Adaptive Filtering for Full-Duplex UWA Systems with Time-Varying Self-Interference Channel

Lu Shen [1], Yuriy Zakharov [2], Benjamin Henson [2], Nils Morozs [2], and Paul Mitchell [2]

[1]University of York
[2]Affiliation not available

October 30, 2023

Abstract

Abstract:

To enable full-duplex (FD) in underwater acoustic (UWA) systems, a high level of self-interference (SI) cancellation (SIC) is required. For digital SIC, adaptive filters are used. In time-invariant channels, the SI can be effectively cancelled by classical recursive least-square (RLS) adaptive filters, such as the sliding-window RLS (SRLS) or exponential-window RLS, but their SIC performance degrades in time-varying channels, e.g., in channels with a moving sea surface. Their performance can be improved by delaying the filter inputs. This delay, however, makes the mean squared error (MSE) unsuitable for measuring the SIC performance. In this paper, we propose a new evaluation metric, the SIC factor (SICF), which gives better indication of the SIC performance compared to MSE. The SICF can be used in experiments and in real FD systems. A new SRLS adaptive filter based on parabolic approximation of the channel variation in time, named SRLS-P, is also proposed. The SIC performance of the SRLS-P adaptive filter and classical RLS algorithms (with and without the delay) is evaluated by simulation and in lake experiments. The results show that the SRLS-P adaptive filter significantly improves the SIC performance, compared to the classical RLS adaptive filters.

Adaptive Filtering for Full-Duplex UWA Systems with Time-Varying Self-Interference Channel

Lu Shen, *Student Member, IEEE*, Yuriy Zakharov, *Senior Member, IEEE*,
Benjamin Henson, *Member, IEEE*, Nils Morozs, *Member, IEEE*, Paul Mitchell, *Senior Member, IEEE*

***Abstract*—To enable full-duplex (FD) in underwater acoustic (UWA) systems, a high level of self-interference (SI) cancellation (SIC) is required. For digital SIC, adaptive filters are used. In time-invariant channels, the SI can be effectively cancelled by classical recursive least-square (RLS) adaptive filters, such as the sliding-window RLS (SRLS) or exponential-window RLS, but their SIC performance degrades in time-varying channels, e.g., in channels with a moving sea surface. Their performance can be improved by delaying the filter inputs. This delay, however, makes the mean squared error (MSE) unsuitable for measuring the SIC performance. In this paper, we propose a new evaluation metric, the SIC factor (SICF), which gives better indication of the SIC performance compared to MSE. The SICF can be used in experiments and in real FD systems. A new SRLS adaptive filter based on parabolic approximation of the channel variation in time, named SRLS-P, is also proposed. The SIC performance of the SRLS-P adaptive filter and classical RLS algorithms (with and without the delay) is evaluated by simulation and in lake experiments. The results show that the SRLS-P adaptive filter significantly improves the SIC performance, compared to the classical RLS adaptive filters.**

***Index Terms*—Adaptive filter, full-duplex, self-interference cancellation, time-varying channel estimation, underwater acoustic communications**

I. Introduction

In recent years, full-duplex (FD) operation of terrestrial radio systems, such as communication systems, has demonstrated an ability to significantly increase the system throughput [1]–[5]. If FD operation can be adopted in underwater acoustic (UWA) systems, e.g., in UWA communication systems, the capacity of the acoustic links can be almost doubled. Active sonar systems can also benefit from the FD operation by expanding the signal family used for transmission. Despite the benefits of FD, it is not widely considered for UWA systems mainly due to the severe self-interference (SI) introduced by the near-end transmission. Various active SI cancellation (SIC) techniques have been proposed for FD terrestrial radio systems. Normally, a certain amount of SI is cancelled in the analogue domain before digital cancellation to avoid the saturation in the analogue-to-digital converter (ADC) [1], [6], [7]. For FD UWA systems, due to the lower frequencies of acoustic signals, high resolution ADCs are available. Thus, digital cancellation can be considered as the main practical approach for SI cancellation in FD UWA systems [8], [9].

One of the major limitations of the digital cancellation performance is due to the hardware imperfection in the transmit and receive chains, among which the non-linearity introduced by the power amplifier (PA) is the dominant factor [10]. A general approach to deal with the PA non-linearity is to estimate the non-linear distortion, e.g. using the Hammerstein model and its extensions, and then compensate it in the received signal [5], [8], [11]. To accurately model the non-linearity, high order basis functions are required. The disadvantages of this approach are the high complexity of the non-linear model and a large number of parameters to be estimated. Another approach is to use the PA output as the reference signal for SIC [9], [12]. In this case, lower complexity linear adaptive filters can be used for the SIC. In [9], we show that, with the use of the PA output as the reference signal, a high level of SIC can be achieved in slow-varying UWA SI channels by using classical recursive least-square (RLS) adaptive filters. The general block diagram of the FD system with digital cancellation using PA output as the reference signal is shown in Fig. 1.

Adaptive filters operate efficiently if the power spectral density of the input signal (regressor) does not have zeros, i.e. the regressor correlation matrix is full rank. This, however, requires sampling the baseband signal at a (symbol) rate, which is lower than the Nyquist frequency. As a result, the performance of the adaptive filter is sensitive to the delay between the regressor (PA output) and the desired (hydrophone) signal. To overcome this problem and ensure robust SIC performance regardless of the delay, the digital cancellation scheme from [9] is extended in [13]; the block diagram of the scheme is shown in Fig. 2. In this scheme, two branches are used with symbol rate sampling in each branch, with odd and even samples, respectively, taken from a twice oversampled baseband signal at the PA output. In this paper, we investigate the SIC performance of this scheme using different adaptive filters.

As observed in our lake experiments, another phenomenon that limits the SI cancellation is the time-varying surface reflections [14]. While a high level of SI cancellation can be achieved for time-invariant SI channels (e.g., in a water tank [9]) using classical RLS adaptive filters, the cancellation performance is limited in experiments with a moving surface. The main limitation is the tracking ability of the classical adaptive filters. The Kalman filter is considered as a good candidate for estimation of time-varying channels [15], [16]. However, for using the Kalman filter, the channel statistics should be known, which is often not the case in practice. To improve SIC performance in fast time-varying channels, other schemes are required.

In time-varying channels, the SIC performance can be significantly improved if the input signals are delayed with respect to the time-varying estimate of the channel response as

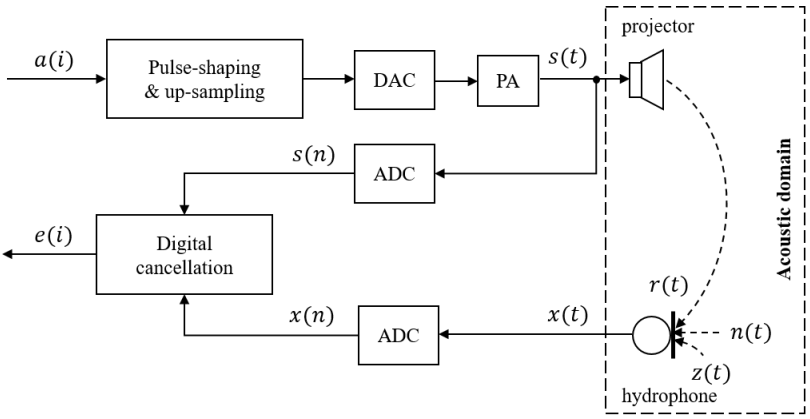

Fig. 1. Block diagram of the FD UWA system with digital cancellation. The system works at two sampling frequencies. The index of the signal sample with the high (passband) sampling rate is denoted by n, and the low (baseband) sampling rate sample index is i. The analogue (passband) signals are: the PA output $s(t)$; the SI $r(t)$; the noise $n(t)$; the far-end signal $z(t)$; and the received (hydrophone) signal $x(t)$. The digital (passband) signals are: the PA output $s(n)$ and the received signal $x(n)$. The digital baseband signals are: the transmitted data symbols $a(i)$ and the residual signal after the digital cancellation $e(i)$. DAC is the digital-to-analogue converter. See more details in [9].

