ROBUST WIENER DESIGN OF ADAPTATION LAWS WITH CONSTANT GAINS

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

*Signals and Systems, Uppsala University, PO Box 528, SE-75120, Uppsala, Sweden.
[1]Ericsson Infotech, PO Box 1038, SE-65115 Karlstad
mikael.sternad@signal.uu.se (corresponding author), fax +46 18 555096

Abstract: Filters can be introduced into LMS-like adaptation algorithms to improve their tracking performance. This paper discusses the systematic model-based design of such filters. Parameter variations in coefficients of linear regression models are modeled as ARIMA-processes. The aim is to provide high performance filtering, prediction or fixed lag smoothing estimates for arbitrary lags. The properties of the time-varying parameters are in general not known exactly, so a robust design for a set of possible models will be of interest. We minimize the average tracking MSE, based on probabilistic descriptions of the model uncertainty. The method is based on a novel signal transformation that recasts the algorithm design into a robust Wiener filtering problem. The performance is illustrated on the tracking of mobile radio channels in IS-136 systems, based on a model of the time-variations affected by parametric uncertainty.

1. INTRODUCTION

Algorithms that estimate time-varying parameters of models and filters in linear regression form are important tools for signal processing, control and digital communication. Kalman estimators are the optimal linear algorithms when the statistics of the parameter variations are known. However, their complexity is often unacceptable in high-speed applications.

The required low complexity of channel estimators in mobile radio systems has motivated us to develop a class of adaptation laws that for channel tracking attain close to the optimal Kalman performance, at a computational complexity close to that of LMS algorithms (Lindbom 1995, Sternad *et.al.* 2000, Lindbom *et.al.* 2000a). The approach also opens up new ways of analyzing adaptation laws for fast variations (Ahlén *et.al.* 2000), that provide accurate estimates of the parameter error variance. These and related algorithms can, for example, be applied effectively on the fading 1900MHz channels of North American IS-136 mobile radio systems (Lindbom *et.al.* 2000b). Design of a related class of algorithms has been investigated by Benveniste *et.al.* (1990) for slowly varying parameters. Section 4 and 5 outline an iterative Wiener design that is effective also for fast variations.

A model-based Wiener design may become sensitive to the assumed model. We here propose a method for decreasing this sensitivity. Using tools from Sternad and Ahlén (1993) and Öhrn *et.al.* (1995), the design equations are in Section 6 modified to minimize the *average* tracking MSE, based on probabilistic descriptions of the uncertainty in the parameter models. In Section 7, time-varying mobile radio channels in IS-136 systems are estimated. The model of the time-variations is there affected by parametric uncertainty in the Doppler frequency and the robust design is performed by using an averaged covariance function.

Notation: Here, $R(q^{-1})$, $\boldsymbol{R}(q^{-1})$ and $\boldsymbol{\mathcal{R}}(q^{-1})$ denote polynomials, polynomial matrices and causal rational matrices, respectively. Conjugate matrices $\boldsymbol{P}_*(q)$ or $\boldsymbol{\mathcal{R}}_*(q)$ are obtained by conjugating complex coefficients, transposing and substituting the forward shift operator q for the backward shift operator q^{-1}.

2. OUTLINE OF THE PROBLEM

A sequence of measurement vectors $\{y_t\}$ of dimension $n_y|1$ is assumed available at the discrete time instants $t = 0, 1, 2, ...$ and to be generated by a linear regression

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

where v_t is a noise vector that is uncorrelated with the $n_y|n_h$ regression matrix φ_t^*. We assume the possibly complex-valued regressors to be known at time t and to be persistently exciting, so that

$$\mathbf{R} \triangleq \mathrm{E}\{\varphi_t \varphi_t^*\} \tag{2}$$

is nonsingular. The covariance matrix $\mathbf{R}$ will here be assumed time-invariant, but it can in practice be slowly varying. The time-varying parameter vector

$$h_t = (h_{0,t} \ \dots \ h_{n_h-1,t})^T \tag{3}$$

is to be estimated, with the order n_h assumed known. Models describing the variation of h_t are sometimes called *hypermodels* (Benveniste *et.al.* 1990). We here use linear time-invariant stochastic models

$$h_t = \boldsymbol{\mathcal{H}}(q^{-1}) e_t \quad , \tag{4}$$

where e_t is white noise with covariance matrix $\mathbf{R}_e$ and where $\boldsymbol{\mathcal{H}}(q^{-1})$ is an $n_h|n_h$ matrix of stable or marginally stable transfer operators. The model (4) is for now assumed known, but will in Section 6 be assumed uncertain. Denote the tracking error by

$$\tilde{h}_{t+k|t} \triangleq h_{t+k} - \hat{h}_{t+k|t} \quad (5)$$

where the estimate $\hat{h}_{t+k|t}$ may be obtained by filtering ($k = 0$), prediction ($k > 0$) or fixed lag smoothing ($k < 0$). Kalman estimators, based on (1) and on state-space realizations of (4), are the linear estimators that minimize the tracking covariance matrix

$$\mathbf{P}_{k,t} \triangleq \mathrm{E}\{\tilde{h}_{t+k|t}\tilde{h}^*_{t+k|t}\} \quad . \quad (6)$$

Since φ^*_t in (1) is time-varying, the Kalman gains will not converge as $t \to \infty$, so on-line Riccati updates are required.

