On Performance of Transform Domain Adaptive Filters with Markov-2 Inputs

Zhao Shengkui, Man Zhihong, Khoo Suiyang
School of Computer Engineering, Nanyang Technological University
Nanyang Avenue, Singapore 639798. Email: zhao0024@ntu.edu.sg

Abstract—In this paper, the analysis for the performance of the discrete Fourier transform LMS adaptive filter (DFT-LMS) and the discrete cosine transform LMS adaptive filter (DCT-LMS) for the Markov-2 inputs is presented. To improve the convergence property of the least mean squares (LMS) adaptive filter, the DFT-LMS and DCT-LMS preprocess the inputs with the fixed orthogonal transforms and power normalization. We derive the asymptotic results for the eigenvalues and eigenvalue distributions of the preprocessed input autocorrelation matrices with DFT-LMS and DCT-LMS for Markov-2 inputs. These results explicitly show the superior decorrelation property of DCT-LMS over that of DFT-LMS, and also provide the upper bounds for the eigenvalue spreads of the finite-length DFT-LMS and DCT-LMS adaptive filters. Simulation results are demonstrated to support the analytic results.

I. INTRODUCTION

The least mean squares (LMS) algorithm is widely used in adaptive filtering for its simplicity and robustness [1]. Since it uses the instantaneous gradient of the mean square error (MSE) performance function to adapt the filter weights, the LMS algorithm always achieves the low computational complexity of $O(N)$ for a filter with length of $N$. However, the LMS algorithm also has the inherent slow convergence properties when the inputs are with large eigenvalue spreads of the autocorrelation matrices. To improve the convergence properties of the LMS algorithm, the colored inputs are expected to be preprocessed such that the eigenvalue spread of the input autocorrelation matrix is reduced to one. That is, the input autocorrelation matrix is proportional to the identity matrix. In that case, the inputs are perfectly uncorrelated and the LMS algorithm achieves its best convergence properties.

To diagonalize the autocorrelation matrix of the input signals, the direct way is to estimate its inverse autocorrelation matrix and apply this inverse autocorrelation matrix to the input signals, which results in the recursive least squares (RLS) algorithm. Though RLS achieves fast convergence rate and low steady-state error, it has main issues of probably suffering from stability problems, even for the use of the lattice structure [2], and poor tracking capabilities in nonstationary environments [3].

The transform-domain adaptive filtering as another solution was first introduced by Narayn [4], which diagonalizes the input signals by using fixed orthogonal transforms. It was pointed out that properly applying the fixed orthogonal transformations and power normalizations to the LMS filter’s inputs reduces the eigenvalue spreads of the autocorrelation matrix within the bound given by the ratio of the maximum and minimum values of the power spectrum of the preprocessed input signals. Thus the convergence property of the LMS adaptive filter is improved. In addition, the fixed orthogonal transforms such as the discrete Fourier transform (DFT) and the discrete cosine transform (DCT) maintain the low computation complexity of $O(N)$. Though other transforms have also been introduced, the discrete Fourier transform LMS filter (DFT-LMS) and the discrete cosine transform LMS filter (DCT-LMS) are more popular. In this paper, we will focus on DFT-LMS and DCT-LMS.

As pointed out in [5], all the transform-domain LMS adaptive filters obtain the identical optimum Wiener solution and minimum MSE. However, different convergence properties are shown by these filters for different characteristics of the inputs. Beaufays [6] first analyzed the DFT-LMS and DCT-LMS for the case of Markov-1 inputs with correlation parameter $\rho \in [0, 1]$ and showed that the eigenvalue spread of the preprocessed input autocorrelation matrix tends to $(1 + \rho)/(1 - \rho)$ with DFT-LMS and to $1 + \rho$ with DCT-LMS. Before any transforms, the eigenvalue spread with LMS algorithm is asymptotically equal to $(1 + \rho)^2/(1 - \rho)^2$. It was concluded that the DCT-LMS always achieves a smaller eigenvalue spread than DFT-LMS, thereby giving more improvement on the convergence speed for Markov-1 inputs.

