

Received February 10, 2020, accepted March 6, 2020, date of publication March 10, 2020, date of current version March 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2979863

A Fast Quasi-Newton Adaptive Algorithm Based on Approximate Inversion of the Autocorrelation Matrix

MOHAMMAD SHUKRI SALMAN¹, (Senior Member, IEEE),

OSMAN KUKRER², (Senior Member, IEEE), AND AYKUT HOCANIN², (Senior Member, IEEE)

¹College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait

²Department of Electrical and Electronic Engineering, Eastern Mediterranean University, 10 Mersin, Turkey

Corresponding author: Mohammad Shukri Salman (mohammad.salman@aum.edu.kw)

ABSTRACT The Newton adaptive filtering algorithm in its original form is computationally very complex as it requires inversion of the input-signal autocorrelation matrix at every time step. Also, it may suffer from stability problems due to the inversion of the input-signal autocorrelation matrix. In this paper, we propose to replace the inverse of the input-signal autocorrelation matrix by an approximate one, assuming that the input-signal autocorrelation matrix is Toeplitz. This assumption would help us in replacing the update of the inverse of the autocorrelation matrix by the update of the autocorrelation matrix itself, and performing the multiplication of $R^{-1}x$ in the update equation by using the Fourier transform. This would increase the stability of the algorithm, in one hand, and decrease its computational complexity, on the other hand. Since the objective of the paper is to enhance the stability of the Newton algorithm, the performance of the proposed algorithm is compared to those of the Newton and the improved quasi-Newton (QN) algorithms in noise cancellation and system identification settings.

INDEX TERMS Impulsive noise, Newton method, noise cancellation, system identification.

I. INTRODUCTION

The least-mean-squares (LMS) algorithm is used in many applications that vary from single-input/single-output (SISO) to multiple-input/multiple-output (MIMO) systems [1]–[9]. However, when the input signal is highly correlated or band-limited, the LMS algorithm as well as other algorithms of the steepest-descent family converge slowly and the capability of such algorithms in tracking nonstationarities deteriorates. In such situations, more sophisticated algorithms such as Newton based algorithms provide improved performance [10]–[13]. Unfortunately, these algorithms, which are usually computationally very complex, suffer from stability problems since they require the inverse of the input-signal autocorrelation matrix in the weight vector update. And sometimes they have low convergence rate. Quasi-Newton algorithms avoid the inversion of the autocorrelation matrix by directly updating the inverse of the autocorrelation matrix, but they still have high computational complexity. In [14],

the authors propose an improved quasi-Newton (QN) algorithm that performs data-selective adaptation whereby the weight vector and the inverse of the input-signal autocorrelation matrix are updated only when the apriori-error exceeds a prespecified error bound. The authors have increased the stability and the convergence rate of the known QN algorithm and significantly reduce its computational complexity. However, the algorithm still requires the update of the inverse autocorrelation matrix when the apriori-error exceeds the prespecified error bound which, in turn, may cause stability problems [14]. Also, its performance may deteriorate depending on the problem setting.

Several methods have been developed for efficient inversion of Toeplitz matrices [15]–[21]. An important class of these methods are based on the Gohberg-Semencul formula [15], [17]. This formula, derived from Trench's method, involves the construction of the inverse from a low number of its columns and the entries of the Toeplitz matrix itself. The inversion formula developed in [21], which is one of the more recent of such formulae, reconstructs the inverse using only the first column of the inverse and entries of

The associate editor coordinating the review of this manuscript and approving it for publication was Jenny Mahoney.

the original matrix. When this inversion technique is closely examined it is noticed that it requires additional computations for constructing the matrices of the formula, which then can be efficiently multiplied using the fast-Fourier transform (FFT). First of all, the first column of the inverse must be computed using the standard procedure for calculating matrix inverses. Then, the equation $Tx = y$, where T is the Toeplitz matrix and y is a vector containing elements of T , must be solved for x . The vector x is then used to construct two of the matrices of the formula. Actually, the solution for x is achieved by using the inverse formula itself, where x is solved by back substitution. In [21] it is stated that the number of real operations for calculating the inverse is $16n \log_2 n$ where n is the matrix dimension. However, this count does not include the above-mentioned additional computations. Therefore, compared with such fast inversion techniques, the proposed approximate inversion method has much lower complexity. It is worth noting that, as far as application to the Newton adaptive filter is concerned, the exact inverse of the autocorrelation matrix is not essential. As will be discussed in Section V, it is sufficient that the eigenvalues of the product of the autocorrelation matrix with its approximate inverse lie within the unit disc. If any eigenvalue is near the boundary, the step size of the Newton algorithm can be adjusted to stabilize the iterations.

Approximate inverses of Toeplitz matrices have been considered as preconditioners in the numerical solution of Toeplitz systems of equations $Tx = y$, usually based on the conjugate gradient (CG) algorithm [22]–[25]. The preconditioning matrix M is chosen such that the preconditioned matrix MT has eigenvalues clustered around unity, enabling fast convergence of the CG algorithm. Results concerning inverses of infinite-dimensional Toeplitz matrices are presented in the literature [26], where the inverse of a Toeplitz matrix generated by a symbol is related to the Toeplitz matrix generated by the inverse of the symbol (plus product of Hankel matrices), where the symbol is positive definite. Approximate inverses of finite-dimensional Toeplitz matrices generated from the inverse of the symbol of the matrix have been proposed as preconditioners in [27]. However, this paper does not attempt to derive results indicating the degree to which the inverse generated in this manner approximates the true inverse, but focuses on showing that the product of the approximate inverse with the matrix has eigenvalues clustered around unity. In particular, no explicit result is given linking the properties of the symbol of the Toeplitz matrix to the distribution of the eigenvalues of the preconditioned matrix MT . This is crucial for the application here, since the eigenvalues must be contained in the unit disc for stability of the Newton algorithm. The main contribution of this work is the derivation of an explicit expression that approximately relates the eigenvalues to the symbol of the Toeplitz matrix (truncated power spectral density (PSD) of the input signal). It is shown that the inverse generated from the inverse of the sampled PSD is a reasonably good approximation for the true inverse in a wide dynamic range of the PSD.

In this paper, we provide a detailed description of our algorithm proposed in [28] which is mainly a new approximate inverse quasi-Newton (AIN) algorithm that replaces the inverse of the input-signal autocorrelation matrix by an approximate one, assuming that the input-signal's instantaneous autocorrelation matrix is Toeplitz. This assumption replaces the update of the inverse autocorrelation matrix by the update of the autocorrelation matrix itself, and allows performing the multiplication of $R^{-1}x$ in the update equation by using the Fourier transform. This increases the stability of the algorithm, on one hand, and decreases its computational complexity, on the other hand. The number of computations needed in the update equation for one iteration becomes much less than that needed in the improved QN algorithm when the filter length is moderately large. Our aim, here, is to enhance the stability of the Newton algorithm, hence the performance of the proposed algorithm is compared to those of the Newton and the improved QN algorithms under different settings. Also, convergence analysis of the algorithm in the mean sense is provided.

The paper is organized as follows: In Section II, the approximate inversion of the Toeplitz matrices is analyzed in detail. In Section III, the Newton adaptive filter is reformulated to use a fixed step-size, instead of the variable step-size in the original Newton, and the inverse of the instantaneous estimate of the autocorrelation matrix $R^{-1}(k)$ instead of $R^{-1}(k-1)$. In Section IV, analysis of the convergence in the mean of Newton is presented. In Section V, the mathematical model of the proposed AIN algorithm is presented. In Section VI, implementation of the approximate inversion technique is described. In Section VII, simulation results are provided and discussed. Finally, the conclusions are drawn.

