LUND UNIVERSITY

Energy Efficient MIMO Channel Pre-processor Using a Low Complexity On-Line Update Scheme

Zhang, Chenxin; Prabhu, Hemanth; Liu, Liang; Edfors, Ove; Öwall, Viktor

Link to publication

Energy Efficient MIMO Channel Pre-processor Using a Low Complexity On-Line Update Scheme

Chenxin Zhang, Hemanth Prabhu, Liang Liu, Ove Edfors, and Viktor Öwall

Department of Electrical and Information Technology

Lund University, Box 118, SE-221 00 Lund, Sweden

Email: {Chenxin.Zhang, Hemanth.Prabhu, Liang.Liu, Ove.Edfors, Viktor.Owall}@eit.lth.se

Abstract—This paper presents a low-complexity energy efficient channel pre-processing update scheme, targeting the emerging 3GPP long term evolution advanced (LTE-A) downlink. Upon channel matrix renewals, the number of explicit QR decompositions (QRD) and channel matrix inversions are reduced since only the upper triangular matrices R and R^{-1} are updated, based on an on-line update decision mechanism. The proposed channel pre-processing updater has been designed as a dedicated unit in a 65 nm CMOS technology, resulting in a core area of $0.242\,\text{mm}^2$ (equivalent gate count of 116 K). Running at a 330 MHz clock, each QRD or R^{-1} update consumes 4 or 2 times less energy compared to one exact state-of-the-art QRD in open literature.

I. INTRODUCTION

Efficient MIMO signal detection relies on the knowledge of channel state information (CSI) and channel pre-processing. For instance, the QRD of channel matrices is a key prerequisite in tree-search based signal detectors such as sphere and K-Best decoders [1]. Channel matrix inversion is crucial to linear detectors such as MMSE [2]. However, despite their importance, channel pre-processing units are often excluded from conventional signal detectors [1], as their computations are considered to be less frequent due to the common assumption of block-stationary data transmissions [3]. Unfortunately, CSI of real-world radio channels is rarely constant because of Doppler induced channel changes and multi-path propagation. Outdated CSI introduces additional interferences to signal detections, which will drastically degrade MIMO performance. Thus, frequent CSI update and the corresponding channel pre-processing are highly desirable in wireless communication systems to provide signal detectors with adequate channel knowledge.

Using the channel's time correlation, tracking of CSI changes can be achieved using low-complexity decision-directed algorithms such as least mean square (LMS), recursive least square (RLS), and Kalman filtering [4] [5]. Nevertheless, continuous CSI tracking has not been widely adopted in practical systems, as each CSI update requires computational intensive channel pre-processing, either QRD or channel matrix inversion, which has comparable complexity to that of MIMO signal detectors [5]. Moreover, considering energy consumption, frequent CSI updates result in an increased power budget for channel pre-processing operations, which is not always affordable in practical systems, especially for portable devices. Although enormous efforts have been made on algorithms, architectures, and arithmetics for QRD and matrix inversion to mitigate computation load and to enhance processing throughput, few studies have addressed their energy efficiency. For instance, QRD of one 4×4 channel matrix may be accomplished in 4 clock cycles with power consumption of 318.66 mW, corresponding to 12.76 nJ energy at 100 MHz clock frequency [6]. Although this QRD scheme achieves 20% higher processing throughput than the state-of-the-art signal detector presented in [1], each decomposition consumes 10 times more energy than one MIMO signal detection.

To reduce energy consumption, QRD update schemes are proposed in open literature to reduce the frequency of brute-force QRD computations by using either LMS [5] or matrix perturbation algorithm [7]. Although those schemes involve low complexity QRD update operations, additional error evaluation criteria are as complex as the QRD itself, which is against the initial intention of the complexity and energy reduction.

To tackle the aforementioned issues, this paper proposes an energy-efficient update scheme for performing channel pre-processing operations upon CSI updates. Since matrix inversion can be efficiently computed via QRD and R^{-1} [6], we focus on those updates to serve both tree-search based and linear signal detectors. To the best of our knowledge, this is the first study in open literature targeting both QRD and R^{-1} update and has hardware implementation and energy figures reported. By taking advantage of time correlation between adjacent channels, exact tone-by-tone QRD computation during successive channel matrix updates can be avoided by holding unitary matrix Q fixed while only updating the upper triangular matrix R. For R^{-1} update, Neumann series approximation is adopted instead of computing exact matrix inversion. To minimize detection errors caused by outdated Q and R matrices, we evaluate the applicability of the update scheme prior to each QRD update. This is accomplished by applying a low-complexity on-line condition check between newly acquired channel matrices and their predecessors. The update scheme is executed only when the calculated condition suffices a certain update criterion (adjusted with respect to the channel scenario and desired design trade-off), otherwise, exact QRD is carried out. To investigate the effectiveness of the proposed algorithm, we conduct simulations based on an LTE-A downlink system and 3GPP channel models with moderate moving speed. Benefiting from LTE-A's scattered pilot allocations, the computational complexity of the on-line update condition evaluation is further reduced by half. Considering a 4x4 MIMO setup, we show that each QRD and R^{-1} update consumes 0.36 nJ and 0.72 nJ respectively, representing 4 and 2 times higher energy efficiency than one exact state-of-the-art QRD. Moreover, with a 50% update level, the performance degradation due to outdated Q-R and R^{-1} matrices is less than 1 dB and 2 dB respectively, when compared to the exact QRD and channel matrix inversion.