For higher-order Markov input signals, the performance of the transform-domain adaptive filters are only experimentally examined. In this paper, we mathematically analyze the performance of DFT-LMS and DCT-LMS with the second-order Markov input signals with two parameters of $\rho_1, \rho_2 \in (-1, 1)$. For this class of input signals, we analyze that the eigenvalues and eigenvalue spreads of the input autocorrelation matrices after DFT/DCT transformations and power normalization eventually converge to some specific values. We derive these values in the expression forms of the correlation parameters $\rho_1, \rho_2$. It will be shown that these expression forms serve as the upper bounds for the input eigenvalue spreads. Therefore, the performance of DFT-LMS and DCT-LMS are clearly indicated by computing the corresponding eigenvalue spreads given a Markov-2 input signal. We find that the DCT-LMS always yields a smaller eigenvalue spread than DFT-LMS for any given Markov-2 input signal. We conclude that DCT-LMS always
outperforms DFT-LMS for Markov-2 inputs. As a result, this is true for Markov-1 inputs since Markov-1 signals are special cases Markov-2 signals. Our analytic results are coincident with the experimental results.

The paper is organized as follows. We first formulate the transform-domain adaptive filters of DFT-LMS and DCT-LMS in Section II. In Section III, the asymptotical eigenvalues and eigenvalue spreads of DFT-LMS and DCT-LMS for Markov-2 inputs are given. Their performance are then discussed and compared. Computer simulations in system modeling are provided to corroborate our analytic results in Section IV. Finally, Section V concludes this paper.

II. TRANSFORM-DOMAIN ADAPTIVE FILTERS

The transform-domain adaptive filters of DFT-LMS and DCT-LMS have more stages than LMS adaptive filter. Consider the tap-delayed inputs \( x_i, x_{i-1}, \ldots, x_{i-N+1} \). LMS uses the tap-delayed inputs directly to update filter coefficients, while DFT-LMS and DCT-LMS first transform the tap-delayed inputs into \( u_{i}(i) \) with an \( N \times N \) discrete Fourier or cosine transform matrix \( T_{k} \), and then normalize the transformed signals with the square root of their estimated power. The resulting less correlated signal \( v_{i}(i) \) is then applied to the LMS filter. The analytic adaptation rules for these two filters can be summarized as follows [6].

DFT/DCT Transform Matrices:

\[
T_{k}(i,l) = \begin{cases} 
\frac{1}{\sqrt{N}} & (DFT) \\
\frac{2}{\sqrt{N}} \kappa \cos \left( \frac{i(l+1/2)\pi}{N} \right) & (DCT) 
\end{cases}
\]

with \( i, l = 0, 1, \ldots, N-1 \) and \( \kappa = 1/\sqrt{2} \) for \( i = 0 \) and 1 otherwise.

Transformation and Power Normalization:

\[
u_{i}(i) = u_{i}(i) \sigma_{i}(i) + \gamma
\]

for \( i = 0, 1, \ldots, N-1 \). The small constant \( \gamma \) is used to eliminate the overflow when the input power is small and the constant \( \beta \) is the convergent factor.

LMS Filtering:

\[
e_{i} = d_{i} - \sum_{i=0}^{N-1} w_{i}(i) v_{i}(i)
\]

\[
w_{i+1}(i) = w_{i}(i) + \mu e_{i} v_{i}(i)
\]

where \( e_{i}, d_{i} \) are filter output error signal and desired output signal respectively, \( v_{i}(i) \) represents the \( i^{th} \) filter coefficient.

