Estimating and Tracking Wireless Channels Under Carrier and Sampling Frequency Offsets

Karel Pärlin [ID], Taneli Riihonen [ID], *Senior Member, IEEE*, Vincent Le Nir [ID], and Marc Adrat

Abstract—This article addresses the challenge of estimating and tracking wireless channels under carrier and sampling frequency offsets, which also incorporate phase noise and sampling time jitter. We propose a novel adaptive filter that explicitly estimates the channel impulse response, carrier frequency offset, and sampling frequency offset by minimizing the mean-square error (MSE) and, when the estimated parameters are time-varying, inherently performs tracking. The proposed filter does not have any requirements for the structure of the waveform, but the digital transmitted waveform must be known to the receiver in advance. To aid practical implementation, we derive upper bounds for the filter's step sizes. We also derive expressions for the filter's steady-state MSE performance, by extending the well-known energy conservation relation method to account for the self-induced nonstationarity and coupling of update equations that are inherent in the proposed filter. Theoretical findings are verified by comparison to simulated results. Proof-of-concept measurement results are also provided, which demonstrate that the proposed filter is able to estimate and track a practical wireless channel under carrier and sampling frequency offsets.

Index Terms—Adaptive filtering, frequency offset, mean-square error, steady-state analysis.

I. INTRODUCTION

OSCILLATOR inaccuracies cause two common impairments in wireless systems — mismatches between transmitter and receiver carrier generators result in a carrier frequency offset, while mismatches between sampling clocks result in a sampling frequency offset. Both of those impairments are further aggravated by the random fluctuations of oscillators and the wireless propagation. The former causes the frequency offsets to vary with time and the latter can have equivalent negative consequences due to Doppler shift. In many cases, time-varying frequency offsets can be damaging or destructive to the performance of wireless systems [1], [2].

As such, estimating and compensating frequency offsets in those cases is essential. Although not the main focus of this work, a popular example are orthogonal frequency division multiplexing (OFDM) systems, where synchronization of the carrier frequency at the receiver must be performed accurately in order to avoid loss of orthogonality between the subcarriers. Those systems can only tolerate carrier offsets that are a fraction of the spacing between the subcarriers without large degradation in performance [3], [4], [5]. The performance of OFDM systems can also degrade due to sampling frequency offsets [4], [6], although this is often less significant. Various methods for joint carrier and sampling frequency offset estimation and compensation in OFDM systems exist [7], [8], [9], [10], [11], [12], [13], [14], not to mention abundant works on only either of the offsets. However, those methods largely rely on the properties that are strictly characteristic to OFDM and are not directly applicable to other applications.

Carrier and sampling frequency offsets also pose a major challenge in known-interference cancellation. The capability to cancel known interference is a fundamental prerequisite of physical layer security schemes that envision preventing eavesdropping by superposing the signal of interest with some interference that is known only to the legitimate receiver. Perfect known-interference cancellation has been for long assumed feasible in theoretical physical layer security works without practical basis [15], [16]. However, lack of proper frequency synchronization actually has a considerable negative effect on the cancellation performance [16], [17]. This is leading to the development of interference cancellation methods with built-in frequency synchronization [18], [19].

Frequency synchronization, as well as time synchronization, is also a key issue in interference alignment and distributed beamforming. Interference alignment and distributed beamforming envision concurrent transmissions that result in a substantial increase in wireless network's total capacity [20] or an increase in range and energy efficiency [21]. In addition, since distributed beamforming entails directing more power in the desired direction, less is scattered in the undesired directions, possibly increasing security [21]. However, again the challenges in realizing the benefits of interference alignment and distributed beamforming include coordinating the transmitters for distributed information sharing plus carrier and sampling synchronization, so that the transmissions combine as necessary at the destination [22].

Bistatic radars are promising supplements to classical monostatic systems, and they too face the challenge of synchronization.

Unlike a monostatic radar, a bistatic radar has a transmitter and receiver on separate platforms which results in various operational advantages like, e.g., additional information about the scene, as the scattering characteristics of objects depend strongly on the line-of-sight vectors to the transmitter and receiver. Another advantage is the potential of cost reduction by using one transmitter, or even illuminators of opportunity, and several passive receivers [23]. However, separation of the transmit and receive platforms necessitates time and frequency synchronization for coherent signal processing and range measurement [24], [25].

Similar challenges arise in the acoustic domain. For example, in underwater acoustic communications the use of wideband modulation and low velocity of acoustic waves mean that Doppler shifts have a significantly larger impact than in the electromagnetic domain and these shifts need to be compensated for [2]. In acoustic echo control, the sampling frequency offsets between separate devices, if not compensated for, can cause poor echo cancellation performance [26], [27].

It is often so that adaptive filters are used in such nonstationary environments and consequently frequency offsets compromise the conventional filters' performance [28]. To that end, various extended adaptive filters have been proposed that are able to track certain nonstationarities or nonlinear impairments. For example, least mean squares (LMS)-type gradient descent has been used for explicit time-delay estimation [29] as well as power amplifier distortion [30] and IQ imbalance compensation [31]. An LMS-type adaptive algorithm has been proposed for joint channel estimation and explicit sampling rate correction in acoustic echo control applications [27]. The adaptive notch filter proposed in [32] is a simple algorithm capable of extracting a nonstationary narrowband signal buried in noise, being essentially a carrier frequency offset tracker. However, a single general algorithm for tracking a channel under both carrier and sampling frequency offsets, without specific requirements on the waveform, is still missing.

The purpose of this article is to present an efficient adaptive algorithm for estimating and tracking a channel under time-varying carrier and sampling frequency offsets when the receiver knows the signal that is to be transmitted, or at least a considerable part of it, in advance. The presented algorithm aims to be waveform-agnostic and not strictly rely on the characteristics of the underlying system. Hence, it is potentially applicable to the aforementioned concepts and beyond. We provide a thorough analysis on the optimal selection of the algorithm parameters (viz. three step sizes) to facilitate rapid convergence, and we carry out theoretical steady-state analysis for the proposed algorithm by extending the well-known energy conservation relation [33]. The extended relation introduces nonstationary a priori errors for each update equation and decouples the errors of separate update equations to account for the algorithm's self-induced nonstationarity. Several supporting simulations are provided, which verify the theoretical results and demonstrate that the algorithm is able to track time-varying frequency offsets. Furthermore, proof-of-concept measurement results are presented, which illustrate that the algorithm is capable of explicitly estimating and tracking a wireless channel and
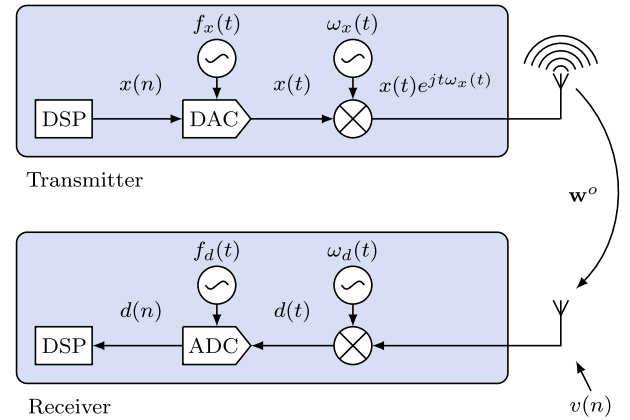

Fig. 1. General system model considered in this work, focusing on the carrier and sampling frequency offsets together with the channel impulse response between a transmitter and a receiver. In this work we assume that the digital transmitted signal $x(n)$ is known to the receiver.

frequency offsets between two radios. The proposed algorithm is positioned with regards to the existing works and comparisons are made throughout.

The rest of this article is organized as follows. Section II introduces a general system model and in Section III the novel adaptive algorithm is presented for estimating and tracking the parameters of the system model. Also, in Section III bounds for the algorithm's step sizes are derived. In Section IV, expressions are derived for the steady-state mean-square error (MSE) of the proposed algorithm, by introducing an energy conservation relation that accounts for the algorithm's self-induced nonstationarity. Section V provides a comparison of the theoretical MSE results to simulations, proof-of-concept experimental results, and a brief comparison. Finally, conclusions of the study are given in Section VI.

Notation: Small boldface letters are used to denote vectors, and capital boldface letters are used to denote matrices, e.g., $\mathbf{w}$ and $\mathbf{R}$. Furthermore, the symbol $*$ denotes Hermitian conjugation for vectors and complex conjugation for scalars. The identity matrix is denoted by $\mathbf{I}$ and a zero vector is denoted by the boldface letter $\mathbf{0}$, both with dimensions compatible to each context. The iteration index is placed as a subscript for vectors and between parentheses for scalars, e.g., $\mathbf{w}_n$ and $v(n)$. All vectors are column vectors, except for two vectors, namely, the input data vector denoted by $\mathbf{x}_n$ and its resampled counterpart $\mathbf{y}_n$, which are taken to be row vectors for convenience of notation. Lastly, $E[\cdot]$ is the statistical expectation operator.

II. SYSTEM MODEL

The system model considered in this work focuses on the time-varying sampling and carrier frequency offsets between a transmitter and a receiver along with the channel that separates the two as illustrated in Fig. 1. The relative sampling frequency offset between the two devices is denoted as $\eta^o + \beta(n)$, where $\eta^o = \Delta T/T_x$ represents the fundamental time-invariant offset with $\Delta T = 1/f_d - 1/f_x$ being the difference between the sampling periods at the receiver and transmitter, f_d is the sampling frequency at the receiver, f_x is the sampling frequency

at the transmitter, and $\beta(n)$ is the time-varying offset, including sampling jitter. The carrier frequency offset is denoted as $\epsilon^o + \phi(n)$, where ϵ^o denotes the fundamental time-invariant offset $\epsilon^o = \omega_d - \omega_x$ between the receiver and transmitter carrier frequencies, ω_d is the carrier frequency at the receiver, ω_x is the carrier frequency at the transmitter, and $\phi(n)$ is the time-varying offset, including phase noise. Lastly, we denote the finite impulse response of the complex-valued channel with order M as $\mathbf{w}^o$.

The transmitter broadcasts a complex signal $x(n)$ that, in its discrete-time form, is known to the receiver. However, due to noise, channel, and mismatches in carrier and sampling frequencies at the transmitter and the receiver, the discrete-time signal at the receiver becomes

$$d(n) = \mathbf{y}_n^o \mathbf{w}^o e^{j \sum_{i=1}^n \epsilon^o + \phi(i)} + v(n), \tag{1}$$

where $v(n)$ is the measurement noise, $\mathbf{y}_n^o$ accounts for sampling $x(t)$ with sampling frequency offset $\eta^o + \beta(n)$ so that

$$\mathbf{y}_n^o = \left[x\left(\sum_{i=1}^{n-M+1} (1 + \eta^o + \beta(i)) \right), \dots, \right.$$
$$\left. x\left(\sum_{i=1}^{n} (1 + \eta^o + \beta(i)) \right) \right] \tag{2}$$

and the multiplicative term $e^{j \sum_{i=1}^n \epsilon^o + \phi(i)}$ accounts for the carrier frequency offset.

This is a general system model that is relevant, e.g., for the following scenarios. Firstly, it holds in cases when known training data is used to estimate the channel impulse response and frequency offsets to improve subsequent information demodulation. Secondly, this general system model rather directly applies to the bistatic or multistatic radar scenario, in which case the receiver is familiar with the transmitted signal, but is interested in tracking the channel and frequency offsets to estimate range/velocity. Thirdly, in case of the known-interference cancellation scenarios, the received noise $v(n)$ can be considered to contain an unknown signal of interest, which is uncorrelated to the known signal $x(n)$ that is suppressed to facilitate processing the signal of interest.

