\Big|^2.
\tag{13}
$$

This expression, however, is computational intensive, as it involves the L matrix multiplications, each requiring $O(U^3)$ operations. In order to reduce the complexity of computing $\tilde{\nu}_{i\,|\,K}^2$, we can use the $K = 1$ term approximation NPI

$$
\tilde{\nu}_{i|1}^2 = E_s\sum_{w=1}^L (d_w^{(i,i)})^{-2}\mathbf{a}_{i,w}^H\mathbf{g}_{i,w} - E_s\Big|\tilde{\mu}_1^{(i)}\Big|^2
\tag{14}
$$

as a substitute for $\tilde{\nu}_{i\,|\,K}^2$. Here, $d_w^{(i,i)}$ is the ith diagonal entry of $\mathbf{D}_w$, $\mathbf{a}_{i,w}^H$ is the ith row of $\mathbf{A}_w$, and $\mathbf{g}_{i,w}$ is the ith column

of $\mathbf{G}_w$. This approximation requires low computational complexity as it involves only L inner products, each requiring U operations. In addition, the larger K is, the closer the approximate inversion in (7) is to the exact inverse (assuming the Neumann series converges). Hence, for $K > 1$, the exact NPI variance would be lower than $\tilde{\nu}_{i\,|\,K}^2$, which reveals that (14) is a pessimistic approximation.

We emphasize that we can further reduce the computational complexity of the NPI approximation in (14). In particular, let $a_w^{(i,j)}$ be the ith entry of the vector $\mathbf{a}_{j,w}$ and $g_w^{(i,j)}$ be the ith entry of the vector $\mathbf{g}_{j,w}$. Since $\mathbf{A}_w = \mathbf{G}_w + N_0 E_s^{-1} \mathbf{I}_U$, we have the following identity:

$$(d_w^{(i,i)})^{-2} \mathbf{a}_{i,w}^H \mathbf{g}_{i,w} = (d_w^{(i,i)})^{-2} a_w^{(i,i)} g_w^{(i,i)} + \sum_{j, i \neq j} (a_w^{(i,j)})^H g_w^{(i,j)}.$$

Since (i) $a_w^{i,j} = g_w^{i,j}, \forall i \neq j$, (ii) $d_w^{(i,i)} = a_w^{(i,i)}$, and (iii) $d_w^{(i,i)} \gg a_w^{(i,j)}$ in the case where $U \ll B$, we can use the approximation $(d_w^{(i,i)})^{-2} \mathbf{a}_{i,w}^H \mathbf{g}_{i,w} \approx (d_w^{(i,i)})^{-1} g_w^{(i,i)}$. Hence, we propose the following low-complexity NPI approximation:

$$\tilde{\nu}_i^2 \approx E_s \sum_{w=1}^{L} (d_w^{(i,i)})^{-1} g_w^{(i,i)} - E_s \left| \tilde{\mu}_1^{(i)} \right|^2. \tag{15}$$

Note that our own simulations show that the low-complexity NPI approximation (15) performs well compared to the exact NPI variance (13). For example, the performance loss caused by the approximation compared to the exact NPI computation for $U = 4$, $B = 8$, and $K = 3$ is less than $0.02\,\text{dB}$ at a BLER of 10^{-2} (cf. Section III-D2 for the simulation settings).

D. Simulation Results

We next demonstrate the advantages and limitations of the proposed approximate matrix inversion approach in terms of computational complexity and error-rate performance. To assess the error-rate performance for practically relevant antenna configurations, we note that the Samsung Full-Dimensional MIMO prototype [33] consists of 64 BS antennas, whereas the massive MIMO research platform developed at Rice University [34] currently consists of 96 BS antennas (with plans for larger array sizes). Hence, we focus our results on the following cases: $B = 64$, $B = 128$, and $B = 256$.

1) Computational complexity: To demonstrate that the proposed approximate inverse exhibits (often significantly) lower complexity than an exact inverse, we chose a Cholesky decomposition-based inverse as a reference (see Section IV-E for algorithm details), as this method exhibits lower complexity compared to other inversion algorithms, including (but not limited to) direct matrix inversion, QR decomposition, or LU factorization [14], [26]. The computational complexity (characterized by the sum of real-valued division[7], addition, and multiplication[8] operations) of an exact Cholesky-based inverse scales with $O(U^3)$, whereas the complexity of a $K = 1$ and $K = 2$ Neumann series expansion scales only with $O(U)$

---

[7]The number of divisions is not significant for the total operation count.

[8]To obtain the real-valued multiplication count, we assumed four real-valued multiplications per one complex-valued multiplication. One could further reduce the number of the real-valued multiplications by using strength-reduction; this approach, however, maintains the trends observed in Fig. 1.
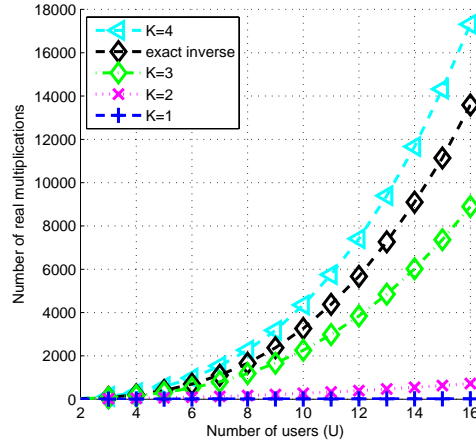

Fig. 1. Number of real-valued multiplications depending on the number of users U. The proposed approximation with $K \leq 3$ requires substantially lower complexity than that of an exact inverse based on the Cholesky decomposition.

and $O(U^2)$, respectively. The computational complexity of $K \geq 3$ is dominated by matrix-by-matrix multiplications, where the number of such operations grows linearly with K. For example, $K = 3$ requires one matrix-by-matrix multiplication, whereas $K = 4$ requires two. In general, a $K \geq 3$ term approximation requires $K - 2$ matrix-by-matrix multiplications. As a result, the complexity of a $K \geq 3$ term approximation is $O((K-2)U^3)$. Hence, we have $O(U^3)$ for $K = 3$, which is equivalent to that of an exact Cholesky-based inverse. Consequently, a Neumann series approximation with $K \geq 3$ does not appear to be advantageous.

The overall operation counts of both methods are dominated by the number of real-valued multiplications and additions, where real-valued multiplication is more expensive than real-valued addition. Since asymptotic complexity scalings do not, in general, reveal the full truth, we count the number of real-valued multiplications of both methods in Fig. 1 for varying numbers of users U. We observe that for $K \leq 3$, the Neumann series approach results in substantially lower complexity than the exact inversion approach. As expected, $K \geq 4$ results in higher complexity than a Cholesky-based exact inversion.

