Adaptive Filtering Based on Time-Averaged MSE for Cyclostationary Signals

Nir Shlezinger, *Student Member, IEEE*, Koby Todros, *Member, IEEE*, and Ron Dabora, *Senior Member, IEEE*

***Abstract*— Adaptive filters are commonly used in many signal processing and communications systems. In many practical digital communications scenarios, including, for example, interference-limited wireless and wireline communications, as well as narrowband power line communications, the considered signals are jointly *cyclostationary*. Yet, most works on adaptive filtering of cyclostationary signals used ad hoc application of adaptive algorithms designed for *stationary* signals, e.g., the least-mean-squares (LMS). It is known that these algorithms may not converge for jointly cyclostationary signals. In this paper, we rigorously study the optimal adaptive filtering of jointly cyclostationary signals. We first identify the relevant objective as the time-averaged mean-squared error criterion (TA-MSE), and obtain an adaptive algorithm as the stochastic approximation of the TA-MSE minimizer. When the considered signals are jointly stationary, the algorithm specializes to the standard LMS algorithm. We provide a comprehensive transient and steady-state performance analysis without imposing a specific distribution on the considered signals, and derive conditions for convergence and stability. The algorithm, which we call *time-averaged LMS*, is applied to practical scenarios in a simulations study, and an excellent agreement between the theoretical and the empirical performance is observed.**

***Index Terms*— Adaptive estimation, cyclostationary signals, power line communications.**

I. Introduction

ADAPTIVE filters play an important role in the implementation of linear estimators. These algorithms adapt the filter coefficients in run-time without a-priori knowledge about the statistical moments of the input and reference signals [2, Part 3], [3, Part III]. A widely used class of adaptation algorithms is the family of least mean-squares (LMS)-type algorithms, that gained popularity due to their low computational and implementation complexities while providing good performance [2, Ch. 9], [3, Ch. 10]. The majority of LMS-type filters are designed for linear estimation of a wide-sense stationary (WSS) reference signal, also called the signal of interest (SOI), based on a statistically dependent WSS input signal.

Manuscript received September 18, 2016; revised January 10, 2017; accepted January 11, 2017. Date of publication January 20, 2017; date of current version April 14, 2017. This work was supported in part by the Ministry of Economy of Israel through the Israeli Smart Grid Consortium and in part by the Israel Science Foundation under Grant 1685/16. This work was presented at the 2016 International Symposium on Wireless Communications Systems, Poznan, Poland, in the paper [1]. The associate editor coordinating the review of this paper and approving it for publication was G. Colavolpe.

The authors are with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Be'er-Sheva 8410501, Israel (e-mail: nirshl@post.bgu.ac.il; todros@ee.bgu.ac.il; ron@ee.bgu.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2017.2655526

Since man-made signals are typically cyclostationary [4, Secs. 5–7], it follows that cyclostationary signals are inherent to most communications scenarios. For example, most digitally modulated communications signals, including, e.g., quadrature amplitude modulated (QAM) signals [5, Ch. 1] and orthogonal frequency division multiplexing (OFDM) signals [6], are cyclostationary. Consequently, in interference-limited communications, which is attracting extensive research attention [7], [8], the additive interfering signal is cyclostationary. Cyclostationarity is also a fundamental characteristic of power line channels [9]. In particular, in narrowband (NB) power line communications (PLC), the additive noise is a cyclostationary random process [10], [11].

The work [12] studied the convergence of the LMS algorithm for linear estimation of a cyclostationary SOI based on a jointly cyclostationary input signal. It was shown that the LMS algorithm is mean convergent only when the step-size of the LMS algorithm approaches zero. In [13] the LMS algorithm was applied to the identification of a linear time-varying system whose coefficients obey a random walk model. The system input was considered to be a Gaussian cyclostationary process and its output was corrupted by an additive WSS Gaussian noise. For this setup, the transient, tracking, and steady-state performance of the LMS algorithm were studied and conditions for convergence were derived. The work [14] studied the performance of an adaptive line enhancer/canceler for a cyclostationary input consisting of a superposition between a fixed amplitude random phase sine wave and a Gaussian process with periodic variance, based on the LMS adaptation. In many practical scenarios involving cyclostationary signals, e.g., interference-limited communications and NB-PLC, the SOI and the input signal are related via a time-varying system with deterministic coefficients, and in particular, a linear periodically time-varying (LPTV) system, whose output is corrupted by a cyclostationary noise.

The optimal linear estimator in the mean-squared error (MSE) sense of a cyclostationary SOI based on a cyclostationary input signal is the cyclic Wiener filter (CWF) [15]. This estimator achieves the minimal MSE (MMSE) at any time instant during the period, and is realized by an LPTV filter. One of the most widely used filter structures for realizing LPTV filters is the frequency shift (FRESH) filter [15] (also referred to as *modulator filter* in [16]). This filter realizes LPTV systems by summing the outputs of multiple linear time-invariant (LTI) filters, applied to frequency-shifted versions of the input signal. Adaptive FRESH filters were considered in several works, e.g., [17]–[19], in which adaptation

of the coefficients of the component LTI filters comprising the FRESH filter was considered. The work [20] studied the performance of adaptive LPTV filters, including adaptive FRESH filters, for interference rejection in the presence of cyclostationary digitally modulated signals. However, we note that in the works [17]–[20], adaptation was implemented using the ad hoc conventional approach for stationary signals.

We note that while any LPTV filter can be implemented by a set of LTI filters and modulators or switches [16], with possibly infinite impulse responses, in practical designs, linear estimators for cyclostationary signals are typically realized using a finite overall number of taps. Consequently, often times in practice, the number of LTI filters used for implementing an LPTV filter is smaller than the number required for implementing the CWF [15], [18], and as a result, linear estimators with time-invariant coefficients typically do not achieve the MMSE at all time instants. In such cases, the *time-averaged* MSE (TA-MSE) objective is preferred upon the instantaneous MSE [15], [18], [21]. This motivates the development of an adaptive LMS-type algorithm for linear estimation in a jointly cyclostationary setup, based on the TA-MSE criterion.

A. Main Contributions and Organization

In this work we rigorously study optimal adaptive filtering for discrete-time (DT) jointly proper-complex [22] wide-sense cyclostationary [5, Ch. 1] SOI and input signal. Our linear estimator setup accommodates various structures of linear estimators for cyclostationary signals, including finite impulse response (FIR) LTI filters and finite-memory LPTV filters. We first highlight the fact that, when the overall number of taps used for realizing the linear estimator is fixed, then the linear MMSE (LMMSE) filter for the estimation of cyclostationary signals may involve time-varying coefficients. Thus, when a linear estimator with time-invariant coefficients is desired, the TA-MSE cost function is typically preferable over the MSE objective. Then, we derive an adaptive algorithm via a stochastic approximation of the steepest descent (SD) algorithm with the TA-MSE cost function, which is referred to as the time-averaged LMS (TA-LMS). When the considered signals are jointly WSS, the TA-LMS specializes to the standard LMS. Differently from previous works which used ad hoc application of LMS variations, our algorithm is rigorously shown to converge for cyclostationary signals. We discuss the relationship of the algorithm with adaptive algorithms designed for stationary signals, and in particular, the block LMS (B-LMS) and the affine projection algorithm (APA).

We carry out a comprehensive performance analysis of the algorithm without imposing a specific distribution on the input signal, and obtain explicit expressions for the mean-square deviation (MSD) in the filter coefficients and for the excess TA-MSE. Additionally, we derive sufficient conditions for convergence and stability. The new theoretical analysis facilitates the selection of the filter parameters, including the step-size and the number of taps. The advantages of our optimal approach in terms of convergence and transient performance are illustrated for channel estimation in NB-PLC scenarios. We show a very good agreement between the theoretical and the empirical performance measures.

The rest of this paper is organized as follows: In Section II the relevant properties of cyclostationary processes are briefly reviewed and the considered estimation problem is stated formally. In Section III the adaptive TA-MSE minimizer is derived. In Section IV the performance analysis of the algorithm is carried out. In Section V the TA-LMS is illustrated in simulation examples. Lastly, in Section VI the main contributions of this work are summarized. Proofs of the propositions and theorems stated in the paper are detailed in the appendix.

II. Problem Formulation

A. Notations

We denote column vectors with lower-case boldface letters, e.g., $\mathbf{x}$; the k-th element ($k \geq 0$) of a vector $\mathbf{x}$ is denoted with $(\mathbf{x})_k$. Matrices are denoted with upper-case boldface letter, e.g., $\mathbf{X}$; the element at the k-th row and the l-th column of $\mathbf{X}$ is denoted by $(\mathbf{X})_{k,l}$. We use $\mathbf{I}_n$ to denote the $n \times n$ identity matrix and $\mathbf{0}_{n\times m}$ to denote the all-zero $n \times m$ matrix. Hermitian transpose, transpose, complex conjugate, and stochastic expectation are denoted by $(\cdot)^H$, $(\cdot)^T$, $(\cdot)^*$, and $\mathbb{E}\{\cdot\}$, respectively. We use $\mathrm{Re}\{x\}$ to denote the real part of the complex number x, $\mathrm{Tr}\{\cdot\}$ to denote the trace operator, $((n))_m$ to denote the remainder of n when divided by m, and $\otimes$ to denote the Kronecker product. We use $\langle\cdot\rangle_N$ to denote time-averaging over $N > 0$ samples, namely, for a sequence $x[n]$, $\langle x[n]\rangle_N \triangleq \frac{1}{N}\sum_{k=0}^{N-1} x[n-k]$. The sets of integers and non-negative integers are denoted by $\mathbb{Z}$ and $\mathbb{N}$, respectively. For an $n \times n$ matrix $\mathbf{X}$, let $\lambda_{\max}(\mathbf{X})$ denote the largest real eigenvalue of $\mathbf{X}$, given that such exists, and $\mathbf{x} = \mathrm{vec}(\mathbf{X})$ denote the $n^2 \times 1$ column vector obtained by stacking the columns of $\mathbf{X}$ one below the other, with the leftmost column being at the top of $\mathbf{x}$. The matrix $\mathbf{X}$ is recovered from $\mathbf{x}$ via $\mathbf{X} = \mathrm{vec}^{-1}(\mathbf{x})$. For an $n \times 1$ vector $\mathbf{y}$ and an $n^2 \times 1$ vector $\mathbf{x}$, we use $\|\mathbf{y}\|^2$ to denote the squared Euclidean norm $\mathbf{y}^H\mathbf{y}$, and $\|\mathbf{y}\|^2_{\mathbf{x}}$ to denote $\mathbf{y}^H\mathrm{vec}^{-1}(\mathbf{x})\,\mathbf{y}$. Note that when $\mathrm{vec}^{-1}(\mathbf{x})$ is a positive-(semi)-definite matrix then $\|\mathbf{y}\|^2_{\mathbf{x}}$ is the weighted squared Euclidean (semi) norm. Lastly, for a sequence of $n\times n$ matrices $\{\mathbf{X}_k\}_{k=l}^m$, $m \geq l$, $\prod_{k=l}^{m}\mathbf{X}_k$ denotes the matrix product $\mathbf{X}_m\mathbf{X}_{m-1}\cdots\mathbf{X}_l$.

B. Cyclostationary Stochastic Processes

A DT proper-complex multivariate process $\mathbf{x}[n]$ is said to be *wide-sense cyclostationary* (referred to henceforth as cyclostationary) if both its mean value and autocorrelation function are periodic with some period, N_0, i.e., $\mathbb{E}\{\mathbf{x}[n]\} = \mathbb{E}\{\mathbf{x}[n+N_0]\}$, and $\mathbf{C}_{\mathbf{x}}[n,l] \triangleq \mathbb{E}\{\mathbf{x}[n+l]\mathbf{x}^H[n]\} = \mathbf{C}_{\mathbf{x}}[n+N_0,l]$, see also [4, Sec. 3.5], [5, Ch. 1]. A pair of jointly proper-complex DT processes $\mathbf{x}_1[n]$ and $\mathbf{x}_2[n]$ are said to be *jointly wide-sense cyclostationary* (referred to henceforth as jointly cyclostationary) with period N_0 if each process is cyclostationary with period N_0 and the cross-correlation function $\mathbb{E}\{\mathbf{x}_1[n+l]\mathbf{x}_2^H[n]\}$ is periodic[1] with period N_0 w.r.t. n [4, Sec. 3.6.2]. A

[1]Note that when $\mathbf{x}_1[n]$ and $\mathbf{x}_2[n]$ are cyclostationary with different periods, say N_1 and N_2, respectively, and the cross correlation is periodic with period N_{12}, then the joint second-order statistical moments have a common period N_0 which is a common multiple of N_1, N_2, and N_{12}. In particular, if $\mathbf{x}_1[n]$ is WSS and $\mathbf{x}_2[n]$ is cyclostationary with period N_2, uncorrelated with $\mathbf{x}_1[n]$, then they are jointly cyclostationary with a period N_2.

TABLE I
COMMON ESTIMATOR STRUCTURES FOR SCALAR CYCLOSTATIONARY SIGNALS

System	Parameters	Receive transformation
Scalar LTI FIR filter [3, Ch. 10.5]	M - FIR length (positive integer)	$(\mathbf{x}[n])_k = r[n-k]$, $k \in \{0, 1, \ldots, M-1\}$
Scalar FRESH filter [17], [18]	K - number of LTI filters (positive integer), L - number of taps in each filter (positive integer), $\{\alpha_k\}_{k=0}^{K-1}$ - cyclic frequencies, $M = KL$ - total number of taps	$(\mathbf{x}[n])_{u \cdot L+v} = r[n-v]e^{j\alpha_u(n-v)}$, $u \in \{0, 1, \ldots, K-1\}$, $v \in \{0, 1, \ldots, L-1\}$

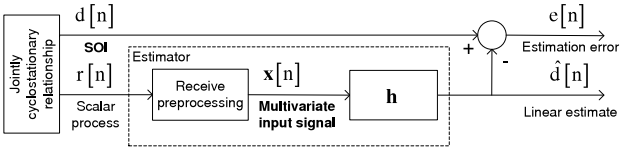


Fig. 1. Linear estimation of scalar jointly cyclostationary signals.

DT cyclostationary process can be obtained by synchronously sampling a continuous-time (CT) cyclostationary process with a sampling period, T_s, which is an integer divisor of some positive integer multiple of the period of the CT process, denoted T_0. When T_s is not synchronized with T_0, i.e., T_0 satisfies $T_0 = kT_s + \epsilon$, where $k \in \mathbb{N}$ and $\epsilon \in (0, T_s)$ is incommensurate with T_s, then the resulting DT process is *almost cyclostationary* rather than cyclostationary [4, Sec. 3.9]. In this work we focus on cyclostationary DT signals, hence, when the signals are obtained by sampling CT cyclostationary processes, we assume synchronous sampling.

C. Problem Formulation

We study the problem of linear estimation of a zero-mean scalar SOI $d[n]$ based on an $M \times 1$ zero-mean multivariate input signal $\mathbf{x}[n]$, where $\mathbf{x}[n]$ and $d[n]$ are jointly cyclostationary with period N_0 and jointly proper-complex [22, Definition 1]. Let $\mathbf{h}$ denote the $M \times 1$ coefficients vector. The linear estimate of $d[n]$ is computed via

$$\hat{d}[n] = \mathbf{h}^H \mathbf{x}[n], \quad n \geq 0. \tag{1}$$

The formulation (1) accommodates a wide range of estimator structures, including scalar LTI filters and scalar FRESH filters, which are very common in linear estimation of cyclostationary signals [23, Ch. 17.5]. These special cases are illustrated in Fig. 1. The multivariate filter input $\mathbf{x}[n]$ is obtained by mapping the scalar input signal $r[n]$. From the setup in Fig. 1, the relationships between the general formulation (1) and the common scalar LTI and scalar FRESH estimator structures can be easily obtained, and Table I details the corresponding transformations to $r[n]$. As any DT LPTV system can be realized as a FRESH filter [16], it follows that (1) can realize *any scalar LPTV filter* of finite dimensions.

