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

LU SHEN, (Graduate Student Member, IEEE), YURIY ZAKHAROV, (Senior Member, IEEE), BENJAMIN HENSON, (Member, IEEE), NILS MOROZS, (Member, IEEE), AND PAUL D. MITCHELL, (Senior Member, IEEE)

Department of Electronic Engineering, University of York, York YO10 5DD, U.K.

Corresponding author: Lu Shen (ls1215@york.ac.uk)

ABSTRACT To enable full-duplex (FD) in underwater acoustic (UWA) systems, a high level of self-interference (SI) cancellation (SIC) is required. This can be achieved by using a combination of SIC methods, including digital SIC. 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 article, 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 to evaluate the performance of digital SIC techniques without the need of implementing a full FD system. 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 can significantly improve 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 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 SI cancellation (SIC) techniques have been proposed for FD terrestrial radio systems. A combination of SIC methods is used, including digital SIC.

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

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 SIC in FD UWA systems [8]–[10].

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 [11]. 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], [12]. 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], [13], [14]. 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.
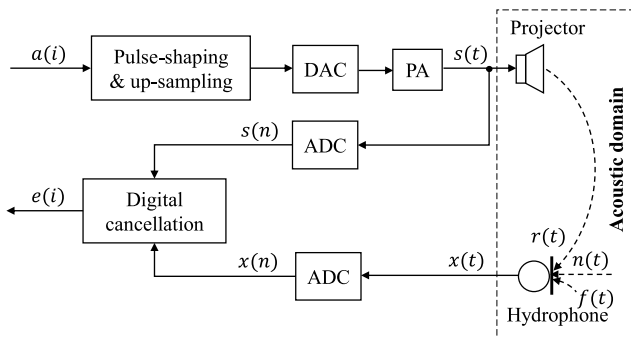

FIGURE 1. Block diagram of the FD UWA system with digital cancellation. The system works at two sampling frequencies. The index of the signal sample with the high (passband) sampling rate is denoted by *n*, and the low (baseband) sampling rate sample index is *i*. The analogue (passband) signals are: the PA output $s(t)$; the SI $r(t)$; the noise $n(t)$; the far-end signal $f(t)$; 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].

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 [15]; 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 article, the SIC performance using different adaptive filters is investigated with this digital cancellation scheme.

Another phenomenon that limits the SIC is the time-varying surface reflections [16], [17]. While a high level of SIC 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 [18], [19]. 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.

To measure the performance of a SI canceller, the best approach would seem to be to implement a whole FD system and investigate its performance. However, so far the FD UWA technology is not mature enough to make this approach practical. It is still a problem to cancel the SI close to the noise level. The signal distortions in the SI channel and their influence onto the SI cancellation performance have not been yet well understood. Such issues have been partly addressed separately or in some combination as the PA nonlinearity [8], [9], the nonlinearity in the preamplifier of the hydrophone [20], passband to baseband conversion [15], modelling of the near-end surface reflections [14], modelling of the self-loop interference through the mounting system [21], acoustic isolation of the projector and hydrophone [22], etc. In this article, we propose an improved technique for dealing with the fast time-variability of the SI channel in digital SI cancellers. Other problems that need to be dealt with are the acoustic cavitation [23] and nonlinear signal distortion in the projector. To build a whole FD system, all these issues should be taken into account together, which will be done in the future.

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 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 SIC performance [5], [24], [25], which can be measured by the mean squared error (MSE) [22]. 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 article, 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 article 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.
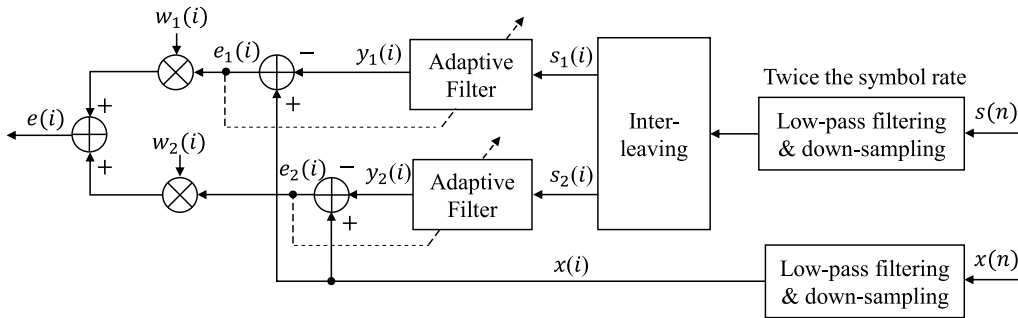4) The proposed algorithm is investigated using numerical simulations and lake experiments, and its performance

FIGURE 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 [15].
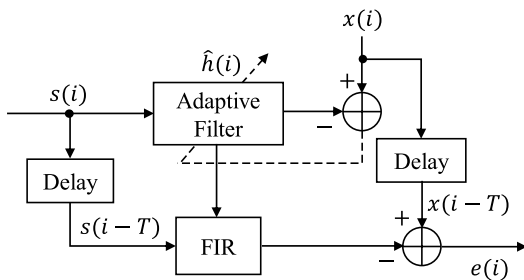

FIGURE 3. Adaptive filter with a delay.

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. Section VII draws conclusions.

Notations: In this article, we use capital and small bold fonts for matrices and vectors, respectively; e.g, $\mathbf{R}$ and $\mathbf{h}$. We also denote the expectation as $\mathrm{E}\{\cdot\}$, 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

The mean squared error (MSE) and the mean squared deviation (MSD) are normally used for evaluating the channel estimation performance [19], [26]. In subsection II-A, we discuss if it is practical to use these two metrics to evaluate the SIC performance in FD systems. In subsection II-B, a new metric, the SICF, is proposed for evaluation of the SIC performance.

A. MSE AND MSD 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:

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

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 operating in the identification mode, $\mathbf{s}(i)$ is the regressor and $x(i)$ is the desired signal [18], [19]. 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 $i-1$. In this case, the FIR filter shown in Fig. 3 is not required since it is the same as the FIR filter within the adaptive filter with the same input. However, if $T > 0$, the regressors of these FIR filters are different, they are $\mathbf{s}(i)$ for the adaptive filter and the delayed regressor $\mathbf{s}(i - T)$ for the FIR filter. 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:

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

The performance of an adaptive filter is most often evaluated using the mean squared error (MSE) [18], [19]. The MSE is defined as:

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

For a classical adaptive filter (with $T = 0$), the SIC performance can be evaluated by computing the MSE. However, by adjusting parameters of an adaptive filter with a delay (non-causal adaptive filter), 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.

From the interference cancellation point of view, the SIC performance can be evaluated by how much the near-end

SI is cancelled. Therefore, everything apart from the near-end SI is treated as the signal of interest (including far-end signal and the noise), which should be recovered. The signal to interference ratio (SIR) at the SI canceller can be written as:

$$\text{SIR}(i) = \frac{\sigma_z^2}{\text{E}\{|\varepsilon(i)|^2\}}, \tag{4}$$

where $\sigma_z^2 = \text{E}\{|z(i)|^2\}$, $z(i)$ is the signal of interest that includes the far-end signal and the noise, and $\varepsilon(i)$ is the residual interference.

