Adaptive Filtering Based on Legendre Polynomials

Lu Shen, *Student Member, IEEE* and Yuriy Zakharov, *Senior Member, IEEE*
Long Shi, *Student Member, IEEE* and Benjamin Henson, *Member, IEEE*

Abstract—In system identification scenarios, classical adaptive filters, such as the recursive least squares (RLS) algorithm, predict the system impulse response. If a tracking delay is acceptable, interpolating estimators capable of providing more accurate estimates of time-varying impulse responses can be used; channel estimation in communications is an example of such applications. The basis expansion model (BEM) approach is known to be efficient for non-adaptive (block) channel estimation in communications. In this paper, we combine the BEM approach with the sliding-window RLS (SRLS) algorithm and propose a new family of adaptive filters. Specifically, we use the Legendre polynomials, thus the name the SRLS-L adaptive filter. The identification performance of the SRLS-L algorithm is evaluated analytically and via simulation. The analysis shows significant improvement in the estimation accuracy compared to the SRLS algorithm and a good match between the theoretical and simulation results. The performance is further investigated in application to the self-interference cancellation in full-duplex underwater acoustic communications, where a high estimation accuracy is required. A field experiment conducted in a lake shows significant improvement in the cancellation performance compared to the classical SRLS algorithm.

Index Terms—Adaptive filter, BEM, full-duplex, time-varying channel estimation, underwater acoustic communications

I. INTRODUCTION

Classical adaptive filtering algorithms such as the recursive least squares (RLS) algorithm, when used in identification scenarios, predict the system response for the next time instant based on input data received at the current and past time instants. In general, for time-varying systems, predictive estimators are less accurate than interpolating estimators. The later, however, are non-causal since they require input data not only from the past but also from future time instants. If an application can accept a tracking delay, the interpolating adaptive filtering can significantly improve the identification performance. Channel estimation in communications is an example of such applications.

Underwater acoustic (UWA) channels are considered to be difficult for communications [1]. They are characterised by multipath propagation with long delay spreads due to multiple reflections from the sea surface and bottom, refraction and low speed of sound. The motion of the seawater and transmitter/receiver result in the Doppler effect, which makes a UWA channel fast time-varying. For reliable UWA communication systems operating in fast-varying channels, the channel estimation is an essential part. RLS based adaptive filters are normally employed for this purpose [2]–[5]. However, the performance of the RLS adaptive filter is limited in fast-varying channels.

Another application where efficient adaptive filtering is the key to the successful operation is the self-interference cancellation (SIC) in full-duplex (FD) UWA systems [6], [7]. FD operation allows simultaneous transmission and reception of signals in the same frequency bandwidth by closely positioned transmit and receive antennas. The key problem in FD implementation is to cancel the strong self-interference (SI) from the near-end transmitter [6], [8]. Classical RLS adaptive filters can efficiently solve this problem in a time-invariant environment [7]. However, in practice, the SI channel can be fast time-varying [9], especially when the transmit and receive antennas are positioned close to the sea/lake surface [10].

In communication systems, basis expansion models (BEMs) are widely used for block (non-adaptive) estimation of time-varying channels, e.g. see [11]–[24]. The most often used BEMs are the Karhunen-Loeve functions [12], [14], discrete prolate spheroidal functions [16], [22]–[24], generalized complex exponentials [15]–[17], B-splines [18]–[21], and algebraic polynomials [25] including Legendre polynomials [26], [27]. With a BEM, estimation of a realization of the random process describing the time-variant channel is transformed into the simpler problem of estimation of a vector of time-invariant expansion coefficients [22].

In [28] a predictive RLS adaptive filter was proposed based on representation of the system time variation with algebraic polynomials. In [29], for the UWA FD application, we have proposed an interpolating adaptive filter based on the sliding-window RLS (SRLS) algorithm. This adaptive filter exploits a parabolic approximation of the time-varying SI channel response; however, it has a high complexity.

In this paper, we combine the SRLS adaptive algorithm with the BEM approach and propose a new family of (interpolating) adaptive filters, designed particularly for estimation of fast time-varying channels. We investigate their performance through theoretical analysis, numerical simulation and lake experiments. The contributions of this paper are as follows:

- 1) A new family of adaptive filters is proposed for improved identification of time-varying systems. The adaptive filters combine the BEM approach and SRLS algorithm.
- 2) Within this family, the SRLS-L adaptive filter is proposed using Legendre algebraic polynomials.
- 3) The analysis of computational complexity of the SRLS-L adaptive filter is given, and techniques are proposed to reduce the complexity.
- 4) The theoretical analysis of the tracking performance of the SRLS-L adaptive filter is presented. This analysis provides a good match with simulation results in system identification scenarios.

- 5) The SRLS-L adaptive filter is practically tested in an SIC scenario for an FD UWA system and shows improvement in the cancellation performance compared to the classical SRLS algorithm.

The rest of the paper is organized as follows. In Section II, the SRLS-L adaptive filter is derived. Section III presents the tracking performance of the SRLS-L adaptive filter. Section IV compares theoretical and simulation results, and presents the performance of the SRLS-L algorithm in an FD scenario. In Section V, conclusions are drawn.

Notations: In this paper, we use capital and small bold fonts for matrices and vectors, respectively; e.g., $\mathbf{R}$ and $\mathbf{h}$. The calligraphic bold font $\mathcal{R}$ also represents a matrix. We denote the complex conjugate as $(\cdot)^*$, transpose of $\mathbf{h}$ as $\mathbf{h}^T$, and the Hermitian transpose of $\mathbf{h}$ as $\mathbf{h}^H$. An $L \times L$ identity matrix is denoted as $\mathbf{I}_L$. The p th column of matrix $\mathcal{R}$ is denoted as $[\mathcal{R}]_{:,p}$. The norm of a vector is denoted as $\|\cdot\|$ and $\|\mathbf{h}\|^2 = \mathbf{h}^H \mathbf{h}$. The order of magnitude is denoted as $\mathcal{O}(\cdot)$. The trace and diagonal of a matrix are denoted as $\text{tr}\{\cdot\}$ and $\text{diag}\{\cdot\}$, respectively. The expectation operation is denoted as $\mathbb{E}\{\cdot\}$.

II. PROPOSED SRLS-L ADAPTIVE FILTER

In this section, we introduce the signal model in subsection II-A. In subsection II-B, we derive the BEM based SRLS algorithm in a general form and specify it for Legendre polynomials. Finally, in subsection II-C we propose techniques for reducing the algorithm complexity by exploiting time-shifted structure of the regressor.

A. Signal model

The observed discrete-time signal at time instant i is given by

$$x(i) = \mathbf{h}^H(i) \mathbf{s}(i) + n(i), \quad (1)$$

where $\mathbf{s}(i) = [s(i), s(i-1), \dots, s(i-L+1)]^T$ is an $L \times 1$ regressor vector, $s(i)$ is the input signal to the system (channel) with a time-varying impulse response $\mathbf{h}(i)$ to be identified, and $n(i)$ is a noise signal. For derivation in this section, we will ignore the noise $n(i)$.

We assume that within a time interval $[i - M_o, i + M_o]$ centred at the time instant i , the time-varying response can be accurately approximated by an algebraic polynomial of the P th order:

$$\mathbf{h}(i+k) = \sum_{p=0}^P \mathbf{c}_p(i) \phi_p(k), \quad k = -M_o, \dots, M_o, \quad (2)$$

where $M_o = (M-1)/2$ and M is the sliding window length; here, for convenience of presentation, we will assume that M is odd, however the results can be easily extended to the case of even M . The $L \times 1$ vectors of expansion coefficients $\mathbf{c}_p(i)$ should be estimated, where L is the length of the impulse response.