2) Error-rate performance: Evidently, the reduction in complexity for $K \leq 3$ Neumann series terms comes at the cost of an approximation error (cf. Section III-B). To characterize the associated performance loss, we now compare the error-rate performance of the proposed approximate matrix inverse with the error-rate performance of the exact inversion for an LTE-based large-scale MIMO uplink system. To this end, we show simulation results of an SC-FDMA LTE uplink system with B antennas at the BS and $U \leq B$ single-antenna users. In particular, we study a challenging communication scenario (from an error-rate perspective) and focus on the MCS (modulation and coding scheme) of the highest rate (i.e., MCS 28) and 20 MHz bandwidth with 1200 subcarriers, as specified by the LTE standard [4]; this mode corresponds to 64-QAM, and a rate ≈ 0.75 3GPP LTE turbo code. In order to generate

channel matrices that reflect a potential[9] real-world scenario, we use the WINNER-Phase-2 model [35]. In addition, we assume a linear antenna array with an antenna spacing of $10/128 \approx 0.0781\,\text{m}$, which resembles that of the real-world channel measurement campaign in [36]. At the BS, we use the exact and approximate soft-output MMSE detectors detailed above. Furthermore, we use a log-MAP LTE turbo decoder that performs 16 (full-)iterations. Further, we define signal-to-noise-ratio (SNR) as BE_s/N_0, which corresponds to the average SNR per receive antenna.

Figures 2(a), 2(b), and 2(c) show the block-error rate (BLER) performance of the proposed approximate detection algorithm compared to that of an exact MMSE detector for $U = 4$, $U = 8$ and $U = 12$, respectively.

We see that for small ratios between BS antennas and users, the MF detector (equivalent to $K = 1$) and the Neumann series approximation for $K = 2$ result in large residual errors.[10] Hence, considering the 10% BLER requirement for LTE [4], the MF detector and $K = 2$ term approximation are not suitable in practice in the considered 64-QAM cases (note that this fact is also reflected by Theorem 1). For a larger number of BS antennas, this error floor can be recovered partially. Our own simulations have shown that the MF detector achieves $< 10^{-2}$ BLER for $U = 4$ and $B = 512$. Furthermore, for 16-QAM, our approximation method requires smaller values of K (see [1], [2] for corresponding 16-QAM simulations in a large-scale MIMO-OFDM setting).

We see that for 64-QAM, the proposed approximate inversion method with $K = 3$ terms is able to approach the performance of the exact detector, i.e., the BLER performance loss is less than $0.25\,\text{dB}$ SNR at 10^{-2} BLER in all of the $K = 3, U = 4$ cases and the $K = 3, U = 8, B = 256$ case. Hence, the proposed approximate inverse for $K = 3$ can deliver the performance of an exact inversion at (often substantially) lower complexity for large ratios between BS antennas and users. For small antenna ratios, however, the approximate inverse with $K = 3$ exhibits an error floor.

We conclude that systems with small ratios between BS and user antennas will need to resort to an exact inverse, while systems with large ratios can take advantage of the proposed approximate inverse. Hence, we next propose corresponding MIMO detection architectures for both, the approximate inverse and an exact Cholesky-based inverse.

IV. VLSI ARCHITECTURE

We now detail two VLSI architectures suitable for large-scale MIMO detection in 3GPP LTE-A. The first design implements the proposed approximate inversion approach and the second design implements an exact inverse; this enables us to perform a fair hardware complexity vs. error rate performance comparison (see Section V for the comparison).
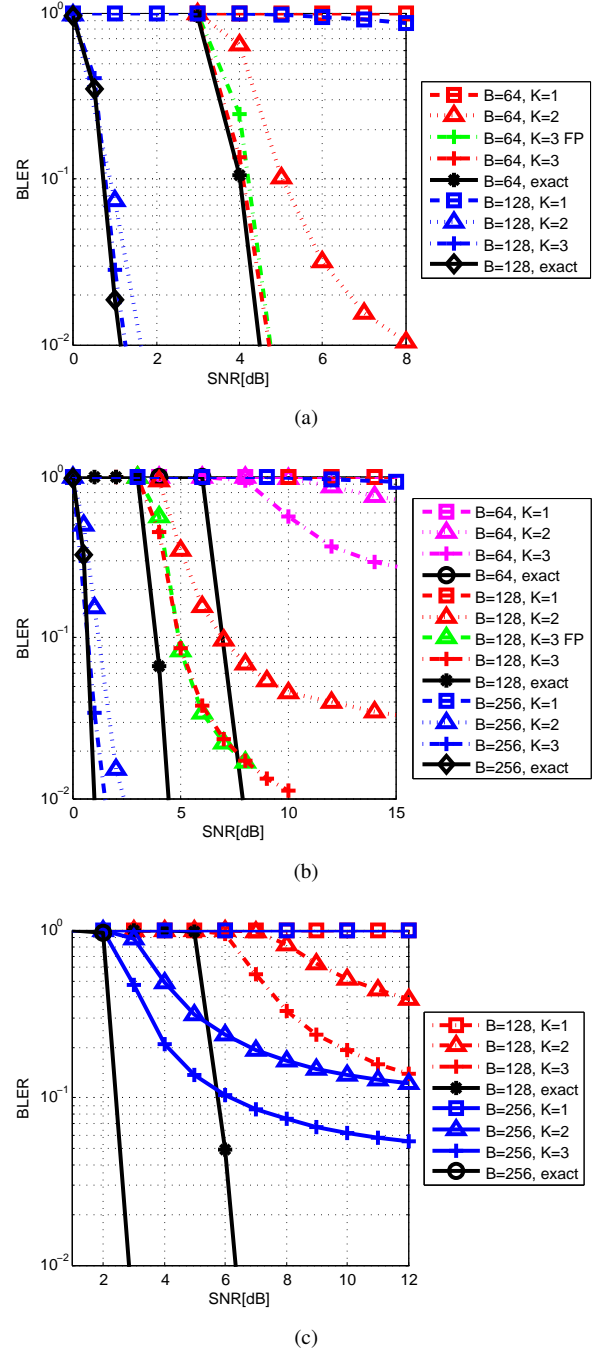

Fig. 2. Block error-rate (BLER) performance comparison for (a) $U = 4$ (b) $U = 8$, and (c) $U = 12$ single-antenna users where $M = 64$ and MCS = 28; 'FP' designates the performance of a fixed-point implementation.

A. Architecture Overview

The proposed general architecture is depicted in Fig. 3 and consists of the following parts. The preprocessing unit performs matched filter computation, i.e., computes $\mathbf{y}_w^{\text{MF}} = \mathbf{H}_w^H \mathbf{y}_w$, the regularized Gram matrix, and the (approximate) inverse. Note that for the approximate inversion unit, we also output $\mathbf{D}_w^{-1}$ and $\mathbf{G}_w$, which are needed to compute the SINR (cf. Section III-C). To achieve the peak throughput specified in LTE-A [6], while being able to handle the (worst) case where the channel estimates change from subcarrier to

---

[9]To the best of our knowledge, no specific channel model for large-scale MIMO systems is available in the open literature.