If the far-end signal and the error signal are not correlated, then the residual interference $\varepsilon(i)$ can be represented as:

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

and substituting (1) and (2) into (5), we have:

$$\varepsilon(i) = [\mathbf{h}(i - T) - \hat{\mathbf{h}}(i)]^H \mathbf{s}(i - T). \tag{6}$$

Assuming that $s(i)$ are uncorrelated for different i and uncorrelated to $\hat{\mathbf{h}}(i)$, we have:

$$\text{E}\{\mathbf{s}(i - T)\mathbf{s}^H(i - T)\} = \sigma_s^2 \mathbf{I}_L, \tag{7}$$

where $\sigma_s^2 = \text{E}\{|s(i)|^2\}$ is the variance of the signal $s(i)$, which is assumed stationary. Then using (6) and (7), we obtain:

$$\text{E}\{|\varepsilon(i)|^2\} = \sigma_s^2 \text{E}\{\|\mathbf{h}(i - T) - \hat{\mathbf{h}}(i)\|_2^2\} \tag{8}$$
$$= \sigma_s^2 \text{MSD}(i - T), \tag{9}$$

where the MSD is defined as:

$$\text{MSD}(i) = \text{E}\{\|\mathbf{h}(i) - \hat{\mathbf{h}}(i + T)\|_2^2\}. \tag{10}$$

Finally, we obtain:

$$\text{SIR}(i) = \frac{\sigma_z^2}{\sigma_s^2} \cdot \frac{1}{\text{MSD}(i - T)}. \tag{11}$$

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 signal of interest (including noise) and near-end interference improves due to the accuracy of the near-end channel estimation. 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 (11) is only applicable if $\hat{\mathbf{h}}(i)$ and $s(i)$ are uncorrelated, which may not be the case for adaptive filters with delay.

B. SIC FACTOR

In [4], [5], [24], [25], the interference cancellation gain, which is defined as the ratio of the near-end SI power to the residual SI power, is used for evaluating the performance of the SI canceller. Note that the residual SI is computed as in (5) assuming that the far-end signal is not correlated with the error signal. This assumption is no longer valid when adaptive filters with delay are used. We now propose the SICF, which is shown to provide 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. This SICF can be used to evaluate the SIC performance without the need of implementing a full FD system.

Here we consider the SI cancellation problem from the far-end signal detection point of view. The higher far-end signal to residual interference ratio at the SI canceller output, the better the SIC performance. In this scenario, the far-end signal is the signal of interest, and everything else is treated as interference (including noise). Since the far-end signal level is typically higher than the receiver's noise floor, the noise is ignored in the derivation below to simplify the expression. Although the noise is ignored in our derivation, the metric SICF is applicable in the case when the noise is present; this can be seen in numerical results presented in Section IV.

Fig. 4 illustrates our description below. We artifically add to the SI signal $r(i)$ a known signal $f(i)$ assumed to be a far-end signal. The level of the signal $\sigma_f^2 = \text{E}\{|f(i)|^2\}$ is chosen to guarantee a predefined input SIR:

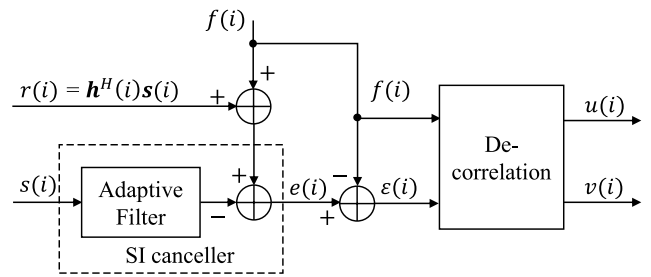$$\text{SIR}_{\text{in}}(i) = \frac{\sigma_f^2}{\text{E}\{|r(i)|^2\}}. \tag{12}$$

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

The SI canceller (shown in Fig. 4) subtracts the SI estimate produced by the adaptive filter from the received signal $r(i) + f(i)$. The canceller output $e(i)$ contains the signal of interest $f(i)$ and a residual signal $\varepsilon(i)$:

$$e(i) = f(i) + \varepsilon(i), \tag{13}$$

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

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

$$\text{SICF}(i) = \frac{\text{SIR}_{\text{out}}(i)}{\text{SIR}_{\text{in}}(i)}. \tag{14}$$

By introducing the artificially added far-end signal, the SICF that we propose evaluates the SI canceller performance taking into account the loss of the far-end signal after SIC.

For classical adaptive filters without delay, the signal of interest $f(i)$ and the residual $\varepsilon(i)$ are uncorrelated, thus $\text{SIR}_{\text{out}}(i)$ can be computed as a ratio of their variances. For adaptive filters with delay, due to the over-fitting in the adaptive filter, in general, these two signals are correlated.

Therefore, in this case, we cannot use their ratio for computing $\text{SIR}_{\text{out}}(i)$, another approach is required.

We now assume that the signal of interest $f(i)$ is attenuated due to the imperfection of the adaptive filter. More specifically, we rewrite (13) as:

$$e(i) = \alpha f(i) + [(1-\alpha)f(i) + \varepsilon(i)] \quad (15)$$
$$= u(i) + v(i), \quad (16)$$

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

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

$$\text{E}\{u(i)v^*(i)\} = \text{E}\{\alpha f(i)[(1-\alpha)f(i) + \varepsilon(i)]^*\} = 0. \quad (17)$$

From (17), we find α as:

$$\alpha = 1 + \frac{1}{\sigma_f^2}\text{E}\{f^*(i)\varepsilon(i)\}. \quad (18)$$

After finding α, the modified signal of interest $u(i)$ and residual interference $v(i)$ can be computed from (15), and the ratio of their variances can now be used for computation of $\text{SIR}_{\text{out}}(i)$.

In experiments, the mathematical expectation in (18) is replaced by the average over a time interval after convergence of the adaptive filter. The output SIR can be computed as:

$$\text{SIR}_{\text{out}} = \frac{\|\mathbf{u}\|_2^2}{\|\mathbf{v}\|_2^2}, \quad (19)$$

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 averaging interval P is preferred to be longer than the coherence time of the SI channel.

Note that the far-end signal we used to compute the SICF is artificially added to the received signal, thus it is known at the receiver. The SICF is intended to be used for adjusting the parameters of the adaptive filters to ensure optimal SIC performance can be achieved. In practical systems with real far-end transmission, the far-end signal is unknown. In that case, the SICF can still be computed with an artificial far-end signal for parameter tuning at the training stage.

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.

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:

$$\mathbf{R}(i)\hat{\mathbf{h}}(i) = \boldsymbol{\beta}(i), \quad (20)$$

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:

$$\mathbf{R}(i) = (\lambda - 1)\mathbf{R}(i-1) + \mathbf{s}(i)\mathbf{s}^H(i), \quad (21)$$
$$\boldsymbol{\beta}(i) = (\lambda - 1)\boldsymbol{\beta}(i-1) + x^*(i)\mathbf{s}(i), \quad (22)$$

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 $\lambda^{|i-p|}, p \leq i$.

For the classical SRLS adaptive filter, the update of $\mathbf{R}(i)$ and $\boldsymbol{\beta}(i)$ can be written as [27], [28]:

$$\mathbf{R}(i) = \mathbf{R}(i-1) + \mathbf{s}(i)\mathbf{s}^H(i) - \mathbf{s}(i-M)\mathbf{s}^H(i-M), \quad (23)$$
$$\boldsymbol{\beta}(i) = \boldsymbol{\beta}(i-1) + x^*(i)\mathbf{s}(i) - x^*(i-M)\mathbf{s}(i-M), \quad (24)$$

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 in the SRLS algorithm with respect to the time varying channel response $\mathbf{h}(i)$.
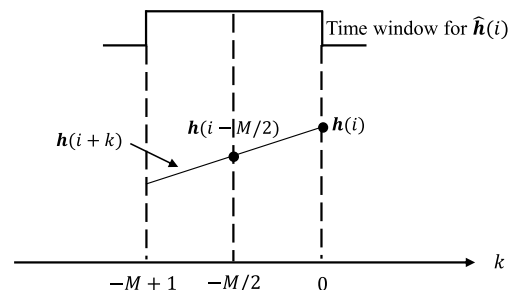

FIGURE 5. Time-varying channel and time windows of the SRLSd algorithm.

