Layered Space-Time Equalization for Wireless MIMO Systems

Xu Zhu and Ross D. Murch, Senior Member, IEEE

Center for Wireless Information Technology
Department of Electrical & Electronic Engineering
The Hong Kong University of Science & Technology
Clear Water Bay, Kowloon, Hong Kong

June 5, 2002

Abstract

In this paper we investigate layered space-time equalization (LSTE) architectures for multiple-input-multiple-output (MIMO) frequency selective channels. At each layer or stage of detection, a MIMO delayed decision feedback sequence estimator (MIMO-DDFSE) is used to tentatively detect a group of selected data streams, among which a sub-group of data streams are output and are canceled from the received signals. With the proposed architectures, the number of the tentatively detected and output data streams can be different at different LSTE stages, while the MIMO-DDFSE can also reduce to the special cases of multiple-input-single-output decision feedback equalizer (MISO-DFE), MISO-DDFSE and MIMO-DFE, allowing tradeoffs between performance and complexity. We also derive the equalizer coefficients, discuss timing recovery and consider channel estimation. Simulation results demonstrate the performance of the proposed LSTE structures, and the tradeoffs between performance and complexity of the multistage structure and the single-stage version. We also demonstrate the impact of imperfect channel estimation, imperfect interference cancellation, the number of receive antennas, filter length, and oversampling on performance.

Keywords: layered space-time equalization (LSTE), MIMO, DDFSE, DFE, channel estimation.

1This work was supported by Hong Kong Research Grants Council HKUST 6048/00E.
2Corresponding Author. Phone: (+852) 2358 7044. Fax: (+852) 2358 1485. Email: eermurch@ee.ust.hk.
I. INTRODUCTION

To achieve high-speed data transmission in future wireless communications, multiple-input-multiple output (MIMO) systems, where multiple antennas are employed at both the transmitter and receiver, have been proposed. They promise improved performance and capacity and are currently the subject of much research activity [1-8].

Extensive investigations [1-6] have been performed on a variety of techniques for MIMO systems, such as Vertical Bell Laboratories Layered Space-Time (V-BLAST) [4] and maximum likelihood detection (MLD) [5-6]. However, most of them are based on the assumption of flat fading channels, and further research is needed on the equalization of MIMO frequency selective channels. In [7] a so-called wideband BLAST structure was developed, where at each stage of detection, a transmit data stream is estimated by a multiple-input-single-output decision feedback equalizer (MISO-DFE). Its counterpart with a MISO delayed decision feedback sequence estimator (MISO-DDFSE) for the coded system was shown in [8] (DFE is a special case of DDFSE [9]).

In this paper we investigate a wide range of layered space-time equalization (LSTE) architectures to combat MIMO frequency selective channels, by employing both space-time equalization and successive interference cancellation. At each stage or layer of LSTE, a MIMO-DDFSE (or MIMO-DFE [10-11]) is employed, and a group of the best data streams in the minimum mean square error (MMSE) sense are selected from the remaining candidates for tentative detection by the MIMO-DDFSE. Among the tentative decisions, a sub-group of the best (also in the MMSE sense) detected data streams are output and are canceled from the received signals.

Our work is different from previous work of [7-8] in that we present a general framework by introducing MIMO equalizers into the LSTE structure. With this architecture, the number of the tentatively detected and output data streams can be different at different LSTE stages. In addition, our use of MIMO-DDFSE allows us to consider the special cases of MIMO-DFE [10-11], MISO-DFE [7] and MISO-DDFSE, allowing a wide range of tradeoffs between complexity and performance. We also derive the equalizer coefficients, discuss timing recovery and consider channel estimation. Simulation results show that the LSTE structure with MIMO-DDFSE has a significant advantage over other structures including the structure proposed by [7], especially with fewer space-time samples. We also show the tradeoffs between performance and complexity of the multistage structure and the single-stage version. The impact of imperfect
channel estimation and imperfect interference cancellation on performance is also investigated. The effects of the number of receive antennas, filter length, and oversampling are also illustrated.

Section II presents the system model, and the proposed LSTE structure is described in section III. Computational complexity of the LSTE is analyzed in section IV, and section V provides the simulation results. Conclusions are drawn in section VI.

II. SYSTEM MODEL

The wireless MIMO system considered here is an $M$-ary complex baseband-equivalent system, with $K$ transmit antennas and $L$ receive antennas. We use $d_k[i]$ to denote the $i$th data symbol transmitted by the $k$th ($k = 1, \ldots, K$) antenna, with symbol period $T$ and unit average symbol energy $E_s = 1$. It is assumed that the data are uncorrelated in space and time, and are transmitted over frequency selective channels which are assumed to be quasi-static, i.e., time invariant over a data packet length [12], where a packet is defined as a single transmission burst.

The channel impulse response (CIR) between the $k$th transmit antenna and the $l$th receive antenna is given by

$$f_{kl}(t) = \sum_{n=0}^{N_p-1} f_{n,kl} \delta(t - \sigma_n)$$

where $N_p$ is the total number of paths, $f_{n,kl}$ and $\sigma_n$ are the channel gain and delay for the $n$th ($n = 0, \ldots, N_p - 1$) path, respectively. It is assumed that paths with different delays and different spatial diversity branches are independent. $f_{n,kl}$ is an independent zero mean complex Gaussian random variable, with a variance following the discrete exponential power delay profile:

$$E[f_{n,kl}^2] = b \exp(-\sigma_n / \sigma)$$

where $\sigma$ is referred to as the RMS delay spread [13], and $b$ is a normalizing factor such that $\sum_{n=0}^{N_p-1} E[f_{n,kl}^2] = 1$. Letting $g(t)$ denote the pulse shape lumping the effect of the transmit filter and receive filter, the overall CIR between the $k$th transmit antenna and the $l$th receive antenna is the convolution result of the physical CIR $f_{kl}(t)$ with $g(t)$, expressed as

$$h_{kl}(t) = \sum_{n=0}^{N_p-1} f_{n,kl} g(t - \sigma_n)$$

Thus, the continuous-time received signal at the $l$th ($l = 1, \ldots, L$) receive antenna is given by

$$x_l(t) = \sum_{k=1}^{K} d_k[i] h_{kl}(t - iT) + n_l(t)$$
where \( n_l(t) \) is the complex Additive White Gaussian Noise (AWGN) at the \( l \)th receive antenna, with a single-sided power spectral density \( N_s \).

Our focus will be on the discrete-time system model and it is obtained by sampling the received signals at rate \( N_s/T \), where \( N_s \) represents the oversampling factor (\( N_s = 1 \) denotes symbol-spaced, and fractionally-spaced otherwise). Letting \( t_{\text{oa}} \) denote the sampling phase for the \( l \)th receive antenna, we define \( h_{\text{oa}}[i,m] = h_{\text{oa}}((iT + mT / N_s + t_{\text{oa}}) \) as the \( m \)th (\( 0 \leq m \leq N_s - 1 \)) sample of \( h_{\text{oa}}(t) \) during the \( i \)th symbol period. It is assumed that \( h_{\text{oa}}[i,m] \) is nonzero only within the interval \(-N_s \leq i \leq N_s \), where \( N_s \) and \( N_c \) are referred to as “anti-causal channel memory” and “causal channel memory”, respectively. Further, we define \( N = N_s + N_c \) as the “overall channel memory” (\( N = N_s \) denotes a causal channel). The noise sample is denoted by \( n_l[i,m] = n_l(iT + mT / N_s + t_{\text{oa}}) \), which has an autocorrelation function given by

\[
E[n_l[i,m]n_l[i2,m2]] = \begin{cases} N_s g((i - i2)T + (m1 - m2)T / N_s) & (l1 = l2) \\ 0 & (l1 \neq l2) \end{cases}
\] (5)

Thus, the \( m \)th sample of \( x_l(i) \) during the \( i \)th symbol period is given by

\[
x_l[i,m] = x_l(iT + mT / N_s + t_{\text{oa}}) = \sum_{l=1}^{N_s} d_{l}[i-j] h_{\text{oa}}[j,m] + n_l[i,m]
\] (6)