II. BACKGROUND

Considering a general baseband model of a spatial-multiplexing $N \times N$ MIMO system, the $N \times 1$ received complex signal vector y is expressed as

$$y = Hx + n, \qquad (1)$$

where x is the N-length transmit vector, n is the i.i.d. complex Gaussian noise vector with zero mean and variance σ^2, and H denotes the $N \times N$ propagation matrix containing uncorrelated complex-valued channel coefficients. Each component of x is obtained by
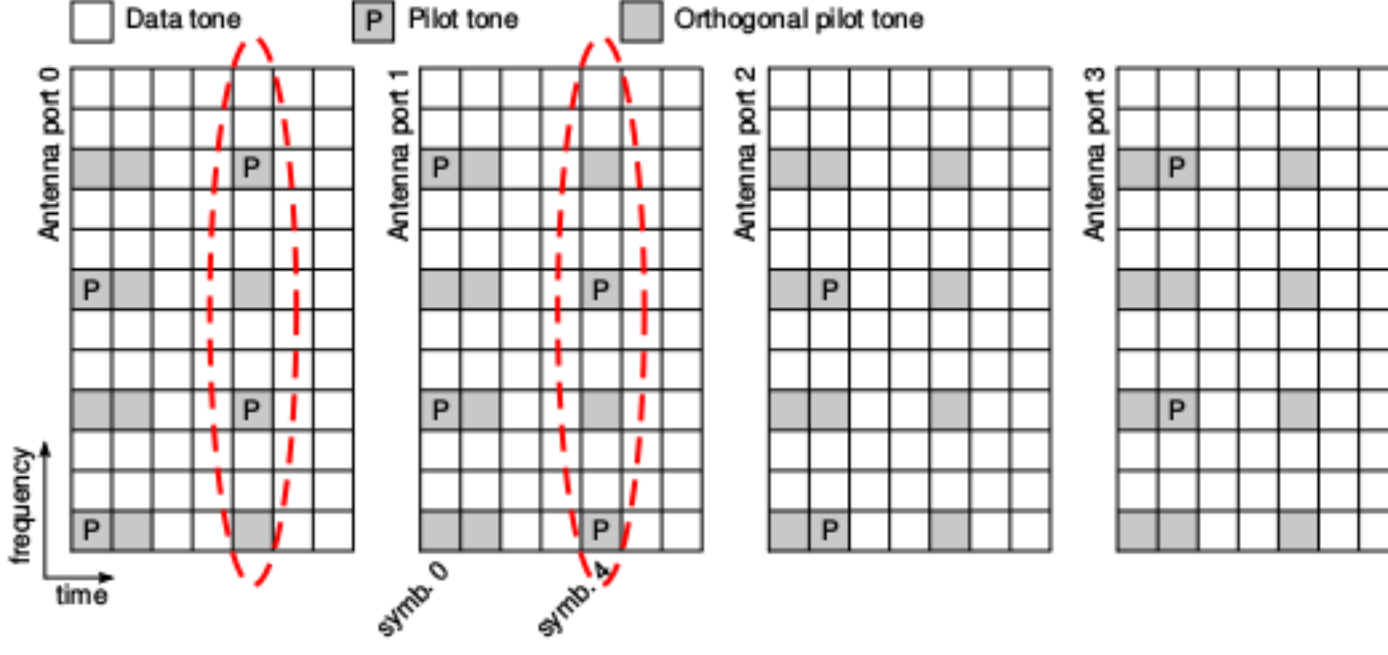
Fig. 1. Pilot pattern for four antenna ports in one LTE-A resource block. OFDM symbol positions for half-H renewals are circled in dashed lines.

Fig. 2. FER performance of different receiver behaviours to CSI updates.

mapping a set of information bits, encoded by error-correcting codes (ECC), onto a Gray-labelled complex constellation such as M-QAM. In this experiment, we use LTE-A downlink as a testbed and resort to frame-error-rate (FER) to evaluate the proposed algorithm. The ECC scheme adopted in simulations is a rate 1/2 parallel concatenated turbo code with coding block length of 5376 and decoding iteration 6. Furthermore, we assume the receiver has perfect channel knowledge.

A. Channel matrix update

To help receivers track frequent channel changes, LTE-A inserts scattered orthogonal pilot tones to multiple antenna ports. Pilot pattern for four antenna ports are sketched in a time-frequency grid, as shown in Fig. 1. One special property to notice, pilot tones allocated in the middle of each LTE-A resource block (RB) are only available for antenna ports 0 and 1. This corresponds to an update of first two columns in channel matrix H, denoted as *half-H renewal* hereafter. Benefiting from this property, the on-line condition check for QRD updates can be simplified as described further in Section III-A.

