# Channel Tracking with WLMS Algorithms: High Performance at LMS Computational Load

Anders Ahlén*, Lars Lindbom[1] and Mikael Sternad*

*Signals and Systems, Uppsala University, PO Box 528, SE-75120, Uppsala, Sweden.
[1]Ericsson Infotech, PO Box 1038, SE-65115 Karlstad
aa@signal.uu.se, Lars.Lindbom@ein.ericsson.se, ms@signal.uu.se

**Abstract** - Adaptation algorithms with constant gains are designed for tracking time-varying parameters of linear regression models with stationary regressors, in particular channel models in mobile radio communications. In a companion paper, an application to channel tracking in the IS-136 TDMA system is discussed. We propose algorithms that are based on two key concepts: First, the design is transformed into a Wiener filtering problem. Second, the parameters are modeled as correlated ARIMA processes with known dynamics. This leads to a new framework for systematic and optimal design of simple adaptation laws based on a priori information.

The simplest adaptation law, named the Wiener LMS algorithm, is presented here. In the design, all parameters are assumed to be governed by the same dynamics and the covariance matrix of the regressors is assumed to be known. The computational complexity is of the same order of magnitude as that of LMS for white regressors. The tracking performance is however substantially improved.

## I. Introduction and Outline

Consider a possibly multivariable time-varying linear regression

$$y_t = \varphi_t^* h_t + v_t \tag{1}$$

describing, for example, a mobile radio channel where the vector $y_t$ represents the received baseband signal at multiple antennas and $v_t$ is noise. The time-varying regression parameters are represented by the column vector $h_t = [h_{0,t} \dots h_{n_h-1,t}]^T$. In channel modelling, the regressor matrix $\{\varphi_t^*\}$ contains transmitted symbols, with zero means. In such cases, $\{\varphi_t^*\}$ is stationary with a known non-singular autocorrelation matrix $\mathbf{R} = E[\varphi_t \varphi_t^*]$. The aim is to estimate $h_{t+k}$ based on measurements $y_t$, for some $k$.

The time-varying Kalman filter constitutes the MSE-optimal algorithm for estimating time-varying regression parameters, based on linear models [2]. Unfortunately, its complexity may often preclude its use.

Standard LMS and RLS algorithms in general suffer from having an incorrect structure. These algorithms do not utilize a priori knowledge about the statistics of $h_t$ in an appropriate way. Furthermore, for fast time-variations an RLS algorithm with forgetting factor must use a short data window. Consequently, the estimated covariance matrix, or Hessian, will be inaccurate. In the problems considered in this paper, the regressors are assumed stationary. If their statistics is unknown, it is then better to estimate $\mathbf{R}$ separately, using a long data window rather than discarding information as in the RLS case.

We focus on adaptation laws with time-invariant gains, where the dynamics of $h_t$ are taken into account in the design. Such designs, have been studied earlier in e.g. [3]. We then recast the tracking problem into a Wiener filter design, [6, 7, 9], which can be solved in a straightforward way by using e.g. a polynomial approach to Wiener filter design [1].

The present paper outlines the algorithm within the class that has the lowest computational complexity: The Wiener LMS algorithm (WLMS) algorithm. If desired, it can be implemented as "LMS/Newton" updates [13], complemented by an additional filtering of each parameter. This filter provides parameter predictions or, if desired, smoothing estimates. In the design, all elements of $h_t$ are assumed to be governed by the same dynamics. The computational complexity becomes a few times higher than that of LMS for white regressors, while the tracking performance is considerably improved. The traditional LMS-algorithm can be regarded as a special case, obtained by assuming $h_t$ to be a random walk process and $\mathbf{R}$ to be diagonal. In [10] and in a companion paper presented at VTC2000 [11], the application to channel tracking for systems specified by the IS-136 TDMA standard is discussed.