We assume knowledge of the period of the second order statistics, N_0. This is the case in many practical scenarios of cyclostationary signals, e.g., interference-limited communications, in which the period is equal to the symbol duration of the interfering signal [5, Ch. 1], and NB-PLC, where the period is equal to the period of the mains voltage [10].

We also note that the most general class of linear estimators can be obtained by letting the coefficients vector $\mathbf{h}$ in (1) to vary in time. Nonetheless, since the time-invariant formulation (1) can be used to realize scalar LTI and LPTV filters, we focus in this work on the adaptive implementation of linear estimators with time-invariant coefficients.

III. ADAPTIVE LINEAR ESTIMATION OF CYCLOSTATIONARY SIGNALS BASED ON TA-MSE

In this section we derive an adaptive linear estimator for cyclostationary signals based on the TA-MSE objective. One of the motivations for using the TA-MSE objective is the fact that a linear estimator with time-invariant coefficients which achieves the MMSE for each time instant does not always exist for cyclostationary signals [15]. Specifically, for the class of linear estimators (1), an LMMSE estimator exists if and only if there exists a vector $\mathbf{h}_{\text{MMSE}}$ that satisfies

$$\mathbb{E}\left\{\mathbf{x}[n]\,\mathbf{x}^H[n]\right\}\mathbf{h}_{\text{MMSE}} = \mathbb{E}\left\{\mathbf{x}[n]\,d^*[n]\right\}, \tag{2}$$

$\forall n \in \{0, 1, \ldots, N_0 - 1\} \triangleq \mathcal{N}_0$. The resulting LMMSE estimator is then given by $\hat{d}[n] = \mathbf{h}_{\text{MMSE}}^H \mathbf{x}[n]$. We therefore conclude that a time-invariant LMMSE estimator does not necessarily exist for cyclostationary signals. One example scenario in which a time-invariant LMMSE estimator does exist, is the scenario considered in [13], when the filter relating the input signal and the SOI is restricted to be deterministic LTI, as was the case in [13, Figs. 2, 4, 6, and 8].

A. The Minimum TA-MSE Linear Estimator

Since an LMMSE estimator with time-invariant coefficients does not always exist for cyclostationary signals, the minimum TA-MSE (MTA-MSE) is commonly used as an alternative performance objective [18], [21]. Define the deterministic quantities: $\underset{\sim}{c}_d \triangleq \left\langle \mathbb{E}\left\{|d[n]|^2\right\}\right\rangle_{N_0}$, $\underset{\sim}{\mathbf{c}}_{\mathbf{x}d} \triangleq \left\langle \mathbb{E}\left\{\mathbf{x}[n]\,d^*[n]\right\}\right\rangle_{N_0}$,

$$\mathbf{C}_{\mathbf{x}}[n] \triangleq \mathbb{E}\left\{\mathbf{x}[n]\,\mathbf{x}^H[n]\right\}, \tag{3a}$$

$$\mathbf{c}_{\mathbf{x}}[n] \triangleq \text{vec}\left(\mathbf{C}_{\mathbf{x}}[n]\right), \tag{3b}$$

and

$$\underset{\sim}{\mathbf{C}}_{\mathbf{x}} \triangleq \left\langle \mathbf{C}_{\mathbf{x}}[n] \right\rangle_{N_0}. \tag{3c}$$

Note that $\underset{\sim}{c}_d$, $\underset{\sim}{\mathbf{c}}_{\mathbf{x}d}$, and $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ are time-invariant since $d[n]$ and $\mathbf{x}[n]$ are jointly cyclostationarity, while $\mathbf{C}_{\mathbf{x}}[n]$ and $\mathbf{c}_{\mathbf{x}}[n]$ are periodic with period N_0. We henceforth assume the following:

AS1 $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ is non-singular.

The TA-MSE associated with the linear estimator (1) is obtained as (see, e.g., [18, eq. (26)], [21, eq. (24)])

$$\left\langle \mathbb{E}\left\{\left|d\left[n\right]-\mathbf{h}^H\mathbf{x}\left[n\right]\right|^2\right\}\right\rangle_{N_0} = \underset{\sim}{c}_d - 2\mathrm{Re}\left\{\underset{\sim}{\mathbf{c}}_{\mathbf{x}d}^H\mathbf{h}\right\} + \mathbf{h}^H\underset{\sim}{\mathbf{C}}_{\mathbf{x}}\mathbf{h} \triangleq J\left(\mathbf{h}\right). \quad (4)$$

Therefore, the optimal filter coefficients vector in the sense of minimum TA-MSE, denoted by $\mathbf{h}_{\mathrm{opt}}$, must satisfy the time-averaged Wiener-Hopf equations (see, e.g., [18, eq. (5)]):

$$\underset{\sim}{\mathbf{C}}_{\mathbf{x}}\mathbf{h}_{\mathrm{opt}} = \underset{\sim}{\mathbf{c}}_{\mathbf{x}d}, \quad (5)$$

resulting in the linear MTA-MSE (LMTA-MSE) estimator

$$\hat{d}_{\mathrm{opt}}[n] \triangleq \mathbf{h}_{\mathrm{opt}}^H\mathbf{x}[n]. \quad (6)$$

The following lemma states an equivalence between the LMMSE and LMTA-MSE estimators.

Lemma 1: Assume that AS1 is satisfied. If (2) *is satisfied then the time-invariant LMMSE estimator exists and it coincides with the LMTA-MSE estimator* (6)*.*

Proof: Applying the time-averaging operator $\langle\cdot\rangle_{N_0}$ to both sides of (2) yields (5). Hence, under the lemma conditions, the LMMSE estimator is also TA-MSE optimal. Next, we compute the Hessian matrix of $J\left(\mathbf{h}\right)$. From [24, p. 284] it follows that the Hessian of $J\left(\mathbf{h}\right)$ stated in (4) is $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$. As $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ is a non-singular average of covariance matrices, it is positive definite. Then, from [25, Ch. 3.1.4] it follows that the positive definiteness of the Hessian of $J\left(\mathbf{h}\right)$ implies that $J\left(\mathbf{h}\right)$ is a strictly convex function of $\mathbf{h}$. This, in turn, implies uniqueness of the minimizing $\mathbf{h}$. □

The property stated in Lemma 1 further motivates the use of the TA-MSE as a design criterion.

B. Adaptive Estimator Based on TA-MSE

The adaptive algorithm adjusts the filter coefficients in runtime for $n > 0$ based on the SOI $d[n]$ and the input signal $\mathbf{x}[n]$ for $n > -N_0$, and on an initial guess of the filter coefficients $\mathbf{h}[0]$. For a fixed step-size μ, the SD update equation is given by [3, Ch. 8.3]

$$\mathbf{h}\left[n+1\right] = \mathbf{h}\left[n\right] - \mu\left.\frac{\partial}{\partial\mathbf{h}^*}J\left(\mathbf{h}\right)\right|_{\mathbf{h}=\mathbf{h}[n]} \overset{(a)}{=} \mathbf{h}\left[n\right] + \mu\left(\underset{\sim}{\mathbf{c}}_{\mathbf{x}d} - \underset{\sim}{\mathbf{C}}_{\mathbf{x}}\mathbf{h}\left[n\right]\right), \quad (7)$$

where (a) follows by applying the gradient to (4). Define the $M \times N_0$ matrix

$$\mathbf{X}\left[n\right] \triangleq \frac{1}{\sqrt{N_0}}\Big[\mathbf{x}\left[n\right], \mathbf{x}\left[n-1\right], \ldots, \mathbf{x}\left[n-N_0+1\right]\Big], \quad (8a)$$

and the $N_0 \times 1$ vector

$$\mathbf{d}\left[n\right] \triangleq \frac{1}{\sqrt{N_0}}\Big[d\left[n\right], d\left[n-1\right], \ldots, d\left[n-N_0+1\right]\Big]^T. \quad (8b)$$

As $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ and $\underset{\sim}{\mathbf{c}}_{\mathbf{x}d}$ represent the *time-averaged* correlations of jointly cyclostationary signals, their corresponding instantaneous stochastic approximations are given by:

$$\hat{\underset{\sim}{\mathbf{C}}}_{\mathbf{x}}[n] \triangleq \frac{1}{N_0}\sum_{k=0}^{N_0-1}\mathbf{x}\left[n-k\right]\mathbf{x}^H\left[n-k\right] = \mathbf{X}\left[n\right]\mathbf{X}^H\left[n\right], \quad (9a)$$

and

$$\hat{\underset{\sim}{\mathbf{c}}}_{\mathbf{x}d}[n] \triangleq \frac{1}{N_0}\sum_{k=0}^{N_0-1}\mathbf{x}\left[n-k\right]d^*\left[n-k\right] = \mathbf{X}\left[n\right]\mathbf{d}^*\left[n\right], \quad (9b)$$

respectively. Note that the estimators in (9) are unbiased since $\mathbb{E}\left\{\mathbf{X}\left[n\right]\mathbf{X}^H\left[n\right]\right\} = \underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ and $\mathbb{E}\left\{\mathbf{X}\left[n\right]\mathbf{d}^*\left[n\right]\right\} = \underset{\sim}{\mathbf{c}}_{\mathbf{x}d}$.

Replacing the time-averaged covariance matrix and cross-covariance vector in (7) with their instantaneous estimates (9), the resulting update equation is for $n \geq 0$:

$$\mathbf{h}[n+1] = \mathbf{h}[n] + \mu\mathbf{X}\left[n\right]\Big(\mathbf{d}^*\left[n\right] - \mathbf{X}^H\left[n\right]\mathbf{h}\left[n\right]\Big). \quad (10)$$

1) Relationship Between the TA-LMS and the LMS, B-LMS, and APA Algorithms: Note that for a period $N_0 = 1$, i.e., when $\mathbf{x}[n]$ and $d[n]$ are jointly WSS [5, p. 21], [23, Property 4], Eq. (10) specializes the conventional LMS, cf. [2, Ch. 9], hence, the conventional LMS is a special case of the TA-LMS when setting the signal period to unity. We also note that, similarly to the TA-LMS, both the B-LMS [2, Ch. 10.1] and the APA [3, Ch. 13.1] replace the correlations with time-averaged empirical estimates. It should be noted, however, that while the TA-LMS is the proper stochastic approximation of the SD algorithm for minimizing the TA-MSE for jointly cyclostationary signals, the B-LMS and the APA are derived via stochastic approximations of the SD and Newton's method, respectively, under the MSE objective for jointly stationary signals. Therefore, without correctly specifying the averaging window and the update rate, their application to cyclostationary signals is a purely ad hoc application which results in sub-optimal performance. Specifically, when the B-LMS block size is set to the period N_0, then the B-LMS and the TA-LMS have a similar update equation, with the difference that the B-LMS updates its coefficients once at each block of N_0 samples, while the TA-LMS updates the coefficients with each incoming sample. This implies that the rate of convergence of the B-LMS is much slower. Finally, we note that when the APA block size is set to the period N_0, and without normalizing the input signal, the APA coincides with the TA-LMS.

2) TA-LMS Complexity Analysis: A direct implementation of the TA-LMS update equation (10) requires a total of $N_0(8M+2)$ real multiplications and $8M \cdot N_0$ real additions per iteration. It can be shown that when the period N_0 satisfies $N_0 > \frac{7M+11}{8}$, the computational complexity of the TA-LMS update equation (10) can be significantly reduced, by recursively updating $\hat{\underset{\sim}{\mathbf{C}}}_{\mathbf{x}}[n]$ and $\hat{\underset{\sim}{\mathbf{c}}}_{\mathbf{x}d}[n]$ at each time n via the recursions $\hat{\underset{\sim}{\mathbf{C}}}_{\mathbf{x}}[n] = \hat{\underset{\sim}{\mathbf{C}}}_{\mathbf{x}}[n-1] + \frac{1}{N_0}\big(\mathbf{x}[n]\mathbf{x}^H[n] - \mathbf{x}[n-N_0]\mathbf{x}^H[n-N_0]\big)$ and $\hat{\underset{\sim}{\mathbf{c}}}_{\mathbf{x}d}[n] = \hat{\underset{\sim}{\mathbf{c}}}_{\mathbf{x}d}[n-1] + \frac{1}{N_0}\big(\mathbf{x}[n]d^*[n] - \mathbf{x}[n-N_0]d^*[n-N_0]\big)$. Under this recursive implementation, each iteration requires a total of $M(6M+8)$ real multiplications and $M(7M+11)$ real additions. For comparison, the standard LMS requires a total of $8M+2$ real multiplications and $8M$

real additions per iteration [3, Ch. 10.3]. The B-LMS with block size N_0 updates its coefficients once at each block with the same update equation as the TA-LMS (10), hence its complexity is equal to that of the TA-LMS divided by N_0, while the complexity of the APA with data reuse factor K is in the order of $O(K^3 + K^2M)$ real multiplications and real additions per iteration [3, Ch. 13.2], due to the additional normalization operation. To consistently compare the computational complexity, we compare the algorithms using the number of multiplications per iteration, as the cost of addition is negligible compared to the multiplication cost [26, Ch. 1.2]. To conclude, the computational complexity of the TA-LMS is $\min(M, N_0)$ times higher than the computational complexity of the standard LMS.

Remark 1: The fact that cyclostationary signals can be transformed into multivariate WSS signals using the decimated components decomposition (DCD) [23, Sec. 17.2], suggests that filters for cyclostationary signals can be designed by applying methods for designing filters for WSS signals to the equivalent multivariate representation, see, e.g., [27]. We note that using this approach for designing adaptive filters results in a slower convergence rate, larger memory requirements, and higher computational complexity compared to the approach followed in the manuscript: Each pair of the multivariate input and SOI samples corresponds to N_0 pairs of input and SOI samples in the original setup. Consequently, the rate of update of the filter coefficients is decreased by N_0, which decreases the rate of convergence by a factor of N_0. At the same time, the number of coefficients of the multivariate filter increases by a factor of N_0^2 compared to (1)*. For standard LMS adaptation, this increases the computational complexity per time instance by a factor of N_0. These effects constitute major drawbacks as adaptive filtering is applied at run-time. Therefore, in this work we consider filters applied directly to the cyclostationary input signal, as in* (1)*.*

IV. Performance Analysis of the TA-LMS Algorithm

In this section, we study the transient and steady-state performance of the TA-LMS algorithm (10) for linear estimation of cyclostationary signals.

A. Data Model and Assumptions

Let $\mathbf{h}_M[n]$ denote the (possibly time-varying) coefficients vector of the LMMSE estimator and $\upsilon[n]$ be a noise process representing the resulting estimation error. Using $\mathbf{h}_M[n]$ and $\upsilon[n]$, the SOI $d[n]$ can be written as

$$d[n] = \mathbf{h}_M^H[n]\mathbf{x}[n] + \upsilon[n]. \tag{11}$$

Note that the stationary linear data model used for the analysis of the standard LMS algorithm [3, Ch. 10.2] is a special case of (11), obtained by setting $N_0 = 1$. Note also that under the data model of (11), the condition in (2) is not necessarily satisfied, i.e., a time-invariant LMMSE estimator may not exist.

Since $d[n]$ and $\mathbf{x}[n]$ are jointly cyclostationary with period N_0, it follows that $\mathbf{h}_M[n]$ is an $M \times 1$ periodic sequence [23, Ch. 17.5.1], i.e., $\mathbf{h}_M[n] = \mathbf{h}_M[n + N_0]$, $\forall n \in \mathbb{Z}$. As $d[n]$ and $\mathbf{x}[n]$ are also zero-mean and jointly proper-complex, it follows from (11) that $\upsilon[n]$ is a zero-mean proper-complex cyclostationary process with period N_0. Furthermore, it follows from the orthogonality principle [3, Ch. 4.2] that $\upsilon[n]$ and $\mathbf{x}[n]$ are orthogonal. Similar to the standard approach used in the analysis of the LMS algorithm for jointly WSS signals, e.g., [2], [3], [29], [30], we make several additional independence assumptions on the signals used in (11):

AS2 The noise signal $\upsilon[n_1]$ and the input signal $\mathbf{x}[n_2]$ are *mutually independent* $\forall n_1, n_2$, see also [3, Ch. 15.2], [29, Sec. B.2], [30, Sec. II.D]. This assumption is satisfied, for example, if the SOI $d[n]$ is of the form $d[n] = \mathbf{g}_{N_0}^H[n]\mathbf{x}[n] + z[n]$, where $\mathbf{g}_{N_0}[n]$ is a deterministic periodically time-varying coefficients vector, and $z[n]$ is a cyclostationary process which is independent of $\mathbf{x}[n_2]$, $\forall n, n_2$. Another scenario in which this assumption is satisfied is when $d[n]$ and $\mathbf{x}[n]$ are jointly Gaussian and temporally uncorrelated, i.e., $\mathbb{E}\{\mathbf{x}[n_1]\mathbf{x}^H[n_2]\} = \mathbf{0}_{M\times M}$, and $\mathbb{E}\{\mathbf{x}[n_1]d^*[n_2]\} = \mathbf{0}_{M\times 1}$, for all $n_1 \neq n_2$. In such a scenario one can verify that $\mathbf{x}[n_1]$ and $\upsilon[n_2]$ are mutually independent $\forall n_1, n_2$.