The orthogonal Legendre algebraic polynomials, which we will be using as an example BEM, are defined as [30]:

$$L_p(\tau) = \frac{1}{2^p p!} \frac{d^p}{d\tau^p} [(\tau^2 - 1)^p], \quad -1 < \tau < 1, \quad p \geq 0. \quad (3)$$

Fig. 1: Legendre polynomials.

Fig. 1 shows the first four Legendre polynomials. The basis functions $\phi_p(k)$ are then given by

$$\phi_p(k) = L_p\left(\frac{2k}{M-1}\right). \quad (4)$$

B. SRLS-L algorithm

We now consider the following vectors:

$$\mathbf{b}_p(i) = \mathbf{S}^T(i) \Phi_p \mathbf{x}^*(i), \quad p = 0, \dots, P, \quad (5)$$

where $\Phi_p = \text{diag}\{\phi_p(M_o), \dots, \phi_p(-M_o)\}$ is an $M \times M$ diagonal matrix, $\mathbf{S}(i) = [\mathbf{s}(i), \dots, \mathbf{s}(i-M+1)]^T$ is the $M \times L$ regressor matrix, and $\mathbf{x}(i) = [x(i), x(i-1), \dots, x(i-M+1)]^T$ is an $M \times 1$ desired signal vector. Equation (5) can be written as:

$$\begin{aligned} \mathbf{b}_p(i) &= \sum_{k=0}^{2M_o} \phi_p(M_o - k) \mathbf{s}(i-k) \mathbf{s}^H(i-k) \mathbf{h}(i-k) \\ &= \sum_{k=-2M_o}^0 \phi_p(M_o + k) \mathbf{R}(i+k) \mathbf{h}(i+k), \end{aligned} \quad (6)$$

where the second equality is obtained by replacing k with $-k$, $\mathbf{R}(i) = \mathbf{s}(i) \mathbf{s}^H(i)$ and, further replacing i with $i + M_o$,

$$\mathbf{b}_p(i + M_o) = \sum_{k=-M_o}^{M_o} \phi_p(k) \mathbf{R}(i+k) \mathbf{h}(i+k). \quad (7)$$

By substituting (2) into (7), we obtain

$$\begin{aligned} \mathbf{b}_p(i + M_o) &= \sum_{k=-M_o}^{M_o} \phi_p(k) \mathbf{R}(i+k) \sum_{q=0}^P \mathbf{c}_q(i) \phi_q(k) \\ &= \sum_{q=0}^P \mathcal{R}_{p,q}(i) \mathbf{c}_q(i), \end{aligned} \quad (8)$$

where

$$\mathcal{R}_{p,q}(i) = \sum_{k=-M_o}^{M_o} \phi_p(k) \mathbf{R}(i+k) \phi_q(k). \quad (9)$$

TABLE I: SRLS-L algorithm

Step	Equation
	for $i > 0$, repeat:
1	Generate vectors $\mathbf{b}_p(i + M_o)$ using (5)
2	Compute matrices $\mathcal{R}_{p,q}(i)$ as in (9)
3	Generate the matrix $\mathcal{R}(i)$ as in (10) and vector $\mathbf{b}(i)$
4	Find a solution $\hat{\mathbf{c}}$ to the system $[\mathcal{R}(i) + \epsilon \mathbf{I}_{(P+1)L}] \mathbf{c}(i) = \mathbf{b}(i)$
5	Compute the estimate $\hat{\mathbf{h}}(i) = \sum_{p=0}^P \hat{\mathbf{c}}_p(i) \phi_p(0)$

By further denoting

$$\mathcal{R}(i) = \begin{bmatrix} \mathcal{R}_{0,0}(i) & \dots & \mathcal{R}_{0,P}(i) \\ \dots & \dots & \dots \\ \mathcal{R}_{P,0}(i) & \dots & \mathcal{R}_{P,P}(i) \end{bmatrix}, \quad (10)$$

$\mathbf{b}(i) = [\mathbf{b}_0^T(i + M_o), \dots, \mathbf{b}_P^T(i + M_o)]^T$ and $\mathbf{c}(i) = [\mathbf{c}_0^T(i), \dots, \mathbf{c}_{P-1}^T(i)]^T$, we obtain a system of equations with respect to the unknown $(P+1)L \times 1$ vector $\mathbf{c}(i)$:

$$\mathcal{R}(i)\mathbf{c}(i) = \mathbf{b}(i). \quad (11)$$

By solving this system, we find an estimate $\hat{\mathbf{c}}(i)$ of the expansion coefficients $\mathbf{c}(i)$ for the representation of the time-varying impulse response $\mathbf{h}(i)$ in (2). We are only interested in the estimate for $k = 0$; therefore, the channel estimate at time instant i is given by

$$\hat{\mathbf{h}}(i) = \sum_{p=0}^P \hat{\mathbf{c}}_p(i) \phi_p(0). \quad (12)$$

The SRLS-L algorithm is summarized in Table I, where $\epsilon > 0$ is a regularization parameter used to stabilize the solution of the system in (11), and $\mathbf{I}_{(P+1)L}$ is an $(P+1)L \times (P+1)L$ identity matrix. Note that the sliding window length M in the SRLS-L algorithm should satisfy the condition $M > (P+1)L$; otherwise, the matrix $\mathcal{R}(i)$ will have rank less than $(P+1)L$ and the system (11) will be ill-conditioned.

C. Complexity of the SRLS-L adaptive filter

The complexity of the SRLS-L algorithm is mainly determined by the computation of the matrices $\mathcal{R}_{p,q}(i)$ in (9) and vectors $\mathbf{b}_p(i + M_o)$ in (5) and solving the system of equations (11). We first show how the complexity of the matrix and vector computation can be reduced and then discuss a reduction in complexity by recursively solving the system of equations.

1) *Matrix $\mathcal{R}_{p,q}(i)$* : The direct computation of the matrix $\mathcal{R}_{p,q}(i)$ would require an order of ML^2 , i.e. $\mathcal{O}(ML^2)$, arithmetic operations. We will show that this can be reduced to $\mathcal{O}(ML)$ operations, and then further reduced to $\mathcal{O}((L+M)\log_2(L+M))$ operations. Firstly, we show that $\mathcal{R}_{p,q}(i)$ can be efficiently computed using elements in $\mathcal{R}_{p,q}(i-1)$.