III. ADAPTIVE ESTIMATION AND TRACKING

In order to derive an algorithm for estimating and tracking the parameters described in the system model, we first define the instantaneous error of the estimation process as

$$e(n) = d(n) - \mathbf{y}_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^n \epsilon(i-1)}, \tag{3}$$

where $\mathbf{w}_{n-1}$, $\epsilon(n-1)$, and $\eta(n-1)$ are respectively the estimates of the channel's impulse response $\mathbf{w}^o$, carrier frequency offset ϵ^o, and sampling frequency offset η^o at iteration n, and $\mathbf{y}_n$ is the result of resampling $x(n)$ with $\eta(n-1)$, so that

$$\mathbf{y}_n = \left[x\left(\sum_{i=1}^{n-M+1} (1 + \eta(i-1)) \right), \dots, \right.$$
$$\left. x\left(\sum_{i=1}^{n} (1 + \eta(i-1)) \right) \right]. \tag{4}$$

The instantaneous error $e(n)$ will contain $v(n)$ and excess noise from the algorithm's operation. In case of known-interference cancellation, the instantaneous error $e(n)$ would additionally contain some unknown signal of interest.

The aim of the adaptive filter is to update iteratively the system model parameter estimates $\mathbf{w}_n$, $\epsilon(n)$, and $\eta(n)$ so that a nonnegative cost function $J(n)$ is reduced successively

$$J(n+1) \leq J(n). \tag{5}$$

This will generally ensure that after every iteration, the adaptive filter improves its estimation of the parameters that we are trying to model.

A. Mean-Square Error

We define the cost function as the mean-square value of the estimation error, i.e., the MSE:

$$J(n) = E\left[|e(n)|^2\right] = E\left[e(n)e^*(n)\right]. \tag{6}$$

We opted for the MSE over other potential error measures, e.g., weighted least squares, because of the simplicity of the resulting algorithm. Note that in practical applications of adaptive filtering, the use of ensemble averaging is not feasible as we are adapting the filter in an on-line manner, based on a single realization of the estimation error, $e(n)$, as it evolves across iteration index n. Therefore, during the derivation of the proposed algorithm, we proceed by ignoring the expectation operation in the cost function (6) as is typical to the stochastic gradient descent method [34].

We apply the method of stochastic gradient descent for a sequential computation of the model parameters, using gradients of the performance surface in seeking its minimum. Even though only one of the estimated parameters, namely the channel impulse response $\mathbf{w}_n$, is complex-valued, then in the following derivation we also consider $\epsilon(n)$ and $\eta(n)$ to be complex-valued, as this will lay a clear consistent foundation for later carrying out the steady-state analysis of the adaptive filter. In order to accommodate for complex-valued $\epsilon(n)$ and $\eta(n)$, we use the real and imaginary part operators, $\Re\{z\}$ and $\Im\{z\}$, where appropriate.

We obtain the gradient vector at any point on the performance surface by differentiating the cost function (6) with respect to the model parameter estimates, resulting in

$$\nabla J(n) = \left[\frac{\partial J(n)}{\partial \mathbf{w}_{n-1}}, \frac{\partial J(n)}{\partial \epsilon(n-1)}, \frac{\partial J(n)}{\partial \eta(n-1)} \right], \tag{7}$$

where

$$\frac{\partial J(n)}{\partial \mathbf{w}_{n-1}} = -\left[\mathbf{y}_n e^{j \sum_{i=1}^n \Re\{\epsilon(i-1)\}} \right]^* e(n), \tag{8a}$$

$$\frac{\partial J(n)}{\partial \epsilon(n-1)} = -\left[\mathbf{y}_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^n \Re\{\epsilon(i-1)\}} j \right]^* e(n), \tag{8b}$$

$$\frac{\partial J(n)}{\partial \eta(n-1)} = -\left[\mathbf{y}_n' \mathbf{w}_{n-1} e^{j \sum_{i=1}^n \Re\{\epsilon(i-1)\}} \right]^* e(n), \tag{8c}$$

and $\mathbf{y}_n'$ is the derivative of $\mathbf{y}_n$.

When using (8b) and (8c) in practice, we are only interested in the partial derivative of a complex function $e(n)$ with respect to the real part of the parameters $\epsilon(n)$ and $\eta(n)$. Therefore, we can simplify the partial derivatives relying on the Cauchy–Riemann equations [35] and consider only the real parts of the partial derivatives so that

$$\frac{\partial J(n)}{\partial \epsilon(n-1)} = -\Im\left\{\left[\mathbf{y}_n \mathbf{w}_{n-1} e^{j\sum_{i=1}^{n}\epsilon(i-1)}\right]^* e(n)\right\}, \quad (9b)$$

$$\frac{\partial J(n)}{\partial \eta(n-1)} = -\Re\left\{\left[\mathbf{y}'_n \mathbf{w}_{n-1} e^{j\sum_{i=1}^{n}\epsilon(i-1)}\right]^* e(n)\right\}. \quad (9c)$$

B. Algorithm

We formulate the updating rules of the algorithm using the stochastic gradient in (7) by moving in the opposite direction of the gradient vector so that

$$\mathbf{w}_n = \mathbf{w}_{n-1} - \mu_w \frac{\partial J(n)}{\partial \mathbf{w}_{n-1}}, \quad (10a)$$

$$\epsilon(n) = \epsilon(n-1) - \mu_\epsilon \frac{\partial J(n)}{\partial \epsilon(n-1)}, \quad (10b)$$

$$\eta(n) = \eta(n-1) - \mu_\eta \frac{\partial J(n)}{\partial \eta(n-1)}, \quad (10c)$$

where $\mathbf{w}_0$, $\epsilon(0)$, and $\eta(0)$ are initial guesses and μ_w, μ_ϵ, and μ_η are fixed positive step size parameters that allow to control the convergence speed and steady-state performance of the algorithm. For computing the gradient vector at every iteration of the algorithm, (9b) and (9c) are to be used in (10b) and (10c). However, for carrying out the steady-state analysis, we will rely on the full complex-valued gradient and use (8b) and (8c) in (10b) and (10c); while (8a) is always used in (10a).

We also acknowledge that the partial derivative (9c) with regards to the sampling rate offset estimate $\eta(n-1)$ includes a time derivative of the resampled signal vector. If the third derivative of $\mathbf{y}_n$ exists, then it is beneficial to use the centered first-order divided difference, which has an approximation error of order two [36, p. 172], so that

$$\mathbf{y}'_n \mathbf{w}_{n-1} \approx \frac{(\mathbf{y}_{n+1} - \mathbf{y}_{n-1})\mathbf{w}_{n-1}}{2(1+\eta(n))}. \quad (11)$$

This is equivalent to considering $\mathbf{w}_{n-1}$ to be time-invariant and taking the centered first-order difference of $(\mathbf{y}_n \mathbf{w}_{n-1})'$. Alternatively, the first-order backward divided difference

$$\mathbf{y}'_n \mathbf{w}_{n-1} \approx \frac{(\mathbf{y}_n - \mathbf{y}_{n-1})\mathbf{w}_{n-1}}{1+\eta(n)} \quad (12)$$

can be used, which does not require computation of $\mathbf{y}_{n+1}$ nor the existence of the third derivative, but has an approximation error of order one.

To produce $\mathbf{y}_n$, the sampling rate of the know signal $\mathbf{x}_n$ needs to be converted. Various methods exist for arbitrary sampling rate conversion (SRC) [37], such as, e.g., the Lagrange interpolator [38], but the used SRC method can be selected independently of the proposed algorithm. If prior knowledge of the estimation parameters is available, then this knowledge may be used to speed up the start-up process of the algorithm. Otherwise, $\mathbf{w}_0$,
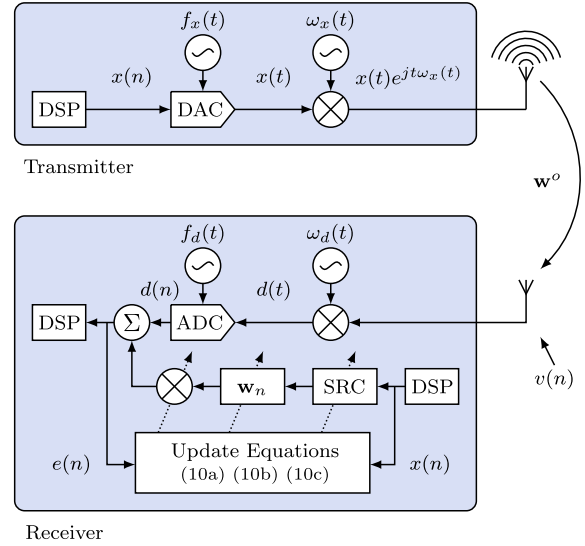

Fig. 2. System model with the proposed adaptive filter.

Algorithm 1: LMS-Type Frequency Offsets Tracking.

1: **Procedure** FO-LMS$(x, d, \mu_w, \mu_\varepsilon, \mu_\eta, M)$
2:    $\mathbf{w}_0 \leftarrow \mathbf{0}_{1,M}$
3:    $\epsilon(0) \leftarrow 0, \eta(0) \leftarrow 0$
4:    $\phi(1) \leftarrow 0, t(1) \leftarrow 0$
5:    **for** $n \leftarrow 1$ to N **do**
6:      $\mathbf{y}_n \leftarrow [x(t(n)), x(t(n) - (1+\eta(n-1))), \ldots,$
             $x(t(n) - (M+1)(1+\eta(n-1)))]$
7:      $e(n) \leftarrow d(n) - \mathbf{y}_n \mathbf{w}_{n-1} e^{j\phi(n)}$
8:      $\mathbf{w}_n \leftarrow \mathbf{w}_{n-1} + \mu_w [\mathbf{y}_n e^{j\phi(n)}]^* e(n)$
9:      $\epsilon(n) \leftarrow \epsilon(n-1) + \mu_\epsilon \Im\{[\mathbf{y}_n \mathbf{w}_{n-1} e^{j\phi(n)}]^* e(n)\}$
10:     $\eta(n) \leftarrow \eta(n-1) + \mu_\eta \Re\{[\mathbf{y}'_n \mathbf{w}_{n-1} e^{j\phi(n)}]^* e(n)\}$
11:     $\phi(n+1) \leftarrow \phi(n) + \epsilon(n)$
12:     $t(n+1) \leftarrow t(n) + (1 + \eta(n))$
13:    **end for**
14: **end procedure**

$\epsilon(0)$, and $\eta(0)$ can be initialized to zero. Conclusively, the adaptive algorithm for iteratively estimating and tracking a wireless channel under carrier and sampling frequency offsets is listed as Algorithm 1 and illustrated in Fig. 2. It should be noted that in order for the algorithm to be able to handle sampling frequency offsets, several filter taps should be allocated, i.e., $M > 1$, even if the channel itself can be modeled by a single complex coefficient. Furthermore, in general there are several equivalent formulations for complex-valued adaptive filters [39, p. 69] and corresponding equivalent formulations exist also for the proposed algorithm. An open-source implementation of the algorithm is available as part of an adaptive filters toolkit.[1]

C. Computational Cost

A useful property of the proposed algorithm, mainly due to the chosen cost function, is its computational simplicity —

---

[1] https://github.com/karel/gr-adapt

each iteration of the algorithm requires only a limited number of straightforward calculations. Evaluation of the proposed algorithm requires $12\,M + 26$ real-valued multiplications and $14\,M + 13$ real-valued additions at each iteration. There can be various ways to perform specific calculations, but the resulting overall filter complexity will be of the same order of magnitude. However, these numbers do not include the arbitrary SRC, which can be implemented in several ways with varying complexity and accuracy. For example, Lagrange interpolation can be implemented with computational complexity growing linearly with the interpolation order [40].