From above listed algorithms, we can see that more complicated forms are introduced due to the transformation and power normalization compared with the LMS adaptive filter. These two added stages ensure that the input signals are orthogonally transformed and the power of the input signals is decomposed and redistributed in the frequency domain. For each frequency bin, the input is normalized by the estimated input power such that all the frequency bins carry equal power. Consequently, the correlation nature of the input signals is destroyed and the LMS filtering step that performs adaptation on the decorrelated input signals improves the convergence speed. However, the DFT and DCT are not perfect decorrelators and they produce power leakage from each frequency bin to the others [6]. Though experiments can be used to test the performance of DFT-LMS and DCT-LMS, mathematic analysis of their performance is also necessary. In the following sections, we formulate the class of Markov-2 input signals, and provide the analysis on the performance improvements of DFT-LMS and DCT-LMS over LMS.

III. PERFORMANCE OF DFT-LMS AND DCT-LMS WITH MARKOV-2 INPUTS

In this section, we introduce the class of Markov-2 input signals and analyze DFT-LMS and DCT-LMS with this type of input signals based on the performance factor of the eigenvalue spread of autocorrelation matrix.

The Markov-2 signals are obtained by passing white noise through two-pole filter of transfer function

\[
H(z) = \frac{1}{(1-\rho_1 z^{-1})(1-\rho_2 z^{-1})}
\]

where \( \rho_1 \) and \( \rho_2 \) are the two poles of the filter. Both poles are required to lie inside the unit circle on the z-plane for stability. The \( N \times N \) autocorrelation matrix of a Markov-2 signal \( x_i \) is then given by

\[
R_{N} = c_{1} R_{K} - c_{2} R_{K}^{2}
\]

with

\[
R_{K} = \begin{pmatrix} 1 & \rho_1 & \rho_1^2 & \cdots & \rho_1^{N-1} \\
\rho_1 & 1 & \rho_1 & \cdots & \rho_1^{N-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho_1^{N-1} & \rho_1^{N-2} & \cdots & 1 & \rho_1 \end{pmatrix}
\]

and

\[
R_{K}^{2} = \begin{pmatrix} 1 & \rho_2 & \rho_2^2 & \cdots & \rho_2^{N-1} \\
\rho_2 & 1 & \rho_2 & \cdots & \rho_2^{N-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho_2^{N-1} & \rho_2^{N-2} & \cdots & 1 & \rho_2 \end{pmatrix}
\]

where \( c_1 \) and \( c_2 \) are constants given by

\[
c_1 = \frac{\rho_1 (1-\rho_2^2)}{(1-\rho_1^2)(1-\rho_2^2)} \quad c_2 = \frac{\rho_1 (1-\rho_1^2)}{(1-\rho_1^2)(1-\rho_2^2)}
\]

Since matrices \( R_{K} \) and \( R_{K}^{2} \) are Toeplitz, the matrix \( R_{N} \) is also Toeplitz. Applying the asymptotic theory of Toeplitz matrix.
can be obtained as
\[ P(0) = \sum r(l)e^{-j\omega l} = \frac{c_1(1 - \rho_1^2)}{1 - 2\rho_1 \cos \omega + \rho_1^2} - \frac{c_2(1 - \rho_2^2)}{1 - 2\rho_2 \cos \omega + \rho_2^2} \]  
(9)

where \( r(l) \) denotes the autocorrelation function given by
\[ r(l) = E[x_i x_{i-l}] = c_i \rho_1^l - c_i \rho_2^l, \]
(10)
with \( r(l) = r(-l) \) for real valued signal \( x_i \). Considering a general case \( \rho_1 > \rho_2 \geq 0 \), the maximum and minimum of power spectrums are obtained when \( \cos \omega \) taking values of 1 and -1 respectively. The asymptotic eigenvalue spread of \( R_N \) is thus given by
\[ \lim_{N \to \infty} \left( \frac{\lambda_{\text{min}}}{\lambda_{\text{max}}} \right) = \frac{P_{\text{max}}(\omega)}{P_{\text{min}}(\omega)} = \frac{\rho_1^2 / (1 - \rho_1^2)^2 - \rho_2^2 / (1 - \rho_2^2)^2}{\rho_1^2 / (1 + \rho_1^2)^2 - \rho_2^2 / (1 + \rho_2^2)^2}. \]
(11)

It can be easily checked that the eigenvalue spread given by (11) can be extremely large when both \( \rho_1 \) and \( \rho_2 \) are close to 1. Therefore, when LMS algorithm is used to adapt the filter coefficients, the convergence rate can be quite slow. Thus, the decorrelation of the input signal is necessary.