[10]Compared to lower modulation orders, such as 16-QAM (not shown here), 64-QAM requires a relatively high SNR to perform well.
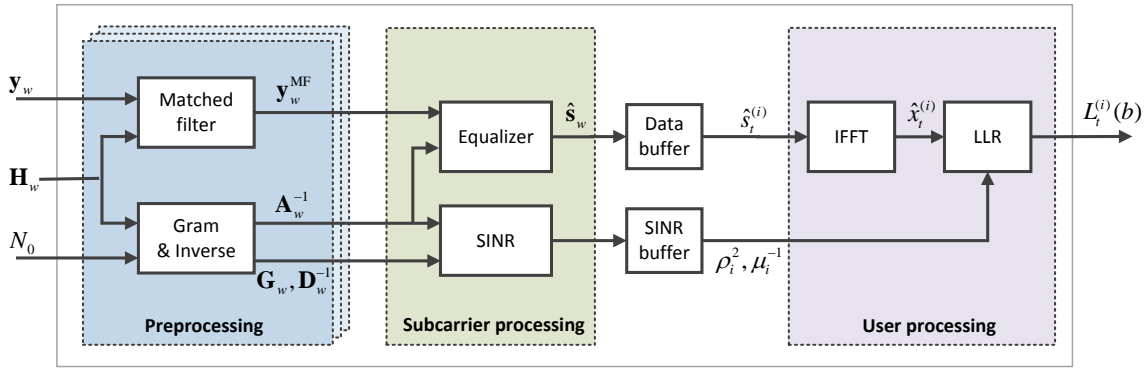
Fig. 3.   High-level VLSI architecture of the large-scale MIMO detection engine for 3GPP LTE-A.

subcarrier and from SC-FDMA symbol to SC-FDMA symbol (see, e.g, [37]), we use multiple instances of the preprocessing unit.[11] The matched filter output, the (approximate) inverse, and the regularized Gram matrix, are then passed to the subcarrier processing unit. This unit performs equalization, i.e., computes $\hat{\mathbf{s}}_w = \mathbf{A}_w^{-1} \mathbf{y}_w^{\mathrm{MF}}$ and the post-equalization SINR (detailed in Section II-B for the exact inverse and in Section III-C for the Neumann series approximation). To perform per-user data detection, a buffer is required that aggregates all equalized symbols and SINR values, which are computed on a per-subcarrier basis. The architecture then performs an IFFT, which transforms the equalized symbols from the subcarrier domain into the user domain (or time domain). The LLR computation unit finally computes, together with the buffered post-equalization NPI values, soft-output information in the form of max-log LLRs (2). We next provide the details for the key blocks of the proposed detector architecture.

B. Approximate Inversion and Matched Filter Units

1) Approximate inverse computation: In order to achieve high throughput, we propose a *single* systolic array that computes both, the regularized Gram matrix and the approximate inverse in four phases. The proposed architecture is detailed in Fig. 4 and is capable of computing inverses for various K-term expansions, i.e., the number of Neumann series terms can be selected at run-time. As shown in Fig. 4, the lower triangular systolic array consists of two distinct processing elements (PEs): (i) PEs on the main diagonal of the systolic array (referred to as PE-D) and PEs on the off-diagonal (referred to as PE-OD). As detailed next, both PEs have different modes in the four computation phases.

In the first phase, the $U \times U$ normalized regularized Gram matrix $\mathbf{A}_w/B = (\mathbf{G}_w + N_0 E_s^{-1} \mathbf{I}_U)/B$ is computed in B clock cycles. Since $\mathbf{A}_w$ is diagonally dominant with diagonal entries close to B, i.e., the number of BS antennas, we reduce its dynamic range by computing a normalized version, whose entries on the main diagonal are close to 1 by the 'scale down' unit shown in Fig. 4; this trick mitigates dynamic-range issues, which are common for matrix inversion circuits

[11]In many practical scenarios, the channel estimates may change only slowly. Hence, one does not need to compute the inverse for every SC-FDMA symbol. This fact could be either exploited to reduce the power consumption or to increase the achievable throughput of our detector designs.

implemented with fixed-point arithmetic. The systolic array also computes $\mathbf{D}_w^{-1} B$ from the diagonal entries of $\mathbf{A}_w/B$. These entries are computed in reciprocal units (denoted by 'inv' in Fig. 4) residing in the PE-D units. The results $\mathbf{D}_w^{-1} B$ and $\mathbf{E}_w/B$ are then stored in register files distributed in the systolic array.

In the second phase, the systolic array computes $-\mathbf{D}_w^{-1} \mathbf{E}_w$, by using the matrices $\mathbf{D}_w^{-1} B$ and $\mathbf{E}_w/B$ computed in the first phase. Since the matrix $-\mathbf{D}_w^{-1} \mathbf{E}_w$ is not Hermitian, the systolic array computes the upper- and lower-triangular parts of $-\mathbf{D}_w^{-1} \mathbf{E}_w$ separately. As $\mathbf{D}_w^{-1}$ is a diagonal matrix, computation of $-\mathbf{D}_w^{-1} \mathbf{E}_w$ only requires a series of scalar multiplications (rather than a matrix multiplication).

In the third phase, the systolic array computes the $K = 2$ term Neumann series approximation, i.e., $\tilde{\mathbf{A}}_{w\,|\,2}^{-1} B = (\mathbf{D}_w^{-1} B - \mathbf{D}_w^{-1} \mathbf{E}_w \mathbf{D}_w^{-1} B)$. To this end, it is important to realize that the matrix $\mathbf{D}_w^{-1} B - \mathbf{D}_w^{-1} \mathbf{E}_w \mathbf{D}_w^{-1} B$ is Hermitian, implying that only the lower triangular part needs to be computed. Furthermore, since $\mathbf{D}_w^{-1} B$ is diagonal, computation of $-\mathbf{D}_w^{-1} \mathbf{E}_w \mathbf{D}_w^{-1} B$ only requires entry-wise multiplications (instead of costly matrix multiplications). These scalar multiplications are carried out by loading $\mathbf{D}_w^{-1} B$ and $-\mathbf{E}_w \mathbf{D}_w^{-1}$ into all PEs and performing a scalar multiplication to compute $\mathbf{D}_w^{-1} \mathbf{E}_w \mathbf{D}_w^{-1} B$. Then, we add $\mathbf{D}_w^{-1} B$ to the result in the diagonal PEs. The result of this phase, i.e., $\mathbf{D}_w^{-1} B - \mathbf{D}_w^{-1} \mathbf{E}_w \mathbf{D}_w^{-1} B$, is stored in the distributed register files.

In the fourth phase, the K-term Neumann series approximation is computed with the results residing in the distributed register files. In particular, the systolic array first performs a matrix multiplication of $-\mathbf{D}_w^{-1} \mathbf{E}_w$ with $\tilde{\mathbf{A}}_{w\,|\,K-1}^{-1} B$, and then adds $\mathbf{D}_w^{-1} B$ to the diagonal PE. The resulting K-term approximation $\tilde{\mathbf{A}}_{w\,|\,K}^{-1} B$ is then stored in the register files. This phase can be repeated for a configurable number of iterations, which allows us to compute an arbitrary K-term approximation.