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. As shown in Fig. 5, if we assume that the channel response varies linearly in the vicinity of i, then its average over the rectangular window is equal to $\mathbf{h}(i - M/2)$. In such a case, $\hat{\mathbf{h}}(i)$ is a more accurate estimate of $\mathbf{h}(i - M/2)$ than $\mathbf{h}(i)$. 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 forgetting factor 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 SIC 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. 6:

$$\mathbf{h}(i + k) = \mathbf{h}_0(i) + \mathbf{h}_1(i)k + \mathbf{h}_2(i)k^2, \quad (25)$$
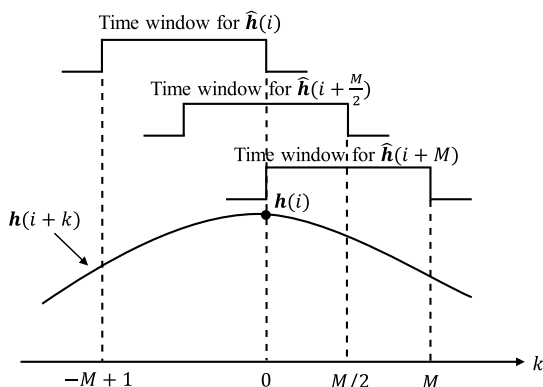

FIGURE 6. Time-varying channel and time windows of the SRLS-P algorithm.

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 (25), 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):

$$\hat{\mathbf{h}}(i + k) = \frac{1}{M}\mathbf{R}^{-1}(i + k) \sum_{m=-M+k+1}^{k} \mathbf{R}_{i+m}\mathbf{h}(i + m), \quad (26)$$

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 (25) into (26) 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:

$$\hat{\mathbf{h}}(i) = \frac{1}{M}\mathbf{R}^{-1}(i)$$
$$\times \sum_{m=-M+1}^{0} \mathbf{R}_{i+m}[\mathbf{h}_0(i) + m\mathbf{h}_1(i) + m^2\mathbf{h}_2(i)]$$
$$= \mathbf{h}_0(i) + \mathbf{A}_1\mathbf{h}_1(i) + \mathbf{A}_2\mathbf{h}_2(i), \quad (27)$$

where

$$\mathbf{A}_1 = \mathbf{R}^{-1}(i) \sum_{m=-M+1}^{0} m\mathbf{R}_{i+m}, \quad (28)$$

$$\mathbf{A}_2 = \mathbf{R}^{-1}(i) \sum_{m=-M+1}^{0} m^2\mathbf{R}_{i+m}. \quad (29)$$

Similarly, we obtain:

$$\hat{\mathbf{h}}(i + M/2) = \mathbf{h}_0(i) + \mathbf{B}_1\mathbf{h}_1(i) + \mathbf{B}_2\mathbf{h}_2(i), \quad (30)$$
$$\hat{\mathbf{h}}(i + M) = \mathbf{h}_0(i) + \mathbf{C}_1\mathbf{h}_1(i) + \mathbf{C}_2\mathbf{h}_2(i), \quad (31)$$

where

$$\mathbf{B}_1 = \mathbf{R}^{-1}(i + M/2) \sum_{m=-M/2+1}^{M/2} m\mathbf{R}_{i+m}, \quad (32)$$

$$\mathbf{B}_2 = \mathbf{R}^{-1}(i + M/2) \sum_{m=-M/2+1}^{M/2} m^2\mathbf{R}_{i+m}, \quad (33)$$

and

$$\mathbf{C}_1 = \mathbf{R}^{-1}(i + M) \sum_{m=1}^{M} m\mathbf{R}_{i+m}, \quad (34)$$

$$\mathbf{C}_2 = \mathbf{R}^{-1}(i + M) \sum_{m=1}^{M} m^2\mathbf{R}_{i+m}. \quad (35)$$

We now arrive at the system of equations:

$$\begin{cases} \mathbf{h}_0(i) + \mathbf{A}_1\mathbf{h}_1(i) + \mathbf{A}_2\mathbf{h}_2(i) = \hat{\mathbf{h}}(i), & (36) \\ \mathbf{h}_0(i) + \mathbf{B}_1\mathbf{h}_1(i) + \mathbf{B}_2\mathbf{h}_2(i) = \hat{\mathbf{h}}(i + M/2), & (37) \\ \mathbf{h}_0(i) + \mathbf{C}_1\mathbf{h}_1(i) + \mathbf{C}_2\mathbf{h}_2(i) = \hat{\mathbf{h}}(i + M), & (38) \end{cases}$$

or, in a compact form,

$$\mathbf{D}\mathbf{z} = \hat{\mathbf{h}}, \quad (39)$$

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

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_L & \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{I}_L & \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{I}_L & \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix}. \tag{40}$$

After solving the system in (39), the estimate of the impulse response is found as the first L elements in the vector $\mathbf{z}$:

$$\tilde{\mathbf{h}}(i) = \hat{\mathbf{h}}_0(i) = [\mathbf{z}]_{1,\dots,L}. \tag{41}$$

Algorithm 1: SRLS-P Algorithm

Input: $\mathbf{s}, \mathbf{x}, L, M, \epsilon$
Output: $\hat{\mathbf{h}}$
Initialization: $\hat{\mathbf{h}}(0) = \mathbf{0}$
for every sample i **do**
    $y(i) = \hat{\mathbf{h}}^H(i-1)\mathbf{s}(i)$
    $e(i) = x(i) - y(i)$
    $\mathbf{R}(i) = \mathbf{S}^H(i)\mathbf{S}(i) + \epsilon\mathbf{I}_L$
    $\boldsymbol{\beta}(i) = \mathbf{S}^H(i)\mathbf{x}(i)$
    $\hat{\mathbf{h}}(i) = \mathbf{R}^{-1}(i)\boldsymbol{\beta}(i)$
    Compute $\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_1, \mathbf{B}_2, \mathbf{C}_1, \mathbf{C}_2$ as in (28), (29), and (32)-(35)
    Generate the matrix $\mathbf{D}$ as in (40) 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 $\mathbf{Dz} = \hat{\mathbf{h}}$
    $\tilde{\mathbf{h}}(i) = \hat{\mathbf{h}}_0(i) = [\mathbf{z}]_{1,\dots,L}$
end

The SRLS-P adaptive algorithm is summarized in Algorithm 1, where ϵ is a regularization parameter, $\mathbf{s}$ is the transmitted signal, $\mathbf{x}$ is the desired signal, $\mathbf{I}_L$ is an $L \times L$ identity matrix, $\mathbf{x}(i) = [x(i), x(i-1), \dots, x(i - M + 1)]^T$ is an $M \times 1$ desired signal vector at the ith time instant.

The complexity of the SRLS-P algorithm will be dominated by the complexity of solving the system of equations in (39). Directly solving the system of equation requires an order of L^3 arithmetic operations. The complexity can be reduced by solving the system of equation recursively based on the solution obtained at the previous time instant using the dichotomous coordinate descent (DCD) algorithm [27]. In such a case, the complexity reduces to an order of $N_u L$ operations, where N_u is the number of DCD updates, which is typically a small number.