*Remarks on the notation.* For any polynomial $P(q^{-1}) = p_0 + p_1 q^{-1} + \dots + p_{n_P} q^{-n_P}$ in the backward shift operator $q^{-1}$ ($q^{-1}x_t = x_{t-1}$), a conjugate polynomial is defined as $P_*(q) \triangleq p_0^* + p_1^* q + \dots + p_{n_P}^* q^{n_P}$ where $q$ is the forward shift operator ($qx_t = x_{t+1}$) and $p^*$ denotes the complex conjugate of $p$.

## II. The Channel Estimator

A WLMS design begins with the selection of a model, sometimes denoted hypermodel, describing the second order statistics of $h_t$. The selection can either be based on *a priori* knowledge, or be regarded as a user choice, reflecting our belief on the nature of the time-variations. Parameter vectors can be modeled as stochastic processes

$$h_t = \boldsymbol{\mathcal{H}}(q^{-1})e_t \tag{2}$$

where $\boldsymbol{\mathcal{H}}$ is a stable or marginally stable transfer function matrix with time-invariant or slowly timevarying parameters. The noise $e_t$ is a white zero mean random vector sequence with covariance matrix $\mathbf{R}_e = Ee_te_t^*$. This hypermodel should capture the essential behavior of the time variability. In radio channel modelling, $\boldsymbol{\mathcal{H}}$ can be adjusted to the autocorrelation function of the fading model [10, 11]. In this paper, we shall consider marginally stable autoregressive integrating moving average (ARIMA) models of order $n_D$, with equal dynamics for all channel taps

$$h_t = \frac{C(q^{-1})}{D(q^{-1})}\mathbf{I}\, e_t = \frac{1 + c_1q^{-1} + \ldots + c_{n_C}q^{-n_D}}{1 + d_1q^{-1} + \ldots + d_{n_D}q^{-n_D}}\mathbf{I}\, e_t \quad , \tag{3}$$

and with real-valued scalar coefficients $\{c_i \; d_i\}$.

Define the tracking error vector

$$\tilde{h}_{t+k|t} \triangleq h_{t+k} - \hat{h}_{t+k|t} \tag{4}$$

where $\hat{h}_{t+k|t}$ is an estimate of $h_{t+k}$ at time $t$ representing filtering ($k = 0$), prediction ($k > 0$) or fixed lag smoothing ($k < 0$). The tracking performance will be measured by

$$\lim_{t\to\infty} \mathrm{tr}\left(E\tilde{h}_{t+k|t}\tilde{h}^*_{t+k|t}\right) \tag{5}$$

where the expectation is taken with respect to $e_t$ in (2) and $v_t$ in (1) after the initial transients ($t \to \infty$).

The considered class of adaptation algorithms has the structure

$$\begin{aligned}
\varepsilon_t &= y_t - \varphi_t^*\hat{h}_{t|t-1} && (6)\\
\hat{h}_{t|t} &= \hat{h}_{t|t-1} + \mu\mathbf{R}^{-1}\varphi_t\varepsilon_t && (7)\\
\hat{h}_{t+k|t} &= \boldsymbol{\mathcal{P}}_k(q^{-1})\hat{h}_{t|t} \quad , && (8)
\end{aligned}$$

where $\mu$ is a scalar gain and $\boldsymbol{\mathcal{P}}_k(q^{-1})$ is a causal and stable rational matrix, which introduces an appropriate amount of coupling in the estimates and provides prediction or smoothing estimates for any horizon $k$. The appropriate tuning of $\boldsymbol{\mathcal{P}}_k(q^{-1})$ will depend on the dynamics of $h_t$ and on the SNR. We shall refer to $\boldsymbol{\mathcal{P}}_k(q^{-1})$ as the *coefficient smoothing-prediction filter*. Note that the one-step prediction estimate $\hat{h}_{t+1|t} = \boldsymbol{\mathcal{P}}_1(q^{-1})\hat{h}_{t|t}$ must always be made available, since $\hat{h}_{t|t-1}$ is required in (6) and (7). We constrain $\boldsymbol{\mathcal{P}}_k(q^{-1})$ to a *diagonal rational matrix.* With a diagonal $\boldsymbol{\mathcal{P}}_k(q^{-1})$, the required number of computations grows only linearly with $n_h$ when $\mathbf{R}$ is diagonal. With $\boldsymbol{\mathcal{P}}_1(q^{-1}) = \mathbf{I}$, (6)-(8) reduces to LMS/Newton algorithm [13]. If furthermore $\mathbf{R}$ is diagonal, it reduces to LMS.