2) Matched filter computation: The matched filter (MF) unit consists of a linear array of U PEs. Each PE is associated with one row of the Hermitian matrix $\mathbf{H}_w^H$, and contains a single multiply accumulate unit (MAC) and a scaling unit to normalize the result to $\mathbf{y}_w^{\mathrm{MF}}/B$. The MF unit reads a new entry of $\mathbf{y}_w$ every clock cycle, and multiplies it with the corresponding entries in $\mathbf{H}_w^H$ in each PE and then, adds it the previous results; the final result is then normalized by $1/B$.
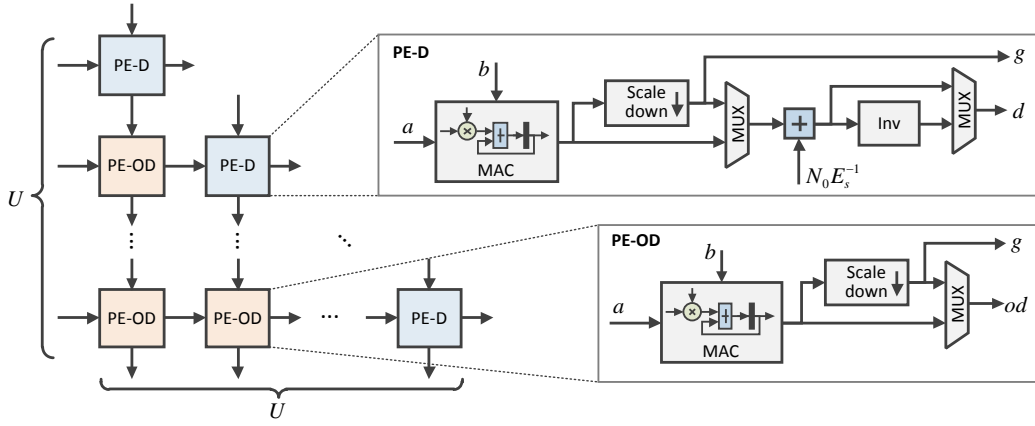
Fig. 4. Architecture details of the Gram matrix computation and approximate matrix inversion unit. The lower-triangular systolic array shown on the left consists of two processing elements (PEs); their architectural details are shown on the right.

C. Equalization and SINR Computation Units

1) Equalization unit: The equalization unit consists of a linear array of U MAC units, and reads the normalized approximate inverse $\widetilde{\mathbf{A}}_{w\,|\,K}^{-1}B$ and the $\mathbf{y}_w^{\mathrm{MF}}/B$ from the matched filter unit. For each clock cycle, this unit takes one column of $\widetilde{\mathbf{A}}_{w\,|\,K}^{-1}B$, multiplies it with one element from $\mathbf{y}_w^{\mathrm{MF}}/B$, and adds the scaled column to the previous results. The unit outputs an equalized symbol $\hat{\mathbf{s}}_w$ every U clock cycles.

2) SINR computation unit: The SINR computation unit simply consists of U MAC units that sequentially compute the approximate effective channel gain $\tilde{\mu}_K^{(i)}$. This unit furthermore computes the approximate NPI (15) using a single MAC unit. Subsequently, the unit multiplies $\tilde{\mu}_K^{(i)}$ with the reciprocal of the approximate NPI $\tilde{\nu}_i^2$ to obtain the post-equalization SINR ρ_i^2. The same unit computes the reciprocal of $\tilde{\mu}_K^{(i)}$ which is used in the LLR computation unit detailed next.

D. IFFT and LLR Computation Units

1) IFFT unit: In order to transform the per-subcarrier data into the user (or time) domain, we deploy a single Xilinx Discrete Fourier Transform IP LogiCORE unit (see [38] for the specifications). This unit supports all forward and inverse DFT modes specified in 3GPP LTE [4], but we only make use of its IDFT capabilities. The IFFT unit reads and outputs data in a serial manner. For an IFFT transform size of 1200 subcarriers, the core can process a new set of data every 3779 clock cycles. This FFT unit achieves more than 317 MHz on a Virtex-7 XC7VX980T FPGA and hence, achieves a throughput beyond 600 Mb/s for 8 users, 64-QAM, and 20MHz bandwidth.

2) LLR computation unit: The LLR computation unit (LCU) generates max-log soft output values given the effective channel gains $\mu^{(i)}$ from the IFFT block and the post-equalization SINR values ρ_i^2 obtained from the SINR block. Since LTE specifies Gray mappings for all modulation schemes (BPSK, QPSK, 16-QAM, and 64-QAM), one can simplify the computation of the max-log LLR values in (2) by rewriting $L_t^{(i)}(b) = \rho_i^2 \lambda_b(\hat{x}_t^{(i)})$ and realizing that $\lambda_b(\cdot)$ is a piecewise linear function that depends on the bit index (see [25] for the details). To this end, the LCU first scales the real and imaginary parts of the equalized time-domain symbol

with the reciprocal of the effective channel gain $1/\mu^{(i)}$. Then, it evaluates the piecewise linear function $\lambda_b(\hat{x}_t^{(i)})$ and scales the result with the post-equalization SINR ρ_i^2. The resulting max-log LLR value is then delivered to the output of the unit. In order to minimize the circuit area, the proposed architecture evaluates each piecewise linear function with logical shifts and additions only. The reciprocals are computed with a lookup table that is stored in B-RAM units (see [1] for architectural details). A single instance of the resulting LCU is able to processes one symbol every clock cycle, resulting in a peak throughput of 1.89 Gb/s for 64-QAM at 317 MHz.

E. Reference Cholesky-based Inversion Unit

In order to enable a fair performance/complexity assessment of the proposed approximate matrix inversion unit, we also implemented a reference unit that performs an exact matrix inversion. This unit simply replaces the approximate inverse unit detailed in Section IV-B. We next summarize the used Cholesky-based inversion algorithm and then, outline the corresponding VLSI architecture.

1) Inversion algorithm: In the proposed exact inversion unit, we compute $\mathbf{A}_w^{-1}$ in three steps: (i) we form the regularized Gram matrix $\mathbf{A}_w = \mathbf{G}_w + N_0 E_s^{-1} \mathbf{I}_U$; (ii) we perform a Cholesky decomposition according to $\mathbf{A}_w = \mathbf{L}_w \mathbf{L}_w^H$, where $\mathbf{L}_w$ is a lower-triangular matrix with real-values on the main diagonal [26]; (iii) we compute the inverse $\mathbf{A}_w^{-1}$ using an efficient forward/backward substitution procedure proposed in [25]. Specifically, we first solve $\mathbf{L}_w \mathbf{u}_i = \mathbf{e}_i$ for $\mathbf{u}_i$, $i = 1, \ldots, U$, where $\mathbf{e}_i$ is the ith unit vector, via forward substitution. We then solve $\mathbf{L}_w^H \mathbf{v}_i = \mathbf{u}_i$ for $\mathbf{v}_i$, $i = 1, \ldots, U$, via back substitution, which leads to the desired inverse $\mathbf{A}_w^{-1} = [\mathbf{v}_1 \cdots \mathbf{v}_U]$. Note that this approach avoids a costly matrix-by-matrix multiplication, which would be needed by directly computing $\mathbf{A}_w^{-1} = (\mathbf{L}_w^H)^{-1} \mathbf{L}_w^{-1}$.

2) Cholesky decomposition architecture: The VLSI architecture for the Cholesky-based inverse differs from the one in Section IV-B. In particular, we deploy three separate units that compute (i) the regularized Gram matrix, (ii) the exact inverse using the above algorithm, and (iii) a forward/backward substitution unit to compute the inverse $\mathbf{A}_w^{-1}$. All units are detailed next and separated by pipeline stages.