We here consider a class of adaptation laws that avoid on-line Riccati updates. Instead, pre-designed linear time-invariant filters $\boldsymbol{\mathcal{M}}_k(q^{-1})$ operate on the negative instantaneous gradient of $|\varepsilon_t|^2$ with respect to $\hat{h}_{t|t-1}$,

$$\varepsilon_t = y_t - \varphi^*_t \hat{h}_{t|t-1} \quad (7)$$

$$\hat{h}_{t+k|t} = \boldsymbol{\mathcal{M}}_k(q^{-1})\varphi_t\varepsilon_t \quad . \quad (8)$$

The LMS algorithm

$$\hat{h}_{t+1|t} = \frac{\mu}{1-q^{-1}}\mathbf{I}\ \varphi_t\varepsilon_t \quad , \quad (9)$$

where $\mu > 0$ is a scalar gain, constitutes a simple special case of the structure (7),(8). The rational matrix $\boldsymbol{\mathcal{M}}_k$ can be selected to asymptotically minimize (6) under various constraints and assumptions.

3. THE LOOP TRANSFORMATION

The algorithm (7),(8) can be expressed as a stable and causal filter, denoted the *learning filter* $\boldsymbol{\mathcal{L}}_k(q^{-1})$, that operates on a signal vector

$$f_t \triangleq \varphi_t\varepsilon_t + \mathbf{R}\hat{h}_{t|t-1} \quad . \quad (10)$$

Since (7),(8) give $\hat{h}_{t|t-1} = q^{-1}\boldsymbol{\mathcal{M}}_1(q^{-1})\varphi_t\varepsilon_t$,

$$\varphi_t\varepsilon_t = (\mathbf{I} + q^{-1}\mathbf{R}\boldsymbol{\mathcal{M}}_1(q^{-1}))^{-1} f_t \quad .$$

Thus, we obtain

$$\hat{h}_{t+k|t} = \boldsymbol{\mathcal{M}}_k(q^{-1})(\mathbf{I}+q^{-1}\mathbf{R}\boldsymbol{\mathcal{M}}_1(q^{-1}))^{-1} f_t \triangleq \boldsymbol{\mathcal{L}}_k(q^{-1}) f_t \, . \quad (11)$$

By (7) and (1),

$$\varphi_t\varepsilon_t = \varphi_t\varphi^*_t\tilde{h}_{t|t-1} + \varphi_t v_t \quad . \quad (12)$$

Adding and subtracting $\mathbf{R}\tilde{h}_{t|t-1}$ on the right-hand side of (12) gives

$$\varphi_t\varepsilon_t = \mathbf{R}h_t - \mathbf{R}\hat{h}_{t|t-1} + (\varphi_t\varphi^*_t - \mathbf{R})\tilde{h}_{t|t-1} + \varphi_t v_t \quad . \quad (13)$$

We now define

$$Z_t \triangleq \varphi_t\varphi^*_t - \mathbf{R} \quad (14)$$

$$\eta_t \triangleq Z_t\tilde{h}_{t|t-1} + \varphi_t v_t \quad (15)$$

which we call the *autocorrelation matrix noise* and the *gradient noise*, respectively. The signal f_t which can be regarded as a fictitious measurement, can then by using (10), (13), (14) and (15), be expressed as

$$f_t = \mathbf{R}h_t + Z_t\tilde{h}_{t|t-1} + \varphi_t v_t = \mathbf{R}h_t + \eta_t \quad , \quad (16)$$

see Fig. 1. The design of our adaptation law (7),(8) has now been transformed into a *Wiener filter design* for $\boldsymbol{\mathcal{L}}_k(q^{-1})$, where η_t plays the role of noise, see Fig. 2.

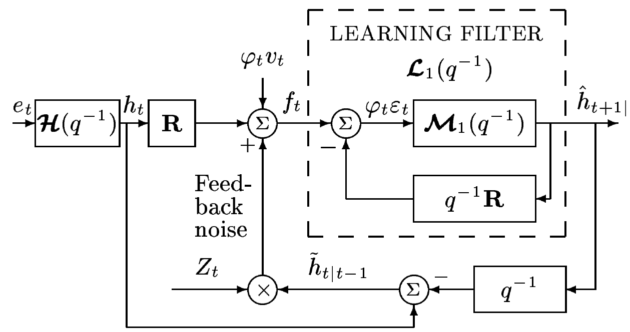

Figure 1: The adaptation algorithm (8) operates in closed loop. This loop can be decomposed into an inner time-invariant feedback of $\mathbf{R}\hat{h}_{t|t-1}$ and an outer time-varying loop via the feedback noise $Z_t\tilde{h}_{t|t-1}$.