Proposition: The following relationship holds for elements $[\mathcal{R}_{p,q}(i)]_{m+1,n+1}$, $m, n = 1, \dots, L-1$, of the matrix $\mathcal{R}_{p,q}(i)$:

$$[\mathcal{R}_{p,q}(i)]_{m+1,n+1} = [\mathcal{R}_{p,q}(i-1)]_{m,n}. \quad (13)$$

Proof: The matrix $\mathcal{R}_{p,q}(i)$ in (9) can be represented as:

$$\mathcal{R}_{p,q}(i) = \sum_{k=-\infty}^{\infty} w_{p,q}(k) \mathbf{R}(i+k), \quad (14)$$

where the weights $w_{p,q}(k)$ are defined as:

$$w_{p,q}(k) = \begin{cases} \phi_p(k)\phi_q(k), & \text{if } -M_o \leq k \leq M_o \\ 0, & \text{otherwise} \end{cases}. \quad (15)$$

This can also be rewritten as:

$$\mathcal{R}_{p,q}(i) = \sum_{k=-\infty}^{\infty} w_{p,q}(k-i) \mathbf{R}(k). \quad (16)$$

The $(m+1, n+1)$ th element of $\mathcal{R}_{p,q}(i)$ is given by

$$[\mathcal{R}_{p,q}(i)]_{m+1,n+1} = \sum_{k=-\infty}^{\infty} w_{p,q}(k-i) [\mathbf{R}(k)]_{m+1,n+1}. \quad (17)$$

Since $[\mathbf{R}(k)]_{m+1,n+1} = s(k-m)s^*(k-n)$, we have

$$[\mathcal{R}_{p,q}(i)]_{m+1,n+1} = \sum_{k=-\infty}^{\infty} w_{p,q}(k-i) s(k-m)s^*(k-n). \quad (18)$$

The (m, n) th element of $\mathcal{R}_{p,q}(i-1)$ is given by

$$[\mathcal{R}_{p,q}(i-1)]_{m,n} = \sum_{k=-\infty}^{\infty} w_{p,q}(k-i+1) [\mathbf{R}(k)]_{m,n}. \quad (19)$$

Since $[\mathbf{R}(k)]_{m,n} = s(k-m+1)s^*(k-n+1)$, we obtain

$$\begin{aligned} & [\mathcal{R}_{p,q}(i-1)]_{m,n} = \\ &= \sum_{k=-\infty}^{\infty} w_{p,q}(k-i+1) s(k-m+1)s^*(k-n+1) \\ &= \sum_{k=-\infty}^{\infty} w_{p,q}(k-i) s(k-m)s^*(k-n) \\ &= [\mathcal{R}_{p,q}(i)]_{m+1,n+1}. \end{aligned} \quad (20)$$

This equality holds for $m, n = 1, \dots, L-1$.

This proposition shows that $(L-1)^2$ elements of the matrix $\mathcal{R}_{p,q}(i)$ at time instant i are the same as elements of the matrix $\mathcal{R}_{p,q}(i-1)$ at time instant $i-1$. Therefore, only $2L-1$ elements of the matrix require computation. Since the matrix $\mathcal{R}_{p,q}(i)$ is Hermitian, only one column of length L should be updated. Thus, the complexity of the computation is reduced from $\mathcal{O}(ML^2)$ to $\mathcal{O}(ML)$.

We now show how the fast Fourier transform (FFT) can be used to further reduce the complexity. We only need to compute the first column of the matrix $\mathcal{R}_{p,q}(i)$. This column can be represented as

$$\begin{aligned} [\mathcal{R}_{p,q}(i)]_{:,1} &= \sum_{k=-M_o}^{M_o} w_{p,q}(k) s(i+k) \mathbf{s}^*(i+k) \\ &= \sum_{k=-M_o}^{M_o} \tilde{w}_{p,q}(k) \mathbf{s}^*(i+k), \end{aligned} \quad (21)$$

where $\tilde{w}_{p,q}(k) = w_{p,q}(k)s(i+k)$. We therefore can think of the column $[\mathcal{R}_{p,q}(i)]_{:,1}$ as a result of convolution of sequences $[s^*(i-M_o-L+1), \dots, s^*(i+M_o)]$ and $\tilde{w}_{p,q}(k)$, $k = -M_o, \dots, M_o$. For this purpose, the FFT of a size $(M+L)$ or bigger can be used as follows. One needs to compute FFTs of these two sequences, multiply them, and compute the inverse FFT. Therefore, the complexity of these computations is about $3(L+M)\log_2(L+M)$ operations. This

is instead of the direct computation, which would require about $2ML$ arithmetic operations.

Thus, depending on the filter length L and the sliding window length M , the complexity of computing the matrix $\mathcal{R}_{p,q}(i)$ at every time instant i can be either $\mathcal{O}(ML)$ or $\mathcal{O}((L+M)\log_2(L+M))$, whatever is smaller.

2) *Vector $\mathbf{b}_p(i+M_o)$* : The direct computation of the vector requires $\mathcal{O}(ML)$ operations. Note that (5) can be thought of as convolution of the sequences $s(i+k)$ of length $L+M$ and $\mathbf{x}_p^*(i) = \Phi_p \mathbf{x}^*(i)$ of length M . This again can be done using the FFT of length $(L+M)$ with a complexity of $\mathcal{O}((L+M)\log_2(L+M))$.

3) *System of equations (11)*: Direct solution of the system would require $\mathcal{O}((P+1)^3 L^3)$ operations. The system at time instant i however can be solved recursively using the solution obtained at time instant $i-1$ for the initialization, e.g. by applying the dichotomous coordinate descent (DCD) algorithm [31], [32]. This will lead to the complexity $\mathcal{O}(N_u(P+1)L)$ where N_u is the number of DCD updates. If $N_u < (P+1)^2 L^2$, this approach reduces the complexity. We will investigate this relationship in Section IV and show that the choice of N_u will depend on the speed of the system response variation and required accuracy of the solution.

III. TRACKING PERFORMANCE OF THE SRLS-L ADAPTIVE FILTERS

In this section, we investigate the steady-state Mean Square Deviation (MSD) performance of the SRLS-L algorithm in time-varying channels.

Consider the signal model (1). The observed signal vector $\mathbf{x}(i) = [x(i), \dots, x(i-M+1)]^T$ over the length M of the sliding window is given by

$$\mathbf{x}(i) = \text{diag}\{\mathbf{H}^H(i)\mathbf{S}^T(i)\} + \mathbf{n}(i), \quad (22)$$

where $\mathbf{H}(i) = [\mathbf{h}(i) \cdots \mathbf{h}(i-M+1)]$ is an $L \times M$ matrix whose columns are the channel response vectors, and $\mathbf{n}(i) = [n(i), \dots, n(i-M+1)]^T$ is an $M \times 1$ noise vector.

1) *Estimation matrix*: We first represent the estimate of the impulse response $\mathbf{h}(i)$ in the form:

$$\hat{\mathbf{h}}(i) = \mathbf{A}_{\text{SRLS-L}}(i)\mathbf{x}^*(i+M_o), \quad (23)$$

where $\mathbf{A}_{\text{SRLS-L}}(i)$ is an $L \times M$ estimation matrix. We have $\hat{\mathbf{c}}_p(i) = \Gamma_p \hat{\mathbf{c}}(i)$, where Γ_p is an $L \times (P+1)L$ selection matrix: $\Gamma_p = [\mathbf{0}_L \cdots \mathbf{I}_L \cdots \mathbf{0}_L]$, $\mathbf{I}_L$ is an $L \times L$ identity matrix, and $\mathbf{0}_L$ is an $L \times L$ zero matrix; in the matrix Γ_p , the identity matrix is positioned after p zero matrices. Taking this into account, from (12) we obtain:

$$\begin{aligned} \hat{\mathbf{h}}(i) &= \sum_{p=0}^P \phi_p(0) \Gamma_p \hat{\mathbf{c}}(i) \\ &= \left[\sum_{p=0}^P \phi_p(0) \Gamma_p \right] \mathcal{R}^{-1}(i) \mathbf{b}(i) \\ &= \Gamma \mathcal{R}^{-1}(i) \mathbf{b}(i), \end{aligned} \quad (24)$$

where $\Gamma = [\phi_0(0)\mathbf{I}_L \cdots \phi_P(0)\mathbf{I}_L]$. Since the vector $\mathbf{b}(i)$ can be represented as $\mathbf{b}(i) = \mathbf{G}(i+M_o)\mathbf{x}^*(i+M_o)$,

where $\mathbf{G}(i) = [\mathbf{S}^T(i)\Phi_0; \cdots; \mathbf{S}^T(i)\Phi_P]$ is an $(P+1)L \times M$ matrix, we obtain the SRLS-L estimation matrix:

$$\mathbf{A}_{\text{SRLS-L}}(i) = \Gamma \mathcal{R}^{-1}(i) \mathbf{G}(i+M_o). \quad (25)$$

