DESIGN AND ANALYSIS OF ADAPTATION ALGORITHMS WITH TIME-INVARIANT GAINS

Mikael Sternad*

Outline of joint work with Lars Lindbom[1] and Anders Ahlén*

*Signals and Systems Group, Uppsala University

[1] Ericsson Infotech, Karlstad

(See www.signal.uu.se/Publications/abstracts/r001.html - r004.html)

UPPSALA UNIVERSITET

Time-Varying Linear Regression Models

$$y_t \;=\; \varphi_t^* h_t + v_t$$

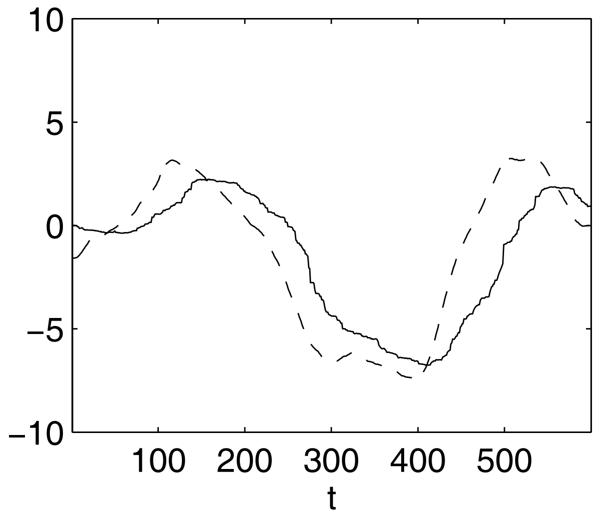
Complex-valued and possibly MIMO, with φ_t^* known at time t.

Example: Mobil radio channel

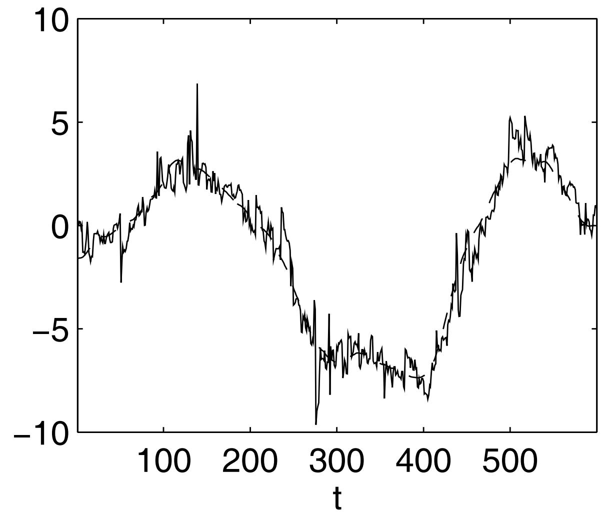
$$y_t \;=\; (u_t \ldots u_{t-M+1}) \begin{pmatrix} h_{0,t} \\ \vdots \\ h_{M-1,t} \end{pmatrix} + v_t$$

Our goal: Estimate vector h_t when $\mathbf{R} = \mathrm{E}\left(\varphi_t \varphi_t^*\right)$ is known.

The LMS Stepsize Dilemma



μ small



μ large

Tracking Algorithms

- Time-varying Kalman Filter (time varying gain)
- RLS with exponential forgetting (time varying gain)
- LMS and LMS-Newton (time invariant gain)
- General Constant Gain Algorithms (NEW) (time invariant gain)

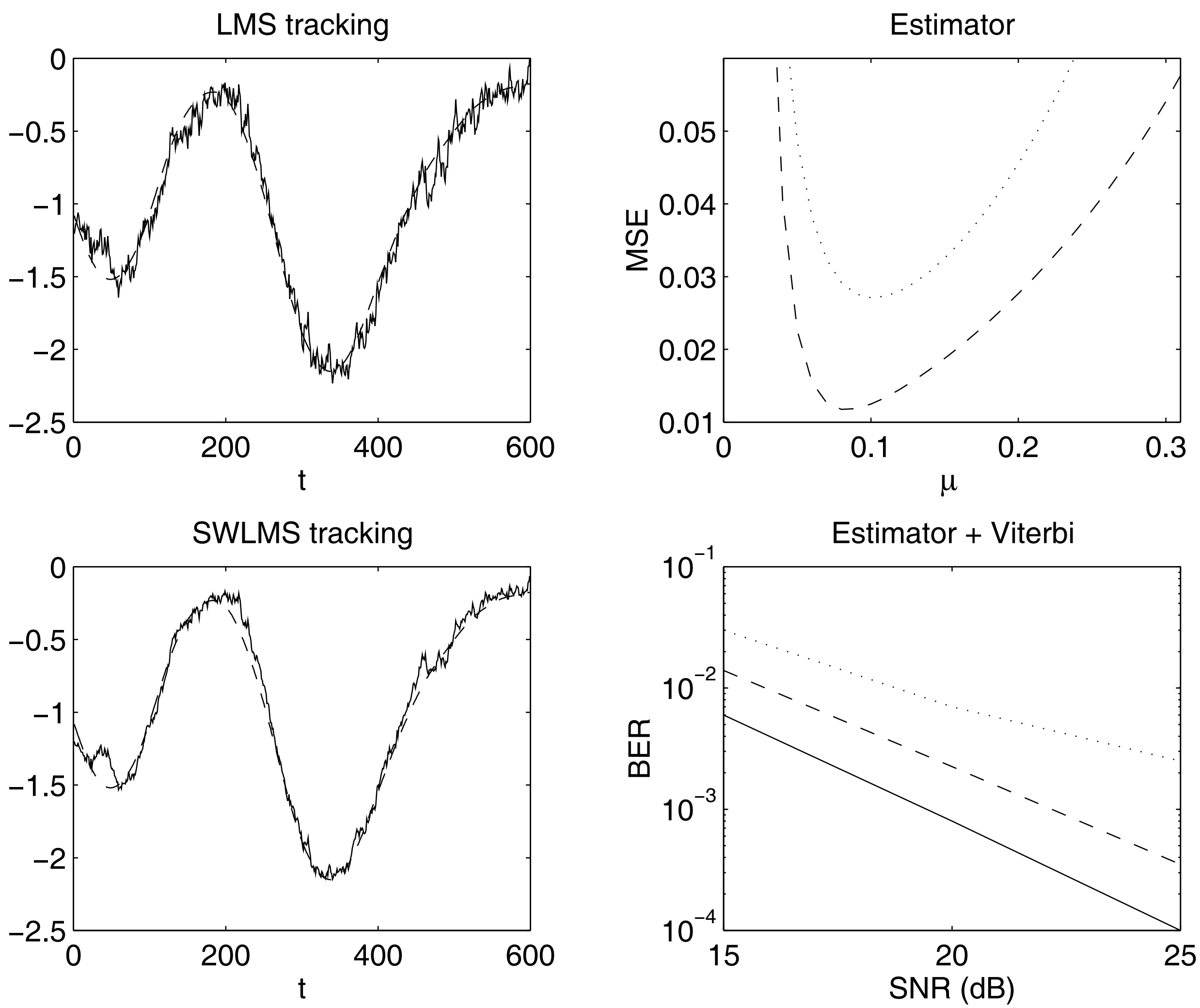
  \+ Structure and tuning can be tailored to ARIMA-type time-variations

  \+ Low complexity

  \+ Performance close to Kalman

  \- At present restricted to slowly time-varying $\mathbf{R} = \mathrm{E}\left(\varphi_t \varphi_t^*\right)$.