AS3 The TA-LMS filter coefficients, $\mathbf{h}[n]$, and the instantaneous input vector, $\mathbf{x}[n]$, are statistically independent, see also [2, p. 392], [3, Ch. 16.4].

AS4 Similar to [28], we assume that $\mathbf{h}[n]$ is independent of $\left\{\mathbf{x}[k]\mathbf{x}^H[k]\right\}_{k=n-N_0+1}^{n}$.

Note that *AS3-AS4* are satisfied asymptotically, when the filter has converged, and the effect of the inputs on the filter coefficients is negligible. Similar independence assumptions are commonly used non-asymptotically, and are known to result in a reliable characterization of the transient performance, see discussions in [2, p. 393] and [3, p. 247]. Additionally, in Subsection V-A we detail an example scenario for which these assumptions are satisfied non-asymptotically. In addition to the independence assumptions *AS2-AS4*, we make the following assumptions:

AS5 The temporal correlation of $\upsilon[n]$ is bounded and spans a finite interval, i.e., $\mathbb{E}\{\upsilon[n+l]\upsilon^*[n]\}$ is bounded $\forall n, l \in \mathbb{Z}$ and also $\exists L_{max} > 0$ such that $\mathbb{E}\{\upsilon[n+l]\upsilon^*[n]\} = 0$ for all $|l| \geq L_{max}$, $\forall n \in \mathbb{Z}$. This assumption is satisfied when, for example, $\mathbf{x}[n]$ and $d[n]$ are periodic moving-average processes [31] with a temporal cross-correlation which has a finite duration.

AS6 Define

$$\eta \triangleq L_{max} + N_0. \tag{12}$$

All the $2N$-th order moments of $\mathbf{x}[n]$, $N \in \{1, 2, \ldots, \eta + 1\}$, are bounded and periodic[2] with period N_0. Specifically, for any set of pairs $\{k_i, n_i\}_{i=0}^{2N-1}$, where $k_i \in \{0, 1, \ldots, M-1\}$ and $n_i \in \mathbb{N}$,

[2]We note that a similar assumption was made in the analysis of LMS with non-Gaussian WSS inputs, which assumed that the fourth-order moments of the input signal are time-invariant [3, eq. (24.9)] and bounded [3, p. 361].

we assume that $\mathbb{E}\left\{\prod_{i=0}^{N-1}\left(\mathbf{x}\left[n_{2i}\right]\right)_{k_{2i}}\left(\mathbf{x}\left[n_{2i+1}\right]\right)_{k_{2i+1}}^{*}\right\}$ is bounded and equal to $\mathbb{E}\left\{\prod_{i=0}^{N-1}\left(\mathbf{x}\left[n_{2i}+N_0\right]\right)_{k_{2i}}\left(\mathbf{x}\left[n_{2i+1}+N_0\right]\right)_{k_{2i+1}}^{*}\right\}$. This assumption is satisfied when, for example, $\mathbf{x}[n]$ is proper-complex Gaussian and cyclostationary with period N_0. This is because in this case the joint probability density function (PDF) of $\mathbf{x}[n_1], \mathbf{x}[n_2], \ldots, \mathbf{x}[n_N]$ is equal to the joint PDF of $\mathbf{x}[n_1+N_0], \mathbf{x}[n_2+N_0], \ldots, \mathbf{x}[n_N+N_0]$, for any integer $N \geq 1$, and for all $\{n_i\}_{i=1}^{N} \in \mathbb{Z}$.

AS7 Define the $M^2 \times M^2$ matrix

$$\mathbf{B}[n] \triangleq \mathbb{E}\left\{\left(\mathbf{X}\left[n\right]\mathbf{X}^{H}\left[n\right]\right)^{T} \otimes \left(\mathbf{X}\left[n\right]\mathbf{X}^{H}\left[n\right]\right)\right\}. \quad (13)$$

We assume that $\mathbf{B}[n] = \mathbf{B}$, i.e., it is time-invariant. The assumption is satisfied, for example, when $\mathbf{x}[n_1]$ and $\mathbf{x}[n_2]$ are mutually independent $\forall n_1 \neq n_2$ and $\mathbf{x}[n]$ is fourth-order cyclostationary. This is shown by defining

$$\mathbf{B}_1\left[n\right] \triangleq \sum_{k_1=0}^{N_0-1}\mathbb{E}\Big\{\left(\mathbf{x}\left[n-k_1\right]\left(\mathbf{x}\left[n-k_1\right]\right)^{H}\right)^{T} \otimes \mathbf{x}\left[n-k_1\right]\left(\mathbf{x}\left[n-k_1\right]\right)^{H}\Big\}, \quad (14a)$$

and

$$\mathbf{B}_2\left[n\right] \triangleq \sum_{k_1=0}^{N_0-1}\mathbb{E}\left\{\mathbf{x}\left[n-k_1\right]\left(\mathbf{x}\left[n-k_1\right]\right)^{H}\right\}^{T} \otimes \sum_{k_2=0,k_2\neq k_1}^{N_0-1}\mathbb{E}\left\{\mathbf{x}\left[n-k_2\right]\left(\mathbf{x}\left[n-k_2\right]\right)^{H}\right\}. \quad (14b)$$

The fourth-order cyclostationarity of $\mathbf{x}[n]$ implies that both $\mathbf{B}_1[n]$ and $\mathbf{B}_2[n]$ are time-invariant. Furthermore, by the mutual independence of $\mathbf{x}[n_1]$ and $\mathbf{x}[n_2]$ for all $n_1 \neq n_2$, it follows that $\mathbf{B}[n] = \frac{1}{N_0^2}\left(\mathbf{B}_1[n] + \mathbf{B}_2[n]\right)$. It follows that *AS7* is satisfied for this scenario.

We emphasize that assumptions *AS2-AS7* are made only to facilitate performance analysis and *are not needed for the derivation of the TA-LMS algorithm in Subsection III-B*, which only assumes that the signals are jointly proper-complex wide-sense cyclostationary.

B. Performance Measures

We begin with definitions for the error measures. Denote the coefficients error vector w.r.t. the MTA-MSE filter, $\mathbf{h}_{\text{opt}}$, obtained from (5), by

$$\bar{\mathbf{h}}\left[n\right] \triangleq \mathbf{h}_{\text{opt}} - \mathbf{h}\left[n\right], \quad (15)$$

the error between the output of the adaptive filter $\hat{d}[n]$ and the SOI $d[n]$ by

$$e[n] \triangleq d\left[n\right] - \hat{d}\left[n\right] = d\left[n\right] - \mathbf{h}^{H}\left[n\right]\mathbf{x}\left[n\right], \quad (16a)$$

and the error between the LMTA-MSE estimator and $d[n]$ at time index n by

$$e_{\text{opt}}[n] \triangleq d\left[n\right] - \mathbf{h}_{\text{opt}}^{H}\mathbf{x}\left[n\right]. \quad (16b)$$

The excess TA-MSE is defined as the difference between the TA-MSE of the adaptive filter and that of the LMTA-MSE estimator, namely,

$$\xi[n] \triangleq \left\langle \mathbb{E}\left\{\left|e\left[n\right]\right|^2\right\} - \mathbb{E}\left\{\left|e_{\text{opt}}\left[n\right]\right|^2\right\}\right\rangle_{N_0}. \quad (17)$$

In *transient performance analysis*, we study the conditions that guarantee convergence of the coefficients vector of the adaptive filter $\mathbf{h}[n]$, as well as characterize the behavior of the algorithm prior to convergence. In the analysis, we consider the following two types of convergence:

Definition 1 (Convergence in the Mean): An adaptive filter with coefficients error vector $\bar{\mathbf{h}}[n]$ *is said to be* mean convergent *if* $\lim_{n\to\infty}\mathbb{E}\left\{\bar{\mathbf{h}}\left[n\right]\right\} = \mathbf{0}_{M\times 1}$ *[3, Ch. 23.2].*

Definition 2 (Mean-Square Stability): An adaptive filter with coefficients error vector $\bar{\mathbf{h}}[n]$ *is said to be* mean-square stable *if* $\mathbb{E}\left\{\left\|\bar{\mathbf{h}}\left[n\right]\right\|^2\right\}$ *is convergent [2, Ch. 9.4] [3, Ch. 23.4].*

In *steady-state performance analysis*, we characterize the asymptotic behavior of the algorithm. Based on Defs. 1 and 2, steady-state is defined as follows:

Definition 3 (Steady-State): An adaptive filter is said to operate in steady-state if it is mean convergent, mean-square stable, and $n \to \infty$ *[3, Definition 15.1].*

The steady-state performance of the TA-LMS algorithm is characterized via the *steady-state excess TA-MSE*, defined as

$$\xi_s \triangleq \lim_{n\to\infty}\xi[n]. \quad (18)$$

C. Performance Analysis

To facilitate the performance analysis of the TA-LMS algorithm, we first introduce several quantities. Then, we present the transient performance analysis followed by the steady-state performance analysis.

1) Definitions: The quantities used in the performance analysis are summarized in Table II.

Lemma 2: The quantities defined in Table II satisfy the following properties:

- P1 *The random vector* $\tilde{\mathbf{g}}[n]$ *defined in (19b) satisfies* $\mathbb{E}\{\tilde{\mathbf{g}}[n]\} = \mathbf{0}_{M\times 1}$.
- P2 *The matrix* $\mathbf{F}[n]$ *defined in (19m) is time-invariant, and we set* $\mathbf{F}[n] = \mathbf{F}$.
- P3 *The random vectors* $\mathbf{z}_{\mathbf{xv}}[n]$ *and* $\mathbf{z}_{\mathbf{xv}}^{s}[n]$ *defined in (19i) and (19j), respectively, are equal* $\forall n \geq \eta$, *where* η *is defined in* (12).
- P4 *The random vectors* $\mathbf{p}_{\mathbf{xv}}[n]$ *and* $\mathbf{z}_{\mathbf{xv}}^{s}[n]$, *defined in (19h) and (19j), respectively, are periodic with period* N_0.

[A proof is given in Appendix A]

2) Transient Performance: We first characterize the time-evolution of the mean of the coefficients error vector:

Lemma 3: The mean of the TA-LMS coefficients error vector is given by

$$\mathbb{E}\left\{\bar{\mathbf{h}}\left[n\right]\right\} = \left(\tilde{\mathbf{R}}_{\mathbf{x}}\right)^{n}\mathbb{E}\left\{\bar{\mathbf{h}}\left[0\right]\right\}, \quad n \geq 0, \quad (20)$$

where $\tilde{\mathbf{R}}_{\mathbf{x}}$ *is defined in (19e).*

[A proof is given in Appendix B]

TABLE II
QUANTITIES USED IN THE PERFORMANCE ANALYSIS OF THE TA-LMS ALGORITHM

Definition		Type
$\mathbf{g}[n] \triangleq \mathbf{h}_M[n] - \mathbf{h}_{\text{opt}}$	$(19a)$	$M \times 1$ vector
$\tilde{\mathbf{g}}[n] \triangleq \frac{1}{N_0} \sum_{k=0}^{N_0-1} \mathbf{x}[n-k]\, \mathbf{x}^H[n-k]\, \mathbf{g}[n-k]$	$(19b)$	$M \times 1$ random vector
$\mathbf{R}_{\mathbf{x}}[n] \triangleq \mathbf{I}_M - \mu \mathbf{X}[n]\, \mathbf{X}^H[n]$	$(19c)$	$M \times M$ random matrix
$\mathbf{L}_k^n \triangleq \prod_{l=k}^{n} \mathbf{R}_{\mathbf{x}}[l]$ for $k \leq n$ and $\mathbf{L}_k^n \triangleq \mathbf{I}_M$ for $k > n$	$(19d)$	$M \times M$ random matrix
$\tilde{\mathbf{R}}_{\mathbf{x}} \triangleq \mathbb{E}\{\mathbf{R}_{\mathbf{x}}[n]\} = \mathbf{I}_M - \mu \underset{\sim}{\mathbf{C}}_{\mathbf{x}}$	$(19e)$	$M \times M$ matrix
$\mathbf{v}[n] \triangleq \frac{1}{\sqrt{N_0}} [v[n], v[n-1], \ldots, v[n-N_0+1]]^T$	$(19f)$	$N_0 \times 1$ random vector
$\mathbf{c}_{\mathbf{v}}[n,l] \triangleq \text{vec}\left(\mathbb{E}\left\{\mathbf{v}[n]\, \mathbf{v}^H[n-l]\right\}\right)$	$(19g)$	$N_0^2 \times 1$ vector
$\mathbf{p}_{\mathbf{xv}}[n] \triangleq \text{vec}\left(\mathbb{E}\left\{\tilde{\mathbf{g}}[n]\, \tilde{\mathbf{g}}^H[n]\right\}\right) + \mathbb{E}\{\mathbf{X}^*[n] \otimes \mathbf{X}[n]\}\, \mathbf{c}_{\mathbf{v}}^*[n,0]$	$(19h)$	$M^2 \times 1$ vector
$\mathbf{z}_{\mathbf{xv}}[n] \triangleq \left(\sum_{k=1}^{n} \mathbf{c}_{\mathbf{v}}^T[n,k]\, \mathbb{E}\left\{\left(\mathbf{L}_{n-k+1}^n \mathbf{X}[n-k]\right)^T \otimes \mathbf{X}^H[n]\right\}\right)^H$	$(19i)$	$M^2 \times 1$ vector
$\mathbf{z}_{\mathbf{xv}}^s[n] \triangleq \left(\sum_{k=1}^{\eta} \mathbf{c}_{\mathbf{v}}^T[n,k]\, \mathbb{E}\left\{\left(\mathbf{L}_{n-k+1}^n \mathbf{X}[n-k]\right)^T \otimes \mathbf{X}^H[n]\right\}\right)^H$	$(19j)$	$M^2 \times 1$ vector
$\mathbf{z}_{\mathbf{xg}}[n] \triangleq \left(\mathbb{E}\left\{\bar{\mathbf{h}}^T[0]\right\} \left(\mathbf{I}_M - \mu \underset{\sim}{\mathbf{C}}_{\mathbf{x}}^T\right)^n \mathbb{E}\left\{\mathbf{R}_{\mathbf{x}}^T[n] \otimes \tilde{\mathbf{g}}^H[n]\right\}\right)^H$	$(19k)$	$M^2 \times 1$ vector
$\mathbf{A} \triangleq \left(\underset{\sim}{\mathbf{C}}_{\mathbf{x}}^T \otimes \mathbf{I}_M\right) + \left(\mathbf{I}_M \otimes \underset{\sim}{\mathbf{C}}_{\mathbf{x}}\right)$	$(19l)$	$M^2 \times M^2$ matrix
$\mathbf{F}[n] \triangleq \mathbb{E}\left\{\mathbf{R}_{\mathbf{x}}^T[n] \otimes \mathbf{R}_{\mathbf{x}}[n]\right\} = \mathbf{I}_{M^2} - \mu \mathbf{A} + \mu^2 \mathbf{B} = \mathbf{F}$ (Lemma 2, *P2*)	$(19m)$	$M^2 \times M^2$ matrix
$\mathbf{H} \triangleq \frac{1}{2} \begin{bmatrix} \mathbf{A} & -\mathbf{B} \\ 2\mathbf{I}_{M^2} & \mathbf{0}_{M^2 \times M^2} \end{bmatrix}$	$(19n)$	$2M^2 \times 2M^2$ matrix
$\mathbf{P} \triangleq \frac{1}{N_0} \sum_{k=0}^{N_0-1} \sum_{l=1}^{N_0} \mathbf{c}_{\mathbf{x}}[k] \left(\mathbf{p}_{\mathbf{xv}}\left[((k-l))_{N_0}\right] + 2\mathbf{z}_{\mathbf{xv}}^s\left[((k-l))_{N_0}\right]\right)^H \mathbf{F}^{l-1}$	$(19o)$	$M^2 \times M^2$ matrix