# III. Wiener LMS Design

The optimization of the adaptation law (6)-(8) will now be solved by an open loop linear Wiener design, which gives several advantages: It provides a systematic design technique, a numerically safe implementation and an opportunity for using tools and design intuition from Wiener filtering.

### III.A An Open-Loop Formulation

Consider the signal prediction error (6) and insert (1) describing $y_t$, to obtain

$$\begin{aligned}
\varepsilon(t) &= \varphi_t^*(h_t - \hat{h}_{t|t-1}) + v_t \\
\varphi_t\varepsilon_t &= \varphi_t\varphi_t^*\tilde{h}_{t|t-1} + \varphi_t v_t \quad . && (9)
\end{aligned}$$

By adding and subtracting $\mathbf{R}\tilde{h}_{t|t-1}$ and defining

$$\begin{aligned}
Z_t &= \varphi_t\varphi_t^* - \mathbf{R} && (10)\\
\eta_t &= Z_t\tilde{h}_{t|t-1} + \varphi_t v_t && (11)\\
f_t &= \mathbf{R}h_t + \eta_t \quad , && (12)
\end{aligned}$$

the vector (9) is now reformulated as

$$\begin{aligned}
\varphi_t\varepsilon_t &= \mathbf{R}\tilde{h}_{t|t-1} + Z_t\tilde{h}_{t|t-1} + \varphi_t v_t \\
&= f_t - \mathbf{R}\hat{h}_{t|t-1} \quad . && (13)
\end{aligned}$$

Here, $f_t$ can be regarded as a fictitious measurement, with $\mathbf{R}h_t$ and $\eta_t$ in (12) being the signal and the noise, respectively. It can be constructed from known signals via (13). In the sequel, the noise terms $\eta_t$ and $Z_t\tilde{h}_{t|t-1}$ will be referred to as the *gradient noise* and the *feedback noise*, respectively. The matrix $Z_t$, of dimension $n_h|n_h$, has zero mean by definition and was referred to as the *autocorrelation matrix noise* in [4].

Based on the relations (10)-(13), we may design a time-invariant stable rational matrix $\boldsymbol{\mathcal{L}}_k(q^{-1})$ that operates on $f_t$ and provides an estimate of $h_{t+k}$

$$\begin{aligned}
f_t &= \mathbf{R}\hat{h}_{t|t-1} + \varphi_t\varepsilon_t \;=\; \mathbf{R}h_t + \eta_t && (14)\\
\hat{h}_{t+k|t} &= \boldsymbol{\mathcal{L}}_k(q^{-1})f_t \quad . && (15)
\end{aligned}$$

The filter $\boldsymbol{\mathcal{L}}_k(q^{-1})$ will be referred to as the *learning filter.* As we shall see below the design of a learning filter is equivalent to the design of (6)-(8).

Three terms influence the tracking performance via $f_t$ in (14): The scaled and rotated parameters $\mathbf{R}h_t$, representing the useful signal, the noise $\varphi_t v_t$, and old tracking errors via the feedback noise $Z_t\tilde{h}_{t|t-1}$.