Before presenting channel pre-processing update schemes, different baseband processing responses to CSI changes are analyzed using 3GPP channel models with various mobile speeds. In this paper, we consider a RB as one basic unit for LTE-A data transmissions, and treat half-H renewals as CSI updates with respect to full-H updates at the beginning of each RB. Using QRD as a case study, *exact QRD update* refers to a case (Case-I) that has exact QRD performed during half-H renewals. In contrast, *no QRD update* case (Case-II) assumes the receiver stays in a static channel environment, thus has no operation applied during CSI updates. Utilizing K-Best signal detector with $K = 10$, FER performance for the aforementioned cases in a 4×4 MIMO downlink with 64-QAM modulation are illustrated in Fig. 2. Two channel models are used in the simulations: extended pedestrian A model with 5 Hz Doppler frequency (EPA-5) and extended vehicular A model with 70 Hz Doppler frequency (EVA-70). Operating at 2.6 GHz carrier frequency, these Doppler frequencies correspond to 2 km/h and 29 km/h moving speed, respectively. Interesting to note that Case-II achieves similar FER performance as Case-I in EPA-5 simulations, thanks to the high channel correlation caused by low mobility. In contrast, a significant performance degradation is observed with EVA-70 model when no operation is performed during CSI updates (Case-II). Channel models with higher mobility, such as extended typical urban model with 300 Hz Doppler frequency, are not considered in this study, as smaller MIMO configurations (e.g., 2×2) or lower modulation schemes (e.g., QPSK) are expected to be used in such channel scenarios to mitigate serious interferences induced by fast channel variations. The same analyses can be applied on R^{-1} update together with linear detectors such as MMSE, and similar results are obtained for the two cases. Hence, we can conclude that frequent QRD and R^{-1}
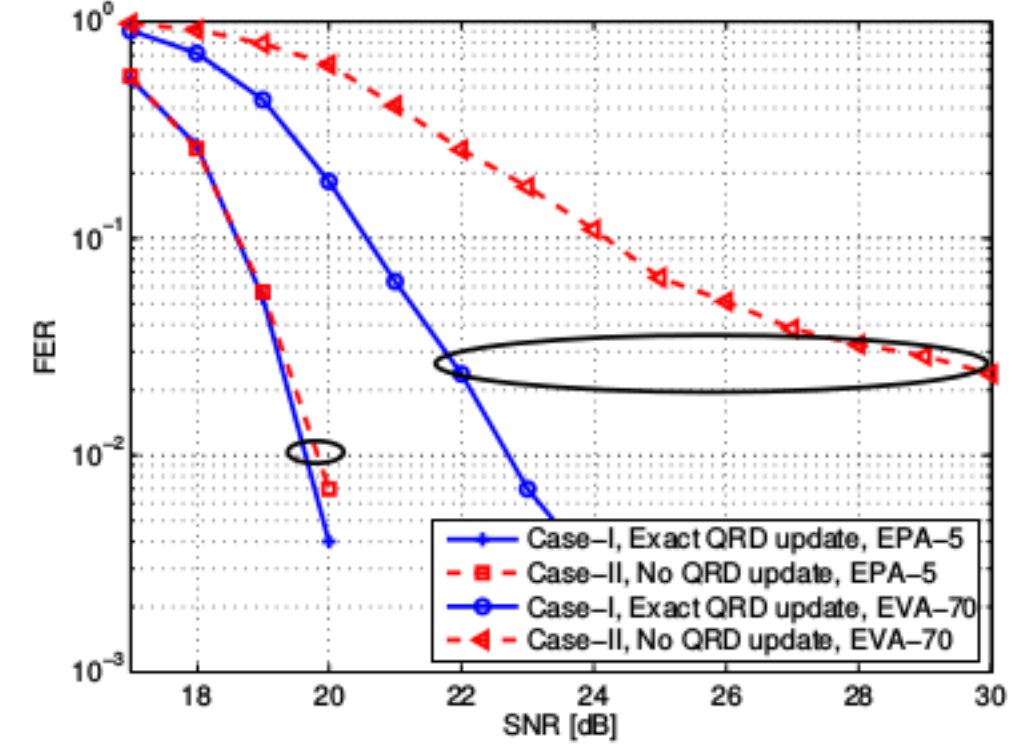
updates only lead to small contributions to FER improvements in static and slow moving channel scenarios. However, channel pre-processing updates need to be performed explicitly upon CSI changes in channel models with moderate and high mobile speed, in order to prevent performance degradation at the cost of computation power.

The above observations motivate us to propose an efficient channel pre-processing update scheme that can bridge the performance gap between Case-I and Case-II during CSI updates, and at the same time, have much lower computational complexity and energy consumption than performing the exact channel pre-processing operations (Case-I).

III. PROPOSED CHANNEL PRE-PROCESSING UPDATE SCHEME

In this section, we start by presenting the proposed update schemes in conjunction with an on-line decision mechanism. Thereafter, update schemes are evaluated using simulated FER and different update criteria are discussed in accordance with performance degradation.

