Adaptive Feedback Cancellation in a Closed-Loop System Identification Framework

Toon van Waterschoot* (ESAT-SCD/SISTA)

SISTA Seminar 19/02/2004

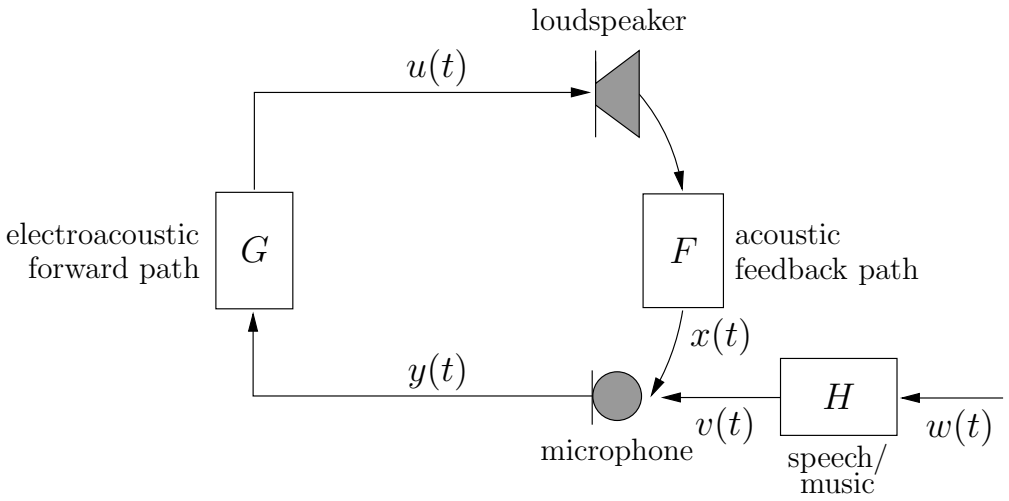
*in collaboration with Geert Rombouts, Ann Spriet and Marc Moonen (promotor)

SISTA

Overview

- Acoustic Feedback Problem
- Prediction Error Method for Closed-Loop System Identification
- Time-varying Autoregressive Noise model
- Simulation Results
- Conclusions

The Acoustic Feedback Problem



System:
$$\begin{cases} y(t) &= F(q)u(t) + v(t) \\ u(t) &= G(q)y(t) \\ v(t) &= H(q)w(t) \end{cases}$$

Closed-Loop Transfer Function:
$$u(t) = \frac{G(q)}{1 - G(q)F(q)}v(t)$$

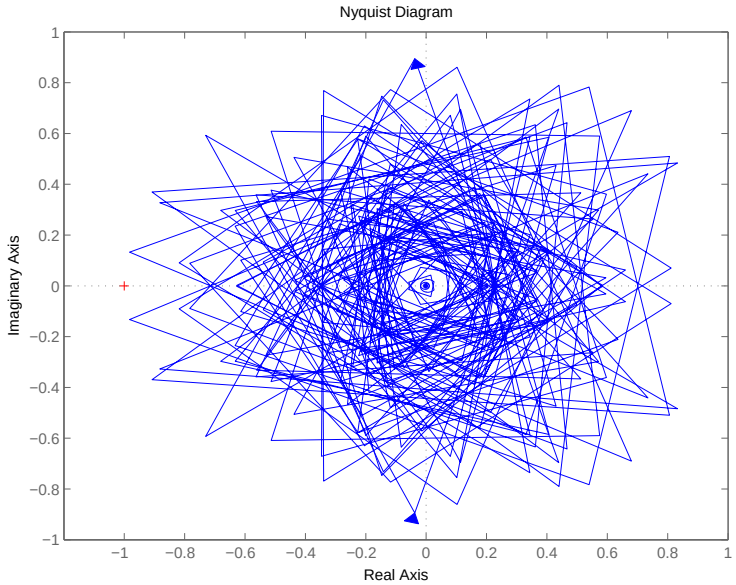
Goal:
$$u(t) = G(q)v(t)$$

Nyquist stability

- Nyquist: instability occurs if $\exists\ \omega$:

$$\begin{cases} |G(e^{i\omega})F(e^{i\omega})| & \geq & 1 \\ \angle G(e^{i\omega})F(e^{i\omega}) & = & n2\pi, n \in \mathbb{Z}. \end{cases}$$