Tracking Performance Example 1:



Tracking Performance Example 2:

Tracking MSE for complex two-tap fast fading (second order AR):

Input (symbol) properties	Kalman	WLMS	LMS	RLS
White and constant modulus:	0.011	0.011	0.020	0.026
White and Gaussian:	0.012	0.015	0.032	0.038
Colored Gaussian $(\lambda_{max}/\lambda_{min} = 10)$:	0.026	0.038	0.085	0.075
Real add+mult/sample for white inputs:	214	30	18	72
For colored inputs $(\mathbf{R} \neq c\mathbf{I})$:	214	44	18	72

Tracking MSE can for WLMS and LMS be predicted exactly for 2-tap FIR and well approximated for higher order FIR models.

Outline of Presentation:

1. Design:
   - Structure of the tracking algorithms
   - Wiener design
   - Simplified algorithms.
   - Iterative design.
2. Analysis:
   - Slow variations
   - Fast variations.
3. D-AMPS 1900 Channel Tracking
4. Summary

Exploiting Prior Information

"Known" time-variability will improve tracking.

Model: $h_t = \boldsymbol{\mathcal{H}}(q^{-1})e_t$, for example

$$h_t = \frac{C(q^{-1})}{D(q^{-1})}\mathbf{I}e_t = \frac{1 + c_1 q^{-1} + \ldots + c_{n_C} q^{-n_D}}{1 + d_1 q^{-1} + \ldots + d_{n_D} q^{-n_D}}\mathbf{I}e_t$$

Examples: Let $C(q^{-1}) = 1$

Random Walk: $D(q^{-1}) = 1 - q^{-1}$

Filtered Random Walk: $D(q^{-1}) = (1 - aq^{-1})(1 - q^{-1})$

Quasi Periodic (AR2): $D(q^{-1}) = 1 - 2\rho\cos\omega_0 q^{-1} + \rho^2 q^{-2}$

Kalman Estimators:

Model:

$$x_{t+1} = \mathbf{F} x_t + \mathbf{G} e_{t+1} \qquad \text{(assumed parameter dynamics)}$$
$$h_t = \mathbf{H} x_t$$
$$y_t = \varphi_t^* h_t + v_t = \varphi_t^* \mathbf{H} x_t + v_t \qquad \text{(linear regression)}$$

The Kalman estimator for scalar y_t is

$$\varepsilon_t = y_t - \varphi_t^* \hat{h}_{t|t-1}$$
$$\hat{x}_{t|t} = \mathbf{F} \hat{x}_{t-1|t-1} + \mathbf{K}_t \varphi_t \varepsilon_t \quad ; \quad \mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}^* / \sigma_{\varepsilon,t}^2$$
$$\hat{h}_{t+k|t} = \mathbf{H} \mathbf{F}^k \hat{x}_{t|t} \qquad k \geq 0 \ .$$

It can be expressed as a time-varying linear filtering of $\varphi_t \varepsilon_t$:

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

The General Constant Gain Structure:

Linear time-invariant filtering of the instantaneous gradient $\varphi_t \varepsilon_t$:

$$\begin{aligned} \varepsilon_t &= y_t - \varphi_t^* \hat{h}_{t|t-1} \\ \hat{h}_{t+k|t} &= \boldsymbol{\mathcal{M}}_k(q^{-1}) \varphi_t \varepsilon_t \end{aligned}$$

where $\boldsymbol{\mathcal{M}}_{k(q^{-1})}$ is optimized based on the model

$$\begin{aligned} y_t &= \varphi_t^* h_t + v_t \quad ; \ \mathrm{E}\, v_t v_t^* = \mathbf{R}_v \\ h_t &= \boldsymbol{\mathcal{H}}(q^{-1}) e_t \quad ; \ \mathrm{E}\, e_t e_t^* = \mathbf{R}_e \end{aligned}$$

"hypermodel"

to minimize

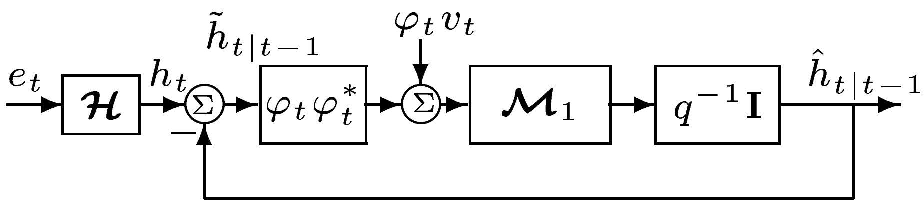
$$\mathbf{P}_k \triangleq \lim_{t\to\infty} \mathrm{E}\, \tilde{h}_{t+k|t} \tilde{h}_{t+k|t}^* \text{, where } \tilde{h}_{t+k|t} \triangleq h_{t+k} - \hat{h}_{t+k|t} \ .$$

(LMS: $\boldsymbol{\mathcal{M}}_k(q^{-1}) = \frac{\mu}{1-q^{-1}} \mathbf{I}$.)

$$\varphi_t \varepsilon_t = \varphi_t(y_t - \varphi_t^* \hat{h}_{t|t-1}) = \varphi_t \varphi_t^* \tilde{h}_{t|t-1} + \varphi_t v_t$$

$$\hat{h}_{t+1|t} = \boldsymbol{\mathcal{M}}_1(q^{-1})\varphi_t \varepsilon_t \qquad \text{(one-step predictor)}$$

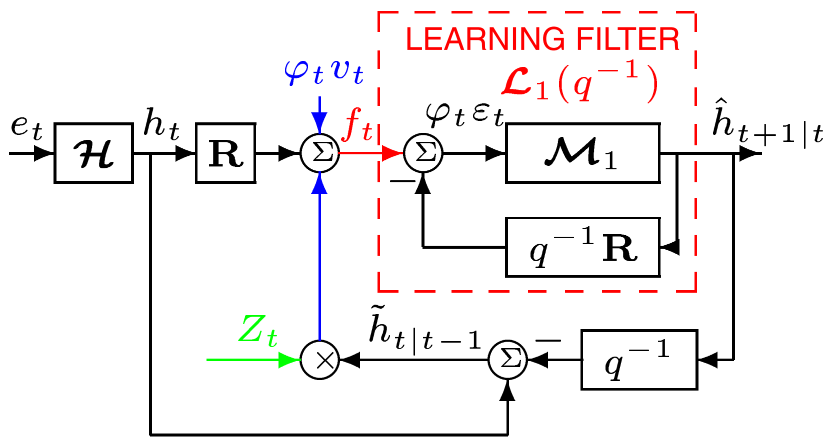
Can be seen as a time-invariant regulator for a time-varying system:



$$\mathbf{R} = \mathrm{E}\,(\varphi_t \varphi_t^*)$$

Add+subtract $\mathbf{R}\tilde{h}_{t|t-1}$: $\varphi_t \varepsilon_t = \mathbf{R}(h_t - \hat{h}_{t|t-1}) + (\varphi_t \varphi_t^* - \mathbf{R})\tilde{h}_{t|t-1} + \varphi_t v_t$.

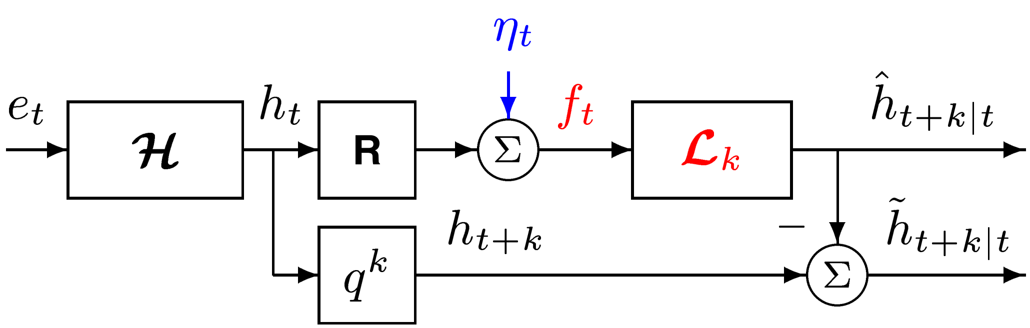
Define $Z_t = \varphi_t \varphi_t^* - \mathbf{R}$. Then, ...