Note that the proof of Lemma 3, detailed in Appendix B, strongly relies of on property *P1* in Lemma 2. Property *P1* is immediate for stationary signals, as for such signals the LMMSE filter and the MTA-MSE filter coincide, thus $\mathbf{g}[n]$ vanishes. For cyclostationary signals, this is not the case. Consequently, although (20) bears similarity to the mean behavior of the LMS for stationary signals [3, Ch. 23.2], it is not a trivial extension of the stationary case. In order to characterize the time-evolutions of the MSD of the coefficients error vector, $\mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n]\right\|^2\right\}$, and of the excess TA-MSE, $\xi[n]$, we first derive a recursive relationship for the mean of the weighted squared Euclidean norm of $\bar{\mathbf{h}}[n]$, which is stated in the following proposition:

Proposition 1: For any $M^2 \times 1$ vector $\mathbf{q}$ such that $\mathbf{Q} = \text{vec}^{-1}\{\mathbf{q}\}$ is Hermitian positive semi-definite, the coefficients error vector (15) *satisfies the following recursion for $n \geq 0$:*

$$\mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n+1]\right\|_{\mathbf{q}}^2\right\} = \mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n]\right\|_{\mathbf{Fq}}^2\right\} + \mu^2 \mathbf{p}_{\mathbf{xv}}^H[n]\mathbf{q} - 2\mu \cdot \text{Re}\left\{\left(\mathbf{z}_{\mathbf{xg}}^H[n] - \mu \cdot \mathbf{z}_{\mathbf{xv}}^H[n]\right)\mathbf{q}\right\}, \quad (21)$$

where $\mathbf{F}$, $\mathbf{p}_{\mathbf{xv}}$, $\mathbf{z}_{\mathbf{xg}}$, and $\mathbf{z}_{\mathbf{xv}}$ are defined in (19m), (19h), (19k), and (19i), respectively.

[A proof is given in Appendix C]

Note that when the period is $N_0 = 1$, i.e., when the considered signals are jointly WSS, then the TA-LMS specializes to the standard LMS and we obtain $\mathbf{g}[n] = \mathbf{0}_{M \times 1}$, $\forall n \in \mathbb{N}$, and consequently, $\mathbf{z}_{\mathbf{xg}}[n]$ defined in (19k) is equal to $\mathbf{0}_{M^2 \times 1}$, $\forall n \in \mathbb{N}$. Additionally, when the noise signal $v[n]$ is temporally uncorrelated, then $\mathbf{z}_{\mathbf{xv}}[n]$ defined in (19i) is equal to $\mathbf{0}_{M^2 \times 1}$, $\forall n \in \mathbb{N}$, and consequently, (21) specializes to the variance relationship for the standard LMS with general WSS inputs [3, eq. (24.11)]. The third summand in (21) is therefore unique for cyclostationary signals and thus (21) cannot be obtained by extending its stationary counterpart [3, eq. (24.11)]. Applying the recursion (21) n times, we obtain a non-recursive relationship for the mean weighted squared Euclidean norm of $\bar{\mathbf{h}}[n]$:

Corollary 1: For any $M^2 \times 1$ vector $\mathbf{q}$ for which $\mathbf{Q} = \text{vec}^{-1}\{\mathbf{q}\}$ is a Hermitian positive semi-definite matrix, the weighted squared Euclidean norm of the coefficients error vector of the TA-LMS algorithm at time $n \geq 1$ is given by the following non-recursive relationship:

$$\mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n]\right\|_{\mathbf{q}}^2\right\} = \mathbb{E}\left\{\left\|\bar{\mathbf{h}}[0]\right\|_{\mathbf{F}^n\mathbf{q}}^2\right\} + \mu^2 \sum_{l=1}^{n} \mathbf{p}_{\mathbf{xv}}^H[n-l]\, \mathbf{F}^{l-1}\mathbf{q} - 2\mu \cdot \text{Re}\left\{\sum_{l=1}^{n} \left(\mathbf{z}_{\mathbf{xg}}^H[n-l] - \mu \cdot \mathbf{z}_{\mathbf{xv}}^H[n-l]\right)\mathbf{F}^{l-1}\mathbf{q}\right\}, \quad (22)$$

where $\mathbf{F}$, $\mathbf{p}_{\mathbf{xv}}$, $\mathbf{z}_{\mathbf{xg}}$, and $\mathbf{z}_{\mathbf{xv}}$ are defined in (19m), (19h), (19k), and (19i), respectively.

Note that Corollary 1 can be used to characterize the time-evolution of the MSD of the coefficients vector by setting $\mathbf{q} = \mathrm{vec}\,(\mathbf{I}_M)$. In the following theorem we characterize the time-evolution of the excess TA-MSE (17):

Theorem 1: For all $n \geq N_0 - 1$, the excess TA-MSE of the TA-LMS algorithm is given by

$$\xi[n] = \frac{1}{N_0} \sum_{k=n-N_0+1}^{n} \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[k] \right\|^2_{\mathbf{c}_{\mathbf{x}}\left[((k))_{N_0}\right]} \right\} + 2\mathrm{Re}\left\{ \mathbb{E}\left\{ \bar{\mathbf{h}}^H[0] \right\} \times \left(\tilde{\mathbf{R}}_{\mathbf{x}} \right)^k \mathbf{C}_{\mathbf{x}}\left[((k))_{N_0}\right] \mathbf{g}\left[((k))_{N_0}\right] \right\}, \tag{23}$$

where $\mathbf{c}_{\mathbf{x}}[n]$, $\tilde{\mathbf{R}}_{\mathbf{x}}$, and $\mathbf{g}[n]$ are defined in (3b)*, (19e), and (19a), respectively, and the term $\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[k] \right\|^2_{\mathbf{c}_{\mathbf{x}}\left[((k))_{N_0}\right]} \right\}$ can be computed using* (22) *with $\mathbf{q} = \mathbf{c}_{\mathbf{x}}\left[((k))_{N_0}\right]$.*

[A proof is given in Appendix D]

Next, we study the mean convergence and mean-square stability of the TA-LMS algorithm. Corollary 2, which follows from Lemma 3, states that the TA-LMS is mean convergent if and only if the step-size μ takes values in the interval $\left(0, \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}\right)$, where $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ is defined in (3c):

Corollary 2: Following Def. 1, the TA-LMS algorithm is mean convergent if and only if the step-size μ satisfies:

$$0 < \mu < \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}. \tag{24}$$

[A proof is given in Appendix E]

In the following theorem we obtain a sufficient condition on the step-size μ under which the TA-LMS algorithm is mean-square stable.

Theorem 2: Let $\mathbf{A}$, $\mathbf{B}$, $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$, and $\mathbf{H}$ be the matrices defined in (19l), (13)*,* (3c)*, and (19n), respectively. Assume that $\mathbf{H}$ has at least one real-valued positive eigenvalue, then the TA-LMS algorithm is both mean convergent and mean-square stable if*

$$0 < \mu < \min\left\{ \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}, \frac{1}{\lambda_{\max}\left(\mathbf{A}^{-1}\mathbf{B}\right)}, \frac{1}{\lambda_{\max}(\mathbf{H})} \right\}. \tag{25a}$$

If $\mathbf{H}$ does not have any real positive eigenvalues then the TA-LMS is both mean convergent and mean-square stable if

$$0 < \mu < \min\left\{ \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}, \frac{1}{\lambda_{\max}\left(\mathbf{A}^{-1}\mathbf{B}\right)} \right\}. \tag{25b}$$

[A proof is given in Appendix F]

3) Steady-State Performance: In order to derive the steady-state excess TA-MSE (18), we first show that it can be obtained as the following limit:

Lemma 4: The steady-state excess TA-MSE (18) *can be written as*

$$\xi_s = \lim_{n\to\infty} \left\langle \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[n] \right\|^2_{\mathbf{c}_{\mathbf{x}}[n]} \right\} \right\rangle_{N_0}. \tag{26}$$

[A proof is given in Appendix G]

The steady-state excess TA-MSE of the TA-LMS algorithm is explicitly stated in the following corollary, which follows from Lemma 4 and Proposition 1:

Corollary 3 (Steady-State Excess TA-MSE): When (25) *is satisfied, the steady-state excess TA-MSE of the TA-LMS algorithm is given by*

$$\xi_s = \mu^2 \cdot \mathrm{Re}\left\{ \mathrm{Tr}\left\{ \mathbf{P}\left(\mathbf{I}_{M^2} - \mathbf{F}^{N_0} \right)^{-1} \right\} \right\}, \tag{27}$$

where $\mathbf{F}$ and $\mathbf{P}$ are defined in (19m) and (19o), respectively.

[A proof is given in Appendix H]

Similarly to the discussion following Proposition 1, it can be shown that when the period is $N_0 = 1$ (the considered signals are jointly WSS) and $v[n]$ is temporally uncorrelated, then (27) specializes to the excess MSE of the standard LMS with general WSS inputs in [3, Th. 24.1].

V. Simulation Examples

In this section we evaluate the performance of the TA-LMS algorithm and demonstrate the theoretical results presented in Section IV. The simulations study consists of two parts: First, in Subsection V-A we compare the theoretical and empirical performance measures when assumptions *AS1-AS7* are satisfied. In Subsection V-B,we consider an NB-PLC channel estimation scenario, in which the SOI and the input signal are jointly cyclostationary, and evaluate the performance of the algorithm. Additionally, for each scenario we compare the TA-LMS performance to that of the standard LMS of *equal complexity* and that of the standard LMS of *equal dimensionality*. All empirical performance measures were obtained by averaging over 10000 Monte Carlo simulations. In all simulation examples the theoretical TA-MSE and steady-state TA-MSE are obtained by summing the TA-MSE of the LMTA-MSE estimator $J\left(\mathbf{h}_{\mathrm{opt}}\right)$ (4) with the excess TA-MSE (23) and the steady-state excess TA-MSE (27), respectively.

A. Example 1: Validating the Theoretical Results When AS1-AS7 are Satisfied

We first demonstrate that the theoretical TA-MSE and MSD in (23) and (22), respectively, accurately match the corresponding empirically computed values, which verifies the theoretical analysis. To that aim, we consider an input signal $\mathbf{x}[n]$ and an SOI $d[n]$ which satisfy *AS1-AS7*: Let $\phi[n]$ be an i.i.d. process, uniformly distributed over $[0, 2\pi)$ for each n, and let $r[n]$ be given by $r[n] = \left(1 + 0.5\sin\left(\frac{2\pi n}{N_0}\right)\right) e^{j\phi[n]}$, with $N_0 = 20$. For implementing FRESH filtering with $K = 3$ and $L = 1$, the vector input signal $\mathbf{x}[n]$ is constructed from $r[n]$ as described in the second row of Table I, by setting $(\mathbf{x}[n])_k = r[n]e^{j\alpha_k n}$, $k \in \{0, 1, 2\}$, and $\{\alpha_0, \alpha_1, \alpha_2\} = \{-1, 0, 1\} \cdot \frac{2\pi}{N_0}$. It is straightforward to confirm that the input vector $\mathbf{x}[n]$ constructed above satisfies the following properties: *1)* For each finite positive integer $i_{\max}$ and a set of integer indexes $\{n_i\}_{i=1}^{i_{\max}}$ such that $n_i \neq n_j$, $\forall i \neq j$, the RVs $\{\mathbf{x}[n_i]\}_{i=1}^{i_{\max}}$ are mutually independent; *2)* The N-th order moment of $\mathbf{x}[n]$ is periodic with period N_0 for any positive integer N; *3)* The matrix $\mathbf{x}[n]\mathbf{x}^H[n]$ is deterministic $\forall n \in \mathbb{N}$; and *4)* The matrix $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ defined in (3c) is non-singular. It follows that the input signal satisfies *AS1*, *AS3*, *AS4*, *AS6*, and *AS7*. Furthermore, we set the coefficients vector $\mathbf{h}_M[n]$ to satisfy $(\mathbf{h}_M[n])_k = \left(1 + \frac{((n))_{N_1}}{N_1}\right) e^{-|k-1|}$, $k \in \{0, 1, 2\}$,

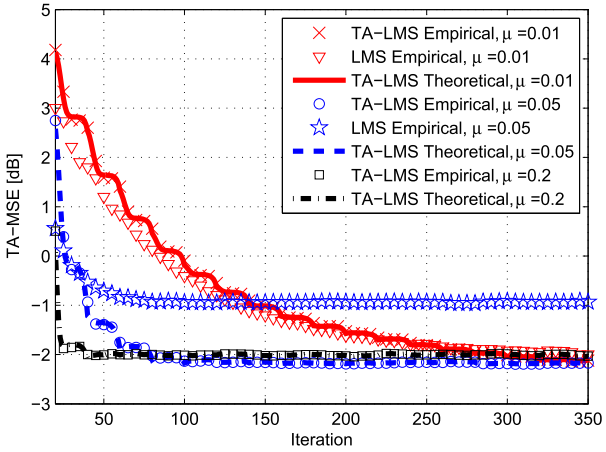


Fig. 2. The theoretical and the empirical TA-MSEs of the TA-LMS algorithm as compared to the standard LMS algorithm of equal complexity for three step-sizes versus the number of iterations when *AS1*-*AS7* are satisfied.

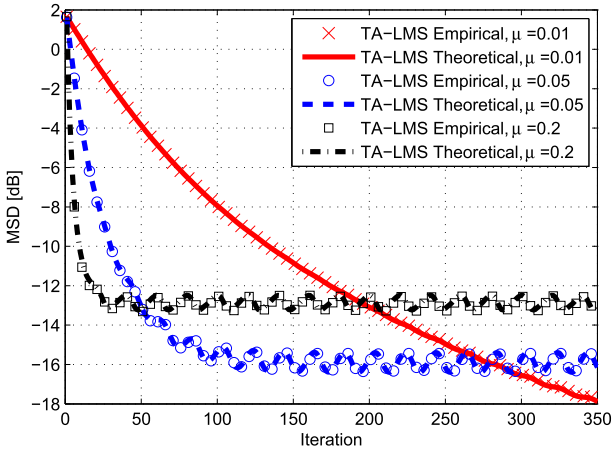


Fig. 3. The theoretical and the empirical MSDs in the filter coefficients of the TA-LMS algorithm for three step-sizes versus the number of iterations when *AS1*-*AS7* are satisfied.

with $N_1 = 10$. The noise process $\upsilon[n]$ in (11) is set to be a zero mean temporally uncorrelated Gaussian process that is mutually independent of $\mathbf{x}[n]$. Note that with this noise setting, *AS2* is satisfied. Define next SNR $\triangleq \frac{\langle \mathbb{E}\{|d[n]-\upsilon[n]|^2\}\rangle_{N_0}}{\langle \mathbb{E}\{|\upsilon[n]|^2\}\rangle_{N_0}}$. Letting the time-varying variance of $\upsilon[n]$ be $\mathbb{E}\left\{|\upsilon[n]|^2\right\} \triangleq \sigma_\upsilon^2[n] = \gamma_\upsilon \left(1 + \frac{1}{((n))_{N_2}+1}\right)$, where $N_2 = 5$ and γ_υ is set to achieve a signal-to-noise ratio (SNR) of 7 dB, and letting $\mathbb{E}\{\upsilon[n+l]\upsilon^*[n]\} = 0$ for all $l \neq 0$, it follows that *AS5* is satisfied. Note that $d[n]$ and $\mathbf{x}[n]$ are jointly cyclostationary with period N_0. We conclude that for this scenario all the assumptions stated in Subsection IV-A are satisfied.