D. Convergence Properties

For a given system with a fixed set of parameters, the choice of step sizes μ_w, μ_ϵ, and μ_η is effectively the only way to affect the performance of the algorithm. For example, in order to speed up the initial adaptation process, it might be desirable to use large step sizes, which minimize the instantaneous error at every iteration as much as possible, yet do not cause the algorithm to diverge. An approximate way of finding the upper bounds for the step sizes of an adaptive filter is by expanding the instantaneous output error by a Taylor series expansion [41], [39, p. 86], which in this case gives

$$e(n+1) = e(n) + \frac{\partial e(n)}{\partial \mathbf{w}_{n-1}} \Delta \mathbf{w}_{n-1} + \frac{\partial e(n)}{\partial \epsilon(n-1)} \Delta \epsilon(n-1)$$
$$+ \frac{\partial e(n)}{\partial \eta(n-1)} \Delta \eta(n-1) + h.o.t., \quad (13)$$

where $\Delta \mathbf{w}_{n-1}$, $\Delta \epsilon(n-1)$, and $\Delta \eta(n-1)$ are the estimate updates and $h.o.t.$ denotes the truncated higher-order terms of the expansion. From (10a), (10b), and (10c), by considering the full complex-valued gradient vector, we get

$$\Delta \mathbf{w}_{n-1} = \mu_w e(n) \left[\mathbf{y}_n e^{j \sum_{i=1}^n \Re\{\epsilon(i-1)\}} \right]^*, \quad (14a)$$

$$\Delta \epsilon(n-1) = \mu_\epsilon e(n) \left[\mathbf{y}_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^n \Re\{\epsilon(i-1)\}} j \right]^*, \quad (14b)$$

$$\Delta \eta(n-1) = \mu_\eta e(n) \left[\mathbf{y}'_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^n \Re\{\epsilon(i-1)\}} \right]^*, \quad (14c)$$

respectively. For sufficiently small $\Delta \mathbf{w}_n$, $\Delta \epsilon(n)$, and $\Delta \eta(n)$, the values of the higher-order terms in (13) can be neglected and, therefore, in the following analysis we approximate the expansion without them. Thus, evaluating the partial derivatives in (13) and substituting in (14a), (14b), and (14c) yields after direct simplification

$$e(n+1) \approx e(n) \cdot \left(1 - \mu_w \|\mathbf{y}_n\|^2 \right.$$
$$\left. - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2 - \mu_\eta |\mathbf{y}'_n \mathbf{w}_{n-1}|^2 \right). \quad (15)$$

In order to ensure convergence, it is essential that the norm of the left hand side is not greater than that of the right hand side so that

$$|e(n+1)| \leq |e(n)| \cdot \left| 1 - \mu_w \|\mathbf{y}_n\|^2 \right.$$
$$\left. - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2 - \mu_\eta |\mathbf{y}'_n \mathbf{w}_{n-1}|^2 \right|. \quad (16)$$

The goal in (16) is reached if the following relation holds:

$$\left| 1 - \mu_w \|\mathbf{y}_n\|^2 - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2 - \mu_\eta |\mathbf{y}'_n \mathbf{w}_{n-1}|^2 \right| \leq 1, \quad (17)$$

which in turn implies the following bounds on the choice of the step sizes μ_w, μ_ϵ, and μ_η:

$$0 < \mu_w \leq \frac{2 - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2 - \mu_\eta |\mathbf{y}'_n \mathbf{w}_{n-1}|^2}{\|\mathbf{y}_n\|^2}, \quad (18a)$$

$$0 < \mu_\epsilon \leq \frac{2 - \mu_w \|\mathbf{y}_n\|^2 - \mu_\eta |\mathbf{y}'_n \mathbf{w}_{n-1}|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2}, \quad (18b)$$

$$0 < \mu_\eta \leq \frac{2 - \mu_w \|\mathbf{y}_n\|^2 - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2}{|\mathbf{y}'_n \mathbf{w}_{n-1}|^2}. \quad (18c)$$

However, the expressions above are merely necessary conditions for the stability of the proposed algorithm. The actual values of the step sizes to achieve stability are slightly smaller than the derived bounds due to the used approximation, i.e., discarding the higher-order terms in the error expansion.

We see that all quantities in (18) are positive, so the convergence properties depend on the slope but not on the sign of the gradient vector, and that the upper bounds are coupled, so the step sizes are to be selected collectively. That is, upper bound for each step size depends on the other two step sizes and convergence can be reached only if the relation in (17) is satisfied. The preceding analysis on the Taylor series expansion of the instantaneous error provides two results. Firstly, the step size bounds that are necessary but not sufficient conditions for the algorithm to converge and, secondly, these bounds can potentially be used to derive a normalized variant of the algorithm. As is, the adaptive filter assumes fixed step sizes, but an approach could also be developed that varies the step sizes to optimize convergence speed and subsequent steady-state performance.

E. Comparison

The application-specific methods for estimating a wireless channel and frequency offsets typically require the waveform to have a certain structure. The most general of those techniques aims to suppress known interference so as to provide physical layer security and relies on the waveform being cyclic with some period L [19]. Evaluation of that method for one cyclic block with length L requires $25\,L + 9$ real-valued multiplications, $18\,L - 1$ real-valued additions, $L + 1$ real-valued divisions, evaluating atan2() $L + 1$ times, and calculating the L-point discrete Fourier transform at least once. This puts the referenced and proposed methods roughly on par in terms of computational complexity for a single data point. However, due to its block-based nature, the reference method can take advantage of parallel processing. Also, methods that rely on features built into the waveform generally require fewer samples than the proposed algorithm to provide accurate parameter estimates. Then again, the repetitive waveform structure required by the reference method could be a vulnerability in physical layer security applications.

IV. STEADY-STATE ANALYSIS

An important performance measure of an adaptive filter, which is typically used in the literature, is its steady-state excess mean-square error (EMSE) [35]. In this section, we will carry out the derivation to express the total EMSE in terms of three EMSEs, each related to an update equation in (10). The analysis developed in this section relies on energy conservation arguments [33] and on decoupling the errors of separate update equations by solving a system of linear equations [42]. In order to accommodate the errors accumulated by the frequency offset update equations, we extend the existing methodology to account for what we will refer to as the self-induced nonstationarity. Furthermore, to make the analysis tractable, we omit the time-varying terms $\phi(n)$ and $\beta(n)$ of the system model here. That is, the focus is on steady-state analysis rather than tracking analysis, considering a quasi-static channel.

A. Self-Induced Nonstationarity

In practice, the frequency offset estimates $\epsilon(n)$ and $\eta(n)$ are bound to differ from the actual parameters ϵ^o and η^o, resulting in estimation errors $\tilde{\epsilon}(n) = \epsilon^o - \epsilon(n)$ and $\tilde{\eta}(n) = \eta^o - \eta(n)$. This is especially so during the start-up phase of the algorithm but also during the steady state, as gradient noise affects the estimates at each iteration. Therefore, the accumulating estimation errors $\sum_{i=1}^{n} \tilde{\epsilon}(i-1)$ and $\sum_{i=1}^{n} \tilde{\eta}(i-1)$ inevitably cause a phase shift and fractional time delay, or self-induced nonstationarity, which the channel estimate $\mathbf{w}_n$ will then try to compensate for. In order to proceed with the steady-state analysis, we first need a way to express how those accumulated estimation errors affect the channel update equation (10a).

Based on (3), we define the total a priori error as

$$e_a(n) = \mathbf{y}_n^o \mathbf{w}^o e^{j \sum_{i=1}^{n} \epsilon^o} - \mathbf{y}_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (19)$$

which is simply the error between the received signal and the estimated signal but discarding the noise term $v(n)$. By shifting both sides of the a priori error equation in phase by $-\sum_{i=1}^{n} \epsilon(i-1)$ and in time by $-\sum_{i=1}^{n} \eta(i-1)$, we get

$$e_a^s(n) = \mathbf{x}_n \mathbf{T}_n \mathbf{w}^o e^{j \sum_{i=1}^{n} \tilde{\epsilon}(i-1)} - \mathbf{x}_n \mathbf{w}_{n-1}, \quad (20)$$

where, for notational simplicity, we have denoted the phase and time shifted a priori error as

$$e_a^s(n) \triangleq e_a\left(n - \sum_{i=1}^{n} \eta(i-1)\right) e^{-j \sum_{i=1}^{n} \epsilon(i-1)} \quad (21)$$

and $\mathbf{T}_n$ is an arbitrary time-shift matrix of size $M \times M$ that, when multiplying with $\mathbf{x}_n$, delays the signal $\mathbf{x}_n$ by $\sum_{i=1}^{n} \tilde{\eta}(i-1)$. From (20), we can define $\mathbf{w}_n^o$ as

$$\mathbf{w}_n^o \triangleq \mathbf{T}_n \mathbf{w}^o e^{j \sum_{i=1}^{n} \tilde{\epsilon}(i-1)}. \quad (22)$$

In order to proceed, we call on the following assumption.

A.1: At the steady state, as $n \to \infty$, the instantaneous estimation errors $\tilde{\epsilon}(n)$ and $\tilde{\eta}(n)$ satisfy the conditions

$$\tilde{\epsilon}(n) \ll 1 \qquad \text{and} \qquad \tilde{\eta}(n) \ll \frac{1}{f_{max}},$$

where f_{max} is the maximum frequency component of $\mathbf{x}_n$.

This is a reasonable assumption because in steady state we expect the estimation errors to vary around zero. Relying on A.1, we can use linear approximation [43] to write (20) as

$$e_a^s(n) \approx [\mathbf{x}_n \mathbf{T}_{n-1} + (\mathbf{x}_n' \mathbf{T}_{n-1}) \circ \tilde{\boldsymbol{\eta}}_{n-1}] \mathbf{w}^o$$
$$\cdot \left[e^{j \sum_{i=1}^{n-1} \tilde{\epsilon}(i-1)} + e^{j \sum_{i=1}^{n-1} \tilde{\epsilon}(i-1)} j\tilde{\epsilon}(n-1) \right] - \mathbf{x}_n \mathbf{w}_{n-1}, \quad (23)$$

where $\circ$ denotes the Hadamard product, i.e., element-wise multiplication, and $\tilde{\boldsymbol{\eta}}_n$ is the row vector

$$\tilde{\boldsymbol{\eta}}_n = [\tilde{\eta}(n - M + 1), \dots, \tilde{\eta}(n-1)].$$

By expanding (23), and ignoring the cross-terms that include both $\tilde{\epsilon}(n-1)$ and $\tilde{\eta}(n-1)$, as they are very small under A.1, (23) can be rewritten as

$$e_a^s(n) \approx \mathbf{x}_n \mathbf{T}_{n-1} \mathbf{w}^o e^{j \sum_{i=1}^{n-1} \tilde{\epsilon}(i-1)}$$
$$+ \mathbf{x}_n \mathbf{T}_{n-1} \mathbf{w}^o e^{j \sum_{i=1}^{n-1} \tilde{\epsilon}(i-1)} j\tilde{\epsilon}(n-1)$$
$$+ \mathbf{x}_n' \mathbf{T}_{n-1} \mathbf{w}^o e^{j \sum_{i=1}^{n-1} \tilde{\epsilon}(i-1)} \tilde{\eta}(n-1) - \mathbf{x}_n \mathbf{w}_{n-1}, \quad (24)$$

which, by substituting in (22) for index $n - 1$, is simply

$$e_a^s(n) \approx \mathbf{x}_n \mathbf{w}_{n-1}^o + \mathbf{x}_n \mathbf{w}_{n-1}^o j\tilde{\epsilon}(n-1)$$
$$+ \mathbf{x}_n' \mathbf{w}_{n-1}^o \tilde{\eta}(n-1) - \mathbf{x}_n \mathbf{w}_{n-1}. \quad (25)$$