The Learning Filter:

We design a stable rational matrix $\mathcal{L}_{k(q^{-1})}$ that estimates h_{t+k} **for any** k, by operating on the "fictitous measuement" f_t:

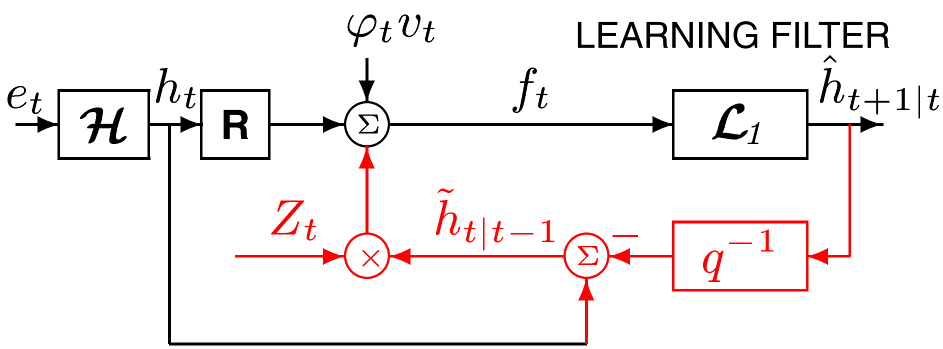
$$\begin{aligned} f_t &= \mathbf{R}\hat{h}_{t|t-1} + \varphi_t \varepsilon_t = \mathbf{R} h_t + \eta_t \\ \hat{h}_{t+k|t} &= \mathcal{L}_k(q^{-1}) f_t \ . \end{aligned}$$



$$\eta_t = \underbrace{Z_t \tilde{h}_{t|t-1}} + \varphi_t v_t$$ "gradient noise"

"feedback noise"

The Feedback Effect on the Gradient Noise



$$Z_t = \varphi_t \varphi_t^* - \mathbf{R}$$

- Stability by Small gain theorem whenever

$$\|q^{-1}\mathcal{L}_1 Z_t \tilde{h}_{t|t-1}\|_p \leq \gamma \|\tilde{h}_{t|t-1}\|_p \;\; ; \;\; \gamma < 1 \;\; .$$

- For "slow variations" (see below), $\varphi_t \varphi_t^* \tilde{h}_{t|t-1} \approx \mathbf{R}\tilde{h}_{t|t-1}$, so $Z_t \tilde{h}_{t|t-1} \approx 0$. The feedback loop can then be neglected.
- In other cases, we assume a low *correlation* (true for white FIR regressors).
- Then, an iterative open-loop Wiener design can be performed.

Insignificant Feedback Noise ⇔ Slow Variations

Degree of nonstationarity (Macchi):

$$\sqrt{\frac{\mathrm{E}\,\|\varphi_t^*(h_t - h_{t-1})\|_2^2}{\mathrm{E}\,|v_t|^2}} \quad . \tag{1}$$

We define regression parameters as slowly time-varying when the feedback noise $Z_t \tilde{h}_{t|t-1}$ can be neglected in an optimal MSE design without affecting the tracking error covariances significantly.

Lemma : Let the learning filter $\boldsymbol{\mathcal{L}}_k(q^{-1})$ be obtained by the Wiener design equations. If $\boldsymbol{\mathcal{H}}(z^{-1})$ is stable or marginally stable, then the relative impact of the feedback noise on the resulting true error will tend to zero as (1) vanishes.

Assumptions for the Wiener Design:

1. Regressors φ_t^* are stationary and known at t and $\mathbf{R}$ is known.
2. The gradient noise is decribed by a known and stable vector-ARMA model:

$$\eta_t = \frac{1}{N(q^{-1})}\boldsymbol{M}(q^{-1})\nu_t \quad .$$

3. Innovation sequence ν_t is uncorrelated with h_{t-i} and with $\hat{h}_{t-i|t-i-1}$, $i \geq 0$
4. We assume

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

where e_t is white, with $\mathrm{E}\, e_t = 0$ and $\mathrm{E}\,[e_t e_t^*] = \mathbf{R}_e$ is nonsingular.

$$\begin{aligned} \boldsymbol{D}(q^{-1}) &= D_u(q^{-1})\boldsymbol{D}_s(q^{-1}) \quad ; \quad D_u \text{ polynomial with zeros on} |z| = 1 \\ &= \mathbf{I} + \mathbf{D}_1 q^{-1} + \ldots \mathbf{D}_{n_D} q^{-n_D} \quad \text{(Marginally stable)} \\ \boldsymbol{C}(q^{-1}) &= \mathbf{I} + \mathbf{C}_1 q^{-1} + \ldots \mathbf{C}_{n_C} q^{-n_C} \quad \text{(Stable)} \end{aligned}$$

Wiener Design of the Learning Filter:

Under Assumptions 1-4, the stable and causal learning filter minimizing $\mathbf{P}_k$ is

$$\boldsymbol{\mathcal{L}}_k^{opt} = \boldsymbol{D}_s^{-1}\boldsymbol{Q}_k\boldsymbol{\beta}^{-1}N\boldsymbol{D}_s\mathbf{R}^{-1} \quad ,$$

given by the spectral factorization

$$\boldsymbol{\beta}\boldsymbol{\beta}_* = \boldsymbol{C}\,\mathbf{R}_e\boldsymbol{C}_*NN_* + \boldsymbol{D}\,\mathbf{R}^{-1}\boldsymbol{M}\boldsymbol{M}_*\mathbf{R}^{-1}\boldsymbol{D}_* \quad ,$$

and the bilateral Diophantine equation

$$q^k\boldsymbol{C}\,\mathbf{R}_e\boldsymbol{C}_*N_* = \boldsymbol{Q}_k\boldsymbol{\beta}_* + q\boldsymbol{D}\,\boldsymbol{L}_{k*} \quad .$$

(Here X_* denote conjugated matrices in q). The solution is unique.

The error $\tilde{h}_{t+k|t}$ is stationary, with finite covariance matrix and zero mean.

$$\hat{h}_{t+k|t} = \boldsymbol{\mathcal{M}}_k(q^{-1})\varphi_t\varepsilon_t = \boldsymbol{D}_s^{-1}\boldsymbol{Q}_k\left[\boldsymbol{\beta} - q^{-1}N\boldsymbol{Q}_1\right]^{-1}N\boldsymbol{D}_s\mathbf{R}^{-1}\varphi_t\varepsilon_t$$

Wiener LMS (WLMS):

We may minimize $\mathrm{tr}\,\mathbf{P}_k = \lim_{t\to\infty} \mathrm{E} \sum_{i=0}^{n_h-1} |h_{i,t+k} - \hat{h}_{i,t+k|t}|^2$
for diagonal hypermodels with equal elements