The estimation error follows from (12) and (15) as

$$\tilde{h}_{t+k|t} = (q^k\mathbf{I} - \boldsymbol{\mathcal{L}}_k(q^{-1})\mathbf{R})h_t - \boldsymbol{\mathcal{L}}_k(q^{-1})\eta_t \quad , \tag{16}$$

where $q^kh_t = h_{t+k}$. The first right-hand term is for $k = 0$ usually called the *lag error.*

If $\eta_t$ is uncorrelated with $\hat{h}_{t-i|t-i-1}, i \geq 0$, then an open-loop Wiener design of $\boldsymbol{\mathcal{L}}_k$ can be performed. If the noise $\eta_t$ is uncorrelated with the signal $h_{t-i}$, such an open-loop design is further simplified. These conditions will not always be fulfilled but they hold approximately in many important situations, since the multiplication by $Z_t$ in (11) acts as a scrambler.

Uncorrelatedness of $\eta_t$ with $h_{t-i}$ and $\hat{h}_{t-i|t-i-1}$ will below be stated as an assumption, under which $\boldsymbol{\mathcal{L}}_k(q^{-1})$ will be optimized by just treating $\eta_t$ in (16) as an additive noise, with known properties.

### III.B The Design Equations

The diagonality constraint imposed on $\boldsymbol{\mathcal{P}}_k(q^{-1})$ in the implementation (6)-(8) will correspond to a related constraint on the structure of the learning filter. The so constrained learning filter will now be optimized for parameter variations described by (3).

The design assumptions are formalized below.

*Assumption A1*: The signal $y_t$ is described by (1), where $\mathbf{R} = E[\varphi_t\varphi_t^*]$ is known, time-invariant, and nonsingular. The second order moments of $h_t$ are described by (3) with known polynomials. Zeros of $C(z^{-1})$ are located in $|z| < 1$, zeros of $D(z^{-1})$ in $|z| \leq 1$ and $e_t$ is white and stationary, with zero mean and a known $\mathbf{R}_e = E[e_te_t^*]$.

*Assumption A2*: The learning filter (15) of dimension $n_h|n_h$ is constrained to have the structure

$$\boldsymbol{\mathcal{L}}_k(q^{-1}) = \frac{Q_k(q^{-1})}{\beta(q^{-1})}\,\mathbf{R}^{-1} \qquad (17)$$

with the polynomial $\beta(z^{-1})$ having all zeros in $|z| < 1$.

*Assumption A3*: The gradient noise $\eta_t$ is uncorrelated with $h_{t-i}$ and with $\hat{h}_{t-i|t-i-1}, i \geq 0$. It is stationary and white, with zero mean and covariance matrix $\mathbf{R}_\eta$.

*Assumption A4*: The parameter-drift-to-noise ratio, defined by

$$\gamma \stackrel{\Delta}{=} \operatorname{tr}\mathbf{R}_e/\operatorname{tr}\mathbf{R}^{-1}\mathbf{R}_\eta\mathbf{R}^{-1} \qquad (18)$$

is known, nonzero and limited ($0 < \gamma < \infty$).

In A2, the inverse of the regressor covariance matrix is included *a priori* in (17). This choice is made for two reasons: First, the formulation (15) becomes equivalent to the algorithm (7) only if $\mathbf{R}^{-1}$ is a right factor of $\boldsymbol{\mathcal{L}}_k$. Second, it assures that the constrained algorithm can attain perfect tracking, with $\boldsymbol{\mathcal{L}}_0 = \mathbf{R}^{-1}$, in the noise-free case. Furthermore, assuming the gradient noise to be white in A3 simplifies the design equations.

In practice, the covariance matrix $\mathbf{R}_\eta$ of $\eta_t$ will depend on the actual choice of estimator. This means that the scalar $\gamma$ has to be adjusted iteratively. We will return to this issue in the next subsection.