II. APPROXIMATE INVERSION OF THE TOEPLITZ MATRICES

Given a Toeplitz autocorrelation matrix (R) corresponding to the autocorrelation sequence (ACS) of a stationary process:

$$r(n) = E \{x(k)x^*(k+n)\}, \quad n = -(N-1), \dots, (N-1) \quad (1)$$

where $r(-n) = r^*(n)$. The matrix R can be written as:

$$\begin{bmatrix} r(0) & r(-1) & \dots & r(-(N-1)) \\ r(1) & r(0) & \dots & r(-(N-2)) \\ r(2) & r(1) & \dots & r(-(N-3)) \\ \vdots & \vdots & \ddots & \vdots \\ r(N-1) & r(N-2) & \dots & r(0) \end{bmatrix} \quad (2)$$

The power spectrum of the signal corresponding to the truncated ACS is,

$$S(\omega) = \sum_{n=-(N-1)}^{N-1} r(n)e^{-jn\omega}. \quad (3)$$

The question is whether the inverse of R can be approximated using the inverse power spectrum defined as:

$$Q(\omega) = \frac{1}{S(\omega)} = \sum_{n=-(N-1)}^{N-1} q(n)e^{-jn\omega}. \quad (4)$$

Let $P = \text{Toep}\{q(n)\}$ generated by the sequence $\{q(n); n = -(N-1), \dots, (N-1)\}$.

It can be asserted that P is an approximation for the inverse of R if the eigenvalues of $(I - PR)$ lie on a disc with radius $\epsilon < 1$: $D_\epsilon(\lambda) = \{\lambda; |\lambda| < \epsilon\}$. Note that as $\epsilon \rightarrow 0$, P approaches the exact inverse R^{-1} , and the smaller ϵ is the better the approximation. Now, to establish the conditions under which this approximation is valid, we consider the eigenvalues of PR ,

$$(PR)x = \lambda x, \quad (5)$$

where $x = [x(0) \ x(1) \ \dots \ x(N-1)]^T$ is the eigenvector corresponding to the eigenvalue λ . Let $y = Rx = [y(0) \ y(1) \ \dots \ y(N-1)]^T$, then, the elements of this vector are given by:

$$y(n) = \sum_{m=0}^{N-1} r(n-m)x(m), \quad n = 0, \dots, (N-1) \quad (6)$$

Define an extended sequence as:

$$\tilde{y}(n) = \sum_{m=0}^{N-1} r(n-m)x(m), \quad -\infty < n < \infty \quad (7)$$

where $r(k) = 0$ if $k \notin [-(N-1), (N-1)]$. Equation (7) is in the form of a convolution sum. Taking the Fourier transform (FT) gives,

$$\tilde{Y}(\omega) = S(\omega)X(\omega). \quad (8)$$

Then, the original sequence $\{y(0), \dots, y(N-1)\}$ can be related to the extended sequence as

$$y(k) = \tilde{y}(k)a(k), \quad (9)$$

where $a(k)$ is:

$$a(k) = \begin{cases} 1, & k = 0, \dots, N-1 \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

Taking the Fourier transform (FT) of (9),

$$Y(\omega) = \frac{1}{2\pi} \int_0^{2\pi} \tilde{Y}(\omega_1)A(\omega - \omega_1)d\omega_1, \quad (11)$$

where $A(\omega)$ is the FT of the sequence $a(k)$

$$A(\omega) = \frac{1 - e^{-jN\omega}}{1 - e^{-j\omega}}. \quad (12)$$

The FT of $y(k)$ then becomes,

$$Y(\omega) = \frac{1}{2\pi} \int_0^{2\pi} S(\omega_1)X(\omega_1) \frac{1 - e^{-jN(\omega - \omega_1)}}{1 - e^{-j(\omega - \omega_1)}} d\omega_1. \quad (13)$$

Consider also the product $Py = z = \lambda x$. Since P is also Toeplitz, the FT of z can be written as:

$$Z(\omega) = \frac{1}{2\pi} \int_0^{2\pi} Q(\omega_2)Y(\omega_2)A(\omega - \omega_2)d\omega_2, \quad (14)$$

where $Q(\omega)$ is the FT of the sequence $q(k)$. Substituting (13) in (14) we get

$$\lambda X(\omega) = \frac{1}{(2\pi)^2} \int_0^{2\pi} S(\omega_1)X(\omega_1) \left\{ \int_0^{2\pi} Q(\omega_2) \times A(\omega_2 - \omega_1)A(\omega - \omega_2)d\omega_2 \right\} d\omega_1. \quad (15)$$

Equation (15) is an integral equation relating the Fourier transforms of the generating sequences $r(n)$ and $q(n)$ to the eigenvalues of the product PR . Let us consider the trivial case where the signal $x(k)$ is white:

$$S(\omega_1) = r(0) \Rightarrow Q(\omega_2) = \frac{1}{r(0)} = q(0). \quad (16)$$

Then,

$$\lambda X(\omega) = \frac{1}{(2\pi)^2} \int_0^{2\pi} X(\omega_1) \left\{ \int_0^{2\pi} A(\omega_2 - \omega_1) \times A(\omega - \omega_2)d\omega_2 \right\} d\omega_1, \quad (17)$$

simplifying (17)

$$\begin{aligned} & \int_0^{2\pi} A(\omega_2 - \omega_1)A(\omega - \omega_2)d\omega_2 \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \int_0^{2\pi} e^{-jn(\omega_2 - \omega_1)} e^{-jm(\omega - \omega_2)} d\omega_2 \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{jn\omega_1} e^{-jm\omega} \int_0^{2\pi} e^{-j(n-m)\omega_2} d\omega_2 \\ &= 2\pi \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{jn\omega_1} e^{-jm\omega} \delta(n-m) \\ &= 2\pi \sum_{n=0}^{N-1} e^{-jn(\omega - \omega_1)} \\ &= 2\pi A(\omega - \omega_1), \end{aligned} \quad (18)$$

where $\delta(k)$ is the Kronecker delta. Substituting (18) in (15) gives

$$\begin{aligned} \lambda X(\omega) &= \frac{1}{2\pi} \int_0^{2\pi} X(\omega_1)A(\omega - \omega_1)d\omega_1 \\ &= X(\omega) \Rightarrow \lambda = 1, \end{aligned} \quad (19)$$

which implies that all the eigenvalues of PR are equal to 1, as expected.

Equation (15) can be discretized using the interpolation relation for $X(\omega_1)$ in (20)

$$\lambda X(\omega_1) = \frac{1}{N} \sum_{k=0}^{N-1} X(\alpha_k)A(\omega_1 - \alpha_k), \quad (20)$$

where $\alpha_k = \frac{2\pi k}{N}, k = 0, \dots, N-1$. Substituting (20) in (15) and setting $\omega = \alpha_l = \frac{2\pi l}{N}, l = 0, \dots, N-1$ yields

$$\lambda X(\alpha_l) = \sum_{k=0}^{N-1} c_{lk}X(\alpha_k), \quad (21)$$

where

$$c_{lk} = \frac{1}{(2\pi)^2 N} \int_0^{2\pi} S(\omega_1)A(\omega_1 - \alpha_k) \times \left\{ \int_0^{2\pi} Q(\omega_2)A(\omega_2 - \omega_1)A(\alpha_l - \omega_2)d\omega_2 \right\} d\omega_1. \quad (22)$$

Letting $X = [X(\alpha_0) \dots X(\alpha_{N-1})]^T$ and $C = [c_{lk}]_{l,k=0,\dots,N-1}$, equation (22) can be written as

$$\lambda X = CX$$

Therefore, the original eigenvalue equation has been transformed to one in the transform domain involving the DFT of the eigenvector x . The eigenvalues of PR are the same as those of the matrix C . An advantage of the transform domain equation is that the eigenvalues are now those of a matrix having elements expressed in terms of the power spectral densities $S(\omega)$ and $Q(\omega)$. Hence, the eigenvalues can be related to the dynamic range of $S(\omega)$.

The integrals in the R.H.S. of (22) can be discretized using interpolation equations (see Appendix) for $S(\omega_1)$ and $Q(\omega_2)$ as described in the following proposition.

Proposition 1: Equation (22) can be rewritten in the following form:

$$c_{lk} = \sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} Q(u_i)S(u_j)\eta(u_i, u_j, \alpha_l, \alpha_k), \quad l, k = 0, \dots, N-1 \quad (23)$$

where $N_s = 2N - 1$, and

$$\eta(u_i, u_j, \alpha_l, \alpha_k) = \frac{1}{NN_s^2} e^{j\frac{1}{2}(N-1)(\alpha_k - \alpha_l)} F_N(u_j - u_i)F_N(u_j - \alpha_k)F_N(u_i - \alpha_l), \quad (24)$$

and

$$F_N(\omega) = \frac{\sin\left[\frac{N\omega}{2}\right]}{\sin\left[\frac{\omega}{2}\right]}.$$

Proof: The proof is given in Appendix.