Notice that the autocorrelation matrix after DFT/DCT transformation and power normalization is not Toeplitz any more, and the power spectrum method cannot be applied here.

The analysis presented in [6] for Markov-1 signals, the matrix (22) is more complicated with two more non-diagonal rows. However, it is still well structured with symmetric property

\[ \frac{\text{lim} \ det(S_N - \lambda I) = 0}{\text{det}(S_N - \lambda I) \text{ is the determinant of matrix and } I \text{ stands for the identity matrix.}} \]
(14)

Problem (14) can be simplified by using matrix theory. Substituting \( B_N \) and (13) into (14), we have
\[ \frac{\text{lim} \ det \left( \left( \text{diag}B_N \right)^{\frac{1}{2}} F_N R_N F_N^H \left( \text{diag}B_N \right)^{\frac{1}{2}} - \lambda I \right) = 0}{\text{det}(S_N - \lambda I) = 0} \]
(15)

It is actually hard to solve equation (15) for the eigenvalues directly. We can simplify equation (15) by multiplying \( F_N^H \left( \text{diag}B_N \right)^{\frac{1}{2}} \) to the left side and \( \left( \text{diag}B_N \right)^{\frac{1}{2}} F_N \) to the right side. Noticing the unitary matrix of \( F_N \), it leads problem (15) to the following generalized eigenvalue problem [17]:
\[ \text{lim} \ det \left( R_N - \lambda D_N \right) = 0, \]
(16)

where matrix \( D_N \) is defined as
\[ D_N = F_N^H \text{diag}B_N F_N. \]
(17)

It can easily be checked that the matrix \( D_N \) is Toeplitz and circulant. Using the DFT transform matrix in (1) and the definition of \( B_N \), the first row of the matrix \( D_N \) is derived as
\[ D_N(0,l) = R_N(0,l) - \frac{1}{N} R_N(0,l) + \frac{1}{N} R_N(0,N-l) \]
\[ = c_1 \left( \rho_1^l - \frac{1}{N} \rho_1^l N \rho_1^{N-l} \right) - c_2 \left( \rho_2^l + \frac{1}{N} \rho_2^l N \rho_2^{N-l} \right), \]
(18)

for \( l = 0, 1, \cdots, N-1 \). For simplicity, we replace \( D_N \) with another asymptotically equivalent Toeplitz matrix \( \tilde{D}_N \):
\[ \tilde{D}_N(0,l) = R_N(0,l) + \frac{1}{N} R_N(0,N-l) \]
\[ = c_1 \left( \rho_1^l + \frac{1}{N} \rho_1^l N \rho_1^{N-l} \right) - c_2 \left( \rho_2^l + \frac{1}{N} \rho_2^l N \rho_2^{N-l} \right). \]
(19)

The asymptotical equivalence of \( D_N \) and \( \tilde{D}_N \) can be verified from the fact that \( \text{det}(N \rho_1^l) \) and \( \text{det}(N \rho_2^l) \) converge to zero when \( N \) goes to infinity independent of \( l \). Now solving (16) is equivalent to solving the following equation
\[ \text{lim} \ det \left( R_N - \lambda \tilde{D}_N \right) = 0. \]
(20)

Multiplying both sides of (20) by \( \text{det}(\tilde{D}_N^{l}) \), we get
\[ \text{det} \left( \tilde{D}_N^{l} R_N - \lambda I \right) = 0 \] by matrix theory. For algebraic simplicity, we rather solve the following inverse problem:
\[ \text{lim} \ det \left( R_N^{-1} \tilde{D}_N - \lambda^{-1} I \right) = 0. \]
(21)

We now try to find a simple form \( \tilde{X}_N \) asymptotically equivalent to \( X_N \). A candidate matrix is shown in (22). Compared to the matrix \( \tilde{X}_N \) provided in [6] for Markov-1 signals, the matrix (22) is more complicated with two more non-diagonal rows. However, it is still well structured with symmetric property

\[ \text{lim} \ det \left( S_N - \lambda I \right) = 0, \]
(14)

where \( \text{det}(\cdot) \) denotes the determinant of matrix and \( I \) stands for the identity matrix.