Define vector \( x[i] \) which contains the \( N_s \) samples of \( x_l[i,m] \), i.e., \( x[i] = [x_l[i,0]...x_l[i,N_s-1]]^T \), where \((.)^T\) denotes the transpose of a matrix or a vector. Thus, we can write

\[
x[i] = \sum_{l=1}^{N_s} H_{l} d_{l}[i] + n[i]
\] (7)

where \( d_{l}[i] = [d_{l}[i+N_s]...d_{l}[i-N_s]]^T \), \( n[i] = [n_l[i,0]...n_l[i,N_s-1]]^T \), and

\[
H_{l} = \begin{bmatrix}
        h_{l}[\text{-}N_s,0] & \ldots & h_{l}[N_s,0] \\
        \vdots & \ddots & \vdots \\
        h_{l}[\text{-}N_s,N_s-1] & \ldots & h_{l}[N_s,N_s-1]
\end{bmatrix}
\] (8)

Further, by using the operator \( \text{vec} \) given by [14]

\[
\text{vec}(A_{1}...A_{L}) = [A_{1}^T...A_{L}^T]^T
\] (9)

we define vector \( x[i] = \text{vec}([x_1[i]...x_L[i]]) \), which contains all the \( N_s L \) samples during the \( i \)th symbol period, and can be expressed as

\[
x[i] = \sum_{l=1}^{L} H_{l} d_{l}[i] + n[i]
\] (10)

where \( H_{l} = \text{vec}([H_{l1}...H_{lL}]) \) is the associate CIR matrix of size \( N_s L \times (N+1) \), and \( n[i] = \text{vec}([n_1[i]...n_L[i]]) \) is the noise vector.
III. LAYERED SPACE-TIME EQUALIZATION

A. Algorithm description

Our proposed layered space-time equalization (LSTE) structure uses both interference cancellation and equalization to perform detection. An example of this structure is shown in fig.1. At the $j$th stage, there are $k^{(j)}$ undetected data streams and the best $m^{(j)}_1$ of these (in the MMSE sense) are selected for tentative detection by the equalizer. The tentative decisions of the best $m^{(j)}_0$ data streams out of $m^{(j)}_1$ (also in the MMSE sense) are then output, and are canceled from the received signals before they are passed to the next stage. Note that at each stage when tentatively detecting the $m^{(j)}_1$ data streams, the remaining $(K^{(j)}-m^{(j)}_1)$ data streams are regarded as interferers, and the number of undetected data streams $k^{(j)}$ can be expressed as $K^{(j)}=K-\sum_{m=1}^{j} m^{(m)}_0$.

The advantage of our approach is that by selecting various $m^{(j)}_1$ and $m^{(j)}_0$ combinations, tradeoffs between performance and complexity can be achieved. For instance, in fig.1a, at each stage, a MIMO equalizer is employed to tentatively detect all the undetected data streams (i.e., $m^{(j)}_1=K^{(j)}$), and $m^{(j)}_0=2$ data streams among them are selected for output. Therefore, $m^{(j)}_1=K^{(j)}=K-2(j-1)$, and the required number of stages is $K/2$. This should be compared to the structures proposed by [7-8], which are based on MISO equalizers with $K^{(j)}=K-j+1$ and $m^{(j)}_1=m^{(j)}_0=1$ for each stage, and therefore need $K$ stages, as shown in fig. 1b.

The $j$th stage of our proposed LSTE structure is illustrated in fig.2, where a MIMO-DDFSE equalizer is employed. The MIMO-DDFSE can reduce to the special cases of MIMO-DFE, MISO-DFE and MISO-DDFSE, which allow a wide range of comparisons. In the MIMO-DDFSE, the $LN_s$ outputs of the receive filters are passed through a feed-forward filter (FFF) which spans $F$ symbol periods with $N_s$ samples per symbol period and therefore has a length $LN_sF$. The FFF performs interference suppression and channel equalization, yielding $m^{(j)}_1$ outputs which correspond to the $m^{(j)}_1$ selected data streams for detection. The FFF outputs are input into a vector DDFSE (VDDFSE). The scalar DDFSE was originally developed in [9] and was shown to be an effective reduced-state MLSE [15]. The VDDFSE is composed of a vector Viterbi algorithm (VA) sequence estimator of memory $\mu^{(j)}$ and a symbol-wise FBF of length $m^{(j)}_1(B-\mu^{(j)})$, which utilizes the tentative decisions of the VA sequence estimator. Given $M$-ary modulated signals (e.g., $M=4$ for QPSK), the number of VA states grows exponentially as $M^{m^{(j)}_1+\mu^{(j)}}$. Therefore, VDDFSE can be viewed as a reduced-state vector MLSE (VMLSE), since
\( \mu^{(j)} \) can be chosen to be smaller than the actual channel memory. Note that when \( \mu^{(j)} = 0 \), DDFSE reduces to DFE, which is a special case of DDFSE with only \( M \) VA states.

We use notation \( \text{EQ}(\mu = \mu^{(j)}, m_i = m_i^{(j)}, m_o = m_o^{(j)}) \) to describe the LSTE structures, where \( \mu \), \( m_i \) and \( m_o \) in the brackets refer to \( \mu^{(j)} \), \( m_i^{(j)} \) and \( m_o^{(j)} \), their values for the first stage, and \( \text{EQ} \) denotes the equalizer type for each stage, which is either a DFE when \( \mu = 0 \) or DDFSE when \( \mu > 0 \). Later stages use the same \( m_o \) and if possible, the same \( m_i \). In particular, at the \( j \)th stage, if the number of the undetected data streams is not smaller than \( m_i \) (i.e., \( K^{(j)} \geq m_i \)), then \( m_i^{(j)} = m_i \), otherwise \( m_i^{(j)} = K^{(j)} \). Further, we let \( m_i^{(j)} \mu^{(j)} = m_j \mu \), such that the number of VA states \( M^{(j)} \mu^{(j)} \) remains constant throughout the detection. Some special examples of the LSTE structures are as follows:

a. \( m_i^{(j)} = m_o^{(j)} \) for each stage: When \( m_i = m_o = K \), with \( \mu = 0 \) our proposed LSTE structure is denoted by DFE \((\mu = 0, m_i = K, m_o = K)\) and reduces to the single-stage MIMO-DFE in [10-11], where all the signals are detected simultaneously. When \( m_i = m_o = 1 \) for all stages, as shown in fig.1b, with \( \mu = 0 \) the LSTE structure is equivalent to the wideband BLAST structure of [7], which is denoted by DFE \((\mu = 0, m_i = 1, m_o = 1)\). The same configuration but with \( \mu > 0 \) corresponds to its DDFSE counterpart.

b. \( m_i^{(j)} < m_o^{(j)} = K^{(j)} \) for each but the last stage: An explicit example is shown in fig.1a, where \( m_i^{(j)} = K^{(j)}, m_o^{(j)} = 2 \) for each stage, and \( K/2 \) stages are required in total. With \( \mu = 0 \) this is denoted by DFE \((\mu = 0, m_i = K, m_o = 2)\). By reducing the number of output data streams to only one for each stage, i.e., \( m_i^{(j)} = K^{(j)}, m_o^{(j)} = 1 \), we have the structure of DFE \((\mu = 0, m_i = K, m_o = 1)\), with a total number of \( K \) stages. The advantage of this is that when all the \( K^{(j)} \) undetected data streams are tentatively detected, and only some of them are output (i.e., \( m_i^{(j)} < m_o^{(j)} = K^{(j)} \), except for the last stage which is \( m_i^{(j)} = m_o^{(j)} = K^{(j)} \)), tentative decisions of the \((m_i^{(j)} - m_o^{(j)})\) residual data streams contribute to performance improvement since the FBF has a higher degree of freedom [10].

c. \( m_i^{(j)} < m_o^{(j)} < K^{(j)} \) for each but the last stage: With \( \mu^{(j)} > 0 \), the use of \( m_i^{(j)} = K^{(j)} \) is constrained by its overwhelming complexity especially at earlier stages, which requires \( M^{(j)} m_o^{(j)} \mu^{(j)} \) VA states. Thus, \( m_i^{(j)} < K^{(j)} \) is used to reduce the complexity, i.e., the number of the tentatively detected data streams is smaller than that of the undetected data streams. For example, with \( K = 4 \) transmit data streams, we can employ four stages. At each of the first three stages two data streams are selected and tentatively detected, and only the better one of them is output, i.e.,
\[m_j^{(0)} = 2, m_j^{(0)} = 1 \text{ for } j \leq 3.\] At the last stage, i.e., \(j = 4\), we have \(m_j^{(0)} = m_j^{(0)} = 1\). In this paper, we consider this structure with \(\mu^{(0)} = 1\) and \(\nu^{(0)} = 2\) for the first three stages and last stage, respectively, which is denoted by DDFSE \((\mu = 1, m_1 = 2, m_0 = 1)\).