*Theorem 1: The Wiener LMS learning filter.* Under Assumptions A1 to A4, the optimal constrained learning filter (17) minimizing (5) is unique. The polynomial $\beta(q^{-1})$ is the stable and monic solution to the polynomial spectral factorization

$$r\beta\beta_* = \gamma CC_* + DD_* \qquad (19)$$

with $r$ being a real-valued positive scalar. The polynomial $Q_k(q^{-1})$ is together with a polynomial $L_{k*}(q)$ the unique solution to the Diophantine equation

$$q^k\gamma CC_* = rQ_k\beta_* + qDL_{k*} \qquad (20)$$

where $Q_k(q^{-1})$ and $L_{k*}(q)$ have degrees

$$\begin{aligned} nQ_k &= \max(n_c - k, n_D - 1) \\ nL_{k*} &= \max(n_c + k, n_\beta) - 1 \end{aligned} \qquad (21)$$

The tracking error $\tilde{h}_{t+k|t}$ will be stationary with zero mean □

*Proof:* See [9].

Learning filters $\boldsymbol{\mathcal{L}}_k(q^{-1})$ determined by the design equations (19) and (20), all have the same denominator polynomial $\beta(q^{-1})$ for any lag $k$. Since $\beta(q^{-1})$ is a stable spectral factor, $\boldsymbol{\mathcal{L}}_k(q^{-1})$ is causal and stable[1].

The spectral factorization can be solved by computing the roots of the right-hand side of (19) and forming $\beta(q^{-1})$ from the factors with stable roots. There also exist several iterative Newton algorithms for spectral factorization, see [5].

For predictors, $k > 0$, the computational complexity of the estimator (the degree of $Q_k$) is independent of the prediction horizon.

Equation (20) is a polynomial Diophantine equation. If equated for equal powers of $q$ and $q^{-1}$, it constitutes a linear system of equations, with equal number of unknowns and equations. It can always be solved[2] with respect to the coefficients of $Q_k(q^{-1})$ and $L_{k*}(q)$ [1]. This operation can be simplified still further, since a closed-form solution exists [9].

We now return to the implementation (6)-(8).

*Lemma 1:* For a learning filter designed according to Theorem 1, the adaptive filter implementation (6)-(8) is equivalent to filtering by (15). The optimal adaptation gain is

$$\mu = Q_0^0 = 1 - \frac{1}{r} \qquad (22)$$

where $Q_0^0$ constitutes the leading coefficient of the polynomial $Q_0(q^{-1})$ and $r$ originates from (19). □

*Proof:* See [9].

*Remark:* Since $\gamma$, defined in (18), belongs to $]0,\infty[$, $r \in ]1,\infty[$, we have $\mu \in ]0,1[$.

[1] By Assumption A1, the right-hand side terms of (19) cannot have common zeros on $|z| = 1$.

[2] Since $\beta(z^{-1})$ is stable, it has all zeros in $|z| < 1$, so $\beta_*(z)$ has all zeros in $|z| > 1$. Since $D(z^{-1})$ has zeros in $|z| \leq 1$, $\beta_*(z)$ and $zD(z^{-1})$ will be coprime.

In order to determine the coefficient smoothing-prediction filter corresponding to $\boldsymbol{\mathcal{L}}_k(q^{-1})$, we note that, by (8) and (15),

$$\hat{h}_{t+k|t} = \boldsymbol{\mathcal{P}}_k(q^{-1})\hat{h}_{t|t} \quad \text{and} \quad \hat{h}_{t|t} = \boldsymbol{\mathcal{L}}_0(q^{-1})f_t \quad .$$

The $k$-step estimate may thus be expressed in two ways, by

$$\hat{h}_{t+k|t} \triangleq \boldsymbol{\mathcal{L}}_k(q^{-1})f_t = \boldsymbol{\mathcal{P}}_k(q^{-1})\boldsymbol{\mathcal{L}}_0(q^{-1})f_t \quad .$$

Thus, $\boldsymbol{\mathcal{P}}_k(q^{-1})$ will be obtained as

$$\boldsymbol{\mathcal{P}}_k(q^{-1}) = \boldsymbol{\mathcal{L}}_k(q^{-1})\boldsymbol{\mathcal{L}}_0^{-1}(q^{-1}) = \frac{Q_k(q^{-1})}{Q_0(q^{-1})}\mathbf{I} \quad , \qquad (23)$$

with $Q_k(q^{-1})$ and $Q_0(q^{-1})$ obtained from Theorem 1. It is shown in [9] that all zeros of $z^{n_\beta}Q_0(z^{-1})$ are located in $|z| < 1$, so $\boldsymbol{\mathcal{P}}_k(q^{-1})$ will always be stable.