A. QRD update scheme

In each channel matrix factorization, namely $H = QR$, the Q matrix contains orthogonalized unit (column) vectors, and R holds norm values of the corresponding columns in Q. By utilizing the time correlation property of H, we assume that 1) the orthogonality of column vectors in Q remains unchanged during successive CSI updates; and 2) any change in channel matrix is represented as norm value variations in R. With these assumptions, the demanding QRD computation can be simplified to updates of matrix R, expressed as

$$H_i = Q_i R_i \approx Q_{i-t} \hat{R}'_i, \quad t = 1 \ldots T, \qquad (2)$$

$$\hat{R}'_i = Q_{i-t}^H H_i, \qquad (3)$$

where i and t are time indexes of channel matrix updates (both full- and half-H renewals), and the constant T defines effective time interval for holding Q matrix fixed in between two QRD computations. Because of the outdated Q matrix, $\hat{R}'_i$ is an approximation of R_i and generally looses its upper triangular matrix form, which is not in favor of tree-search based signal detectors. Hence, post-processing of $\hat{R}'_i$ is needed to transform it back to the expected matrix form. This is accomplished by forcing non-zero imaginary elements on the main diagonal and non-zero entries below the main diagonal to zero. This post-processed $\hat{R}'_i$ matrix is hereafter denoted as $\hat{R}_i$.

Comparing to an exact QRD, the total error induced by the update scheme includes inaccuracies caused by the outdated Q matrix and zero forcing applied on $\hat{R}'_i$. To minimize these errors, an on-line condition check is applied prior to each QRD update. During this check, vector closeness between H_i columns and their predecessors in H_{i-t} is evaluated. QRD update scheme is applied only when the vector closeness is below a certain threshold value $\mathcal{T}$, otherwise exact QRD is performed. The closeness evaluation is in essential the
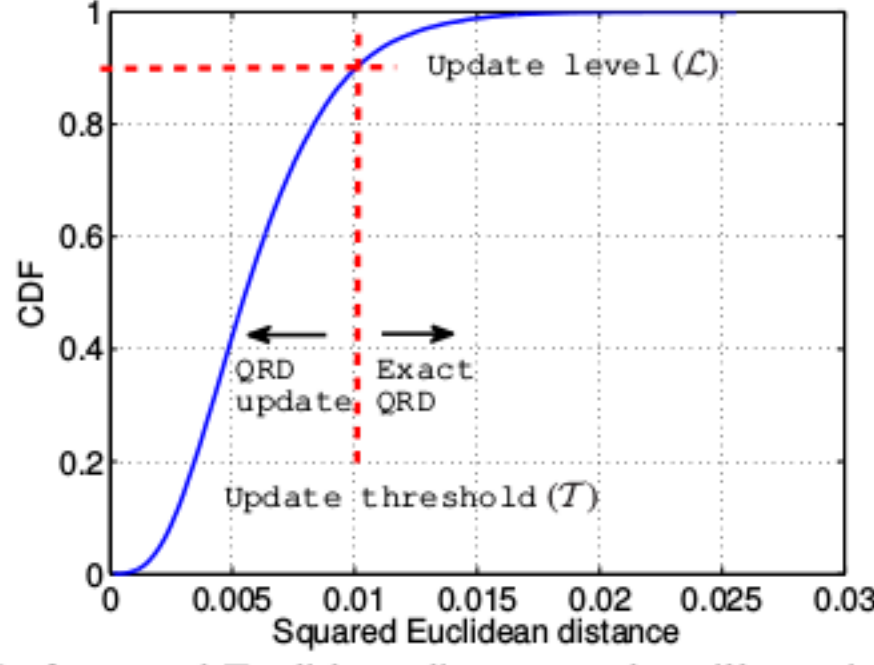
Fig. 3. CDF of squared Euclidean distance and an illustration of $\mathcal{L} = 90\%$ for 3GPP EVA-70 channel model.

calculation of maximum squared Euclidean distance (ED) between column vectors, formulated as

$$d^2 = \max \| H_i^n - H_{i-t}^n \|^2, \quad n = 1 \ldots N, \tag{4}$$

where n is the index of column vectors in H. The threshold value is determined by adjusting update level $\mathcal{L}$ in ED distributions (decided off-line and obtained by Monte-Carlo simulations), where $\mathcal{L}$ indicates an average ratio of QRD updates to total number of channel matrices during half-H renewals. As an illustration, the cumulative distribution function (CDF) of ED values simulated for 3000 EVA-70 channel realizations is shown in Fig. 3. Together with the CDF, $\mathcal{T}$ and $\mathcal{L}$ values for an 90% update level are marked. H matrices with ED values larger than $\mathcal{T}$ refer to the channel matrices that require exact QRDs. Selections of update levels $\mathcal{L}$ are elaborated in Section III-C accompanying with design trade-offs between performance degradation and complexity-energy reduction. Making use of the special pilot allocation in LTE-A (Section II-A), we choose to compute exact QRD in the beginning of each resource block and apply QRD update during half-H renewals. According to this setup, the effective time interval T in (2) is set to 1, and the ED calculations (4) are reduced to the first two columns only, i.e., $n = \{1, 2\}$.