Corollary: When $S(\omega)$ is constant (white signal case), it can be shown using (18) that $c_{lk} = 0$ for $l \neq k$ and $c_{lk} = 1$ for $l = k$. This further implies that

$$\sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} \eta(u_i, u_j, \alpha_l, \alpha_k) = \begin{cases} 1, & l = k \\ 0, & l \neq k. \end{cases} \quad (25)$$

which corresponds to the result $\lambda = 1$ for the white signal case.

Equation (25) means that if the dynamic range of $S(\omega)$ is not large, the eigenvalues of PR can be approximated by the diagonal elements c_{ll} . A formal proof of this statement seems extremely difficult. However, a heuristic argument may be used to support it. Consider the function $F_N(\omega)$ in (24) which has an impulse-like shape for large N . Hence, for instance, the term $F_N(u_j - u_i)$ has significant values when u_j is very close to u_i . Evidently, as a result of the product of the three terms, the kernel $\eta(u_i, u_j, \alpha_l, \alpha_k)$ will have significant magnitudes only when $\alpha_l = \alpha_k$. Therefore, letting $Q(u_i) = \frac{1}{S(u_i)}$ we get

$$\lambda_l \cong \sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} \frac{S(u_j)}{S(u_i)} \eta_0(u_i, u_j, \alpha_l), \quad l = 0, \dots, N-1 \quad (26)$$

where $\eta_0(u_i, u_j, \alpha_l) = \eta(u_i, u_j, \alpha_l, \alpha_l)$.

The eigenvalues of PR can now be related to the dynamic range of the FT $S(\omega)$ under a simplifying assumption by the following Lemma.

Lemma: If $S(\omega)$ is smooth enough with bounded low order derivatives, the eigenvalues in (26) can be written as

$$\lambda_l \cong 1 + \left(\frac{S'(\alpha_l)}{S(\alpha_l)} \right)^2 \gamma_N(\alpha_l) \quad (27)$$

where

$$\gamma_N(\alpha_l) = \sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} (u_i - \alpha_l)(u_i - u_j)\eta_0(u_i, u_j, \alpha_l) \quad l = 0, \dots, N-1 \quad (28)$$

Proof: Consider the truncated Taylor series expansions of $S(\omega)$ and $Q(\omega) = \frac{1}{S(\omega)}$ around α_l evaluated at u_i and u_j ,

$$S(u_j) \cong S(\alpha_l) + S'(\alpha_l)(u_j - \alpha_l) + \frac{1}{2}S''(\alpha_l)(u_j - \alpha_l)^2 \quad (29)$$

$$Q(u_i) \cong Q(\alpha_l) + Q'(\alpha_l)(u_i - \alpha_l) + \frac{1}{2}Q''(\alpha_l)(u_i - \alpha_l)^2 \quad (30)$$

where $\phi'(u)$ and $\phi''(u)$ denote the first and second derivatives of $\phi(u)$ with respect to u , respectively. Then, (26) can be written as

$$\lambda_l \cong Q(\alpha_l)S(\alpha_l) + \left[Q(\alpha_l)S'(\alpha_l) + Q'(\alpha_l)S(\alpha_l) \right] \sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} (u_i - \alpha_l)\eta_0(u_i, u_j, \alpha_l)$$

$$\begin{aligned}
 &+ [Q(\alpha_l)S''(\alpha_l) + Q''(\alpha_l)S(\alpha_l)] \\
 &\sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} (u_i - \alpha_l)^2 \eta_0(u_i, u_j, \alpha_l) \\
 &+ Q'(\alpha_l)S'(\alpha_l) \sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} (u_i - \alpha_l) \\
 &(u_i - u_j) \eta_0(u_i, u_j, \alpha_l) \tag{31}
 \end{aligned}$$

where third order terms are omitted, and symmetry of the summation terms with respect to u_i and u_j has been utilized. Now, making use of the fact that $Q(\omega) = \frac{1}{S(\omega)}$, the following can be shown to hold

$$Q(\alpha_l)S'(\alpha_l) + Q'(\alpha_l)S(\alpha_l) = 0 \tag{32}$$

$$\begin{aligned}
 \frac{1}{2} [Q(\alpha_l)S''(\alpha_l) + Q''(\alpha_l)S(\alpha_l)] &= -Q'(\alpha_l)S'(\alpha_l) \\
 &= \left(\frac{S'(\alpha_l)}{S(\alpha_l)} \right)^2 \tag{33}
 \end{aligned}$$

using (32) and (33) in (31) and after simplification, (31) reduces to (27).

In (27) the factors $\gamma_N(\alpha_l)$, which depend only on N , take small values ($O(1/N)$) when α_l is away from 0 or 2π , but increase abruptly to large values as $\alpha_l \rightarrow 0, 2\pi$. In the latter case, since $S'(0) = S'(2\pi) = 0$, $S'(\alpha_l)$ may be expected to be very small in the vicinity of 0 or 2π if $S(\omega)$ is sufficiently smooth, thus making the second term in (27) small enough. Therefore, if the maximum of $\frac{S'(\alpha_l)}{S(\alpha_l)}$ occurs at a frequency sufficiently away from 0 or 2π , then the eigenvalues are likely to be clustered around unity if

$$\max_{\alpha_l \in [0, \pi]} \left| \frac{S'(\alpha_l)}{S(\alpha_l)} \right| \ll O(\sqrt{N}) \tag{34}$$

The following proposition gives an upper bound for the eigenvalues in terms of the dynamic range of the PSD $S(\omega)$ for the special case where $S(\omega)$ has only extremum in $(0, 2\pi)$.

Proposition 2: For the case where $S(\omega)$ has only one extremum in $[0, 2\pi]$ at the frequency u_k such that $S(u_k) = S_{max}$, $S(0) = S_{min}$ (see Fig. 1), and that $\left| \frac{S'(\alpha_l)}{S(\alpha_l)} \right|$ attains its maximum at a frequency sufficiently away from 0 or 2π , the eigenvalues approximately satisfy

$$|\lambda_l - 1| \leq \left(\frac{2 \ln \rho}{u_k} \right)^2 \gamma_N(\alpha_m) \tag{35}$$

where $\rho = \frac{S_{max}}{S_{min}}$ and $\alpha_m, m \in [0, N - 1]$ is the frequency nearest u_m . Note that u_k in (35) is by definition the difference in the frequencies at which S_{max} and S_{min} occur, hence cannot be equal to zero.

Proof: Consider the first two terms of the Taylor series expansion of $S(u_j)$ about $S(u_i)$ in (23)

$$|\lambda_l - 1| \leq \gamma_N(\alpha_l) \max_{\alpha_l \in [0, \pi]} \left(\frac{S'(\alpha_l)}{S(\alpha_l)} \right)^2 \tag{36}$$

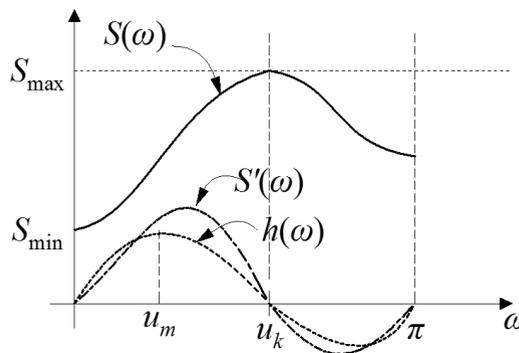


FIGURE 1. The FT $S(\omega)$ with one extremum in $(0, \pi)$.

in (37) the fact that $\gamma_N(\alpha_l)$ is almost constant when α_l is away from 0 or 2π , is used. Letting $h(\omega) = \frac{S'(\omega)}{S(\omega)}$, then

$$\int_0^{u_k} h(\omega) d\omega = \int_{S_{min}}^{S_{max}} \frac{dS(\omega)}{S(\omega)} = \ln \left[\frac{S_{max}}{S_{min}} \right] = \ln \rho \tag{37}$$

If $S(\omega)$ is smooth enough, it may be assumed that the maximum (h_{max}) of $h(\omega)$ occurs in the range $(0, u_k)$. Letting $h(u_m) = h_{max}$ and using a piecewise linear approximation for $h(\omega)$ in $[0, u_k]$ (i.e. from 0 to u_m and from u_m to u_k), the integral in (37) may be approximated as

$$\int_0^{u_k} h(\omega) d\omega \approx \frac{1}{2} u_k h_{max} \tag{38}$$

Solving for h_{max} from (37) and (38) and substituting in (33) for the maximum, (32) is obtained.