Finally, taking $\tilde{\mathbf{w}}_n = \mathbf{w}_n^o - \mathbf{w}_n$ to be the estimation error of the channel and reversing the phase and time shift introduced in (20), the a priori error can be expressed as

$$e_a^n(n) \approx \mathbf{y}_n \tilde{\mathbf{w}}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)}$$
$$+ \mathbf{y}_n \mathbf{w}_{n-1}^o j\tilde{\epsilon}(n-1) e^{j \sum_{i=1}^{n} \epsilon(i-1)}$$
$$+ \mathbf{y}_n' \mathbf{w}_{n-1}^o \tilde{\eta}(n-1) e^{j \sum_{i=1}^{n} \epsilon(i-1)} \quad (26)$$

and we denote the three terms on the right-hand side as the a priori errors of the three update equations so that

$$e_{w,a}^n(n) = \mathbf{y}_n \tilde{\mathbf{w}}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (27a)$$

$$e_{\epsilon,a}^n(n) = \mathbf{y}_n \mathbf{w}_{n-1}^o j\tilde{\epsilon}(n-1) e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (27b)$$

$$e_{\eta,a}^n(n) = \mathbf{y}_n' \mathbf{w}_{n-1}^o \tilde{\eta}(n-1) e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (27c)$$

where the superscript ^{n} denotes this first set of definitions for the a priori errors.

B. Mean-Square Performance

Following the well-known energy conservation relation method [28], we also define the following second set of a priori errors

$$e_{w,a}(n) = \mathbf{y}_n (\mathbf{w}_n^o - \mathbf{w}_{n-1}) e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (28a)$$

$$e_{\epsilon,a}(n) = \mathbf{y}_n \mathbf{w}_{n-1} \tilde{\epsilon}(n-1) e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (28b)$$

$$e_{\eta,a}(n) = \mathbf{y}_n' \mathbf{w}_{n-1} \tilde{\eta}(n-1) e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \quad (28c)$$

so that the total error $e(n)$ is the sum of the a priori errors and the measurement noise

$$e(n) = e_{w,a}(n) + e_{\epsilon,a}(n) + e_{\eta,a}(n) + v(n). \quad (29)$$

Similarly, we define the a posteriori errors

$$e_{w,p}(n) = \mathbf{y}_n \tilde{\mathbf{w}}_n e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \tag{30a}$$

$$e_{\epsilon,p}(n) = \mathbf{y}_n \mathbf{w}_n \tilde{\epsilon}(n) e^{j \sum_{i=1}^{n} \epsilon(i-1)}, \tag{30b}$$

$$e_{\eta,p}(n) = \mathbf{y}'_n \mathbf{w}_n \tilde{\eta}(n) e^{j \sum_{i=1}^{n} \epsilon(i-1)}. \tag{30c}$$

A.2: The noise sequence $v(n)$ is stationary, with variance σ_v^2, and statistically independent of the a priori errors $e_{w,a}(n)$, $e_{\epsilon,a}(n)$, and $e_{\eta,a}(n)$.

Under the above justifiable assumption, we find that the MSE is equivalently given by

$$MSE = \zeta + \sigma_v^2 = \zeta^w + \zeta^\epsilon + \zeta^\eta + \sigma_v^2, \tag{31}$$

where

$$\zeta^w = \lim_{n \to \infty} E\left[|e_{w,a}(n)|^2\right], \tag{32a}$$

$$\zeta^\epsilon = \lim_{n \to \infty} E\left[|e_{\epsilon,a}(n)|^2\right], \tag{32b}$$

$$\zeta^\eta = \lim_{n \to \infty} E\left[|e_{\eta,a}(n)|^2\right]. \tag{32c}$$

Employing the energy conservation relation method and relying on the two sets of a priori errors, it is shown in the Appendix that the following equations hold:

$$\mu_w E\left[\|\mathbf{y}_n\|^2 |e_{w,a}(n)|^2\right] + \mu_w E\left[\|\mathbf{y}_n\|^2 |e_{\epsilon,a}(n)|^2\right]$$
$$+ \mu_w E\left[\|\mathbf{y}_n\|^2 |e_{\eta,a}(n)|^2\right] + \mu_w E\left[\|\mathbf{y}_n\|^2 |v(n)|^2\right]$$
$$+ E\left[\frac{M}{\mu_w \|\mathbf{y}_n\|^2} |e_{\epsilon,a}^n(n)|^2\right] + E\left[\frac{M}{\mu_w \|\mathbf{y}_n\|^2} |e_{\eta,a}^n(n)|^2\right]$$
$$= 2E\left[|e_{w,a}(n)|^2\right], \tag{33a}$$

$$\mu_\epsilon E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2 |e_{w,a}(n)|^2\right] + \mu_\epsilon E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2 |e_{\epsilon,a}(n)|^2\right]$$
$$+ \mu_\epsilon E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2 |e_{\eta,a}(n)|^2\right] + \mu_\epsilon E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2 |v(n)|^2\right]$$
$$= 2E\left[|e_{\epsilon,a}(n)|^2\right], \tag{33b}$$

$$\mu_\eta E\left[|\mathbf{y}'_n \mathbf{w}_{n-1}|^2 |e_{w,a}(n)|^2\right] + \mu_\eta E\left[|\mathbf{y}'_n \mathbf{w}_{n-1}|^2 |e_{\epsilon,a}(n)|^2\right]$$
$$+ \mu_\eta E\left[|\mathbf{y}'_n \mathbf{w}_{n-1}|^2 |e_{\eta,a}(n)|^2\right] + \mu_\eta E\left[|\mathbf{y}'_n \mathbf{w}_{n-1}|^2 |v(n)|^2\right]$$
$$= 2E\left[|e_{\eta,a}(n)|^2\right]. \tag{33c}$$

This system of equations can now be solved for the EMSEs ζ^w, ζ^ϵ, and ζ^η. To do so, we consider the following two cases.

1) Using Separation Principle: One way to solve the equations in (33) is by imposing the following assumption.

A.3: In steady state, $\|\mathbf{y}_n\|^2$, $|\mathbf{y}_n \mathbf{w}_{n-1}|^2$, and $|\mathbf{y}'_n \mathbf{w}_{n-1}|^2$ are statistically independent of $|e_{w,a}|^2$, $|e_{\epsilon,a}(n)|^2$, and $|e_{\eta,a}(n)|^2$.

This assumption is reasonable at the steady state since the behavior of the a priori errors is less likely to be sensitive to the input data. It is similar to the separation principle assumption made in, e.g., [33], [42], and allows us to write

$$E\left[\|\mathbf{y}_n\|^2 |e_{w,a}(n)|^2\right] = E\left[\|\mathbf{y}_n\|^2\right] E\left[|e_{w,a}(n)|^2\right],$$

$$E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2 |e_{\epsilon,a}(n)|^2\right] = E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2\right] E\left[|e_{\epsilon,a}(n)|^2\right],$$

$$E\left[|\mathbf{y}'_n \mathbf{w}_{n-1}|^2 |e_{\eta,a}(n)|^2\right] = E\left[|\mathbf{y}'_n \mathbf{w}_{n-1}|^2\right] E\left[|e_{\eta,a}(n)|^2\right].$$

Furthermore, we make the following assumptions.

A.4: In steady state for a static channel, as $n \to \infty$, the channel estimate is close to the actual channel $\mathbf{w}_{n-1} \to \mathbf{w}^o$.

A.5: In steady state, for sufficiently small η^o, the following equalities hold: $\|\mathbf{y}_n\|^2 = \|\mathbf{x}_n\|^2$, $|\mathbf{y}_n \mathbf{w}_{n-1}|^2 = |\mathbf{x}_n \mathbf{w}^o|^2$ and $|\mathbf{y}'_n \mathbf{w}_{n-1}|^2 = |\mathbf{x}'_n \mathbf{w}^o|^2$.

A.6: In steady state, the two sets of a priori errors are equivalent, i.e., $E|e_{w,a}^n(n)|^2 = E|e_{w,a}(n)|^2$, $E|e_{\epsilon,a}^n(n)|^2 = E|e_{\epsilon,a}(n)|^2$, $E|e_{\eta,a}^n(n)|^2 = E|e_{\eta,a}(n)|^2$.

Using the assumptions A.3 through A.6, and solving (33) for ζ^w, ζ^ϵ, and ζ^η, we obtain the following expressions for the EMSEs of the proposed algorithm:

$$\zeta^w = \frac{2\mu_w \operatorname{Tr}(\mathbf{R})\sigma_v^2}{\gamma}$$
$$+ \frac{M\frac{\mu_\epsilon \operatorname{Tr}(\mathbf{RQ})}{\mu_w \operatorname{Tr}(\mathbf{R})}\sigma_v^2 + M\frac{\mu_\eta \operatorname{Tr}(\mathbf{PQ})}{\mu_w \operatorname{Tr}(\mathbf{R})}\sigma_v^2}{\gamma}, \tag{35a}$$

$$\zeta^\epsilon = \frac{2\mu_\epsilon \operatorname{Tr}(\mathbf{RQ})\sigma_v^2}{\gamma}, \tag{35b}$$

$$\zeta^\eta = \frac{2\mu_\eta \operatorname{Tr}(\mathbf{PQ})\sigma_v^2}{\gamma}, \tag{35c}$$

where the denominator γ is

$$\gamma = 4 - 2\mu_w \operatorname{Tr}(\mathbf{R}) - 2\mu_\epsilon \operatorname{Tr}(\mathbf{RQ}) - 2\mu_\eta \operatorname{Tr}(\mathbf{PQ})$$
$$- M\frac{\mu_\epsilon \operatorname{Tr}(\mathbf{RQ})}{\mu_w \operatorname{Tr}(\mathbf{R})} - M\frac{\mu_\eta \operatorname{Tr}(\mathbf{PQ})}{\mu_w \operatorname{Tr}(\mathbf{R})} \tag{36}$$

and $\mathbf{R}$ is the covariance matrix $\mathbf{R} = E[\mathbf{x}_n^* \mathbf{x}_n]$, $\mathbf{P}$ is the covariance matrix $\mathbf{P} = E[(\mathbf{x}'_n)^* \mathbf{x}'_n]$, and $\mathbf{Q} = \mathbf{w}^o (\mathbf{w}^o)^*$.

Note that, in order for the algorithm to remain stable, the denominator of the EMSEs needs to be positive. If we consider an approximation of the denominator without the self-induced nonstationarity terms, i.e., the last two terms in (36), then this result has an equivalent implication to that of the simple approximation (18), which we derived using the Taylor series expansion of the instantaneous error.

2) Assuming Gaussian White Input Signals: For Gaussian white input signals (with $\mathbf{R} = \sigma_x^2 \mathbf{I}$), relying on A.4 and A.5, (33) can be more accurately solved by resorting to the following independence assumption.

A.7: At steady state, the estimation errors $\tilde{\mathbf{w}}_n$, $\tilde{\epsilon}(n)$, and $\tilde{\eta}(n)$ are all statistically independent of $\mathbf{x}_n$, $\mathbf{x}_n \mathbf{w}^o$, and $\mathbf{x}'_n \mathbf{w}^o$.