Note that for the classical SRLS algorithm, the channel response estimate can be represented in a similar form, $\hat{\mathbf{h}}(i) = \mathbf{A}_{\text{SRLS}}(i)\mathbf{x}^*(i)$, where the estimation matrix is given by

$$\mathbf{A}_{\text{SRLS}}(i) = [\mathbf{S}^H(i)\mathbf{S}(i)]^{-1} \mathbf{S}^T(i). \quad (26)$$

2) *MSD*: The MSD is given by

$$\epsilon(i) = \mathbb{E} \left\{ \|\hat{\mathbf{h}}(i) - \mathbf{h}(i)\|^2 \right\}. \quad (27)$$

Assuming that the noise $\mathbf{n}(i)$ is independent of the channel response $\mathbf{h}(i)$ and signal $s(i)$, and using (22), the MSD can be represented as

$$\epsilon(i) = \epsilon_a(i) + \epsilon_n(i), \quad (28)$$

where $\epsilon_a(i)$ is an approximation (modelling) component:

$$\begin{aligned} \epsilon_a(i) &= \mathbb{E} \left\{ \|\mathbf{A}_{\text{SRLS-L}}(i) \text{diag}[\mathbf{H}^T(i+M_o)\mathbf{S}^H(i+M_o)] \right. \\ &\quad \left. - \mathbf{h}(i)\|^2 \right\}, \end{aligned} \quad (29)$$

and $\epsilon_n(i)$ is a noise component:

$$\epsilon_n(i) = \mathbb{E} \left\{ \|\mathbf{A}_{\text{SRLS-L}}(i)\mathbf{n}^*(i+M_o)\|^2 \right\}. \quad (30)$$

The MSD approximation component $\epsilon_a(i)$ depends on the speed (derivatives) of the system response time variation and approximation properties of the basis functions; this component does not depend on the noise. The MSD noise component $\epsilon_n(i)$ depends on the noise and does not depend on the system response.

3) *MSD approximation component*: The term $\epsilon_a(i)$ is the approximation error defined by the speed of the channel variations and approximation properties of the estimator. We assume that elements of the vector $\mathbf{h}(i)$ are independent zero-mean stationary random processes over time i , also independent of the noise and regressor $s(i)$. The speed of the channel variation is defined by a power spectral density $G_h^{(\ell)}(\omega)$, where ℓ indicates the ℓ th tap $[\mathbf{h}(i)]_\ell$ of the system response vector $\mathbf{h}(i)$.

We now consider the case $L = 1$ and obtain: $\Gamma = [\phi_0(0), \dots, \phi_P(0)]$ is an $1 \times (P+1)$ vector; $\mathcal{R}(i)$ is an $(P+1) \times (P+1)$ matrix with elements $\mathcal{R}_{p,q}(i) = \sum_{k=-M_o}^{M_o} \phi_p(k)\phi_q(k)|s(i+k)|^2$; $\mathbf{G}(i+M_o)$ is an $(P+1) \times M$ matrix with elements $[\mathbf{G}(i+M_o)]_{p,m} = s(i+M_o-m+1)\phi_p(M_o-m+1)$. If the signal $s(i)$ has a constant magnitude u , $|s(i)| = u$, then the matrix elements are $\mathcal{R}_{p,q}(i) = u^2 \sum_{k=-M_o}^{M_o} \phi_p(k)\phi_q(k)$, and if, in addition, the basis functions are orthogonal, i.e. $\mathcal{R}_{p,q}(i) = 0$ for $p \neq q$, we obtain

$$\hat{h}(i) = \sum_{p=0}^P \hat{c}_p(i) \phi_p(0), \quad (31)$$

where

$$\hat{c}_p(i) = \frac{1}{E_p} \sum_{k=-M_o}^{M_o} h(i+k) \phi_p(k) \quad (32)$$

and $E_p = \sum_{k=-M_o}^{M_o} \phi_p^2(k)$.

In Appendix A, we find an approximate expression for $\epsilon_a(i)$ using the continuous-time approximation of (32):

$$\hat{c}_p(i) = \frac{1}{E_p} \int_{-T/2}^{T/2} h(t) \phi_p(t) dt, \quad (33)$$

$$E_p = \int_{-T/2}^{T/2} \phi_p^2(t) dt, \quad (34)$$

where $T = MT_s$ and T_s is the sampling interval when converting continuous-time signals into discrete-time signals. Specifically, if the power spectral density of $h(t)$ is $G_h(\omega) = (2/T)G_{\tilde{h}}(\omega T/2)$, then

$$\begin{aligned} \epsilon_a(i) &\approx \frac{1}{2\pi} \sum_{p=0}^P \sum_{q=0}^P I_{p,q} \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) \tilde{L}_p^*(\omega) \tilde{L}_q(\omega) d\omega \\ &- \frac{1}{2\pi} \sum_{p=0}^P (2p+1) L_p(0) \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) \tilde{L}_p^*(\omega) d\omega \\ &+ \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) d\omega, \end{aligned} \quad (35)$$

where $I_{p,q}$ is given in (51) and $\tilde{L}_p(\omega)$ is the Fourier transform of the p th Legendre polynomial given in (55) [33]. The integrals in (35) can be computed numerically for given P , M and $G_{\tilde{h}}(\omega)$.

In Section IV, we will show by simulation that the expression (35) provides a good prediction for the MSD approximation component $\epsilon_a(i)$.

4) *MSD noise component*: Taking the expectation in (30) over the noise distribution and assuming that the noise samples $n(i)$ are independent zero-mean with variance σ_n^2 , we arrive at

$$\epsilon_n(i) = \sigma_n^2 \mathbb{E} \left\{ \text{tr} [\mathbf{A}_{\text{SRLS-L}}^H(i) \mathbf{A}_{\text{SRLS-L}}(i)] \right\}. \quad (36)$$

Using (25), this can be further written as

$$\begin{aligned} \epsilon_n(i) &= \sigma_n^2 \mathbb{E} \left\{ \text{tr} \left[\mathbf{G}^H(i + M_o) [\mathcal{R}^{-1}(i)]^H \mathbf{\Gamma}^T \right. \right. \\ &\quad \times \left. \left. \mathbf{\Gamma} \mathcal{R}^{-1}(i) \mathbf{G}(i + M_o) \right] \right\}. \end{aligned} \quad (37)$$

Taking into account the property $\text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]$ and the fact that $\mathcal{R}(i) = \mathbf{G}(i + M_o) \mathbf{G}^H(i + M_o)$, we obtain

$$\epsilon_n(i) = \sigma_n^2 \mathbb{E} \left\{ \text{tr} \left[\mathbf{\Gamma} [\mathcal{R}^{-1}(i)]^H \mathbf{\Gamma}^T \right] \right\}. \quad (38)$$