Problem (14) can be simplified by using matrix theory. Substituting \( B_N \) and (13) into (14), we have
\[ \text{lim} \ det \left( \left( \text{diag}B_N \right)^{\frac{1}{2}} F_N R_N F_N^H \left( \text{diag}B_N \right)^{\frac{1}{2}} - \lambda I \right) = 0. \]
and \((N-4)\times(N-4)\) central identity submatrix. The elements of the first two rows and the last two rows of the matrix are just inverted in directions.

\[
\hat{X}_N = \begin{pmatrix}
1 & \sum_{i=1}^{N-2} \rho_i^4 & \sum_{i=1}^{N-2} \rho_i^4 & \ldots & \rho_0 + \rho_N \\
0 & 1 & -\rho_0 \sum_{i=1}^{N-2} \rho_i^4 & \ldots & -\rho_0 \rho_1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & 1 & 0 \\
0 & 0 & \ldots & 0 & 1 \\
0 & 0 & \ldots & 0 & 0 \\
\rho_1 + \rho_N & \sum_{i=1}^{N-2} \rho_i^4 & \sum_{i=1}^{N-2} \rho_i^4 & \ldots & 1 + \rho_0
\end{pmatrix}
\]

(22)

To use the matrix \(\hat{X}_N\) for computation of the eigenvalues, the asymptotic equivalence \(\hat{X}_N \approx X_N \approx R_N^T D_N\) must be proved. For mathematical simplicity, we show that the difference matrix \(\phi_N = R_N \hat{X}_N - D_N\) is asymptotically a rank zero matrix based on the facts \(|\rho| < 1\) and \(|\rho_2| < 1\). Following Theorem 1, we have \(\hat{X}_N \approx R_N^T D_N\). The interested readers are referred to [8] for the details of proof.

Notice that the asymptotic eigenvalues of \(S_N\) are just the inverse eigenvalues of \(\hat{X}_N\). The eigenvalues of \(\hat{X}_N\) is to be computed from the equation

\[
\lim_{N \to \infty} \det(\hat{X}_N - \lambda I) = 0.
\]

(23)

From the matrix \(\hat{X}_N\) given in (22), we know that the \((N-4)\) eigenvalues of \(\hat{X}_N\) are given by the eigenvalues of the \((N-4)\times(N-4)\) central submatrix of \(\hat{X}_N\), which are equal to 1. The other four eigenvalues of \(\hat{X}_N\) are computed and given by

\[
\tilde{\lambda}_{\pm} = 1 \pm 0.5(\rho_1 + \rho_2)(1 - \rho_1 \rho_2) \pm r
\]

(24)

\[
\tilde{\lambda}_{\pm} = 1 \pm 0.5(\rho_1 + \rho_2)(1 - \rho_1 \rho_2) \pm r
\]

(24)

with \(r = 0.5\sqrt{(\rho_1 + \rho_2)^2 (1 + \rho_1^2 \rho_2^2) - 2 \rho_1 \rho_2 (\rho_1^2 + \rho_2^2)}\).

Using the theorem of strong asymptotic equivalence, we are ready to give the eigenvalues of \(S_N\). Four of them converges to the inverse of \(\tilde{\lambda}_{\pm}\) and \(\tilde{\lambda}_{\pm}\), and the rest are equal to 1. Considering the case \(\rho_1, \rho_2 \in [0, 1]\), the maximum and minimum eigenvalues are obvious. The eigenvalue spread of the autocorrelation matrix after DFT and power normalization is thus equal

\[
\lim_{N \to \infty} \frac{\tilde{\lambda}_{\max}}{\tilde{\lambda}_{\min}} = \frac{1 + 0.5(\rho_1 + \rho_2)(1 - \rho_1 \rho_2) \pm r}{1 - 0.5(\rho_1 + \rho_2)(1 - \rho_1 \rho_2) \mp r}
\]

(25)

with \(r\) given as above. It is easy to check for the special case of Markov-1 signals, that is one of the correlation factors \(\rho_1\) and \(\rho_2\) is set to zero; the result of (25) becomes the same form as provided in [6].