$$h_t = \frac{C(q^{-1})}{D(q^{-1})}\mathbf{I}\, e_t$$

with a structurally constrained learning filter

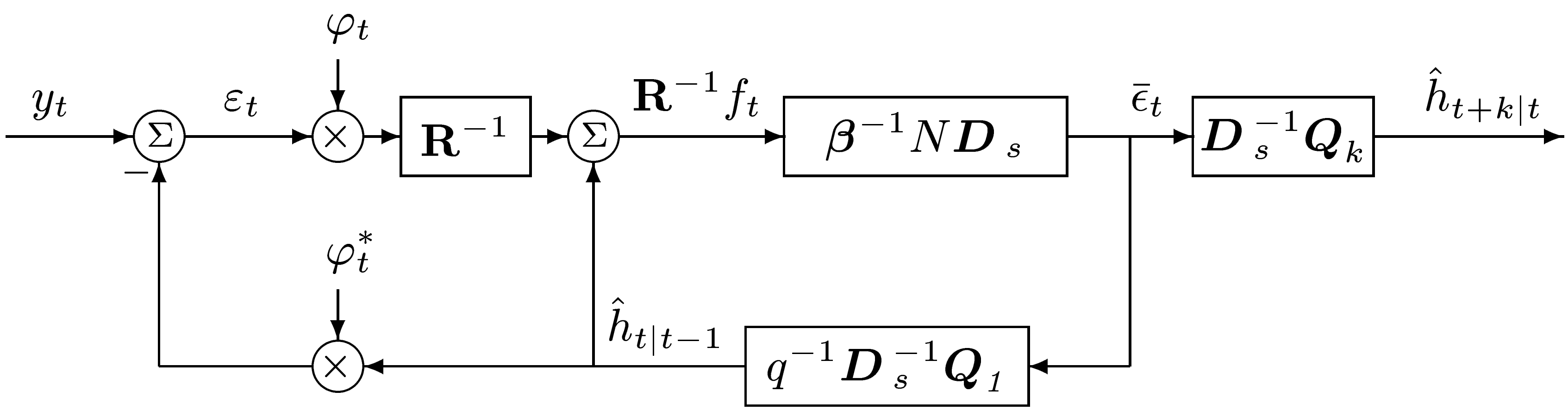
$$\hat{h}_{t+k|t} = \frac{Q_k(q^{-1})}{\beta(q^{-1})}\mathbf{R}^{-1} f_t \quad .$$

If η_t is **white** with covariance $\mathbf{R}_\eta$ and $\gamma \triangleq \mathrm{tr}\,\mathbf{R}_e/\mathrm{tr}\,\mathbf{R}^{-1}\mathbf{R}_\eta\mathbf{R}^{-1}$, solve

$$\begin{aligned} r\beta\beta_* &= \gamma CC_* + DD_* \\ q^k\gamma CC_* &= rQ_k\beta_* + qDL_{k*} \quad . \end{aligned}$$

For random walk models and white regressors $(\mathbf{R} = c\mathbf{I})$, WLMS reduces to LMS.

Realization of Constant Gain Algorithms:

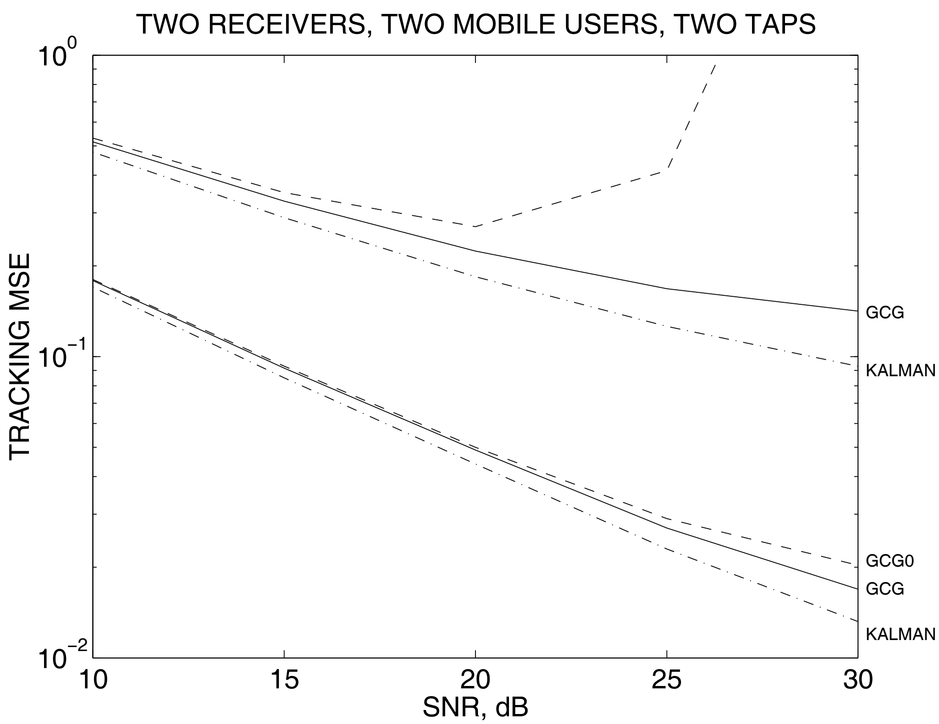


Numerically well-behaved. All blocks are internally stable.

(Low complexity computation of $\mathbf{R}^{-1}\varphi_t^*$ when regressors are autoregressive:
See Farhang-Boroujeny, IEEE SP pp1987-2000 1997.)

Iterative Design

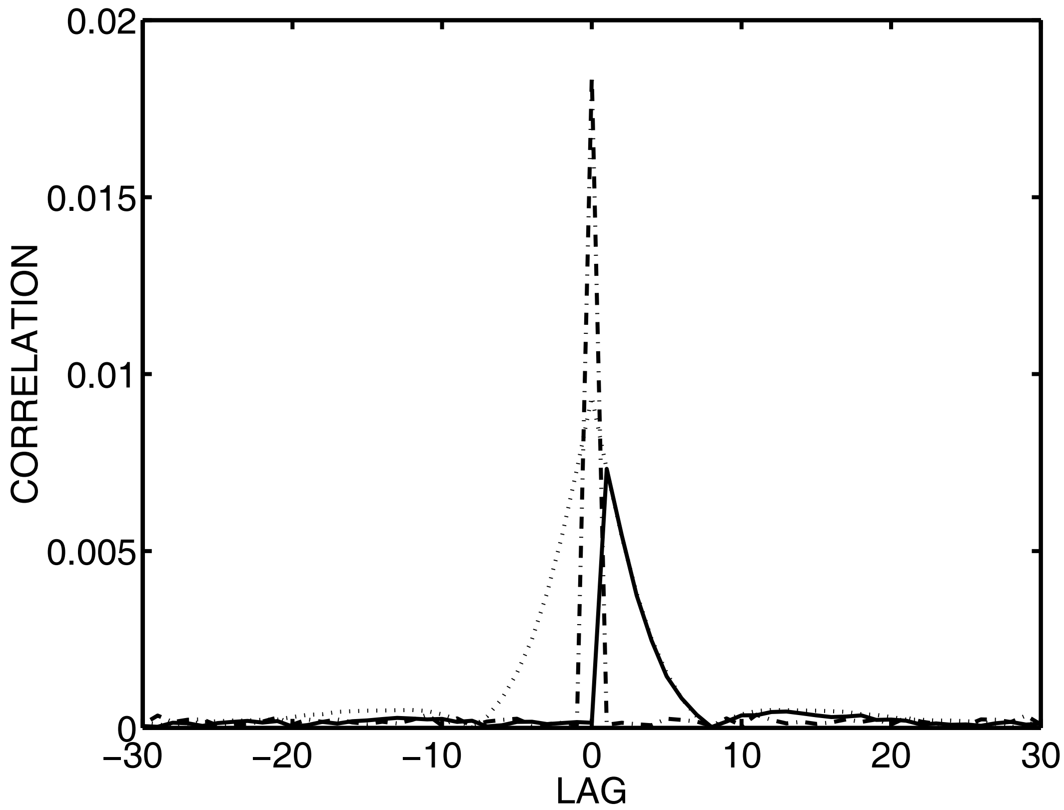
1. Assume $\eta_t = \varphi_t v_t$ (slow variations) and design $\mathcal{L}_1$.
2. Estimate covariance matrix of the gradient noise η_t by theory or simulation.
3. Re-design $\mathcal{L}_1$ if required. Otherwise, obtain the desired $\mathcal{L}_k$.



Example: MIMO D-AMPS with one fast mobile (225km/h) and one slow (45km/h) in upper curves.

Dashed curves: iteration 1 (assuming slow variations).
Solid: after iterations.

Iterative Design Example: Correlations



Example: MIMO D-AMPS

Solid: $\mathrm{E}\left(\eta_t \tilde{h}^*_{t+\tau}\right)$

(Small for $\tau \leq 0$).