This is an extension of the assumption, which is widely used for analysing the performance of adaptive filters [33]. Relying on the independence assumption A.7 and following the same reasoning that is used for analysing the steady-state performance of the LMS adaptive filter [35, p. 296], it can be verified that

$$E\left[\|\mathbf{x}_n\|^2 |e_{k,a}(n)|^2\right] = (M+1)\sigma_x^2 \zeta^k, \tag{37a}$$

$$E\left[|\mathbf{x}_n \mathbf{w}^o|^2 |e_{k,a}(n)|^2\right] = \left(1 + \frac{1}{M}\right)\sigma_x^2 \|\mathbf{w}^o\|^2 \zeta^k, \tag{37b}$$

$$E\left[|\mathbf{x}'_n \mathbf{w}^o|^2 |e_{k,a}(n)|^2\right] \approx \left(2 + \frac{2}{M}\right)\sigma_x^2 \|\mathbf{w}^o\|^2 \zeta^k, \tag{37c}$$

where k is either w, ϵ, or η, and we have used that

$$\sigma_{x'}^2 = \frac{\sigma_{x(n)}^2 + \sigma_{x(n-1)}^2}{(\Delta n)^2} = 2\sigma_x^2,$$

where we are first relying on the independence of successive samples of $x(n)$ and consequently on the samples being identically distributed as well. The last equation in (37) does not hold precisely but is an approximation, because the derivative itself is not identically and independently distributed. Still, for a channel $\mathbf{w}^o$ with a relatively flat frequency response, this approximation can be practical, as will be seen in the results section. Using A.4, A.5, A.7, and solving (33) for ζ^w, ζ^ϵ, and ζ^η, we obtain

$$\zeta^w = \frac{2\mu_w M \sigma_x^2 \sigma_v^2}{\gamma} \tag{38a}$$
$$+ \frac{\frac{\mu_\epsilon}{\mu_w M}\|\mathbf{w}^o\|^2\sigma_v^2 + \frac{\mu_\eta}{\mu_w M}2\|\mathbf{w}^o\|^2\sigma_v^2}{\gamma},$$

$$\zeta^\epsilon = \frac{2\mu_\epsilon \sigma_x^2 \|\mathbf{w}^o\|^2 \sigma_v^2}{\gamma}, \tag{38b}$$

$$\zeta^\eta = \frac{4\mu_\eta \sigma_x^2 \|\mathbf{w}^o\|^2 \sigma_v^2}{\gamma}, \tag{38c}$$

where the denominator γ is the same for all equations:

$$\gamma = 4 - 2\mu_w(M+1)\sigma_x^2$$
$$- 2\mu_\epsilon\left(1 + \frac{1}{M}\right)\sigma_x^2\|\mathbf{w}^o\|^2 - \frac{\mu_\epsilon}{\mu_w}\left(1 + \frac{1}{M}\right)\|\mathbf{w}^o\|^2$$
$$- 4\mu_\eta\left(1 + \frac{1}{M}\right)\sigma_x^2\|\mathbf{w}^o\|^2 - 2\frac{\mu_\eta}{\mu_w}\left(1 + \frac{1}{M}\right)\|\mathbf{w}^o\|^2. \tag{39}$$

V. NUMERICAL RESULTS

In order to verify the theoretical steady-state MSE expressions and evaluate the performance of the proposed algorithm, the theoretical results are herein first compared to steady-state simulations, where the channel and frequency offsets are assumed to be known to the algorithm and time-varying terms are omitted, and then time-varying simulations together with proof-of-concept RF measurements are presented.

A. Steady-State Results

In Fig. 3, the steady-state theoretical MSEs obtained from expressions (35) and (38) are compared with the MSE observed in simulations. The simulations are run with different channel weight vectors $\mathbf{w}^o$, each of length $M = 3$ with a rather flat frequency response. The input signal $\mathbf{x}_n$ is Gaussian of unit variance and the noise $v(n)$ is Gaussian with variance $\sigma_v^2 = 10^{-3}$. From here on out, in order to make the results relatable, we refer to the sampling frequency offset as $\Delta f = f_d - f_x$ instead of η^o. The simulated frequency offsets are $\epsilon^o = 6\,\text{kHz}$ and $\Delta f = 5\,\text{Hz}$, which, considering a carrier frequency of $2.4\,\text{GHz}$ and sampling frequency of $2\,\text{MHz}$, is equivalent to a $2.5\,\text{ppm}$ oscillator inaccuracy. Note that the MSE expressions do not depend
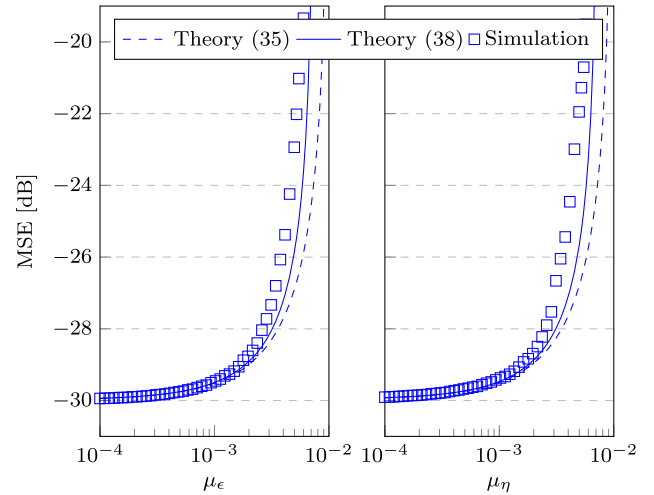

Fig. 3. Simulated and theoretical MSE curves relying on the separation principle and Gaussian input versus μ_ϵ for $\mu_w = 0.0025$ and $\mu_\eta = 0$ on the left and versus μ_η for $\mu_w = 0.005$ and $\mu_\epsilon = 0$ on the right.

on the frequency offset values, since the offsets themselves inherently do not affect the energy conservation relation. This is in alignment with our extensive simulation results for practical ranges of ϵ^o and Δf (that are not shown herein), as steady-state MSE is indifferent w.r.t. the offset values. Thus, the simulated results are only plotted for these two example frequency offsets.

Each simulation result is the steady-state statistical average of 1024 runs, with 5000 iterations in each run. The average of the last 2500 entries of the ensemble-average curve is then used as the simulated MSE value. Oversampling is used to prevent interpolation errors from skewing the simulation results. In Fig. 3, the analysis focuses separately on either frequency offset estimation combined with the channel estimation. The comparison shows that both expressions are in good match with simulation results at small values of μ_ϵ and μ_η. However, (38) gives a better match with the simulation results for larger μ_ϵ and μ_η values, which supports the use of A.7.

Figs. 4 and 5 compare the theoretical MSE obtained from (38) with the simulated MSE for various μ_w over a range of μ_ϵ or μ_η. Again, the results show a good match between theoretical and simulated results, especially at smaller step size values, when the steady-state assumptions are better justified. However, in general the sampling frequency offset update equation is not well suited for operating with disproportionally selected step sizes — carrier frequency offset can usually be recovered, but if the signals become unaligned in time because of persisting large estimation errors in sampling frequency offset, then this can be difficult to recover from.

Furthermore, Figs. 4 and 5 also illustrate the relevance of the the step sizes' upper bound (18). For visual clarity, only a single upper bound is calculated and plotted by taking the two step sizes, which are varied, to be equal in (18). As the step sizes approach the upper bound, performance of the filter deteriorates, and, since the filter leaves the steady state, the match between theoretical and simulated MSE results also declines. Finally, Fig. 6 presents a comparison of the theoretical and
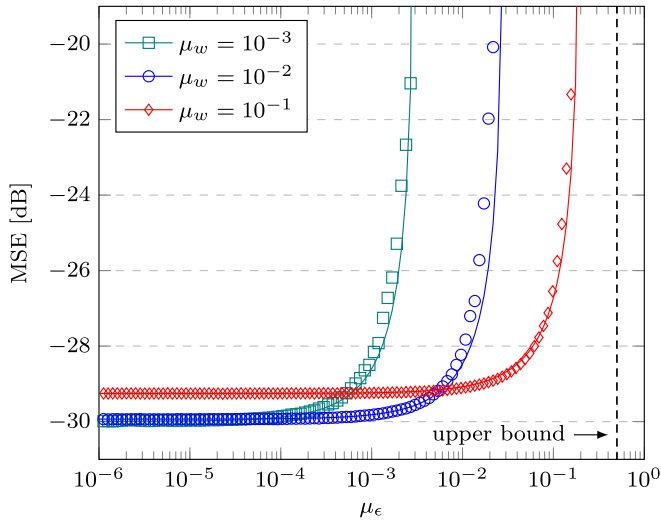
Fig. 4. Simulated (only markers) and theoretical (solid lines) MSE curves at various μ_w versus μ_ϵ for $\mu_\eta = 0$. The dashed vertical line indicates the upper bound for the two step sizes when $\mu_w = \mu_\epsilon$.
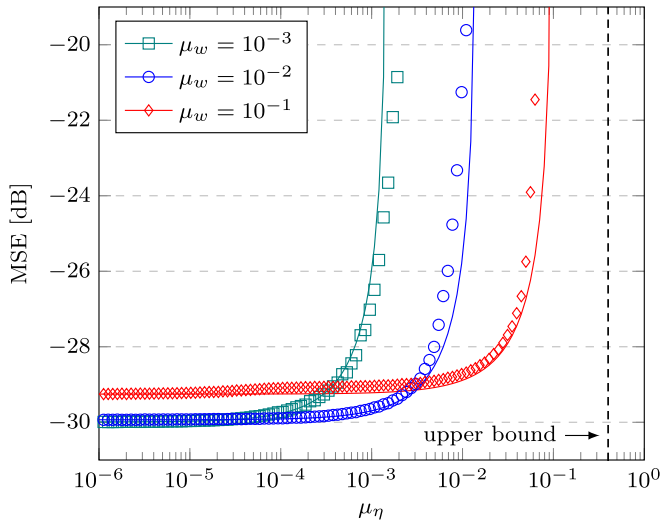

Fig. 6. Simulated (only markers) and theoretical (solid lines) MSE curves at various μ_ϵ versus μ_η for $\mu_w = 10^{-3}$.

Fig. 5. Simulated (only markers) and theoretical (solid lines) MSE curves at various μ_w versus μ_η for $\mu_\epsilon = 0$. The dashed vertical line indicates the upper bound for the two step sizes when $\mu_w = \mu_\eta$.

Fig. 7. Simulated results illustrating the filter's ability to track time-varying frequency offsets at step sizes $\mu_w = 10^{-3}$, $\mu_\epsilon = 10^{-6}$, and $\mu_\eta = 10^{-6}$.

simulated MSEs of the proposed algorithm when all of the system parameters are simultaneously estimated. As shown by all the foregoing numerical results in Figs. 3–6, the theoretical results match very well with the simulations.