IV. BASEBAND SIMULATION

In this section, we first show that the delayed RLS algorithms provide improvement in the MSD performance when identifying time-varying channels 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 [29]. The power delay profile c_ℓ is generated as:

$$c_\ell = e^{-\gamma\ell}, \quad \ell = 0, \dots, L - 1, \tag{42}$$

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 [30]. We assume a sampling frequency $f_s = 1$ 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.

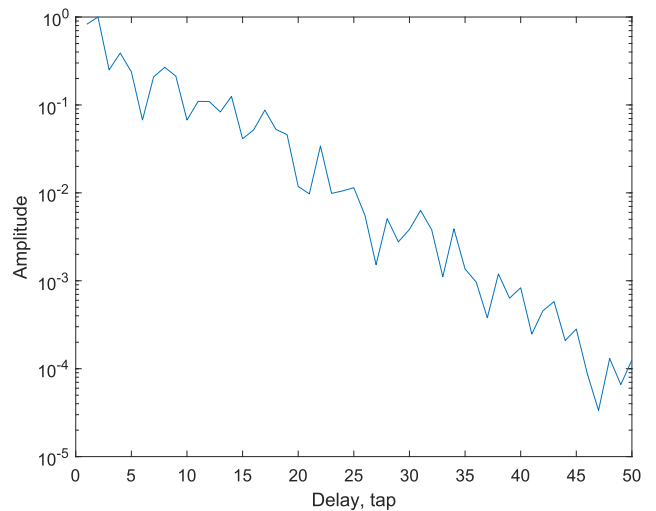In Fig. 7, a snapshot of the channel impulse response generated through the aforementioned process is shown.

FIGURE 7. A snapshot of the channel impulse response.

A. MSD PERFORMANCE OF RLS ALGORITHMS WITH A DELAY

Fig. 8 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 MSD performance is averaged over 20 simulation trials. The choice of M depends on the channel variation speed, the level of noise and other interference. The faster the channel variation, the smaller M should be chosen. The higher the noise level, the higher M is required. The SRLSd algorithm can provide a significant improvement in the MSD performance compared to the SRLS algorithm ($T = 0$). It can be seen that the optimal delay is $T = M/2$. The minimum MSD is achieved at $T = M/2 = 50$ for $M = 100$.
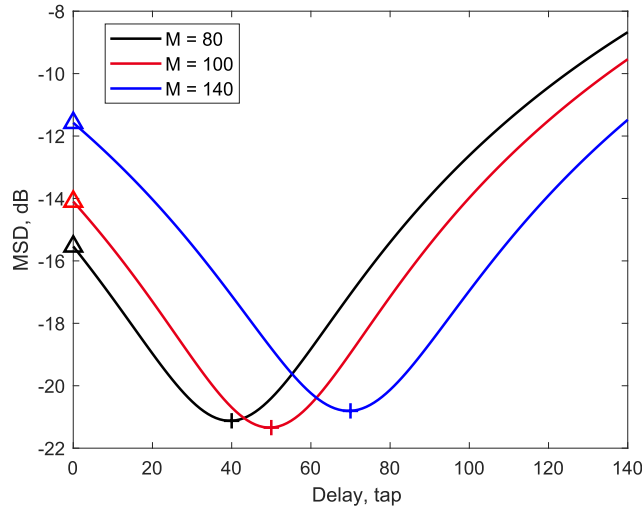
FIGURE 8. MSD performance of the SRLS (Delay $T = 0$) and SRLSd algorithms. The optimal delay minimising the MSD is $T = M/2$.

FIGURE 9. MSD performance of the ERLS (Delay $T = 0$) and ERLSd algorithms.

Fig. 8 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. 9, we observe that the MSD performance of the ERLS algorithm can also be improved by introducing a delay. As can be seen, in this simulation scenario, the minimum MSD is achieved for $\lambda = 0.955$ and $T = 37$. For $\lambda = 0.94$ and $\lambda = 0.97$, the minimum MSD is achieved at $T = 31$ and $T = 45$, respectively. For the ERLS algorithm, from Fig. 9, one can arrive at the following approximate expression for the optimal delay T_{opt}:

$$T_{\text{opt}} \approx \frac{\beta}{\sqrt{1 - \lambda}}, \qquad (43)$$

where $\beta = 7.8$. Note that (43) cannot provide the optimal delay precisely, it can only be used as a reference.

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. 10 presents the MSE, MSD and SIC performance of the adaptive filters in slow and fast varying channels.

The MSE, MSD and SIC performance are computed over the steady-state part of the learning curve from 1000 to 5000 samples. The average interval for the SIC factor computation is 4 s. These three evaluation metrics are all averaged over 20 simulation trials. We consider the case when the
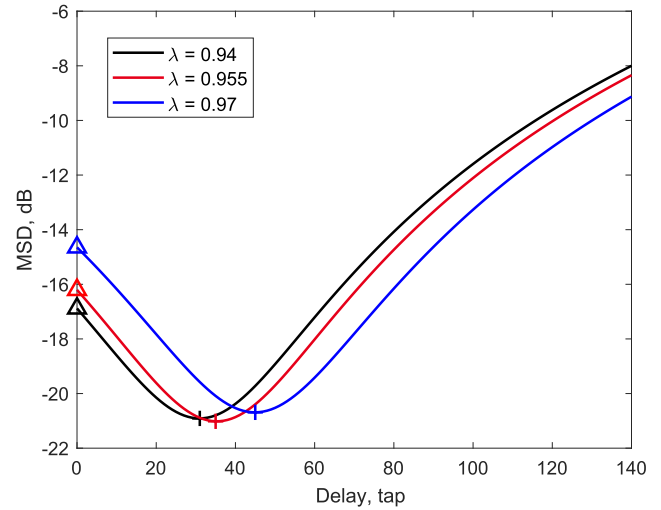
power of the far-end signal is significantly higher than the noise power, thus the noise is not added to the far-end signal. The far-end signal to SI ratio is set to -43 dB.

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 3 L 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. 10 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. 10 (e) and (f), we show the SICF of the adaptive filters together with the inverse MSD. It is seen that the SICF and the inverse MSD for the SRLS adaptive filter are nearly the same, as expected from (11). For the adaptive filters with delay, there is some discrepancy between them for small M. 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 SRLSd AND SRLS-P ALGORITHMS

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) when using the SRLSd and SRLS-P algorithms.