Dash-dotted: $\mathrm{E}\left(\eta_t \eta^*_{t+\tau}\right)$

Dotted:

$\mathrm{E}\left(\tilde{h}_{t|t-1} \tilde{h}^*_{t+\tau|t+\tau-1}\right)$

Iterative Design Example: Performance

SNR	$\omega_{\text{D},2}T$	Kalm.	GCG	WLMS	RLS	LMS
10	0.10	0.477	**0.516**	1.045	1.43	1.58
30	0.10	0.093	**0.142**	0.488	0.82	1.00
10	0.02	0.170	**0.179**	0.247	0.33	0.413
30	0.02	0.013	**0.017**	0.028	0.077	0.115
	# real. mult.	5440	**416**	272	1564	132

Tracking of 8 parameters. Second order oscillative hypermodels, known and diagonal.

$\mathbf{R}_e$ is 2×2 block diagonal.

4-step predictors are calculated. (one-step predictions used in RLS and LMS).

Kalman predictors estimate 16 complex-valued states.

Iterative Design Example: Modelling

$$\begin{pmatrix} y_t^1 \\ y_t^2 \end{pmatrix} = \begin{pmatrix} B_t^{11}(q^{-1}) & B_t^{12}(q^{-1}) \\ B_t^{21}(q^{-1}) & B_t^{22}(q^{-1}) \end{pmatrix} \begin{pmatrix} u_t^1 \\ u_t^2 \end{pmatrix} + \begin{pmatrix} v_t^1 \\ v_t^2 \end{pmatrix}$$

where y_t^i is the sampled baseband signal at receiver i. Two-tap channels: $B_t^{ij}(q^{-1}) = b_{0,t}^{ij} + b_{1,t}^{ij} q^{-1}$.

$$\varphi_t^* = \begin{pmatrix} u_t^1 & u_{t-1}^1 & u_t^2 & u_{t-1}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & u_t^1 & u_{t-1}^1 & u_t^2 & u_{t-1}^2 \end{pmatrix}$$

$$h_t = (\, b_{0,t}^{11} \; b_{1,t}^{11} \; b_{0,t}^{12} \; b_{1,t}^{12} \; b_{0,t}^{21} \; b_{1,t}^{21} \; b_{0,t}^{22} \; b_{1,t}^{22} \,)^T \; ; \; v_t = [v_t^1 \; v_t^2]^T \text{ white.}$$

$\{u_t^i\}$ are white complex-valued QPSK symbols with $\mathbf{R} = \mathbf{I}_8$.

Fading model $\boldsymbol{D}(q^{-1})h_t = e_t$, where $\mathbf{R}_e$ is 2×2 block diagonal.

$$\boldsymbol{D}(q^{-1}) = \mathrm{diag}[\, \boldsymbol{D}_{11}(q^{-1}) \; \boldsymbol{D}_{12}(q^{-1}) \; \boldsymbol{D}_{21}(q^{-1}) \; \boldsymbol{D}_{22}(q^{-1}) \,]$$

$$\boldsymbol{D}_{ij}(q^{-1}) = [1 - 2\rho\cos(\omega_{\mathrm{D},j}T/\sqrt{2})q^{-1} + \rho^2 q^{-2}]\mathbf{I}_2$$

ANALYSIS AND APPLICATION OF ADAPTATION ALGORITHMS WITH TIME-INVARIANT GAINS

Mikael Sternad*

Outline of joint work with Lars Lindbom[1] and Anders Ahlén*

*Signals and Systems Group, Uppsala University

[1] Ericsson Infotech, Karlstad

(See www.signal.uu.se/Publications/abstracts/r001.html - r004.html)

UPPSALA UNIVERSITET

Outline of Presentation:

1. Design (Seminar 1):
   - Structure of the tracking algorithms
   - Wiener design
   - Simplified algorithms.
   - Iterative design.
2. Analysis:
   - Slow variations
   - Fast variations.
3. D-AMPS 1900 Channel Tracking
4. Summary

Time-Varying Linear Regression Models

$$y_t \;=\; \varphi_t^* h_t + v_t$$

Complex-valued and possibly MIMO, with φ_t^* known at time t.

Example: Mobil radio channel

$$y_t \;=\; (u_t \ldots u_{t-M+1}) \begin{pmatrix} h_{0,t} \\ \vdots \\ h_{M-1,t} \end{pmatrix} + v_t$$

Our goal: Estimate vector h_t when $\mathbf{R} = \mathrm{E}\left(\varphi_t \varphi_t^*\right)$ is known.

Approaches to analysis of Adaptation Algorithms:

$$
\begin{aligned}
y_t &= \varphi_t^* h_t + v_t && \text{System} \\
\varepsilon_t &= y_t - \varphi_t^* \hat{h}_{t|t-1} && \text{Prediction error} \\
\hat{h}_{t+1|t} &= f(\varepsilon_t) && \text{Adaptation Law}
\end{aligned}
$$

1. Time-varying systems, products of matrices

   Ewada and Macchi (Automatica 1985, AC 1986), Farden (ASSP 1981),

   Guo and Ljung (AC 1995), Moustakides (IJACSP 1998).

2. Slowly varying parameters and low adaptation gain

   Benveniste et.al. 1990, Kushner and Shwartz (IT 1984), Haykin 1996, Macchi 1995.

3. Independent consecutive regression vectors

   Widrow et al (Proc. IEEE 1976, IT 1984), Gardner (1984, 1987) Haykin 1996.

The General Constant Gain Structure:

Linear time-invariant filtering of the instantaneous gradient $\varphi_t \varepsilon_t$:

$$\begin{aligned} \varepsilon_t &= y_t - \varphi_t^* \hat{h}_{t|t-1} \\ \hat{h}_{t+k|t} &= \boldsymbol{\mathcal{M}}_k(q^{-1}) \varphi_t \varepsilon_t \end{aligned}$$

where $\boldsymbol{\mathcal{M}}_{k(q^{-1})}$ is optimized based on the model

$$\begin{aligned} y_t &= \varphi_t^* h_t + v_t \quad ; \ \mathrm{E}\, v_t v_t^* = \mathbf{R}_v \\ h_t &= \boldsymbol{\mathcal{H}}(q^{-1}) e_t \quad ; \ \mathrm{E}\, e_t e_t^* = \mathbf{R}_e \end{aligned}$$ “hypermodel”

to minimize

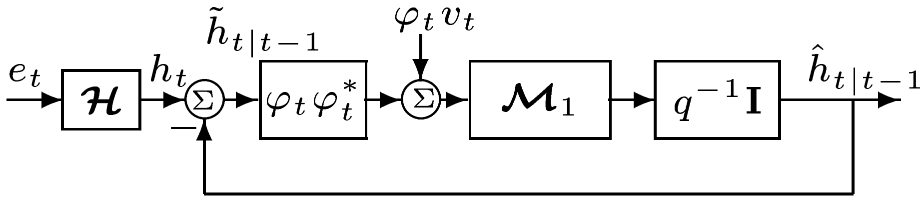
$$\mathbf{P}_k \triangleq \lim_{t \to \infty} \mathrm{E}\, \tilde{h}_{t+k|t} \tilde{h}_{t+k|t}^* \text{, where } \tilde{h}_{t+k|t} \triangleq h_{t+k} - \hat{h}_{t+k|t} \ .$$

(LMS: $\boldsymbol{\mathcal{M}}_k(q^{-1}) = \frac{\mu}{1-q^{-1}} \mathbf{I}$.)

$$\varphi_t \varepsilon_t = \varphi_t(y_t - \varphi_t^* \hat{h}_{t|t-1}) = \varphi_t \varphi_t^* \tilde{h}_{t|t-1} + \varphi_t v_t$$

$$\hat{h}_{t+1|t} = \boldsymbol{\mathcal{M}}_1(q^{-1})\varphi_t \varepsilon_t \qquad \text{(one-step predictor)}$$