**B. Equalizer Coefficients**

Our equalizer coefficients for the DDFSE with a VA memory \(\mu^{(0)}\) and FBF memory \((B - \mu^{(0)})\) are the same as the coefficients for the DFE with an FBF memory \(B\). Our derivation is an extension of that in [10] to the general case where an arbitrary number of data streams are selected for tentative detection, based on the MMSE criterion.

Assuming perfect cancellation of previously output signals, at a particular stage, during the \(i\)-th symbol period, the signals spanned by the FFF are denoted by \(\pi[i] = \text{vec}(\{x[i]...x[i-F+1]\})\), which can be expressed in matrix notation as:

\[
\pi[i] = \sum_k \mathbf{H}_k \mathbf{d}[i] + \mathbf{n}[i]
\]

(11)

where \(k\) denotes the summation over the undetected data streams, \(\mathbf{d}[i] = [d[i+N_0]...d[i-N_c-F+1]]^T\), \(\mathbf{n}[i] = \text{vec}(\{n[i]...n[i-F+1]\})\), and

\[
\mathbf{H}_k = \begin{bmatrix}
H_k & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & H_k
\end{bmatrix}
\]

(12)

which is an \((LN,F) \times (N + F)\) block Toeplitz matrix.

Without loss of generality, it is assumed that at a particular stage, \(\mu^{(0)} = \mu\), \(m_j^{(0)} = m_j\), and \(m_j^{(0)} = m_0\). Let \(d_{km}[i]\) denote the signal transmitted by the \(k_m\)-th \((k_m \in \{1,...,K\})\) antenna, which corresponds to the \(m\)-th \((m = 1,...,m_j)\) output branch of VDDFSE. Also, let \(\hat{d}_{km}[i]\) and \(\hat{d}'_{km}[i]\) denote the associated VA state transition element and the tentative decision feedback (FB) signal of VDDFSE, respectively. Accordingly, we define \(s[i] = [d_{k1}[i] \ldots d_{km_j}[i]]^T\) as the transmit data vector under detection. \(\hat{s}[i]\) and \(\hat{s}'[i]\) are defined similarly, which denote the VA state transition vector and FBF input vector, composed of \(\hat{d}_{km}[i]\) and \(\hat{d}'_{km}[i]\), respectively. Further, we define \(\hat{s}_{va}[i] = \text{vec}(\hat{s}[i-1]...\hat{s}[i-\mu])\) as the VA state transition sequence of length \(m_1\mu\), and \(\hat{s}'_{va}[i] = \text{vec}(\hat{s}'[i-\mu-1]...\hat{s}'[i-B])\) as the tentative decision feedback sequence of length \(m_1(B - \mu)\). Note that tentative decision feedback sequence \(\hat{s}'_{va}[i]\) is maintained for each VA state, rather than a single tentative decision feedback sequence for all. That is, each \(\hat{s}_{va}[i]\) has a corresponding
\( \hat{s}_{FB}[i] \). Combination of \( \hat{s}_{VA}[i] \) and \( \hat{s}_{FB}[i] \) yields the most recent \( m_B \) signals in VDDFSE, which are referred to as the “past” signals, and are expressed as \( \hat{s}_{past}[i] = vec\left( \hat{e}_{VA}[i], \hat{e}_{FB}[i] \right) \).

As illustrated in fig.2, let \( W \) denote the FFF weight matrix of size \( N_m \times m_L \), and \( \tilde{V} \) denote the overall equivalent CIR (ECIR) matrix seen by the VDDFSE block, which is given by \( \tilde{V} = vec\left( V_{VA}, V_{FB} \right) \) where \( V_{VA} \) and \( V_{FB} \) are the ECIR matrices corresponding to the VA state transition sequence \( \hat{s}_{VA}[i] \) and tentative decision feedback sequence \( \hat{s}_{FB}[i] \), respectively. \( V = vec\left( V_{VA}, V_{FB} \right) \) is the combined ECIR matrix associated with the “past” \( m_B \) signals, which is of size \( m_B \times m_L \). Assuming all the detected signals at each stage have the same decision delay \( \Delta \), the soft estimate of the \( i \)th detected signal vector is defined as

\[
\tilde{s}[i-\Delta] = W H x[i] - V H \hat{s}_{past}[i-\Delta] = y[i] - z_{VA}[i] - z_{FB}[i]
\]

(14)

where \( y[i] = W H x[i] \) is the FFF output, \( z_{FB}[i] = V_{FB} H \hat{s}_{FB}[i-\Delta] \) is the FBF output, as shown in fig.2, and \( z_{VA}[i] = V_{VA} H \hat{s}_{VA}[i-\Delta] \) is the VA state transition output.

The total path metric of VDDFSE for the \( i \)th detected signal vector is given by

\[
\Lambda_i = \left[ \tilde{y}[i-\Delta] - c \right]^H = \sum_{m=1}^{m_L} \left| \tilde{d}_{m,i}[i] - c_m \right|^2 = \sum_{m=1}^{m_L} P_m
\]

(15)

where \( c_m \) is the \( m \)th element of a candidate vector \( e \), \( \tilde{d}_{m,i}[i] \) denotes the \( m \)th element of \( \tilde{s}[i] \), (i.e., the soft estimate of \( d_{m,i}[i] \)), and \( p_m = \left| \tilde{d}_{m,i}[i] - c_m \right|^2 \) represents the path metric of the \( m \)th branch of VDDFSE. The selected data streams are tentatively detected by minimizing \( \sum_i \Lambda_i \).

We define the difference between \( s[i] \) and its soft estimate \( \tilde{s}[i] \) as the error vector \( e[i] = s[i] - \tilde{s}[i] \). The optimum weight matrices \( W \) and \( V \) are determined to minimize

\[
J = trace\left( E[i] e[H][i] \right) = trace(R_m)
\]

(16)

where \( (.)^H \) denotes the complex-conjugate transpose, and \( R_m = E[i] e[H][i] \) is the error autocorrelation matrix. Letting \( MSE_{\tilde{d}_{m,i}} = E[\tilde{d}_{m,i}[i] - d_{m,i}[i]]^2 \) be the mean square error (MSE) with respect to \( d_{m,i}[i] \), which is the \( m \)th diagonal element of \( R_m \), (16) is equivalent to \( J = \sum_m MSE_{\tilde{d}_{m,i}} \).

Letting \( \hat{H}_{\tilde{d}_{m,i}} \) denote the estimate of \( \bar{H}_{\tilde{d}_{m,i}} \), and (\( \hat{H}_{\tilde{d}_{m,i}} \)), be the \( i \)th column of \( \hat{H}_{\tilde{d}_{m,i}} \), we define

\[
\hat{H}_{past} = \left[ (\hat{H}_{\tilde{d}_{1,i}})_{\Delta s_1 \times N_d} \cdots (\hat{H}_{\tilde{d}_{m,i}})_{\Delta s_1 \times N_d} \right]
\]

(17a)
\[
\mathbf{\hat{H}}_{\text{past}} = \left[ (\mathbf{H}_{k})_{\Delta s+2+N_s} \ldots (\mathbf{H}_{k})_{\Delta s+2+N_s} \ldots (\mathbf{H}_{k})_{\Delta s+B+1+N_s} \ldots (\mathbf{H}_{k})_{\Delta s+B+1+N_s} \right] \quad (17b)
\]

which are the CIR matrices associated with \( s(i-\Delta) \) and \( \hat{s}_{\text{past}}(i-\Delta) \), respectively. Our equalizer design and data selection are under the assumption of correct decision feedback, perfect channel estimation and perfect interference cancellation of the previously output signals. By the orthogonality principle, \( E[\mathbf{x}[i]|\mathbf{x}_{\text{past}}^H[i]] = \mathbf{0} \). Some derivation similar to [10] yields the optimum FFF matrix \( \mathbf{W} \) as

\[
\mathbf{W} = \mathbf{R}^{-1}E[\mathbf{x}[i]|\mathbf{x}_{\text{past}}^H[i]] - \mathbf{R}^{-1}\mathbf{H}_{\text{pres}}
\]

where \( \mathbf{R} \) is the autocorrelation matrix of the input signal vector clear of the “past” data, given by

\[
\mathbf{R} = E[\mathbf{x}[i]|\mathbf{x}_{\text{past}}^H[i]] - E[\mathbf{x}_{\text{past}}[i-\Delta]|\mathbf{x}_{\text{past}}^H[i]] = (\Sigma_k \mathbf{H}_{k}^H + \mathbf{R}_{\text{nc}}) - \mathbf{\hat{H}}_{\text{past}}^H \mathbf{\hat{H}}_{\text{past}}
\]

where \( \mathbf{R}_{\text{nc}} = E[\mathbf{x}[i]|\mathbf{x}_{\text{past}}^H[i]] \) denotes the noise autocorrelation matrix whose elements are given by (5).