- instability leads to howling



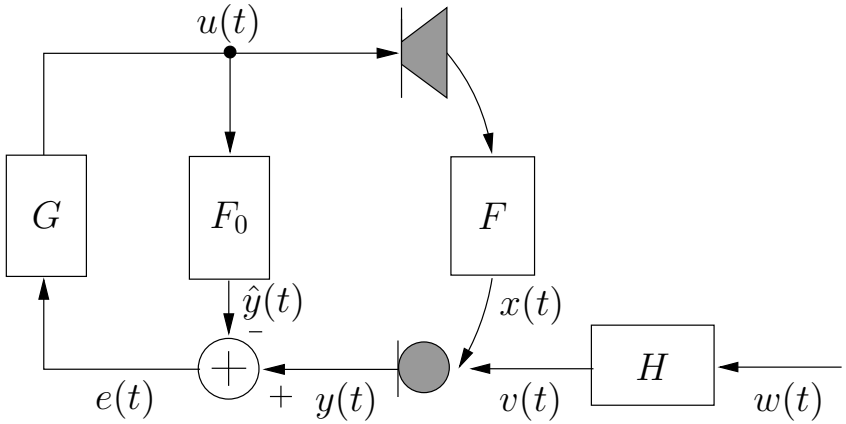
- system operation close to instability leads to ringing and excessive reverberation (typically avoided by choosing gain margin $GM > 2dB$)
- consequence: restriction on forward path amplification,

  e.g. when $G(q) = Kq^{-d}$ choose $K \leq \dfrac{10^{-\frac{GM}{20}}}{\max\limits_{\omega} |F(e^{i\omega})|}$

Feedback Cancellation: Idea

Include cancellation filter $F_0(q)$:

$$\frac{U(e^{j\omega})}{V(e^{j\omega})} = \frac{G(e^{j\omega})}{1 - G(e^{j\omega})(F(e^{j\omega}) - F_0(e^{j\omega}))}$$

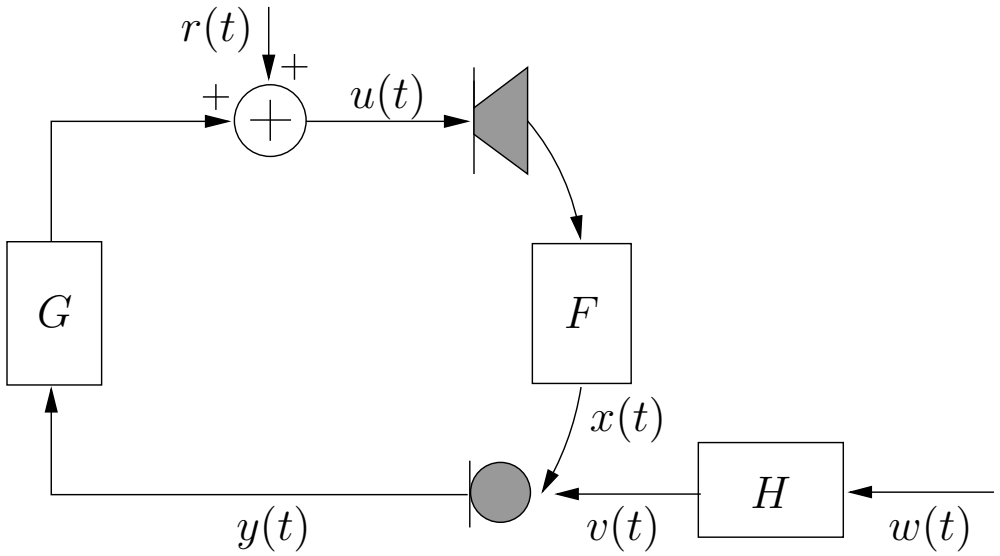


- to increase gain margin
  - $\longrightarrow$ $|F_0(e^{j\omega})|$ should model peaks of $|F(e^{j\omega})|$
  - $\longrightarrow$ fixed cancellation filter
- to reduce signal distortion
  - $\longrightarrow$ $F_0(q)$ should be a consistent estimate of $F(q)$
  - $\longrightarrow$ adaptive cancellation filter