From (25) we can easily notice that the DFT is not a perfect decorrelator. The eigenvalue spread with DFT-LMS can be high enough to limit its uses for strongly correlated input signals (\(\rho_1\) and \(\rho_2\) are close to 1). In the next section, we will present the performance analysis for DCT-LMS which is superior to DFT-LMS.

B. DCT-LMS with Markov-2 Input Signals

To find the eigenvalue distribution of the autocorrelation matrix \(S_N\) after DCT and power normalization in DCT-LMS, we actually solve the similar problem as provided in Section A for DFT-LMS. But here we have a different matrix \(S_N\), which is given by

\[
S_N = (\text{diag}B_N)^{1/2} B_N (\text{diag}B_N)^{1/2}
\]

(26)

with \(B_N = C_N R_N C_N^H\). For real matrix \(C_N\), the Hermitian transpose performs matrix transpose only. The problem of deriving the eigenvalues of \(S_N\) is to solve the equation

\[
\lim_{N \to \infty} \det(S_N - \lambda I) = 0.
\]

(27)

From the derivation for DFT-LMS, we know that solving (27) is equivalent to solving

\[
\lim_{N \to \infty} \det(X_N^* D_N - \lambda^{-1} I) = 0
\]

(28)

with \(D_N = C_N^H \text{diag}B_N C_N\). Since the DCT transform matrix \(C_N\) is not with the properties of \(F_N\), \(D_N\) is not Toeplitz and circulant anymore. So solving for \(D_N\) becomes a difficult task. Therefore, we will take the steps as follows. First, we find a simple and analytical form \(\tilde{X}_N\) for \(X_N \approx R_N^* D_N\) with the aid of computer simulations. Then we give proofs of the asymptotic equivalence in strong sense

\[
\tilde{X}_N \approx X_N = R_N^* D_N.
\]

(29)

The candidate of \(\tilde{X}_N\) through computation and simulation is found to be:

\[
\tilde{X}_N = \begin{pmatrix}
1 & \rho_1 & \sum_{i=1}^{N-1} \rho_i & \sum_{i=1}^{N-1} \rho_i & \ldots & \sum_{i=1}^{N-1} \rho_i & 0 \\
-\rho_1 & 1 & -\rho_1 & -\rho_1 & \ldots & -\rho_1 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 1 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 1 \\
\sum_{i=1}^{N-1} \rho_i & \sum_{i=1}^{N-1} \rho_i & \ldots & \rho_0 & \rho_0 & \ldots & 1 + \rho_0
\end{pmatrix}
\]

(30)

Before we start to compute the eigenvalues of \(\tilde{X}_N\), it must be shown that indeed \(\tilde{X}_N \approx R_N^* D_N\). This is equivalent to proving that \(R_N \tilde{X}_N \approx D_N\) or \(C_N R_N \tilde{X}_N C_N^H \approx \text{diag}B_N\). The readers are referred to [8] for the details of proof.