To test the accuracies of (27) and (35), we consider a banded matrix, where $r(0) = 1$ and $r(\pm n) = \beta^n$ for $n = 1, 2$. This matrix has been chosen, because the PSD $S(\omega)$ has one extremum in $(0, \pi)$, which was the assumption made in deriving (35), and is sufficiently smooth. The PSD is given by $S(\omega) = 1 + 2\beta \cos(\omega) + 2\beta^2 \cos(2\omega)$ for $\beta \leq \sqrt{\frac{3}{8}}$ and $u_k \simeq \cos^{-1}(-0.25\beta^{-1})$. Fig. 2 shows the computed and theoretical maximum eigenvalues of $(I - PR)$ versus $\rho = \frac{S_{max}}{S_{min}}$. The computed eigenvalues are calculated by directly finding the approximate inverse P , and the theoretical ones are calculated using (27). The upper bounds calculated using (35) are also displayed on Fig. 2. These computations indicate that the analytical results are sufficiently accurate under the conditions assumed in deriving them.

III. NEWTON ADAPTIVE ALGORITHM

In this section, an overview of the Newton adaptive filter is provided. Consider the Wiener-Hopf equation at time-step k ,

$$R(k)w(k) = p(k) \tag{39}$$

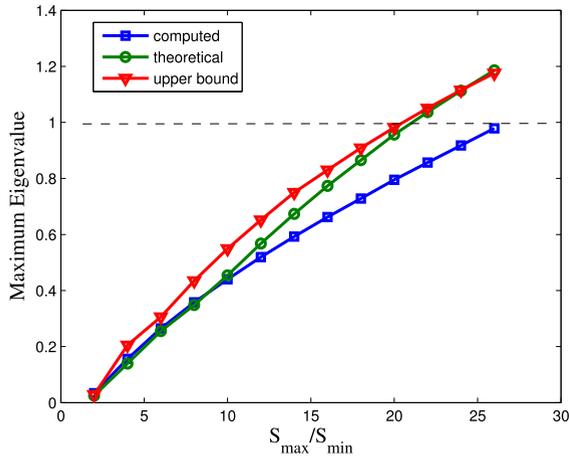


FIGURE 2. Theoretical and computed maximum eigenvalues of $(I - PR)$ versus $\rho = \frac{S_{max}}{S_{min}}$.

It should be highlighted that it is possible to augment the Wiener-Hopf equation, in (39), as follows

$$\sum_{m=1}^N r_{xx}(n-m)w(m) = p(n), \quad n=1, \dots, M, \quad M > N \quad (40)$$

In this case the autocorrelation matrix would be $M \times N$, and would have full column rank. The equations would have a least-squares solution and the inverse of the autocorrelation matrix would be replaced by the Moore-Penrose pseudo-inverse. However, approximation of the inverse based on the power spectrum would not be possible.

In (39), $R(k)$ is the instantaneous estimate of the autocorrelation matrix, $w(k)$ is the filter weights vector and $p(k)$ is the instantaneous estimate of the cross-correlation between the desired signal and the input vector. The correlations are estimated recursively as

$$R(k) = \beta R(k-1) + x(k)x^T(k) \quad (41)$$

$$p(k) = \beta p(k-1) + d(k)x(k) \quad (42)$$

where $\beta < 1$ is the forgetting factor and is usually chosen very close to unity. Substituting (41) and (42) in (39) and letting $\Delta w(k) = w(k) - w(k-1)$ leads to

$$\begin{aligned} & \left[\beta R(k-1) + x(k)x^T(k) \right] \Delta w(k) \\ &= x(k) \left[d(k) - x^T(k)w(k-1) \right] \end{aligned} \quad (43)$$

which simplifies to

$$R(k)\Delta w(k) = x(k)e(k) \quad (44)$$

where $e(k)$ is the a-priori estimation error. The update equation for the weight vector is obtained as

$$w(k) = w(k-1) + \mu R^{-1}(k)x(k)e(k) \quad (45)$$

where μ is a step-size which may be required to stabilize the recursion in the update equation. Equation (45) is known as the Newton adaptive algorithm [29].

IV. ANALYSIS OF CONVERGENCE IN THE MEAN

The update equation in (45) can be rewritten as

$$\begin{aligned} w(k) = & \left[I - \mu R^{-1}(k)x(k)x^T(k) \right] w(k-1) \\ & + \mu R^{-1}(k)x(k)d(k) \end{aligned} \quad (46)$$

Equations (41) and (42) can be solved as

$$R(k) = \sum_{i=0}^k \beta^{k-i} x(i)x^T(i) \quad (47)$$

$$p(k) = \sum_{i=0}^k \beta^{k-i} x(i)d(i) \quad (48)$$

Let

$$x(k)x^T(k) = R_{xx} + \Delta_k, \quad x(k)d(k) = p_{xd} + \delta_k \quad (49)$$

where R_{xx} and p_{xd} are the theoretical autocorrelation matrix and cross-correlation vector, respectively, so that $E\{\Delta_k\} = 0, E\{\delta_k\} = 0$. Substituting (49) in (48) and (47)

$$\begin{aligned} R(k) = \sum_{i=0}^k \beta^{k-i} (R_{xx} + \Delta_k) &= \frac{1 - \beta^{k+1}}{1 - \beta} R_{xx} + \tilde{\Delta}_k \\ &= \alpha_k R_{xx} + \tilde{\Delta}_k \end{aligned} \quad (50)$$

$$\begin{aligned} p(k) = \sum_{i=0}^k \beta^{k-i} (p_{xd} + \delta_k) &= \frac{1 - \beta^{k+1}}{1 - \beta} p_{xd} + \tilde{\delta}_k \\ &= \alpha_k p_{xd} + \tilde{\delta}_k \end{aligned} \quad (51)$$

where

$$\tilde{\Delta}_k = \sum_{i=0}^k \beta^{k-i} \Delta_k, \quad \tilde{\delta}_k = \sum_{i=0}^k \beta^{k-i} \delta_k$$

Substituting (50) and (51) in (46) yields, after some manipulation,

$$\begin{aligned} w(k) = & \left[I - \mu \alpha_k^{-1} \left(I + \alpha_k^{-1} R_{xx}^{-1} \tilde{\Delta}_k \right)^{-1} \left(I + R_{xx}^{-1} \Delta_k \right) \right] \\ & \times w(k-1) \\ & + \mu \alpha_k^{-1} \left(I + \alpha_k^{-1} R_{xx}^{-1} \tilde{\Delta}_k \right)^{-1} R_{xx}^{-1} x(k)d(k) \end{aligned} \quad (52)$$

The term $\left(I + \alpha_k^{-1} R_{xx}^{-1} \tilde{\Delta}_k \right)^{-1}$ can be approximated as

$$\left(I + \alpha_k^{-1} R_{xx}^{-1} \tilde{\Delta}_k \right)^{-1} \simeq I - \alpha_k^{-1} R_{xx}^{-1} \tilde{\Delta}_k \quad (53)$$

where the first two terms in the expansion $(I + A)^{-1} = I - A + A^2 - \dots$ have been used, which is valid if the spectrum of A is contained within the unit disc ($|\sigma(A)| < 1$). The approximation is good if the eigenvalues of A have magnitudes much less than unity. Note that if R_{xx} is a proper autocorrelation matrix, R_{xx}^{-1} will be a well-behaved matrix. The matrix $\tilde{\Delta}_k$ is a low-pass-filtered version of the fluctuation Δ_k with a very small cut-off frequency if β is close to unity. Therefore, since $E\{\Delta_k\} = 0$ it may be expected that the elements of $\tilde{\Delta}_k$ would be small enough, so that the eigenvalues of $R_{xx}^{-1} \tilde{\Delta}_k$ would be sufficiently small.