In particular, when \( N_s = 1 \) and \( g(nT) = \begin{cases} 1 & (n = 0) \\ 0 & (n \neq 0) \end{cases}, \) \( \mathbf{R}_{\text{nc}} \) reduces to an identity matrix \( \mathbf{R}_{\text{nc}} = N_q \mathbf{I} \). The optimum ECIR matrix \( \mathbf{V} \) can be expressed as

\[
\mathbf{V} = \mathbf{E}[\mathbf{x}_{\text{past}}[i-\Delta]|\mathbf{y}_{\text{past}}^H[i]] \mathbf{W} = \mathbf{\hat{H}}_{\text{past}}^H \mathbf{W}
\]

The resulting error autocorrelation matrix \( \mathbf{R}_{\text{er}}, \) as a function of \( \Delta \) is given by

\[
\mathbf{R}_{\text{er}} = \mathbf{I} - \mathbf{\hat{H}}_{\text{past}}^H \mathbf{R}^{-1} \mathbf{\hat{H}}_{\text{pres}}
\]

where \( \mathbf{I} \) denotes an \( N_s \times N_s \) identity matrix. The MSE with respect to \( d_{\text{in}}[i] \) is the \( m \)th diagonal element of \( \mathbf{R}_{\text{er}}, \) which is expressed as

\[
\text{MSE}_{\text{in}} = E[(\mathbf{x}_{\text{in}}[i] - d_{\text{in}}[i])^2] = 1 - (\mathbf{\hat{H}}_{\text{in}})^H_{\Delta s + 1 + N_s} \mathbf{R}^{-1} (\mathbf{\hat{H}}_{\text{in}})_{\Delta s + 1 + N_s}
\]

The \( m \) data streams which minimize \( J = \sum_{m=1}^{m} \text{MSE}_{\text{in}} \) are selected for tentative detection, and \( m_{\text{opt}} \) of them with the smallest MSEs are output and are canceled from the input signals. The decision delay can be optimized, as will be discussed in section III-D, to also minimize \( J = \sum_{m=1}^{m} \text{MSE}_{\text{in}} \). The DDFSE output SINR with respect to \( d_{\text{in}}[i] \) can be expressed by \( \text{MSE}_{\text{in}} \) as

\[
\text{SINR}_{\text{in}} = \frac{1}{\text{MSE}_{\text{in}}} - 1 = (\mathbf{\hat{H}}_{\text{in}})^H_{\Delta s + 1 + N_s} \mathbf{R}_{\text{nc}} (\mathbf{\hat{H}}_{\text{in}})_{\Delta s + 1 + N_s}
\]

where \( \mathbf{R}_{\text{nc}} = \mathbf{R} - (\mathbf{\hat{H}}_{\text{in}})_{\Delta s + 1 + N_s} (\mathbf{\hat{H}}_{\text{in}}^H)_{\Delta s + 1 + N_s} \). Thus, minimizing the MSE is equivalent to maximizing the output SINR.
When \( m_{ij}^{(i)} < K^{(j)} \), the \((K^{(j)} - m_{ij}^{(j)})\) remainders that are not selected for tentative detection are regarded as “interferers”, and the FFF suppresses both interference and noise, yielding a high output SINR. Hence, although the FFF output noise (including interference) is colored, it has little effect on the performance of the VA sequence estimator, which is derived upon the assumption of white noise.

Note that when \( N_u = N_v = 0 \), the channel is flat fading, and \( H_k \) reduces to vector \( h_j \). In addition, if \( R_m = N_0 I \), \( R \) reduces to a block diagonal matrix. Letting \( \hat{h}_k \) denote the estimate of \( h_k \), it can be shown that the optimum FFF matrix \( W \) is given by

\[
W = \left[ \sum_k \hat{h}_k \hat{h}_k^H + N_0 I \right]^{-1} \left[ \hat{h}_1 \ldots \hat{h}_{N_t} \right]
\]  

(24)

The resulting ECIR matrix \( V \) is zero. The MSE with respect to \( d_{m,i} \) is given by

\[
MSE_{d_{m,i}} = 1 - \left[ \sum_k \hat{h}_k \hat{h}_k^H + N_0 I \right]^{-1} \hat{h}_{m,i}
\]  

(25)

Therefore, it can be deduced that in flat fading channels, all the LSTE structures reduce to their narrow–band counterparts, which have an FFF spanning only one symbol period, and no FBF.

C. Channel Estimation

In the work presented, a least-squares (LS) technique is used for channel estimation, where a training sequence of \( L_t \) symbols is prefixed to each packet. Assuming less than 100% excess bandwidth, the CIR can be fully represented by its \( T/2 \)-spaced samples [18]. For the simplicity of presentation, we take samples of the received signals generated by the training sequences at times \( t = iT/2 \). Letting \( c_k[i] \) denote the \( i \)th training symbol transmitted by the \( k \)th antenna, define \( r_{i,m} \) as the \( m \)th \((m = 0,1)\) sample of the signal at the \( l \)th receiver filter output during the \( i \)th symbol period, which is given by:

\[
r_{i,m} = \sum_{k=1}^{K} \sum_{j=N_u}^{N_v} c_k[i-j] h_{k,j,m} + n_{i,m}
\]  

(26)

where \( h_{k,j,m} = h_{k,j+mT/2} \). With given \( m \), the \((L_t - N)\) samples of \( r_{i,m} \) used for channel estimation are written in a vector form of \( r_{i,m} = [r_{i,N_u,m}, \ldots, r_{i,L_t - N_u - 1,m}]^T \) and expressed as

\[
r_{i,m} = \sum_{k=1}^{K} C_k h_{k,m} + n_{i,m}
\]  

(27)

where \( h_{k,m} = [h_{k,-N_u,m}, \ldots, h_{k,N_v,m}]^T \), \( n_{i,m} = [n_{i,N_u,m}, \ldots, n_{i,L_t - N_u - 1,m}]^T \), and

\[
C_k = \begin{bmatrix}
    c_k[N] & \ldots & c_k[0] \\
    \vdots & \ddots & \vdots \\
    c_k[L_t-1] & \ldots & c_k[L_t-N-1]
\end{bmatrix}
\]  

(28)
By defining a length-$K(N+1)$ vector $h_i[m] = \text{vec}(\{h_{1i}[m], \ldots, h_{K_i}[m]\})$, and a size- $(L_r - N) \times K(N+1)$ matrix $C = [c_1 \ldots c_N]$, (27) can be written as

$$r_i[m] = C h_i[m] + n_i[m]$$

(29)

The LS estimate of CIR vector $h_i[m]$ is given by

$$\hat{h}_i[m] = C^* r_i[m]$$

(30)

where $C^* = (C^H C)^{-1} C^H$ is the pseudo-inverse of $C$. This work should be repeated $2L$ times for each receive antenna and each sampling instant. To guarantee the existence of $C^*$, $C$ must be full column rank. Consequently, the minimum required length of the training sequence is $L_r = N + K(N+1)$, as shown in [19]. The continuous-time CIR can be estimated by interpolation.

The estimation error with respect to $h_i[m]$ is defined as $e_i[m] = h_i[m] - \hat{h}_i[m]$, and we measure the channel estimation accuracy by using the sum of the error variances, normalized by the sum of the channel gain variances, that is,

$$\sigma^2 = \frac{1}{L} \sum_{j=1}^{L} \frac{\|e_i[j]\|^2}{\sum_{i=1}^{K} \|h_i[j]\|^2} = \frac{1}{L} \frac{1}{\sum_{i=1}^{K} \|h_i[j]\|^2} \sum_{j=1}^{L} \text{trace}(R_i'[j])$$

(31)

where $R_i'[j] = e_i[j] e_i^H[j]$ is the error autocorrelation matrix, and $R_i'[j] = h_i[j] h_i^H[j]$ is the channel gain autocorrelation matrix. With given $C$ and AWGN noise vector $n_i[m]$, we have $R_i'[j] = N_0 (C^H C)^{-1}$, and (31) can be written as

$$\sigma^2 = 2LN_0 \frac{\text{trace} \left( C^H C \right)}{\sum_{j=1}^{L} \sum_{i=1}^{K} \text{trace}(R_i'[j])}$$

(32)

It suggests that the estimation error variance is proportional to the noise power spectral density, i.e., proportional to the reciprocal of SNR. Thus, selection of the training sequences can be achieved simply by minimizing $\text{trace} \left( C^H C \right)$, which has been discussed in [19] and numerical results were given with a signal constellation of BPSK. In this paper, however, we do not consider the design of training sequences, since our focus is on different LSTE structures.

