Soft-decision-Directed MIMO Channel Estimation Geared to Pipelined Turbo Receiver Architecture

Daejung Yoon
Dept of Electrical and Computer Engineering
University of Minnesota
Minneapolis, Minnesota 55455
Email: yoonx053@umn.edu

Jaekyun Moon
Dept of Electrical and Computer Engineering
University of Minnesota
Minneapolis, Minnesota 55455
Email: moon@umn.edu

***Abstract*—*We consider channel estimation specific to turbo equalization for multiple-input multiple-output (MIMO) wireless communication. We develop soft-decision-driven sequential algorithms geared to a specific pipelined turbo equalizer architecture operating on orthogonal frequency division multiplexing (OFDM) symbols. One interesting feature of the pipelined turbo equalizer is that multiple soft-decisions become available at various processing stages. A tricky issue is the fact that these multiple decisions from different pipeline stages have correlated decision errors as well as varying levels of reliability. This paper establishes an optimization strategy for the channel estimator to track the target channel while dealing with observation sets with different qualities. The resulting algorithm is basically a linear sequential estimation algorithm and, as such, is Kalman-like in nature. The main difference here, however, is that the proposed algorithm must deal with the inherent correlation that exist among the multiple module outputs that cannot easily be removed by the traditional innovation approach. The proposed algorithm continuously monitor the quality of the feedback decisions and incorporate it in the channel estimation process. The proposed channel estimation schemes show certain performance and complexity advantages over existing EM-based algorithms.*

I. Introduction

Combining the multiple-input multiple-output (MIMO) antenna method with orthogonal frequency division multiplexing (OFDM) and spatial multiplexing is a well-established wireless communication technique. Bit-interleaved coded modulation (BICM) [1] used in conjunction with MIMO-OFDM and spatial multiplexing (SM) is particularly effective in exploring both spatial diversity and frequency selectivity without significant design efforts on specialized codes [2], [3]. Turbo equalization, also known as iterative detection and decoding (IDD) in wireless applications [5], is well-suited for BICM-MIMO-OFDM for high data rate transmission with impressive performance potentials [4], [5].

A critical issue in realizing the full performance of a MIMO-OFDM turbo receiver is significant performance degradation due to imperfect channel state information (CSI). The detrimental impact of imperfect CSI on MIMO detection is well known. (see, for example, [6], [7]). Previous works have identified desirable training patterns or pilot tones for estimating channel responses for MIMO systems [8]–[10]. However, the achievable data rate is substantially reduced when the number of channel parameters to be estimated increases (e.g., caused by an increased number of antennas).

Decision-directed (DD) channel estimation algorithms can be applied to the turbo receivers to improve channel estimation accuracy [11]–[13]. However, inaccurate feedback decisions degrade the estimator performance [14]. Maximum-a-posteriori (MAP)-based DD algorithms discussed in [11], [12] can improve the estimation accuracy, but they require additional information like the channel probability density function. The DD channel estimation algorithm jointly working with IDD has been actively researched [15]–[17]. Among the existing research works, several papers have been devoted to iterative expectation-maximization (EM) channel estimation algorithms using extrinsic or a posteriori information fed back from the outer decoder [15]–[17].

As alternative approach to iterative EM channel estimation, Kalman-based channel estimators have been discussed that are effective against the error propagation problem [18]. This work has introduced a soft-input channel estimator that adaptively updates the channel estimates depending on feedback decision quality. The decision quality is important for the decision-directed estimation due to error propagation. The soft-input channel estimator of [18] evaluates the feedback decision quality by tracking the noise variance including potential soft-decision error impact in order to realize the robust updating process of the Kalman filter. However an issue there is that the decision errors in the turbo equalizer can be highly correlated, in which case Kalman-based estimators lose optimality.

In this work, we develop a Kalman-like channel estimators for MIMO-OFDM based on a specific pipelined turbo equalizer receiver architecture. Before setting up the Kalman estimator, a novel method to reduce decision error correlation is introduced. The proposed method constructs a refined innovation sequence by irregularly puncturing certain soft decisions that are deemed to be correlated with the previous decisions. The resulting algorithm is basically a linear sequential estimation algorithm and, as such, is Kalman-like in nature. We also weigh the estimated channel responses in the detection process according to the reliability level of the estimation.

In demonstrating the viability of the proposed schemes, a SM-MIMO-OFDM system is constructed to comply with the same preamble and pilot tone structure of the IEEE 802.11n

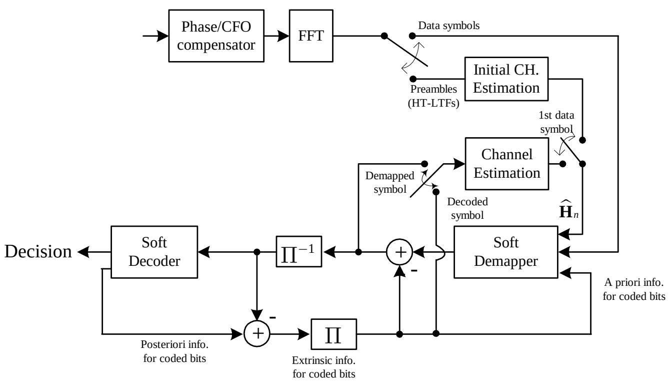


Fig. 1: Block diagram of the turbo receiver and the soft-decision-directed channel estimator

high speed WLAN standard [19]. Section II discusses the channel and system model, and briefly touches upon the high-throughput pipelined IDD architecture. Section III discusses an initial channel estimation based on the preambles and presents the proposed soft-DD channel estimation methods. Section IV presents packet error rate (PER) simulation result for performance evaluation, and finally conclusions are drawn in Section V.