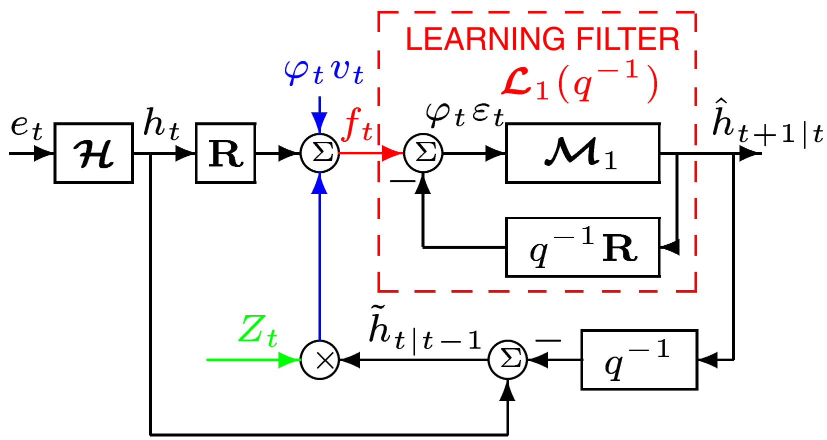
Can be seen as a time-invariant regulator for a time-varying system:



$$\mathbf{R} = \mathrm{E}\,(\varphi_t \varphi_t^*)$$

Add+subtract $\mathbf{R}\tilde{h}_{t|t-1}$: $\varphi_t \varepsilon_t = \mathbf{R}(h_t - \hat{h}_{t|t-1}) + (\varphi_t \varphi_t^* - \mathbf{R})\tilde{h}_{t|t-1} + \varphi_t v_t.$

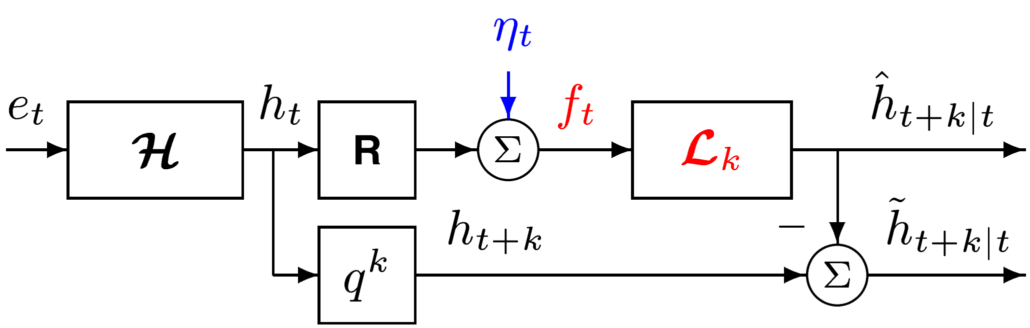
Define $Z_t = \varphi_t \varphi_t^* - \mathbf{R}$. Then, ...



The Learning Filter:

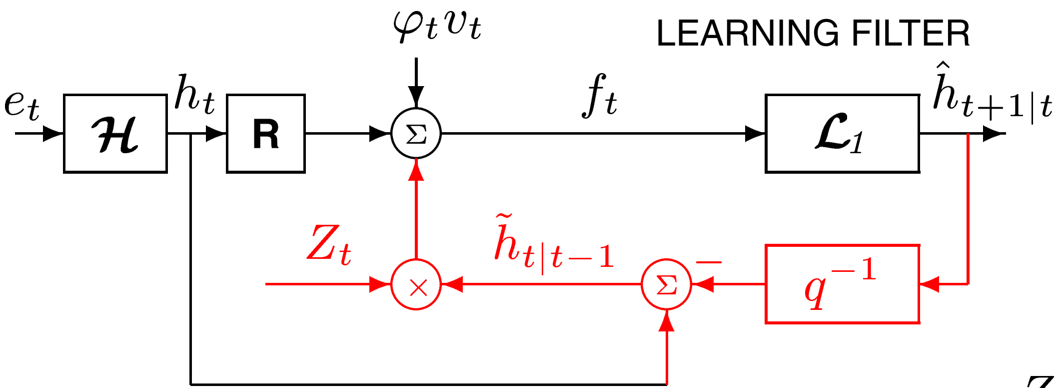
We design a stable rational matrix $\mathcal{L}_{k(q^{-1})}$ that estimates h_{t+k} **for any** k, by operating on the "fictitous measuement" f_t:

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



$$\eta_t = \underbrace{Z_t \tilde{h}_{t|t-1}}_{\text{"feedback noise"}} + \varphi_t v_t \qquad \text{"gradient noise"}$$

Analysis of Adaptation Laws with Constant Gains



$$Z_t = \varphi_t \varphi_t^* - \mathbf{R}$$

- When can the feedback loop be neglected?
- How to quantify feedback effects?
- Less conservative stability conditions than Small gain theorem?

Basic Assumtions of our Analysis

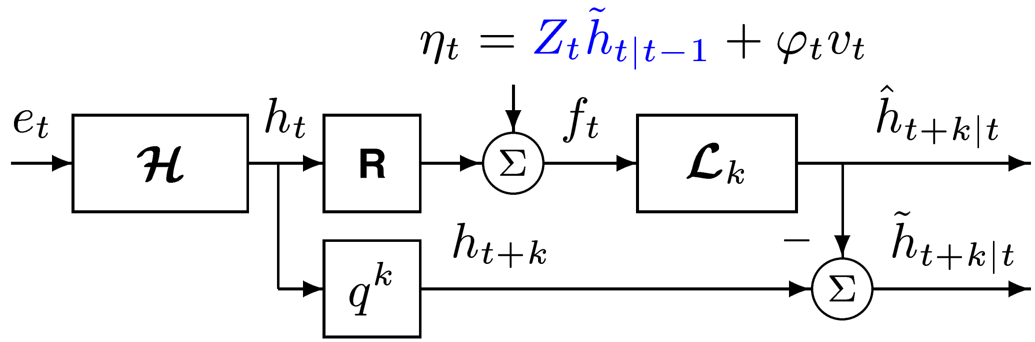
Assumption 1:

- The noise v_t is stationary and zero mean
- The regressor matrix φ_t^* is stationary with zero first and third order moments and finite higher order moments.
- The parameter model $\boldsymbol{\mathcal{H}}(q^{-1})$ is stable or marginally stable.
- The parameter innovations e_t are stationary, white and zero mean.
- e_t, v_t, and φ_t^* are mutually independent with bounded covariance matrices

$$\mathbf{R} \triangleq \mathrm{E}\,\varphi_t\varphi_t^* \;\; ; \;\; \mathbf{R}_e \triangleq \mathrm{E}\,e_t e_t^* \;\; ; \;\; \mathbf{R}_v \triangleq \mathrm{E}\,v_t v_t^*$$

respectively, with $\mathbf{R}$ being nonsingular.

The Estimation Error



$$\tilde{h}_{t+k|t} = \underbrace{(\mathbf{I} - q^{-k}\mathcal{L}_k\mathbf{R})h_{t+k}}_{\text{Lag Error}} - \underbrace{\mathcal{L}_k\varphi_t v_t}_{\text{Noise}} - \underbrace{\mathcal{L}_k Z_t \tilde{h}_{t|t-1}}_{\text{Feedback Effects}} \quad .$$

$$\mathbf{P}_k = \lim_{t\to\infty}\left(\mathbf{V}^k_{h,t} + \mathbf{V}^k_{\varphi v,t} + \mathbf{V}^k_{Z\tilde{h},t} + \underbrace{\mathbf{V}^k_{hZ\tilde{h},t} + \mathbf{V}^k_{\varphi v Z\tilde{h},t}}_{\text{Cross-terms}}\right)$$

Slow Variations 1

Degree of nonstationarity (Macchi):

$$\sqrt{\frac{\mathrm{E}\,\|\varphi_t^*(h_t - h_{t-1})\|_2^2}{\mathrm{E}\,|v_t|^2}} \quad . \tag{1}$$

We define regression parameters as slowly time-varying when the feedback noise $Z_t\tilde{h}_{t|t-1}$ can be neglected in an optimal MSE design without affecting the tracking error covariances significantly.