Now we solve for the eigenvalues from the equation

\[
\lim_{N \to \infty} \det(\tilde{X}_N - \lambda I) = 0
\]

(31)

where \(\tilde{X}_N\) is from (30) and \(\tilde{\lambda}\) are equal to the inverse eigenvalues of \(S_N\). The \((N-4)\) eigenvalues of \(\tilde{X}_N\) are given
by the eigenvalues of the \((N-4)\times(N-4)\) central identity submatrix of \(X_N\), which are equal to 1. The rest four of them are computed from (31) and given as
\[
\lambda_{1,2} = 1 + 0.5\left[\rho_1 + \rho_2 \right] \pm \sqrt{\frac{1}{4} \left[1 + (\rho_1 + \rho_2) (1 - \rho_1 \rho_2) - \rho_1 \rho_2 s_{N-2} \right]^2 - \left[\rho_1 \rho_2 (1 + \rho_1 \rho_2) s_{N-2} - s_{N-4} \right]} + r_i
\]

(32)
with \(r_i = \sqrt{\frac{1}{4} \left(1 + (\rho_1 + \rho_2) (1 - \rho_1 \rho_2) - \rho_1 \rho_2 s_{N-2} \right)^2 - \left[\rho_1 \rho_2 (1 + \rho_1 \rho_2) s_{N-2} - s_{N-4} \right]}\),
\[
\lambda_{3,4} = 1 + 0.5\left[\rho_1 + \rho_2 \right] \pm \sqrt{\frac{1}{4} \left[1 + (\rho_1 + \rho_2) (1 - \rho_1 \rho_2) - \rho_1 \rho_2 s_{N-2} \right]^2 - \left[\rho_1 \rho_2 (1 + \rho_1 \rho_2) s_{N-2} - s_{N-4} \right]} + r_i
\]
(33)
where \(r_i\) is the same as given in the computation for DFT-LMS. Comparing (35) with (24), we can see that the eigenvalues of (35) computed for DCT-LMS are the two eigenvalues of (24) computed for DFT-LMS, with interestingly the same case for Markov-1 input signals [6].

The asymptotic eigenvalue spread of \(S_N\) for the case \(\rho_1, \rho_2 \in [0, 1)\) is then given by:
\[
\lim_{N \to \infty} \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{1 + 0.5\left[\rho_1 + \rho_2 \right] \left[1 - \rho_1 \rho_2 \right]}{1 + 0.5\left[\rho_1 + \rho_2 \right] \left[1 - \rho_1 \rho_2 \right] + r_i}.
\]
(36)
By comparing (36) with (25), it is easy to see that DCT-LMS is more powerful for Markov-2 input signals than DFT-LMS since it has smaller eigenvalue spread of the autocorrelation matrix.

The derivations show that in steady state of solving the original problem with infinite size, we actually manipulate a finite and exact eigenvalue problem. The expressions of eigenvalue spread in (25) and (36) provide upper bounds for DFT-LMS and DCT-LMS with Markov-2 input signals, thereby providing a good understanding of the performance of finite length DFT-LMS and DCT-LMS.

### V. Simulation Results

To have a conceptual comparison for the performance of LMS, DFT-LMS and DCT-LMS adaptive filters, we have listed in Table I that the eigenvalue spread for selected parameters computed based on the analyzed results (11), (25) and (36) under assumption of large \(N\). It is observed that DFT-LMS speeds up the convergence of filter coefficients by moderate factors, while DCT-LMS speeds up the convergence by large factors. The difference of their performance becomes significant for highly correlated input signals. Table II validates that our analyzed results provide the upper bounds for the eigenvalue spread of the autocorrelation matrix for DFT-LMS and DCT-LMS with Markov-2 (including Markov-1 as special case) input signals.

### V. Conclusions

In this paper, we have presented the analytic results of the eigenvalue distribution of autocorrelation matrix for the adaptive filters DFT-LMS and DCT-LMS with the class of Markov-2 input signals. It has been shown that without DFT and DCT, the LMS performs adaptation on very large asymptotic eigenvalue spread of the autocorrelation matrix (11). While DFT-LMS and DCT-LMS perform adaptation on the reduced eigenvalue spreads which are expressed in (25) and (36). Since the convergence behavior of the analyzing algorithms is mainly affected by the eigenvalue spread, these results confirm that DCT-LMS provides better performance than DFT-LMS and LMS for the class of Markov-2 (including Markov-1 as special case) input signals. Further research will be focused on the higher order (\(\geq 3\)) Markov input signals.

### REFERENCES