II. Channel and System Model

We assume a SM-MIMO-OFDM transmitter where a data bit sequence is encoded by a convolutional channel encoder, and the encoded bit stream is divided to N_t spatial streams by a serial-to-parallel demultiplexer specified in the IEEE 802.11n spec. Each spatial stream is interleaved separately, and the interleaved streams are modulated using an M-ary quadrature amplitude modulation (M-QAM) symbol set $\mathcal{A}$. Since Q binary bits are mapped to an M-QAM symbol, a binary vector $\mathbf{b} = [b_0, b_1, \cdots, b_{QN_t-1}]^T$ is mapped to a transmitted symbol vector $\mathbf{s} = [s_1, s_2, \cdots, s_{N_t}]^T, (s_i \in \mathcal{A})$ is given from a set of $\mathcal{A}^{Nt}$, where $\mathcal{A}^{Nt}$ is the Cartesian product of M-QAM constellations. The M-QAM symbol sequence in each spatial stream is transmitted by an OFDM transmitter utilizing a fixed number of frequency subcarriers. For a particular subcarrier for the n^{th} OFDM symbol, the received signal at the discrete Fourier transform (DFT) output can be written as

$$\mathbf{z}_n = \mathbf{H}\mathbf{s}_n + \mathbf{n}_n \ , \tag{1}$$

where $\mathbf{z}_n = [z_1(n), z_2(n), \cdots, z_{N_r}(n)]^T$ is the received signal vector observed at N_r receive antennas, and $\mathbf{H}$ is the channel response matrix associated with all wireless links connecting N_t transmit antennas with N_r receive antennas antennas, and $\mathbf{n}$ is a vector of uncorrelated, zero-mean additive white Gaussian noise (AWGN) samples with equal variance set to $\mathcal{N}_o$.

The IDD technique described in [4] that performs turbo equalization for MIMO systems is assumed at the receiver. The extrinsic (EXT) information on the coded-bit stream is exchanged in the form of log-likelihood ratio (LLR) between

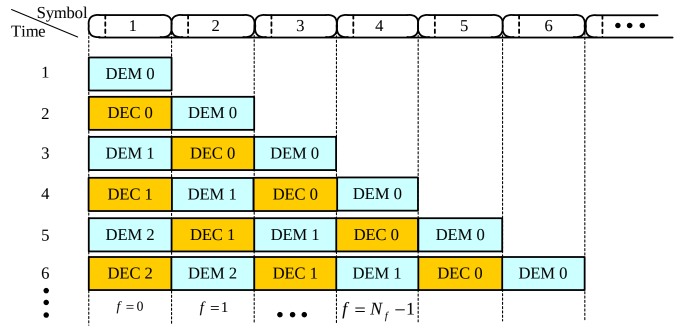


Fig. 2: Block diagram of the proposed optimum channel estimation algorithm geared to the pipelined IDD

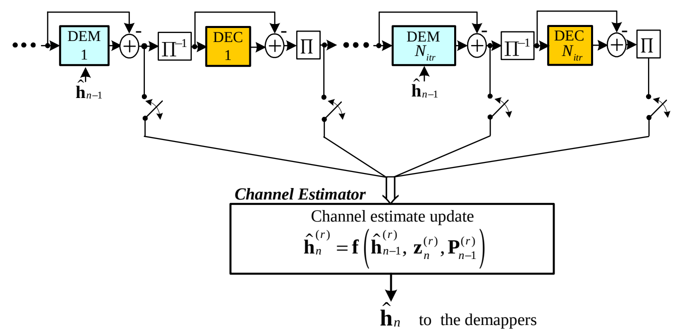


Fig. 3: Block diagram of the proposed channel estimation algorithm geared to the pipelined IDD

the soft-input soft-output (SISO) decoder and the SISO demapper as shown in Fig. 1. The demapper takes advantage of the reliable soft-symbol information made available by the outer SISO decoder. A soft-output Viterbi algorithm (SOVA) is used for the SISO decoder implementation [23]. Each data packet transmitted typically contains many OFDM symbols, and they are processed sequentially by the demapper and the decoder as they arrive at the receiver. In this way, there is no need for a buffer memory at the receiver to hold the entire packet. The feedback decisions used for channel estimation must be a interleaved coded-bit decisions. The EXT information from the demapper and the a-priori information to the demapper are matched with the interleaved coded-bit decisions and available for the channel estimation block as in Fig. 1.