(a) MSD in slow varying channels

(b) MSD in fast varying channels

(c) MSE in slow varying channels

(d) MSE in fast varying channels

(e) SICF and inverse MSD in slow varying channels

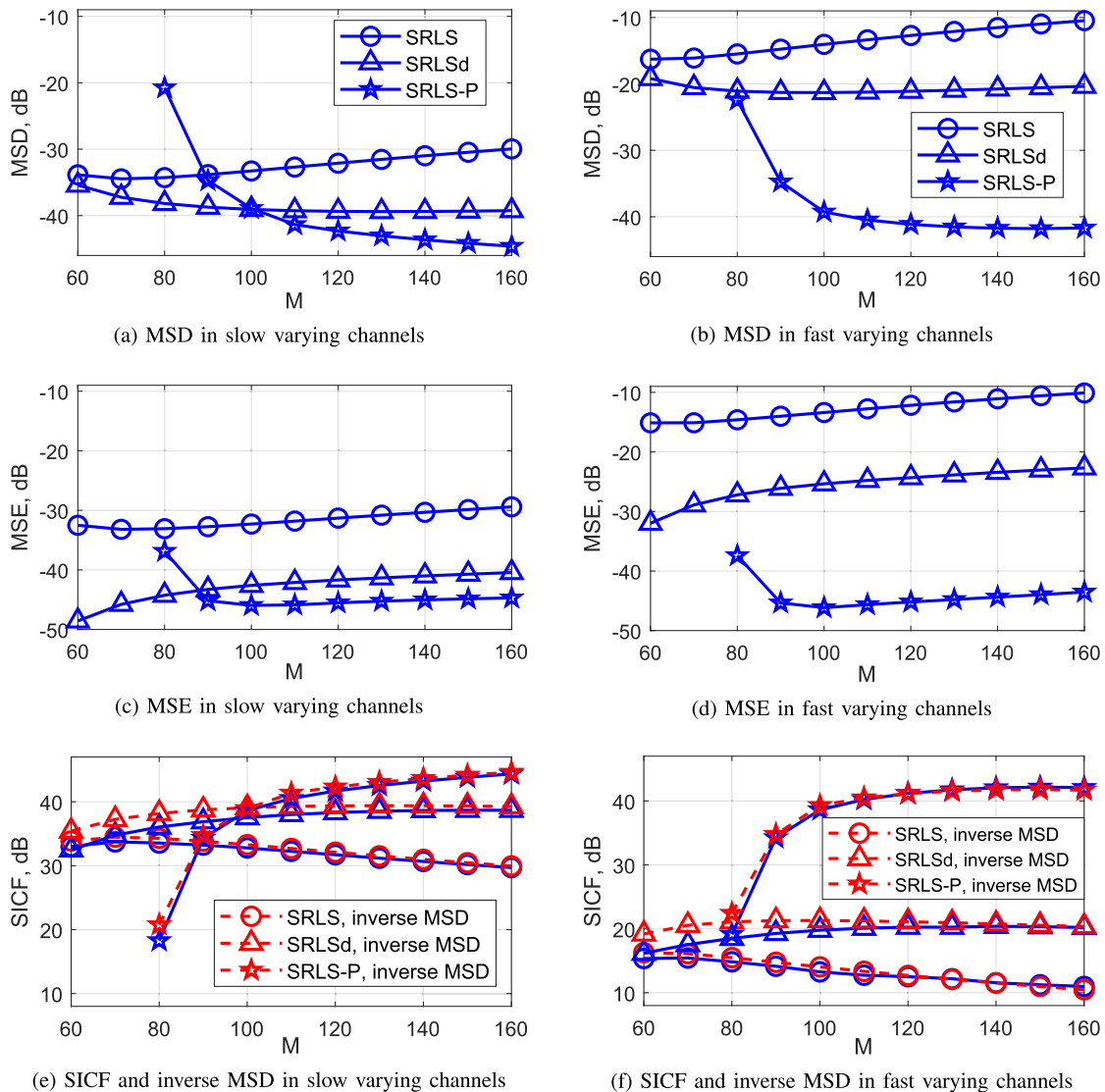(f) SICF and inverse MSD in fast varying channels

FIGURE 10. MSD, MSE and SIC performance of the SRLS, SRLSd, and SRLS-P adaptive filters in slow and fast varying SI channels.

Fig. 11 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 received signal is generated by adding the far-end signal and noise to the SI channel output. Samples of the noise are generated as Gaussian random zero-mean numbers. The noise variance σ_n^2 is set according to the SI to noise ratio (SNR$_{\text{SI}}$), which is defined as

$$\text{SNR}_{\text{SI}} = \frac{\text{E}\{|x(i)|^2\}}{\sigma_n^2}. \qquad (44)$$

We use the BPSK direct sequence spread spectrum signal as the far-end signal. The chip rate is 1 kHz, the spreading factor is 250. The far-end channel is assumed to be a single path channel. The far-end signal level is defined by the far-end SNR as σ_f^2/σ_n^2. Here we set SNR$_{\text{SI}} = 43$ dB, and the far-end SNR varies from 10 dB to 19 dB. The SICF is computed over

the steady-state period from 2 to 15 s, which is about ten times longer than the time correlation of the SI channel.

The performance of the SRLSd algorithm is shown in Fig. 11 (a), (c) and (e). Fig. 11 (a) shows the detection performance of the far-end data after SIC, which is an important indicator of the performance of an FD communication system. 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.
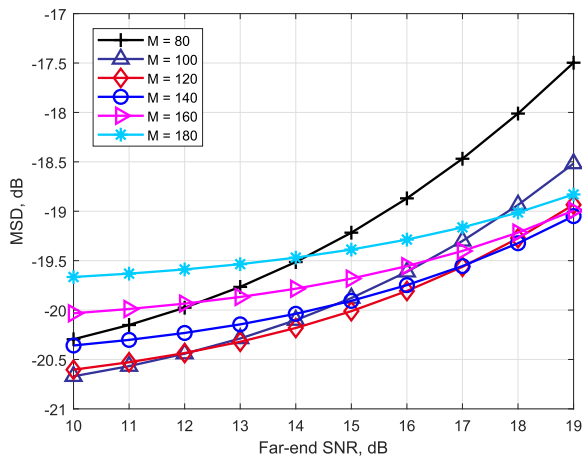
The SICF indicates that the best performance is achieved with $M = 140$ when the far-end SNR lower than 14 dB and with $M = 160$ when the far-end SNR between 14 dB and 19 dB. It is clear that the SICF provides a better indication of the optimal M for the detection performance.
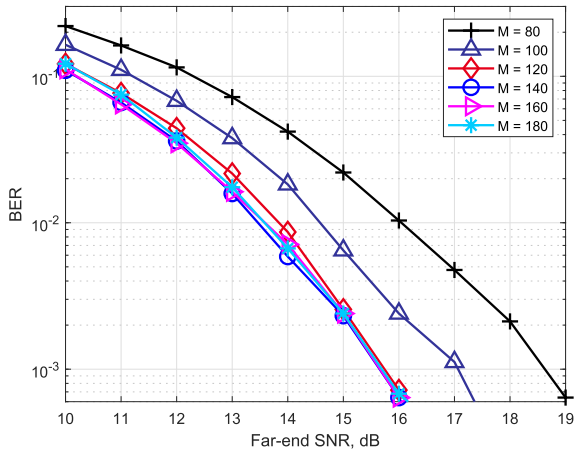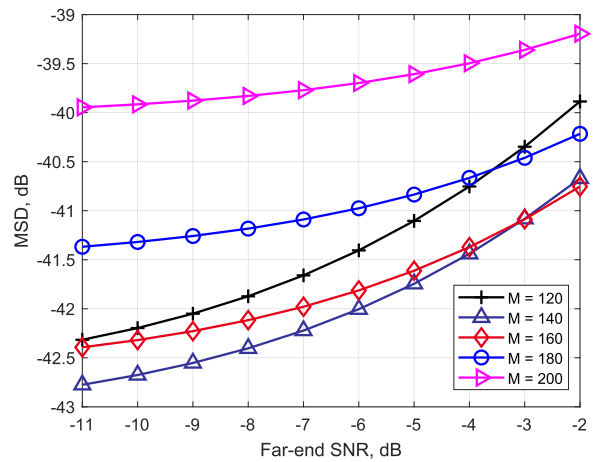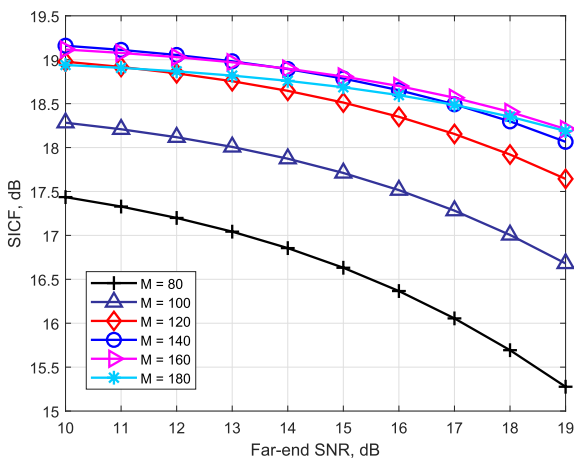
(a) BER performance for the SRLSd algorithm
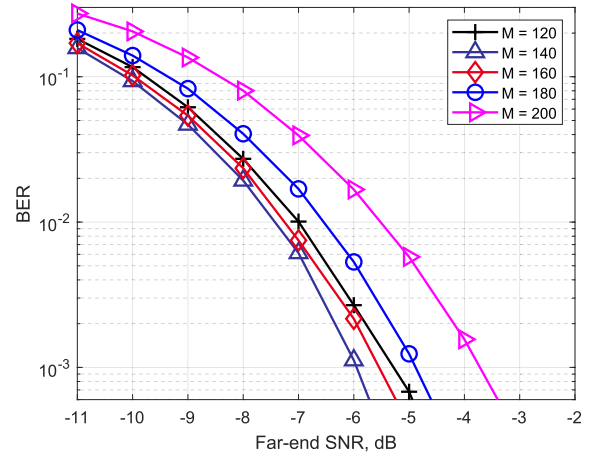
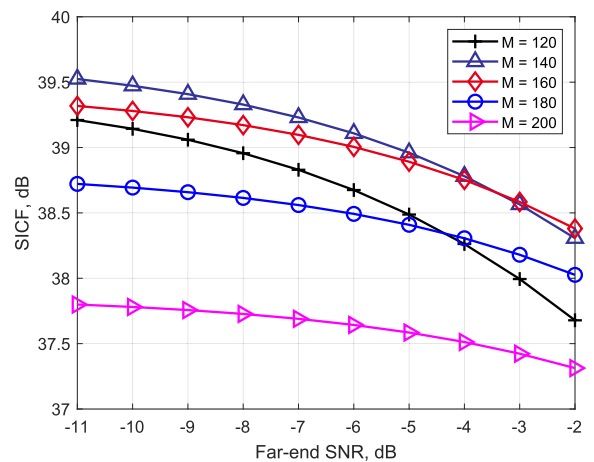(b) BER performance for the SRLS-P algorithm