Now, substituting (53) in (52) and neglecting second order terms after expanding the matrix product, we get

$$w(k) = \left[I - \mu \alpha_k^{-1} \left(I + R_{xx}^{-1} \left(\Delta_k - \alpha_k^{-1} \tilde{\Delta}_k \right) \right) \right] w(k-1) + \mu \alpha_k^{-1} \left(I - \alpha_k^{-1} R_{xx}^{-1} \tilde{\Delta}_k \right) R_{xx}^{-1} x(k) d(k) \quad (54)$$

Taking expectation and after invoking the independence assumption [10], we get

$$\bar{w}(k) = \left[1 - \mu \alpha_k^{-1} \right] \bar{w}(k-1) + \mu \alpha_k^{-1} R_{xx}^{-1} p_{xd} \quad (55)$$

where $\bar{w}(k)$ is the expected value of the tap-weight vector.

It can be deduced that, for uniform convergence we must have $0 < \mu < \sup_{k \geq 0} (\alpha_k) = 1$. It can easily be verified that as $k \rightarrow \infty$ the mean weight vector converges to the optimum $w_{opt} = R_{xx}^{-1} p_{xd}$.

V. A NEW NEWTON ADAPTIVE ALGORITHM

Newton adaptive algorithm in its original form is computationally very complex as it requires the inversion of the autocorrelation matrix at every time step. We propose to replace the inverse by an approximate one, obtained from the Toeplitz approximation of $R(k)$, as described in section II. It was shown that the approximate inverse satisfies

$$P(k)R(k) = U(k) \quad (56)$$

where $U(k)$ is a matrix having eigenvalues concentrated around unity when the eigenvalue spread of $R(k)$ is less than a certain value (see (35)). The weight update equation of the proposed algorithm would be

$$w(k) = w(k-1) + \mu P(k)x(k)e(k) \quad (57)$$

which, by making use of (56), can be written as

$$w(k) = w(k-1) + \mu U(k)R^{-1}(k)x(k)e(k) \quad (58)$$

Following a similar procedure for the original Newton algorithm we get

$$\bar{w}(k) = \left[I - \mu \alpha^{-1} U(k) \right] \bar{w}(k-1) + \mu \alpha^{-1} U(k) w_{opt} \quad (59)$$

Then, for convergence, it is sufficient that $0 < \mu < \frac{1}{\lambda_{max}(U(k))}$ where $\lambda_{max}(U(k))$ is that eigenvalue of $U(k)$ having the maximum magnitude. Letting $\lambda_{max}(U(k)) = 1 + \epsilon$, where ϵ depends on the eigenvalue spread of $R(k)$. (see Fig. 2) then for stability it is sufficient to choose $\mu < \frac{1}{(1+\epsilon)}$.

VI. IMPLEMENTATION OF THE APPROXIMATE INVERSION TECHNIQUE

In this section, implementing the multiplication of $R^{-1}(k)x(k)$, or equivalently $P(k)x(k)$, using the DFT method is described in detail. As shown in section II, The main idea is to obtain an approximate inversion of $R(k)$ (i.e., $P(k)$) and apply transform techniques to carry out the multiplication $P(k)x(k)$ in the update equation. Now, considering the

sequence $\{q(n); n = -(N-1), \dots, (N-1)\}$ given in (4), the symmetric sequence $g_q(k)$ can be constructed as:

$$g_q(k) = \begin{cases} q_i, & 0 \leq i \leq (N-1) \\ q_i^*, & -(N-1) \leq i < 0 \end{cases} \quad (60)$$

The n^{th} element of the vector $p_f(k) = P(k)x(k)$ can be written as:

$$p_{f,n}(k) = \sum_{m=1}^N q_{n,m} x_{m-1}(k), \quad n = 1, 2, \dots, N. \quad (61)$$

Rewriting (61) in terms of the sequence in (60) gives

$$p_{f,n}(k) = \sum_{m=0}^{N-1} g_{n-m-1} x_m(k), \quad n = 1, 2, \dots, N, \quad (62)$$

Equation (62) represents the convolution sum. Now, taking $(2N-1)$ -point DFT of both sides of (62) at time k

$$P_{fe}(l) = G(l)X_e(l), \quad l = 1, 2, \dots, 2N-1, \quad (63)$$

where $P_{fe}(l)$ is the DFT of the zero-padded sequence $\{p_{fe,n}(k); n = 1, 2, \dots, 2N-1\}$:

$$p_{fe,n}(k) = \begin{cases} p_{f,n}(k), & n = 1, 2, \dots, N \\ 0, & n = N+1, \dots, 2N-1, \end{cases} \quad (64)$$

and $X_e(l)$ is the DFT of $x_e(k) = [x(k) \mathbf{0}]$ where $\mathbf{0}$ is an $(N-1)$ -dimensional zero vector. The sequence $\{p_{f,n}(k); n = 1, 2, \dots, N\}$ can now be recovered from the inverse DFT of $P_{fe}(l)$.

By applying this method, the computational complexity of the Newton algorithm will be significantly reduced as shown in Table 1.

TABLE 1. Computational complexity of AIN-LMS, Newton-LMS and improved QN-LMS algorithms.

Algorithm	Mul./Div.	Add./Sub.
AIN Algorithm	$0.5N^2 + N[5 + 3\log_2(N)]$	$0.5N^2 + N[1.5 + 9\log_2(N)]$
Newton Algorithm	$6N^2 + 3N + 2$	$5N^2 + 3N + 3$
improved QN	$N^2 + 4N + 1$	$2N^2 + N$

VII. SIMULATION RESULTS

To test the performance of the proposed algorithm, we compare its performance to the original Newton and the improved QN [14] algorithms in additive white Gaussian noise (AWGN) and additive white impulsive noise (AWIN) [30], [31] for noise cancellation and system identification settings. In all the experiments, the input signal is generated using a first-order autoregressive process (AR(1)) with $\rho = 0.6$ ($x(k) = 0.6x(k-1) + v_o(k)$) where $v_o(k)$ is assumed to be a white Gaussian process with zero mean and variance ($\sigma_{v_o}^2 = 0.15$). All the algorithms were implemented with the following parameters: For the proposed AIN and Newton algorithms: $\beta = 1$ and $\mu = 1$. For the improved QN algorithm: $\gamma = \sqrt{6\sigma_v^2}$ and $\zeta = 10^{-3}$. All the experiments were implemented with 1000 Monte-Carlo independent runs.

A. ADAPTIVE NOISE CANCELLATION

Some adaptive applications such as noise cancellation require large filter length (in this part we use filter length of $N = 30$ taps). In such applications, the computational complexity of the proposed algorithm becomes more prominent over the aforementioned algorithms. A block diagram of a noise cancellation setting is shown in Fig. 3.

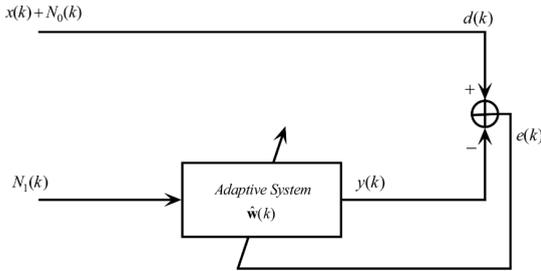


FIGURE 3. A Block diagram of the adaptive noise cancellation model.

First, an AWGN process is added to the input signal and the steady-state mean-square-errors (mse's) of all the methods were equated to compare their convergence rates. Fig. 4 shows that the proposed algorithm converges to the same steady-state mse (mse = -13.5dB) of the Newton and improved QN algorithms. Even though the improved QN algorithm starts with a faster convergence rate (due to the initialization of the algorithm), it slows down to converge to the same mse with the same rate as the other algorithms (all the algorithms converge at approximately 2500 iterations). The advantage of the proposed algorithm over the Newton and improved QN algorithms, here, appears in terms of reduction in the computational complexity as shown in Table 1.