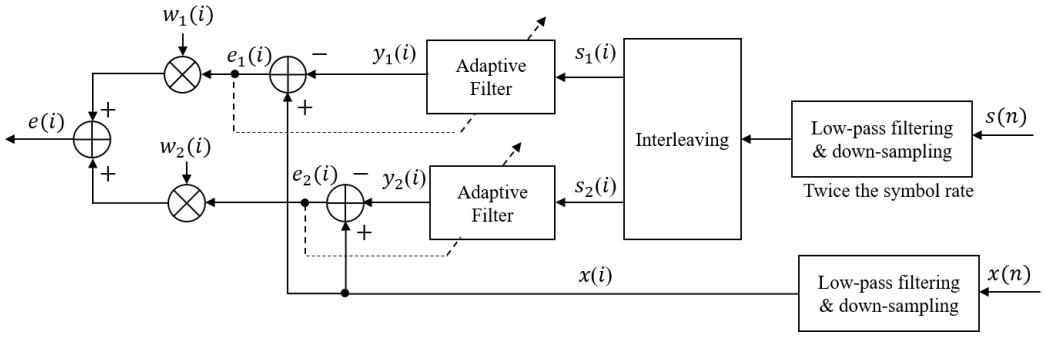

Fig. 2. Block diagram of the digital cancellation scheme. The PA output $s(n)$ is down-sampled to twice the symbol rate and interleaved into two branches, $s_1(i)$ contains odd samples and $s_2(i)$ contains even samples; $x(i)$ is the baseband received signal; $e_1(i)$ and $e_2(i)$ represent residual signals in the two branches; $w_1(i)$ and $w_2(i)$ are weights computed as suggested in [13].

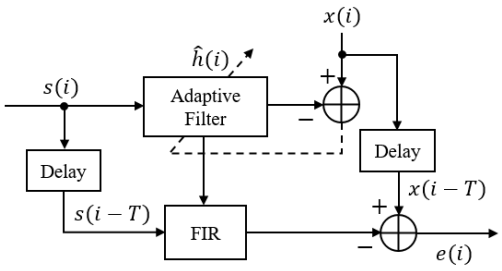

Fig. 3. Adaptive filter with a delay.

shown in Fig. 3. However, to our knowledge, this opportunity for FD systems has not been investigated yet.

Introducing a delay between the channel estimate and the inputs to the adaptive filter results in a problem in measuring the cancellation performance. The residual SI power is normally used to characterise the FD system performance [5], [17], [18]. The mean squared error (MSE) is used to measure the residual SI power [19]. However, the MSE in an adaptive filter with a delay is unsuitable for this purpose, since, in this case, unlike the classical RLS algorithms, the same data is used for channel estimation and computation of the MSE, resulting in over-fitting. Therefore, another measure of SIC performance is required when using adaptive filters with a delay.

In this paper, we propose and investigate the SIC factor (SICF) for measuring the cancellation performance and a new adaptive algorithm for FD UWA systems with time-varying SI channels. The contributions of this paper are as follows.

1) The SICF is proposed for evaluation of the SIC performance in digital SI cancellers.
2) The dependence between the delay of the input signals and the SIC performance for the exponential window RLS (ERLS) and sliding window RLS (SRLS) adaptive filters is investigated.
3) The new adaptive filter (SRLS-P) is proposed, which is derived based on parabolic approximation of the channel variation in time.
4) The proposed algorithm is investigated using numerical simulations and lake experiments, and its performance is compared with that of the classical RLS adaptive algorithms.

The rest of the paper is organized as follows. In Section II,

the new evaluation metric SICF is described. In Section III, the SRLS-P adaptive filter is derived. Section IV and Section V present simulation result in baseband and passband scenarios, respectively. Section VI compares the SIC performance provided by the adaptive filters using experimental data. In Section VII, we draw the conclusions.

Notations: In this paper, we use capital and small bold fonts for matrices and vectors, respectively; e.g, $\mathbf{R}$ and $\mathbf{h}$. We also denote the transpose of $\mathbf{x}$ as $\mathbf{x}^T$, and the Hermitian transpose of $\mathbf{h}$ as $\mathbf{h}^H$.

II. Evaluation of SIC performance

Consider the SIC scheme shown in Fig. 3. In this scheme, $x(i)$ is a baseband version of the signal received by the hydrophone, and it is modelled as:

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

where $\mathbf{h}(i)$ is the baseband SI channel response at time instant i, $s(i)$ is the baseband version of the PA output signal, $\mathbf{s}(i) = [s(i), \ldots, s(i-L+1)]^T$, and L is the channel length. The signal $z(i)$ contains the far-end signal, as well as noise signals such as the ambient noise, ADC noise, etc. In terms of an adaptive filter operated in the identification mode, $\mathbf{s}(i)$ is the regressor and $x(i)$ is the desired signal [15], [16]. Using these signals, the adaptive filter produces an estimate $\hat{\mathbf{h}}(i+T)$ of $\mathbf{h}(i)$. Note that, in classical adaptive filters, $T = 0$ and it is assumed that the estimate $\hat{\mathbf{h}}(i)$ is obtained using the regressor and desired signal up to time instant $i-1$. Based on this channel estimate, the SI is cancelled by recovering the SI signal as $\hat{\mathbf{h}}^H(i)\mathbf{s}(i-T)$ and subtracting it from the received signal:

$$e(i) = x(i-T) - \hat{\mathbf{h}}^H(i)\mathbf{s}(i-T). \tag{2}$$

The performance of an adaptive filter is most often evaluated using the mean squared error (MSE) and mean squared deviation (MSD) [15], [16]. The MSE is defined as:

$$\mathrm{MSE}(i) = \mathrm{E}\{|e(i)|^2\}, \tag{3}$$

where $\mathrm{E}\{\cdot\}$ denotes the expectation. The MSD is defined as:

$$\mathrm{MSD}(i) = \mathrm{E}\{\|\mathbf{h}(i) - \hat{\mathbf{h}}(i+T)\|_2^2\}. \tag{4}$$

For a classical adaptive filter (with $T = 0$), the SIC performance is normally evaluated by computing the MSE. However, by adjusting parameters of an adaptive filter with a delay, it is possible to make the MSE even lower than the 'noise-plus-far-end-signal' floor, although this does not mean that the SIC performance is good. It means that not only the SI is cancelled, but also a part of the far-end signal (i.e., the signal of interest) is also cancelled. Essentially, the adaptive filter is over-fitted, since, due to the delay, the same data is used for training the adaptive filter and for the MSE computation. In these scenarios, the MSE becomes an unreliable metric for assessment of the SIC performance.

Using (1) and (2), the residual signal $e(i)$ can be represented as:

$$e(i) = \varepsilon(i) + z(i-T), \tag{5}$$

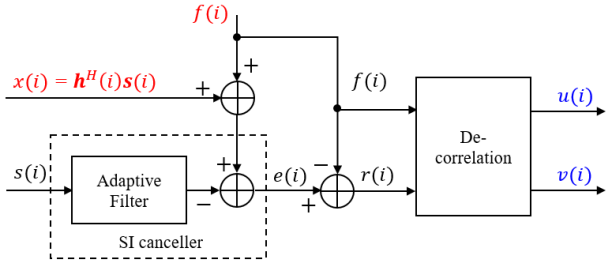

Fig. 4. Block diagram of FD system with SI cancellation.

where $\varepsilon(i) = [\mathbf{h}(i-T) - \hat{\mathbf{h}}(i)]^H\mathbf{s}(i-T)$. Then the SIC performance can be evaluated by computing the output SNR of the FD system:

$$\mathrm{SNR}_{\mathrm{out}}(i) = \frac{\sigma_z^2}{\mathrm{E}\{|\varepsilon(i)|^2\}}, \tag{6}$$

where $z(i)$ is treated as the signal of interest and $\varepsilon(i)$ is an interference. Assuming that $s(i)$ are uncorrelated for different i and uncorrelated to $\hat{\mathbf{h}}(i)$, we have:

$$\begin{aligned} \mathrm{E}\{|\varepsilon(i)|^2\} &= \sigma_s^2 \mathrm{E}\{||\mathbf{h}(i-T) - \hat{\mathbf{h}}(i)||_2^2\} &(7)\\ &= \sigma_s^2 \mathrm{MSD}(i-T), &(8)\end{aligned}$$

where $\sigma_s^2 = \mathrm{E}\{|s(i)|^2\}$ is the variance of the signal $s(i)$, which is assumed stationary. Then, finally, we obtain:

$$\mathrm{SNR}_{\mathrm{out}}(i) = \frac{\sigma_z^2}{\sigma_s^2} \cdot \frac{1}{\mathrm{MSD}(i-T)}. \tag{9}$$

Thus, the MSD is a useful characteristic of an adaptive filter operating within an SI canceller. It shows how much the ratio between powers of the far-end and near-end signals improves due to the accuracy of channel estimation. It can also be seen that the adjustment of parameters of an adaptive filter to minimize the MSD will also result in the maximum output SNR. However, the MSD computation requires knowledge of the true channel response $\mathbf{h}(i)$, which is unavailable in most practical scenarios. Another important issue is that (9) is only applicable if $\hat{\mathbf{h}}(i)$ and $s(i)$ are uncorrelated, which might not be the case for adaptive filters with a delay.