The pipelined architecture is adopted to reduce the iteration latency [20], [21]. Fig. 2 illustrates OFDM symbols processed in pipelined IDD, and Fig. 3 shows the structure of the pipelined IDD receiver and its interface with the channel estimator. Multiple demapper-decoder pairs process multiple OFDM symbols at different iteration stages. Let N_{itr} denote the number of the IDD iterations required to achieve satisfactory error rate performance. Then, the N_{itr}-stage pipelined IDD receiver is equipped with N_{itr} demappers and N_{itr} decoders that are serially connected as in Fig. 3. Instead of feeding back EXT information from the decoder to the demapper, the decoder forwards its EXT information output to the demapper in the next iteration stage. Simultaneously, the demapper and the decoder in the previous iteration stage start to process a new OFDM symbol. The pipelined IDD operation is functionally equivalent to the original IDD scheme [20].

The EXT LLRs feedback from the pipelined demappers and decoders are utilized for the channel estimation. Let N_{sym} indicate the number of the total OFDM symbols in a packet, and N_f denotes the number of the data symbols which can be candidate feedbacks for the channel estimation. If a receiver requires N_{itr} IDD iterations, then a maximum of $2N_{itr}$ OFDM symbols are processed in the pipelined IDD receiver as illustrated in Fig. 3. Because the LLR outputs from the initial demapper and decoder have low reliability, they are not used for the channel estimation. Let index n indicate the processing time in pipeline IDD. In this pipelined IDD setup, when $2 \leq n \leq 2N_{itr}$, the channel estimator can get $(n-2)$ feedback decisions (i.e. $N_f = n-2$). When the number of the processed symbols increases to $2N_{itr}$ ($2N_{itr} \leq n \leq N_{sym}$), N_f is equal to $2N_{itr}-2$. After all the OFDM symbols in the packet have arrived at the receiver front-end, it will take sometime until all symbols will clear out of the pipeline. For $n \geq N_{sym}$, N_f is equal to $N_{sym}+2N_{itr}-n$.

III. Sequential and Soft-Decision-Directed Channel Estimation

First, we give a quick review on the practical initial channel estimation method based on preamble in the packet head, using the IEEE 802.11n standard as an example, and then present then the proposed sequential channel estimation algorithms.

A. Initial Channel Estimation Based on Training Symbols

The IEEE 802.11n standard specification provides a special training sequence named high-throughput long training fields (HT-LTFs) for the initial channel estimation purposes [19]. These HT-LTFs are inserted before the data fields in each packet. The transmitter sends an orthogonal symbol matrix $\mathbf{S}_{tr}$ representing the HT-LTFs sequence. A MIMO transmitter sends $\mathbf{S}_{tr}$ through a MIMO channel and the receiver observes a signal $\mathbf{Z}_{tr}$ as given in (1). The initially estimated $N_r \times N_t$ channel matrix can be performed by the least square method as

$$\begin{aligned} \widehat{\mathbf{H}}_{init} &= \mathbf{Z}_{tr}\mathbf{S}_{tr}^T\left(\mathbf{S}_{tr}\mathbf{S}_{tr}^T\right)^{-1} \\ &= \frac{1}{N_{tr}}\mathbf{Z}_{tr}\mathbf{S}_{tr}^T, \end{aligned} \quad (2)$$

where N_{tr} is the number of training symbols. The need for direct matrix inversion is avoided due to the orthogonal constraint imposed on $\mathbf{S}_{tr}$ (i.e. $\mathbf{S}_{tr}\mathbf{S}_{tr}^T = N_{tr}\mathbf{I}_{N_t}$). Using (2), the initial channel state information is obtained for each frequency tone.

B. Derivation of the Weighted Kalman-Based Sequential Channel Estimation Algorithm

The sequential form of the estimator is useful to improve the quality of channel estimate as the observation symbols arrive in a sequential fashion, as OFDM symbols do in the system of our interest. It is assumed that the channel is quasi-static over N_f OFDM symbol periods. For the pipelined IDD receiver at hand, the observation equation is set up at the r^{th} receiver (RX) antenna as

$$\mathbf{z}_n^{(r)} = \mathbf{S}_n\mathbf{h}^{(r)} + \mathbf{n}_n^{(r)}, \quad (3)$$

where $\mathbf{z}_n^{(r)}$ is the received signal vector $[z_0^{(r)}[n], ..., z_{N_f-1}^{(r)}[n]]^{\mathrm{T}}$, $\mathbf{S}_n$ is a $N_f \times N_t$ matrix, $\mathbf{h}^{(r)}$ is a $N_t \times 1$ vector that is a multi-input-single-output (MISO) channel vector specific to the r^{th} RX antenna. Truly, other antenna observations are useless for the MISO channel estimation related with the r^{th} receiver. The goal is to do a sequential estimation of $\mathbf{h}^{(r)}$ as n progresses. The estimation process is done in parallel to obtain channel estimates for all N_r RX antennas. With an understanding that we focus on a specific RX antenna, the RX antenna index r is dropped for notation simplification.

A mean symbol decision $\tilde{s}$ is defined as the average of the constellation symbols according to the EXT probabilities (i.e. $\tilde{s} = \sum_{s_i \in \mathcal{A}} s_i P(s_i)$). The EXT probabilities can be simply found by converting the available EXT LLRs into probabilities.