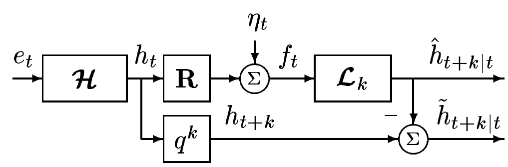

Figure 2: The filter design problem. The vector h_{t+k} is to be estimated from f_t, such that the steady state tracking error covariance matrix of the parameter error $\tilde{h}_{t+k|t}$ is minimized.

The gradient noise η_t is affected by the term $Z_t\tilde{h}_{t|t-1}$, here called the *feedback noise*. It is shown in Ahlén *et.al.* (2000) that the feedback noise is negligible either when h_t has small increments or when the noise v_t has high variance. Such situations are denoted "slow variations" (Ahlén *et.al.*, 2000; Macchi, 1995). The optimal learning filter will then operate in open loop, with $\eta_t \approx \varphi_t v_t$. Stability and convergence in MSE is then guaranteed by stability of the learning filter, which follows directly from a Wiener design. (While the learning filter $\boldsymbol{\mathcal{L}}_k(q^{-1})$ must be stable, the filter $\boldsymbol{\mathcal{M}}_1(q^{-1})$ need not be stable, since it works within the inner feedback loop of Fig. 1.) An iterative design must bee performed when $Z_t\tilde{h}_{t|t-1}$ cannot be neglected, see Section 5.

4. LEARNING FILTER OPTIMIZATION

The criterion (6) for $t \to \infty$ could be minimized directly by adjusting $\boldsymbol{\mathcal{L}}_k(q^{-1})$, if $\boldsymbol{\mathcal{H}}(q^{-1})$ in (4) and the properties of η_t were known exactly. Using a polynomial approach to Wiener filtering, the learning filter is here designed under the constraint of stability, and under the following assumptions.

Assumption A1: The sequence $\{\varphi_t^*\}$ is stationary and known, with $\mathbf{R}$ known and nonsingular $\square$
Assumption A2: The gradient noise η_t is white and stationary with zero mean and known covariance matrix $\mathbf{R}_\eta$. The correlation of η_t with h_{t-i} and with $\hat{h}_{t-i|t-i-1}, i \geq 0$ is negligible. $\square$
Assumption A3: The time-varying parameters are described by a known vector-ARIMA process

$$\boldsymbol{D}(q^{-1})h_t = \boldsymbol{C}(q^{-1})e_t \quad , \qquad (17)$$

with $\mathbf{R}_e = \mathrm{E}\, e_t e_t^*$ nonsingular, where $\boldsymbol{D}(q^{-1}) = D_u(q^{-1})\boldsymbol{D}_s(q^{-1})$. Moreover, $\boldsymbol{C}(q^{-1})$ and $\boldsymbol{D}_s(q^{-1})$ are monic and stably invertible, while the polynomial $D_u(q^{-1})$ has zeros on the unit circle $\square$

Assumption A3 implies that e.g. random walks, integrated random walks and filtered random walk models can be considered, but that the unstable dynamics $D_u(q^{-1})$ must then affect all the elements of h_t. We can now present the optimal learning filter.

Theorem 1: Under Assumptions A1-A3, the stable and causal learning filter minimizing the asymptotic parameter covariance matrix (6) is

$$\hat{h}_{t+k|t} = \boldsymbol{\mathcal{L}}_k^{opt} f_t = \boldsymbol{D}_s^{-1}\boldsymbol{Q}_k\boldsymbol{\beta}^{-1}\boldsymbol{D}_s\mathbf{R}^{-1}f_t \quad , \qquad (18)$$

where the polynomial matrix $\boldsymbol{\beta}(q^{-1})$ of dimension $n_h|n_h$ and degree $n_\beta = \max(n_C, n_D)$ is the stable left spectral factor obtained from

$$\boldsymbol{\beta}\boldsymbol{\beta}_* = \boldsymbol{C}\mathbf{R}_e\boldsymbol{C}_* + \boldsymbol{D}\mathbf{R}^{-1}\mathbf{R}_\eta\mathbf{R}^{-1}\boldsymbol{D}_* \quad . \qquad (19)$$

The unique solution to the Diophantine equation

$$q^k\boldsymbol{C}\mathbf{R}_e\boldsymbol{C}_* = \boldsymbol{Q}_k\boldsymbol{\beta}_* + q\boldsymbol{D}\boldsymbol{L}_{k*} \qquad (20)$$

provides polynomial matrices $\boldsymbol{Q}_k(q^{-1})$ and $\boldsymbol{L}_{k*}(q)$ of dimension $n_h|n_h$, with generic degrees

$$n_Q = \max(n_C - k, n_D - 1) \quad , \quad n_L = \max(n_C + k, n_\beta) - 1 \qquad (21)$$

respectively. The estimation error $\tilde{h}_{t+k|t}$ will be stationary with finite covariance matrix and zero mean.
Proof: See Sternad *et.al.* (2000), where a generalization to colored gradient noise is also presented $\square$