Figs. 2 and 3 depict the theoretical TA-LMS and MSD, respectively, along with their empirically computed values versus the number of iterations. To demonstrate the advantages of the TA-LMS over the standard LMS in a jointly cyclostationary environment, we include the TA-MSE of the LMS in Fig. 2. To make a fair comparison, the number of taps for the standard LMS was selected such that both algorithms have the same computational complexity. Note that

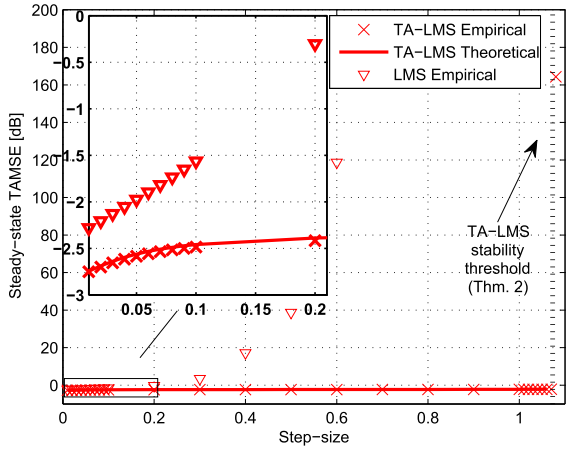


Fig. 4. The theoretical and the empirical steady-state TA-MSEs of the TA-LMS algorithm as compared to the standard LMS algorithm of equal dimensionality versus the step-size when *AS1*-*AS7* are satisfied.

for $M = 3$, the TA-LMS requires $6M^2 + 8M = 78$ real multiplications per iteration; Accordingly, we set the number of taps in the standard LMS to $M_{\text{LMS}} = 9$, thus requiring $8M_{\text{LMS}} + 2 = 74$ real multiplications per iteration. It follows that the LMS algorithm implements a FRESH filter with $K_{\text{LMS}} = 3$ branches where at each branch there is an FIR filter with $L_{\text{LMS}} = 3$ taps. The input to the LMS is obtained by setting $(\mathbf{x}_{\text{LMS}}[n])_{k\cdot L_{\text{LMS}}+l} = r[n-l]e^{j\alpha_k(n-l)}$, $k, l \in \{0, 1, 2\}$. The results are presented for step-sizes $\mu = \{0.01, 0.05, 0.2\}$. The TA-MSE performance of the LMS for $\mu = 0.2$ is not depicted in Fig. 2 since it diverges for this step-size. We note that when the step-size is sufficiently small, i.e., $\mu = 0.01$, then the LMS has roughly the same TA-MSE as the TA-LMS. However, when the step-size increases to $\mu = 0.05$ the TA-MSE performance of the standard LMS deteriorates while the TA-LMS demonstrates rapid convergence to lower TA-MSE. Increasing the step-size further to $\mu = 0.2$, results in the LMS becoming non-convergent while the TA-LMS demonstrates very fast convergence to low TA-MSEs. Note that the MSD of the standard LMS is not depicted in Fig. 3 since maintaining the same computational complexity results in different number of filter coefficients to be estimated by the LMS and TA-LMS estimators. In Fig. 4, the theoretical steady-state TA-MSE of the TA-LMS, computed using (27), is compared to its empirical value for various step-sizes and to the theoretical stability threshold, computed using (25). In order to demonstrate that the performance gain of the TA-LMS over the standard LMS of equal complexity does not result from the reduced dimensionality, we also depict in Fig. 4 the steady-state TA-MSE of the standard LMS with the same number of taps as the TA-LMS, i.e., $M_{\text{LMS}} = 3$. Observing Fig. 4, we note that the stability of the algorithm is accurately predicted using the step-size region in (25), and that the stability region of the LMS is substantially smaller. Since for different algorithms, equal step-size does not imply equal convergence rate, we also depict in Fig. 5 the empirical steady-state TA-MSE versus the convergence rate of the TA-LMS, the LMS, and the B-LMS, for equal number of taps, i.e., same dimensionality. We consider a filter to converge

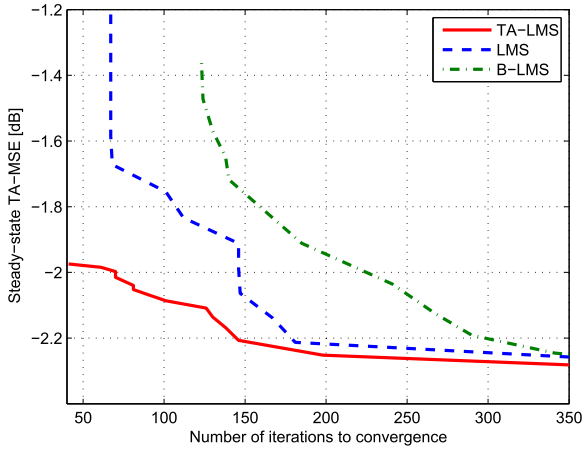


Fig. 5. The empirical steady-state TA-MSEs of the TA-LMS algorithm as compared to the standard LMS and the B-LMS algorithms of equal dimensionality versus the convergence rate when *AS1*-*AS7* are satisfied.

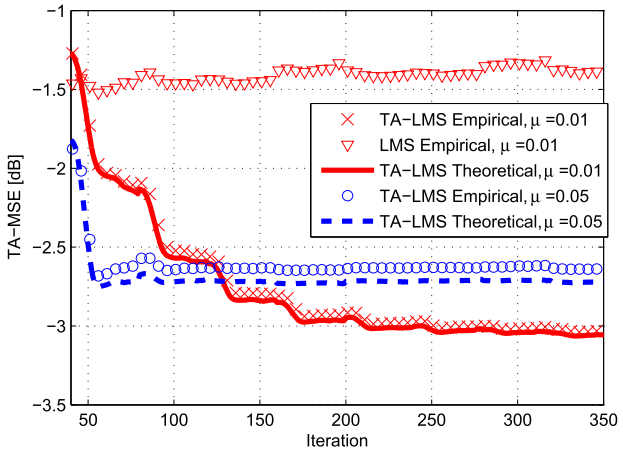


Fig. 7. The theoretical and the empirical TA-MSEs of the TA-LMS algorithm as compared to the LMS algorithm of equal complexity for $\mu = \{0.01, 0.05\}$ versus the number of iterations in a channel estimation scenario.

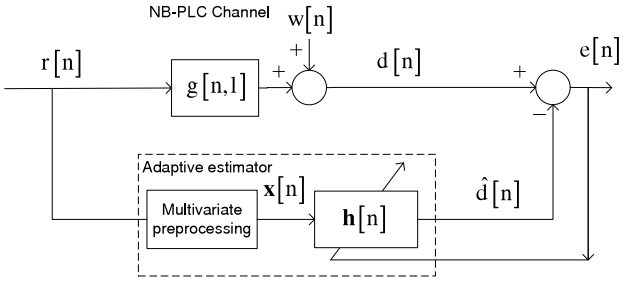


Fig. 6. A schematic description of the adaptive channel estimation in NB-PLC scenario.

when the TA-MSE variations are smaller than 2%. We observe in Fig. 5 that for slow convergence rate, i.e., $\mu \to 0$, all tested adaptive algorithms obtain nearly the same performance, with only a slight advantage to the TA-LMS. However, for fast convergence rates, the TA-LMS significantly outperforms the LMS, while the convergence rate of the B-LMS is very slow at all values of steady-state TA-MSE.

To conclude, in this example we observe an excellent agreement between the theoretical and empirical performance measures of the TA-LMS. Furthermore, the TA-LMS demonstrates clear and substantial superiority over the conventional LMS in terms of both steady-state TA-MSE and convergence speed.

B. Example 2: Application to Channel Estimation in NB-PLC

We now consider the application of the TA-LMS algorithm in a practical NB-PLC channel estimation scenario. Let $r[n]$ be the input signal to the NB-PLC channel and $d[n]$ be the resulting channel output signal, which corresponds to the SOI in this scenario. The signal $r[n]$ is a passband OFDM signal s.t. each OFDM symbol consists of 36 data subcarriers, each modulated via a QPSK constellation, to which a cyclic prefix of four samples is added in the time domain. In [19, Sec. II-C] it is shown that $r[n]$ is cyclostationary with period $N_0 = 40$. The NB-PLC channel output is modeled by $d[n] = \sum_{l=0}^{\infty} g[n,l] r[n-l] + w[n]$ [32, Sec. III-C], where $g[n,l]$ is a real passband LPTV filter with period N_0, generated based on the theory of transmission lines as in [32] and following the IEEE P1901.2 standard [11]. The noise $w[n]$ is a real passband additive cyclostationary Gaussian noise (ACGN) having the same period N_0, generated using the model [10, Sec. IV.A], with a set of typical parameters which corresponds to low voltage site 11 (LV11) in [11, Appendix G]. This scenario is depicted in Fig 6. The SOI $d[n]$ is estimated via a FRESH filter with $K = 3$ branches corresponding to frequency shifts $\{\alpha_0, \alpha_1, \alpha_2\} = \{-1, 0, 1\} \cdot \frac{2\pi}{N_0}$. Each branch has an FIR filter with $L = 3$ taps and the overall dimensionality is $M = 9$. The input to the adaptive filter, $\mathbf{x}[n]$, is obtained from $r[n]$ as described in the second row of Table I. The MMSE filter $\mathbf{h}_M[n]$ in (11) is obtained from the orthogonality principle, i.e., $\mathbb{E}\{\mathbf{x}[n] d^*[n]\} = \mathbb{E}\left\{\mathbf{x}[n] \mathbf{x}^H[n]\right\} \mathbf{h}_M[n]$, and $v[n]$ is obtained as the estimation error for $\mathbf{h}_M[n]$. The SNR for the simulations is defined as $\text{SNR} \triangleq \frac{\langle \mathbb{E}\{|d[n]-w[n]|^2\} \rangle_{N_0}}{\langle \mathbb{E}\{|w[n]|^2\} \rangle_{N_0}}$. As NB-PLC channels are typically characterized by low SNR, we used SNR = 0 dB in the simulations.

Figs. 7 and 8 depict the theoretical TA-MSE and MSD, respectively, compared to their empirical values versus the number of iterations for step-sizes $\mu = \{0.01, 0.05\}$. Fig. 7 also depicts the TA-MSE of the standard LMS, for the same computational complexity. Note that for $M = 9$, the TA-LMS requires $6M^2 + 8M = 558$ real multiplications per iteration, thus we set the number of taps in the standard LMS to $M_{\text{LMS}} = 69$, thus requiring $8M_{\text{LMS}} + 2 = 554$ real multiplications per iteration. Therefore, the standard LMS implements a FRESH filter with $K_{\text{LMS}} = 3$ branches s.t. each branch has an FIR filter with $L_{\text{LMS}} = 23$ taps. The input $\mathbf{x}[n]$ is obtained from $r[n]$ as described in the second row of Table I. The TA-MSE performance of the LMS for $\mu = 0.05$ is not depicted in Fig. 6 since the algorithm does not converge for this step-size and for the same computational complexity. Moreover, we note that even when the step-size is as small as $\mu = 0.01$, the standard LMS obtains considerably

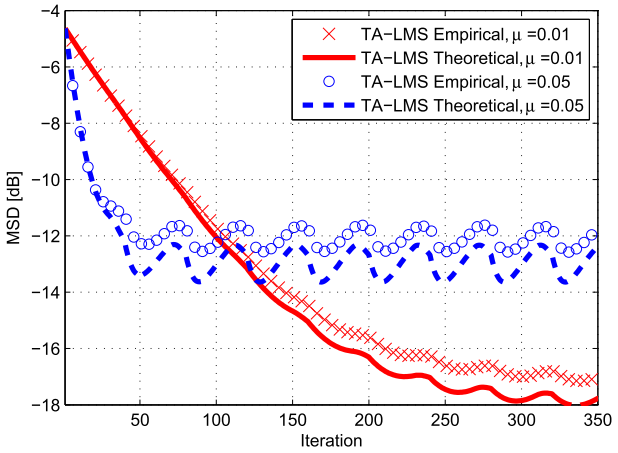


Fig. 8. The theoretical and the empirical MSDs in the filter coefficients of the TA-LMS algorithm for $\mu = \{0.01, 0.05\}$ versus the number of iterations in a channel estimation scenario.

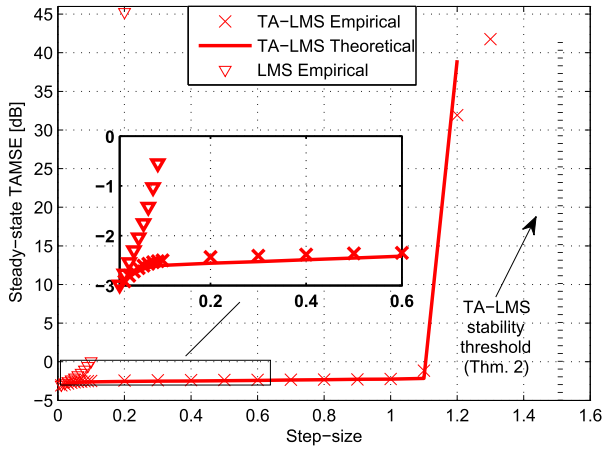


Fig. 9. The theoretical and empirical steady-state TA-MSEs of the TA-LMS algorithm as compared to the standard LMS algorithm of equal dimensionality versus the step-size in a channel estimation scenario.

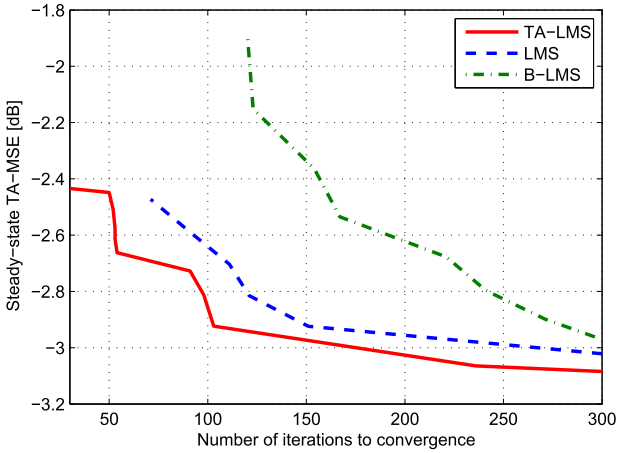


Fig. 10. The empirical steady-state TA-MSEs of the TA-LMS algorithm as compared to the standard LMS and the B-LMS algorithms of equal dimensionality versus the convergence rate in a channel estimation scenario.

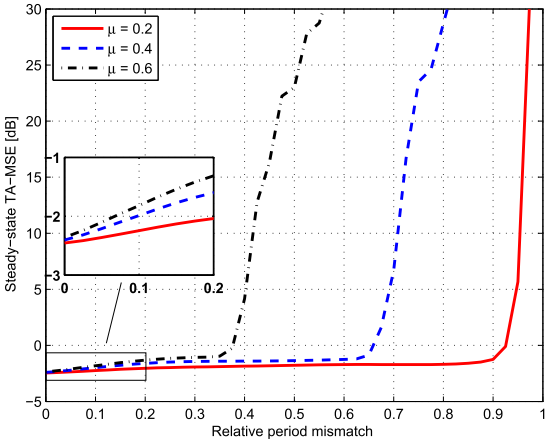


Fig. 11. The empirical steady-state TA-MSEs of the TA-LMS algorithm for step-sizes $\mu \in \{0.2, 0.4, 0.6\}$ versus the period error in a channel estimation scenario.

higher TA-MSE compared to the TA-LMS. Lastly, we note that although the entire set of assumptions *AS1-AS7* is not satisfied for this scenario, there is a very good agreement between the theoretical and empirical performance measures.