1) Innovation Sequence Setup: The pipeline architecture can be viewed as a buffer large enough to accommodate N_f OFDM symbols, but we take into account in our channel estimator design the different levels of reliability for the soft decisions coming out of the demapper or decoder modules at different iteration stages. First defining the soft decision error $\mathbf{E} \triangleq \mathbf{S} - \widetilde{\mathbf{S}}$, (3) can be rewritten as

$$\mathbf{z}_n = \{\widetilde{\mathbf{S}}_n + \mathbf{E}_n\}\mathbf{h} + \mathbf{n}_n. \quad (4)$$

Let $\mathbf{y}_n$ be an innovation sequence for the observation $\mathbf{z}_n$. An innovation sequence is a white sequence that is a causal and casually invertible linear transformation of the observation sequence [22]. We write

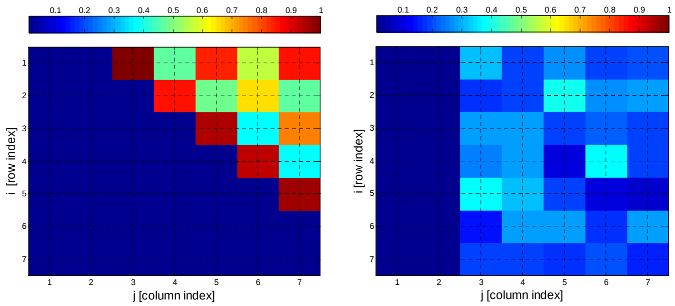
$$\begin{aligned} \mathbf{y}_n &\triangleq \mathbf{z}_n - \widetilde{\mathbf{S}}_n\hat{\mathbf{h}}_{n-1} & (5) \\ &= \widetilde{\mathbf{S}}_n(\mathbf{h} - \hat{\mathbf{h}}_{n-1}) + \mathbf{E}_n\mathbf{h} + \mathbf{n}_n. & (6) \end{aligned}$$

Ideally, the vector sequence $\mathbf{y}_n$ would represent an innovation sequence in the sense that any given component of the vector $\mathbf{y}_{n-k}$ is orthogonal to any component of $\mathbf{y}_n$ as long as $k \neq 0$. In this scenario we would have

$$\begin{aligned} E &\quad [y_{n-k}[i]y_n^H[j]] = E\left[\mathbf{e}_{n-k}[i]\widehat{\mathbf{h}}_{n-k-1}\mathbf{h}^H\mathbf{e}_n^H[j]\right] \\ = &\quad \begin{cases} \sum_{t=1}^{N_t}\hat{\rho}_{n-1}^{(t)}\sigma_s^2[t,i] & \text{when} \quad k=0 \text{ and } i=j \\ \epsilon(\approx 0) & \text{otherwise}, \end{cases} \end{aligned} \quad (7)$$

where $\widetilde{\mathbf{s}}_n[i]$ and $\widetilde{\mathbf{e}}_n[i]$ indicate the i^{th} row vectors respectively of $\widetilde{\mathbf{S}}_n$ and $\widetilde{\mathbf{E}}_n$, and $\hat{\rho}_{n-1}^{(t)} \triangleq E[\hat{h}_{n-1}^{(t)}h^{(t)H}]$ and $\sigma_s^2 \triangleq E[|s-\tilde{s}|^2]$ indicating the symbol decision error variance. In deriving (7), we assumed: $E[y_{n-k}[i](\mathbf{h}-\widehat{\mathbf{h}}_{n-1})^H] = \mathbf{0}$, $E[s[i]\,e[j]^H] = 0$ for any k, i and j. In order for (7) to hold, the following must be true:

1) Links in the MISO channel are uncorrelated.
2) The channel estimate and decision error are independent.



(a) (b)

Fig. 4: Innovation sequence correlation example : (a) $E[\mathbf{y}_{n-2}\mathbf{y}_n^H]$ vs (b) $E[\mathbf{y}_{n-2}^{'}\mathbf{y}_n^{'H}]$,$c = 0.8$ (normalized by $E[|y_n[0]|^2]$, averaging 50 error packets)

3) The decision errors are uncorrelated.
(i.e. $E[e_{n-k}[i]e_n^H[j]] = \epsilon, k \neq 0$ or $i \neq j$)

Under these three assumptions, the vector $\mathbf{y}_n$ reasonably represents an innovation sequence.

2) Innovation Sequence with Punctured Feedback: Assumption (1) is reasonable, if the antenna is physically separated by a half of frequency wavelength. However, assumption (2) and (3) are problematic. A poor channel estimate generates a poor decision, which goes against assumption (2). Also, with such a poor estimate, the poor decision has dominant error caused by the estimation error, which is against assumption (3). We can also see that assumptions (2) and (3) are related to each other because of the interplay between detection and channel estimation. Once a sequence begins to be correlated, the Kalman filter is not optimum any more, and the correlated error circulates the IDD and channel estimator loop. Our goal here is to provide a refined innovation sequence to avoid this error propagation.

The estimator utilizes multiple soft feedback decisions from the pipelined IDD blocks. In terms of the maximum likelihood estimation, the more observations we have, the better performance is guaranteed, essentially due to noise averaging. Thus, using both the demapper and decoder feedback are definitely beneficial to the estimator. However, correlated decision errors from the multiple feedback paths hurt the optimality of the Kalman-based estimator.