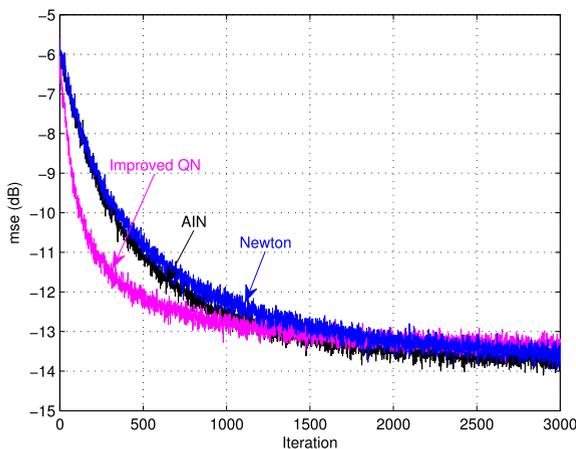


FIGURE 4. Ensemble mse for AIN, Improved QN and Newton algorithms in AWGN (Noise Cancellation).

In the second part of this experiment, we compare the performance of the algorithms in impulsive noise. We use an AWIN process with zero mean and variance ($\sigma_v^2 = 4 \times 10^{-4}$). The impulsive noise process generated by the probability density function [30]: $f = (1 - \epsilon) G(0, \sigma_n^2) + \epsilon G(0, \kappa \sigma_n^2)$ with variance σ_f^2 given as: $\sigma_f^2 = (1 - \epsilon) \sigma_n^2 + \epsilon \kappa \sigma_n^2$, where

$G(0, \sigma_n^2)$ is a Gaussian probability density function with zero mean and variance σ_n^2 that represents the nominal background noise. $G(0, \kappa \sigma_n^2)$ represents the impulsive component of the noise model, where ϵ is the probability and $\kappa \geq 1$ is the strength of the impulsive components, respectively. The white impulsive noise process is generated with the parameters: $\epsilon = 0.2$, $\kappa = 100$ and $\sigma_n^2 = 4 \times 10^{-4}$. Fig. 5 shows that the proposed AIN algorithm converges to the same mse as that of the Newton algorithm (mse = -20dB) with the same convergence rate (both algorithms converge at approximately 2500 iterations) with lower computational complexity for the AIN algorithm. The improved QN algorithm converges to higher mse than the other algorithms (mse = -17dB). It should be mentioned that, the mse of the improved QN cannot go below certain levels and this could be due to its structure that provides a fast convergence behavior at the beginning.

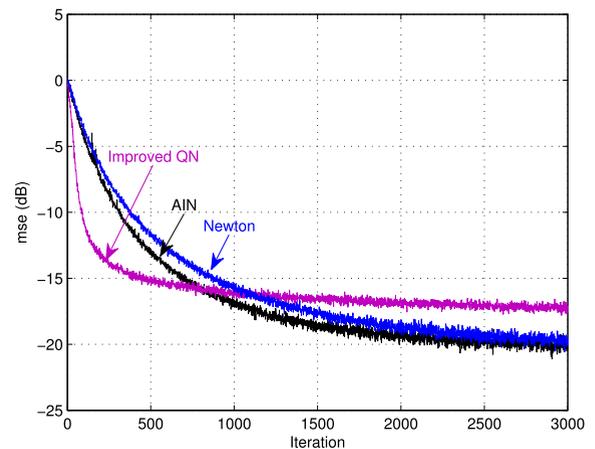


FIGURE 5. Ensemble mse for AIN, Improved QN and Newton algorithms in AWIN (Noise Cancellation).

B. ADAPTIVE SYSTEM IDENTIFICATION

To show the robustness of the proposed algorithm due to the change in the experimental setting, we now compare the performance of the aforementioned algorithms in a system identification setting. A block diagram of a system identification setting is shown in Fig. 6. Again the input signal is

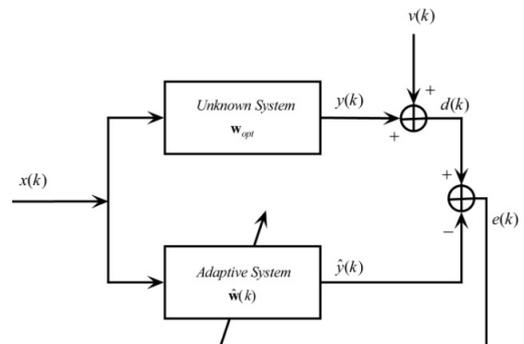


FIGURE 6. A Block diagram of the adaptive system identification model.

generated using the same AR(1) process used in the previous section with the same parameters. The unknown system was assumed to be a low-pass filter with length of $N = 12$ taps and a frequency magnitude shown in Fig. 7.

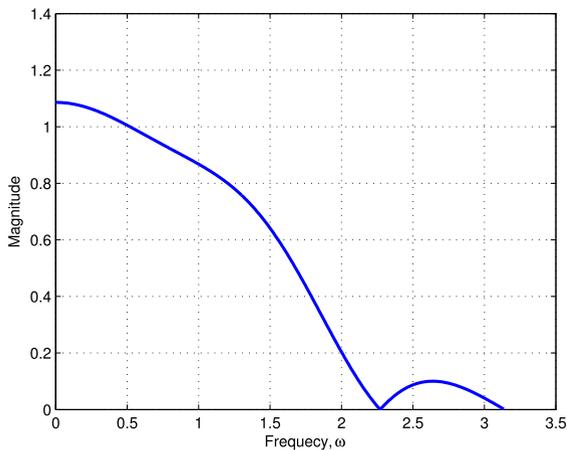


FIGURE 7. The magnitude response of the unknown system.

In the first part of this experiment, an AWGN process is added to the input signal and the steady-state mse's of all the methods were equated to compare their convergence rates. Fig. 8 shows that the proposed algorithm converges to the same steady-state mse (mse = -20dB) of the Newton and improved QN algorithms at the same rate (the AIN, Newton and improved QN algorithms converge at approximately 600 iterations).

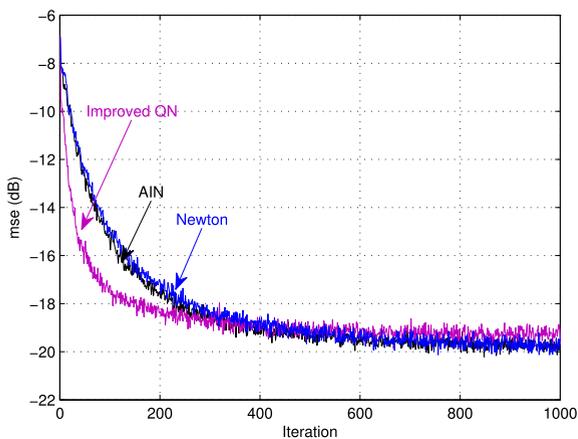


FIGURE 8. Ensemble mse for AIN, Improved QN and Newton algorithms in AWGN (System Identification).

In the second part, the convergence rates of all the methods were equated and their steady-state mse's were compared. The noise again here is assumed to be an AWIN with the same parameters in the second experiment of Section VII-A. Fig. 9 shows that the proposed AIN algorithm converges to the same mse as that of the Newton algorithm (mse = -20dB) with the same convergence rate (but lower computational complexity). However, the improved QN algorithm converges to a higher mse than the other algorithms (mse = -17.5dB).

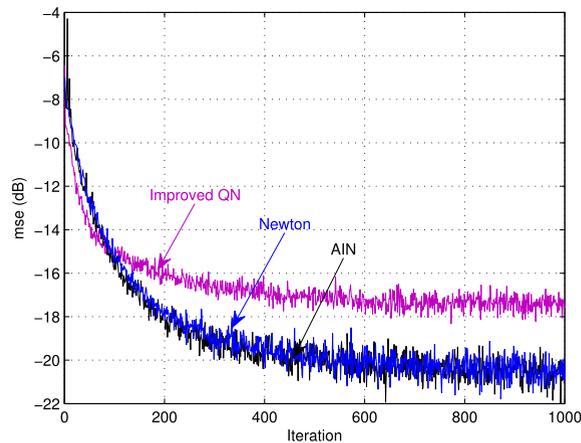


FIGURE 9. Ensemble mse for AIN, Improved QN and Newton algorithms in AWIN (System Identification).