B. Time-Varying Results

In this subsection, using simulations, we analyze the performance of the proposed filter when the frequency offsets are time-varying, i.e., we focus on the effect of $\phi(n)$ and $\beta(n)$ on the algorithm's performance. Fig. 7 illustrates the filter's ability to track long-term changes in the time-varying terms. The simulations are started with perfect knowledge about the initial state of the channel and with zero frequency offsets. Then both frequency offsets are varied over time either gradually
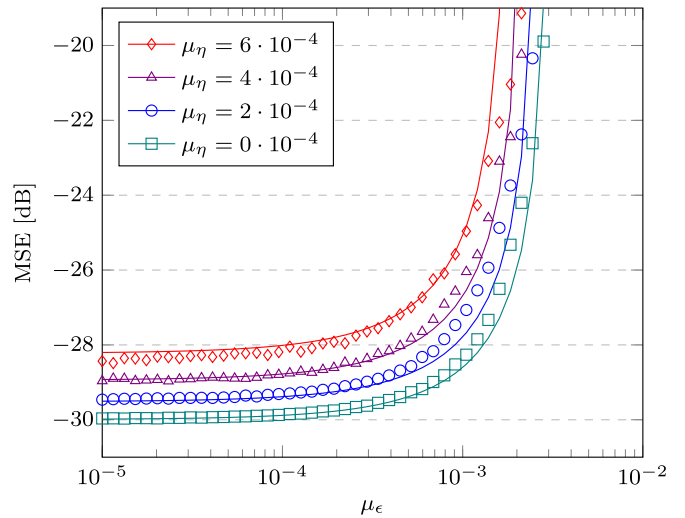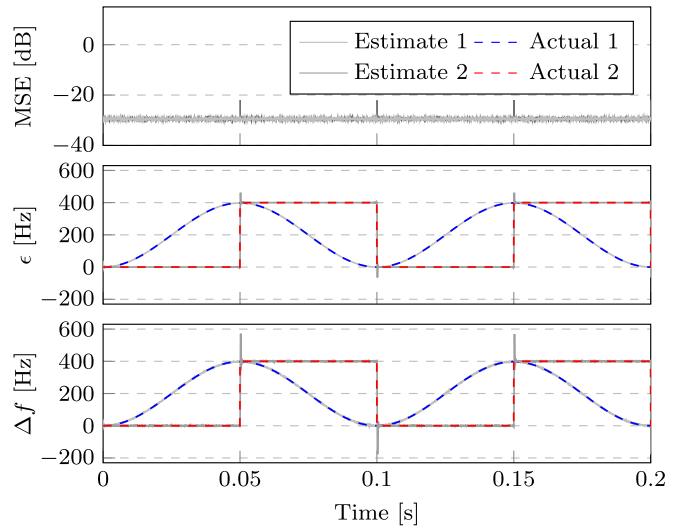
or abruptly as shown in Fig. 7 with the dashed lines. Other simulation parameters are kept the same as previously, including the ensemble averaging. The simulation results indicate that the adaptive filter is able to track those changes, regardless of whether the parameters change gradually or abruptly. As a result, the MSE is stable over time, except for a brief readjustment period during the abrupt frequency offset changes, which is expected.

In contrast, Fig. 8 demonstrates the filter's tracking performance under short-term changes, i.e., phase noise and sampling time jitter. Both are modelled as first-order autoregressive processes with the process parameters α_ϕ and α_β close to one and the variances being σ_ϕ^2 and σ_β^2 (the exact values of which are given in Fig. 8). The algorithm is run for 10^6 iterations and, again, the simulations are started with perfect knowledge of the initial state of the channel, yet without knowledge about
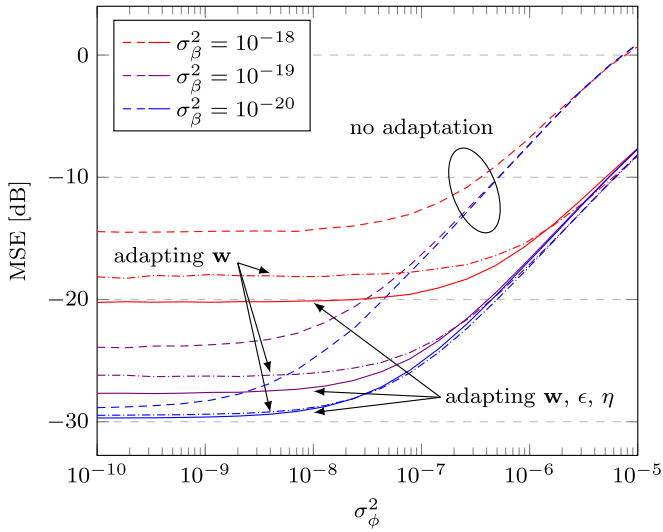
Fig. 8. Simulated performance for various sampling jitter and phase noise variances when $\epsilon^o = 0$ and $\eta^o = 0$ at step sizes $\mu_w = 0$, $\mu_\epsilon = 0$, and $\mu_\eta = 0$ (dashed lines), $\mu_w = 10^{-3}$, $\mu_\epsilon = 0$, and $\mu_\eta = 0$ (dash dotted lines), and $\mu_w = 10^{-3}$, $\mu_\epsilon = 10^{-5}$, and $\mu_\eta = 10^{-5}$ (solid lines).
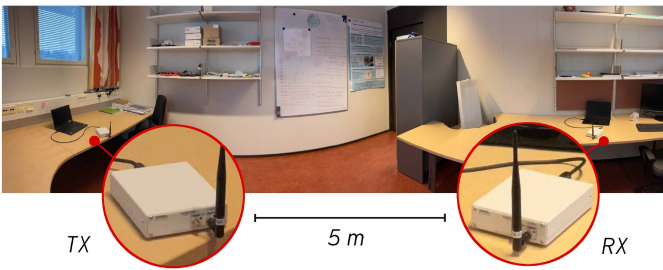

Fig. 9. Experiment setup in an office room with two USRP-2900s.

the noise processes. The simulations illustrate three cases: no adaptation at all, adaptation of only the channel estimate $\mathbf{w}_n$, and adaptation of all the parameters. The case without adaptation serves as a baseline for the MSE performance in the given noisy circumstances, while the other cases illustrate the benefits of adapting the channel and frequency offset estimates. The results show that, even though excessive phase noise and sampling jitter can degrade the algorithm's performance, adapting all the parameters still has a clear benefit compared to limited or no adaptation.

C. Experimental Results

The experiment is carried out indoors using two USRP-2900 software-defined radios with dipole antennas. The radios have internal temperature-compensated crystal oscillators with frequency accuracy of couple parts per million, presenting a fair scenario for analyzing the algorithm. The radios are positioned in the opposite corners of an office room with about five meters line-of-sight distance between them as shown in Fig. 9. As such, the experimental setup is static, with only the inherent oscillator drifts contributing a slowly time-varying component. The measurements are done in a relatively quiet section of the $2.4\,\mathrm{GHz}$ ISM frequency band, so that signals from other wireless devices
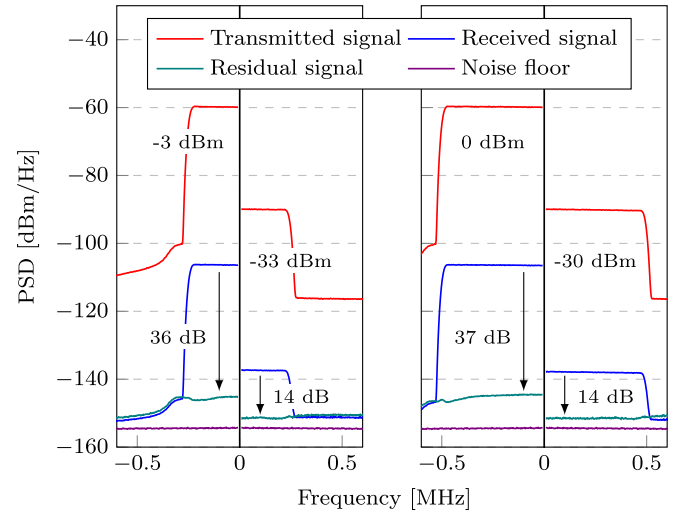

Fig. 10. Power spectral densities of the transmitted, received, and residual signals along with the noise floor at the receiver in steady state, i.e., discarding the start-up phase of the algorithm.

do not affect the measurements, and using a sampling rate of $2\,\mathrm{MHz}$. The transmitter broadcasts a bandlimited Gaussian noise signal, which is known to the receiver entirely. As such, the experiment illustrates the known-interference cancellation scenario, where the residual error signal could contain a signal of interest. Two signals bandwidths, $1\,\mathrm{MHz}$ and $0.5\,\mathrm{MHz}$, are used with transmit powers $-60\,\mathrm{dBm/Hz}$ or $-90\,\mathrm{dBm/Hz}$. The receiving node receives the bandlimited noise signal over the air and records it. The algorithm is then run offline on the recordings.

Length of the estimated channel vector $\mathbf{w}_n$ is taken to be $M = 9$, which is more than sufficient for this scenario, and all of the estimated parameters are initialized to zero. For the algorithm to converge, it is required that the known and received signal streams be coarsely aligned in time (i.e., the difference in the two streams' starts may not exceed $M - 1$ samples). That coarse alignment is provided by onset detection — comparing the received signal's energy to a threshold. Fig. 10 shows the measured signal spectra at different stages of the system model. It can be observed that suppression of the known interference is not significantly affected by its bandwidth. Furthermore, when the received known interference is substantially above the noise floor then the MSE, i.e., the residual signal, is much higher than the measurement noise floor. This is caused by the nonlinearities induced in the USRP-2900 RF front-ends, which the algorithm does not account for. When the received known interference is not so powerful, those nonlinearities do not affect cancellation. Based on measurements at other received known-interference power levels that are omitted for brevity, in this scenario the algorithm requires that the signal be at least $4\,\mathrm{dB}$ above the noise floor in order to provide stable parameter estimates.

Finally, Fig. 11 demonstrates the algorithm's performance for the purpose of known-interference cancellation while estimating and tracking the channel together with the frequency offsets (Residual 2 and 3) as opposed to estimating and tracking the channel without compensating for the frequency offsets (Residual 1). It is evident that explicit adaptation of frequency
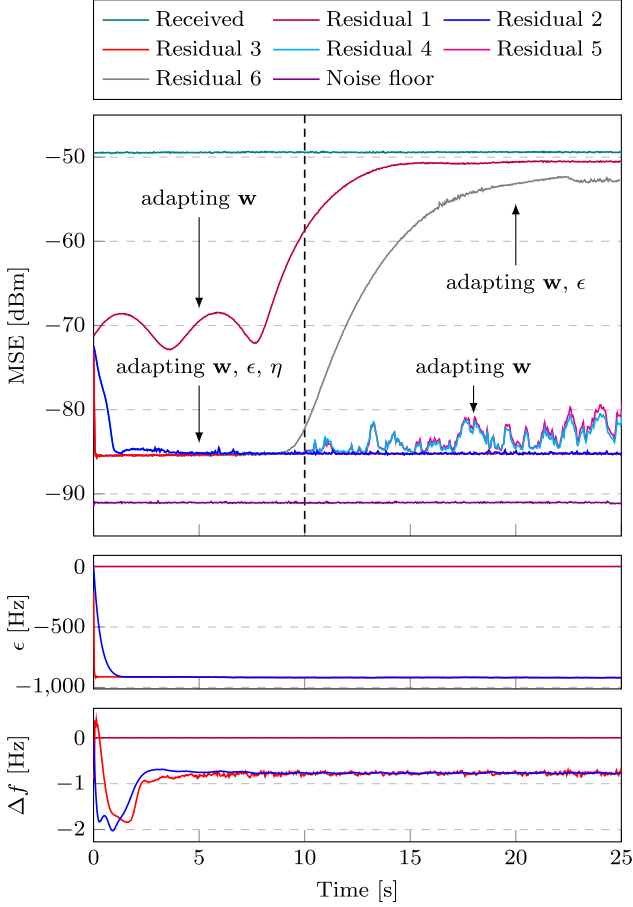
Fig. 11. Proposed algorithm's MSE progression in time for the 0.5 MHz −60 dBm/Hz signal. Residual 1 is without frequency offset compensation ($\mu_\epsilon = 0$ and $\mu_\eta = 0$); Residual 2 uses larger step sizes for frequency offset updates ($\mu_\epsilon = 5 \cdot 10^{-5}$ & $\mu_\eta = 5 \cdot 10^{-5}$) and Residual 3 uses smaller step sizes ($\mu_\epsilon = 2 \cdot 10^{-6}$ & $\mu_\eta = 2 \cdot 10^{-6}$); Residual 4 and 5 supplement cases 2 and 3, as the frequency offset estimations are stopped after 10 s. Residual 6 illustrates the situation without sampling frequency offset compensation.

offsets gives better short-term and long-term performance. The results also show how continuous frequency offsets tracking is necessary in practice (Residual 4, 5 and 6), due to their time-varying nature. Again, it is clear that the experimental MSE does not reach the noise floor, as excessive phase noise, sampling time jitter, and nonlinear distortions degrade the performance of the algorithm. Nevertheless, the experimental results demonstrate the efficiency of the proposed algorithm in estimating and compensating for time-varying carrier and sampling frequency offsets of an unknown channel.