In our set-up, before applying Kalman-like channel estimation, the channel estimator attempts to "puncture" out the correlated inputs. The ultimate purpose of puncturing is to forcefully set up an innovation sequence by reducing correlation between successive observations. First we observe that there is no significant correlation between the demapper and decoder outputs thanks to the de-/interleaver. An issue is the demapper-demapper or decoder-decoder output correlations for a given received signal (OFDM symbol), especially when a packet is bad (certain tones causing errors despite persistent IDD efforts). In the pipelined IDD setup, it takes $n = 2$ time steps for a demapper decision to shift to the next-stage demapper, and likewise for the decoder outputs. Consequently, components in observation vectors with even time difference has correlation as seen in Fig.4-(a) between $\mathbf{y_n}$ and $\mathbf{y_n - 2}$. This can cause biased channel estimation. The correlation between the previous demapper output and the current demapper output (or between the previous decoder output and the current decoder output) is very roughly found as

$$\beta_n(f) \triangleq \langle y_{n-2}[f-2], \;\; y_n[f]\,\rangle \tag{8}$$

where $< a, b >$ indicates an inner product operation. An improved innovation sequence is defined based on the component correlation as

$$\mathbf{y}_n^{'} = \{y_n[f] \mid \;\; y_n[0],\, \delta_f\,(y_n[f]) == 1 \;\;\}\,, \tag{9}$$

where $y_n[0]$ is a new input element in $\mathbf{y_n}$, which is automatically included in the refined innovation vector, while other components are selected or punctured by the condition $\delta_f\,(y_n[f])$ set as

$$\delta_f\,(y_n[f]) = \begin{cases} 1, & if\;\;|\beta_n(f)| \;\leq\; c\mathcal{N}_o \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

where c is a parameter to control the threshold ($c \geq 0$). Denote the size of the punctured innovation sequence $\mathbf{y}_n^{'}$ as $N_d (\leq N_f)$, and let index d indicate its components.

Fig.4 shows the example of the correlations in the innovation sequence before and after the refinement through puncturing: $E[\mathbf{y}_{n-2}\mathbf{y}_n^H]$ vs $E[\mathbf{y}_{n-2}^{'}\mathbf{y}_n^{'H}]$. The sequence $\mathbf{y}_n^{'}$ may have a smaller number of observation samples, but its correlation is low as seen in Fig.4-(b), which is useful to maintain the optimality of the Kalman filter. The parameter c is a very important parameter that controls trade-off : if c is large, the number of observation increases, which is beneficial in terms of ML estimation. However a large c can feed biased decision errors to the Kalman-based estimator.

It is interesting to note that a puncturing mechanism to construct $\mathbf{y}_n^{'}$ results in innovating the sequence $\mathbf{y}_n$ over time n. In other words, puncturing irregularly (or randomly) contributes to innovating a sequence. Another interesting observation we make is that irregular puncturing activity become more pronounced in broken (bad) packets. Once the decisions are incorrect, correlation between the components of $\mathbf{y}_n$ appears, and puncturing becomes active. In order to salvage a bad packet from biased errors, the puncturing forcefully attempts to innovate the sequence $\mathbf{y}_n$. The puncturing process in this context can also be viewed as an effort to remove redundant information to circulate in iterative signal processing. We observe that although the puncturing cannot completely remove the correlated errors, a significant portion of the biased-error gets eliminated before the channel estimation step resumes.

C. Kalman-Based Sequential Channel Estimation Algorithm with Punctured Innovation Sequence

The weighted innovation sequence is set up. Then, a linear channel estimator using an weighted innovation sequence can be specified as a matrix $\mathbf{A}$, that is $\hat{\mathbf{h}} = \mathbf{A}\mathbf{y}_n^{'}$. The Kalman estimator is now derived as

$$\begin{aligned}\hat{\mathbf{h}}_n &= \widehat{E}[\mathbf{h}|\mathbf{y}_1^{'}, \mathbf{y}_2^{'}, ..., \mathbf{y}_n^{'}] \\ &= \widehat{E}[\mathbf{h}|\mathbf{y}_1^{'}, \mathbf{y}_2^{'}, ..., \mathbf{y}_{n-1}^{'}] + \widehat{E}[\mathbf{h}|\mathbf{y}_n^{'}] \\ &= \hat{\mathbf{h}}_{n-1} + \mathbf{A}_n\mathbf{y}_n^{'} \end{aligned} \tag{11}$$

where $\widehat{E}[\mathbf{a}|\mathbf{b}]$ denotes the optimal linear estimator of $\mathbf{a}$ given $\mathbf{b}$.

To find the linear estimator matrix $\mathbf{A}_n$, the orthogonality principle is applied:

$$\begin{aligned}\overline{(\mathbf{h} - \mathbf{A}_n\mathbf{y}_n^{\prime})\,\mathbf{y}_n^{\prime H}} = \mathbf{0} \\ \mathbf{A}_n\overline{\mathbf{y}_n^{\prime}\mathbf{y}_n^{\prime H}} = \overline{\mathbf{h}\mathbf{y}_n^{\prime H}},\end{aligned} \tag{12}$$

where an overbar also indicates statistical expectation. The right-hand-side of the last line in (12) is given by

$$\overline{\mathbf{h}\mathbf{y}_n^{\prime H}} = \underbrace{\overline{(\mathbf{h} - \hat{\mathbf{h}}_{n-1})(\mathbf{h} - \hat{\mathbf{h}}_{n-1})^H}}_{\triangleq \mathbf{P}_{n-1}}\,\widetilde{\mathbf{S}}_n^H, \tag{13}$$