B. R^{-1} update scheme

R^{-1} update is computed based on the result of QRD. By adopting Neumann series approximation, the explicit matrix inversion (from R to R^{-1}) is avoided. Given the fact that matrix $\hat{R}_i$ is an updated version of R_{i-1}, the convergence condition of the Neumann series is satisfied, expressed as

$$\lim_{k \to \infty} (I_N - R_{i-1}^{-1} \hat{R}_i)^k \approx 0_N. \tag{5}$$

Assuming the inverse of R_{i-1} is computed together with the exact QRD, R^{-1} update ($\hat{R}_i^{-1}$) can be calculated using Neumann series

$$\hat{R}_i^{-1} = \sum_{j=0}^{\infty} (I_N - R_{i-1}^{-1} \hat{R}_i)^j R_{i-1}^{-1}. \tag{6}$$

Since linear detectors do not utilize the property of upper triangular matrices, no post-processing is needed for $\hat{R}_i^{-1}$. For practical implementations, the infinite matrix summation in (6) needs to be truncated to a fixed term L, such as

$$\hat{R}_i^{-1} \approx \sum_{j=0}^{L} (I_N - R_{i-1}^{-1} \hat{R}_i)^j R_{i-1}^{-1}. \tag{7}$$

Setting $L = 1$, an inverse approximation can be written as

$$\hat{R}_i^{-1} \approx R_{i-1}^{-1} + (I_N - R_{i-1}^{-1} \hat{R}_i) R_{i-1}^{-1},$$
$$= (2I_N - R_{i-1}^{-1} \hat{R}_i) R_{i-1}^{-1}. \tag{8}$$

Considering the relatively low complexity, the first order Neumann series approximation (8) is used in this experiment to compute $\hat{R}_i^{-1}$.
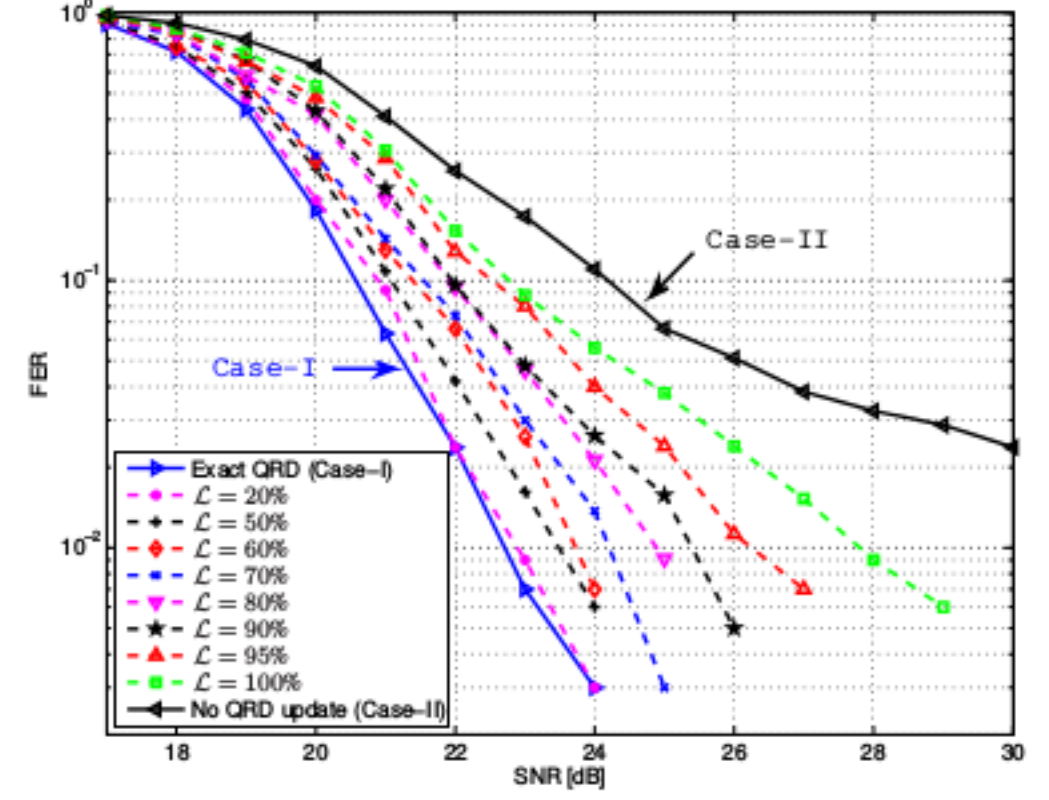
(a) QRD update
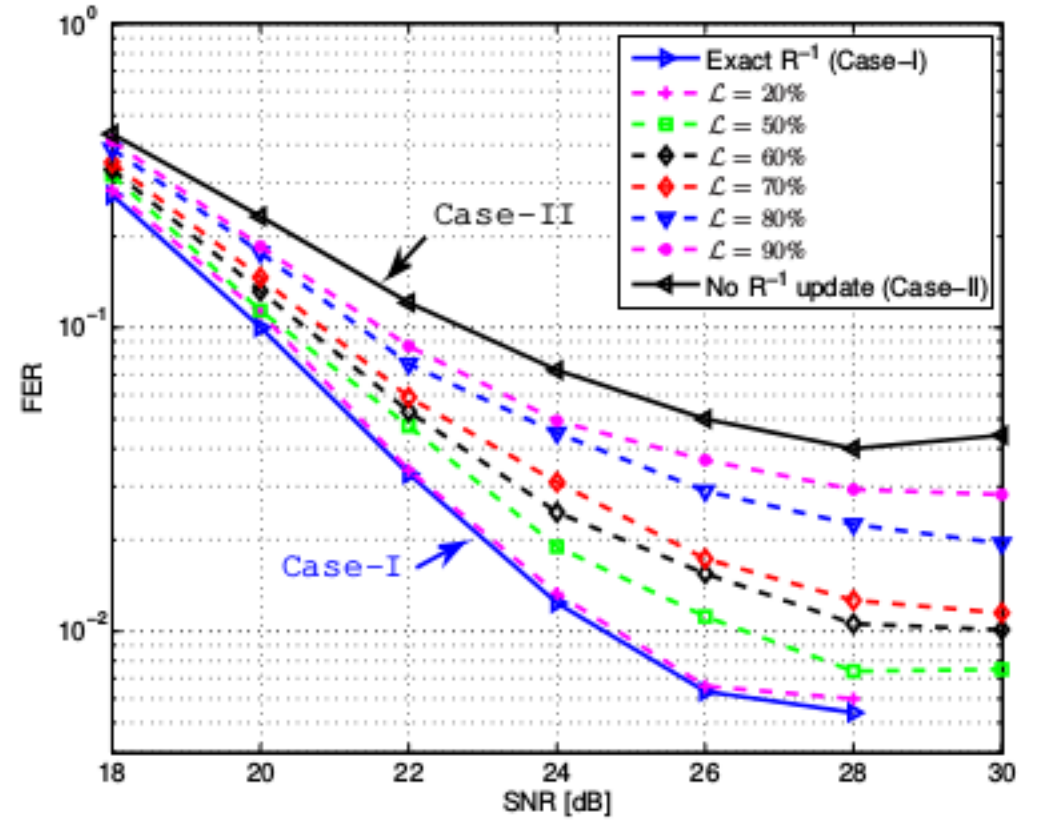
(b) R^{-1} update