D. Timing Recovery

Timing recovery plays an important role in equalizers, which includes selection of both sampling phase (non-integer) and decision delay (integer). Here we propose a simple timing recovery method that handles sampling phase and decision delay separately. In [7], an
exhaustive search is used to jointly determine the sampling phase and decision delay in terms of minimizing the bit error rate (BER), which requires significant searching for the low BER scenario. In [17], the timing recovery method is based on approximately maximizing the equalizer output SINR, assuming all the postcursor interference has been canceled. However, it is derived for a MISO equalizer, and a short length FBF may not have the capability to remove all the postcursor interference.

Unlike the method of [17], our sampling method for a symbol-spaced LSTE structure is designed for a MIMO equalizer, independent of the decision delay, and different receive filters exploit different sampling phases. Our sampling phase \( t_{\ell t} \) for the \( \ell \)th receive filter maximizes the output SINR of the \( \ell \)th antenna (or SINR of \( \ell \)th input of equalizer) in a straightforward expression as:

\[
SINR_{\ell t} = \frac{\sum_{k=1}^{K} \left| \hat{h}_a (t_{\ell t}) \right|^2}{\sum_{k=1}^{K} \sum_{j=0}^{N_{tj}} \left| \hat{h}_a (jT + t_{\ell t}) \right|^2 + N_o}
\]

where \( \hat{h}_a (jT + t_{\ell t}) \) is the channel estimate with sampling phase \( t_{\ell t} \), the numerator denotes the total received power of the desired signals after sampling at the \( \ell \)th antenna, and the first term in the denominator is the ISI power.

The decision delay is also important, especially for an equalizer with short-length filters [16]. In [10], the decision delay \( \Delta \) of a single stage MIMO-DFE is fixed to be \( \Delta = F + N_c - B - 1 \). In [7] it is determined by an exhaustive search and remains constant for all stages of the so-called wideband BLAST structure. As described in section III-B, at each stage, our method of finding the decision delay is to minimize (16), i.e., minimize the sum of MSEs of the tentatively detected data streams. Different decision delays are used for different stages. An appropriate range is required within which the decision delay is searched. It was shown by [16] that the optimum decision delay converges to \( (F-1) \) (i.e., \( \Delta \leq F - 1 \)) as the FBF length gets large. Another restriction \( \Delta \leq F + N_c - B - 1 \) was also given by [16] to guarantee that the coefficients derivation always makes sense. In order to satisfy both the constraints, the decision delay \( \Delta \) should be optimized in the range of \( 0 \leq \Delta \leq \Delta_{\text{max}} \), where

\[
\Delta_{\text{max}} = \begin{cases} 
F - 1 & (N_c \geq B) \\
F + N_c - B - 1 & (N_c < B) 
\end{cases}
\]

Because our timing recovery scheme finds the sampling phase and decision delay separately, it is easier to perform than the other approaches.
IV. COMPUTATIONAL COMPLEXITY

We approximately evaluate and compare the computational complexity of LSTE structures by only counting the number of complex multiplications. We divide the complexity into two parts: 1) Initialization complexity in which complexity associated with calculating coefficients and data selection is found, and 2) Signal processing complexity in which complexity associated with processing the symbols is found. In our analysis, it is assumed that $B > 0$.

A. Initialization: Assuming quasi-static channels, computation of the LSTE coefficients is needed only once for every packet.

1) Computation of $r$ (19): computation of $r$ for all stages has been included in that of $\sum_{k=1}^{K} \tilde{H}_i \tilde{H}_j^*$, which requires complexity of $L' N_1^2 (N+1)(N+2)K/2$ for $N \leq F - 1$, and $L' N_1^2 F (2N-F+3)K/2$ for $N > F - 1$.

2) Data selection: At the $j$th stage, the computation of the MSE for the tentatively detected signal given by (22) requires the inverse of $r$, with approximately $L' N_1^2 F^3 / 6$ complex multiplications [10]. In order to select $m_{ij}$ data streams from the $K^{(i)}$ undetected data streams, one needs to compute $r^{-1}$ for $\left( \begin{array}{c} K^{(i)} \\ m_{ij} \end{array} \right)$ times. In addition, the optimization of the decision delay requires searching in the range given by (34) for $r = \Delta_{\text{max}} + 1$ times. Therefore, a total of $r \left( \begin{array}{c} K^{(i)} \\ m_{ij} \end{array} \right)$ times manipulations are needed to find the optimum $r^{-1}$. With given $r$.

3) FFF and ECIR matrices: No additional computation is needed for the FFF matrix $w$ (18), which has been included in computation of MSEs. The computation of the ECIR matrix $\tilde{v}$ (13) reduces to that of $v$. At the $j$th stage, after the optimum $w$ is determined, the ECIR matrix $v$ is computed only once by using (20), with $m_{ij}^2 BLN, F$ complex multiplications.

B. Signal Processing: We investigate the complexity of signal processing required for each symbol period. The number of times that this work is repeated within a packet length is simply the number of symbols per packet.

1) FFF: At the $j$th stage, the FFF has $m_{ij}^2 LN, F$ taps, and therefore a total number of $m_{ij}^2 LN, F$ complex multiplications are needed.
2) VDDFSE: Since the FBF in VDDFSE can be implemented by using additions, we only count the complexity of VA. The total path metric for VA given by (15) needs to be calculated, by separately computing the path metric \( p_m \) for each VDDFSE branch. Thus, VDDFSE reduces to \( m^{(i)} \) scalar DDFSEs with \( M^{m^{(i)}} \mu \) VA states each. Each VDDFSE branch requires \( M^{m^{(i)}} \mu \) complex multiplications and the overall complexity is \( m^{(i)} M^{m^{(i)}} \mu \).

The above complexity analysis is summarized in table I for easy reference. A numerical example of complexity is also provided in table II, where different \( T \)-spaced (\( N_s = 1 \)) LSTE structures are employed, with \( K = 4 \), \( L = 4 \) antennas, both the FFF and FBF spanning 4 symbol periods, respectively. QPSK modulation (i.e., \( M = 4 \)) is used, and each packet consists of 80 symbols. We count the number of complex multiplications for both initialization and processing over each packet. When a specific packet rate (e.g., in GSM the packet duration is 576.92 \( \mu s \), and therefore the packet rate is 1733 packets/sec) is considered, the complexity figure in operations per second can be obtained by multiplying the number of complex multiplications by the packet rate. The notations of the demonstrated LSTE structures here have been introduced in detail in section III-A. Note that all the DDFSE’s have \( M^{m^{(i)}} = 16 \) VA states throughout, which means that \( \mu^{(i)} = 2 \) for \( m^{(i)} = 1 \) and \( \mu^{(i)} = 1 \) for \( m^{(i)} = 2 \). In particular, in DDFSE \((\mu = 1, m_1 = 2, m_o = 1)\), \( \mu^{(i)} = 1 \) is used for the first three stages and \( \mu^{(i)} = 2 \) for the last stage. While in DDFSE \((\mu = 2, m_1 = 1, m_o = 1)\), \( \mu^{(i)} = 2 \) for all stages.

In table II, we take the single-stage DFE \((\mu = 0, m_1 = 4, m_o = 4)\) as a reference to compare with other LSTE structures, which has the least complexity. DDFSE \((\mu = 1, m_1 = 2, m_o = 1)\) has the most complexity of 613%. The other DFE structures have complexity within 3 times of the reference. It should be noticed that the 4-stage DFE \((\mu = 0, m_1 = 1, m_o = 1)\) requires more initialization complexity than the 4-stage DFE \((\mu = 0, m_1 = 4, m_o = 1)\), since at the \( j \)th stage, the former has \( \kappa^{(i)} \) different combinations for data selection, while the latter has only one. On the other hand, since DFE \((\mu = 0, m_1 = 4, m_o = 1)\) requires 2.5 times more processing complexity than DFE \((\mu = 0, m_1 = 1, m_o = 1)\), the overall complexity of these two structures is comparable. It is also important to consider the relative weights of the initialization and signal processing complexity in comparison with the packet length.