Under Assumptions A1-A3, the (generalized) innovations model of $f_t = \mathbf{R}h_t + \eta_t$ can be expressed as

$$f_t = \mathbf{R}\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{\beta}(q^{-1})\epsilon_t \qquad (22)$$

where ϵ_t is the white zero mean innovation sequence with unit covariance matrix. By defining the signal

$$\bar{\epsilon}_t \triangleq \frac{1}{D_u(q^{-1})}\epsilon_t = \boldsymbol{\beta}^{-1}(q^{-1})\boldsymbol{D}_s(q^{-1})\mathbf{R}^{-1}f_t \quad , \qquad (23)$$

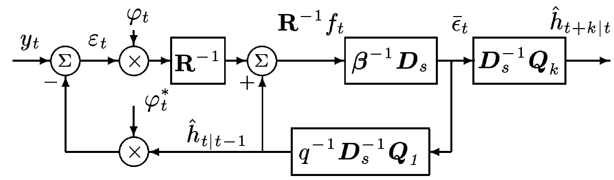

Figure 3: The Wiener optimized tracking algorithm.

the adaptation law (18) can be realized as in Fig. 3. The polynomial matrix $\boldsymbol{Q}_k(q^{-1})$ can be obtained from closed-form expressions, see Sternad *et.al.* (2000). In particular, $\boldsymbol{Q}_1(q^{-1}) = q(\boldsymbol{\beta}(q^{-1}) - \boldsymbol{D}(q^{-1})\beta_0)$, with β_0 being the leading coefficient matrix of $\boldsymbol{\beta}(q^{-1})$.

With this expression and (11),(18), the Wiener optimized filter matrix $\boldsymbol{\mathcal{M}}_k(q^{-1})$ in (8) can for white gradient noise be shown to be given by

$$\boldsymbol{\mathcal{M}}_k^{opt}(q^{-1}) = \boldsymbol{D}^{-1}(q^{-1})\boldsymbol{Q}_k(q^{-1})\beta_0^{-1}\mathbf{R}^{-1} \quad . \qquad (24)$$

The filter has the inverse regressor covariance matrix $\mathbf{R}^{-1}$ as a right factor. The optimized tracker can thus be seen as a generalization of the LMS-Newton adaptation law (Widrow and Stearns 1985).

5. ITERATIVE WIENER DESIGN

For slow time-variations, the feedback noise $Z_t\tilde{h}_{t|t-1}$ is negligible, so we may perform a one-shot design by assuming $\eta_t = \varphi_t v_t$. Otherwise, the properties of η_t depend on $\boldsymbol{\mathcal{L}}_1(q^{-1})$ via (15). The multiplication by Z_t in (15), see also Fig. 1, acts as a scrambler. For FIR models (1) with white regressors, it will reduce the correlation between the feedback noise and $\tilde{h}_{t|t-1}$. Assumption A2 will then hold for white regressor elements and a design for fast parameter variations can be obtained iteratively. (See Sternad *et.al.* (2000) for a design example.)

1. Design a one-step predictor for slow variations, i.e. use $\mathbf{R}_\eta = \mathrm{E}\{\varphi_t v_t v_t^* \varphi_t^*\}$ to design $\boldsymbol{\mathcal{L}}_1(q^{-1})$. Verify that the closed loop around $\boldsymbol{\mathcal{L}}_1(q^{-1})$ of Fig. 1 is stable. If not, scale up $\mathbf{R}_\eta$ to decrease the gain of $\boldsymbol{\mathcal{L}}_1(q^{-1})$.

2. Based on a simulation of φ_t, v_t, h_t and of $\hat{h}_{t|t-1}$, estimate $\mathbf{R}_\eta$ from $\hat{\eta}_t = \varphi_t\varepsilon_t - \mathbf{R}(h_t - \hat{h}_{t|t-1})$ (see (10),(16)), by using sample averages over $\hat{\eta}_t$.

3. Design a new estimator $\boldsymbol{\mathcal{L}}_1(q^{-1})$.

Repeat 2. and 3. until the difference in consecutive $\hat{h}_{t+1|t}$ becomes small. Then, design $\boldsymbol{\mathcal{L}}_k(q^{-1})$ for lag k.

It will be possible to find an initial stable solution under mild conditions. If $\boldsymbol{\mathcal{H}}(q^{-1})$ is stable, then $\boldsymbol{\mathcal{L}}_1(\omega) \to 0 \;\forall\; \omega$ when the assumed noise power is increased. If Z_t has bounded elements, then the small gain theorem implies that the closed loop of Fig. 1

can be stabilized by assuming a sufficiently high noise power in the design of $\mathcal{L}_1(q^{-1})$.

6. ROBUST WIENER DESIGN

The statistics of time-varying parameters will rarely be exactly known. Small uncertainties can be disregarded, but large uncertainties should be taken into account in a model-based design.