(c) MSD performance for the SRLSd algorithm

(d) MSD performance for the SRLS-P algorithm

(e) SIC performance for the SRLSd algorithm

(f) SIC performance for the SRLS-P algorithm

FIGURE 11. BER, MSD and SIC performance for the SRLSd and SRLS-P algorithms in the fast varying SI channel.

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.

In Fig. 11(b), (d) and (f), the BER, MSD and SIC performance of the SRLS-P algorithm are shown. The far-end SNR now varies from -11 dB to -2 dB. We consider much lower far-end signal level compared to that used for the SRLSd algorithm to generate the BER curves, as the SIC performance is significantly improved with the SRLS-P algorithm. It is seen that the optimal detection performance is achieved with $M = 140$. The dependence between M and the BER performance is consistent with that of the MSD and the SICF. In overall, the SRLS-P algorithm with optimal M outperforms the SRLSd adaptive filter by around 20 dB in terms of the MSD and SICF. It is observed that the BER curve with the optimal M is also shifted in the far-end SNR by about the same value.

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 1.2 kHz signal bandwidth is transmitted. The symbol rate is $f_d = 1$ kHz. The BPSK symbols 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 $\mathrm{SNR_{SI}} = 100$ dB and consider the far-end SNR between 0 dB and 15 dB.

Fig. 12 shows the SIC performance of the SRLS, SRLSd and SRLS-P adaptive filters. The SIC factor is computed over the time interval from 2 s to 10 s, i.e., the average interval for computing the SICF is 8 s. 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 ($M = 240$).
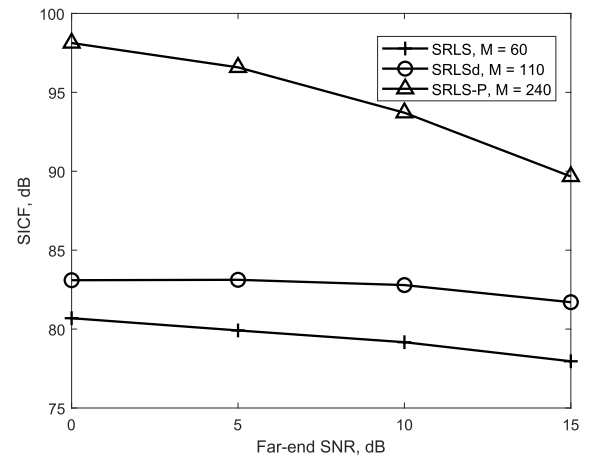

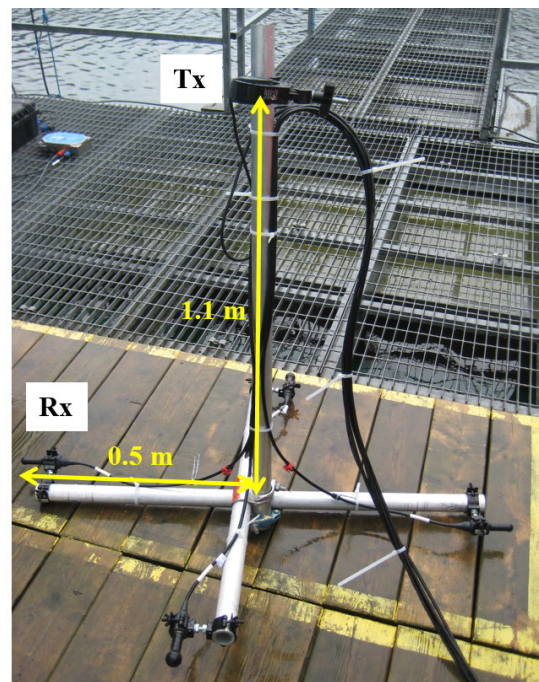**FIGURE 12.** SIC performance of adaptive filters in the passband simulation.

FIGURE 13. The configuration of the lake experiments (Tx: transducer; Rx: hydrophone).

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. In the experiment, a Zoom F4 multitrack recorder [31] with a high-resolution 24-bit ADC is used to record the PA output and the hydrophone output. The PA output is fed to the recorder through an attenuator to avoid truncation of the signal or causing damage to the recorder due to the high voltage level.