Estimate $F(q)$ using the Prediction Error Method

- direct approach (no probe signal required): identify $F(q)$ from

$$y(t) = F(q)u(t) + H(q)w(t)$$



- indirect approach (injection of probe signal $r(t)$ and knowledge of forward path $G(q)$ required):

  with $F^c(q) = \frac{F(q)}{1-G(q)F(q)}$ and $H^c(q) = \frac{H(q)}{1-G(q)F(q)}$, identify $F^c(q)$ from

  $$y(t) = F^c(q)r(t) + H^c(q)w(t)$$

  and calculate $F(q)$ as $F(q) = F^c(q)(1 + F^c(q)G(q))^{-1}$

- joint input-output approach (injection of probe signal $r(t)$ is usual but not required):

  with $F^c(q) = \frac{F(q)}{1-G(q)F(q)}$, $H^c(q) = \frac{H(q)}{1-G(q)F(q)}$ and $S(q) = \frac{1}{1-G(q)F(q)}$,

  identify $\begin{bmatrix} F^c(q) \\ S(q) \end{bmatrix}$ from

  $$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} F^c(q) \\ S(q) \end{bmatrix} r(t) + \begin{bmatrix} H^c(q) \\ G(q)H^c(q) \end{bmatrix} w(t)$$

  and calculate $F(q)$ as $F(q) = F^c(q)(S(q))^{-1}$

Prediction Error Method: Direct approach (1)

Assume FIR model structure $\hat{F}(q) = \hat{f}_0(t) + \hat{f}_1(t)q^{-1} + \ldots + \hat{f}_{n_{\hat{F}}}(t)q^{-n_{\hat{F}}}$ and sufficient order case ($n_{\hat{F}} = n_F$).

$$
\begin{aligned}
\hat{\mathbf{f}}_{\mathbf{LS}}(t) &= \arg\min_{\hat{\mathbf{f}}} \| \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(1) \end{bmatrix} - \begin{bmatrix} u(t) & \ldots & u(t-n_{\hat{F}}) \\ u(t-1) & \ldots & u(t-n_{\hat{F}}-1) \\ \vdots & \ddots & \vdots \\ u(1) & \ldots & u(-n_{\hat{F}}+1) \end{bmatrix} \cdot \begin{bmatrix} \hat{f}_0(t) \\ \vdots \\ \hat{f}_{n_{\hat{F}}}(t) \end{bmatrix} \|^2 \\
&= \arg\min_{\hat{\mathbf{f}}} \|\mathbf{y}_{t\times 1} - \mathbf{U}_{t\times(n_{\hat{F}}+1)}\hat{\mathbf{f}}(t)_{(n_{\hat{F}}+1)\times 1}\|^2 \\
&= (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{y} \\
&= \mathbf{f} + (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{v}
\end{aligned}
$$

Bias term in the time domain: $\hat{\mathbf{f}}_{\mathbf{LS}}(t) - \mathbf{f} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{v}$

Prediction Error Method: Direct approach (2)

If a noise model $\hat{H}(q)$ is included in the PEM identification, the bias term in the frequency domain may be expressed as

$$\hat{F}(e^{j\omega}) - F(e^{j\omega}) = (H(e^{j\omega}) - \hat{H}(e^{j\omega}))\Phi_{wu}\Phi_u^{-1}$$

Since $\Phi_{wu} \neq 0$ due to one-sided correlation between $u(t)$ and $w(t)$:

$$u(t) = \frac{G(q)H(q)}{1 - G(q)F(q)}w(t),$$

$\hat{F}(q)$ will be a biased estimate of $F(q)$ unless

$$\hat{H}(q) = H(q), \quad \forall t.$$

Note that $\Phi_{wu}\Phi_u^{-1}$ can be seen as the inverse signal-to-noise ratio at the input of the system $F(q)$.

Noise model: Time-Varying Autoregressive (TVAR) modelling

Audio signals are commonly modelled as Autoregressive (AR) sequences, which leads to a Linear Prediction (LP) estimation problem.

However, audio signals are only stationary over short data windows (on average $20ms$ for speech). Hence for consistent estimation of the source signal model $H(q)$, a time-varying autoregressive (TVAR) model should be identified:

- either by tracking the TVAR coefficients using an adaptive algorithm (assuming 'slow' time-variation and short data windows, which is not always possible when concurrently estimating $H(q)$ and $F(q)$),
- or by basis expansion of
  - the TVAR coefficients (e.g. on a Fourier, Legendre, wavelet, ... basis)
  - the model structure (e.g. on a Kautz, Laguerre basis),
- or by regarding the TVAR coefficients as stochastic variables and estimating them using Kalman filtering.