The regularized Gram matrix is computed as a sum of outer products, i.e., as $\mathbf{G}_w = \sum_{i=1}^{B} \mathbf{r}_i \mathbf{r}_i^H$, where $\mathbf{r}_i$ designates the ith row of $\mathbf{H}_w$. Since the Gram matrix is symmetric, it can be computed efficiently with a triangular systolic array of multiply and accumulate units (MACs), similar to the array detailed in Section IV-B. The Gram computation unit reads one row of $\mathbf{H}_w$ at a time and is able to output a Gram matrix every B^{th} clock cycle. To obtain the regularized Gram matrix $\mathbf{A}_w$, we add $N_0 E_s^{-1}$ to the diagonal of $\mathbf{G}_w$ in the final clock cycle.

We then perform the Cholesky decomposition of $\mathbf{A}_w$ with a lower-triangular systolic array to obtain the lower-triangular matrix $\mathbf{L}_w$. The systolic array consists of two distinct processing elements (PEs): (i) the PEs on the main diagonal and (ii) the PEs on the off-diagonal. The data flow is similar to the linear systolic array (the "obvious case") proposed in [39]. The difference is that our design processes an incoming column of $\mathbf{A}_w$ with multiple PEs, whereas an incoming column is processed with a single PE in [39]. As a result, our design is able to achieve the peak throughput requirements of LTE-A. In our design, the pipeline of one column of PEs is 16 stages deep and streams out one column of $\mathbf{L}_w$ every clock cycle (after a latency of $16(U-1)$ clock cycles). Consequently, the achieved throughput corresponds to one Cholesky decomposition every U clock cycles.

3) Forward/backward-substitution architecture: The forward/backward substitution unit (FBSU) receives a lower-triangular matrix $\mathbf{L}_w$ as input, and computes $\mathbf{A}_w^{-1} = (\mathbf{L}_w^H)^{-1} \mathbf{L}_w^{-1}$ as outlined in Section IV-E1. The FBSU consists of three major components: (i) a forward substitution unit (FSU), which solves for $\mathbf{L}_w \mathbf{u}_i = \mathbf{e}_i$, (ii) a backward substitution unit (BSU), which solves for $\mathbf{L}_w^H \mathbf{v}_i = \mathbf{u}_i$, and (iii) a Hermitian transpose unit, which computes $\mathbf{L}_w^H$. Since the computations for the FSU and the BSU are symmetric, we implement the forward substitution architecture and re-use it for the backward substitution, by reversing the order of the columns of the matrix $\mathbf{L}_w^H$ and vector $\mathbf{u}_i$ before reading them into the BSU. To simplify notation, we assume that the equation to be solved by the forward substitution corresponds to $\mathbf{L}\mathbf{x} = \mathbf{b}$ for some $\mathbf{x}$ and $\mathbf{b}$. Since the forward substitution of solving the equation $\mathbf{L}\mathbf{x}_i = \mathbf{b}_i$ for each $\mathbf{b}_i$ $(i = 1, \dots, U)$ is independent, we use U processor elements (PEs) to solve for all $\mathbf{x}_i$ in parallel. Each PE is implemented using a fully pipelined architecture, which consists of U stages of computation logic. Each stage contains two multiplexers, a complex-valued multiplier, and a complex-valued subtraction. In each stage, either $\Delta_i = b_i - \sum_j L_{i,j} x_j$ or $\Delta_i / L_{i,i}$ is computed according to the control signals. Therefore, for an input matrix $\mathbf{L}_w$ of dimension U, the FSU uses U^2 complex-valued multipliers; the entire FBSU utilizes $2U^2$ complex-valued multipliers. The matrix conjugate unit is implemented using multiplexers and U FIFOs (realized by on-chip B-RAMs in the FPGA). The conjugate matrix $\mathbf{L}_w^H$ is also reordered based on the pattern of the input sequence of the BSU.

V. IMPLEMENTATION RESULTS AND TRADE-OFFS

The approximate detection engine for 3GPP-LTE and the exact Cholesky-based detector have been implemented on a

TABLE I
IMPLEMENTATION RESULTS ON A XILINX VIRTEX-7 XC7VX980T FPGA

Antenna configuration[a] Inversion algorithm[b]	128×8		64×4	
	$K = 3$	Cholesky	$K = 3$	Cholesky
Clock frequency [MHz]	317	317	317	317
Throughput [Mb/s]	603	603	301	301
LUT slices	168 125 (28%)	208 161 (34%)	34 631 (6%)	78 756 (12.9%)
FF slices	193 451 (16%)	213 226 (17.4%)	39 492 (3.2%)	39 602 (3.2%)
DSP48 units	1 059 (30%)	1 447 (40.2%)	233 (7%)	329 (9.14%)
Block RAMs	18 (0.6%)	65 (2.17%)	12 (0.4%)	32 (1.07%)

[a] 128×8 refers to $B = 128$ BS antennas and $U = 8$ single-antenna users.
[b] $K = 3$ designates the approximate inversion with 3 Neumann series terms.

Xilinx Virtex-7 XC7VX980T FPGA. The fixed-point parameters, FPGA implementation results, and the associated performance/complexity trade-offs are presented next.

A. Fixed-Point Design Parameters

In order to minimize the hardware complexity, fixed-point arithmetic is used in the entire design. The associated fixed-point parameters were determined via extensive simulations. In the following, the word-lengths refer to the real or imaginary part of a complex-valued number.

The channel matrices $\mathbf{H}_w$, the receive-vectors $\mathbf{y}_w$, and the noise variance $N_0 E_s^{-1}$, are all quantized to 15 bit. The word-length of the output of the Gram matrix and inversion unit are also set to 15 bit; equivalently, the matched filter unit has 15 bit at the input and output. For both matrix inversion circuits, all multiplications have been mapped onto Xilinx DSP48 slices. In order to achieve sufficient precision at minimum implementation complexity, the MAC registers within the DSP48 units are set to 22 bit. The LUT in the reciprocal unit consists of 1024 addresses with 12 bit outputs. Hence, it can be implemented efficiently using a single block-RAM (B-RAM) available on the FPGA. The equalizer module uses a 15 bit input and its output, which is stored in the data buffer, is quantized to 12 bit. The buffer stores (complex-valued) data for 1200 subcarriers and U users. The SINR computation module has a 15 bit input and 12 bit output. The input and output of the IFFT unit are 12 bit; the precision of the internal multipliers is set to 18 bit. The inputs of the LLR computation are quantized to 12 bit and the computed LLRs are represented by 8 bit.

The resulting fixed-point performance is shown in Fig. 2 (labeled by 'FP') for 64×4 and 128×8 systems. As it can be seen, the fixed-point implementation is virtually indistinguishable from the floating-point golden model. In particular, the implementation loss is less than 0.05 dB SNR at 10% BLER.

B. FPGA Implementation Results

Table I summarizes the key (post-place-and-route) implementation results of the proposed approximate and exact soft-output data detector for LTE-based massive MIMO wireless

systems. We parameterized the architecture for U and B to explore the impact on the required FPGA resources and the corresponding throughput. The implementation results for antenna configurations of 128×8 and 64×4 are detailed in Table I. In order to support $75\,\mathrm{Mb/s}$ data rate for each LTE-A user in $20\,\mathrm{MHz}$ bandwidth, we use multiple instances of the preprocessing unit. Specifically, we used 8 and 5 instances of approximate matrix inversion units for the 128×8 and 64×4 system, respectively. For the exact inverse, we used 6 and 3 regularized Gram matrix units for the 128×8 and 64×4 system, respectively. In addition, we used one Cholesky decomposition unit and one forward and backward substitution unit for both cases to meet the data rate requirements.