One can then use a gain scheduling approach or a minimax robust design, as investigated in Lindbom *et.al.* 2000b). Another alternative is to optimize the parameter error covariance *on average* over a set of possible dynamics for h_t. Such an approach to the design of robust Wiener filters was originally suggested by Speyer and Gustafsson (1975) The minimization of averaged quadratic criteria has been developed into a systematic design methodology in Sternad and Ahlén (1993), Öhrn *et.al.* (1995) and Öhrn (1996) and it can be applied directly here. Assume that a set of hypermodels is described by the probabilistic *extended design model*

$$h_t = (\mathcal{H}^o(q^{-1}) + \Delta\mathcal{H}(q^{-1}))\, e_t \; . \qquad (25)$$

Here, $\mathcal{H}^o(q^{-1})$ is the nominal model, while the error model $\Delta\mathcal{H}(q^{-1})$ represents the set of possible deviations, described by a set of time-invariant transfer functions, parametrized by random coefficients.

Define the operation $\bar{\mathrm{E}}(\cdot)$ as an average over all random coefficients parameterizing the error model. We assume that

$$\bar{\mathrm{E}}\,(\Delta\mathcal{H}(q^{-1})) = 0$$

so the nominal model is defined as being the average model of the set. For a set of models (25), it will be possible to minimize the average of the asymptotic tracking error covariance matrix (6)

$$\bar{\mathbf{P}}_k \;\triangleq\; \bar{\mathrm{E}}\left(\lim_{t\to\infty} \mathbf{P}_{k,t}\right) \;, \qquad (26)$$

if A3 is substituted by the following assumption:

Assumption A4: The set of hypermodels can be represented by

$$h_t = \frac{1}{\tilde{D}_u(q^{-1})}\,(\mathcal{H}^o_s(q^{-1}) + \Delta\mathcal{H}_s(q^{-1}))\, e_t \;, \qquad (27)$$

where the known polynomial $\tilde{D}_u(z^{-1})$ has zeros on $|z| = 1$, $\mathcal{H}^o_s(q^{-1})$ is a known, stable and stably invertible rational matrix and $\Delta\mathcal{H}_s(q^{-1})$ is a set of stable random rational matrices that are independent of e_t, with zero mean and known second order statistics. The noise e_t is white, with nonsingular covariance matrix $\mathbf{R}_e = \mathrm{E}\{e_t e_t^*\}$ □

The known polynomial $\tilde{D}_u(z^{-1})$ must thus include all marginally stable factors of transfer function denominators appearing in the set (25). If the marginally stable modes are unknown, there exists no single learning filter which can provide a finite average covariance matrix $\bar{\mathbf{P}}_k$.

Under Assumptions A1,A2 and A4, a robust design can be obtained by a modification of Theorem 1.[1] The key modification of the derivation in Sternad *et.al.* (2000), is the use of an averaged measurement spectrum

$$\bar{\mathrm{E}}\,(\phi_f) \triangleq \bar{\mathrm{E}}\,(\mathbf{R}\,(\mathcal{H}^o + \Delta\mathcal{H})\,\mathbf{R}_e\,(\mathcal{H}^o + \Delta\mathcal{H})_*\,\mathbf{R}) + \mathbf{R}_\eta \qquad (28)$$

Introduce the *averaged hypermodel* $\tilde{\boldsymbol{D}}^{-1}(q^{-1})\tilde{\boldsymbol{C}}\,(q^{-1})$, with $\tilde{\boldsymbol{D}}\,(q^{-1}) = \tilde{D}_u(q^{-1})\tilde{\boldsymbol{D}}_s(q^{-1})$ where $\tilde{\boldsymbol{D}}_s(q^{-1})$ and $\tilde{\boldsymbol{C}}\,(q^{-1})$ are assumed stable. It is a (generalized) innovation description, having spectral density equal to that of the average of the set (25),

$$\tilde{\boldsymbol{D}}^{-1}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\tilde{\boldsymbol{D}}_*^{-1} \;\triangleq\; \bar{\mathrm{E}}\,(\mathcal{H}^o + \Delta\mathcal{H})\,\mathbf{R}_e\,(\mathcal{H}^o + \Delta\mathcal{H})_*$$
$$= \frac{1}{\tilde{D}_u}\left(\mathcal{H}^o_s\mathbf{R}_e\mathcal{H}^o_{s*} + \bar{\mathrm{E}}\,(\Delta\mathcal{H}_s\mathbf{R}_e\Delta\mathcal{H}_{s*})\right)\frac{1}{\tilde{D}_{u*}} \;. \qquad (29)$$

By defining an averaged spectral factorization

$$\tilde{\beta}\tilde{\beta}_* = \tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_* + \tilde{\boldsymbol{D}}\mathbf{R}^{-1}\mathbf{R}_\eta\mathbf{R}^{-1}\tilde{\boldsymbol{D}}_* \;, \qquad (30)$$

the averaged measurement spectrum becomes

$$\bar{\mathrm{E}}\,(\phi_f) = \mathbf{R}\tilde{\boldsymbol{D}}^{-1}\tilde{\beta}\tilde{\beta}_*\tilde{\boldsymbol{D}}_*^{-1}\mathbf{R} \;.$$