As a final point it must be noted that in our comparisons we have not taken channel estimation into account since it is common to all the LSTE structures. However it needs to be considered if the total complexity is to be determined and as shown in (30), channel estimation
requires computation of \( c^* = (c^H c)^t c^H \). This task includes computation of matrix product \( c^H c \), inverse of \( c^H c \) and matrix product \( (c^H c)^t c^H \), which approximately costs \( K^2(N+1)^2(L_r - N) \), \( K^3(N+1)^2/6 \) and \( K^2(N+1)^2(L_r - N) \) complex multiplications respectively. With given \( c \), one should compute \( h_t(n|m) (m = 0, 1; l = 1, \ldots, L) \) in (30) \( 2L \) times, each requiring \( K(N+1)(L_r - N) \) multiplications.

V. SIMULATION RESULTS

A. Setup

We use simulation results to demonstrate the performance of our proposed LSTE structures, where the symbol-spaced (T-spaced) LSTE structures shown in table II are focused on, with \( K = 4 \) transmit antennas and \( L = 4 \) receive antennas. Results for \( L = 6 \) and fractionally-spaced (T/2-spaced) LSTE structures are also provided. As stated in section IV, the VA sequence estimator employs a memory of \( \mu^{(j)} = 2 \) for \( m^{(j)} = 1 \) and \( \mu^{(j)} = 1 \) for \( m^{(j)} = 2 \), respectively.

Both the transmit filter and the receive filter use a square-root raised-cosine pulse, with a roll-off factor of 0.35. Each data packet is composed of 80 QPSK (i.e., \( M = 4 \)) data symbols, over which the channel is assumed to be static. The frequency selective channels in our simulations contain \( N_p = 50 \) paths and span from time zero to five times the RMS delay spread (i.e., \( 5\sigma \)), by following the discrete power delay profile of (2). The delay for the \( n \)th \((n = 0, \ldots, N_p - 1)\) path can be expressed as \( \sigma_n = 5\sigma n / N_p \). We consider the causal channel with \( N_a = 0 \) and \( N_c = 8 \), which is implemented by truncating the convolution result of the physical (untruncated) CIR with the pulse shape. This truncated overall CIR contains more than 90% of the energy on average. The channel is estimated by using randomly generated training sequences, and further details follow. The SNR is defined as the spatial average ratio of the received signal power (from all \( K \) transmit antennas) to noise power at each receive antenna, which can be expressed as \( \text{SNR} = KE_s / N_0 \), since the channel delay profile is normalized to unity on average. Up to 50000 simulation runs are used. The T-spaced LSTE structure follows the sampling method proposed in section III-D, and the decision delay is also determined by using the method described in section III-D.

B. Results

In all the BER results that follow, the DFE and DDFSE based LSTE structures are shown with solid lines and dashed lines, respectively.
In fig.3, the average BER performances of the \( T \)-spaced \((N_c=1)\) LSTE structures are shown, where both the FFF and FBF span 4 symbol periods, with \( K=4 \) transmit antennas and \( L=4 \) receive antennas, a delay spread of 1.25 symbol periods \((1.25T)\) and perfect channel state information (CSI), i.e., perfect channel estimation. The best performing LSTE structure is DDFSE \((\mu=1,m_I=2,m_o=1)\), and DFE \((\mu=0,m_I=1,m_o=1)\) is the worst one. The advantage of DDFSE \((\mu=1,m_I=2,m_o=1)\) over DDFSE \((\mu=2,m_I=1,m_o=1)\) can be interpreted by noting that the former performs vector VA instead of scalar VA for multiple data streams, and has a higher degree of freedom in the FFB than the latter. Note that the single-stage DFE \((\mu=0,m_I=4,m_o=4)\) provides similar performance to DDFSE \((\mu=2,m_I=1,m_o=1)\) at high SNRs, with much less complexity. The 4-stage DFE \((\mu=0,m_I=4,m_o=1)\) and 2-stage DFE \((\mu=0,m_I=4,m_o=2)\) use the same equalizer coefficients as DFE \((\mu=0,m_I=4,m_o=4)\) but only output a part of the data streams at the first stage, and therefore provide better performance than DFE \((\mu=0,m_I=4,m_o=4)\).

To investigate the impact of imperfect channel estimation on performance, we first demonstrate in fig.4 the impact of training sequence length \( L_t \) on the channel estimation error variance defined in \((31)\). \( K=4 \) transmit antennas and \( L=4 \) receive antennas are also used, with an RMS delay spread \( \sigma = 1.25T \). As derived in section III-C, the error variance is inversely proportional to the SNR. In our system configuration, the minimum required length of training sequence is \( L_t = N + K(N+1) = 44 \) (see section III-C). By increasing the training sequence length by only one, i.e., \( L_t = 45 \), the error variance is reduced by nearly 10 times, and it is reduced by about 50 times when \( L_t \) is increased to 50. However, less improvement is achieved by further increasing \( L_t \), which implies that a moderate training sequence length will provide reasonable channel estimation accuracy. It can also be shown that the number of receive antennas and RMS delay spread have little effect on the estimation error variance.

Fig.5 illustrates the impact of imperfect channel estimation on performance of the single-stage DFE \((\mu=0,m_I=4,m_o=4)\) and 4-stage DFE \((\mu=0,m_I=4,m_o=1)\), with the same configuration as fig.3. Different training sequence lengths of \( L_t = 44,50 \) are used, compared to the performance with perfect CSI. Obviously, both DFE \((\mu=0,m_I=4,m_o=4)\) and DFE \((\mu=0,m_I=4,m_o=1)\) with \( L_t = 44 \) provide much worse performance than with perfect CSI. At SNR=24dB the BERs of DFE \((\mu=0,m_I=4,m_o=4)\) and DFE \((\mu=0,m_I=4,m_o=1)\) degrade by about 7 and 11 times from their BERs with perfect CSI, respectively. This suggests that the multistage structure is more...
vulnerable to the channel estimation error, which introduces more error propagation due to
imperfect interference cancellation. With an increased training sequence length of \( L_t = 50 \),
however, at SNR=24dB the relative BER errors generated by imperfect CSI for both the
structures are below 40%. Also, it is as expected that the higher the SNR, the smaller the relative
error. To further demonstrate the effect of imperfect interference cancellation on the
performance of the multistage structure, we also provide the results of DFE \( (\mu = 0, m_f = 4, m_o = 1) \)
with both perfect CSI and perfect interference cancellation. At SNR=24dB, the BER degradation
of DFE \( (\mu = 0, m_f = 4, m_o = 1) \) due to imperfect interference cancellation is about 4 times compared
to perfect cancellation.

In the multistage LSTE structures, the performance for each stage or layer is also of interest.
Fig.6 shows the BER performance of the first stage and last stage (4\(^{th}\) stage) of
DFE \( (\mu = 0, m_f = 4, m_o = 1) \). With both perfect interference cancellation and perfect CSI, the BER of
the last stage is much lower than that of the first stage, benefiting from a higher degree of
freedom. However, it is shown that with imperfect interference cancellation, the last stage
provides worse performance than the first stage in a range of low SNR since high noise power
worsens the effect of error propagation. At a higher SNR, the last stage performs similarly to the
first stage. Imperfect CSI increases the range by yielding more error propagation. In particular,
the threshold value of SNR is about 16dB, 20dB and 24dB with perfect CSI, \( L_t = 50 \) and \( L_t = 44 \),
respectively, beyond which the first stage and last stage of DFE \( (\mu = 0, m_f = 4, m_o = 1) \) have similar
results. Similar trends can be deduced and shown for the stages in between. This suggests that
the overall performance of the LSTE structure is mainly determined by its first layer, and
imperfect channel estimation further increases the performance difference between layers.

In the remainder of the results, we focus on performance of LSTE structures with perfect CSI.