We compared the proposed algorithm to the method in [19] using a separate set of measurements with a cyclic bandlimited Gaussian noise waveform having period L. The two algorithms achieved a similar level of MSE eventually as long as the period L was chosen so that the carrier frequency offset remained within the reference algorithm's estimation range. As such, only the proposed algorithm's results are presented in the figures for brevity. The reference algorithm does have an advantage over the proposed algorithm in that it provides estimates of the channel and frequency offsets quicker. However, this advantage

of the reference method relies on the assumptions that the used waveform is cyclic with period L and the combination of period L and sampling rate is appropriate for the frequency offsets. The latter of which significantly limits the acceptable range of L. The proposed algorithm, however, is not limited to cyclic waveforms and, as such, is also free from the related estimation range and accuracy limitations.

VI. Conclusion

This article proposed an adaptive filter for jointly and explicitly estimating the channel impulse response, carrier frequency offset, and sampling frequency offset between a transmitter and receiver pair. The proposed algorithm relies on the stochastic gradient descent method minimizing the mean-square error and is therefore computationally simple, yet effective. Compared to existing methods, the proposed adaptive filter facilitates estimating the channel and frequency offsets without requirements on the used waveform. Stability and convergence of the algorithm depend on the proper selection of step sizes in relation to the other system parameters. Hence, upper bounds for the step sizes were derived and presented. Furthermore, this article also provides a theoretical steady-state analysis of the proposed adaptive filter. Novel expressions for the excess mean-square error were derived by extending the energy conservation relation to account for the self-induced nonstationarity inherent in the proposed adaptive filter. Validity of the theoretical expressions was corroborated through comparison to simulations. Also, simulation results were presented for time-varying and noisy frequency offsets. Finally, the algorithm was validated on measurement data.

Appendix

The following analysis extends the energy conservation relation [33], which is established by expressing the update equations in (10) in terms of the estimation errors $\tilde{\mathbf{w}}_n$, $\tilde{\epsilon}(n)$, and $\tilde{\eta}(n)$. Subtracting both sides of (10a) from $\mathbf{w}_n^o$, both sides of (10b) from ϵ^o, and both sides of (10c) from η^o, we get

$$\tilde{\mathbf{w}}_n = \mathbf{w}_n^o - \mathbf{w}_{n-1} - \mu_w \left[\mathbf{y}_n e^{j \sum_{i=1}^{n} \epsilon(i-1)} \right]^* e(n), \quad (40\text{a})$$

$$\tilde{\epsilon}(n) = \tilde{\epsilon}(n-1)$$
$$- \mu_\epsilon \left[\mathbf{y}_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)} j \right]^* e(n), \quad (40\text{b})$$

$$\tilde{\eta}(n) = \tilde{\eta}(n-1)$$
$$- \mu_\eta \left[\mathbf{y}_n' \mathbf{w}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)} \right]^* e(n). \quad (40\text{c})$$

Furthermore, by multiplying both sides of equation (40a) with $\mathbf{y}_n e^{j \sum_{i=1}^{n} \epsilon(i-1)}$ from the left, (40b) with $\mathbf{y}_n \mathbf{w}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)}$, and (40c) with $\mathbf{y}_n' \mathbf{w}_{n-1} e^{j \sum_{i=1}^{n} \epsilon(i-1)}$, we see that the a priori (28) and a posteriori (30) estimation errors are related via

$$e_{w,p}(n) = e_{w,a}(n) - \mu_w \|\mathbf{y}_n\|^2 e(n), \quad (41\text{a})$$

$$e_{\epsilon,p}(n) = e_{\epsilon,a}(n) - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2 j^* e(n), \quad (41\text{b})$$

$$e_{\eta,p}(n) = e_{\eta,a}(n) - \mu_\eta |\mathbf{y}_n' \mathbf{w}_{n-1}|^2 e(n). \quad (41\text{c})$$

Equations (40) and (41) provide an alternative representation of the adaptive filter in terms of the error quantities. This is useful, as it will allow relating the steady-state behavior of these errors. So, rearranging (41a), (41b), and (41c) allows us to express the total error $e(n)$ separately in terms of the three sets of a priori and a posteriori errors:

$$e(n) = \frac{1}{\mu_w \|\mathbf{y}_n\|^2} \left[e_{w,a}(n) - e_{w,p}(n) \right], \quad (42a)$$

$$e(n) = \frac{1}{\mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2} \left[e_{\epsilon,a}(n) - e_{\epsilon,p}(n) \right], \quad (42b)$$

$$e(n) = \frac{1}{\mu_\eta |\mathbf{y}_n' \mathbf{w}_{n-1}|^2} \left[e_{\eta,a}(n) - e_{\eta,p}(n) \right]. \quad (42c)$$

Substituting the right-hand sides of the above into (40a), (40b), and (40c), gives respectively

$$\tilde{\mathbf{w}}_n = \mathbf{w}_n^o - \mathbf{w}_{n-1} - \frac{\mathbf{y}_n^*}{\|\mathbf{y}_n\|^2} \left[e_{w,a}(n) - e_{w,p}(n) \right], \quad (43a)$$

$$\tilde{\epsilon}(n) = \tilde{\epsilon}(n-1) - \frac{(\mathbf{y}_n \mathbf{w}_{n-1})^*}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2} \left[e_{\epsilon,a}(n) - e_{\epsilon,p}(n) \right], \quad (43b)$$

$$\tilde{\eta}(n) = \tilde{\eta}(n-1) - \frac{(\mathbf{y}_n' \mathbf{w}_{n-1})^*}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2} \left[e_{\eta,a}(n) - e_{\eta,p}(n) \right], \quad (43c)$$

where on each side those identities, we have a combination of a priori and a posteriori errors, while the step sizes cancel out. By evaluating the energies of both sides, we find that the following energy equalities hold:

$$\|\tilde{\mathbf{w}}_n\|^2 + \frac{|e_{w,a}(n)|^2}{\|\mathbf{y}_n\|^2} = \|\mathbf{w}_n^o - \mathbf{w}_{n-1}\|^2 + \frac{|e_{w,p}(n)|^2}{\|\mathbf{y}_n\|^2}, \quad (44a)$$

$$|\tilde{\epsilon}(n)|^2 + \frac{|e_{\epsilon,a}(n)|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2} = |\tilde{\epsilon}(n-1)|^2 + \frac{|e_{\epsilon,p}(n)|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2}, \quad (44b)$$

$$|\tilde{\eta}(n)|^2 + \frac{|e_{\eta,a}(n)|^2}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2} = |\tilde{\eta}(n-1)|^2 + \frac{|e_{\eta,p}(n)|^2}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2}. \quad (44c)$$

Comparing (44a) with (44b) and (44c), we see that the main difference concerns the interpretation of the terms $\mathbf{w}_n^o - \mathbf{w}_n$ and $\mathbf{w}_n^o - \mathbf{w}_{n-1}$. While the term on the left-hand side of (44a) can be recognized as $\tilde{\mathbf{w}}_n$, just like the terms on the left-hand sides of (44b) and (44c), the second difference is not $\tilde{\mathbf{w}}_{n-1}$ since, due to the self-induced nonstationarity, $\tilde{\mathbf{w}}_{n-1}$ is defined as $\tilde{\mathbf{w}}_{n-1} = \mathbf{w}_{n-1}^o - \mathbf{w}_{n-1}$ in terms of $\mathbf{w}_{n-1}^o$ and not $\mathbf{w}_n^o$.

In order to explain the relevance of the energy relation equations to the steady-state analysis of the adaptive filter, we first need to relate $\|\mathbf{w}_n^o - \mathbf{w}_{n-1}\|^2$ to $\|\tilde{\mathbf{w}}_{n-1}\|^2$. To do so, we can write

$$\|\mathbf{w}_n^o - \mathbf{w}_{n-1}\|^2 = \left\| \mathbf{w}_{n-1}^o + \mathbf{w}_{n-1}^o j \tilde{\epsilon}(n-1) \right.$$

$$\left. + \frac{\mathbf{y}_n^*}{\|\mathbf{y}_n\|^2} \mathbf{y}_n' \mathbf{w}_{n-1}^o \tilde{\eta}(n-1) - \mathbf{w}_{n-1} \right\|^2. \quad (45)$$

Recall that the first three terms on the right-hand side within the squared norm constitute $\mathbf{w}_n^o$ by means of linear approximation

as in the derivation of (26). Based on (45), we get

$$\|\mathbf{w}_n^o - \mathbf{w}_{n-1}\|^2 = \|\tilde{\mathbf{w}}_{n-1}\|^2 + \left\| \mathbf{w}_{n-1}^o j \tilde{\epsilon}(n-1) \right\|^2$$

$$+ \left\| \frac{\mathbf{y}_n^*}{\|\mathbf{y}_n\|^2} \mathbf{y}_n' \mathbf{w}_{n-1}^o \tilde{\eta}(n-1) \right\|^2. \quad (46)$$

The last two terms on the right-hand side of which can be related to $|e_{\epsilon,a}^n(n)|^2$ and $|e_{\eta,a}^n(n)|^2$ by writing

$$\|\mathbf{w}_n^o - \mathbf{w}_{n-1}\|^2 = \|\tilde{\mathbf{w}}_{n-1}\|^2 + \frac{M|e_{\epsilon,a}^n(n)|^2}{\|\mathbf{y}_n\|^2} + \frac{M|e_{\eta,a}^n(n)|^2}{\|\mathbf{y}_n\|^2}. \quad (47)$$

Substituting (47) into (44a), taking the expectation of both sides of (44a), (44b), and (44c), using that $E\|\tilde{\mathbf{w}}_n\|^2 = E\|\tilde{\mathbf{w}}_{n-1}\|^2$, $E|\tilde{\epsilon}(n)|^2 = E|\tilde{\epsilon}(n-1)|^2$, and $E|\tilde{\eta}(n)|^2 = E|\tilde{\eta}(n-1)|^2$ in steady state as $n \to \infty$, gives the following fundamental variance relations:

$$E\left[\frac{|e_{w,a}(n)|^2}{\|\mathbf{y}_n\|^2} \right] = E\left[\frac{M|e_{\epsilon,a}^n(n)|^2}{\|\mathbf{y}_n\|^2} \right]$$

$$+ E\left[\frac{M|e_{\eta,a}^n(n)|^2}{\|\mathbf{y}_n\|^2} \right] + E\left[\frac{|e_{w,p}(n)|^2}{\|\mathbf{y}_n\|^2} \right], \quad (48a)$$

$$E\left[\frac{|e_{\epsilon,a}(n)|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2} \right] = E\left[\frac{|e_{\epsilon,p}(n)|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2} \right], \quad (48b)$$

$$E\left[\frac{|e_{\eta,a}(n)|^2}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2} \right] = E\left[\frac{|e_{\eta,p}(n)|^2}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2} \right]. \quad (48c)$$