VIII. CONCLUSION

In this paper, a new approximate inverse quasi-Newton (AIN) algorithm is proposed. The proposed AIN algorithm replaces the inverse of the input-signal autocorrelation matrix by an approximate one, assuming that the input-signal autocorrelation matrix is Toeplitz. Analysis of the proposed AIN algorithm is also provided. The performance of the proposed AIN algorithm is compared to those of the Newton and the improved QN algorithms in noise cancellation and system identification settings. The proposed AIN algorithm has lower computational complexity than that of the Newton algorithm without affecting the performance. The computational complexity of the proposed AIN algorithm is less than that of the improved QN algorithm when the filter length is relatively large with much better performance in terms of mse if the noise is impulsive (this could be due to the FFT process which helps in suppressing the impulsive components (outliers) of the noise process).

APPENDIX

The interpolation equations for $S(\omega_1)$ and $Q(\omega_2)$ are

$$S(\omega_1) = \frac{1}{N_s} \sum_{j=0}^{N_s-1} S(u_j) \tilde{A}_s(\omega_1 - u_j),$$

$$u_j = \frac{2\pi j}{N_s}, \quad j = 0, \dots, N_s - 1 \quad (65)$$

$$Q(\omega_2) = \frac{1}{N_s} \sum_{i=0}^{N_s-1} Q(u_i) \tilde{A}_s(\omega_2 - u_i),$$

$$u_i = \frac{2\pi i}{N_s}, \quad i = 0, \dots, N_s - 1 \quad (66)$$

In (65) and (66) $N_s = 2N - 1$ and

$$\tilde{A}_s(\omega) = \sum_{n=-N+1}^{N-1} e^{jn\omega} \quad (67)$$

It should be noted that $S(\omega)$ and $Q(\omega)$ are the DFTs of sequences defined for $n = -N + 1, \dots, N - 1$. Hence the interpolation function should be modified accordingly as

in (67). Substitution of (66) in the inner integral (wrt ω_2) in (22) yields

$$\int_0^{2\pi} Q(\omega_2)A(\omega_2 - \omega_1)A(\alpha_l - \omega_2)d\omega_2 = \frac{1}{N_s} \sum_{i=0}^{N_s-1} Q(u_i) \int_0^{2\pi} \tilde{A}_s(\omega_2 - u_i)A(\omega_2 - \omega_1)A(\alpha_l - \omega_2)d\omega_2 \quad (68)$$

The integral on the R.H.S. of (68) can be evaluated by using (12) and (67) as

$$\begin{aligned} \gamma_1(\alpha_l, u_i, \omega_1) &= \int_0^{2\pi} \tilde{A}_s(\omega_2 - u_i)A(\omega_2 - \omega_1)A(\alpha_l - \omega_2)d\omega_2 \\ &= \int_0^{2\pi} \sum_{n=-N+1}^{N-1} \sum_{m=0}^{N-1} \sum_{k=0}^{N-1} e^{-jn(\omega_2-u_i)} e^{-jm(\alpha_l-\omega_2)} e^{-jk(\omega_2-\omega_1)} d\omega_2 \\ &= \sum_{n=-N+1}^{N-1} \sum_{m=0}^{N-1} \sum_{k=0}^{N-1} e^{jnu_i} e^{-jm\alpha_l} e^{jk\omega_1} \int_0^{2\pi} e^{-j(n-m+k)\omega_2} d\omega_2 \\ &= 2\pi \sum_{n=-N+1}^{N-1} \sum_{m=0}^{N-1} \sum_{k=0}^{N-1} e^{jnu_i} e^{-jm\alpha_l} e^{jk\omega_1} \delta(n-m+k) \\ &= 2\pi \sum_{n=-N+1}^{N-1} \sum_{m \in I_0(n)} e^{jnu_i} e^{-jm\alpha_l} e^{j(m-n)\omega_1} \end{aligned} \quad (69)$$

where

$$I_0(n) = \begin{cases} [n, \dots, N-1], & n \geq 0 \\ [0, \dots, N-1+n], & n < 0 \end{cases} \quad (70)$$

Substitution of (69) in (22) and rearranging gives

$$c_{lk} = \frac{1}{(2\pi)^2 NN_s} \sum_{i=0}^{N_s-1} Q(u_i) \int_0^{2\pi} S(\omega_1)A(\omega_1 - \alpha_k)\gamma_1(\alpha_l, u_i, \omega_1)d\omega_1 \quad (71)$$

By using (65), the integral in (71) can be written as

$$\begin{aligned} \gamma_2(u_i, \alpha_l, \alpha_k) &= \int_0^{2\pi} S(\omega_1)A(\omega_1 - \alpha_k)\gamma_1(\alpha_l, u_i, \omega_1)d\omega_1 \\ &= \frac{1}{N_s} \sum_{j=0}^{N_s-1} S(u_j) \int_0^{2\pi} \tilde{A}_s(\omega_1 - u_j) \\ &\quad \times A(\omega_1 - \alpha_k)\gamma_1(\alpha_l, u_i, \omega_1)d\omega_1 \end{aligned} \quad (72)$$

Therefore, the η coefficients in (23) become,

$$\eta(u_i, u_j, \alpha_l, \alpha_k) = \frac{1}{(2\pi)^2 NN_s^2} \int_0^{2\pi} \tilde{A}_s(\omega_1 - u_j) \times A(\omega_1 - \alpha_k)\gamma_1(\alpha_l, u_i, \omega_1)d\omega_1 \quad (73)$$

Using (69), the integral in (73) becomes

$$\begin{aligned} &2\pi \sum_{v=-N+1}^{N-1} e^{jvu_j} \sum_{\mu=0}^{N-1} e^{j\mu\alpha_k} \sum_{n=-N+1}^{N-1} e^{jnu_j} \\ &\quad \sum_{m \in I_0(n)} e^{-jm\alpha_l} \int_0^{2\pi} e^{-j(v+\mu+n-m)\omega_1} d\omega_1 \\ &= (2\pi)^2 \sum_{v=-N+1}^{N-1} e^{jvu_j} \sum_{\mu=0}^{N-1} e^{j\mu\alpha_k} \sum_{n=-N+1}^{N-1} e^{jnu_j} \\ &\quad \sum_{m \in I_0(n)} e^{-jm\alpha_l} \delta(v + \mu + n - m) \end{aligned} \quad (74)$$

The last two summations in (74) can be written as

$$\begin{aligned} &\sum_{n=-N+1}^{N-1} e^{jnu_i} \sum_{m \in I_0(n)} e^{-jm\alpha_l} \delta(v + \mu + n - m) \\ &= \sum_{n \in I_1(v+\mu)} e^{-jnu_i} e^{-j(v+\mu+n)\alpha_l} \end{aligned} \quad (75)$$

where $I_1(v + \mu) = [-(v + \mu), \dots, N - 1 - (v + \mu)]$ and $v + \mu \in [0, \dots, N - 1]$. Substitution in (74) results in

$$\begin{aligned} \eta(u_i, u_j, \alpha_l, \alpha_k) &= \frac{1}{NN_s^2} \sum_{v=-N+1}^{N-1} \sum_{\mu \in I_2(v)} \\ &\quad \sum_{n \in I_1(v+\mu)} e^{jv(u_j-\alpha_l)} e^{j\mu(\alpha_k-\alpha_l)} e^{jn(u_i-\alpha_l)} \end{aligned} \quad (76)$$

where

$$I_2(v) = \begin{cases} [0, \dots, N-1-v], & v \geq 0 \\ [-v, \dots, N-1], & v < 0 \end{cases} \quad (77)$$

The summation in (76) can be readily performed and simplified to yield (24)