In Fig. 9 the theoretical steady-state TA-MSE of the TA-LMS computed using (27) versus the step-size is compared to its empirical value, and to the stability threshold computed using (25). Fig. 9 also depicts the empirical steady-state TA-MSE of the standard LMS with the same number of taps as the TA-LMS, i.e., $M_{\text{LMS}} = 9$. Observe that the theoretical steady-state TA-MSE provides an accurate prediction of the empirical steady-state TA-MSE. Moreover, observe that the stability region of the standard LMS is substantially smaller compared to that of the TA-LMS, and that within this region, i.e., $\mu < 0.1$, the TA-LMS achieves lower steady-state TA-MSE as compared the standard LMS with equal dimensionality. Finally, we observe that the convergence threshold of Thm. 2 provides a good prediction of the stability threshold of the TA-LMS algorithm for this scenario. The gap between simulation and analysis arises from the fact that, unlike the scenario considered in Subsection V-A, assumptions *AS1-AS7* are not satisfied here. Next, we compare in Fig. 10 the steady-state TA-MSE versus the convergence rate of the TA-LMS, the LMS, and the B-LMS for equal number of taps, i.e., same dimensionality. Observing Fig. 10, we note that the TA-LMS outperforms the LMS and the B-LMS for all convergence rates, and that the performance gap increases as the convergence rate increases. We also note that the TA-LMS is the only algorithm which is able to achieve fast convergence rates.

Lastly, we study the effect of imperfect knowledge of the period of the second order statistical moments of the signals on the performance of the TA-LMS. To that aim, we apply the TA-LMS with period N_0' and numerically evaluate the performance of the TA-LMS algorithm when N_0' differs from N_0 (the actual period of the joint second order statistical moments of $\mathbf{x}[n]$ and $d[n]$). In Fig. 11 we compare the steady-state TA-MSE of the TA-LMS algorithm versus the relative period mismatch, defined as $\frac{N_0 - N_0'}{N_0}$, for $N_0' \in \{1, 2, \ldots, N_0\}$ and for step sizes $\mu \in \{0.2, 0.4, 0.6\}$. Observe that the imperfect knowledge of the periodicity of the signals has a more dominant effect in larger step-sizes: For $\mu = 0.6$ the TA-LMS fails to converge for relative period mismatch larger than 0.38, while for $\mu = 0.2$ the TA-LMS still

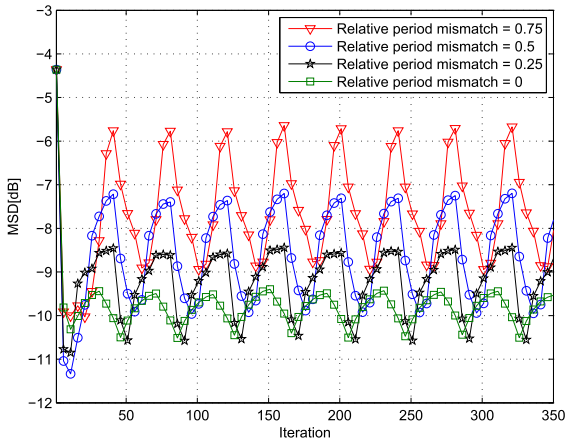


Fig. 12. The empirical MSDs of the TA-LMS algorithm for period errors $\{0, 0.25, 0.5, 0.75\}$ and $\mu = 0.2$ versus the number of iterations in a channel estimation scenario.

converges at relative period mismatch as large as 0.9. The relationship between the steady-state TA-MSE, the relative period mismatch, and the step-size, can be explained by noting that for $N_0' = 1$, the TA-LMS specializes the LMS. Since the stability region of the LMS is much smaller compared to the TA-LMS (see, e.g., Fig. 9), it follows that the stability region is expected to decrease as the relative period mismatch increases, which agrees with the results depicted in Fig. 11. Note that for relative period mismatch smaller than 0.38, there is only a slight loss in steady-state TA-MSE, e.g., for relative period period mismatch of 0.2, the TA-LMS converges for all the observed step-sizes, and the steady-state TA-MSE increases by less than 1 dB. In order to further examine the effect of relative period mismatch when convergence is obtained, we depict in Fig. 12 the MSD of the TA-LMS for relative period mismatch $\{0, 0.25, 0.5, 0.75\}$ at step-size $\mu = 0.2$. While convergence is obtained for all tested values of relative period mismatch (as also observed in Fig. 11), it can be seen that the average MSD and the variations of the MSD are increased when the relative period mismatch increases (recall that the MSD converges to a periodic sequence, see Appendix F). This simulations study demonstrates that while the performance of the TA-LMS is affected by imperfect knowledge of the period, the algorithm is, in general, very robust to small errors in the period, especially at small step-sizes.

VI. Conclusions

In this work, we rigorously studied optimal adaptive linear estimation of cyclostationary signals. We obtained the proper stochastic approximation of the TA-MSE minimizer and characterized its transient and steady-state performance. Sufficient conditions for convergence were provided. The performance analysis does not impose any specific distribution on the SOI and input signal. Naturally, when the signals are jointly WSS the TA-LMS algorithm specializes to the standard LMS algorithm. The simulation results showed that the theoretical analysis accurately describes the empirical performance of the algorithm, both in the transient phase and in steady-state. The advantages of our optimal approach over the ad hoc application of adaptive filters designed for stationary signals were demonstrated in a practical NB-PLC scenario.

Appendix

We begin by stating the following equalities and properties that will be used in the sequel:

R1 For any matrix triplet $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ of compatible dimensions, it holds that [33, Ch. 9.2]

$$\mathrm{vec}\left(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\right) = \left(\mathbf{A}_3^T \otimes \mathbf{A}_1\right)\mathrm{vec}\left(\mathbf{A}_2\right). \tag{28}$$

R2 For any pair of square matrices $\mathbf{A}_1, \mathbf{A}_2$ of identical dimensions, it holds that [33, Ch. 9.2]

$$\mathrm{Tr}\left\{\mathbf{A}_1^T\mathbf{A}_2\right\} = \mathrm{vec}\left(\mathbf{A}_1\right)^T\mathrm{vec}\left(\mathbf{A}_2\right). \tag{29}$$

A. Proof of Lemma 2

Recall that $\mathbf{x}[n]$ and $\upsilon[n]$ are zero mean. Applying the Hermitian transpose to (11), followed by a left multiplication of both sides by $\mathbf{x}[n]$, and then applying the stochastic expectation and using *AS2* and (19a) yields $\mathbb{E}\left\{\mathbf{x}\left[n\right]d^*\left[n\right]\right\} = \mathbb{E}\left\{\mathbf{x}\left[n\right]\mathbf{x}^H\left[n\right]\right\}\left(\mathbf{h}_{\mathrm{opt}} + \mathbf{g}[n]\right)$. Applying the time-averaging operator results in

$$\underset{\sim}{\mathbf{c}}_{\mathbf{x}d} = \underset{\sim}{\mathbf{C}}_{\mathbf{x}}\mathbf{h}_{\mathrm{opt}} + \left\langle \mathbb{E}\left\{\mathbf{x}\left[n\right]\mathbf{x}^H\left[n\right]\right\}\mathbf{g}\left[n\right]\right\rangle_{N_0}. \tag{30}$$

Since $\mathbf{h}_{\mathrm{opt}}$ is the TA-MSE optimal filter, by plugging (5) into (30) and recalling the definition of $\tilde{\mathbf{g}}[n]$ in (19b), it follows that $\mathbb{E}\left\{\tilde{\mathbf{g}}\left[n\right]\right\} = \left\langle \mathbb{E}\left\{\mathbf{x}\left[n\right]\mathbf{x}^H\left[n\right]\right\}\mathbf{g}\left[n\right]\right\rangle_{N_0} = \mathbf{0}_{M\times 1}$. This proves *P1*.

Property *P2* follows from the definition of $\mathbf{R}_{\mathbf{x}}[n]$ (19c) combined with *AS7*, as $\mathbf{F}[n]$ (19m) is given by $\mathbf{F}[n] = \mathbf{I}_{M^2} - \mu\mathbf{A} + \mu^2\mathbf{B}$, where $\mathbf{A}$ and $\mathbf{B}$ are defined in (19l) and (13), respectively.

Property *P3* follows from *AS5* since $\mathbf{c}_{\mathbf{v}}\left[n, l\right]$ defined in (19g) satisfies $\mathbf{c}_{\mathbf{v}}\left[n, l\right] = \mathbf{0}_{N_0^2\times 1}$, $\forall |l| > \eta$, where η is defined in (12). Therefore, it follows that $\mathbf{z}_{\mathbf{xv}}[n]$ and $\mathbf{z}^s_{\mathbf{xv}}[n]$ defined in (19i) and (19j), respectively, are equal $\forall n \geq \eta$.

Lastly, from the cyclostationarity of $\upsilon[n]$ combined with *AS6* it follows that $\mathbf{p}_{\mathbf{xv}}\left[n\right]$ and $\mathbf{z}^s_{\mathbf{xv}}[n]$, defined in (19h) and (19j), respectively, are periodic with period N_0 This proves *P4*. □

B. Proof of Lemma 3

By the TA-LMS update equation (10) and the definition of $\bar{\mathbf{h}}[n]$ in (15) it follows that

$$\bar{\mathbf{h}}[n+1] = \bar{\mathbf{h}}[n] - \mu\left(\mathbf{X}\left[n\right]\mathbf{d}^*\left[n\right] - \mathbf{X}\left[n\right]\mathbf{X}^H\left[n\right]\mathbf{h}\left[n\right]\right). \tag{31}$$

Using (8), (11), and (19b) one can verify that

$$\mathbf{X}\left[n\right]\mathbf{d}^*\left[n\right] = \mathbf{X}\left[n\right]\mathbf{X}^H\left[n\right]\mathbf{h}_{\mathrm{opt}} + \tilde{\mathbf{g}}\left[n\right] + \mathbf{X}\left[n\right]\mathbf{v}^*\left[n\right], \tag{32}$$

for $n \geq 0$. Plugging (32) into (31) yields

$$\bar{\mathbf{h}}[n+1] = \mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}\left[n\right] - \mu\left(\tilde{\mathbf{g}}\left[n\right] + \mathbf{X}\left[n\right]\mathbf{v}^*\left[n\right]\right), \tag{33}$$

where $\mathbf{R}_{\mathbf{x}}[n]$ is defined in (19c). The noise signal $v[n]$ is zero-mean and by *AS2* is independent of $\mathbf{x}[n]$. Therefore, applying the stochastic expectation to (33) yields

$$\begin{aligned}\mathbb{E}\left\{\bar{\mathbf{h}}[n]\right\} &= \mathbb{E}\left\{\mathbf{R}_{\mathbf{x}}[n-1]\bar{\mathbf{h}}[n-1]\right\} - \mu\mathbb{E}\left\{\tilde{\mathbf{g}}[n-1]\right\} \\ &\overset{(a)}{=} \tilde{\mathbf{R}}_{\mathbf{x}}\mathbb{E}\left\{\bar{\mathbf{h}}[n-1]\right\}, \end{aligned} \tag{34}$$

where (a) follows from the independence assumption *AS4*, the definition of $\tilde{\mathbf{R}}_{\mathbf{x}}$ in (19e), and from property *P1* in Lemma 2. Repeating the recursion in (34) n times yields (20). □

C. Proof of Proposition 1

Let $\mathbf{Q}$ be the $M \times M$ Hermitian matrix obtained via $\mathbf{Q} \triangleq \mathrm{vec}^{-1}(\mathbf{q})$. Note that

$$\begin{aligned}&\left\|\bar{\mathbf{h}}[n+1]\right\|_{\mathbf{q}}^2 \\ &\triangleq \bar{\mathbf{h}}^H[n+1]\,\mathbf{Q}\,\bar{\mathbf{h}}[n+1] \\ &\overset{(a)}{=} \bar{\mathbf{h}}^H[n]\,\mathbf{R}_{\mathbf{x}}[n]\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n] + \mu^2\cdot\left\|\tilde{\mathbf{g}}[n] + \mathbf{X}[n]\,\mathbf{v}^*[n]\right\|_{\mathbf{q}}^2 \\ &\quad - 2\mu\cdot\mathrm{Re}\left\{\left(\mathbf{v}^T[n]\,\mathbf{X}^H[n] + \tilde{\mathbf{g}}^H[n]\right)\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\}, \end{aligned} \tag{35}$$

where (a) follows from (33). The noise signal $v[n]$ is zero-mean and by *AS2* is independent of $\mathbf{X}[n]$. Therefore, applying the stochastic expectation to (35) results in

$$\begin{aligned}&\mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n+1]\right\|_{\mathbf{q}}^2\right\} \\ &\quad= \mathbb{E}\left\{\bar{\mathbf{h}}^H[n]\,\mathbf{R}_{\mathbf{x}}[n]\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\} \\ &\qquad + \mu^2\cdot\mathbb{E}\left\{\left\|\tilde{\mathbf{g}}[n]\right\|_{\mathbf{q}}^2\right\} + \mu^2\mathbb{E}\left\{\left\|\mathbf{X}[n]\,\mathbf{v}^*[n]\right\|_{\mathbf{q}}^2\right\} \\ &\qquad - 2\mu\cdot\mathrm{Re}\left\{\mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\}\right\} \\ &\qquad - 2\mu\cdot\mathrm{Re}\left\{\mathbb{E}\left\{\tilde{\mathbf{g}}^H[n]\,\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\}\right\}. \end{aligned} \tag{36}$$

Next, we explicitly compute each of the terms in the right hand side (RHS) of (36). First, using the definition of $\mathbf{F}$ in (19m), then from (28) it follows that $\mathrm{vec}\left(\mathbb{E}\left\{\mathbf{R}_{\mathbf{x}}[n]\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\right\}\right) = \mathbf{F}\mathbf{q}$. Thus,

$$\begin{aligned}&\mathbb{E}\left\{\bar{\mathbf{h}}^H[n]\,\mathbf{R}_{\mathbf{x}}[n]\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\} \\ &\qquad\overset{(a)}{=} \mathbb{E}\left\{\bar{\mathbf{h}}^H[n]\,\mathbb{E}\left\{\mathbf{R}_{\mathbf{x}}[n]\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\right\}\bar{\mathbf{h}}[n]\right\} \\ &\qquad\overset{(b)}{=} \mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n]\right\|_{\mathbf{F}\mathbf{q}}^2\right\}, \end{aligned} \tag{37}$$

where (a) follows from assumption *AS4*, and (b) follows from the definition of the weighted Euclidean norm in Subsection II-A. Next, note that by property *R2* (29)

$$\mathbb{E}\left\{\left\|\tilde{\mathbf{g}}[n]\right\|_{\mathbf{q}}^2\right\} = \left(\mathrm{vec}\left(\mathbb{E}\left\{\tilde{\mathbf{g}}[n]\,\tilde{\mathbf{g}}^H[n]\right\}\right)\right)^H\mathbf{q}. \tag{38}$$

Similarly,

$$\begin{aligned}\mathbb{E}\left\{\left\|\mathbf{X}[n]\,\mathbf{v}^*[n]\right\|_{\mathbf{q}}^2\right\} &\overset{(a)}{=} \mathbf{c}_{\mathbf{v}}^T[n,0]\,\mathbb{E}\left\{\mathbf{X}^T[n]\otimes\mathbf{X}^H[n]\right\}\mathbf{q} \\ &\overset{(b)}{=} \left(\mathbb{E}\left\{\mathbf{X}^*[n]\otimes\mathbf{X}[n]\right\}\mathbf{c}_{\mathbf{v}}^*[n,0]\right)^H\mathbf{q}, \end{aligned} \tag{39}$$

where (a) follows as $\mathbf{v}[n]$ is independent of $\mathbf{X}[n]$ by *AS2*, from properties *R1-R2* (28)-(29), and from the definition of $\mathbf{c}_{\mathbf{v}}[n,l]$ in (19g), and (b) follows since conjugate transposition is distributive over the Kronecker product [33, Ch. 9.2]. Combining (38) and (39), and using (19h), yields

$$\mathbb{E}\left\{\left\|\tilde{\mathbf{g}}[n]\right\|_{\mathbf{q}}^2\right\} + \mathbb{E}\left\{\left\|\mathbf{X}[n]\,\mathbf{v}^*[n]\right\|_{\mathbf{q}}^2\right\} = \mathbf{p}_{\mathbf{x}\mathbf{v}}^H[n]\mathbf{q}. \tag{40}$$