We now propose the SICF, which provides a good indication of the cancellation performance. It does not require the knowledge of the true channel response, and can be used in practice for adaptive filters with and without the delay.

The following scenario is considered. There is only the near-end transmission, the noise is ignored, and a known 'far-end' signal is artificially added to the received signal. Fig. 4 illustrates our description below. We assume that the SI signal $x(i) = \mathbf{h}^H(i)\mathbf{s}(i)$ is noise-free. We add to the SI signal a known signal $f(i)$ assumed to be a far-end signal. The level of the signal $\sigma_f^2 = \mathrm{E}\{|f(i)|^2\}$ is chosen to guarantee a predefined input SNR:

$$\mathrm{SNR}_{\mathrm{in}}(i) = \frac{\sigma_f^2}{\mathrm{E}\{|x(i)|^2\}}. \tag{10}$$

The SI canceller (shown in Fig. 4) subtracts the SI estimate produced by the adaptive filter from the received signal $x(i)+$

$f(i)$. The canceller output $e(i)$ contains the signal of interest $f(i)$ and a residual signal $r(i)$:

$$e(i) = f(i) + r(i), \tag{11}$$

and since both signals $e(i)$ and $f(i)$ are available after the cancellation, the residual signal $r(i)$ can be computed as $r(i) = e(i) - f(i)$.

Here we measure the SIC performance as a factor of improvement in the signal-to-noise ratio due to the SI cancellation and compute the SICF as:

$$\mathrm{SICF}(i) = \frac{\mathrm{SNR}_{\mathrm{out}}(i)}{\mathrm{SNR}_{\mathrm{in}}(i)}. \tag{12}$$

The purpose of a SI canceller is to 'clean' the input signal for further application of a signal processing algorithm, e.g. a communication detector. Such algorithms are typically optimized under the assumption that the signal of interest and the noise (the residual signal $r(i)$, in our case) are uncorrelated. If the signal of interest $f(i)$ and the residual $r(i)$ are uncorrelated, the output SNR, $\mathrm{SNR}_{\mathrm{out}}(i)$, can be computed as a ratio of their variances. However, due to the over-fitting in the adaptive filter, in general, these two signals are correlated.

We now assume that the signal of interest $f(i)$ is attenuated due to the imperfection of the adaptive filter. More specifically, we can rewrite (11) as:

$$\begin{aligned} e(i) &= \alpha f(i) + [(1-\alpha)f(i) + r(i)] && (13)\\ &= u(i) + v(i), && (14) \end{aligned}$$

where the modified signal of interest $u(i) = \alpha f(i)$ and noise component $v(i) = (1-\alpha)f(i) + r(i)$ are uncorrelated.

We now find the coefficient α that zeroes the correlation between $u(i)$ and $v(i)$:

$$\mathrm{E}\{u(i)v^*(i)\} = \mathrm{E}\{\alpha f(i)[(1-\alpha)f(i) + r(i)]^*\} = 0. \tag{15}$$

From (15), we find α as:

$$\alpha = 1 + \frac{1}{\sigma_f^2}\mathrm{E}\{f^*(i)r(i)\}. \tag{16}$$

After finding α, the modified signal of interest $u(i)$ and noise $v(i)$ can be computed from (13).

In experiments, the mathematical expectation is replaced by the average over a time interval after convergence of the adaptive filter. The output SNR can be computed as:

$$\mathrm{SNR}_{\mathrm{out}} = \frac{\|\mathbf{u}\|_2^2}{\|\mathbf{v}\|_2^2}, \tag{17}$$

where $\mathbf{u} = [u(0), \ldots, u(P-1)]^T$ is a $P \times 1$ vector of the signal of interest, $\mathbf{v} = [v(0), \ldots, v(P-1)]^T$, and P is the averaging interval. The average interval P is preferred to be longer than the time correlation of the SI channel.

III. Proposed SRLS-P adaptive filter

In this section, we review the ERLS and SRLS adaptive filters, consider their delayed versions, and propose a new adaptive filter based on the SRLS algorithm and parabolic approximation of channel variation in time; we call it the SRLS-P adaptive filter.

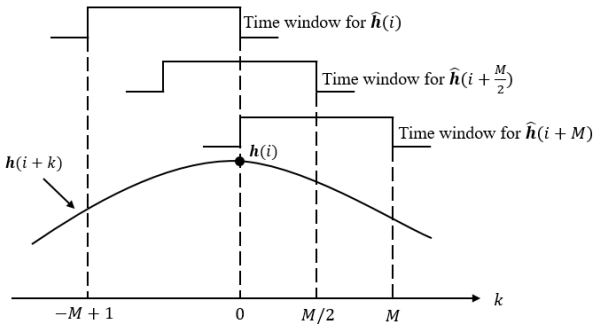

Fig. 5. Time-varying channel and time windows of the SRLS algorithm.

A. Classical ERLS and SRLS adaptive filters

At every time instant i, an RLS adaptive filter updates the solution vector $\hat{\mathbf{h}}(i)$ according to the normal equation:

$$\mathbf{R}(i)\hat{\mathbf{h}}(i) = \boldsymbol{\beta}(i), \tag{18}$$

where $\mathbf{R}(i)$ is an $L \times L$ autocorrelation matrix, $\boldsymbol{\beta}(i)$ is an $L \times 1$ cross-correlation vector, and L is the filter length. The autocorrelation matrix and cross-correlation vector are approximated by averaging in time.

For the classical ERLS adaptive filter, $\mathbf{R}(i)$ and $\boldsymbol{\beta}(i)$ can be updated as:

$$\mathbf{R}(i) = (\lambda-1)\mathbf{R}(i-1) + \mathbf{s}(i)\mathbf{s}^H(i), \tag{19}$$

$$\boldsymbol{\beta}(i) = (\lambda-1)\boldsymbol{\beta}(i-1) + x^*(i)\mathbf{s}(i), \tag{20}$$

where λ is the forgetting factor, $\mathbf{s}(i) = [s(i), s(i-1), \ldots, s(i-L+1)]^T$ is the regressor at the ith time instant, and $x(i)$ is the ith sample of the desired signal. The weights of the time average window is the exponential $e^{|i-p|\lambda}$, $p \le i$.

For the classical SRLS adaptive filter, the update of $\mathbf{R}(i)$ and $\boldsymbol{\beta}(i)$ can be written as [20], [21]:

$$\mathbf{R}(i) = \mathbf{R}(i-1) + \mathbf{s}(i)\mathbf{s}^H(i) - \mathbf{s}(i-M)\mathbf{s}^H(i-M), \tag{21}$$

$$\boldsymbol{\beta}(i) = \boldsymbol{\beta}(i-1) + x^*(i)\mathbf{s}(i) - x^*(i-M)\mathbf{s}(i-M), \tag{22}$$

where M is the sliding window length. The time average window is a constant over the time interval $[i-M+1,\ i]$, and zero otherwise. Fig. 5 shows the position of the time window (top) in the SRLS algorithm with respect to the time varying channel response $\mathbf{h}(i)$.

B. Delayed ERLS and SRLS adaptive filters

Since $\mathbf{R}(i)$ and $\boldsymbol{\beta}(i)$ are obtained by averaging in time, the current channel estimate $\hat{\mathbf{h}}(i)$ can be seen as an average of the true channel response over past time instants. If the SI channel is time-invariant, $\hat{\mathbf{h}}(i)$ can be an accurate estimate of $\mathbf{h}(i)$. However, for a time-varying channel, $\hat{\mathbf{h}}(i)$ is not an accurate estimate of $\mathbf{h}(i)$.

For the SRLS adaptive filter, the channel estimate $\hat{\mathbf{h}}(i)$ can be seen as an average of $\mathbf{h}(i)$ over the past M time instants. If we assume that the channel response varies linearly in the vicinity of i, then its average over the rectangular window centred at i (middle time window in Fig. 5) is equal to $\mathbf{h}(i)$.