We now assume that $\mathcal{R}_{p,p}(i) \approx \sigma_s^2 E_p \mathbf{I}_L$, $\sigma_s^2 = \mathbb{E}\{|s(i)|^2\}$, $E_p = \sum_{k=-M_o}^{M_o} \phi_p^2(k)$, and that matrices $\mathcal{R}_{p,q}(i)$ with $p \neq q$ are zeros. Then $\mathcal{R}^{-1}(i)$ is also a diagonal matrix and we obtain

$$\epsilon_n(i) \approx \frac{L \sigma_n^2}{\sigma_s^2} \sum_{p=0}^P \frac{\phi_p^2(0)}{E_p}. \quad (39)$$

For the case $P = 0$, we then have

$$\epsilon_n(i) \approx \frac{L}{M} \cdot \frac{\sigma_n^2}{\sigma_s^2}. \quad (40)$$

IV. NUMERICAL RESULTS

In this section, in subsection IV-A, we investigate by simulation the identification performance of the SRLS-L algorithm in time-varying channels; we compare it with the theoretical results presented in Section III for a single path ($L = 1$) and multipath ($L = 10$) channels. In subsection IV-B, we show the effectiveness of using the DCD algorithm for solving the system of equations in (11). Finally, subsection IV-C presents results of applying the proposed SRLS-L algorithm to SI cancellation for FD UWA system.

A. Comparison of theoretical and simulation results

We will be assuming that the sampling rate of all the signals is 1 kHz; thus the adaptive filter taps are separated by a 1 ms interval. The ℓ th tap of the time-varying channel response $\mathbf{h}(i)$ is modelled as a stationary random process with a power spectral density $G_h(2\pi f)$, which is uniform within the frequency interval $f \in [-f_{\max}, f_{\max}]$ and independent of random processes describing the other taps. The maximum frequency $f_{\max} = 1$ Hz is used to simulate fast time-varying channels.

Two types of input (regressor) signals are considered: (i) binary input with $s(i) \in \{-1, +1\}$ being independent random binary numbers, $\mathbb{E}\{s(i)\} = 0$; (ii) Gaussian input with real and imaginary parts of $s(i)$ being independent random numbers with the standard Gaussian distribution. The signal-to-noise ratio (SNR) is defined as

$$\text{SNR} = \frac{\mathbb{E}\{|\mathbf{h}^H(i) \mathbf{s}(i)|^2\}}{\sigma_n^2} = \frac{L \sigma_h^2 \sigma_s^2}{\sigma_n^2},$$

where $\sigma_s^2 = \mathbb{E}\{|s(i)|^2\}$ and $\sigma_h^2 = (2\pi)^{-1} \int_{-2\pi f_{\max}}^{2\pi f_{\max}} G_h(\omega) d\omega$. The performance of the SRLS-L adaptive filter is evaluated in low and high SNR scenarios. The low SNR level is set to SNR = 30 dB; this SNR can be viewed as an upper limit for channel estimation in communications. The high SNR is set to SNR = 100 dB; the case of the high SNR level is important for FD applications, where the near-end SI level is significantly higher than the receiver's noise floor [7].

The simulation results are obtained by averaging over 100 simulation trials. In every simulation trial, the signal length is 5 s (5000 samples), and the MSD is averaged over the last 4000 samples.

1) *Single path channel* ($L = 1$): Fig. 2 shows the MSD performance for the binary input signal. We first notice that the SRLS-L algorithm significantly outperforms the classical SRLS algorithm. The typical dependence of the MSD on the length M of the sliding window has a minimum. For values of M lower than the minimum, the MSD is mostly defined by the noise component given by (38); as can be seen in (40), this noise component is approximately inversely proportional to M , and therefore the dependence is almost linear on a plot with logarithmic axis. For values of M higher than the minimum, the MSD is mostly defined by the approximation component given by (35). For binary signals and $L = 1$, the MSD approximation component ϵ_a should provide the best match to the theoretical analysis presented in Section III. Indeed we can observe that this is the case, the simulation

Fig. 2: MSD performance of the SRLS-L adaptive filter in the single path channel ($L = 1$) for the binary input signal.

Fig. 3: MSD performance of the SRLS-L adaptive filter in the single path channel ($L = 1$) for the Gaussian input signal.

results are very close to the theoretical ones. It is seen that, in this scenario, the zero-order ($P = 0$) and first order ($P = 1$) Legendre polynomials provide the same performance; similarly, the second-order ($P = 2$) and third order ($P = 3$) Legendre polynomials provide the same performance. This behaviour directly follows from analysis of (35) and (38). However, increasing the polynomial degree from $P = 0$ to $P = 3$ does bring an MSD performance improvement of about 5 dB and 10 dB for SNR = 30 dB and SNR = 100 dB, respectively.

Fig. 3 shows the MSD performance for the Gaussian input signal. Again, it can be seen that the SRLS-L algorithm significantly outperforms the classical SRLS algorithm. We observe that, in this scenario, the SRLS-L algorithm with the first order ($P = 1$) Legendre polynomials outperform that with the zero-order polynomials ($P = 0$), while the MSD for the case $P = 1$ matches very well to the theoretical prediction. This can be explained as follows.

For the case $P = 0$, the channel estimate is given by

$$\begin{aligned}\hat{h}(i) &= \hat{c}_0(i) = |s(i)|^{-2} b_0(i + M_o) \\ &= |s(i)|^{-2} \sum_{k=-M_o}^{M_o} \phi_0(k) |s(i+k)|^2 h(i+k).\end{aligned}\quad (41)$$

Let the channel response vary linearly in time, $h(i+k) = c_1 \phi_1(k)$, $k = -M_o, \dots, M_o$, and $h(i) = 0$. If the signal envelope $|s(i+k)| = u$ is constant, then $\hat{h}(i)$ in (41) is equal to zero due to the orthogonality of the basis functions: $c_1 \sum_{k=-M_o}^{M_o} \phi_0(k) \phi_1(k) = 0$. In this case, we expect that the theoretical results should match well to the simulation results, and this is what we observe in Fig. 2. If $|s(i+k)|$ is not a constant, then in general $\hat{h}(i) = c_1 |s(i)|^{-2} \sum_{k=-M_o}^{M_o} |s(i+k)|^2 \phi_0(k) \phi_1(k) \neq 0$ since, with the weight $|s(i+k)|^2$ in the sum, the basis functions no longer orthogonal. Thus, due to the non-constant signal envelope, for $P = 0$, there are discrepancies between the simulation results and theoretical results obtained under the assumption of a constant signal

envelope; this is what we observe in Fig. 3. With increasing M , the discrepancy is reduced since the sum $\sum_{k=-M_o}^{M_o} |s(i+k)|^2 \phi_0(k) \phi_1(k)$ will tend to zero; this can also be observed in Fig. 3.

For the case $P = 1$, the channel estimate is given by

$$\hat{h}(i) = \frac{\mathcal{R}_{1,1}b_0(i+M_o) - \mathcal{R}_{0,1}b_1(i+M_o)}{\det(\mathcal{R})}, \quad (42)$$

where $\det(\mathcal{R})$ is the determinant of $\mathcal{R}$. Let again $h(i+k) = c_1\phi_1(k)$. In this case, the first term of the numerator in (42) is given by $c_1 \sum_{k=-M_o}^{M_o} \phi_1^2(k) |s(i+k)|^2$. It is easy to check that the second term in (42), $\mathcal{R}_{0,1}b_1(i+M_o)$, is of exactly the same value. Therefore, independent of $|s(i+k)|$, we obtain $\hat{h}(i) = 0 = h(i)$. Thus, for $P = 1$, the estimate does not depend on the signal envelope, and the simulation and theoretical results should be close; this is what we observe in Fig. 3.