REFERENCES

- [1] C. V. Sinn and J. Gotze, "Comparative study of techniques to compute FIR filter weights in adaptive channel equalization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 6, Apr. 2003, pp. 217–220.
- [2] M. S. Salman, "Sparse leaky-LMS algorithm for system identification and its convergence analysis," *Int. J. Adapt. Control Signal Process.*, vol. 28, no. 10, pp. 1065–1072, Oct. 2014.
- [3] A. Gohn and J. Kim, "Implementation of LMS adaptive filter algorithm based on FPGA," in *Proc. IEEE 62nd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2019, pp. 207–210.
- [4] Z. Szadkowski, "Least mean squares filters suppressing the radio-frequency interference in AERA cosmic ray radio detection," *IEEE Trans. Nucl. Sci.*, vol. 67, no. 1, pp. 405–413, Jan. 2020, doi: 10.1109/TNS.2019.2948319.
- [5] B. Allen and M. Ghavami, *Adaptive Array Systems: Fundamentals and Applications*. West Sussex, U.K.: Wiley, 2005.
- [6] Y. Hao, C. Chi, L. Qiu, and G. Liang, "Sparsity-based adaptive line enhancer for passive sonars," *IET Radar, Sonar Navigat.*, vol. 13, no. 10, pp. 1796–1804, Oct. 2019.
- [7] Z. Zhang, L. Xiao, X. Su, J. Zeng, and X. Xu, "A channel estimation method based on the improved LMS algorithm for MIMO-OFDM systems," in *Proc. 12th Int. Symp. Med. Inf. Commun. Technol. (ISMICT)*, Mar. 2018, pp. 1–5.
- [8] J. Chueh, Y. Li, and B. Vucetic, "Design criteria on convergence analysis of LMS algorithms for layered space-time MIMO systems," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 4, Nov./Dec. 2004, pp. 2557–2561.

- [9] B. V. Veen, O. Leblond, V. Mani and D. Sebal, "Distributed adaptive algorithms for large dimensional MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, vol. 13, May 1998, pp. 1489–1492.
- [10] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
- [11] C. G. Tsinos and P. S. R. Diniz, "Data-selective LMS-Newton and LMS-Quasi-Newton algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4848–4852.
- [12] M. L. R. De Campos and A. Antoniou, "A new Quasi-Newton adaptive filtering algorithm," *IEEE Trans. Circuits Syst. II. Analog Digit. Signal Process.*, vol. 44, no. 11, pp. 924–934, Nov. 1997.
- [13] M. L. R. de Campos and A. Antoniou, "A robust Quasi-Newton adaptive filtering algorithm," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Nov. 1994, pp. 229–232.
- [14] M. Z. A. Bhotto and A. Antoniou, "Improved Quasi-Newton adaptive-filtering algorithm," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 8, pp. 2109–2118, Aug. 2010.
- [15] I. C. Gohberg and A. A. Semencul, "The inversion of finite Toeplitz matrices and their continual analogues," *Matematicheskie Issledovaniya*, vol. 7, no. 2, pp. 201–223, 1972.
- [16] G. Labahn and T. Shalom, "Inversion of toeplitz matrices with only two standard equations," *Linear Algebra its Appl.*, vol. 175, pp. 143–158, Oct. 1992.
- [17] P. Xie and Y. Wei, "The stability of formulae of the Gohberg–Semencul–Trench type for Moore–Penrose and group inverses of Toeplitz matrices," *Linear Algebra its Appl.*, vol. 498, pp. 117–135, Jun. 2016.
- [18] J. Jia, T. Sogabe, and M. E. A. El-Mikkawy, "Inversion of k -Tridiagonal matrices with Toeplitz structure," *Comput. Math. Appl.*, vol. 65, pp. 116–125, Jan. 2013.
- [19] M. El-Mikkawy and F. Atlan, "A new recursive algorithm for inverting general k -Tridiagonal matrices," *Appl. Math. Lett.*, vol. 44, pp. 34–39, Jun. 2015.
- [20] J. Jia and S. Li, "Symbolic algorithms for the inverses of general k -Tridiagonal matrices," *Comput. Math. Appl.*, vol. 70, no. 12, pp. 3032–3042, Dec. 2015.
- [21] Y. Zheng, Z. Fu, and S. Shon, "A new Toeplitz inversion formula, stability analysis and the value," *J. Nonlinear Sci. Appl.*, vol. 10, no. 03, pp. 1089–1097, 2017.
- [22] R. H. Chan and M. K. Ng, "Conjugate gradient methods for Toeplitz systems," *SIAM Rev.*, vol. 38, no. 3, pp. 427–482, Sep. 1996.
- [23] F.-R. Lin and W.-K. Ching, "Inverse Toeplitz preconditioners for hermitian Toeplitz systems," *Numer. Linear Algebra with Appl.*, vol. 12, nos. 2–3, pp. 221–229, Mar. 2005.
- [24] W. Hu, A. Abubakar, T. M. Habashy, and J. Liu, "Preconditioned non-linear conjugate gradient method for frequency domain full-waveform seismic inversion," *Geophys. Prospecting*, vol. 59, no. 3, pp. 477–491, May 2011.
- [25] M. Hanke and J. G. Nagy, "Toeplitz approximate inverse preconditioner for banded toeplitz matrices," *Numer. Algorithms*, vol. 7, no. 2, pp. 183–199, Sep. 1994.
- [26] A. Bottcher and S. M. Grudsky, *Spectral Properties Banded Toeplitz Matrices*, Philadelphia, PA, USA: SIAM, 2005.
- [27] R. H. Chan and K.-P. Ng, "Toeplitz preconditioners for Hermitian Toeplitz systems," *Linear Algebra And its Appl.*, vol. 190, pp. 181–208, 1993.
- [28] M. S. Salman, O. Kukrer, and A. Hocanin, "A fast implementation of quasi-Newton LMS algorithm using FFT," in *Proc. 2nd Int. Conf. Digit. Inf. Commun. Technol. Appl. (DICTAP)*, Bangkok, Thailand, May 2012, pp. 510–513.
- [29] G.-O. Glentis, K. Berberidis, and S. Theodoridis, "Efficient least squares adaptive algorithms for FIR transversal filtering," *IEEE Signal Process. Mag.*, vol. 16, no. 4, pp. 13–41, Jul. 1999.

- [30] H. Delic and A. Hocanin, "Robust detection in DS-CDMA," *IEEE Trans. Veh. Technol.*, vol. 51, no. 1, pp. 155–170, Jan. 2002.
- [31] H. Li, X.-L. Li, M. Anderson, and T. Adali, "A class of adaptive algorithms based on entropy estimation achieving CRLB for linear non-Gaussian filtering," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2049–2055, Apr. 2012.



MOHAMMAD SHUKRI SALMAN

(Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical and electronic engineering from Eastern Mediterranean University (EMU), Gazimagusa, TRNC, in 2006, 2007, and 2011, respectively. From 2011 to 2015, he has worked as an Assistant Professor with the Department of Electrical and Electronics Engineering, Mevlana University, Turkey. He is currently an Associate Professor with the Electrical Engineering Department, American University of the Middle East, Kuwait. He has over 60 refereed journals and conference publications. His research interests include signal processing, adaptive filters, image processing, and sparse signal representation. He has served as a Guest Editor, the General Chair, the Program Chair, and a TPC Member for many international journals and conferences.



OSMAN KUKRER (Senior Member, IEEE)

received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Middle East Technical University (METU), Ankara, Turkey, in 1979, 1982, and 1987, respectively. From 1979 to 1985, he was a Research Assistant with the Department of Electrical and Electronics Engineering, METU. From 1985 to 1986, he was with the Department of Electrical and Electronics Engineering, Brunel University, London, U.K. He is currently a Professor with the Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Gazimagusa, Turkey. He has coauthored more than 50 technical articles in international journals and conferences. His research interests include power electronics, control systems, and signal processing. He is also a member of the Chamber of Electrical Engineers in North Cyprus.



AYKUT HOCANIN (Senior Member, IEEE)

received the B.Sc. degree in electrical and computer engineering from Rice University, Houston, TX, USA, in 1992, the M.E. degree from Texas A&M University, College Station, TX, USA, in 1993, and the Ph.D. degree in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 2000. He joined the Electrical and Electronic Engineering Department, Eastern Mediterranean University, Gazimagusa, North Cyprus, in 2000. He has served as the Chairman of the Electrical and Electronic Engineering Department, from 2007 to 2014. He has been the Dean of the Engineering Faculty, since 2014. His current research interest includes receiver design for wireless systems and adaptive filtering. He is a Senior Member of the ACM.

...