In such a case, $\hat{\mathbf{h}}(i+M/2)$ is a more accurate estimate of $\mathbf{h}(i)$ than $\hat{\mathbf{h}}(i)$. Therefore, using the delay $T = M/2$ in the scheme shown in Fig. 3 should provide an improvement in the SIC performance compared to the case $T = 0$. In Section IV, we demonstrate that this is indeed the case. For the ERLS adaptive filter, the time window is infinite in length, and it is more difficult to determine the optimal delay which provides the highest level of cancellation. Moreover, in Section IV, we also show that even for the same λ, different channel realisations require different T. Therefore, our proposed adaptive filter is based on the SRLS algorithm, for which the optimal delay is well defined. We call the ERLS and SRLS algorithms with delays as ERLSd and SRLSd, respectively, to distinguish them from the classical RLS algorithms.

C. SRLS-P adaptive filter

Compared to the SRLS algorithm, the SRLSd adaptive filter improves the MSD performance, and, as a result, it improves the SI cancellation performance by applying the current channel estimate found at the ith time instant to the delayed regressor $\mathbf{s}(i-M/2)$, corresponding to the middle of the averaging time window of length M. It changes the way the SI signal is reconstructed, but the channel estimates are computed in the same way as in the classical SRLS adaptive filter.

In fast time-varying channels, the channel estimation performance provided by the SRLSd algorithm is still limited, since the channel estimate can be viewed as simply an average of the true channel response over the past M time instants. To improve the tracking ability in fast time-varying channels, we propose the SRLS-P adaptive filter. The key idea of the algorithm is the parabolic interpolation of the channel time variation using the estimates $\hat{\mathbf{h}}(i)$ provided by the SRLS algorithm.

We assume that the time-varying channel response is a second-order algebraic polynomial within a short time interval around the time instant i, as shown in Fig. 5:

$$\mathbf{h}(i+k) = \mathbf{h}_0(i) + \mathbf{h}_1(i)k + \mathbf{h}_2(i)k^2, \tag{23}$$

where $k = -M+1, \ldots, M$, and $\mathbf{h}_0(i)$, $\mathbf{h}_1(i)$ and $\mathbf{h}_2(i)$ are three $L \times 1$ vectors to be estimated. From (23), it can be seen that $\mathbf{h}(i) = \mathbf{h}_0(i)$, and thus an estimate of $\mathbf{h}_0(i)$ can be used as an estimate of the channel response $\mathbf{h}(i)$ at time instant i.

The channel estimate $\hat{\mathbf{h}}(i+k)$ computed by the SRLS algorithm in scenarios without noise can be expressed as (see Appendix):

$$\hat{\mathbf{h}}(i+k) = \frac{1}{M}\mathbf{R}^{-1}(i+k) \sum_{m=-M+k+1}^{k} \mathbf{R}_{i+m}\mathbf{h}(i+m), \tag{24}$$

where $\mathbf{R}(i) = \mathbf{S}^H(i)\mathbf{S}(i)$ is the $L \times L$ auto-correlation matrix of the regressor, $\mathbf{S}(i) = [\mathbf{s}(i), \mathbf{s}(i-1), \ldots, \mathbf{s}(i-M+1)]^T$ is an $M \times L$ observation matrix, $\mathbf{s}(i)$ is the regressor at the ith time instant and $\mathbf{R}_{i+m} = \mathbf{s}(i+m)\mathbf{s}^H(i+m)$.

By substituting (23) into (24) for $k = 0$, $k = M/2$, and $k = M$, we obtain a system of equations with respect to the unknown $3L \times 1$ vector $\mathbf{z} = [\mathbf{h}_0(i); \mathbf{h}_1(i); \mathbf{h}_2(i)]$. By solving the system, we obtain an estimate $\hat{\mathbf{h}}_0(i)$ of $\mathbf{h}_0(i)$, which is also the new channel estimate $\tilde{\mathbf{h}}(i)$ of $\mathbf{h}(i)$.

More specifically, we have:

$$\begin{aligned}\hat{\mathbf{h}}(i) &= \frac{1}{M}\mathbf{R}^{-1}(i) \\ &\times \sum_{m=-M+1}^{0} \mathbf{R}_{i+m}[\mathbf{h}_0(i) + m\mathbf{h}_1(i) + m^2\mathbf{h}_2(i)] \\ &= \mathbf{h}_0(i) + \mathbf{A}_1\mathbf{h}_1(i) + \mathbf{A}_2\mathbf{h}_2(i),\end{aligned} \tag{25}$$

where

$$\mathbf{A}_1 = \mathbf{R}^{-1}(i) \sum_{m=-M+1}^{0} m\mathbf{R}_{i+m}, \tag{26}$$

$$\mathbf{A}_2 = \mathbf{R}^{-1}(i) \sum_{m=-M+1}^{0} m^2\mathbf{R}_{i+m}. \tag{27}$$

Similarly, we obtain:

$$\hat{\mathbf{h}}(i+M/2) = \mathbf{h}_0(i) + \mathbf{B}_1\mathbf{h}_1(i) + \mathbf{B}_2\mathbf{h}_2(i), \tag{28}$$

$$\hat{\mathbf{h}}(i+M) = \mathbf{h}_0(i) + \mathbf{C}_1\mathbf{h}_1(i) + \mathbf{C}_2\mathbf{h}_2(i), \tag{29}$$

where

$$\mathbf{B}_1 = \mathbf{R}^{-1}(i+M/2) \sum_{m=-M/2+1}^{M/2} m\mathbf{R}_{i+m}, \tag{30}$$

$$\mathbf{B}_2 = \mathbf{R}^{-1}(i+M/2) \sum_{m=-M/2+1}^{M/2} m^2\mathbf{R}_{i+m}, \tag{31}$$

and

$$\mathbf{C}_1 = \mathbf{R}^{-1}(i+M) \sum_{m=1}^{M} m\mathbf{R}_{i+m}, \tag{32}$$

$$\mathbf{C}_2 = \mathbf{R}^{-1}(i+M) \sum_{m=1}^{M} m^2\mathbf{R}_{i+m}. \tag{33}$$

We now arrive at the system of equations:

$$\begin{cases} \mathbf{h}_0(i) + \mathbf{A}_1\mathbf{h}_1(i) + \mathbf{A}_2\mathbf{h}_2(i) = \hat{\mathbf{h}}(i), & (34) \\ \mathbf{h}_0(i) + \mathbf{B}_1\mathbf{h}_1(i) + \mathbf{B}_2\mathbf{h}_2(i) = \hat{\mathbf{h}}(i+M/2), & (35) \\ \mathbf{h}_0(i) + \mathbf{C}_1\mathbf{h}_1(i) + \mathbf{C}_2\mathbf{h}_2(i) = \hat{\mathbf{h}}(i+M)\ , & (36) \end{cases}$$

or, in a compact form,

$$\mathbf{D}\mathbf{z} = \hat{\mathbf{h}}, \tag{37}$$

where $\hat{\mathbf{h}} = [\hat{\mathbf{h}}(i);\ \hat{\mathbf{h}}(i+M/2);\ \hat{\mathbf{h}}(i+M)]$ and

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_L & \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{I}_L & \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{I}_L & \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix}. \tag{38}$$

After solving the system in (37), the estimate of the impulse response is found as the first L elements in the vector $\mathbf{z}$:

$$\tilde{\mathbf{h}}(i) = \hat{\mathbf{h}}_0(i) = [\mathbf{z}]_{1,\ldots,L}. \tag{39}$$

The SRLS-P adaptive algorithm is summarized in Algorithm 1, where ε is a regularization parameter, $\mathbf{s}$ is an $N \times 1$ vector of the transmitted signal, N is the number of samples in the transmitted signal, $\mathbf{x}$ is an $N \times 1$ vector of

Algorithm 1: SRLS-P algorithm

Input: $\mathbf{s}$, $\mathbf{x}$
Output: $\tilde{\mathbf{h}}$
Initialization: $\hat{\mathbf{h}}(0) = \mathbf{0}$
for every sample i **do**
  $y(i) = \hat{\mathbf{h}}^H(i-1)\mathbf{s}(i)$
  $e(i) = x(i) - y(i)$
  $\mathbf{R}(i) = \mathbf{S}^H(i)\mathbf{S}(i) + \varepsilon\mathbf{I}_L$
  $\boldsymbol{\beta}(i) = \mathbf{S}^H(i)\mathbf{x}(i)$
  $\hat{\mathbf{h}}(i) = \mathbf{R}^{-1}(i)\boldsymbol{\beta}(i)$
  Compute $\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_1, \mathbf{B}_2, \mathbf{C}_1, \mathbf{C}_2$ as in (26), (27), and (30)-(33)
  Generate the matrix $\mathbf{D}$ as in (38) and vector $\hat{\mathbf{h}} = [\hat{\mathbf{h}}(i);\ \hat{\mathbf{h}}(i+M/2);\ \hat{\mathbf{h}}(i+M)]^T$
  Solve the system of equations $\mathbf{Dz} = \hat{\mathbf{h}}$
  $\tilde{\mathbf{h}}(i) = \hat{\mathbf{h}}_0(i) = [\mathbf{z}]_{1,\ldots,L}$
end

the desired signal samples, $\mathbf{I}_L$ is an $L \times L$ identity matrix, $\mathbf{x}(i) = [x(i), x(i-1), \ldots, x(i-M+1)]^T$ is an $M \times 1$ desired signal vector at the ith time instant.