For the low SNR, the difference in the performance between the SRLS-L with $P = 2$ and $P = 3$ is small, and both versions of the SRLS-L algorithm show a close match between the simulation and theoretical results. For the high SNR, however, the SRLS-L algorithm with $P = 3$ matches well to the theoretical results and outperforms the SRLS-L algorithm with $P = 2$.

2) *Multipath Channel ($L = 10$)*: Fig. 4 shows the MSD performance for the case $L = 10$ and Gaussian input signals. We do not show results for the binary input since, for multipath channels, they are very close to results for the Gaussian signal shown in Fig. 4. Again, it can be seen that the SRLS-L algorithm significantly outperforms the classical SRLS algorithm. Despite the fact that the MSD approximation error were derived for the case of $L = 1$ and constant envelope of the regressor signal, still there is an almost perfect match between the simulation and theoretical results for $P = 1$ and $P = 3$. For $P = 0$ and $P = 2$, there is not such a good match, but it improves as M increases.

B. SRLS-L-DCD algorithm

As was discussed in subsection II-C, the complexity of the SRLS-L adaptive filter can be reduced if the system of equations at step 4 in Table I is solved using the DCD algorithm. We name this version of the SRLS-L algorithm as the SRLS-L-DCD algorithm.

Details on how the DCD algorithm can be incorporated into the SRLS adaptive filter can be found in [31]. The DCD algorithm solves the system of equations recursively where initialisation of the solution at time instant i is the solution found at time instant $i - 1$. A special version of the coordinate descent method is used, called the leading DCD. The parameter N_u defines the maximum number of ‘successful’ DCD iterations per time instant. At the ‘successful’ iterations, the solution is updated; one update requires $\mathcal{O}((P+1)L)$ arithmetic operations. The parameter N_u defines the upper limit to the complexity of the algorithm, since iterations without updates require negligible computation. We can expect to require a larger N_u , if the channel is fast varying, in which

Fig. 4: MSD performance of the SRLS-L adaptive filter in the multipath channel ($L = 10$) for the Gaussian input signal.

case the solution significantly changes between time instants $i - 1$ and i , and if the SNR is high, in which case the change should be estimated with higher precision. This is also seen from the simulation results below.

In this subsection, we investigate the performance and complexity of the SRLS-L-DCD algorithm in comparison with the original SRLS-L algorithm. We consider the low SNR ($\text{SNR} = 30$ dB) and high SNR ($\text{SNR} = 100$ dB) cases for fast ($f_{\max} = 1$ Hz) and slow ($f_{\max} = 0.1$ Hz) time-varying channels. The low SNR case represents scenarios of channel estimation in communications, while the high SNR represents FD scenarios. The same channel model is used as the one described in subsection IV-A. The filter length is $L = 10$ taps and the input signal has the Gaussian distribution. The MSD performance of the adaptive filters is averaged over 100 simulation trials. The sliding-window length M is chosen to provide the best MSD performance for the SRLS-L adaptive filter.

As an illustration, Fig. 5 shows the MSD performance for

(a) SNR = 30 dB, fast varying channel

(b) SNR = 100 dB, slow varying channel

Fig. 5: MSD performance of the SRLS-L and SRLS-L-DCD adaptive algorithms.

the low SNR fast-varying channel and high SNR slow varying channel. It can be seen that both the algorithms provide similar MSD performance.

Table II presents results of the comparison. Several conclusions can be drawn by analysing these results. As can be seen, the MSDs for the two algorithms differ by less than 1 dB, so that the estimation performance is almost the same.

At lower SNR, the number of updates N_u required to achieve this MSD is reduced and, for all the cases, $N_u \ll (P+1)^2 L^2$. Therefore, at lower SNR, the SRLS-L-DCD algorithm is computationally significantly more efficient than the SRLS-L algorithm. Note that for typical channel estimation scenarios in communications, the SNR can even be lower than SNR = 30 dB, and thus the complexity will be further reduced.

At the high SNR, for the slow varying channel, we still have $N_u \ll (P+1)^2 L^2$, so that the SRLS-L-DCD algorithm is still computationally more efficient than the SRLS-L algorithm. In the fast varying channel, at the high SNR and $P = 2$, the

Fig. 6: Transmit (Tx) and receive (Rx) antennas in the lake experiment.

complexity of the two algorithms is comparable. However, even in this extreme case, for $P = 0$ and $P = 1$, the SRLS-L-DCD algorithm requires less computations. Note that SNR = 100 dB is an upper limit for SIC in FD systems; in many cases, it will be lower. Moreover, the DCD algorithm is especially efficient when implemented on real-time hardware platforms, such as an FPGA [32]. Therefore, for practical implementation, the SRLS-L-DCD algorithm is preferable over the SRLS-L algorithm.

C. Full-duplex self-interference cancellation

In this section, the SRLS-L adaptive filter is used for the SI channel estimation in an FD UWA system. We conducted a lake experiment to evaluate the performance of the SRLS-L adaptive filter in such a practical scenario. In the lake experiment, the true channel response is unknown, and therefore the MSD performance cannot be measured. Instead we measure the SIC factor which is the ratio of signal to interference ratios for the far-end signal before and after the SI cancellation. A detailed description of the measurement procedure can be found in [29].

The antenna configuration in the lake experiment is shown in Fig. 6. The distance between the transmitter (Tx) and receiver (Rx) is 7 cm. They are positioned at a depth of 4 m. The lake depth is around 8 m. In the experiment, BPSK signals are transmitted in the frequency bandwidth from 11.5 kHz to 12.5 kHz; a root raised cosine filter with a roll-off factor 0.2 is used for the pulse shaping. At the receiver, the received signal after analog-to-digital conversion (ADC) is down shifted in frequency, low-pass filtered and down sampled to the sampling rate 1 kHz. These samples are applied to the adaptive filter as the desired signal. The same operation is performed on the signal applied to the transmit antenna; the down sampled samples are used as the (regressor) input to the adaptive filter. More details of the processing can be found in [7].

Fig. 7 shows the impulse response measured in this experiment. It contains several strong taps, which are slowly varying in time; these are reflections from stationary parts of the experimental equipment and the lake bottom. The first fast time-varying tap is due to reflection from the time-varying lake surface. Based on our observation, during the experiment, the amplitude of the lake surface waves varies from 5 cm to

TABLE II: Number of updates in the SRLS-L-DCD algorithm

SNR	Fast varying channel						Slow varying channel					
	30 dB			100 dB			30 dB			100 dB		
P	0	1	2	0	1	2	0	1	2	0	1	2
M	35	85	365	25	25	35	35	85	365	25	35	75
N_u	8	16	28	8	100	800	2	4	10	8	80	80
$(P+1)^2 L^2$	100	400	900	100	400	900	100	400	900	100	400	900
MSD (SRLS-L), dB	-31.1	-37.6	-41.0	-35.0	-66.1	-93.9	-34.1	-38.3	-41.8	-54.8	-98.7	-103.7
MSD (SRLS-L-DCD), dB	-30.9	-37.5	-40.4	-34.7	-65.8	-93.1	-33.8	-38.2	-41.8	-54.6	-98.3	-103.4

Fig. 7: Impulse response variation in the FD lake experiment.

Fig. 8: Impulse response variation in the simulation similar to the FD lake experiment.