where $\mathbf{P}_{n-1}$ is defined as the channel estimation error variance matrix, and the term $\overline{\mathbf{y}_n^{\prime}\mathbf{y}_n^{\prime H}}$ in (12) can be written as

$$\begin{aligned}\overline{\mathbf{y}_n^{\prime}\mathbf{y}_n^{\prime H}} &= \widetilde{\mathbf{S}}_n\overline{(\mathbf{h} - \hat{\mathbf{h}}_{n-1})(\mathbf{h} - \hat{\mathbf{h}}_{n-1})^H}\,\widetilde{\mathbf{S}}_n^H \\ &+ \underbrace{\overline{\mathbf{E}_n\mathbf{h}\mathbf{h}^H\mathbf{E}_n^H}}_{\triangleq \mathbf{Q}_n} + \mathcal{N}_o\mathbf{I}_{N_d}.\end{aligned} \tag{14}$$

Now using (12), (13) and (14), the matrix $\mathbf{A}_n$ can be obtained as

$$\begin{aligned}\mathbf{A}_n &= \overline{\mathbf{h}\mathbf{y}_n^H}(\overline{\mathbf{y}_n\mathbf{y}_n^H})^{-1} \\ &= \mathbf{P}_{n-1}\widetilde{\mathbf{S}}_n^H(\widetilde{\mathbf{S}}_n\mathbf{P}_{n-1}\widetilde{\mathbf{S}}^H + \mathbf{Q}_n + \mathcal{N}_o\mathbf{I}_{N_d})^{-1}.\end{aligned} \tag{15}$$

The next steps to complete the process are to express $\mathbf{P}_{n-1}$ and $\mathbf{Q}_n$ in a recursive fashion. Noticing $(\mathbf{h} - \hat{\mathbf{h}}_n) = \mathbf{h} - (\hat{\mathbf{h}}_{n-1} + \mathbf{A}_n\mathbf{y}_n^{'})$ from (11), the channel estimation error variance at time n can be rewritten as

$$\begin{aligned}\mathbf{P}_n &= \overline{\{\mathbf{h} - (\hat{\mathbf{h}}_{n-1} + \mathbf{A}_n\mathbf{y}_n^{\prime})\}\{\mathbf{h} - (\hat{\mathbf{h}}_{n-1} + \mathbf{A}_n\mathbf{y}_n^{\prime})\}^H} \\ &= (\mathbf{I}_{N_t} - \mathbf{A}_n\widetilde{\mathbf{S}}_n)\mathbf{P}_{n-1},\end{aligned} \tag{16}$$

where we utilized the relation $\overline{\mathbf{y}_n^{\prime}\mathbf{y}_n^{\prime H}}\mathbf{A}_n^H = \widetilde{\mathbf{S}}_n\mathbf{P}_{n-1}^H$ which is obvious from (12) and (13). Also note $\mathbf{P}_n$ is a symmetric matrix of which pivot has non-negative real values, based on its definition.

Finally, $\mathbf{Q}_n$ needs to be found. The symbol decision error variance $\sigma_s^2 = E[|s - \tilde{s}|^2]$ can be found by using the EXT probabilities (i.e. $\sigma_s^2 = \sum_{s_i \in \mathcal{A}} |s_i - \tilde{s}|^2 P(s_i)$). However, finding $\mathbf{Q}_n$ is a bit tricky as the channel state information are unknown to obtain the $\rho^{(t)}$ value in the receiver. The actual channel multiplication matrix $\mathbf{h}\mathbf{h}^H$ is not known to the receiver, instead, the channel correlation matrix is found from $\overline{\mathbf{h}\mathbf{h}^H} = \overline{\{(\mathbf{h} - \hat{\mathbf{h}}_n) + \hat{\mathbf{h}}_n\}\{(\mathbf{h} - \hat{\mathbf{h}}_n) + \hat{\mathbf{h}}_n\}^H}$, which reduces to $\overline{\mathbf{h}\mathbf{h}^H} = \mathbf{P}_n + \hat{\mathbf{h}}_n\hat{\mathbf{h}}_n^H$. Under the reasonable assumption of $\overline{(|s_j - \tilde{s}_j)(|s_i - \tilde{s}_i|)^H} = 0$ when $i \neq j$, we have the $N_d \times N_d$ diagonal matrix $\mathbf{Q}_n$ is given as

$$\begin{aligned}\mathbf{Q}_n &= \overline{\mathbf{E}_n\left(\mathbf{P}_n + \hat{\mathbf{h}}\hat{\mathbf{h}}_n^H\right)\mathbf{E}_n^H} \\ &= \mathrm{diag}\Bigg[\sum_{t=1}^{N_t}\left(p_n(t,t) + |\hat{h}_t[n-1]|^2\right)\sigma_s^2(n,0,t), \cdots, \\ &\qquad \sum_{t=1}^{N_t}\left(p_n(t,t) + |\hat{h}_t[n-1]|^2\right)\sigma_s^2(n,N_d-1,t)\Bigg]\end{aligned} \tag{17}$$

where $\hat{h}_t[n-1]$ is from the previous estimate $\hat{\mathbf{h}}_{n-1}$, $p_n(t,t)$ is the t^{th} diagonal element of $\mathbf{P}_{n-1}$, and $\sigma_s^2(n,j,t)$ is the decision error variance of the (j,t) element of $\widetilde{\mathbf{S}}_n$.