The square polynomial matrix $\tilde{\beta}(q^{-1})$ is monic and $\det\tilde{\beta}(z^{-1})$ is stable since the right hand sides of (28)-(30) are nonsingular on the unit circle, due to the assumed non-singularity of $\mathbf{R}_e$ and stability of $\boldsymbol{C}\,(q^{-1})$. The Diophantine equation (20) is in a similar way modified by substituting $\tilde{\boldsymbol{C}}\,(q^{-1})\tilde{\boldsymbol{C}}_*(q)$ for $\boldsymbol{C}\,(q^{-1})\mathbf{R}_e\boldsymbol{C}_*(q)$ and $\tilde{\boldsymbol{D}}\,(q^{-1})$ for $\boldsymbol{D}\,(q^{-1})$:

$$q^k\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_* = \tilde{\boldsymbol{Q}}_k\tilde{\beta}_* + q\tilde{\boldsymbol{D}}\tilde{\boldsymbol{L}}_{k*} \;. \qquad (31)$$

The averaged robust design is summarized as follows:

Theorem 2: For the model set (25), with specified second order moments, a learning filter can be designed under Assumptions A1, A2 and A4 which minimizes the average covariance matrix (26), by obtaining a polynomial spectral factor $\tilde{\beta}(q^{-1})$ from (30) and a polynomial matrix $\tilde{\boldsymbol{Q}}_k(q^{-1})$, together with $\tilde{\boldsymbol{L}}_{k*}(q)$, as the unique solution of the Diophantine equation (31). The robust learning filter is then given by

$$\mathcal{L}_k^{rob} = \tilde{\boldsymbol{D}}_s^{-1}\tilde{\boldsymbol{Q}}_k\tilde{\beta}^{-1}\tilde{\boldsymbol{D}}_s\mathbf{R}^{-1} \qquad (32)$$

□

At frequencies where $\bar{\mathrm{E}}\left(\Delta\mathcal{H}(e^{-j\omega})\mathbf{R}_e\Delta\mathcal{H}_*(e^{j\omega})\right)$ is

[1] More general cases are covered by Theorem 4.1 in Öhrn (1996) which take into account also uncertainty in the transducer, here corresponding to $\mathbf{R}$, and in the noise model.

significant as compared to the nominal spectral density $\mathcal{H}^o(e^{-j\omega})\mathbf{R}_e\mathcal{H}^o_*(e^{j\omega})$, the averaged hypermodel (29) will have higher gain than the nominal hypermodel. At such frequencies, the principal gains of the robust learning filter (32) will be larger than the principal gains of the nominal filter (18), since the average signal-to-noise ratio is higher than the nominal SNR. Note that only second order moments

$$\bar{\mathrm{E}}\,(\Delta\mathcal{H}_s(q^{-1})\mathbf{R}_e\Delta\mathcal{H}_{s*}(q))$$

need to be specified, since the type of distribution, and higher order moments, will not affect the design. The required function on the right-hand side of (29) can be obtained by *averaging over hypermodels*: Draw n samples from the set of their stable parts $\{\mathcal{H}_{s_i} = \mathcal{H}^o_{s_i} + \Delta\mathcal{H}_{s_i}\}_{i=1}^n$. The corresponding spectra (or covariance functions) can then be averaged to obtain (28),(29).

7. EXAMPLE: TRACKING OF IS-136 CHANNELS WITH UNCERTAIN DOPPLER ESTIMATES

In IS-136 mobile radio systems, the symbol-spaced sampled baseband channels can be described by FIR filters with one or two time-varying taps

$$y_t = h_{0,t}u_t + h_{1,t}u_{t-1} + v_t = \varphi_t^* h_t + v_t \ , \qquad (33)$$

where y_t is the received scalar complex-valued signal while v_t is noise and co-channel interference. The transmitted symbols u_t are here assumed known, although they would in reality partly be estimated by the receiver. They have variance σ_u^2 and are mutually uncorrelated, so $\mathbf{R} = \sigma_u^2\mathbf{I}_2$ is known exactly. For a mobile terminal, the channel coefficients $h_{0,t}$ and $h_{1,t}$ will be subject to fading characterized by the maximum Doppler frequency $\omega_{\mathrm{D}} = 2\pi v_o/\lambda$, where v_o denotes the speed of the mobile and λ is the carrier wavelength, which in the following is assumed to be 16cm (∼1900 MHz). For the purpose of our investigation, we shall use Jakes' (1974) fading model, which assumes an infinite number of nearby scatterers and is parametrized by ω_{D}. When the vehicle velocity is constant, the channel coefficient vector $h_t = (h_{0,t}\ h_{1,t})^T$ will then be a stationary, complex circular Gaussian process with zero mean and covariance function

$$r_h(\ell) = \mathrm{E}\{h_t h_{t-\ell}^*\} = \mathbf{R}_h J_0(\Omega_{\mathrm{D}}\,\ell) \quad \ell = 0, \pm 1 \cdots \ . \qquad (34)$$