The configuration and experimental setup are shown in Fig. 13 and 14, respectively. The lake depth at the experi-
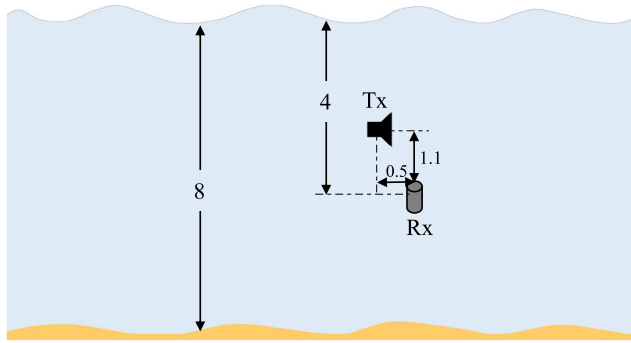
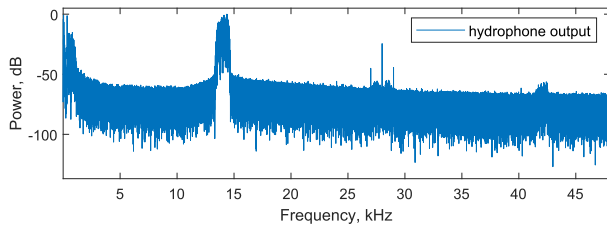**FIGURE 14.** The experimental setup. The distances are shown in meters.

FIGURE 15. Power spectrum of the hydrophone output.

FIGURE 16. Lake surface at the field.

FIGURE 17. SI channel estimate for the lake experiment.

FIGURE 18. SIC performance of adaptive filters for the lake experiment.

mental 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. The experimental site is positioned in the middle of the lake. In Fig. 16, we show a picture of the lake surface taken during the experiment. It was observed during the experiment that the amplitude of the surface waves varied from 5 cm to 10 cm. More information on the experimental site can be found in [32].

In the experiment, we transmit a 15 s BPSK signal at the carrier frequency $f_c = 14$ kHz with a bandwidth of 1.2 kHz; the symbol rate is $f_d = 1$ kHz; the pulse shaping roll-off factor is 0.2. The sampling rate is $f_s = 96$ kHz. At 14 kHz, the transmit voltage response of the transducer [33] is 118 dB re μPa/V at 1 m. During the experiment, the sound pressure level at 1 m range is around 166 dB re μPa.

In Fig. 17, 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 direct path and multiple fast time-varying paths due to reflections from the mounting system and from the lake surface and bottom. The direct path is the one associated with the highest amplitude (at tap 12). Apart from the direct path, there are also a few relatively stable reflections from the structure we used to fix the transducer and hydrophone (shown in Fig. 13). Assuming the sound speed is 1500 m/s, the delay between the direct path and the first surface reflection should be around 3.4 ms. This is consistent with the channel estimates, as the first surface reflection arrives at the 16th tap. The rest of the multipath components are due to multiple reflections from the surface, bottom and the mounting system.
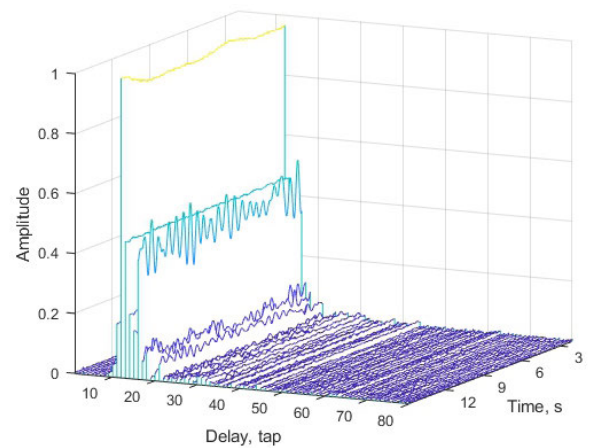
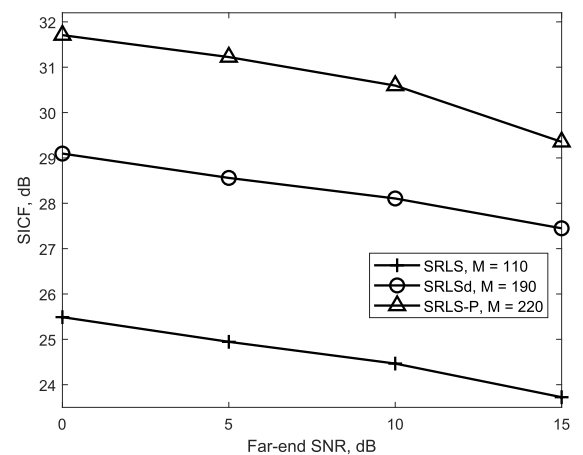In the experiment, the SI to noise ratio is around 48 dB as shown in Fig. 15. This SNR level is mostly defined by the electrical noise and acoustic noise coming to the water from the metallic construction. The filter length is $L = 80$, which

is long enough to cover the channel delay spread, including the direct path and multiple reflections from the surface and bottom. The SICF is computed over the time interval from 2 s to 15 s. Fig. 18 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 SICF.

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.

However, it is seen that even with the SRLS-P adaptive filter, the level of the residual SI is still higher than the level of the far-end signal. At 0 dB far-end SNR, with 32 dB of SICF, the residual SI is 16 dB higher than the far-end signal. At 15 dB far-end SNR, the SICF is around 29 dB, and the residual SI is 4 dB higher than the far-end signal. However, with such a level of the SI cancellation it becomes possible to detect far-end signals with specific modulation techniques, such as the spread-spectrum modulation as demonstrated in Section IV.

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. 17), has shown that $f_{max} > 2$ Hz. For the further reflections from the lake surface and bottom, as can be seen in Fig. 17, 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 [34]. 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. 10). 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. CONCLUSION AND FURTHER WORK

In this article, 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.

Although the SIC performance achieved by the SRLS-P adaptive filter is greatly improved in the simulations, the improvement of that in the experiment is not that high due to too fast surface variations. As further work, we will look into modelling the time-varying channel with higher order polynomials to improve the approximation accuracy. A full FD setup will also be considered.

APPENDIX

We now derive the presentation (26) 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

$$
\begin{aligned}
\hat{\mathbf{h}}(i) &= [\mathbf{S}^H(i)\mathbf{S}(i)]^{-1}\mathbf{S}^H(i)\mathbf{x}(i), \\
&= \mathbf{R}^{-1}(i)\mathbf{S}^H(i)\,\mathrm{diag}\{\mathbf{S}(i)\mathbf{H}(i)\}, \\
&= \mathbf{R}^{-1}(i)\mathbf{S}^H(i)\sum_{m=0}^{M-1}\mathbf{e}_m\mathbf{e}_m^T\mathbf{S}(i)\mathbf{H}(i)\mathbf{e}_m, \\
&= \frac{1}{M}\mathbf{R}^{-1}(i)\sum_{k=0}^{M-1}\mathbf{S}^H(i)\mathbf{e}_m\mathbf{e}_m^T\mathbf{S}(i)\mathbf{H}(i)\mathbf{e}_m, \quad (45)
\end{aligned}
$$

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, $\mathbf{x}(i) = \mathrm{diag}\{\mathbf{S}(i)\mathbf{H}(i)\}$. Equation (45) can be further written as:

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

$$= \frac{1}{M} \mathbf{R}^{-1}(i) \sum_{m=0}^{M-1} \mathbf{R}_{i-m} \mathbf{h}(i-m)$$

$$= \frac{1}{M} \mathbf{R}^{-1}(i) \sum_{m=-M+1}^{0} \mathbf{R}_{i+m} \mathbf{h}(i+m), \tag{46}$$

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:

$$\hat{\mathbf{h}}(i+k) = \frac{1}{M} \mathbf{R}^{-1}(i+k) \sum_{m=-M+k+1}^{k} \mathbf{R}_{i+m} \mathbf{h}(i+m). \tag{47}$$

REFERENCES

[1] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2010, pp. 1–12.