Fig. 7 demonstrates the performance of the \( T \)-spaced LSTE structures, in terms of BER versus
the RMS delay spread normalized to the symbol period, which varies from 0 to 2, with a fixed
SNR=20dB. \( K=4 \) transmit antennas and \( L=4 \) receive antennas are employed, with the FFF and
FFB configuration \( (F=4, B-\mu=4) \) and \( (F=1, B-\mu=0) \). It should be noticed that \( (F=1, B-\mu=0) \) denotes the
narrow-band LSTE structure without an FBF. In this case, the single-stage
DFE \( (\mu = 0, m_f = 4, m_o = 1) \) has the worst performance. With four stages, DFE \( (\mu = 0, m_f = 1, m_o = 1) \)
reduces to the narrow-band BLAST structure of [4] (but based on the MMSE criterion), and
both DFE \( (\mu = 0, m_f = 4, m_o = 1) \) and DDFSE \( (\mu = 1, m_f = 2, m_o = 1) \) have the same performance as
DFE ($\mu = 0, m_1 = 1, m_0 = 1$), since they have no advantage in the FBF degree of freedom. The 2-stage DFE ($\mu = 0, m_1 = 4, m_0 = 2$) provides results in between. As shown in section III-B, for $\sigma = 0$ (i.e., flat fading), the performance of the wideband LSTE structure with an FBF converges to that of its narrow-band counterpart, irrespective of the number of equalizer taps. An explicit interpretation for this is that the narrow-band structure can exploit sufficient channel memory, and no additional equalizer tap is needed. However, the wideband LSTE structure with an FBF is by far superior to its narrow-band counterpart in frequency selective channels ($\sigma > 0$). An interesting observation is that the wideband LSTE structure provides better performance than in flat fading within a range of small delay spread. For example, with ($F=4, B-\mu=4$), DDFSE ($\mu = 1, m_1 = 2, m_0 = 1$) achieves the lowest BER at $\sigma/T = 0.5$, which is about 10 times better than its BER at $\sigma = 0$.

In fig.7, among all the wideband LSTE structures with ($F=4, B-\mu=4$), DDFSE ($\mu = 1, m_1 = 2, m_0 = 1$) provides the best performance, especially when $0.3 < \sigma/T < 1$. The 4-stage DFE ($\mu = 0, m_1 = 4, m_0 = 1$) always has better performance than the single-stage DFE ($\mu = 0, m_1 = 4, m_0 = 4$), especially at lower delay spreads. DFE ($\mu = 0, m_1 = 4, m_0 = 1$) also outperforms DFE ($\mu = 0, m_1 = 1, m_0 = 1$) of [7] for $\sigma/T > 0.5$ with less complexity, since the latter requires more combinations for data selection. The advantage of the DDFSE structure over DFE is attributed to the utilization of channel memory in the VA sequence estimator. The wideband LSTE structures with an FBF, especially the DDFSE versions with $B > \mu > 0$, see their performance improve as they better exploit channel memory and as a result, achieve frequency (or path) diversity. However, as the delay spread increases and exceeds the feedback span, less channel energy is captured due to channel memory truncation and hence the performance starts to drop at some point. The narrow-band structures without an FBF readily see their performance degrade from a delay spread of zero.

Fig.8 illustrates the advantage of the 4-stage DFE ($\mu = 0, m_1 = 4, m_0 = 1$) over the single-stage DFE ($\mu = 0, m_1 = 4, m_0 = 4$) in terms of both performance and complexity, where DFE ($\mu = 0, m_1 = 4, m_0 = 4$) uses the equalizer with ($F=6, B-\mu=4$), and DFE ($\mu = 0, m_1 = 4, m_0 = 1$) employs ($F=2, B-\mu=4$) and ($F=4, B-\mu=4$). The percentages on the right of the graph are the computational complexities normalized to the complexity of DFE ($\mu = 0, m_1 = 4, m_0 = 4$) with ($F=6, B-\mu=4$). It is shown that the 4-stage DFE with ($F=4, B-\mu=4$) outperforms the single-stage DFE with ($F=6, B-\mu=4$) in the range of $\sigma/T < 1.2$, and provides similar performance otherwise, both having similar complexities. Further, it should be stressed that the 4-stage DFE with
provides better performance than the single-stage DFE with \((F = 6, B - \mu = 4)\) in a smaller range of \(\sigma / T < 0.57\), at only 33\% of the complexity required by the latter. Therefore, it can be deduced that the multistage structure has improved performance over the single-stage structure when the RMS delay spread is below a certain threshold value, irrespective of the complexity. By increasing the number of equalizer taps and/or the number of receive antennas (i.e., increasing the complexity) of the multistage structure, the range of RMS delay spread is increased, over which the multistage structure outperforms a given single-stage structure. This is because with a given delay spread, using more equalizer taps and/or receive antennas reduces the error propagation in a multistage structure. In addition, by comparing the performance of DFE \((\mu = 0, m_t = 4, m_o = 1)\) with \((F = 2, B - \mu = 4)\) and \((F = 4, B - \mu = 4)\) in fig.8, it can be deduced that with a shorter-length FFF, the range of RMS delay spread is reduced, over which the wideband structure performs better than or similarly to its narrow-band counterpart in flat fading. Fig.9 has the same configuration as fig.7 except with an increased number of receive antennas \(L=6\) and a fixed SNR=14dB, which demonstrates significant performance improvement over fig.7. In this case, DFE \((\mu = 0, m_t = 4, m_o = 1)\) and DDFSE \((\mu = 1, m_t = 2, m_o = 1)\) outperform the other structures and provide similar performance to each other. At a larger delay spread of \(\sigma / T > 0.7\), however, DFE \((\mu = 0, m_t = 4, m_o = 1)\) is superior to DDFSE \((\mu = 1, m_t = 2, m_o = 1)\), with a higher degree of freedom in the FBF. This might be interpreted by noting [5-6] that in flat fading channels, the diversity order of maximum likelihood detection (MLD) depends only on the number of receive antennas \(L\), while the diversity order of other detection techniques like BLAST depend on the difference between the number of receive antennas and the number of transmit antennas, i.e., \((K-L)\) [5-6]. It is shown in [6] that when the number of receive antennas is similar to or less than the number of transmit antennas, MLD has a significant advantage over the narrow-band BLAST. With an increased number of receive antennas, the performance of BLAST approaches that of MLD, since much of the diversity gain has been obtained by receive diversity. There exists some similarity for the wideband structures in frequency selective channels. When \(K = L = 4\), the benefits of DDFSE with the VA sequence estimator are more significant, which is regarded as the reduced-state maximum likelihood sequence estimator. By increasing the number of receive antennas, the diversity gain achieved by VA is less significant, and therefore a DFE structure with a higher FBF degree of freedom may outperform a DDFSE structure with the VA sequence estimator but a lower FBF degree of freedom, especially when the delay
spread is large. For the same configuration like \( (m_i=1,m_o=1) \), however, DDFSE always outperforms its DFE counterpart in frequency selective channels.

Fig.10 shows the BER versus SNR performance of the \( T/2 \)-spaced \((N_s=2)\) LSTE structures with the same configuration as fig.3, except that the FFF spans only one-half of the time interval spanned by the FFF in fig.3, i.e., \( F=2 \) rather than \( F=4 \). In addition, the sampling phase for the \( T/2 \)-spaced LSTE structure is fixed at zero since a fractionally-spaced equalizer is less sensitive to sampling. With the same number of filter taps and hence similar complexity, the \( T/2 \)-spaced structures significantly outperform their \( T \)-spaced counterparts. Fig.10 shows similar trends to fig.3 for the comparison among different LSTE structures. DDFSE \((\mu=1,m_i=2,m_o=1)\) still has the lowest BER. DFE \((\mu=0,m_i=4,m_o=1)\) provides similar performance to DDFSE \((\mu=1,m_i=2,m_o=1)\) with much lower complexity, similar to the case in fig.9. In addition, comparing the performance of DFE \((\mu=0,m_i=1,m_o=1)\) in fig.10 with that of DDFSE \((\mu=2,m_i=1,m_o=1)\) in fig.3, it can be deduced that DFE outperforms its DDFSE counterpart by oversampling. The performance loss of DDFSE could be attributed to channel memory truncation. Similar conclusions can be drawn by increasing the number of receive antennas, which suggests a duality between oversampling in time and oversampling in space (by increasing the number of receive antennas).

VI. CONCLUSION

In this paper we have described and compared a wide range of LSTE structures over MIMO frequency selective channels, in terms of both performance and complexity. Extensive simulation results show that the proposed multistage LSTE structure has improved performance over the single-stage structure when the RMS delay spread is below a certain threshold value, irrespective of the complexity, which allows tradeoffs between performance and complexity.