TABLE III: SIC factor in the lake experiment

Adaptive filter	M	SIC factor, dB
SRLS	120	45.4
SRLS-L, $P = 0$	195	49.1
SRLS-L, $P = 1$	245	51.0
SRLS-L, $P = 2$	415	51.9
SRLS-L, $P = 3$	505	52.3

10 cm. Our analysis shows that the maximum frequency in the spectrum is $f_{\max} \approx 1.5$ Hz for the first surface reflection. Further taps are due to more complicated reflections (bottom-surface, surface-equipment, etc.), which include the surface reflection; thus, they are also fast varying. We set the filter length to $L = 80$.

In the experiment, the signal (self-interference) to noise ratio is around 55 dB. The SIC factor is computed over a 10 s interval after the convergence of the adaptive filter. Table III shows the SIC factor and the optimal value of M providing the best cancellation for each of the adaptive algorithms. With $P = 3$, the SRLS-L algorithm improves the SIC factor by 6.9 dB compared to the classical SRLS algorithm.

TABLE IV: MSD in the simulation similar to the FD lake experiment

Adaptive filter	M	MSD, dB
SRLS	110	-43.4
SRLS-L, $P = 0$	175	-48.5
SRLS-L, $P = 1$	235	-53.4
SRLS-L, $P = 2$	465	-55.7
SRLS-L, $P = 3$	575	-56.7

We now run a simulation with the time-varying impulse response shown in Fig. 8. This impulse response has the same power delay profile as that in the lake experiment (Fig. 7), the first 15 taps are time-invariant and further taps are realizations of random processes with the uniform power spectral density, $f_{\max} = 1.5$ Hz.

Table IV shows the simulation results. By comparing Table IV with Table III, we can see that values of M providing the best SIC and MSD performance are close to each other. The MSD values in the simulation match very well to the SIC values in the experiment.

V. CONCLUSIONS

In this paper, we combine the BEM approach with the SRLS adaptive filter and propose the SRLS-L adaptive filter for identification of time-varying systems (channels). Instead of using the adaptive filter as a predictor, the proposed algorithm interpolates the channel variations using basis functions; more specifically, the proposed adaptive filter exploits the Legendre algebraic polynomials. We have proposed techniques to reduce the complexity of the proposed algorithm using the FFT and DCD algorithms. We have presented analytical results that allow computation of the MSD performance of the proposed algorithm for a given power spectral density of the channel variation. We have shown that the analytical results match well to the simulation results when using odd degrees of Legendre polynomials and match asymptotically for large M in other cases. We have conducted a lake experiment to evaluate the SI

cancellation performance of the proposed adaptive filter within an FD UWA system, and found that a good improvement is achieved against the classical SRLS algorithm.

NOTE

The authors were not aware of the work in [34], [35] when drafting this version of the paper. This paper will be updated to compare our results with the results in [34], [35].

APPENDIX A: MSD APPROXIMATION COMPONENT

Let $h(t)$ be a realization of a zero-mean real-valued stationary random process of continuous time t with a variance $\sigma_h^2 = E\{|h(t)|^2\}$. Let $\{\phi_p(t)\}_{p=0}^P$ be a set of $(P+1)$ basis functions on interval $[-T/2, T/2]$, and $h(t)$ be approximated as:

$$\hat{h}(t) = \sum_{p=0}^P \hat{c}_p \phi_p(t), \quad t \in \left[-\frac{T}{2}, \frac{T}{2}\right], \quad (43)$$

where $\hat{c}_p$ are expansion coefficients given by

$$\hat{c}_p = \frac{1}{E_p} \int_{-T/2}^{T/2} h(t) \phi_p(t) dt, \quad (44)$$

$$E_p = \int_{-T/2}^{T/2} |\phi_p(t)|^2 dt. \quad (45)$$

We will consider the Legendre algebraic polynomials defined in (3) on the interval $[-1, 1]$, so that the basis functions $\{\phi_p(t)\}_{p=0}^P$ are defined as: $\phi_p(t) = L_p(2t/T)$. Note that [30]:

$$\int_{-1}^1 L_p^2(\tau) d\tau = \frac{2}{2p+1}, \quad (46)$$

and therefore $E_p = T/(2p+1)$. Our purpose is to find

$$\begin{aligned} & E\{|\hat{h}(0) - h(0)|^2\} \\ &= E\{|\hat{h}(0)|^2\} - 2E\{h(0)\hat{h}(0)\} + E\{|h(0)|^2\}. \end{aligned} \quad (47)$$

The first term in (47) can be represented as:

$$E\{|\hat{h}(0)|^2\} = E\left\{\left|\sum_{p=0}^P \hat{c}_p \phi_p(0)\right|^2\right\}. \quad (48)$$

From (44), we have

$$\hat{c}_p = \frac{2p+1}{2} \int_{-1}^1 \tilde{h}(\tau) L_p(\tau) d\tau, \quad (49)$$

where the random process $\tilde{h}(t)$ is a time-compressed version of $h(t)$: $\tilde{h}(\tau) = h(T\tau/2)$. Thus, (48) can be written as

$$\begin{aligned} E\{|\hat{h}(0)|^2\} &= E\left\{\sum_{p=0}^P \frac{(2p+1)L_p(0)}{2} \int_{-1}^1 \tilde{h}(u) L_p(u) du\right. \\ &\times \left.\sum_{q=0}^P \frac{(2q+1)L_q(0)}{2} \int_{-1}^1 \tilde{h}^*(v) L_q(v) dv\right\} \\ &= \sum_{p=0}^P \sum_{q=0}^P I_{p,q} \int_{-1}^1 \int_{-1}^1 E\{\tilde{h}(u) \tilde{h}^*(v)\} L_p(u) L_q(v) dudv \\ &= \sum_{p=0}^P \sum_{q=0}^P I_{p,q} \int_{-1}^1 \int_{-1}^1 \rho_{\tilde{h}}(u-v) L_p(u) L_q(v) dudv, \end{aligned} \quad (50)$$

where $\rho_{\tilde{h}}(u-v) = E\{\tilde{h}(u) \tilde{h}^*(v)\}$ is the correlation function of the random process $\tilde{h}(\tau)$ and

$$I_{p,q} = (p+1/2)(q+1/2)L_p(0)L_q(0). \quad (51)$$

We substitute

$$\rho_{\tilde{h}}(u-v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) e^{j\omega(u-v)} d\omega \quad (52)$$

into (50), and arrive at

$$\begin{aligned} E\{|\hat{h}(0)|^2\} &= \sum_{p=0}^P \sum_{q=0}^P \frac{I_{p,q}}{2\pi} \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) \\ &\times \int_{-1}^1 \int_{-1}^1 L_p(u) e^{j\omega u} du L_q(v) e^{-j\omega v} dv d\omega, \end{aligned} \quad (53)$$

where $G_{\tilde{h}}(\omega)$ is the power spectral density of the random process $\tilde{h}(\tau)$. We then obtain:

$$E\{|\hat{h}(0)|^2\} = \frac{1}{2\pi} \sum_{p=0}^P \sum_{q=0}^P I_{p,q} \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) \tilde{L}_p^*(\omega) \tilde{L}_q(\omega) d\omega, \quad (54)$$

where $\tilde{L}_p(\omega) = \int_{-1}^1 L_p(\tau) e^{-j\omega\tau} d\tau$ is the Fourier transform of the Legendre polynomial given by [33]:

$$\tilde{L}_p(\omega) = \sum_{n=0}^p \frac{(p+n)!}{2^n n! (p-n)!} \frac{1}{(j\omega)^{n+1}} [(-1)^{p+n} e^{j\omega} - e^{-j\omega}]. \quad (55)$$