Incorporating the noise model by prefiltering

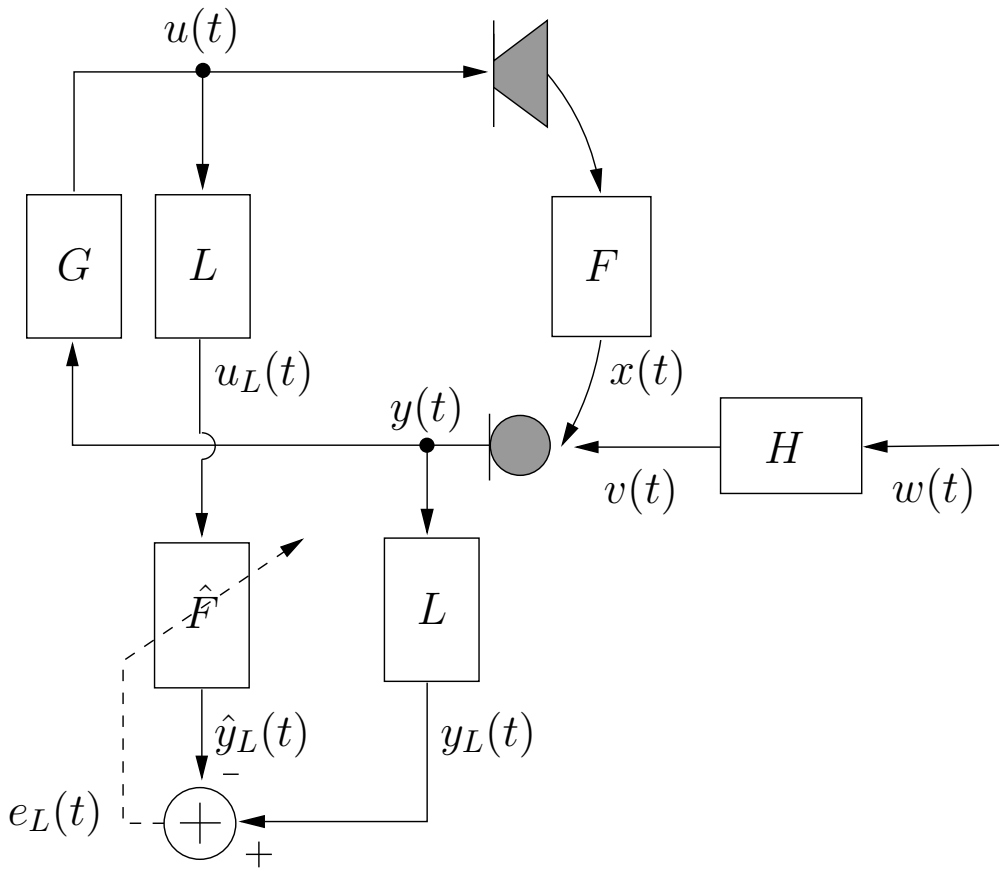
The noise model $\hat{H}(q)$ can be incorporated in the PEM identification by prefiltering the loudspeaker and microphone signals $u(t)$ resp. $y(t)$ with a (time-varying) filter $L(q) = \hat{H}^{-1}(q)$:

$$\begin{aligned}
\hat{\mathbf{f}}_{\mathbf{LS},L}(t) &= \arg\min_{\hat{\mathbf{f}}} \|\mathbf{L}_{t\times(t+n_L)}\mathbf{y}_{(t+n_L)\times 1} - \mathbf{L}_{t\times(t+n_L)}\mathbf{U}_{(t+n_L)\times(n_{\hat{F}}+1)}\hat{\mathbf{f}}(t)_{(n_{\hat{F}}+1)\times 1}\|^2 \\
&= (\mathbf{U}^T\mathbf{L}^T\mathbf{L}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{L}^T\mathbf{L}\mathbf{y} \\
&= \mathbf{f} + (\mathbf{U}^T\mathbf{L}^T\mathbf{L}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{L}^T\mathbf{L}\mathbf{v}
\end{aligned}$$

where prefiltering matrix $\mathbf{L}_{t\times(t+n_L)}$ is defined as

$$\mathbf{L}_{t\times(t+n_L)} = \begin{bmatrix} l_0(t) & l_1(t) & \dots & l_{n_L}(t) & 0 & \dots & 0 \\ 0 & l_0(t-1) & \dots & l_{n_L-1}(t-1) & l_{n_L}(t-1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & l_{n_L}(1) \end{bmatrix}.$$

If $L(q)$ is a consistent estimate of the true inverse source signal model $H^{-1}(q)$, $\forall t$, then $\mathbf{L}\mathbf{v} = \mathbf{w}$ and the bias term $(\mathbf{U}^T\mathbf{L}^T\mathbf{L}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{L}^T\mathbf{L}\mathbf{v}$ will dissapear as $t \longrightarrow \infty$.