Putting it all together, the proposed Kalman estimator is summarized as a set of equations : (17), (15), (16) and (11) by the processing order. To start the Kalman estimator, $\hat{\mathbf{h}}_{-1}$ corresponding to the initial time $n = 0$ can be given by the initial channel estimator in (2). Also the initial matrix $\mathbf{P}_{-1}$ can be derived from the MMSE analysis [24] as $\mathbf{P}_{-1} = \mathrm{diag}[|\hat{\mathrm{h}}_{-1}^{(\mathrm{t})}|^2/(\gamma|\hat{\mathrm{h}}_{-1}^{(\mathrm{t})}|^2 + 1)]$ for $t = 1, .., N_t$ where $\gamma = E_s/(N_t\mathcal{N}_o)$. Remind that the MISO channel estimation algorithm can be extended to the MIMO channel estimation by repeating the (17), (15), (16) and (11) to each RX antenna.

IV. Performance Evaluations

The performance of the proposed algorithm is investigated through the packet error rate (PER) analysis. We present results for a 3×3 16-QAM SM-MIMO-OFDM system. The transmitter transmits 1000-byte-long packets, and PER performances are evaluated down to a 1% PER level at which actual WLAN systems reasonably operate. The SISO MMSE-demapper is used [5]. The convolutional code is used with the coding rate $1/2$ with generator polynomials $g_o = 133_8$ and $g_1 = 171_8$ complying with IEEE 802.11n specifications [19]. The MIMO multi-path channel is modeled based on an exponentially decaying power-profile with delay spread $T_{rms} = 50ns$ and is assumed uncorrelated across the links established over different pairs of TX- and RX-antennas.

The proposed algorithm is also compared with a perfect-CSI scenario. For performance comparisons, the proposed optimum Kalman-based estimator is compared to an EM algorithm with comparable complexity. The DD EM estimator introduced as a variant of the EM estimator in [15] is applied to the setup of (3) as

$$\hat{\mathbf{h}}_{o,n}^{(r)} = \left(\widetilde{\mathbf{S}}_n^H\widetilde{\mathbf{S}}_n\right)^{-1}\widetilde{\mathbf{S}}_n^H\mathbf{z}_n^{(r)}, \tag{18}$$

and this estimate is blended with the training-based channel estimate by a combining method (i.e. $\hat{h}_n^{(t,r)} = a_n\hat{h}_{tr}^{(t,r)} + b_n\hat{h}_o^{(t,r)}[n]$) [16]. For finding the coefficients, the channel estimates are modeled as $\hat{h}_{tr} = h + \eta_{tr}$ and $\hat{h}_o[n] = h + \eta_o[n]$ respectively, and the coefficients a_n and b_n are obtained by the rule [16] : $\min_{a_n, b_n} E\left[|a_n\eta_{tr} + b_n\eta_o|^2\right]$ subject to $a_n + b_n = 1$. The EM channel estimation algorithm needs to update the

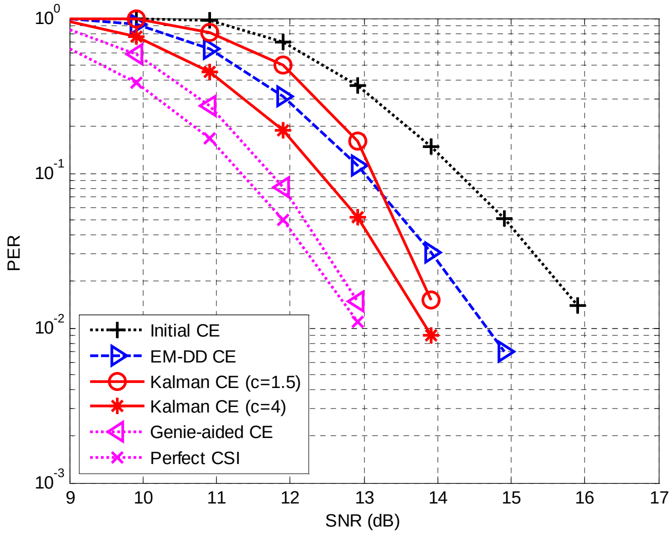


Fig. 5: PER simulations in the 3x3 SM-MIMO-OFDM turbo receiver (7 iterations)

noise variance for the detector. The EM noise variance update method is presented in [15] as

$$\widehat{\mathcal{N}}_o[n] = \frac{1}{N_r N_d} \sum_{r=1}^{N_r} \sum_{d=0}^{N_d-1} \left(z_n^{(r)} - \widetilde{\mathbf{S}}_n \hat{\mathbf{h}}_n^{(r)} \right)^H \left(z_n^{(r)} - \widetilde{\mathbf{S}}_n \hat{\mathbf{h}}_n^{(r)} \right). \tag{19}$$

Fig. 5 shows the results. The proposed Kalman estimator with the threshold parameter $c = 4$ has a $0.8dB$ gap to the perfect-CSI performance at a 10^{-2} PER. At $c = 1.5$, the proposed scheme dose not perform as well at high PERs but tends to approach the performance of $c = 4$ at lower PERS (and possibly is even better at very low PERS, judging from the trend). When assuming the proposed estimator has perfect information feedback ('Genie-aided CE(matrix)'), the proposed algorithm essentially achieves the perfect CSI performance. The EM-DD ('EM-DD CE') curve has a $0.6dB$ loss relative to the proposed estimator with $c = 4$.