As shown in Table I, all designs are capable of running at $317\,\mathrm{MHz}$ and the critical path is the routing between different blocks of the detector. For the 128×8 and 64×4 systems, the proposed units can achieve $603\,\mathrm{Mb/s}$ and $301\,\mathrm{Mb/s}$, respectively. For the 64×4 system, the design meets the $300\,\mathrm{Mb/s}$ peak data rate requirement specified in LTE-A with 4 users and 20MHz bandwidth. In addition, our design can scale beyond LTE-A specifications, i.e., the proposed designs can support up to 8 users and still achieve a $75\,\mathrm{Mb/s}$ per-user requirement.

In terms used resources on the Virtex-7 XC7VX980T FPGA, the approximate soft-output data detector is smaller than the Cholesky-based unit. There are notable saving in logic slices and DSP48 units. For 64×4, $K = 3$ uses 56% fewer LUT slices and 29% fewer DSP48 units compared to that of the Cholesky-based unit. For 128×8, $K = 3$ uses 19% fewer LUT slices and 26% fewer DSP48 units compared to that of the Cholesky-based unit. We emphasize that the savings in hardware resources become significantly larger as the number of users U increases.

C. Performance/Complexity Trade-off

Based on the simulated BLER results in Fig. 2 and the associated FPGA implementation results, we are now ready to characterize the error-rate performance vs. hardware complexity trade-offs associated with the detector containing the proposed approximate matrix inversion and the Cholesky-based exact inversion. To this end, we show the associated hardware complexity against the minimum SNR required to achieve 10% BLER in Fig. 5. Since both designs are dominated by multipliers, we define the hardware complexity as the number of multipliers required to achieve a $75\,\mathrm{Mb/s}$ per-user throughput.

From Fig. 5, we observe that the hardware complexity of the Cholesky-based detector is larger than that of the approximate inversion circuit for $K = 3$ and $K = 2$. In addition, for large ratios between the number of BS antennas to the number of users B/U, we clearly see that the SNR performance of the approximate inverse with $K = 3$ and the exact inverse are very similar. For small ratios B/U, however, the performance difference between the approximate inverse and the exact inverse is rather large, which is reflected in the analysis shown Section III-B. Hence, the ratio B/U determines whether an approximate or exact inversion is beneficial in a practical large-scale MIMO system. Note that for 128×8 and 64×4, the approximate inverse with $K = 2$ is unable to achieve 10%
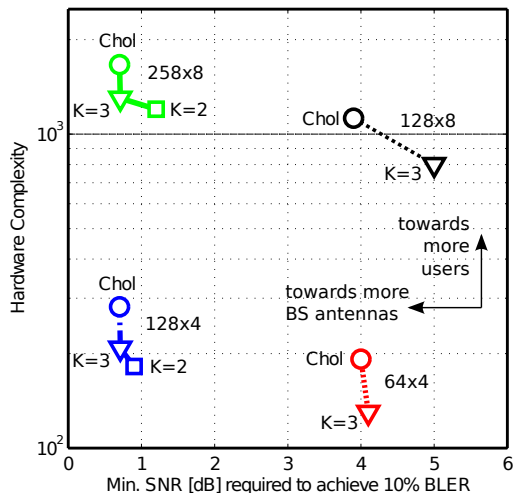

Fig. 5. Performance/complexity trade-off. Hardware complexity is defined as the number of DSP48E1 slices required to achieve the LTE-A uplink 75 Mb/s per-user peak throughput.

BLER (cf. Fig. 2). We note that when considering 16-QAM modulation (rather than 64-QAM modulation, as shown here), the approximate inversion for $K = 2$ is capable of achieving similar performance as the exact inverse (see [1], [2] for corresponding simulation results).

D. Related FPGA Designs for Linear Data Detection

A host of FPGA designs for linear data detection in conventional (small-scale) MIMO systems have been proposed in the literature [28]–[30], [40]–[44]. Unfortunately, all these designs differ in various ways. First, the corresponding architectures rely on different matrix inversion algorithms, such as the QR decomposition [29], [40], [43], [44], Gram-Schmidt orthogonalization [30], [45], LU decomposition [25], direct matrix inversion [42], divide-and-conquer methods [41], [46]. Second, all FPGA implementations do not generate soft outputs, with the exception of [47]. Third, the designs were implemented on different FPGA types.

Since the soft-output detector implementations proposed in this paper are for large-scale MIMO systems having hundreds of BS antennas and none of the small-scale MIMO detector designs in [28], [29], [40]–[46] was implemented on a Xilinx Virtex-7 FPGA, a fair comparison of our design with the above-mentioned implementations is difficult. Hence, we decided to resort to the comparison with our own reference circuit, i.e., the Cholesky-based inverse, as shown in Section V-C.

VI. CONCLUSIONS

We have proposed a new soft-output data detector for large-scale (or massive) MIMO-based 3GPP LTE-Advanced (LTE-A) systems. The proposed solution is capable of performing high throughput detection in single-carrier frequency division multiple access (SC-FDMA)-based large-scale MIMO systems equipped with hundreds of antennas at the base station (BS). In order to achieve low computational complexity, we have proposed a new approximate linear detector relying on a Neumann series approximation of the matrix inverse.

We have designed two reference VLSI architectures, one relying on the approximate inverse, the other on an exact Cholesky-based matrix inversion. Both architectures have been successfully implemented on a state-of-the-art Xilinx Virtex-7 FPGA, are suitable for systems equipped with 128 BS antennas or fewer while serving up to 8 users, and achieve more than $600\,\text{Mb/s}$, exceeding the peak data rates specified in the 3GPP LTE-A uplink for $20\,\text{MHz}$ bandwidth. Our FPGA implementation results reveal that for systems with a large ratio between the number of BS antennas and the number of users, the approximate matrix inversion is able to significantly reduce the hardware implementation complexity (compared to that of the exact inversion) with only a slight error-rate performance degradation. For systems with small ratios between the number of BS antennas and the number of users (as it is the case in, e.g., conventional, small-scale MIMO systems) one must resort to an exact inverse in order to avoid poor error-rate performance. This behavior is in accordance with the analytical results we have developed for the approximate matrix inverse. In summary, our FPGA implementation results demonstrate the practical feasibility of high-throughput data detection for 3GPP LTE-based large-scale MIMO systems. We finally note that a corresponding high-throughput ASIC design has recently been published in [48].

There are many avenues for future work. The development of detection algorithms that are able to perform iterative detection and decoding (as, e.g., in [25]) in large-scale MIMO systems is left for future work. Furthermore, the design of high-performance, near-optimal detection methods (e.g., based on the algorithms in [17], [49]) that require low computational complexity for large-dimensional antenna configurations and for SC-FDMA is a challenging open research problem.

APPENDIX A
PROOF OF THEOREM 1

To prove Theorem 1, we need the following three Lemmata.