Here, $\mathbf{R}_h \triangleq \mathrm{E}\{h_t h_t^*\}$, $J_0(\cdot)$ denotes the Bessel function of the first kind and zero order and $\Omega = \omega T$, $\Omega_{\mathrm{D}} = \omega_{\mathrm{D}} T$. The symbol time T is 41.15μs in IS-136. This yields the classical Rayleigh fading spectrum

$$\phi_h(\Omega) = \begin{cases} \frac{2}{\sqrt{\Omega_D^2-\Omega^2}}\mathbf{R}_h & |\Omega| < \Omega_{\mathrm{D}} \\ 0 & |\Omega| > \Omega_{\mathrm{D}}\ . \end{cases} \qquad (35)$$

When Ω_{D} is known, the model (4), (17) can be adjusted to the autocorrelation function (34). Perfect adjustment would require models of infinite degree, but good performance can be obtained with simple models. In the following, third order autoregressive models (AR$_3$ models) $(1/D(q^{-1}))\mathbf{I}_2$ will be fitted to the relevant covariance function.

The maximum normalized Doppler frequency Ω_{D} can be estimated from data, but such estimates will be imperfect. We therefore here investigate the design of an algorithm for tracking h_t that is robustified against uncertainties in the Doppler estimate. The Doppler spectrum averaged with respect to the uncertain Ω_{D} constitutes our averaged model (29):

$$\tilde{\phi}_h(\Omega) \triangleq \mathrm{E}_{\Omega_{\mathrm{D}}}(\phi_h(\Omega)) = \int_{-\pi}^{\pi}\phi_h(\Omega)p(\Omega_{\mathrm{D}})\,d\Omega_{\mathrm{D}}\ . \qquad (36)$$

Here, $p(\Omega_{\mathrm{D}})$ denotes the assumed probability density function of the normalized maximum Doppler frequency Ω_{D}. When assuming Jakes model giving (34), the covariance function corresponding to (36) is

$$\tilde{r}_h(\ell) \triangleq \int_{-\pi}^{\pi}\mathbf{R}_h J_0(\Omega_{\mathrm{D}}\,\ell)p(\Omega_{\mathrm{D}})\,d\Omega_{\mathrm{D}}\ . \qquad (37)$$

In Figure 4, an element of the averaged covariance function (37) is displayed for a uniformly distributed probability density function, with different uncertainty regions. A wider uncertainty region will increase the damping of the averaged covariance function, yielding a spectrum with a less pronounced peak.

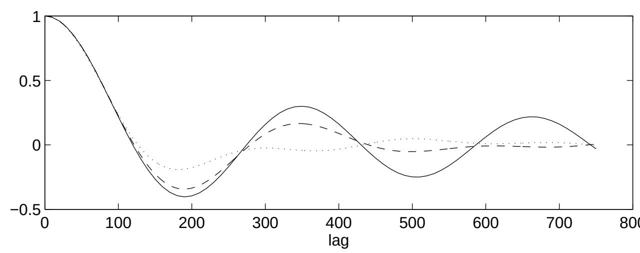

Figure 4: Auto-covariance function $r_h(\ell) = J_0(\Omega_{\mathrm{D}}\ell)$ with $\Omega_{\mathrm{D}} = 0.02$ (solid) and the averaged covariance function (37), with $\Omega_{\mathrm{D}} \in U[0.01\ \ 0.03]$ (dotted) and $\Omega_{\mathrm{D}} \in U[0.015\ \ 0.025]$ (dashed).

A (stable) averaged AR model can now be adjusted to $\tilde{r}_h(\ell)$ (Lindbom *et.al.* 2000b) and a robust tracking algorithm is then designed, as outlined in Section 6, using this model[2]. Assume the signal-to-noise ratio to be 15dB and let $\mathbf{R}_h = \mathbf{I}_2$. Use a nominal $\Omega^o_{\mathrm{D}} = 0.02$, while the true Doppler frequency is varied. A nominal Wiener design is now performed by adjusting an AR$_3$ model to (34) and then using Theorem 1 once, to design an algorithm that minimizes the one-step prediction error. (Iterations are not needed in this case.) Then, robust designs are performed based on

[2] In the work Lin *et. al.* (1995), a windowed LS algorithm was designed to take uncertainties about the Doppler frequency into account. The averaged covariance function (37) was there utilized in the choice of the adaptation window.

AR_3 models adjusted to (37).

The effect of using the robust designs is presented in Figure 5, where it is compared to the nominal design. We also compare to a perfectly matched AR_3 model-based design, based on the correct Doppler frequency. To compute the one-step prediction tracking MSE $\mathrm{tr}\,\mathbf{P}_1$, we use a novel analytical expression which is *exactly* valid for two-tap FIR channels with white inputs, and which gives very good approximations for higher order FIR models. It is derived in Ahlén *et.al.* (2000), and utilized extensively also in Lindbom *et.al.* (2000b). The averaged AR_3 model is matched to the covariance functions described by the dotted and the dashed lines in Figure 4, for lags < 200. The left hand diagram of Figure 5 reflects the performance when the uncertainty level of Ω_D is assumed moderate: $\Omega_D \in U[0.015\ \ 0.025]$, whereas the right hand diagram displays the performance for a larger uncertainty interval, $\Omega_D \in U[0.01\ \ 0.03]$.