Fig. 4. Simulated FERs for the proposed channel pre-processing update schemes in a 4×4 MIMO LTE-A downlink.

C. Algorithm evaluation

To testify the effectiveness of the proposed schemes, FER performance is simulated using a simplified 5MHz bandwidth LTE-A test environment. In each simulation, N_f LTE-A subframes, i.e., 14 OFDM symbols each containing 300 data subcarriers, are transmitted over a 4×4 MIMO 3GPP EVA-70 channel. N_f is dynamically adjusted to take account of different FERs with respect to SNR values. With the target level of FER $= 10^{-2}$, which is a commonly used criterion in practice, N_f varies between 500 and 6000. To evaluate the QRD update scheme, a K-Best signal detector with $K = 10$ is used together with 64-QAM modulation in FER simulations, whereas R^{-1} update test uses a linear MMSE detector with 16-QAM.

Simulated FERs, as shown in Fig. 4, clearly show that both update schemes fill up the performance gap between the two cases (Case-I and Case-II). Various design trade-offs can be obtained by varying $\mathcal{L}$ values. Generally, higher update levels provide higher complexity-energy reduction at the cost of performance degradation, whereas better FER performance can be achieved with lower update levels. Interesting to note that proposed schemes with $\mathcal{L} = 20\%$ exhibit similar FER performance as Case-I. Besides, a substantial FER improvement is presented in comparison to Case-II, e.g., for QRD updates shown in Fig. 4(a), the interference noise floor observed at FER $= 10^{-2}$ in Case-II is completely removed. With the adjustable update levels, the proposed schemes can be configured on-demand to suffice different system requirements, e.g., performance- or energy-oriented design criterion. Additionally, when implementing these schemes on reconfigurable platforms, the update level can be adjusted dynamically to adapt to instantaneous channel condition. Hence, the proposed update schemes are highly flexible.
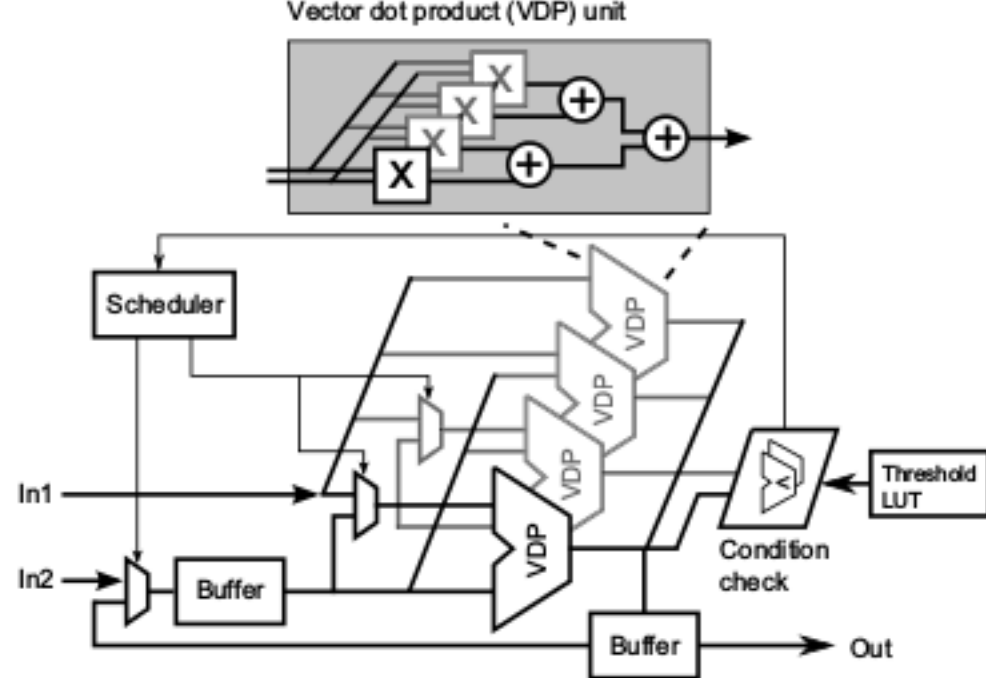
Fig. 5. Hardware architecture of the proposed update schemes.

IV. IMPLEMENTATION RESULTS AND COMPARISON

In this section, we present hardware implementation results of the QRD and R^{-1} update schemes. Since the proposed algorithms only involve vector processing, both QRD and the update schemes have the potential to be mapped onto one platform and to share computational resources. Therefore, hardware area is not our primary concern in this paper. Instead, we mainly focus on energy reduction in comparison with conventional QRD designs, and discuss design trade-offs between energy and FER performance.

A hardware architecture for the proposed update schemes is shown in Fig. 5. Resorting to the use of four complex-valued vector dot product (VDP) units, each R and R^{-1} update ((3) and (8)) for 4×4 channel matrices takes 8 and 16 clock cycles respectively, including on-line update condition check (4). Threshold values are stored in a look-up table (LUP), which are preloaded during system initialization or configured on-the-fly to adapt to the current channel scenario. The proposed hardware architecture is synthesized using Synopsys Design Compiler with a 65 nm low-leakage standard cell CMOS library. Presented by synthesis results, as summarized in Table I, the total core area is $0.242\,\text{mm}^2$ equivalent to a 116 K gate count. Operating at a normal 1.2 V core voltage, the maximum clock frequency is 330 MHz, resulting in an update throughput of 41.25 MQRD/s and 20.625 MR^{-1}/s. Besides, energy consumption of each QRD and R^{-1} update is 0.36 nJ and 0.72 nJ, respectively.

In Table I, we compare our implementation results with recently reported QRD designs. To ensure a fair comparison, we consider normalized throughput (N.T.) and energy (N.E.) to take account of various design parameters, such as throughput, gate count, and technology. It clearly shows that (Table I) the proposed QRD update scheme achieves 14% higher throughput (normalized) than the best QRD design reported in literature [6]. More importantly, each QRD and R^{-1} update consumes 4 and 2 times less energy (normalized) comparing to the most energy efficient QRD design [8]. Considering the simulation setup, i.e., 5MHz LTE-A with 300 data subcarriers per OFDM symbol, a 38% energy reduction can be obtained with QRD updates during half-H renewals when using 50% update level, while performance degradation is less than 1 dB at FER = 10^{-2}.

In order to choose an appropriate $\mathcal{L}$ value, a design trade-off analysis is carried out between FER and energy consumption. In this analysis, we evaluate FER performance by recording the SNR value of each $\mathcal{L}$ at the target level of FER = 10^{-2}. Taking the exact QRD case as a reference, performance degradation of the proposed schemes is measured as SNR value deviations. As an illustration, Fig. 6 shows the energy-performance trade-off curve for QRD update simulations (Fig. 4(a)). The same analysis can be applied to the R^{-1} case. It is worth to note that upto 60% energy reduction can be achieved with the proposed QRD update scheme whereas the SNR degradation is less than 1 dB.