Lemma : Let the learning filter $\boldsymbol{\mathcal{L}}_k(q^{-1})$ be obtained by the Wiener design equations. Under Assumption 1, the relative impact of the feedback noise on the resulting true error will then tend to zero as (1) vanishes.

Slow Variations 2

Analysis for slow variations now becomes simple!

Steady-state error covariances: $(h_t = (1/D)\boldsymbol{C}\, e_t)$.

$$\mathbf{P}_k = \lim_{t\to\infty} (\, \mathbf{V}^k_{h,t} + \mathbf{V}^k_{\varphi v,t} \,)$$

$$= \frac{1}{2\pi j} \oint \left(\mathbf{I} - z^{-k}\boldsymbol{\mathcal{L}}_k \mathbf{R}\right) \frac{\boldsymbol{C}\,\mathbf{R}_e \boldsymbol{C}_*}{DD_*} \left(\mathbf{I} - z^{k}\boldsymbol{\mathcal{L}}_{k*} \mathbf{R}\right) \frac{dz}{z}$$

$$+ \frac{1}{2\pi j} \oint \boldsymbol{\mathcal{L}}_k \frac{\boldsymbol{M}\mathbf{R}_\nu \boldsymbol{M}_*}{NN_*} \boldsymbol{\mathcal{L}}_{k*} \frac{dz}{z}$$

Lag error gives finite contribution whenever

$$\mathbf{I} - q^{-k}\boldsymbol{\mathcal{L}}_k \mathbf{R}$$

contains all marginally stable factors of $D(q^{-1})$ in all numerators.

Slow Variations: LMS

Stability and bounded estimation errors are for stable $\mathcal{H}(q^{-1})$ assured by stability of the learning filter. For LMS,

$$(1 - q^{-1})\hat{h}_{t+1|t} = \mu\varphi_t\varepsilon_t \; = \; \mu(f_t - \mathbf{R}\hat{h}_{t|t-1}) \quad \Rightarrow$$

$$\hat{h}_{t+1|t} \; = \; \mathcal{L}_{1}(q^{-1}) f_t = (\mathbf{I} - (\mathbf{I} - \mu\mathbf{R})q^{-1})^{-1}\mu f_t$$

Let $\lambda_{\max}$ be the largest eigenvalue of $\mathbf{R}$. If $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$,

$$\hat{h}_{t+1|t} = \mathbf{U}(\mathbf{I} - (\mathbf{I} - \mu\mathbf{\Lambda})q^{-1})^{-1}\mathbf{U}^*\mu \, f_t \; .$$

Stability of $\mathcal{L}_1$: The classical condition for convergence in the mean

$$0 \; < \; \mu \; < \; \frac{2}{\lambda_{\max}}$$

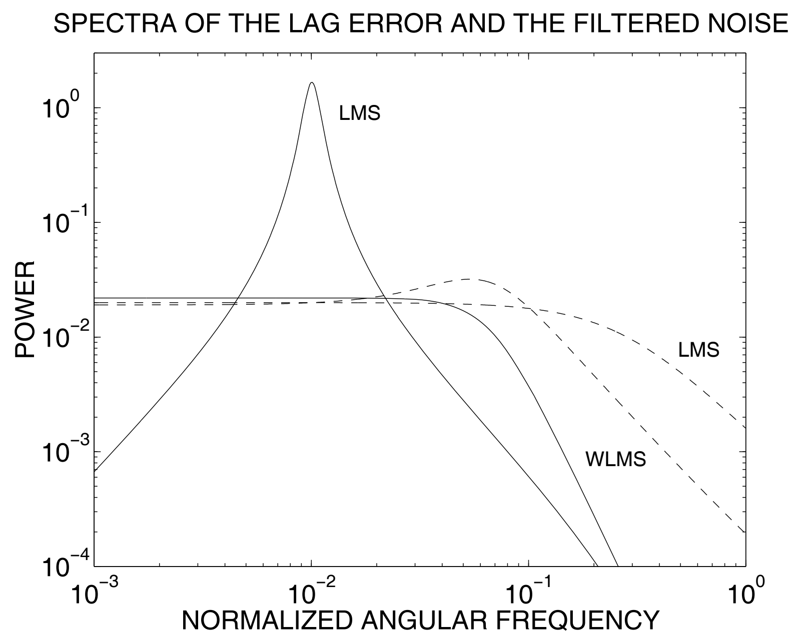
Slow Variations: Example.

$$y_t = h_{0,t}u_t + h_{1,t}u_{t-1} + v_t \quad ; \quad h_t = 2p\cos\omega_o h_{t-1} - p^2 h_{t-2} + e_t$$

$p = 0.999$	ω_o	0.001	0.005	0.01	0.02	0.10
DNS:	(1)	.0141	.0510	.1005	.2002	.9996
LMS:	$\mathrm{tr}\,\mathbf{P}_1$	**.0011**	**.0027**	**.0045**	**.0075**	**.0360**
Measured:		*.0012*	*.0030*	*.0052*	*.0099*	*.0650*
	$\mathrm{tr}\,\mathbf{V}^1_{Z\tilde{h}}$	*.0001*	*.0003*	*.0007*	*.0020*	*.0278*
WIENER	$\mathrm{tr}\,\mathbf{P}_1$	**.0007**	**.0013**	**.0019**	**.0028**	**.0061**
DESIGN:		*.0007*	*.0014*	*.0021*	*.0031*	*.0076*
	$\mathrm{tr}\,\mathbf{V}^1_{Z\tilde{h}}$	*.0000*	*.0001*	*.0002*	*.0003*	*.0015*

Slow Variations: Example.

$$\mathbf{P}_k = \frac{1}{2\pi j}\oint \left(\mathbf{I} - z^{-k}\boldsymbol{\mathcal{L}}_k\mathbf{R}\right)\frac{\boldsymbol{C}\,\mathbf{R}_e\boldsymbol{C}_*}{DD_*}\left(\mathbf{I} - z^{k}\boldsymbol{\mathcal{L}}_{k*}\mathbf{R}\right)\frac{dz}{z} + \frac{1}{2\pi j}\oint \boldsymbol{\mathcal{L}}_k\frac{\boldsymbol{M}\mathbf{R}_\nu\boldsymbol{M}_*}{NN_*}\boldsymbol{\mathcal{L}}_{k*}\frac{dz}{z}$$



Lag error (solid) and the filtered noise (dashed), which equals $0.01|\boldsymbol{\mathcal{L}}_1(\omega)|^2$, for Wiener estimators (WLMS) and for LMS, with $\omega_o = 0.01$.

FIR Models with Rapid Parameter Variations 1.

Scalar FIR model with white inputs:

$$y_t = h_{0,t}u_t + h_{1,t}u_{t-1} + \ldots + h_{\mathrm{M}-1,t}u_{t-\mathrm{M}+1} + v_t$$

Approximation 1:

$$\mathrm{tr}\,\mathrm{E}\,Z_\tau^* Z_t \tilde{h}_{t|t-1}\tilde{h}^*_{\tau|\tau-1} = \mathrm{tr}\,\mathrm{E}\,[Z_\tau^* Z_t]\mathrm{E}\,[\tilde{h}_{t|t-1}\tilde{h}^*_{\tau|\tau-1}] \quad . \qquad (2)$$

Approximation 2: $Z_t\tilde{h}_{t|t-1}$ is uncorrelated with $\varphi_\tau v_\tau$ and $h_\tau, \forall\tau$.