Lemma 2. *Let the scalars $x^{(k)}$ and $y^{(k)}$ for $k = 1,\ldots,B$ be i.i.d. circularly symmetric complex Gaussian with unit variance. Then, $\mathsf{E}\left[\left|\sum_{k=1}^{B} x^{(k)} y^{(k)}\right|^4\right] = 2B(B+1)$.*

Proof: We have

$$
\mathsf{E}\left[\left|\sum_{k=1}^{B} x^{(k)} y^{(k)}\right|^4\right] = \mathsf{E}\left[\left(\sum_{k=1}^{B} x^{(k)} y^{(k)} \sum_{k=1}^{B}\left(x^{(k)} y^{(k)}\right)^*\right)^2\right]
$$
$$
= \binom{B}{2}\mathsf{E}\left[|x^{(k)}|^2|y^{(k)}|^2\right] + 4\mathsf{E}\left[|x^{(k)}|^4|y^{(k)}|^4\right]
$$
$$
= 2B(B-1) + 4B = 2B^2 + 2B.
$$

The above steps can be summarized as follows. After expanding the quadratic expression, the non-zero terms can be written as $|x^{(k)}|^4|y^{(k)}|^4$ and $|x^{(k)}|^2|y^{(k)}|^2$, where $k = 1,\ldots,B$. Then, there are B terms of the form $|x^{(k)}|^4|y^{(k)}|^4$ and $\binom{B}{2}$ of the form $|x^{(k)}|^2|y^{(k)}|^2$. The facts that $\mathsf{E}\left[|x^{(k)}|^4\right] = \mathsf{E}\left[|y^{(k)}|^4\right] = 2$ and $\mathsf{E}\left[|x^{(k)}|^2\right] = \mathsf{E}\left[|y^{(k)}|^2\right] = 1$ concludes the proof. ∎

Lemma 3. *Let $B > 4$ and $x^{(k)}$, $k = 1,\ldots,B$ be i.i.d. circularly symmetric complex Gaussian with unit variance and $g = \sum_{k=1}^{B} |x^{(k)}|^2$. Then,*

$$
\mathsf{E}\left[\left|g^{-1}\right|^4\right] = ((B-1)(B-2)(B-3)(B-4))^{-1}. \quad (16)
$$

Proof: We first rewrite g as $2^{-1}\sum_{k=1}^{2B}|s^{(k)}|^2$ where $s^{(k)}$, $k = 1,\ldots,2B$, are i.i.d. zero-mean real-valued Gaussian with unit variance. Then, $2g^{-1}$ is an inverse chi-square random variable with $2B$ degrees of freedom. The inverse chi-square distribution with $2B$ degrees of freedom $\chi(2B)$ corresponds to an inverse-Gamma distribution with $2B$ degrees-of-freedom. The 4$^{\text{th}}$ moment of this inverse chi-square distribution is given by $\frac{1}{16}(B-1)(B-2)(B-3)(B-4)$ [50] and, hence, we obtain (16). ∎

Lemma 4. *Let $B > 4$ and the entries of $\mathbf{H}_w \in \mathbb{C}^{B\times U}$ be i.i.d. circularly symmetric complex Gaussian with unit variance. Then, we have*

$$
\mathsf{E}\left[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\right]
$$
$$
\leq (U^2 - U)\sqrt{\frac{2B(B+1)}{(B-1)(B-2)(B-3)(B-4)}}
$$

Proof: The regularized Gram matrix corresponds to $\mathbf{A}_w = \mathbf{D}_w + \mathbf{E}_w = \mathbf{G}_w + N_0 E_s^{-1}\mathbf{I}_{U\times U}$. Thus, each element on the ith row and j^{th} column of $\mathbf{A}_w$, $a_w^{(i,j)}$ can be written as:

$$
a_w^{(i,j)} = \begin{cases} g_w^{(i,j)} = \sum_{k=1}^{B}\left(h_w^{(k,i)}\right)^* h_w^{(k,j)}, & i \neq j \\ g_w^{(i,i)} + N_0 E_s^{-1} = \sum_{k=1}^{B}\left|h_w^{(k,i)}\right|^2 + N_0 E_s^{-1}, & i = j, \end{cases}
$$

with $g_w^{(i,j)}$ corresponding to the ith row and j^{th} column of the Gram matrix $\mathbf{G}_w$. We now have the following inequality:

$$
\mathsf{E}\left[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\right] = \mathsf{E}\left[\sum_{i=1}^{i=U}\sum_{j=1,i\neq j}^{j=U}\left|\frac{g_w^{(i,j)}}{a_w^{(i,i)}}\right|^2\right]
$$
$$
\leq \sum_{i=1}^{i=U}\sum_{j=1,i\neq j}^{j=U}\mathsf{E}\left[\left|\frac{g_w^{(i,j)}}{g_w^{(i,i)}}\right|^2\right],
$$

which is obtained by omitting the non-negative regularization term $N_0 E_s^{-1}$. By applying the Cauchy-Schwarz inequality, we can bound $\mathsf{E}\left[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\right]$ from above as

$$
\mathsf{E}\left[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\right] \leq \sum_{i=1}^{i=U}\sum_{j=1,i\neq j}^{j=U}\sqrt{\mathsf{E}\left[\left|g_w^{(i,j)}\right|^4\right]\mathsf{E}\left[\left|\left(g_w^{(i,i)}\right)^{-1}\right|^4\right]}.
$$

Application of Lemmata 2 and 3 to the first and second expected values, respectively, we obtain

$$
\mathsf{E}\left[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\right] \leq \sum_{i=1}^{i=U}\sum_{j=1,i\neq j}^{j=U}\sqrt{\frac{2B(B+1)}{(B-1)(B-2)(B-3)(B-4)}}
$$
$$
= (U^2 - U)\sqrt{\frac{2B(B+1)}{(B-1)(B-2)(B-3)(B-4)}}.
$$
∎

We are now in position to prove Theorem 1. To this end, we start by using Markov's inequality to obtain the following straightforward inequality:

$$\Pr\big\{\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^K \geq \alpha\big\} = \Pr\Big\{\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2 \geq \alpha^{\frac{2}{K}}\Big\}$$
$$\leq \alpha^{-\frac{2}{K}}\mathsf{E}\big[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\big].$$

With $\Pr\big\{\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^K < \alpha\big\} = 1 - \Pr\big\{\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^K \geq \alpha\big\}$ and by using the upper bound for $\mathsf{E}\big[\|\mathbf{D}_w^{-1}\mathbf{E}_w\|_F^2\big]$ from Lemma 4, we finally obtain (11).

REFERENCES

[1] M. Wu, B. Yin, A. Vosoughi, C. Studer, J. R. Cavallaro, and C. Dick, "Approximate matrix inversion for high-throughput data detection in the large-scale MIMO uplink," in *Proc. IEEE ISCAS*, Beijing, China, May 2013, pp. 2155–2158.

[2] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and C. Dick, "Implementation trade-offs for linear detection in large-scale MIMO systems," in *Proc. IEEE ICASSP*, Vancouver, BC, May 2013, pp. 2679–2683.

[3] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. New York, USA: Cambridge University Press, 2008.

[4] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding (Release 9)*. 3GPP Organizational Partners TS 36.212 Rev. 8.3.0, May 2008.