These equalities are given in terms of the a priori and a posteriori errors. However, we know from (41) how those errors are related. Therefore, using (41) the above collapse to the following error variance relations in terms of the a priori errors and noise only:

$$E\left[\frac{|e_{w,a}(n)|^2}{\|\mathbf{y}_n\|^2} \right] = E\left[\frac{M|e_{\epsilon,a}^n(n)|^2}{\|\mathbf{y}_n\|^2} \right] + E\left[\frac{M|e_{\eta,a}^n(n)|^2}{\|\mathbf{y}_n\|^2} \right]$$

$$+ E\left[\frac{1}{\|\mathbf{y}_n\|^2} |e_{w,a}(n) - \mu_w \|\mathbf{y}_n\|^2 e(n)|^2 \right], \quad (49a)$$

$$E\left[\frac{|e_{\epsilon,a}(n)|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2} \right] = E\left[\frac{|e_{\epsilon,a}(n) - \mu_\epsilon |\mathbf{y}_n \mathbf{w}_{n-1}|^2 j^* e(n)|^2}{|\mathbf{y}_n \mathbf{w}_{n-1}|^2} \right], \quad (49b)$$

$$E\left[\frac{|e_{\eta,a}(n)|^2}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2} \right] = E\left[\frac{|e_{\eta,a}(n) - \mu_\eta |\mathbf{y}_n' \mathbf{w}_{n-1}|^2 e(n)|^2}{|\mathbf{y}_n' \mathbf{w}_{n-1}|^2} \right]. \quad (49c)$$

Expanding the above, rearranging, and dividing by μ_w, μ_ϵ, and μ_η respectively, we get

$$\mu_w E\left[\|\mathbf{y}_n\|^2 |e(n)|^2 \right] + E\left[\frac{M|e_{\epsilon,a}^n(n)|^2}{\mu_w \|\mathbf{y}_n\|^2} \right]$$

$$+ E\left[\frac{M|e_{\eta,a}^n(n)|^2}{\mu_w \|\mathbf{y}_n\|^2} \right] = 2\Re\{ E\left[e_{w,a}^*(n) e(n) \right] \}, \quad (50a)$$

$$\mu_\epsilon E\left[|\mathbf{y}_n \mathbf{w}_{n-1}|^2 |e(n)|^2 \right] = 2\Re\{ E\left[e_{\epsilon,a}^*(n) e(n) \right] \}, \quad (50b)$$

$$\mu_\eta E\left[|\mathbf{y}_n' \mathbf{w}_{n-1}|^2 |e(n)|^2 \right] = 2\Re\{ E\left[e_{\eta,a}^*(n) e(n) \right] \}. \quad (50c)$$

Finally, substituting (29) into the equations in (50) while also relying on A.2, we arrive at the equations in (33).

REFERENCES

[1] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

[2] B. S. Sharif, J. Neasham, O. R. Hinton, and A. E. Adams, "A computationally efficient Doppler compensation system for underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 52–61, Jan. 2000.

[3] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 191–193, Apr. 1995.

[4] B. Stantchev and G. Fettweis, "Time-variant distortions in OFDM," *IEEE Commun. Lett.*, vol. 4, no. 10, pp. 312–314, Oct. 2000.

[5] L. Weng et al., "Effect of carrier frequency offset on channel estimation for SISO/MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1854–1863, May 2007.

[6] T. Pollet, P. Spruyt, and M. Moeneclaey, "The BER performance of OFDM systems using non-synchronized sampling," in *Proc. IEEE Glob. Commun. Conf.*, 1994, pp. 253–257.

[7] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct. 1994.

[8] M. Sliskovic, "Carrier and sampling frequency offset estimation and correction in multicarrier systems," in *Proc. Glob. Telecommun. Conf.*, 2001, vol. 1, pp. 285–289.

[9] S.-Y. Liu and J.-W. Chong, "A study of joint tracking algorithms of carrier frequency offset and sampling clock offset for OFDM-based WLANs," in *Proc. IEEE Int. Conf. Commun., Circuits Syst.*, 2002, vol. 1, pp. 109–113.

[10] H. Nguyen-Le, T. Le-Ngoc, and C. C. Ko, "RLS-based joint estimation and tracking of channel response, sampling, and carrier frequency offsets for OFDM," *IEEE Trans. Broadcast.*, vol. 55, no. 1, pp. 84–94, Mar. 2009.

[11] K. Shi, E. Serpedin, and P. Ciblat, "Decision-directed fine synchronization in OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 408–412, Mar. 2005.

[12] M. Morelli and M. Moretti, "Fine carrier and sampling frequency synchronization in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1514–1524, Apr. 2010.

[13] Y.-H. Kim and J.-H. Lee, "Joint maximum likelihood estimation of carrier and sampling frequency offsets for OFDM systems," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 277–283, Jun. 2011.

[14] X. Wang and B. Hu, "A low-complexity ML estimator for carrier and sampling frequency offsets in OFDM systems," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 503–506, Mar. 2014.

[15] H. Viswanathan, S. Venkatesan, and H. Huang, "Downlink capacity evaluation of cellular networks with known-interference cancellation," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 802–811, Jun. 2003.

[16] W. Guo, C. Li, H. Zhao, R. Wen, and Y. Tang, "Comprehensive effects of imperfect synchronization and channel estimation on known interference cancellation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 457–470, Jan. 2020.

[17] W. Guo, H. Zhao, W. Ma, C. Li, Z. Lu, and Y. Tang, "Effect of frequency offset on cooperative jamming cancellation in physical layer security," in *Proc. IEEE Globecom Workshops*, 2018, pp. 1–5.

[18] C. Li, Y. Liu, Q. Xu, and Y. Tang, "Self-interference cancellation with frequency offset and nonlinear distortion suppression for cooperative jamming communications," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 2091–2094, Nov. 2019.

[19] W. Guo, H. Zhao, and Y. Tang, "Testbed for cooperative jamming cancellation in physical layer security," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 240–243, Feb. 2020.

[20] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the *K*-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[21] H.-M. Wang, Q. Yin, and X.-G. Xia, "Distributed beamforming for physical-layer security of two-way relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3532–3545, Jul. 2012.

[22] R. Mudumbai, D. R. B. Iii, U. Madhow, and H. V. Poor, "Distributed transmit beamforming: Challenges and recent progress," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 102–110, Feb. 2009.

[23] I. Walterscheid et al., "Bistatic SAR experiments with PAMIR and TerraSAR-X–setup, processing, and image results," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3268–3279, Aug. 2010.

[24] M. Weib, "Synchronisation of bistatic radar systems," in *Proc. Int. Geosc. Remote Sens. Symp.*, 2004, vol. 3, pp. 1750–1753.

[25] J. Yi, X. Wan, D. Li, and H. Leung, "Robust clutter rejection in passive radar via generalized subband cancellation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1931–1946, Aug. 2018.

[26] R. Gitlin and J. Thompson, "A phase adaptive structure for echo cancellation," *IEEE Trans. Commun.*, vol. 26, no. 8, pp. 1211–1220, Aug. 1978.

[27] M. Pawig, G. Enzner, and P. Vary, "Adaptive sampling rate correction for acoustic echo control in voice-over-IP," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 189–199, Jan. 2010.

[28] N. R. Yousef and A. H. Sayed, "Ability of adaptive filters to track carrier offsets and channel nonstationarities," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1533–1544, Jul. 2002.

[29] H.-C. So, P.-C. Ching, and Y.-T. Chan, "A new algorithm for explicit adaptation of time delay," *IEEE Trans. Signal Process.*, vol. 42, no. 7, pp. 1816–1820, Jul. 1994.

[30] Z. Li, Y. Xia, W. Pei, K. Wang, and D. P. Mandic, "An augmented nonlinear LMS for digital self-interference cancellation in full-duplex direct-conversion transceivers," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4065–4078, Aug. 2018.

[31] X. Zhang, Y. Xia, C. Li, L. Yang, and D. P. Mandic, "Complex properness inspired blind adaptive frequency-dependent I/Q imbalance compensation for wideband direct-conversion receivers," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5982–5992, Sep. 2020.

[32] M. Meller, "Cheap cancellation of strong echoes for digital passive and noise radars," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2654–2659, May 2012.

[33] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.

[34] S. S. Haykin, *Adaptive Filter Theory*, 5th ed. London, U.K.: Pearson Education, 2014.

[35] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley, 2003.

[36] K. Atkinson and W. Han, *Theoretical Numerical Analysis*, vol. 39. New York, NY, USA: Springer, 2005.

[37] T. Ramstad, "Digital methods for conversion between arbitrary sampling frequencies," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 3, pp. 577–591, Jun. 1984.

[38] T. I. Laakso, "Splitting the unit delay," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.

[39] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ, USA: Wiley, 2009.

[40] C. Candan, "An efficient filtering structure for Lagrange interpolation," *IEEE Signal Process. Lett.*, vol. 14, no. 1, pp. 17–19, Jan. 2007.

[41] E. Soria-Olivas, J. Calpe-Maravilla, J. F. Guerrero-Martinez, M. Martinez-Sober, and J. Espi-Lopez, "An easy demonstration of the optimum value of the adaptation constant in the LMS algorithm [FIR filter theory]," *IEEE Trans. Educ.*, vol. 41, no. 1, p. 81, Feb. 1998.

[42] N. Zhang, J. Ni, J. Chen, and Z. Li, "Steady-state mean-square error performance analysis of the tensor LMS algorithm," *IEEE Trans. Circuits Syst. II:Exp. Briefs*, vol. 68, no. 3, pp. 1043–1047, Mar. 2021.

[43] N. D. Dalt, M. Harteneck, C. Sandner, and A. Wiesbauer, "On the jitter requirements of the sampling clock for analog-to-digital converters," *IEEE Trans. Circuits Syst. I*, vol. 49, no. 9, pp. 1354–1360, Sep. 2002.

Karel Pärlin received the M.Sc. degree in electrical engineering from the Tallinn University of Technology, Tallinn, Estonia, in 2017. He is currently working toward the D.Sc. degree in communication engineering with Tampere University, Tampere, Finland. His research interests include adaptive signal processing, signal processing for communications, and physical layer security.

Taneli Riihonen (Senior Member, IEEE) received the D.Sc. degree in electrical engineering from Aalto University, Espoo, Finland, in 2014. He is currently a tenure-track Associate Professor with the Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland. His research interests include physical-layer OFDM(A), multiantenna, multihop, and full-duplex wireless techniques with current research interest includes the evolution of beyond 5G systems.

Marc Adrat received the Diploma and Dr.-Ing. degrees in electrical engineering from RWTH Aachen University, Aachen, Germany, in 1997 and 2003, respectively. He is currently the Head of the Software Defined Radio (SDR) Research Group, Fraunhofer FKIE, Wachtberg, Germany. His research interests include digital signal processing for mobile tactical radio communications, and emerging technologies like in-band full-duplex communications. Since more than 10 years, he is a Guest Lecturer with RWTH Aachen University for a course on channel coding.

Vincent Le Nir received the Ph.D. degree in electronics from the National Institute of Applied Sciences, France, in 2004. He is currently a Senior Researcher with the Royal Military Academy, Brussels, Belgium. His research interests include digital communications and signal processing in the wireless and wireline domains, MIMO communications, space–time coding, OFDM and multicarrier-code-division multiple-access, turbo-equalization, software defined, and cognitive radio.