[2] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput.*, Nov. 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 *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2011, pp. 301–312.

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

[5] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 375–386, Sep. 2013.

[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 *Proc. 5th IEEE Workshop Wireless Mesh Netw.*, Jun. 2010, pp. 1–6.

[7] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Proc. 45th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Nov. 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 full-duplex underwater acoustic systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 1, pp. 192–196, Jan. 2020.

[10] B. A. Jebur, C. T. Healy, C. C. Tsimenidis, J. Neasham, and J. Chambers, "In-band full-duplex interference for underwater acoustic communication systems," in *Proc. OCEANS*, Jun. 2019, pp. 1–6.

[11] 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 Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[12] L. Ding, "Digital predistortion of power amplifiers for wireless applications," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, GA, USA, 2004.

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

[14] C. Healya, B. Jeburb, C. Tsimenidisa, J. Neashama, J. Chambersc, and C. Healy, "Full-duplex channel analysis for underwater acoustic communications," in *Proc. Underwater Acoust. Conf. Exhib.*, Crete, Greece, Jul. 2019, pp. 251–258.

[15] L. Shen, B. Henson, Y. Zakharov, and P. Mitchell, "Robust digital self-interference cancellation for full-duplex UWA systems: Lake experiments," in *Proc. Underwater Acoust. Conf. Exhib.*, Crete, Greece, Jul. 2019, pp. 243–250.

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

[17] M. Towliat, Z. Guo, L. J. Cimini, X.-G. Xia, and A. Song, "Self-interference channel characterization in underwater acoustic in-band full-duplex communications using OFDM," 2020, *arXiv:2005.10933*. [Online]. Available: http://arxiv.org/abs/2005.10933

[18] P. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation* (Kluwer International Series in Engineering and Computer Science). New York, NY, USA: Springer, 2008.

[19] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[20] L. Shen, B. Henson, Y. Zakharov, and P. D. Mitchell, "Adaptive nonlinear equalizer for full-duplex underwater acoustic systems," *IEEE Access*, vol. 8, pp. 108169–108178, 2020.

[21] G. Qiao, Y. Zhao, S. Liu, and N. Ahmed, "The effect of acoustic-shell coupling on near-end self-interference signal of in-band full-duplex underwater acoustic communication modem," in *Proc. 17th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2020, pp. 606–610.

[22] 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. 62324–62334, 2018.

[23] L. Bjørnø, *Applied Underwater Acoustics*. Amsterdam, The Netherlands: Elsevier Science, 2017.

[24] 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 Commun. Mag.*, vol. 54, no. 9, pp. 80–87, Sep. 2016.

[25] M. Emara, P. Rosson, K. Roth, and D. Dassonville, "A full duplex transceiver with reduced hardware complexity," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[26] A. Sayed, *Adaptive Filters*. Hoboken, NJ, USA: Wiley, 2011.

[27] Y. V. Zakharov, G. P. White, and J. Liu, "Low-complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3150–3161, Jul. 2008.

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

[29] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, Jan. 2009.

[30] D. B. Percival, "Simulating Gaussian random processes with specified spectra," *Comput. Sci. Statist.*, vol. 24, pp. 534–538, 1993.

[31] *Zoom F4 Multitrack Recorder*. Accessed: May 20, 2020. [Online]. Available: https://www.zoom-na.com/products/field-video-recording/field-recording/zoom-f4-multitrack-field-recorder

[32] *Acoustic Calibration Service*. Accessed: Jun. 13, 2020. [Online]. Available: http://www.neptune-sonar.co.uk/services/acoustic-calibration-service/

[33] *Low Frequency Underwater Transducer*. Accessed: May 20, 2020. [Online]. Available: https://www.benthowave.com/products/BII-7530lowfrequencytransducer.html

[34] Y. V. Zakharov, T. C. Tozer, and J. F. Adlard, "Polynomial spline-approximation of clarke's model," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1198–1208, May 2004.

LU SHEN (Graduate Student Member, IEEE) received the M.Sc. degree (Hons.) in digital signal processing from the University of York, York, U.K., in 2016. She is currently pursuing the Ph.D. degree in electronic engineering with the Communication Technologies Research Group, Department of Electronic Engineering, University of York. Her research interests include adaptive signal processing and underwater acoustic communications.

YURIY ZAKHAROV (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the Power Engineering Institute, Moscow, Russia, in 1977 and 1983, respectively. From 1977 to 1983, he was with the Special Design Agency with the Moscow Power Engineering Institute. From 1983 to 1999, he was with the N. N. Andreev Acoustics Institute, Moscow. From 1994 to 1999, he was with the Nortel as a DSP Group Leader. Since 1999, he has been with the Communication Technologies Research Group, Department of Electronic Engineering, University of York, U.K., where he is currently a Reader. His research interests include signal processing, communications, and underwater acoustics.

BENJAMIN HENSON (Member, IEEE) received the M.Eng. degree in electronic engineering and the M.Sc. degree in natural computation from the University of York, York, U.K., in 2001 and 2011, respectively, and the Ph.D. degree in electronic engineering from the Communication Technologies Research Group, University of York, in 2018. From 2002 to 2008, he was an Engineer with Snell and Wilcox Ltd., designing broadcast equipment. From 2008 to 2009, he worked for SRD Ltd., on imaging sonar designs. From 2011 to 2013, he worked on laser measurement equipment for Renishaw plc. He is currently a Research Associate with the Department of Electronic Engineering, Communication Technologies Research Group, University of York. His research interests include signal and image processing and acoustics.

NILS MOROZS (Member, IEEE) received the M.Eng. and Ph.D. degrees in electronic engineering from the University of York, in 2012 and 2015, respectively. His Ph.D. research was part of the EU FP7 ABSOLUTE project, where he developed LTE-compliant dynamic spectrum access methods for disaster relief and temporary event networks. Afterwards, he worked as a Researcher in Wi-Fi and wireless convergence at BT in Martlesham, U.K. He is currently a Research Fellow with the Department of Electronic Engineering, University of York, working on channel modeling and protocol design for underwater acoustic sensor networks. His research interests include the development of protocols and architectures for wireless radio and acoustic networks.

PAUL D. MITCHELL (Senior Member, IEEE) received the M.Eng. and Ph.D. degrees from the University of York, in 1999 and 2003, respectively. His Ph.D. research was on medium access control for satellite systems, which was supported by the British Telecom. He has over 20 years research experience in wireless communications and industrial experience gained at BT and DERA (now QinetiQ). He has been a member of academic staff with the Department of Electronic Engineering, University of York, since 2005, where he is currently a Full Professor. His primary research interests include underwater acoustic communication networks, terrestrial wireless sensor networks, and communication protocols, including the development of novel medium access control and routing strategies. Other related interests include machine learning, traffic modelling, queuing theory, satellite and mobile communication systems. He is the author of over 140 refereed journal articles and conference papers and he has served on numerous international conference programme committees including ICC and VTC. He was a General Chair of the International Symposium on Wireless Communications Systems in 2010. He currently serves as an Associate Editor of the *IET Wireless Sensor Systems Journal*, the *International Journal of Distributed Sensor Networks*, and *MDPI Electronics* and has experience as a guest editor and as a reviewer for a number of IEEE, ACM, and IET journals. He has secured more than €2.3M + €4.7M funding as principal and co-investigator. His current projects include research council grants on smart dust for large scale underwater wireless sensing, full-duplex underwater acoustic communications, as well as industrial projects. He is a member of IET.

• • •