IV. Baseband Simulation

In this section, we first show that the delayed RLS algorithms provide improvement in the MSD performance and then investigate the dependence of the performance on the delay. It will be shown that, for the SRLSd algorithm, the optimal delay is $T = M/2$, as discussed in Section III-B. However, for the ERLSd algorithm, there is no one-to-one relationship between the optimal delay and the forgetting factor λ.

We show that the MSE is useful for characterising the SIC performance if $T = 0$, i.e., for classical RLS algorithms. However, if $T > 0$, the MSE is not a useful characteristic for this purpose. We then show that the proposed SICF metric is suitable for characterising the SIC performance for both the cases, in particular by comparing it with the bit error rate (BER) performance of a far-end transmission.

In the simulation, we set the filter length to $L = 50$, and model the SI channel as follows. Every element $[\mathbf{h}(i)]_\ell$ of $\mathbf{h}(i)$ is a stationary random process with a power spectral density $c_\ell G(f)$, where $G(f)$ is uniform within a frequency interval $[-f_{\max}, f_{\max}]$, and c_ℓ is the variance of the ℓth channel tap. The UWA channel normally has a decaying power delay profile due to the spreading and absorbtion loss [22]. The power delay profile c_ℓ is generated as:

$$c_\ell = e^{-\gamma\ell}, \qquad \ell = 0, \ldots, L-1, \tag{40}$$

and γ is chosen to control the ratio between the variance of latest arrivals ($\ell = L-1$) and that of the first arrivals ($\ell = 0$). In this scenario, γ is chosen to make this ratio equal to 80 dB.

The random processes $[\mathbf{h}(i)]_\ell$ are independent for different ℓ, and they are generated using the FFT-method [23]. We assume a sampling frequency of 1 kHz, so that one channel tap delay is 1 ms. The parameter $f_{\max}$ determines the maximum speed of the channel variation. To model fast time-varying channels, we use $f_{\max} = 1$ Hz; for slow time-varying channels, $f_{\max} = 0.1$ Hz.

In Fig. 6, a snapshot of the channel impulse response generated through the aforementioned process is shown, which has

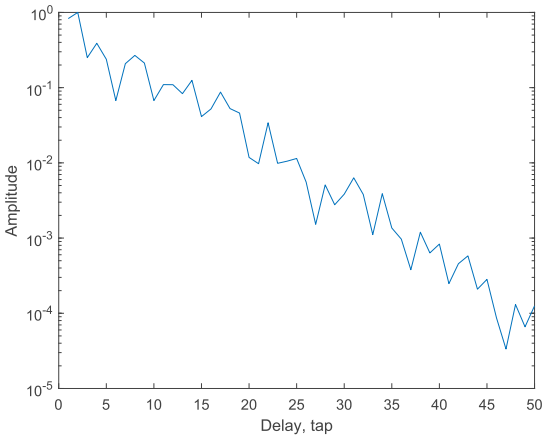

Fig. 6. A snapshot of the channel impulse response.

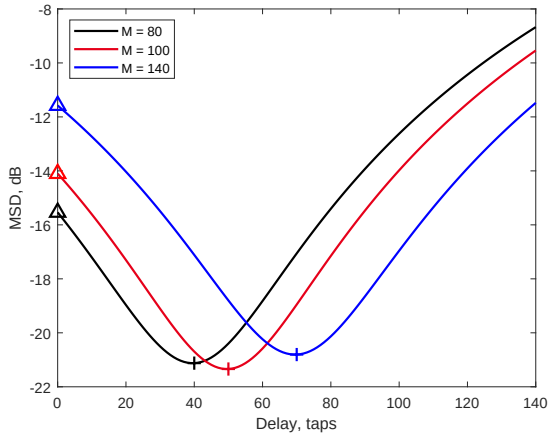

Fig. 7. MSD performance of the SRLS (Delay $T = 0$) and SRLSd algorithms. The optimal delay minimising the MSD is $T = M/2$.

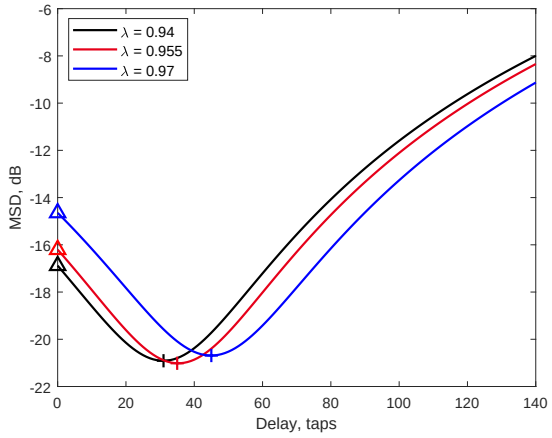

Fig. 8. MSD performance of the ERLS (Delay $T = 0$) and ERLSd algorithms.

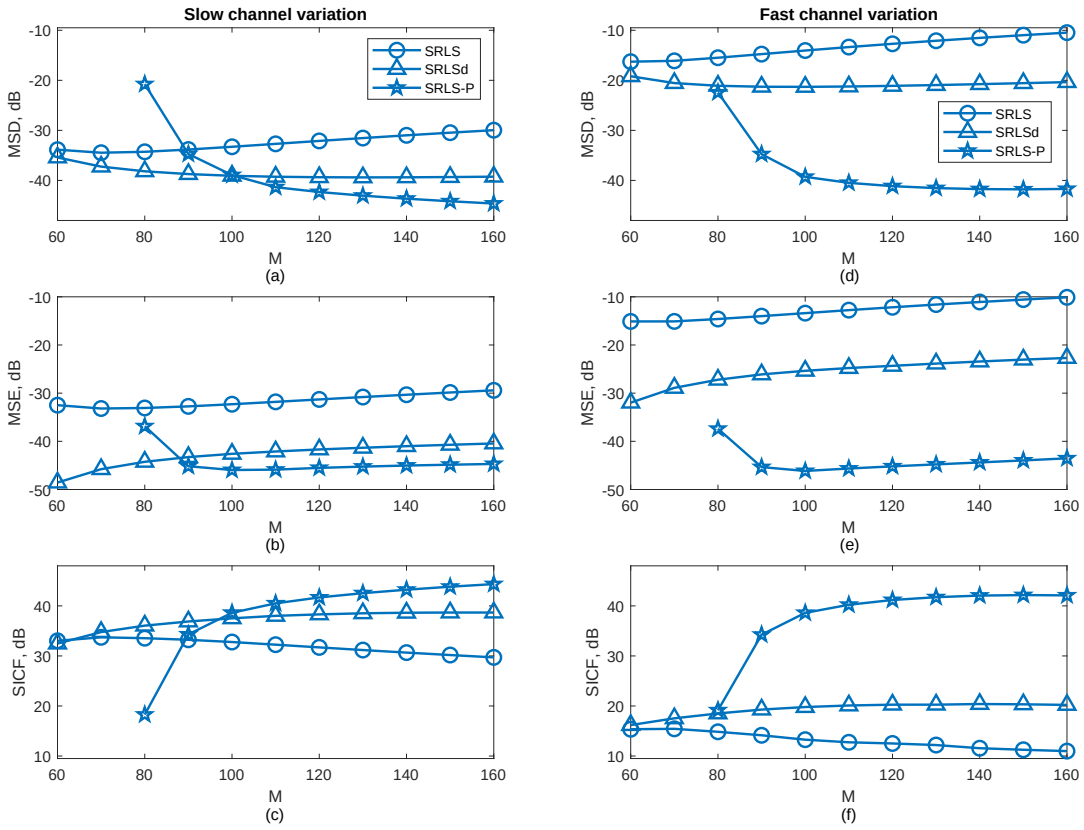

Fig. 9. MSD, MSE and SIC performance of the SRLS, SRLSd, and SRLS-P adaptive filters in slow and fast varying SI channels: (a) MSD in slow varying channels; (b) MSE in slow varying channels; (c) SIC in slow varying channels; (d) MSD in fast varying channels (e) MSE in fast varying channels; (f) SIC in fast varying channels.

a similar shape as the SI channel impulse response obtained in our water tank experiments [9].