To obtain the fourth term in (36), note that repeating the recursion in (33) n times yields $\bar{\mathbf{h}}[n] = \mathbf{L}_0^{n-1}\bar{\mathbf{h}}[0] - \mu\sum_{k=1}^{n}\mathbf{L}_{n-k+1}^{n-1}\left(\mathbf{X}[n-k]\,\mathbf{v}^*[n-k] + \tilde{\mathbf{g}}[n-k]\right)$, where $\mathbf{L}_k^n$ is defined in (19d). As (19d) implies that $\mathbf{R}_{\mathbf{x}}[n]\mathbf{L}_k^{n-1} = \mathbf{L}_k^n$, $\forall k \le n$, it follows that

$$\begin{aligned}&\mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\} \\ &= \mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{L}_0^n\bar{\mathbf{h}}[0]\right\} \\ &\quad - \mu\sum_{k=1}^{n}\mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{L}_{n-k+1}^n\mathbf{X}[n-k]\,\mathbf{v}^*[n-k]\right\} \\ &\quad - \mu\sum_{k=1}^{n}\mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{L}_{n-k+1}^n\tilde{\mathbf{g}}[n-k]\right\}. \end{aligned} \tag{41}$$

The first and the last elements in the RHS of (41) are zero since $\mathbf{v}[n]$ is zero-mean and independent of $\{\mathbf{X}[l]\}_{l=0}^n$ by *AS2*. As for the second element, we note that

$$\begin{aligned}&\mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{L}_{n-k+1}^n\mathbf{X}[n-k]\,\mathbf{v}^*[n-k]\right\} \\ &\qquad\overset{(a)}{=} \mathbf{c}_{\mathbf{v}}^T[n,k]\,\mathbb{E}\left\{\left(\mathbf{L}_{n-k+1}^n\mathbf{X}[n-k]\right)^T\otimes\mathbf{X}^H[n]\right\}\mathbf{q}, \end{aligned} \tag{42}$$

where (a) follows since $\mathbf{v}[n]$ and $\mathbf{X}[n]$ are mutually independent processes by *AS2*, and from properties *R1-R2* (28)-(29). Plugging (42) into (41) and using (19i) yields

$$\mathbb{E}\left\{\mathbf{v}^T[n]\,\mathbf{X}^H[n]\,\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\} = -\mu\cdot\mathbf{z}_{\mathbf{x}\mathbf{v}}^H[n]\mathbf{q}. \tag{43}$$

Lastly, we write

$$\begin{aligned}&\mathbb{E}\left\{\tilde{\mathbf{g}}^H[n]\,\mathbf{Q}\mathbf{R}_{\mathbf{x}}[n]\bar{\mathbf{h}}[n]\right\} \\ &\qquad\overset{(a)}{=} \left(\mathbb{E}\left\{\left(\mathbf{R}_{\mathbf{x}}[n]\,\mathbb{E}\left\{\bar{\mathbf{h}}[n]\right\}\right)\otimes\tilde{\mathbf{g}}^*[n]\right\}\right)^T\mathbf{q} \\ &\qquad\overset{(b)}{=} \left(\mathbb{E}\left\{\mathbf{R}_{\mathbf{x}}[n]\otimes\tilde{\mathbf{g}}^*[n]\right\}\mathbb{E}\left\{\bar{\mathbf{h}}[n]\right\}\right)^T\mathbf{q} \overset{(c)}{=} \mathbf{z}_{\mathbf{x}\mathbf{g}}^H[n]\,\mathbf{q}, \end{aligned} \tag{44}$$

where (a) follows from *AS4* and from (28); (b) follows from the mixed-product property of the Kronecker product [33, Ch. 9.2] by writing $\tilde{\mathbf{g}}^*[n] = \tilde{\mathbf{g}}^*[n]\cdot 1$, and (c) follows from Lemma 3 and from definitions (19e) and (19k). Plugging (37), (40), (43) and (44) into (36) yields (21). □

D. Proof of Theorem 1

Plugging the relationship (11) into the definitions of $e[n]$ and $e_{\mathrm{opt}}[n]$ in (16) results in $e[n] = \bar{\mathbf{h}}^H[n]\,\mathbf{x}[n] + \mathbf{g}^H[n]\,\mathbf{x}[n] + v[n]$, and $e_{\mathrm{opt}}[n] = \mathbf{g}^H[n]\,\mathbf{x}[n] + v[n]$. Therefore,

$$\begin{aligned}\mathbb{E}\left\{|e[n]|^2\right\} - \mathbb{E}\left\{\left|e_{\mathrm{opt}}[n]\right|^2\right\} &\overset{(a)}{=} \mathbb{E}\left\{\left|\left(\bar{\mathbf{h}}[n] + \mathbf{g}[n]\right)^H\mathbf{x}[n]\right|^2\right\} \\ &\quad - \mathbb{E}\left\{\left|\mathbf{g}^H[n]\,\mathbf{x}[n]\right|^2\right\}, \end{aligned} \tag{45}$$

where (a) follows since by *AS2*, $\mathbf{x}[n]$ is zero-mean and independent of $\upsilon[n]$, and by *AS3*, $\mathbf{x}[n]$ is also independent of $\bar{\mathbf{h}}[n]$. thus

$$\xi[n] = \left\langle \mathbb{E}\left\{ \left| \bar{\mathbf{h}}^H[n]\,\mathbf{x}[n] \right|^2 \right\} \right\rangle_{N_0} + 2\mathrm{Re}\left\{ \left\langle \mathbb{E}\left\{ \bar{\mathbf{h}}^H[n]\,\mathbf{x}[n]\,\mathbf{x}^H[n]\,\mathbf{g}[n] \right\} \right\rangle_{N_0} \right\}. \quad (46)$$

Note that the first summand in the RHS of (46) satisfies

$$\mathbb{E}\left\{ \left| \bar{\mathbf{h}}^H[n]\,\mathbf{x}[n] \right|^2 \right\} \overset{(a)}{=} \mathbb{E}\left\{ \bar{\mathbf{h}}^H[n]\,\mathbb{E}\left\{ \mathbf{x}[n]\,\mathbf{x}^H[n] \right\} \bar{\mathbf{h}}[n] \right\} \overset{(b)}{=} \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[n] \right\|^2_{\mathbf{c}_{\mathbf{x}}[n]} \right\}, \quad (47)$$

where (a) follows from the independence assumption *AS3*, and (b) follows from definitions (3a) and (3b). Next, we note that it follows from the independence assumption *AS3*, from Lemma 3, and from the definition of $\mathbf{C}_{\mathbf{x}}[n]$ (3a) that

$$\mathbb{E}\left\{ \bar{\mathbf{h}}^H[n]\,\mathbf{x}[n]\,\mathbf{x}^H[n]\,\mathbf{g}[n] \right\} = \mathbb{E}\left\{ \bar{\mathbf{h}}^H[0] \right\} \left(\tilde{\mathbf{R}}_{\mathbf{x}} \right)^n \mathbf{C}_{\mathbf{x}}[n]\,\mathbf{g}[n]. \quad (48)$$

Plugging (47) and (48) into (46) for $n \geq N_0 - 1$ results in

$$\xi[n] \overset{(a)}{=} \frac{1}{N_0} \sum_{k=n-N_0+1}^{n} \left(\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[k] \right\|^2_{\mathbf{c}_{\mathbf{x}}\left[((k))_{N_0}\right]} \right\} + 2\mathrm{Re}\left\{ \mathbb{E}\left\{ \bar{\mathbf{h}}^H[0] \right\} \times \left(\tilde{\mathbf{R}}_{\mathbf{x}} \right)^k \mathbf{C}_{\mathbf{x}}\left[((k))_{N_0}\right] \mathbf{g}\left[((k))_{N_0}\right] \right\} \right), \quad (49)$$

where (a) follows since $\mathbf{c}_{\mathbf{x}}[n]$, $\mathbf{C}_{\mathbf{x}}[n]$, and $\mathbf{g}[n]$ are all periodic with period N_0. □

E. Proof of Corollary 2

We first note that $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ defined in (3c) is obtained as the average of N_0 Hermitian and positive semi-definite covariance matrices. Thus, $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ is also Hermitian positive semi-definite, and its eigenvalues $\{\lambda_k\}_{k=1}^M$ are all real-valued and non-negative. From *AS1* it follows that the eigenvalues of $\underset{\sim}{\mathbf{C}}_{\mathbf{x}}$ are strictly positive. From (19e) and (20) it follows that $\lim_{n\to\infty} \mathbb{E}\left\{ \bar{\mathbf{h}}[n] \right\} = \mathbf{0}_{M\times 1}$ for every $\mathbf{h}[0]$ if and only if $\lim_{n\to\infty} \left(\tilde{\mathbf{R}}_{\mathbf{x}} \right)^n = \mathbf{0}_{M\times M}$. Since $\tilde{\mathbf{R}}_{\mathbf{x}} = \mathbf{I}_M - \mu \underset{\sim}{\mathbf{C}}_{\mathbf{x}}$, its eigenvalues are $\{1 - \mu\cdot\lambda_k\}_{k=1}^M$, where $\lambda_k > 0$, $\forall k \in \{1, 2, \ldots, M\}$, we conclude that the TA-LMS algorithm is mean convergent if and only if $|1 - \mu\cdot\lambda_k| < 1$ for all $k \in \{1, 2, \ldots, M\}$ [34, Ch. 7.10], which occurs if and only if $0 < \mu < \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}$. □

F. Proof of Theorem 2

The following proof requires analysis of *linear non-homogeneous state-space equations* to show that (25) guarantees convergence, while the proof of the stationary counterpart in [3, Th. 24.1] required only the analysis of homogeneous state-space equations. Define $\mathbf{z}^t_{\mathbf{xv}}[n] \triangleq \mathbf{z}_{\mathbf{xv}}[n] - \mathbf{z}^s_{\mathbf{xv}}[n]$, where $\mathbf{z}_{\mathbf{xv}}[n]$ and $\mathbf{z}^s_{\mathbf{xv}}[n]$ are defined in (19i) and (19j), respectively. For $n \geq N_0$, repeating the recursion (21) N_0 times we obtain

$$\begin{aligned} &\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[n] \right\|^2_{\mathbf{q}} \right\} \\ &= \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[n - N_0] \right\|^2_{\mathbf{F}^{N_0}\mathbf{q}} \right\} \\ &\quad + \mu^2 \cdot \mathrm{Re}\left\{ \sum_{l=1}^{N_0} \left(\mathbf{p}^H_{\mathbf{xv}}[n-l] + 2\left(\mathbf{z}^s_{\mathbf{xv}}[n-l] \right)^H \right) \mathbf{F}^{l-1}\mathbf{q} \right\} \\ &\quad - 2\mu \cdot \mathrm{Re}\left\{ \sum_{l=1}^{N_0} \left(\mathbf{z}^H_{\mathbf{xg}}[n-l] - \mu\left(\mathbf{z}^t_{\mathbf{xv}}[n-l] \right)^H \right) \mathbf{F}^{l-1}\mathbf{q} \right\}, \end{aligned} \quad (50)$$

where $\mathbf{p}_{\mathbf{xv}}[n]$, $\mathbf{F}$, and $\mathbf{z}_{\mathbf{xg}}[n]$ are defined in (19h), (19m), and (19k), respectively. Define

$$\bar{\mathbf{h}}_k[n] \triangleq \bar{\mathbf{h}}[n\cdot N_0 + k], \quad k \in \mathcal{N}_0. \quad (51a)$$

We show that if (25) is satisfied, then $\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n] \right\|^2 \right\}$ converges $\forall k \in \mathcal{N}_0$. To that aim, define

$$\bar{\mathbf{F}} \triangleq \mathbf{F}^{N_0}, \quad (51b)$$

$$\mathbf{a}_k \triangleq \left(\sum_{l=1}^{N_0} \left(\mathbf{p}_{\mathbf{xv}}\left[((k-l))_{N_0}\right] + 2\mathbf{z}^s_{\mathbf{xv}}\left[((k-l))_{N_0}\right] \right)^H \mathbf{F}^{l-1} \right)^H, \quad (51c)$$

and

$$\mathbf{b}_k[n] \triangleq \left(\sum_{l=1}^{N_0} \left(\mathbf{z}^H_{\mathbf{xg}}[(n+1)N_0 + k - l] - \mu\left(\mathbf{z}^t_{\mathbf{xv}}[(n+1)N_0 + k - l] \right)^H \right) \mathbf{F}^{l-1} \right)^H. \quad (51d)$$

From *AS5-AS6* it follows that $\mathbf{a}_k$ and $\mathbf{b}_k[n]$ are bounded $\forall k \in \mathcal{N}_0$. From *P4* in Lemma 2 it follows that $\mathbf{p}_{\mathbf{xv}}[n]$ and $\mathbf{z}^s_{\mathbf{xv}}[n]$ are periodic with period N_0, hence, (50) can be written as

$$\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n+1] \right\|^2_{\mathbf{q}} \right\} = \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n] \right\|^2_{\bar{\mathbf{F}}\mathbf{q}} \right\} + \mu^2 \cdot \mathrm{Re}\left\{ \mathbf{a}^H_k \mathbf{q} \right\} - 2\mu \cdot \mathrm{Re}\left\{ \mathbf{b}^H_k[n]\,\mathbf{q} \right\}. \quad (52)$$

The MSD is obtained from $\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}[n] \right\|^2_{\mathbf{q}} \right\}$ by setting $\mathbf{q} = \mathrm{vec}^{-1}(\mathbf{I}_M)$. Following [3, Ch. 24.2], we use (52) to formulate M^2 state-space recursions for each $k \in \mathcal{N}_0$ as follows:

$$\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n+1] \right\|^2_{\bar{\mathbf{F}}^l\mathbf{q}} \right\} = \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n] \right\|^2_{\bar{\mathbf{F}}^{l+1}\mathbf{q}} \right\} + \mu^2 \cdot \mathrm{Re}\left\{ \mathbf{a}^H_k \bar{\mathbf{F}}^l \mathbf{q} \right\} - 2\mu \cdot \mathrm{Re}\left\{ \mathbf{b}^H_k[n]\,\bar{\mathbf{F}}^l\mathbf{q} \right\}, \quad (53)$$

$l \in \{0, 1, \ldots, M^2 - 1\}$. Let $\{\alpha_l\}_{l=0}^{M^2-1}$ be the coefficients of the characteristic polynomial of $\bar{\mathbf{F}}$ [34, p. 492]. It follows from the Cayley-Hamilton theorem [34, p. 532] and from the linearity of the weighted Euclidean norm [3, eq. (23.31)] that $\mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n] \right\|^2_{\bar{\mathbf{F}}^{M^2}\mathbf{q}} \right\} = -\sum_{l=0}^{M^2-1} \alpha_l \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n] \right\|^2_{\bar{\mathbf{F}}^l\mathbf{q}} \right\}$. Therefore, by defining the $M^2 \times 1$ vectors $\underset{\sim}{\bar{\mathbf{h}}}_k[n]$, $\underset{\sim}{\mathbf{a}}_k$, and $\underset{\sim}{\mathbf{b}}_k[n]$, s.t. $\left(\underset{\sim}{\bar{\mathbf{h}}}_k[n]\right)_l \triangleq \mathbb{E}\left\{ \left\| \bar{\mathbf{h}}_k[n] \right\|^2_{\bar{\mathbf{F}}^l\mathbf{q}} \right\}$, $(\underset{\sim}{\mathbf{a}}_k)_l \triangleq \mathbf{a}^H_k \bar{\mathbf{F}}^l\mathbf{q}$, and $(\underset{\sim}{\mathbf{b}}_k[n])_l \triangleq \mathbf{b}^H_k[n]\,\bar{\mathbf{F}}^l\mathbf{q}$, $l \in \{0, 1, \ldots, M^2 - 1\}$, and the $M^2 \times M^2$ matrix $\underset{\sim}{\bar{\mathbf{F}}}$

such that $\underset{\sim}{\bar{\mathbf{F}}}^T$ is the companion matrix of the characteristic polynomial of $\bar{\mathbf{F}}$ [34, p. 648], the state-space recursions (53) can be written as a set of N_0 multivariate difference equations

$$\underset{\sim}{\bar{\mathbf{h}}}_k[n+1] = \underset{\sim}{\bar{\mathbf{F}}}\underset{\sim}{\bar{\mathbf{h}}}_k[n] + \mu^2 \cdot \mathrm{Re}\{\underset{\sim}{\mathbf{a}}_k\} - 2\mu \cdot \mathrm{Re}\{\underset{\sim}{\mathbf{b}}_k[n]\}, \quad (54)$$

$k \in \mathcal{N}_0$, $n \geq 0$. Note that $\forall k \in \mathcal{N}_0$, (54) represents an $M^2 \times M^2$ multivariate LTI system with input signal $\mu^2\mathrm{Re}\{\underset{\sim}{\mathbf{a}}_k\} - 2\mu\mathrm{Re}\{\underset{\sim}{\mathbf{b}}_k[n]\}$ and output signal $\underset{\sim}{\bar{\mathbf{h}}}_k[n]$. Since from property 2 in Lemma 2 it follows that $\mathbf{z}^t_{\mathbf{xv}}[n] = \mathbf{z}_{\mathbf{xv}}[n] - \mathbf{z}^s_{\mathbf{xv}}[n] = \mathbf{0}_{M^2\times 1}$ for all $n > \eta$, where η is a fixed positive integer defined in (12), we conclude from (51d) that for $0 < \mu < \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}$

$$\lim_{n\to\infty} \mathbf{b}_k^H[n] = \lim_{n\to\infty} \sum_{l=1}^{N_0} \mathbf{z}^H_{\mathbf{xg}}[(n+1)N_0 + k - l]\,\mathbf{F}^{l-1} \overset{(a)}{=} \mathbf{0}_{M^2\times 1}, \quad (55)$$

where (a) follows since by (19k), $\lim_{n\to\infty} \mathbf{z}_{\mathbf{xg}}[n] = \mathbf{0}_{M^2\times 1}$ for $0 < \mu < \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}$. It therefore follows from [3, Ch. 23.4] that if $0 < \mu < \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}$ and the eigenvalues of $\underset{\sim}{\bar{\mathbf{F}}}$ are inside the unit circle then $\underset{\sim}{\bar{\mathbf{h}}}_k[n]$ is bounded and tends to a steady-state value for $n \to \infty$, thus $\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2\right\}$ is convergent.