When the Doppler frequency is known, the Wiener design (dash-dotted) is much superior to LMS tracking by (9) (dotted). Its performance is almost equal to that of a time-varying Kalman predictor designed for the same AR_3 model. For uncertain Ω_D the averaged robust design improves both the worst-case performance and the average performance significantly. The average MSE performance (area under solid line) is 33% higher with an averaged robust design, than for a known Ω_D (area under dash-dotted line). It would be 80% higher for the nominal design (dashed). The effect is significant also for the more moderate uncertainties (left-hand diagram).

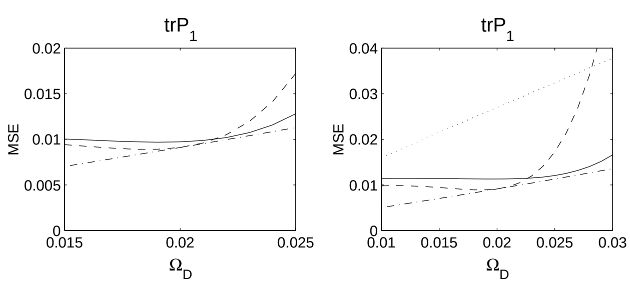

Figure 5: One-step prediction MSE performance of Wiener designs based on AR_3 fading models: Averaged robust (solid), nominal design for $\Omega_D = 0.02$ (dashed) and optimally matched to known model (dash-dotted) . The assumed uncertainty $\Omega_D \in U[0.015\ \ 0.025]$ (left) and $\Omega_D \in U[0.01\ \ 0.03]$ (right). Also shown is an LMS design tuned for a known Ω_D (upper dotted in right figure).

As a generalization of the design above, deviations from the idealized Jakes' model can be introduced. They can be regarded as unstructured uncertainty, which could also be incorporated in an averaged robust design, using methods described in Sternad and Ahlén (1993) and Öhrn (1996).

REFERENCES

Ahlén, A., L.Lindbom and M.Sternad (2000). Tracking of time-varying systems, Part II: Analysis of stability and performance of adaptation algorithms with time-invariant gains. Report R002, Signals and Systems, Uppsala Univ. www.signal.uu.se/Publications/abstracts/r002.html

Benveniste, A., M.Métivier and P.Priouret (1990). *Adaptive Algorithms and Stochastic Approximations.* Springer.

Jakes, W.C. (1974). Multipath Interference. In W. C. Jakes ed. *Microwave Mobile Communications.* Wiley, New York, NY. Reissued by IEEE.

Lindbom L. (1995). *A Wiener filtering approach to the design of tracking algorithms.* PhD Thesis, Dept. of Technology, Uppsala University, Sweden.

Lindbom L., M. Sternad and A. Ahlén (2000a). Tracking of time-varying mobile radio channels. Part I: The Wiener LMS algorithm. To appear in *IEEE Trans. Com.* www.signal.uu.se/Publications/abstracts/r003.html

Lindbom, L., A.Ahlén M.Sternad and M. Falkenström (2000b). Tracking of time-varying mobile radio channels. Part II: A case study. To appear in *IEEE Trans. Com.*. www.signal.uu.se/Publications/abstracts/r004.html

Macchi, O. (1995). *Adaptive Processing: The Least Mean Squares Approach with Applications in Transmission.* Wiley, New York.

Lin, J., J. G. Proakis, F. Ling and H. Lev-Ari (1995). Optimal channel tracking of time-varying channels: A frequency domain approach for known and new algorithms. *IEEE Trans. Select. Areas Commun.*, **vol. 13**, 141-154.

Speyer J.L. and D. E. Gustafson (1975). An approximation method for estimation in linear systems with parameter uncertainty. *IEEE Transactions on Automatic Control*, **vol. 20**, pp. 354-359, 1975.

Sternad M. and A. Ahlén (1993). Robust filtering and feedforward control based on probabilistic descriptions of model errors. *Automatica*, **vol 29**, pp. 661-679.

Sternad, M., L.Lindbom, and A.Ahlén (2000). Tracking of time-varying systems. Part I: Wiener design of algorithms with constant gains. Submitted. www.signal.uu.se/Publications/abstracts/r001.html

Widrow B. and S. D. Stearns (1985) *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ.

Öhrn K., A. Ahlén and M. Sternad (1995). A probabilistic approach to multivariable robust filtering and open-loop control. *IEEE Trans. Aut. Contr.*, **vol. 40**, pp. 405-417.

Öhrn, K. (1996) *Design of Multivariable Cautious Wiener Filters.* PhD Thesis Dept. Tech, Uppsala Univ., Sweden.