We have also investigated the impact of imperfect channel estimation and imperfect interference cancellation on performance. It is shown that the multistage structure is more susceptible to channel estimation error, especially with a short training sequence length. A moderate-length training sequence provides reasonable channel estimation quality. In addition, the later stages of the multistage LSTE structure have worse performance than its first stage at a lower SNR, and perform similarly at a higher SNR, due to imperfect interference cancellation.

The space-time equivalence in improving the performance is also shown. With fewer space-time samples, the benefits of a DDFSE structure with the VA sequence estimator are more
significant. Otherwise, a DFE structure with a higher FBF degree of freedom may have more advantages in terms of both performance and complexity, especially with a larger delay spread.

ACKNOWLEDGEMENT

The authors would like to thank the associate editor and anonymous reviewers for their valuable comments and suggestions.

REFERENCES


---

Stage 1

MIMO Equalizer

\[ K^{(1)} = K \]

\[ m_{0}^{(1)} = K \]

Interference Cancellation

\[ m_{0}^{(2)} = 2 \]

---

Stage \((K/2-1)\)

MIMO Equalizer

\[ K^{(K/2-1)} = 4 \]

\[ m_{0}^{(K/2-1)} = 4 \]

Interference Cancellation

\[ m_{0}^{(K/2)} = 2 \]

---

Stage \(K\)

MIMO Equalizer

\[ K^{(K)} = 2 \]

\[ m_{0}^{(K)} = 2 \]

---

(a)

Stage 1

MISO Equalizer

\[ K^{(1)} = K \]

\[ m_{0}^{(1)} = 1 \]

Interference Cancellation

\[ m_{0}^{(2)} = 2 \]

---

Stage \((K-1)\)

MISO Equalizer

\[ K^{(K-1)} = 2 \]

\[ m_{0}^{(K-1)} = 1 \]

Interference Cancellation

\[ m_{0}^{(K)} = 1 \]

---

Stage \(K\)

MISO Equalizer

\[ K^{(K)} = 1 \]

\[ m_{0}^{(K)} = 1 \]

---

(b)

Fig. 1. Block diagrams of LSTE with (a) \( m_{0}^{(j)} = K^{(j)} = K - 2(j - 1) \) and \( m_{0}^{(j)} = 2 \) for the \( j \)th \((j=1, \ldots, K/2)\) stage (b) \( K^{(j)} = K - j + 1 \) and \( m_{0}^{(j)} = m_{0}^{(j)} = 1 \) for the \( j \)th \((j=1, \ldots, K)\) stage (\( K^{(j)}, m_{0}^{(j)} \) and \( m_{0}^{(j)} \) denote the numbers of undetected data streams, tentatively detected data streams and output data streams at the \( j \)th stage)

---

Fig. 2. Block diagram of the \( j \)th stage of LSTE

---

VDDSE

---

\[ x[i] \]

\[ y[i] \]

\[ W \]

\[ V \]

Weights Calculation

\[ m' \]

\[ x[i - \Delta] \]

\[ \xi \]

Memory \( \mu^{(j)} \)

\[ m_0' \]

\[ m_0'' \]

Length \( LN_F \)

\[ LN_i \]

\[ \hat{s}[i - \Delta] \]

\[ m'' \]

Outputs

---

\[ \hat{y}[i] \]

---

\[ FFB \]

---

\[ VA \]

---

\[ FDFE \]
Fig. 3. T-spaced LSTE with 4 transmit, 4 receive antennas, RMS delay spread $\sigma = 1.25T$ and perfect CSI (both the FFF and FBF span 4 symbol periods)

Fig. 4. Impact of training sequence length $L_t$ on channel estimation error variance
Fig. 5 Impact of imperfect channel estimation and imperfect interference cancellation on the performance of single-stage DFE (µ = 0, m_I = 4, m_O = 4) and multistage DFE (µ = 0, m_I = 4, m_O = 1)

Fig. 6 Impact of imperfect channel estimation and imperfect interference cancellation on the first-stage and last-stage performance of multistage DFE (µ = 0, m_I = 4, m_O = 1)
Fig. 7 $T$-spaced LSTE with 4 transmit, 4 receive antennas, SNR=20dB and perfect CSI

Fig. 8. Illustration of the advantages of multistage DFE ($\mu = 0, m_I = 4, m_O = 1$) over single-stage DFE ($\mu = 0, m_I = 4, m_O = 4$) in terms of performance and complexity (relative complexity is noted on the right of the graph)
Fig. 9 $T$-spaced LSTE with 4 transmit, 6 receive antennas, SNR=14dB and perfect CSI

Fig. 10 $T/2$-spaced LSTE with 4 transmit, 4 receive antennas, RMS delay spread $\sigma=1.25T$ and perfect CSI (the FFF and FBF span 2 and 4 symbol periods, respectively)
Table I: Complexity in terms of complex multiplications ($B=0$)

<table>
<thead>
<tr>
<th>Classification</th>
<th>Tasks</th>
<th>Multiplications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization (per packet)</td>
<td>$\mathbf{R} = \sum_{i=1}^{K} \hat{\mathbf{H}}_k \hat{\mathbf{H}}_i$</td>
<td>$\begin{cases} \begin{align*} L^2 N_0^2 (N+1)(N+2)K/2 &amp; \quad (N \leq F-I) \ L^2 N_0^2 F(2N-F+3)K/2 &amp; \quad (N &gt; F-I) \end{align*} \end{cases}$</td>
</tr>
<tr>
<td>Data Selection $\mathbf{R}^{-1}$</td>
<td>$\frac{1}{6} L^2 N_0^2 F^3 \sum_j \begin{pmatrix} K^{(j)} \ m_j^{(j)} \end{pmatrix}$</td>
<td></td>
</tr>
<tr>
<td>$\left(\hat{\mathbf{H}}<em>k \right)</em>{s+n_0} R^{-1} \left(\hat{\mathbf{H}}<em>k \right)</em>{s+n_0}$ (including FFF matrix $\mathbf{W}$)</td>
<td>$LN_0^2 F (LN_0^2 F+1) \sum_j \begin{pmatrix} K^{(j)} \ m_j^{(j)} \end{pmatrix} m_j^{(j)}$</td>
<td></td>
</tr>
<tr>
<td>ECIR matrix ($\mathbf{V}$)</td>
<td>$\sum_j m_j^{(j)} M^{m_j^{(j)} m_j^{(j)}+1}$</td>
<td></td>
</tr>
<tr>
<td>Signal Processing (per symbol period)</td>
<td>FFF</td>
<td>$LN_0^2 F \sum_j m_j^{(j)}$</td>
</tr>
<tr>
<td></td>
<td>VDDFSE</td>
<td>$\sum_j m_j^{(j)} M^{m_j^{(j)} m_j^{(j)}+1}$</td>
</tr>
</tbody>
</table>

Table II: Complexity with $K=L=4$, $N_s=1$, $F=4$, $B=4$, $N_c=8$, $N_a=0$, 80 symbols per packet for $T$-spaced LTSE (number of complex multiplications per packet)

<table>
<thead>
<tr>
<th>LSTE Structures</th>
<th>Initialiation</th>
<th>Signal Processing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFE ($\mu=0, n_s=4, m_0=4$)</td>
<td>10.0K (100%)</td>
<td>6.4K (100%)</td>
<td>16.4K (100%)</td>
</tr>
<tr>
<td>DFE ($\mu=0, n_s=1, m_0=4$)</td>
<td>40.4K (403%)</td>
<td>6.4K (100%)</td>
<td>46.8K (285%)</td>
</tr>
<tr>
<td>DFE ($\mu=0, n_s=4, m_0=2$)</td>
<td>15.2K (151%)</td>
<td>9.6K (150%)</td>
<td>24.8K (151%)</td>
</tr>
<tr>
<td>DFE ($\mu=0, n_s=4, m_0=1$)</td>
<td>25.6K (256%)</td>
<td>16.0K (250%)</td>
<td>41.6K (254%)</td>
</tr>
<tr>
<td>DDFSE ($\mu=2, n_s=1, m_0=1$)</td>
<td>40.5K (404%)</td>
<td>25.6K (400%)</td>
<td>66.1K (402%)</td>
</tr>
<tr>
<td>DDFSE ($\mu=1, n_s=2, m_0=1$)</td>
<td>55.9K (557%)</td>
<td>44.8K (700%)</td>
<td>100.7K (613%)</td>
</tr>
</tbody>
</table>