A. MSD performance of RLS algorithms with a delay

Fig. 7 shows the normalized MSD ($\mathrm{MSD}(i)/||\mathbf{h}(i)||_2^2$) as a function of the delay T against M for the SRLSd algorithm. The SRLSd algorithm can provide a significant improvement in the MSD performance compared to the SRLS algorithm ($T = 0$). It can also be seen that the optimal delay is $T = M/2$. The minimum MSD is achieved at $T = M/2 = 50$ for $M = 100$. Fig. 7 also shows that with further increase in the delay T, the MSD increases and, as expected, reaches the same level at $T = M$ as at $T = 0$.

In Fig. 8, we observe that the MSD performance of the ERLS algorithm can also be improved by introducing a delay, but the criterion of selecting the optimal delay is not obvious. As can be seen that, in this simulation scenario, the minimum MSD is achieved for $\lambda = 0.955$ and $T = 37$.

To test if the dependencies between the optimal delay and the window parameters can be applied generally, we ran 1000 simulation trials to find the distribution of the optimal delay for the SRLSd and ERLSd adaptive filters, with $M = 100$ and $\lambda = 0.955$. The results show that, for the SRLSd algorithm, the optimal delay is always $T = M/2$ in all simulation trials. However, for the ERLSd adaptive filter, the minimum MSD is obtained at $T = 37$ in 91.5% of the trials, while, in the other trials, the optimal delay is $T = 36$ or $T = 38$.

B. MSE, MSD and SIC performance of SRLS, SRLSd and SRLS-P algorithms

Fig. 9 presents the MSE, MSD and SICF performance of the adaptive filters in slow and fast varying channels. We consider the case when the power of the far-end signal is significantly higher than the noise power, thus the noise is not added to the far-end signal. The far-end signal to SI ratio is set to -43 dB.

We can see that, for the SRLS algorithm ($T = 0$), the optimal sliding window length M found from the MSE and MSD curves is about the same ($M = 60$ or 70). However, for the other algorithms with $T > 0$, the optimal M corresponding to the minimum MSE and MSD are different.

The SRLS-P adaptive filter has a significantly improved MSD performance compared to the SRLSd algorithm, which in turn outperforms the SRLS algorithm. Note that, in the SRLS-P algorithm, there are $3L$ unknown parameters to be estimated. Therefore, since the estimation interval in the SRLS-P algorithm is $2M$, the estimation requires the window length to be at least $M = 3L/2 = 75$; this explains the increase of the MSD at low M.

The results in Fig. 9 show that the MSE is lower than the far-end signal to SI ratio for the SRLSd adaptive filter with $M < 80$. This indicates that the far-end signal is partly cancelled, therefore the MSE is not useful as a performance measure here. In Fig. 9, the SIC performance of the adaptive filters is also shown. We can see that the optimal M for the SICF and MSD curves are very close but not exactly the same. We will show in the next section that the proposed SICF metric provides a better indication of performance of the SI canceller than the MSD.

C. MSD, SIC and BER performance of the SRLSd algorithm

We now investigate the relationship between the MSD, SIC and BER performance provided by the SI canceller in fast time-varying channel ($f_{\max} = 1$ Hz) based on the SRLSd algorithm. Fig. 10 shows these three characteristics for different values of M. We run 500 simulation trials, and in each trial a new time-varying channel is generated. The length of the realization is 15s. The symbol rate is 1 kHz. The received signal is generated by adding the far-end signal and noise to the SI channel output. Samples of the far-end signal and noise are generated as Gaussian random zero-mean numbers. The noise variance σ_n^2 is set according to the SI to noise ratio ($\mathrm{SNR_{SI}}$), which is defined as

$$\mathrm{SNR_{SI}} = \frac{\mathrm{E}\{|x(i)|^2\}}{\sigma_n^2}. \tag{41}$$

The far-end signal level is defined by the far-end SNR as σ_f^2/σ_n^2. Here we set $\mathrm{SNR_{SI}} = 43$ dB, and the far-end SNR varies from 10 dB to 19 dB. The SICF is computed over the convergence period from 2 to 15 s, which is about ten times longer than the time correlation of the SI channel.

The BER shows the detection performance of the far-end data after SIC, which is an important indicator of the FD system performance. The best detection performance is achieved with $M = 140$ or $M = 160$ when the far-end SNR lower than 16 dB. The BER slightly degrades for $M = 120$, and further degrades for smaller M. However, the MSD gives a different indication as the minimum MSD is achieved with $M = 100$ or $M = 120$ when the far-end SNR is lower than 16 dB.

The SICF indicates that the best performance is achieved with $M = 140$ when the far-end SNR lower than 14 dB and with $M = 160$ when the far-end SNR between 14 dB and 19 dB. It is clear that the SIC provides a better indication of the optimal M for the detection performance. More importantly, in practice, the MSD is difficult to compute since the true channel response is unknown, whereas the proposed SICF metric is computed without such knowledge as explained in Section II.

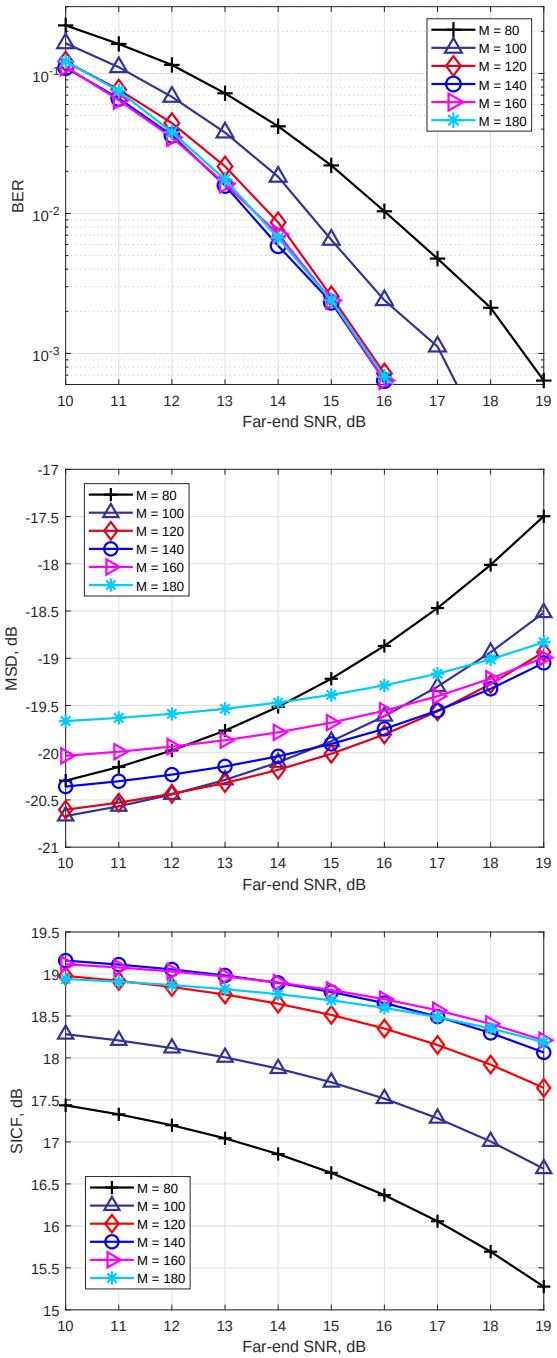

Fig. 10. MSD, SICF and BER performance in the fast varying SI channel.

V. Passband Simulation Results

In this section, we investigate the SIC performance of the SRLS, SRLSd and SRLS-P adaptive filters in scenarios with time-varying SI channels. We use the SIC scheme shown in Fig. 2. The SI channel has one direct path between the projector and hydrophone and one path due to reflection from a time-varying surface. The reflected path is 20 dB weaker than the direct path. The surface is modelled as a sinusoid wave of 0.5 m amplitude and 3 s period. The projector and hydrophone are vertically separated by a distance of 0.5 m, their depths are 9.5 m and 10 m, respectively. We will show that the SIC performance can be significantly improved by the SRLS-P adaptive filter which accurately models the channel variation caused by the time-varying surface reflection.