(Independence between Z_t and $\tilde{h}_{t|t-1}$ would imply (2), but would be a much stronger assumption. Under Approximation 2, the cross-terms are neglected.)

A WLMS tracking structure which gives a finite lag error is assumed:

$$\boldsymbol{\mathcal{L}}_k(q^{-1}) = \frac{Q_k(q^{-1})}{\beta(q^{-1})}\frac{1}{\sigma_u^2}\mathbf{I} = \sum_{i=0}^{\infty} \mathrm{L}_i^k \mathbf{I}\, q^{-i} \quad .$$

FIR Models with Rapid Parameter Variations 2.

Result, under Assumption 1 and the above assumtions:
A finite steady state mean square parameter error exists if and only if

$$\mathcal{G}(z^{-1}) = \frac{1}{1 - m\sigma_u^4 \sum_{i=0}^{\infty} |\mathrm{L}_i^1|^2 z^{-i-1}} \tag{3}$$

is stable, where

$$m \triangleq \underbrace{\frac{\mathrm{E}\,|u_t|^4}{(\mathrm{E}\,|u_t|^2)^2}}_{\kappa_u\text{, Pearson kurtosis.}} + \mathsf{M} - 2 \tag{4}$$

The k-step estimation error is then given by

$$\mathrm{tr}\,\mathbf{P}_k = \mathrm{tr}\,\mathbf{V}_h^k + \mathrm{tr}\,\mathbf{V}_{\varphi v}^k + \mathrm{tr}\,\mathbf{V}_{Z\tilde{h}}^k$$

FIR Models with Rapid Parameter Variations 3.

where

$$\operatorname{tr} \mathbf{V}_h^k = \left\| \frac{\beta(q^{-1}) - q^{-k} Q_k(q^{-1})}{\beta(q^{-1})} \, h_{t+k} \right\|_2^2$$

$$\operatorname{tr} \mathbf{V}_{\varphi v}^k = \mathsf{M} \frac{\sigma_v^2}{\sigma_u^2} \Sigma_k$$

$$\operatorname{tr} \mathbf{V}_{Z\tilde{h}}^k = m \operatorname{tr} \mathbf{P}_1 \Sigma_k$$

in which

$$\Sigma_k \triangleq \frac{1}{2\pi \jmath} \oint_{|z|=1} \left| \frac{Q_k(z^{-1})}{\beta(z^{-1})} \right|^2 \frac{dz}{z} \quad ,$$

$$\operatorname{tr} \mathbf{P}_1 = \frac{\operatorname{tr} \mathbf{V}_h^1 + \mathsf{M} \frac{\sigma_v^2}{\sigma_u^2} \Sigma_1}{1 - m\Sigma_1} \quad .$$

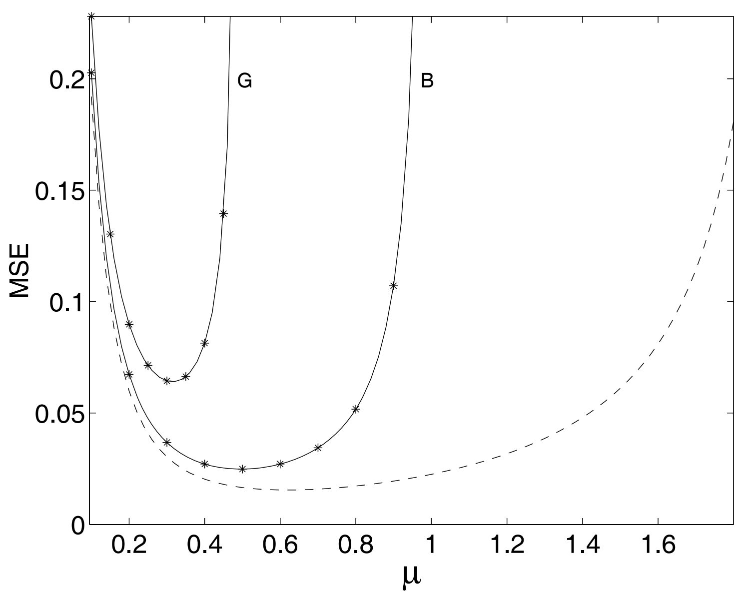
Here, $\sigma_u^2 = E|u_t^2|$ and $\sigma_v^2 = E|v_t^2|$.

LMS Example

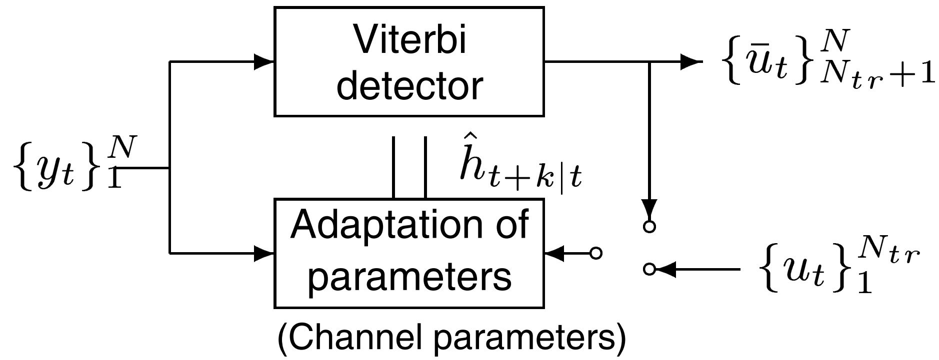
FIR system with

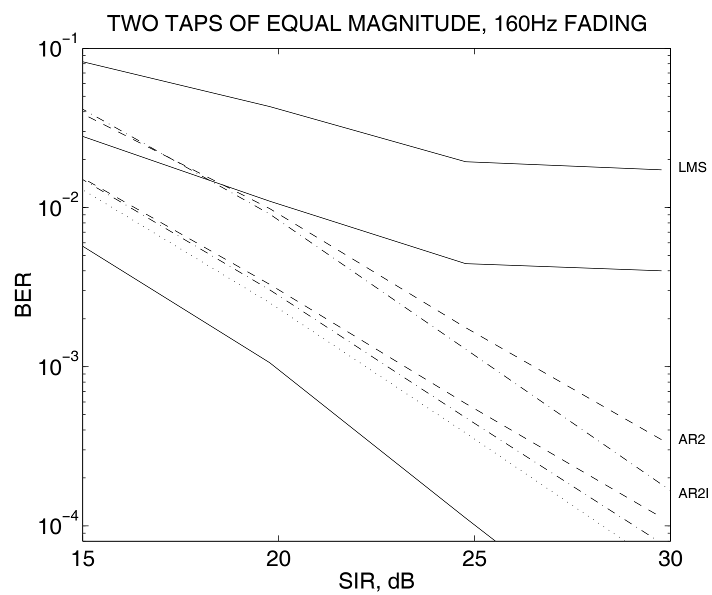
$$h_t = 2p \cos \omega_o h_{t-1} - p^2 h_{t-2} + e_t \; ; \; \omega_o = 0.050 \, , \; p = 0.995 \, .$$

Output SNR of 20 dB, with $|h_t|^2 = 1$. Tracking MSE for two-tap system by theory (solid) and by simulation $(*)$. Two-tap FIR systems with white real-valued binary (B) and Gaussian (G) regressors. Dashed curve neglects the feedback noise.



Adaptive Channel Tracking in D-AMPS





Summary

- A novel formalism for analysis and design of adaptive algorithms for linear regression models.
- Level of design complexity and computational complexity is controlled by selecting models for the parameters h_t and the gradient noise η_t.
- The WLMS principle is standard in all D-AMPS 1900 handsets and base stations by Ericsson. Will also be of use in EDGE.
- FIR systems with white regressors can be analyzed under approximations that are much milder than an assumption of independent regression vectors.
- An exact tracking analysis for fast variations with colored regressors might require considerably more complicated tools.