The second term in (47) can be represented as:

$$\begin{aligned} E\{h(0)\hat{h}(0)\} &= E\left\{h(0) \sum_{p=0}^P c_p \phi_p(0)\right\} \\ &= E\left\{\tilde{h}(0) \sum_{p=0}^P \frac{(2p+1)L_p(0)}{2} \int_{-1}^1 \tilde{h}(u) L_p(u) du\right\} \\ &= \sum_{p=0}^P \frac{(2p+1)L_p(0)}{2} \int_{-1}^1 \rho_{\tilde{h}}(u) L_p(u) du \\ &= \frac{1}{4\pi} \sum_{p=0}^P (2p+1)L_p(0) \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) \int_{-1}^1 L_p(u) e^{j\omega u} du d\omega \\ &= \frac{1}{4\pi} \sum_{p=0}^P (2p+1)L_p(0) \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) \tilde{L}_p^*(\omega) d\omega. \end{aligned} \quad (56)$$

The third term in (47) is the variance of $h(t)$ given by:

$$\sigma_h^2 = \sigma_{\tilde{h}}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\tilde{h}}(\omega) d\omega. \quad (57)$$

Thus, for a given power spectral density $G_h(\omega) = (T/2)G_{\tilde{h}}(T\omega/2)$, defining how fast $h(t)$ varies in time, the MSD approximation component can be found using (54), (56) and (57) as shown in (35), where the integrals can be computed numerically.

REFERENCES

- [1] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, 2009.
- [2] M. Stojanovic, J. A. Catipovic, and J. G. Proakis, "Phase-coherent digital communications for underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 1, pp. 100–111, 1994.
- [3] H. Yeo, B. Sharif, O. Hinton, and A. Adams, "Improved RLS algorithm for time-variant underwater acoustic communications," *Electronics Letters*, vol. 36, no. 2, pp. 191–192, 2000.
- [4] Y. Zhang, Y. V. Zakharov, and J. Li, "Soft-decision-driven sparse channel estimation and turbo equalization for MIMO underwater acoustic communications," *IEEE Access*, vol. 6, pp. 4955–4973, 2018.
- [5] Y. V. Zakharov and J. Li, "Sliding window adaptive filter with diagonal loading for estimation of sparse UWA channels," in *IEEE OCEANS, Shanghai*, 2016, pp. 1–5.
- [6] G. Qiao, S. Gan, S. Liu, L. Ma, and Z. Sun, "Digital self-interference cancellation for asynchronous in-band full-duplex underwater acoustic communication," *Sensors*, vol. 18, no. 6, pp. 1700–1716, 2018.
- [7] L. Shen, B. Henson, Y. Zakharov, and P. Mitchell, "Digital self-interference cancellation for underwater acoustic systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 1, pp. 192–196, 2020.
- [8] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, 2014.
- [9] D. Bharadia, E. McMillin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, 2013, pp. 375–386.
- [10] C. T. Tindle, G. B. Deane, and J. C. Preisig, "Reflection of underwater sound from surface waves," *The Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 66–72, 2009.
- [11] M. Niedzwiecki, "Functional series modeling approach to identification of nonstationary stochastic systems," *IEEE Transactions on Automatic Control*, vol. 33, no. 10, pp. 955–961, 1988.
- [12] G. W. Wornell, "A Karhunen-Loeve-like expansion for $1/f$ processes via wavelets," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 859–861, 1990.
- [13] M. K. Tsatsanis and G. B. Giannakis, "Time-varying system identification and model validation using wavelets," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3512–3523, 1993.
- [14] M. Visintin, "Karhunen-Loeve expansion of a fast Rayleigh fading process," *Electronics Letters*, vol. 32, no. 18, pp. 1712–1713, 1996.
- [15] M. K. Tsatsanis and G. B. Giannakis, "Modelling and equalization of rapidly fading channels," *Int. J. Adaptive Contr. Signal Processing*, vol. 10, no. 2–3, pp. 159–176, 1996.
- [16] Z. Tang and G. Leus, "Time-multiplexed training for time-selective channels," *IEEE Signal Processing Letters*, vol. 14, no. 9, pp. 585–588, 2007.
- [17] G. Leus, "On the estimation of rapidly time-varying channels," *Proc. EUSIPCO'04, Vienna, Austria*, pp. 2227–2230, Sep. 2004.
- [18] Y. V. Zakharov, T. C. Tozer, and J. F. Adlard, "Polynomial spline-approximation of Clarke's model," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1198–1208, 2004.
- [19] Y. V. Zakharov, V. M. Baronkin, and J. Zhang, "Optimal and mismatched detection of QAM signals in fast fading channels with imperfect channel estimation," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 617–621, 2009.
- [20] H. Mai, Y. V. Zakharov, and A. G. Burr, "Iterative channel estimation based on B-splines for fast flat fading channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1224–1229, Apr. 2007.
- [21] Y. V. Zakharov and T. C. Tozer, "Local spline approximation of time-varying channel model," *Electronics Letters*, vol. 37, no. 23, pp. 1408–1409, 2001.
- [22] T. Zemen and C. F. Mecklenbrauker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3597–3607, 2005.
- [23] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty, V: The discrete case," *Bell Syst. Tech. J.*, vol. 43, no. 6, pp. 3009–3058, 1964.
- [24] P. S. Rossi and R. R. Muller, "Slepian-based two-dimensional estimation of time-frequency variant MIMO-OFDM channels," *IEEE Signal Processing Letters*, vol. 15, pp. 21–24, 2008.
- [25] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 862–873, 1999.
- [26] C. Shin, J. G. Andrews, and E. J. Powers, "An efficient design of doubly selective channel estimation for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 10, pp. 3790–3802, 2007.
- [27] T. Hrycak, S. Das, and G. Matz, "Inverse methods for reconstruction of channel taps in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2666–2671, 2012.
- [28] M. Niedzwiecki and T. Klaput, "Fast recursive basis function estimators for identification of time-varying processes," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1925–1934, 2002.
- [29] L. Shen, B. Henson, Y. Zakharov, and P. Mitchell, "Adaptive filtering for fast time-varying full-duplex underwater acoustic systems," *TechRxiv*. Preprint. <https://doi.org/10.36227/techrxiv.11967813.v1>.
- [30] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic press, 2014.
- [31] Y. V. Zakharov, G. P. White, and J. Liu, "Low-complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3150–3161, 2008.
- [32] J. Liu, Y. V. Zakharov, and B. Weaver, "Architecture and FPGA design of dichotomous coordinate descent algorithms," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 11, pp. 2425–2438, 2009.
- [33] A. S. Fokas, A. Iserles, and S. A. Smitheman, "The unified method in polygonal domains via the explicit Fourier transform of Legendre polynomials," *Unified Transforms (eds. AS Fokas and B. Pelloni)*, (SIAM, Philadelphia, PA, 2014), 2014.
- [34] M. Niedzwiecki and M. Ciolek, "Generalized Savitzky–Golay filters for identification of nonstationary systems," *Automatica*, vol. 108, p. 108477, 2019.
- [35] M. J. Niedzwiecki, M. Ciolek, and A. Gańcza, "A new look at the statistical identification of nonstationary systems," *Automatica*, vol. 118, p. 109037, 2020.