TABLE I
IMPLEMENTATION COMPARISON WITH CONVENTIONAL QRD

Reference	[6]	[8]	This work	
Algorithm	QRD	QRD	QRD update	
	Givens	Interpolation	R	R^{-1}
Matrix size	4×4	4×4	4×4	
Technology [nm]	180	90	65	
Gate count [KGE]	111	317.95	116	
Max. freq. [MHz]	100	140.65	330	
Proc. cycles	4	4	8	16
Throughput [MQRD/s]	12.5	35.16	41.25	20.625
Energy [nJ/QRD]	12.76	1.39	0.36	0.72
N.T.[1] [MQRD/s/KGE]	0.31	0.15	0.35	0.18
N.E.[2] [nJ/QRD]	2.05	1.45	0.36	0.72

[1] N.T.: Throughput×(Technology/65nm)/(Gate count).
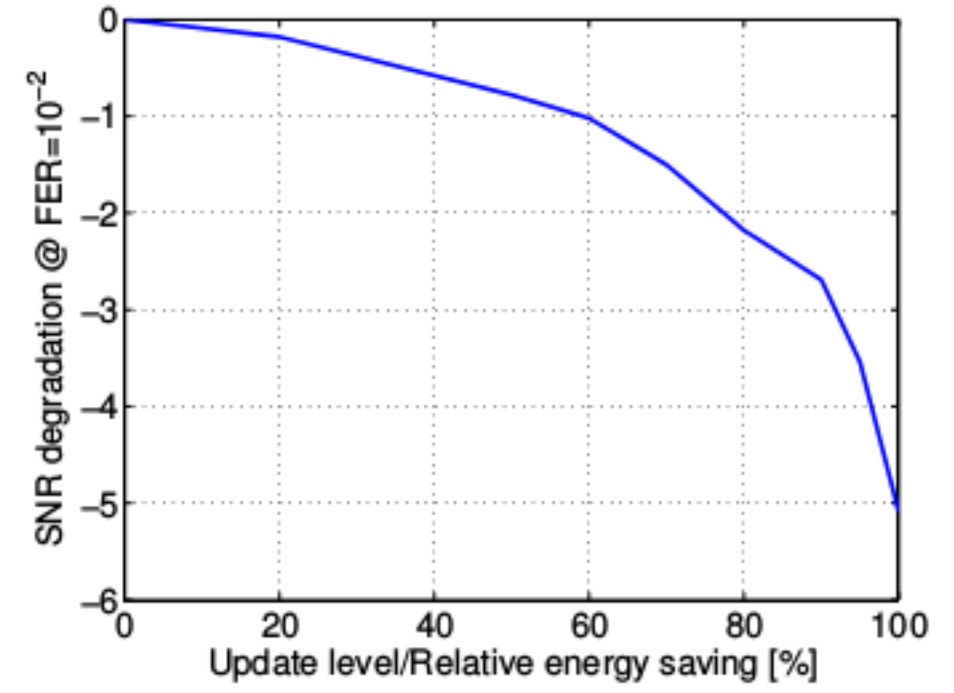[2] N.E.: Energy×(1.2V/Voltage)^2×(65nm/Technology).

Fig. 6. Energy-performance trade-off for QRD update scheme.

V. CONCLUSION

In this paper, an energy efficient update scheme for channel pre-processing operations is presented to reduce the number of explicit QRD computations and channel matrix inversions during successive channel matrix updates. In 4×4 MIMO systems, each QRD and R^{-1} update consumes 4 and 2 times less energy than the state-of-the-art QRD design. Moreover, benefiting from the flexible update criterion, the proposed schemes can be configured on-the-fly to adapt to different channel conditions or system requirements. Effectiveness of the presented update schemes is demonstrated by conducting simulations under 3GPP channel models with moderate mobility.

REFERENCES

[1] L. Liu, J. Lofgren, and P. Nilsson, "Area-Efficient Configurable High-Throughput Signal Detector Supporting Multiple MIMO Modes," *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2012.

[2] C. Senning *et al.*, "Systolic-array based regularized QR-decomposition for IEEE 802.11n compliant soft-MMSE detection," in *International Conference on Microelectronics (ICM)*, Dec. 2010.

[3] S. Haene *et al.*, "OFDM channel estimation algorithm and ASIC implementation," in *European Conference on Circuits and Systems for Communications (ECCSC)*, July 2008, pp. 270–275.

[4] S. Gifford, C. Bergstrom, and S. Chuprun, "Adaptive and linear prediction channel tracking algorithms for mobile OFDM-MIMO applications," in *IEEE Military Communications Conference (MILCOM)*, vol. 2, Oct. 2005.

[5] L. Gor and M. Faulkner, "Power Reduction through Upper Triangular Matrix Tracking in QR Detection MIMO Receivers," in *IEEE 64th Vehicular Technology Conference (VTC)*, Sept. 2006, pp. 1–5.

[6] Z.-Y. Huang and P.-Y. Tsai, "Efficient Implementation of QR Decomposition for Gigabit MIMO-OFDM Systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 10, Oct. 2011.

[7] S. Aubert, J. Tournois, and F. Nouvel, "On the implementation of MIMO-OFDM schemes using perturbation of the QR decomposition: Application to 3GPP LTE-A systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011.

[8] P.L. Chiu *et al.*, "Interpolation-Based QR Decomposition and Channel Estimation Processor for MIMO-OFDM System," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 5, May 2011.