In the special case of considering first and second order hypermodels, the spectral factorization (19) has an analytical solution [6, 9], which provides $r$, and $\beta(q^{-1})$. For example, it is shown in [9] that when (3) is given by

$$h_t + d_1 h_{t-1} + d_2 h_{t-2} = e_t \quad , \qquad (24)$$

the polynomial $Q_k(q^{-1})$ for $k \geq 0$ can be explicitely calculated according to

$$Q_k(q^{-1}) = \mu(1 \ q^{-1}) \begin{pmatrix} -d_1 & 1 \\ -d_2 & 0 \end{pmatrix}^k \begin{pmatrix} 1 \\ p \end{pmatrix} \qquad (25)$$

where $p$ is defined as

$$p = \frac{d_1 d_2(1-\mu)}{1 + d_2(1-\mu)} \quad , \qquad (26)$$

and where $\mu$ is obtained from $r$ via (22), or used as a design variable. Smoothing estimators can be obtained via the backward recursion in $k$,

$$\begin{aligned} Q_{k-1}(q^{-1}) &= q^{-1}Q_k(q^{-1}) + D(q^{-1})\frac{L_0^{k*}}{r} && (27) \\ L_{k-1*}(q^{-1}) &= q^{-1}(L_{k*}(q) - L_0^{k*}\beta_*(q)), && (28) \end{aligned}$$

with $1/r = 1-\mu$, $L_{0*}(q) = Q_{1*}(q)$ and $L_0^{k*}$ being the leading coefficient of $L_{k*}(q)$. Thus, the filter $\boldsymbol{\mathcal{P}}_k(q^{-1})$ is obtained by simple algebraic expressions. This design is denoted the *Simplified WLMS algorithm.*

*Example.* Consider a parameter vector $h_t$ with first order low-pass dynamics

$$h_t = a h_{t-1} + e_t$$

where $0 < a \leq 1$. With $d_2 = 0$ and $d_1 = -a$, (25) gives $Q_1(q^{-1}) = \mu a$, $Q_0(q^{-1}) = \mu$, so the one-step predictor becomes $\hat{h}_{t+1|t} = a\hat{h}_{t|t}$. Thus, the optimal filter estimate (6),(7) is given by

$$\begin{aligned} \varepsilon_t &= y_t - \varphi_t^* a \hat{h}_{t-1} \\ \hat{h}_t &= a\hat{h}_{t-1} + \mu \mathbf{R}^{-1}\varphi_t \varepsilon_t \quad . \end{aligned}$$

For diagonal $\mathbf{R}$ and $a < 1$, this is LMS with leakage [13] which for random walk dynamics ($a = 1$) reduces to the ordinary LMS algorithm □

The WLMS algorithm can also be generalized to structures with more degrees of freedom, which may offer higher performance in some applications, and are better equipped to handle large spreads in the properties of the elements of $h_t$. These generalizations, which remove the structural constraints on learning filters and allow for colored gradient noise $\eta_t$, are presented in [6] and [7], respectively.

### III.C Iterative Design.

How is the design procedure to be applied in practice, when the properties of the gradient noise $\eta_t$ may be hard to know in advance? An obvious approach is to iterate the design a few times [7]: A design may first be based on preliminary assumptions on the gradient noise level $\operatorname{tr}\mathbf{R}_\eta$, e.g. by neglecting the feedback noise and assuming $\eta_t = \varphi_t v_t$. A better estimate of the actual gradient noise level can then be obtained. In some situations, exact analytical expressions for the tracking MSE can be used, see the discussion below. In others, the gradient noise must be investigated by simulation. We then compute $\mathbf{R}_\eta$ from (11) and calculate a modified tracking algorithm.