V. Conclusions

A sequential soft-decision-directed channel estimator for SM-MIMO-OFDM systems has been proposed for the specific pipelined turbo-receiver architecture. The algorithm deals with observation sets with varying levels of reliability. In coping with decision errors that propagate in the pipeline, we have introduced a novel method of innovating a correlated sequence via puncturing. Based on the refined innovation sequence, a Kalman-like estimator has been constructed. The proposed algorithm establishes improved Kalman-like channel estimation where the traditional innovation approach cannot create a true innovation sequence due to decision error propagation.

References

[1] G. Caire, G. Taricco, and E. Biglieri,"Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927-946, May, 1998.

[2] A. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," *Proc. of IEEE Vehicular Technology Conference*, pp. 473-478, Boston, Sept., 2000.

[3] D. Park and B. Lee, "Design criteria and performance of space-frequency bit-interleaved coded modulations in frequencyselective Rayleigh fading channels," *Journal of Commun. and Networks*, vol. 5, no. 2, pp. 141-149, June, 2003.

[4] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization : an iterative equalization and decoding technique for coded data transmission," *IEEE Signal Processing Mag.*, vol. 21, pp. 67-80, Jan. 2004.

[5] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean square error equalization using a prori information," *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 673-683, Mar, 2002.

[6] Y. Huang and J. Ritcey, "EXIT chart analysis of BICM-ID with imperfect channel state information," *IEEE Commun. Letters*, vol. 7, no. 9, pp. 434-436, Sept., 2003.

[7] Y. Huang and J. Ritcey, "16-QAM BICM-ID in fading channels with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1000-1007, Sept., 2003.

[8] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67-75, Jan., 2002.

[9] X. Ma, L. Yang, and G. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453-466, Mar., 2005.

[10] B. Hassibi and B. Hochwald, "Optimal training in space time systems," *Proc. 34th Asilomar Conf. on Signals, Systems and Computers*, pp. 743-747, Oct., 2000.

[11] X. Deng, A. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," *Proc. IEEE Int. Conf. Commun.*, vol. 4, pp. 2326-2329, Anchorage, AK, May, 2003.

[12] J. Gao and H. Liu, "Decision-directed estimation of MIMO time-varying Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1412-1417, Jul., 2005.

[13] M. Loncar, R. Muller, J. Wehinger, and T. Abe, "Iterative joint detection, decoding, and channel estimation for dual antenna arrays in frequency selective fading," *Proc. Int. Symposium on Wireless Personal Multimedia Commun.*, Honolulu, HI, Oct., 2002.

[14] M. Tuchler, R. Otnes, and A. Schmidbauer, "Performance of soft iterative channel estimation in turbo equalizer," *Proc. IEEE ICC2002 Int. Conf.* vol. 3, pp. 1858-1862, New York, NY, Apr. 2002.

[15] X. Wautelet, C. Herzet, A. Dejonghe, J. Louveaux, and L. Vandendorpe, "Comparision of EM-based algorithms for MIMO channel estimation," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 216-226, Jan., 2007.

[16] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select Areas Commun.*, vol. 19, no. 8, pp. 1450-1460, Aug., 2001.

[17] M. Khalighi and J. Boutros, "Semi-blind channel estimation using the EM algorithm in iterative MIMO APP detectors," *IEEE Wireless Commun.*, vol. 5, no. 11, pp. 3165-3173, Nov., 2006.

[18] S. Song, A. Singer, and K. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Sig. Processing*, vol. 52, no. 10, pp. 2885-2894, Oct., 2004.

[19] *IEEE P802.11n/D1.0 : Draft Amendment to STANDARD FOR 2 Information Technology-Telecommunications and 3 information exchange between systems-Local and 4 Metropolitan networks-Specific requirements-Part 5 11: Wireless LAN Medium Access Control (MAC) 6 and Physical Layer (PHY) specifications: 7 Enhancements for Higher Throughput.*

[20] S. Abbasfar, "Turbo-like codes; design for high speed decoding,"Springer, Netherlands, 2007.

[21] S. Lee, N. Shanbhag, and A. Singer, "Area-efficient high-throughput VLSI architecture for MAP-based turbo equalizer," *Proc. IEEE signal procssing system: design and implementation* pp. 87-92, Aug. 2003, Seoul, Korea.

[22] H. Stark and J. Woods, "Probability and random processes with applications to signal processing," Upper Saddle River, NJ, Prentice-Hall, 2002.

[23] J. Hagenauer and P. Hoeher, "A Viterbi Algorithm with Soft-Decision Outputs and its Applications," *Globecom 1989*, vol. 3, pp. 1680-1686, Dallas,TX, Nov., 1989

[24] S. Kay, "Fundamentals of statistical signal processing-estimation theory," Englewood Cliffs, NJ, Prentice-Hall, 1993.