So far we have shown that if $0 < \mu < \frac{2}{\lambda_{\max}(\underset{\sim}{\mathbf{C}}_{\mathbf{x}})}$ and all the eigenvalues of $\underset{\sim}{\bar{\mathbf{F}}}$ are inside the unit circle then $\bar{\mathbf{h}}[n]$ is mean-square stable. We now show that the latter condition is equivalent to the constraints on $\mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{H}$ in (25). Note that it follows from [3, Pg. 346] that the eigenvalues of $\underset{\sim}{\bar{\mathbf{F}}}$ are the eigenvalues of $\bar{\mathbf{F}}$. Also note that from (51b) it follows that all the eigenvalues of $\bar{\mathbf{F}}$ are inside the unit circle if and only if all the eigenvalues of $\mathbf{F}$ are inside the unit circle. Recall that $\mathbf{F} = \mathbf{I}_{M^2} - \mu\mathbf{A} + \mu^2\mathbf{B}$, where $\mu > 0$, $\mathbf{A}$, defined in (19l), is positive-definite, and $\mathbf{B}$, defined in (13), is positive semi-definite and finite (by *AS6*). It therefore follows from [30, Appendix A] that when $\mathbf{H}$ has at least one real-valued positive eigenvalue then the eigenvalues of $\mathbf{F}$ are guaranteed to be inside the unit circle if $\mu < \min\left\{\frac{1}{\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B})}, \frac{1}{\lambda_{\max}(\mathbf{H})}\right\}$, and when $\mathbf{H}$ has no real-valued positive eigenvalues, the eigenvalues of $\mathbf{F}$ are guaranteed to be inside the unit circle if $\mu < \frac{1}{\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B})}$. Combining this with (24) yields (25). □

G. Proof of Lemma 4

Note that in steady-state

$$\lim_{n\to\infty} \left\langle \mathbb{E}\left\{\bar{\mathbf{h}}^H[n]\,\mathbf{x}[n]\,\mathbf{x}^H[n]\,\mathbf{g}[n]\right\}\right\rangle_{N_0} \overset{(a)}{=} \lim_{n\to\infty} \left\langle \mathbb{E}\left\{\bar{\mathbf{h}}^H[n]\right\}\mathbb{E}\left\{\mathbf{x}[n]\,\mathbf{x}^H[n]\right\}\mathbf{g}[n]\right\rangle_{N_0} \overset{(b)}{=} 0,$$

where (a) follows from *AS3*, and (b) follows from Def. 1 and Def. 3, as the filter is in steady-state, and as $\mathbb{E}\{\mathbf{x}[n]\,\mathbf{x}^H[n]\}$ is finite $\forall n \in \mathbb{Z}$. The steady-state excess TA-MSE, obtained by taking $n \to \infty$ in (46), is therefore given by $\xi_s = \lim_{n\to\infty}\left\langle \mathbb{E}\left\{\left|\bar{\mathbf{h}}^H[n]\,\mathbf{x}[n]\right|^2\right\}\right\rangle_{N_0}$. Plugging (47) into the expression for ξ_s yields (26). □

H. Proof of Corollary 3

Assuming that μ satisfies (25), by Lemma 4 the steady-state excess TA-MSE is given by

$$\xi_s = \lim_{n\to\infty}\left\langle \mathbb{E}\left\{\left\|\bar{\mathbf{h}}[n]\right\|^2_{\mathbf{c}_{\mathbf{x}}[n]}\right\}\right\rangle_{N_0} \overset{(a)}{=} \frac{1}{N_0}\sum_{k=0}^{N_0-1}\lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{\mathbf{c}_{\mathbf{x}}[k]}\right\}, \quad (56)$$

where (a) follows from the definition of $\bar{\mathbf{h}}_k[n]$ in (51a), the periodicity of $\mathbf{c}_{\mathbf{x}}[n]$, i.e., $\mathbf{c}_{\mathbf{x}}[n] = \mathbf{c}_{\mathbf{x}}\left[((n))_{N_0}\right]$, and from the fact that when the limit exists then any subsequence converges to the limit [35, Definition 3.5]. Next, recalling the definitions of $\mathbf{a}_k$, $\mathbf{b}_k[n]$, and $\bar{\mathbf{F}}$ stated in (51), it follows from (52) and (55) that $\forall k \in \mathcal{N}_0$ in the steady-state it holds that

$$\lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{\mathbf{q}}\right\} = \lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{\bar{\mathbf{F}}\mathbf{q}}\right\} + \mu^2\cdot\mathrm{Re}\left\{\mathbf{a}_k^H\mathbf{q}\right\}.$$

Thus, since (25) guarantees that $\lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{\mathbf{q}}\right\}$ and $\lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{\bar{\mathbf{F}}\mathbf{q}}\right\}$ both exist and are finite, then from the linearity of the weighted Euclidean norm [3, eq. (23.31)] we have $\lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{(\mathbf{I}_{M^2}-\bar{\mathbf{F}})\mathbf{q}}\right\} = \mu^2\cdot\mathrm{Re}\left\{\mathbf{a}_k^H\mathbf{q}\right\}$, $\forall k \in \mathcal{N}_0$. Setting $\mathbf{q} = \left(\mathbf{I}_{M^2} - \bar{\mathbf{F}}\right)^{-1}\mathbf{c}_{\mathbf{x}}[k]$ results in

$$\lim_{n\to\infty}\mathbb{E}\left\{\left\|\bar{\mathbf{h}}_k[n]\right\|^2_{\mathbf{c}_{\mathbf{x}}[k]}\right\} = \mu^2\cdot\mathrm{Re}\left\{\mathbf{a}_k^H\left(\mathbf{I}_{M^2} - \bar{\mathbf{F}}\right)^{-1}\mathbf{c}_{\mathbf{x}}[k]\right\}. \quad (57)$$

Plugging (57) into (56) results in

$$\xi_s = \mu^2\frac{1}{N_0}\sum_{k=0}^{N_0-1}\mathrm{Re}\left\{\mathbf{a}_k^H\left(\mathbf{I}_{M^2} - \bar{\mathbf{F}}\right)^{-1}\mathbf{c}_{\mathbf{x}}[k]\right\} = \mu^2\mathrm{Re}\left\{\frac{1}{N_0}\sum_{k=0}^{N_0-1}\mathrm{Tr}\left\{\mathbf{c}_{\mathbf{x}}[k]\mathbf{a}_k^H\left(\mathbf{I}_{M^2} - \bar{\mathbf{F}}\right)^{-1}\right\}\right\} \overset{(a)}{=} \mu^2\mathrm{Re}\left\{\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{I}_{M^2} - \mathbf{F}^{N_0}\right)^{-1}\right\}\right\}, \quad (58)$$

where (a) follows from the linearity of the trace operator [33, Ch, 1.1] and from the definitions of $\mathbf{a}_k$, $\bar{\mathbf{F}}$, and $\mathbf{P}$ in (51c), (51b), and (19o), respectively. □

REFERENCES

[1] N. Shlezinger, K. Todros, and R. Dabora, "Adaptive LMS-type filter for cyclostationary signals," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Poznań, Poland, Sep. 2016, pp. 37–41.

[2] S. O. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2003.

[3] A. H. Sayed, *Adaptive Filters*. Piscataway, NJ, USA: IEEE Press, 2008.

[4] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, pp. 639–697, Apr. 2006.

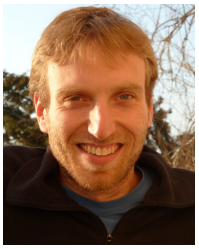
[5] W. A. Gardner, Ed., *Cyclostationarity in Communications and Signal Processing*. Piscataway, NJ, USA: IEEE Press, 1994.

[6] R. W. Heath, Jr., and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.

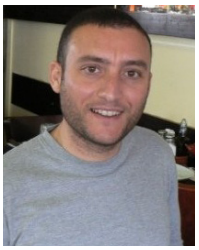
[7] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2014.

[8] W. Nam, D. Bai, J. Lee, and I. Kang, "Advanced interference management for 5G cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 52–60, May 2014.

[9] H. C. Ferreira, L. Lampe, J. Newbury, and T. G. Swart, Eds., *Power Line Communications—Theory and Applications for Narrowband and Broadband Communications Over Power Lines*. Hoboken, NJ, USA: Wiley, 2010.
[10] M. Nassar, J. Lin, Y. Mortazavi, A. Dabak, I. H. Kim, and B. L. Evans, "Local utility power line communications in the 3–500 kHz band: Channel impairments, noise, and standards," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 116–127, Aug. 2012.
[11] *P1901.2/D0.09.00 Draft Standard for Low Frequency (Less Than 500 kHz) Narrow Band Power Line Communications for Smart Grid Applications*, IEEE Standards Association, Piscataway, NJ, USA, Jun. 2013.
[12] D. C. McLernon, "Analysis of LMS algorithm with inputs from cyclostationary random processes," *Electron. Lett.*, vol. 27, no. 2, pp. 136–138, Jan. 1991.
[13] N. J. Bershad, E. Eweda, and J. C. M. Bermudez, "Stochastic analysis of the LMS and NLMS algorithms for cyclostationary white Gaussian inputs," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2238–2249, May 2014.
[14] N. J. Bershad, E. Eweda, and J. C. M. Bermudez, "Stochastic analysis of an adaptive line enhancer/canceler with a cyclostationary input," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 104–119, Jan. 2016.
[15] W. A. Gardner, "Cyclic Wiener filtering: Theory and method," *IEEE Trans. Commun.*, vol. 41, no. 1, pp. 151–163, Jan. 1993.
[16] D. McLernon, "One-dimensional linear periodically time-varying structures: Derivations, interrelationships and properties," *IEEE Proc.-Vis., Image Signal Process.*, vol. 146, no. 5, pp. 245–252, Oct. 1999.
[17] J. Zhang, K. M. Wong, Z. Q. Luo, and P. C. Ching, "Blind adaptive FRESH filtering for signal extraction," *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1397–1402, May 1999.
[18] O. A. Y. Ojeda and J. Grajal, "Adaptive-fresh filters for compensation of cycle-frequency errors," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 1–10, Jan. 2010.
[19] N. Shlezinger and R. Dabora, "Frequency-shift filtering for OFDM signal recovery in narrowband power line communications," *IEEE Trans. Commun.*, vol. 62, no. 4, pp. 1283–1295, Apr. 2014.
[20] J. H. Reed and T. C. Hsia, "The performance of time-dependent adaptive filters for interference rejection," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 8, pp. 1373–1385, Aug. 1990.
[21] G. Gelli, L. Paura, and A. M. Tulino, "Cyclostationarity-based filtering for narrowband interference suppression in direct-sequence spread-spectrum systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1747–1755, Dec. 1998.
[22] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[23] G. B. Giannakis, "Cyclostationary signal analysis," in *The Digital Signal Processing Handbook*. Boca Raton, FL, USA: CRC Press, 1998, pp. 17.1–17.31.
[24] L. L. Scharf and P. J. Schreier, *Statistical Signal Processing of Complex-Valued Data*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[26] R. P. Brent and P. Zimmermann, *Modern Computer Arithmetic*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[27] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationary input sequences," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 3, pp. 630–639, Apr. 1992.
[28] H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 90–102, Jan. 2004.
[29] W. A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique," *Signal Process.*, vol. 6, no. 2, pp. 113–133, Apr. 1984.
[30] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data-normalized adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, Mar. 2003.
[31] A. Dandawate and G. B. Giannakis, "Modeling (almost) periodic moving average processes using cyclic statistics," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 673–684, Mar. 1996.
[32] N. Shlezinger and R. Dabora, "On the capacity of narrowband PLC channels," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1191–1201, Apr. 2015.
[33] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Lyngby, Denmark: Technical Univ. of Denmark, 2008.
[34] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA, USA: SIAM, 2000.
[35] W. Rudin, *Principles of Mathematical Analysis*. New York, NY, USA: McGraw-Hill, 1976.



Nir Shlezinger (S'14) received the B.Sc. and M.Sc. degrees in electrical and computer engineering from Ben-Gurion University of the Negev, Israel, in 2011 and 2013, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering. From 2009 to 2013, he was an Engineer with Yitran Communications. His research interests include information theory and signal processing for communications.



Koby Todros (M'07) was born in Ashkelon, Israel, in 1974. He received the B.Sc. degree, the M.Sc. (summa cum laude) degree, and the Ph.D. (summa cum laude) degree in electrical engineering from the Ben-Gurion University of the Negev, Be'er-Sheva, Israel, in 2000, 2006, and 2011, respectively. He was a Senior Signal Processing Engineer with WideMed Ltd., from 2000 to 2004, where he held the algorithm group manager position, from 2004 to 2006. During 2007, he was a Signal Processing Consultant with EarlySense Ltd. He was a Post-Doctoral Fellow with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA, from 2010 to 2012. Since 2013, he has been a faculty member with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev. His research interests include statistical signal processing and estimation theory with focus on measure-transformed lower-order multivariate analysis, uniformly optimal estimation in the non-Bayesian theory, performance bounds for parameter estimation, blind source separation, and biomedical signal processing. He received the Pharan and Lev-Zion fellowships for excellent doctoral students in 2007 and 2009, respectively.

Ron Dabora (M'07–SM'14) received the B.Sc. and M.Sc. degrees from Tel Aviv University in 1994 and 2000, respectively, and the Ph.D. degree from Cornell University in 2007, all in electrical engineering. From 1994 to 2000, he was an Engineer with the Ministry of Defense of Israel, and from 2000 to 2003, he was with the Algorithms Group, Millimetrix Broadband Networks, Israel. From 2007 to 2009, he was a Post-Doctoral Researcher with the Department of Electrical Engineering, Stanford University. Since 2009, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel. His research interests include network information theory, wireless communications, and power line communications. He served as a TPC Member in a number of international conferences, including the WCNC, the PIMRC, and the ICC. From 2012 to 2014, he served as an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, and he currently serves as a Senior Area Editor for the IEEE SIGNAL PROCESSING LETTERS.