[5] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley Publishing, 2009.

[6] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures (Release 10)*. 3GPP Organizational Partners TS 36.213 version 10.10.0, Jul. 2013.

[7] *IEEE Draft Standard Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 4: Enhancements for Higher Throughput*. P802.11n_D3.00, Sept. 2007.

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

[9] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.

[11] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3266–3239, Sept. 2012.

[12] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *arXiv preprint: 1112.3810v2*, May 2012.

[13] 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, 2002.

[14] A. Burg, "VLSI circuits for MIMO communication systems," Ph.D. dissertation, ETH Zürich, Switzerland, 2006.

[15] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.

[16] K. Wong, C. Tsui, R. Cheng, and W. Mow, "A VLSI architecture of a K-best lattice decoding algorithm for MIMO channels," in *IEEE ISCAS*, vol. 3, Scottsdale, AZ, May 2002, pp. 273–276.

[17] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, 2003.

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

[19] J. Jaldèn and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[20] D. Seethaler, J. Jaldén, C. Studer, and H. Bolcskei, "On the complexity distribution of sphere decoding," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5754–5768, Sept. 2011.

[21] T. Datta, N. Ashok Kumar, A. Chockalingam, and B. Sundar Rajan, "A novel MCMC algorithm for near-optimal detection in large-scale uplink mulituser MIMO systems," in *Proc. IEEE ITA*, San Diego, CA, Feb. 2012, pp. 69 –77.

[22] R. Prasad, *OFDM for Wireless Communications Systems*. Norwood, MA, USA: Artech House, Inc., 2004.

[23] "3GPP TS 36.211 Evolved Universal Terrestrial Radio Access (E-UTRA) physical channels and modulation (release 8)," 3rd Generation Partnership Project.

[24] F. Khan, *LTE for 4G mobile broadband: air interface technologies and performance*. Cambridge University Press, 2009.

[25] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, Jul. 2011.

[26] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins Univ. Press, 1996.

[27] M. P. Fossorier, F. Burkert, S. Lin, and J. Hagenauer, "On the equivalence between SOVA and max-log-MAP decodings," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 137–139, May 1998.

[28] A. Burg, S. Haene, D. Perels, P. Luethi, N. Felber, and W. Fichtner, "Algorithm and VLSI architecture for linear MMSE detection in MIMO-OFDM systems," in *Proc. IEEE ISCAS*, Island of Kos, Greece, May 2006, pp. 4102–4105.

[29] R. M. Rao, H. Tarn, R. Mazahreh, and C. Dick, "A low complexity square root MMSE MIMO decoder," in *Proc. 44th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2010, pp. 1463–1467.

[30] P. Luethi, C. Studer, S. Duetsch, E. Zgraggen, H. Kaeslin, N. Felber, and W. Fichtner, "Gram-Schmidt-based QR decomposition for MIMO detection: VLSI implementation and comparison," in *Proc. IEEE APCCAS*, Macao, China, Nov. 2008, pp. 830–833.

[31] G. Stewart, *Matrix Algorithms: Basic decompositions*, 1998.

[32] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *49th Ann. Allerton Conf. on Commun., Control., and Comput.*, Monticello, IL, Sept. 2011, pp. 545–550.

[33] *New SID Proposal: Study on Full Dimension MIMO for LTE*. 3GPP TSG RAN Meeting 58, Dec. 2012.

[34] C. Shepard, H. Yu, and L. Zhong, "ArgosV2: a flexible many-antenna research platform," in *Proc. 19th Annual International Conference on Mobile Computing and Networking (MobiCom)*. Miami, Florida, USA: ACM, 2013, pp. 163–166.

[35] L. Hentilä, P. Kyösti, M. Käske, M. Narandzic, and M. Alatossava. (2007, Dec.) Matlab implementation of the WINNER phase II channel model ver 1.1. [Online]. Available: https://www.ist-winner.org/phase_2_model.html

[36] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, "Channel measurements for large antenna arrays," in *Proc. IEEE ISWCS*, Aug. 2012.

[37] M. Simko, D. Wu, C. Mehlfuehrer, J. Eilert, and D. Liu, "Implementation aspects of channel estimation for 3GPP LTE terminals," in *Proc. 11th European Wireless Conference - Sustainable Wireless Technologies (European Wireless)*, Vienna, Austria, Apr. 2011, pp. 440–444.

[38] Xilinx Inc., "IP LogiCORE Discrete Fourier Transform," DFT IP for 3GPP LTE systems, DS615 v3.1, Mar. 2011.

[39] R. Schreiber and W.-P. Tang, "On systolic arrays for updating the Cholesky factorization," *BIT Numerical Mathematics*, vol. 26, no. 4, pp. 451–466, Dec. 1986.

[40] M. Myllyla, J. Hintikka, J. R. Cavallaro, M. Juntti, M. Limingoja, and A. Byman, "Complexity analysis of MMSE detector architectures for MIMO OFDM systems," in *Proc. 39th Asilomar Conf. on Signals, Systems and Computers*, Nov. 2005, pp. 75–81.

[41] S. Eberli, D. Cescato, and W. Fichtner, "Divide-and-conquer matrix inversion for linear MMSE detection in SDR MIMO receivers," in *Proc. 26th Norchip Conference*, Nov. 2008, pp. 162–167.

[42] D. Wu, J. Eilert, and D. Liu, "Implementation of a high-speed MIMO soft-output symbol detector for software defined radio," *J. Signal Process. Syst.*, vol. 63, no. 1, pp. 40–60, Apr. 2011.

[43] M. Karkooti, J. R. Cavallaro, and C. Dick, "FPGA implementation of matrix inversion using QRD-RLS algorithm," in *Proc. 44th Asilomar Conf. on Signals, Systems and Computers*, Nov. 2005, pp. 1625–1629.

[44] J. Cong, B. Liu, S. Neuendorffer, J. Noguera, K. Vissers, and Z. Zhang, "High-level synthesis for FPGAs: From prototyping to deployment," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 30, no. 4, pp. 473–491, Apr. 2011.

[45] H. S. Kim, W. Zhu, J. Bhatia, K. Mohammed, A. Shah, and B. Daneshrad, "An efficient FPGA based MIMO-MMSE detector," in *Proc. EUSIPCO*, Sept. 2007, pp. 1131–1135.

[46] J. Eilert, D. Wu, and D. Liu, "Implementation of a programmable linear MMSE detector for MIMO-OFDM," in *Proc. IEEE ICASSP*, Mar. 2008, pp. 5396–5399.

[47] D. Wu, J. Eilert, R. Asghar, and D. Liu, "VLSI implementation of a fixed-complexity soft-output MIMO detector for high-speed wireless," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, pp. 58:1–58:13, Apr. 2010.

[48] B. Yin, M. Wu, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, "A 3.8 Gb/s large-scale MIMO detector for 3GPP LTE-Advanced," in *Proc. IEEE ICASSP*, Florcence, Italy, May 2014.

[49] C. Studer and H. Bölcskei, "Soft–input soft–output single tree-search sphere decoding," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4827–4842, Oct. 2010.

[50] J. D. Cook, "Inverse Gamma distribution," online: http://www.johndcook.com/inverse_gamma.pdf, Tech. Rep., 2008.