In the simulation, a 10 s signal with BPSK (binary phase-shift keying) modulation at a 12 kHz carrier frequency and with 1 kHz signal bandwidth is transmitted. The BPSK sym-

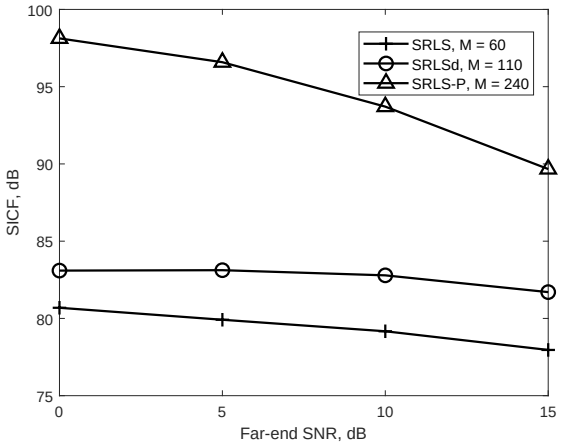

Fig. 11. SIC performance of adaptive filters in the passband simulation.

bols are pulse shaped using the root-raised cosine filter with a roll-off factor of 0.2. The sampling rate of the passband signal is 96 kHz.

The received signal at the hydrophone is generated by adding the far-end signal and noise to the SI channel output. Here we set $\mathrm{SNR}_{\mathrm{SI}}$ = 100 dB and consider the far-end SNR between 0 dB and 15 dB.

Fig. 11 shows the SIC performance of the SRLS, SRLSd and SRLS-P adaptive filters, which is computed over the time interval from 2 s to 10 s (after the convergence of adaptive filters). For each adaptive filter, the parameter M is adjusted to provide the highest SICF. The filter length is $L = 40$, which is long enough to cover both the main path and the surface reflection. For the SRLS adaptive filter, around 81 dB of SIC can be achieved at 0 dB far-end SNR ($M = 60$). The SICF is improved by 3 dB when the SRLSd adaptive filter ($M = 110$) is used, and it is further improved to 98 dB (by 14 dB) with the SRLS-P adaptive filter ($M = 240$).

VI. Experimental Results

In this section, we investigate the SIC performance of the SRLS, SRLSd and SRLS-P adaptive filters in the lake experiment with the SIC scheme shown in Fig. 2. The configuration and experimental setup are shown in Fig. 12 and 13, respectively. The lake depth at the experimental site is around 8 m. The distance between the projector and the hydrophone is around 1.3 m. The hydrophone is placed at 4 m depth. In the experiment, we transmit a 15 s BPSK signal at the carrier frequency 14 kHz and with a bandwidth of 1 kHz; the pulse shaping roll-off factor is 0.2. The amplitude of the lake surface wave varies from 5 cm to 10 cm during the transmission.

In Fig. 14, we show the SI channel estimates obtained with the SRLS-P adaptive filter, which provides the highest SICF among the adaptive filters we considered. It can be seen that the SI channel consists of a strong and stable main path and multiple fast time-varying multipaths reflected from the lake surface and bottom.

In the experiment, the SI to noise ratio is around 48 dB. The filter length is $L = 80$, which is long enough to cover the

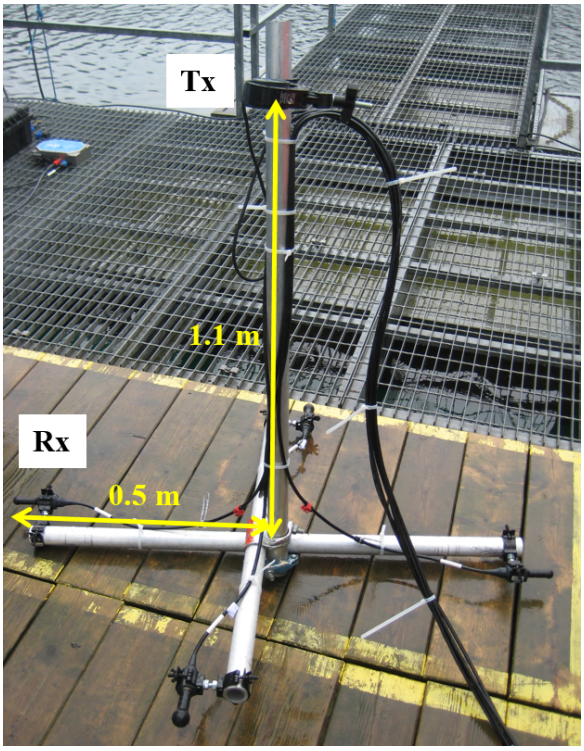

Fig. 12. The configuration of the lake experiments.

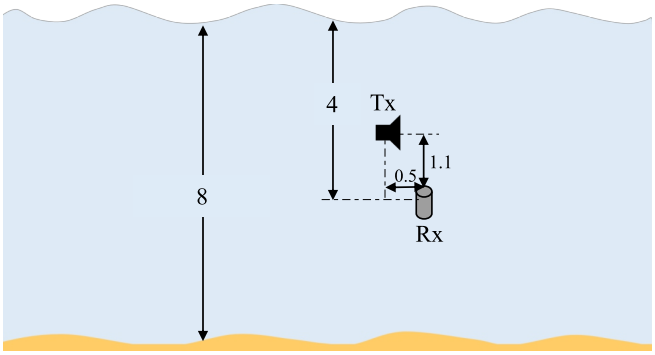

Fig. 13. The experimental setup. The distance is shown in meters.

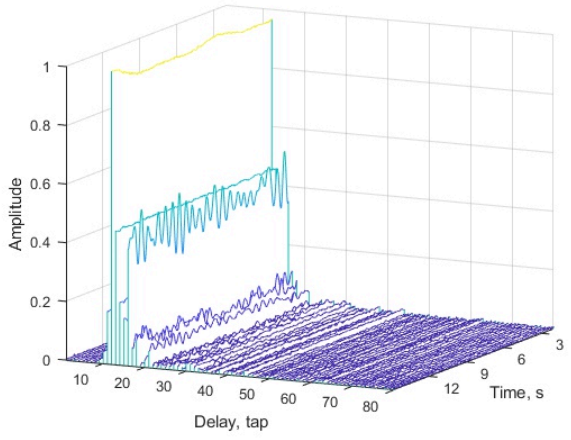

Fig. 14. SI channel estimate for the lake experiment.

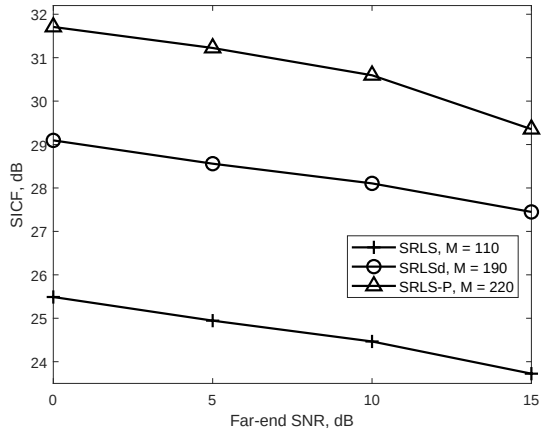

Fig. 15. SIC performance of adaptive filters for the lake experiment.

channel delay spread, including the direct path and multiple reflections from the surface and bottom. The SIC performance is computed over the time interval from 2 s to 15 s. Fig. 15 shows the SIC performance of three adaptive filters with the optimal sliding window lengths M. For the SRLS adaptive filter, at 0 dB far-end SNR, 25.5 dB of SIC is achieved when $M = 110$. The SICF is improved to 29 dB when the SRLSd adaptive filter with $M = 190$ is used. The SRLS-P adaptive filter with $M = 220$ achieves 32 dB of SIC.

The experimental results demonstrate that the SRLS-P adaptive filter provides the best SIC performance among the three adaptive filters. More than 6 dB improvement in the SICF can be achieved by using the SRLS-P adaptive filter compared to that of the SRLS adaptive filter.