Since the bandwidth of the learning filter is controlled by one scalar parameter, $\gamma$, an alternative is to just use it as a tuning knob, to obtain a desired tradeoff between noise sensitivity and tracking ability.

The feedback noise $Z_t\tilde{h}_{t|t-1}$ will often be uncorrelated with $\hat{h}_{t-i|t-i-1}$, but when $Z_t \neq 0$ it will never be *independent*, due to the feedback loop in Figure 1. The loop could become unstable. As discussed in [8], the gain of $\boldsymbol{\mathcal{L}}_1(q^{-1})$ cannot be allowed to be arbitrarily large and the small gain theorem [12] will provide (conservative) sufficient conditions for stability.

In [8], three important scenarios are discussed in which an exact stability, convergence and performance analysis is possible assuming $v_t$, $\varphi_t$ and $e_t$ to be mutually independent.

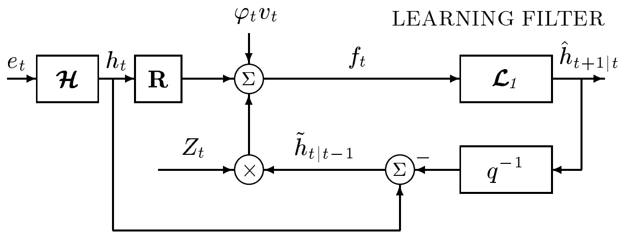


Figure 1: The feedback loop via the feedback noise $Z_t\tilde{h}_{t|t-1}$ may significantly affect the variance of the fictitious measurement $f_t$, and causes dependence with $\tilde{h}_{t|t-1}$.

*1. "Slowly" varying parameters (vanishing feedback noise).* We then have a true open-loop situation. When the power of $e_t$ becomes small relative to the power of $\varphi_t v_t$, then the impact of the feedback noise $Z_t \tilde{h}_{t|t-1}$ on our obtained tracking MSE vanishes. This situation occurs either when the parameters $h_t$ vary slowly, or when the noise level is high[3]. Then, $\eta_t \approx \varphi_t v_t$, and $\eta_t$ will be white whenever $v_t$ or $\varphi_t$ is a white sequence.

*2. Independent consecutive regression matrices.* If $\varphi_t^*$ and $\varphi_s^*$ are independent for $t \neq s$, then the feedback noise $Z_t \tilde{h}_{t|t-1}$ will be white with zero mean and its covariance can be derived exactly.

*3. FIR models of order 2 with white and symmetrically distributed regressors with constant modulus.* The performance can then be predicted without approximations from theoretical expressions, for arbitrarily fast variations of $h_t$. This is the case, for example, for channel models in IS-136 [11].

## IV. Simulation Example

In a scalar FIR system $y_t = h_{0,t}u_t + h_{1,t}u_{t-1} + v_t$ where $v_t$ is white noise with variance 0.03, the parameter evolution is described by

$$\begin{pmatrix} h_{0,t} \\ h_{1,t} \end{pmatrix} = \frac{1}{1 - 2\rho\cos\omega_0 q^{-1} + \rho^2 q^{-2}} \begin{pmatrix} e_{0,t} \\ e_{1,t} \end{pmatrix} \quad (29)$$

where $\mathbf{R}_e = 10^{-5}\mathbf{I}$, $\rho = 0.995$ and $\omega_0 = 0.015$ (SNR 21dB). The steady-state tracking performance for $k = 0$ has been compared by simulation for the time-varying Kalman filter, the (simplified) WLMS algorithm, LMS and RLS with exponential forgetting. The 4-state Kalman estimator and the WLMS algorithms are both based on the known hypermodel (29) and a known $\mathbf{R}$. The step-size in LMS and the forgetting factor in RLS were optimized by simulation. The regressors $u_t$ with variance 1 are either white and binary (B) or Gaussian (G). For Gaussian signals, we investigate two cases: white $u_t$, resulting in $\mathbf{R} = \mathbf{I}$ and colored regressors, resulting in a covariance matrix with eigenvalue spread $\chi(\mathbf{R}) = 10$.

| | $\chi(\mathbf{R})$ | Kalman | WLMS | LMS | RLS |
|---|---|---|---|---|---|
| B | 1 | 0.011 | 0.0115 | 0.020 | 0.026 |
| G | 1 | 0.012 | 0.015 | 0.032 | 0.038 |
| G | 10 | 0.026 | 0.038 | 0.085 | 0.075 |
| | #mult. | 260 | 36 | 18 | 88 |
| | $\mathbf{R}^{-1} = \mathbf{I}$: | 260 | 24 | 18 | 88 |

Table 1: Steady state mean square tracking error and number of real multiplications per time step obtained by optimized Kalman, WLMS, LMS and RLS adaptation algorithms, for binary (B) and Gaussian (G) regressors with different eigenvalue spread $\chi(\mathbf{R})$. Last line: Complexity for $\mathbf{R}^{-1} = \mathbf{I}$.

The results are summarized in Table 1, where we also compare the number of real multiplications per time step.[4] It can be noted that the WLMS design attains almost the same performance as the optimal time-varying Kalman estimator at a much lower computational complexity, not much above that of the LMS algorithm.

[3] Another case when $Z_t \tilde{h}_{t|t-1}$ vanishes completely is when $\varphi_t^*$ is scalar and has constant modulus. Then, $Z_t = 0$. This will be the case when tracking flat fading channels in mobile radio systems using e.g. PSK symbol alphabets.

[4] Multiplications between complex numbers are counted as four real multiplications, while multiplications or divisions between a real and a complex number are counted as two real multiplications.

## References


[1] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations," *IEEE Trans. on Signal Proc.*, vol. 39, pp. 2387–2399, 1991.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Prentice Hall, Englewood Cliffs, NJ, 1979.

[3] A. Benveniste, "Design of adaptive algorithms for tracking of time–varying systems," *Int. J. of Adaptive Control and Signal Processing*, vol. 2, pp. 3–29, 1987.

[4] W. A. Gardner, "Nonstationary learning characteristics of the LMS algorithm," *IEEE Trans. on Circuits and Systems*, vol. 34, pp. 1199–1207, 1987.

[5] D. Henrion *et. al.* "Polynomial toolbox, version 2," See http://www.math.utwente.nl/polbox.

[6] L Lindbom, *A Wiener filtering approach to the design of tracking algorithms: with application in mobile radio communications.* PhD Thesis, Department of Technology, Uppsala University, Sweden, 1995. www.signal.uu.se/Publications/abstracts/a951.html

[7] M. Sternad, L Lindbom, and A. Ahlén "Tracking of time-varying systems, Part I: Wiener design of adaptation algorithms with time-invariant gains". Submitted. www.signal.uu.se/Publications/abstracts/r001.html

[8] A. Ahlén, L. Lindbom and M. Sternad, "Tracking of time-varying systems, Part II: Analysis of stability and performance". Submitted. www.signal.uu.se/Publications/abstracts/r002.html

[9] L. Lindbom, M. Sternad and A. Ahlén, "Tracking of time-varying mobile radio channels. Part I: The Wiener LMS algorithm," Submitted. www.signal.uu.se/Publications/abstracts/r003.html

[10] L. Lindbom, A. Ahlén, M. Sternad and M. Falkenström, "Tracking of time-varying mobile radio channels. Part II: A case study on D-AMPS 1900 channels," Submitted. www.signal.uu.se/Publications/abstracts/r004.html

[11] M. Sternad, L. Lindbom and A. Ahlén, "Tracking of time-varying mobile radio channels with WLMS algorithms: A case study on D-AMPS 1900 channels," *IEEE VTC2000-Spring*, Tokyo, May 16-18 2000.

[12] J. M. Maciejowski, *Multivariable Feedback Design.* Addison Wesley, Reading, Mass, 1989.

[13] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1985.