It can be seen that the improvement in SICF for the lake experiment is lower than that achieved in the passband simulation. The power spectral density computed for the first reflection from the lake surface (with an amplitude of about 0.4 as seen in Fig. 14), has shown that $f_{\max} > 2$ Hz. For the further reflections from the lake surface and bottom, as can be seen in Fig. 14, the variation speed is even higher. With $M = 220$, the product of the estimation window length (0.44 s) by $f_{\max}$ is already close to one, which is less than the Nyquist lower boundary. With such settings, one cannot expect high accuracy of estimating the SI channel due to high modelling errors [24]. Still, the SRLS-P algorithm shows improvement by 5.5 to 6 dB against the SRLS algorithm and by 1.5 to 2.5 dB against the SRLSd algorithm.

The estimation accuracy could have been improved using lower M. However, for the identifiability, the number of available signal samples ($2M$) should be higher than the number of unknown parameters ($3L$), i.e. $M > 3L/2$. For M very close to the boundary $3L/2$, the algorithm performance is limited (see Fig. 9). Reduction in L allows smaller M, but, in this case, the SIC performance will be limited by the SI arrivals being truncated by the filter.

VII. Conclusions

In this paper, the SICF has been proposed as a practical measure of the SIC performance in FD UWA systems. The SICF has been investigated in comparison with the MSE, MSD and BER. It is shown through numerical simulation that the proposed metric provides a good indication of the SI canceller performance.

To improve the SIC performance of the RLS adaptive filters, we have considered their delayed versions, the SRLSd and ERLSd adaptive filters. The dependence of the SIC performance on the delay of the input signals for these adaptive filters has been investigated using numerical simulations. We have shown that, for the SRLSd adaptive filter, the optimal delay is the half of the sliding window length. For the ERLSd adaptive filter, the relationship between the optimal delay and the forgetting factor can differ for different channel realizations, although, with an optimal delay, the ERLS adaptive filter can provide the same level of SIC performance as the SRLSd adaptive filter.

We have proposed the SRLS-P adaptive filter, which is based on the SRLS algorithm and modelling the channel response variation within a short time interval as a second-order algebraic polynomial. The SIC performance of the SRLS-P adaptive filter has been investigated and compared with that of the SRLS and SRLSd adaptive filters using numerical and lake experiments. The SRLS-P algorithm achieves the highest SICF among these adaptive filters.

Acknowledgement

The work of Y. Zakharov, B. Henson, N. Morozs and P. Mitchell was supported in part by the U.K. EPSRC through Grants EP/P017975/1 and EP/R003297/1.

Appendix

We now derive the presentation (24) for the channel estimate $\hat{\mathbf{h}}(i)$ obtained by the SRLS algorithm.

In the SRLS adaptive filter, without the noise, the estimate at time instant i is given by

$$\begin{aligned}\hat{\mathbf{h}}(i) &= [\mathbf{S}^H(i)\mathbf{S}(i)]^{-1}\mathbf{S}^H(i)\mathbf{x}(i),\\ &= \mathbf{R}^{-1}(i)\mathbf{S}^H(i)\,\mathrm{diag}\{\mathbf{S}(i)\mathbf{H}(i)\},\\ &= \mathbf{R}^{-1}(i)\mathbf{S}^H(i)\sum_{m=0}^{M-1}\mathbf{e}_m\mathbf{e}_m^T\mathbf{S}(i)\mathbf{H}(i)\mathbf{e}_m,\\ &= \frac{1}{M}\mathbf{R}^{-1}(i)\sum_{k=0}^{M-1}\mathbf{S}^H(i)\mathbf{e}_m\mathbf{e}_m^T\mathbf{S}(i)\mathbf{H}(i)\mathbf{e}_m, \qquad (42)\end{aligned}$$

where $\mathbf{e}_m$ is a column vector of zero elements, apart from the mth element which equals one, $\mathbf{S}(i) = [\mathbf{s}(i), \ldots, \mathbf{s}(i-M+1)]^T$ is an $M \times L$ observation matrix, $\mathbf{H}(i) = [\mathbf{h}(i), \ldots, \mathbf{h}(i-M+1)]$ is an $L \times M$ channel matrix, and $\mathbf{h}(i)$ is the true channel impulse response at the ith time instant. Here, we used

the fact that, in the absence of noise, $\mathbf{x}(i) = \text{diag}\{\mathbf{S}(i)\mathbf{H}(i)\}$. Equation (42) can be further written as:

$$\begin{aligned}\hat{\mathbf{h}}(i) &= \frac{1}{M}\mathbf{R}^{-1}(i)\sum_{m=0}^{M-1}\mathbf{s}^*(i-m)\mathbf{s}^T(i-m)\mathbf{H}(i)\mathbf{e}_m \\ &= \frac{1}{M}\mathbf{R}^{-1}(i)\sum_{m=0}^{M-1}\mathbf{s}^*(i-m)\mathbf{s}^T(i-m)\mathbf{h}(i-m) \\ &= \frac{1}{M}\mathbf{R}^{-1}(i)\sum_{m=0}^{M-1}\mathbf{R}_{i-m}\mathbf{h}(i-m) \\ &= \frac{1}{M}\mathbf{R}^{-1}(i)\sum_{m=-M+1}^{0}\mathbf{R}_{i+m}\mathbf{h}(i+m),\end{aligned} \tag{43}$$

where we use $\mathbf{H}(i)\mathbf{e}_m = \mathbf{h}(i-m)$, $\mathbf{S}^H(i)\mathbf{e}_m = \mathbf{s}^*(i-m)$, and denote $\mathbf{R}_{i-m} = \mathbf{s}^*(i-m)\mathbf{s}^T(i-m)$. By replacing i with $i+k$, this can also be rewritten as:

$$\hat{\mathbf{h}}(i+k) = \frac{1}{M}\mathbf{R}^{-1}(i+k)\sum_{m=-M+k+1}^{k}\mathbf{R}_{i+m}\mathbf{h}(i+m). \tag{44}$$

REFERENCES

[1] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *16th Annual International Conference on Mobile Computing and Networking*, 2010, pp. 1–12.

[2] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *44th Asilomar Conference on Signals, Systems and Computers*, 2010, pp. 1558–1562.

[3] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *17th Annual International Conference on Mobile Computing and Networking*, 2011, pp. 301–312.

[4] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 2012.

[5] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, 2013, pp. 375–386.

[6] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, and G. Dejean, "Rethinking indoor wireless mesh design: Low power, low frequency, full-duplex," in *IEEE 5th Workshop on Wireless Mesh Networks*, 2010, pp. 1–6.

[7] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *45th Asilomar Conference on Signals, Systems and Computers*, 2011, pp. 2002–2006.

[8] 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.

[9] 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*, 2019, in press.

[10] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3821–3836, 2014.

[11] L. Ding, "Digital predistortion of power amplifiers for wireless applications," Ph.D. dissertation, Georgia Institute of Technology, 2004.

[12] S. Li and R. D. Murch, "An investigation into baseband techniques for single-channel full-duplex wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4794–4806, 2014.

[13] L. Shen, B. Henson, Y. Zakharov, and P. Mitchell, "Robust digital self-interference cancellation for full-duplex UWA systems: Lake experiments," in *Underwater Acoustics Conference and Exhibition, Crete, Greece*, July 2019, pp. 243–250.

[14] L. Shen, B. Henson, Y. V. Zakharov, and P. Mitchell, "Two-stage self-interference cancellation in full-duplex underwater acoustic systems," in *MTS/IEEE Oceans - Marseille*, June 2019, pp. 1–6.

[15] P. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, ser. Kluwer international series in engineering and computer science. Springer US, 2008.

[16] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 2002.

[17] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: Pushing the limits," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 80–87, 2016.

[18] M. Emara, P. Rosson, K. Roth, and D. Dassonville, "A full duplex transceiver with reduced hardware complexity," in *IEEE Global Communications Conference*, 2017, pp. 1–6.

[19] G. Qiao, S. Gan, S. Liu, and Q. Song, "Self-interference channel estimation algorithm based on maximum-likelihood estimator in in-band full-duplex underwater acoustic communication system," *IEEE Access*, vol. 6, pp. 62 324–62 334, 2018.

[20] 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.

[21] Y. V. Zakharov and J. Li, "Sliding-window homotopy adaptive filter for estimation of sparse UWA channels," in *IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2016, pp. 1–4.

[22] 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.

[23] D. B. Percival, "Simulating gaussian random processes with specified spectra," *Computing Science and Statistics*, pp. 534–534, 1993.

[24] 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.