$u(t)$
G
L
F
$u_L(t)$
$x(t)$
$y(t)$
H
$v(t)$
$w(t)$
$\hat{F}$
L
$\hat{y}_L(t)$
$y_L(t)$
$e_L(t)$
$-$
$+$

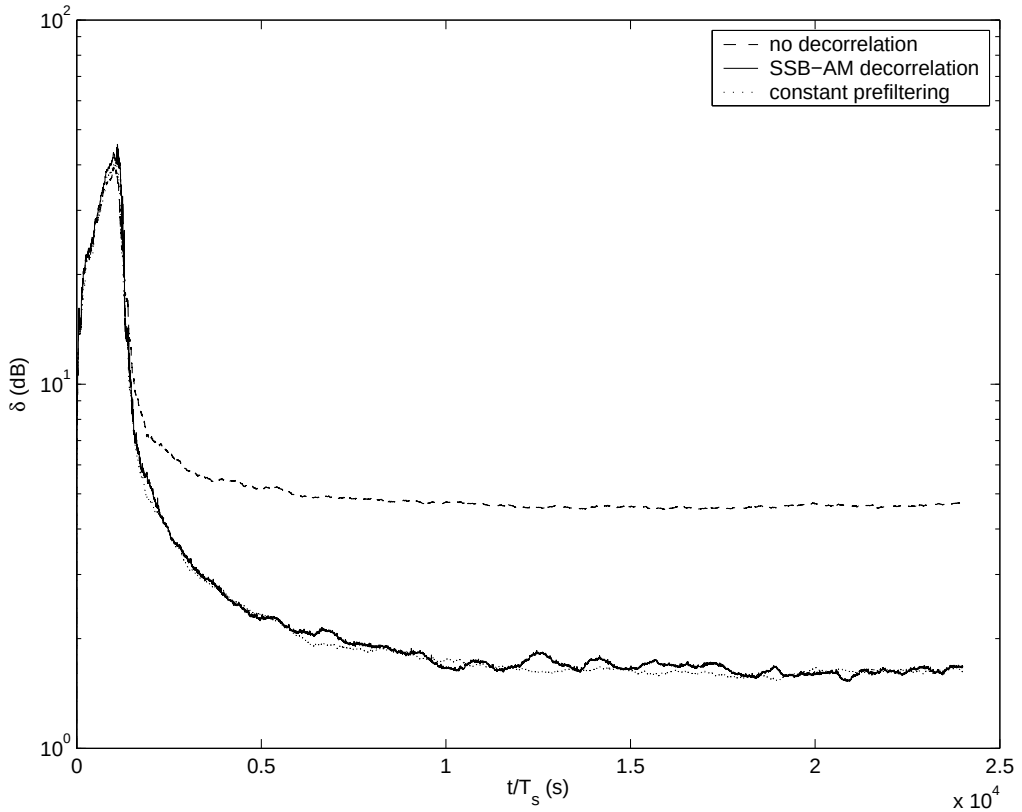
Simulations: TVAR prefiltering (1)

- Simulation parameters: $f_s = 8kHz$, $F(q)$ a pre-measured room impulse response, $n_F + 1 = 1000$, forward path $G(q) = Kq^{-1}$, $N = 24000$ samples, $GM = 3dB$, $\hat{F}(q)$ an exponentially windowed RLS adaptive filter with $\lambda = 0.9997$
- Performance measure: normalized bias $\delta(t) = 20 \log_{10} \frac{\|\hat{\mathbf{f}}_{\mathbf{LS}}(t) - \mathbf{f}\|}{\|\mathbf{f}\|}$
- Reference algorithm: decorrelation of $u(t)$ and $y(t)$ by SSB-AM frequency shifting[1]
- Prefiltering techniques (source signal $v(t)$ assumed to be known):
  - constant prefiltering
  - piecewise constant prefiltering with $N_{win} = 3000$ samples
  - BEM-TVAR prefiltering: basis expansion of source signal on a Fourier basis with 51 basis functions over time windows of $N_{win} = 3000$ samples
- Source signals:
  - stationary AR sequence of order 12 exhibiting a long-term speech spectrum
  - true speech signal

[1] C. P. Janse et al., "Signal amplifier system with improved echo cancellation", U.S. Patent 5,748,751, May 1998

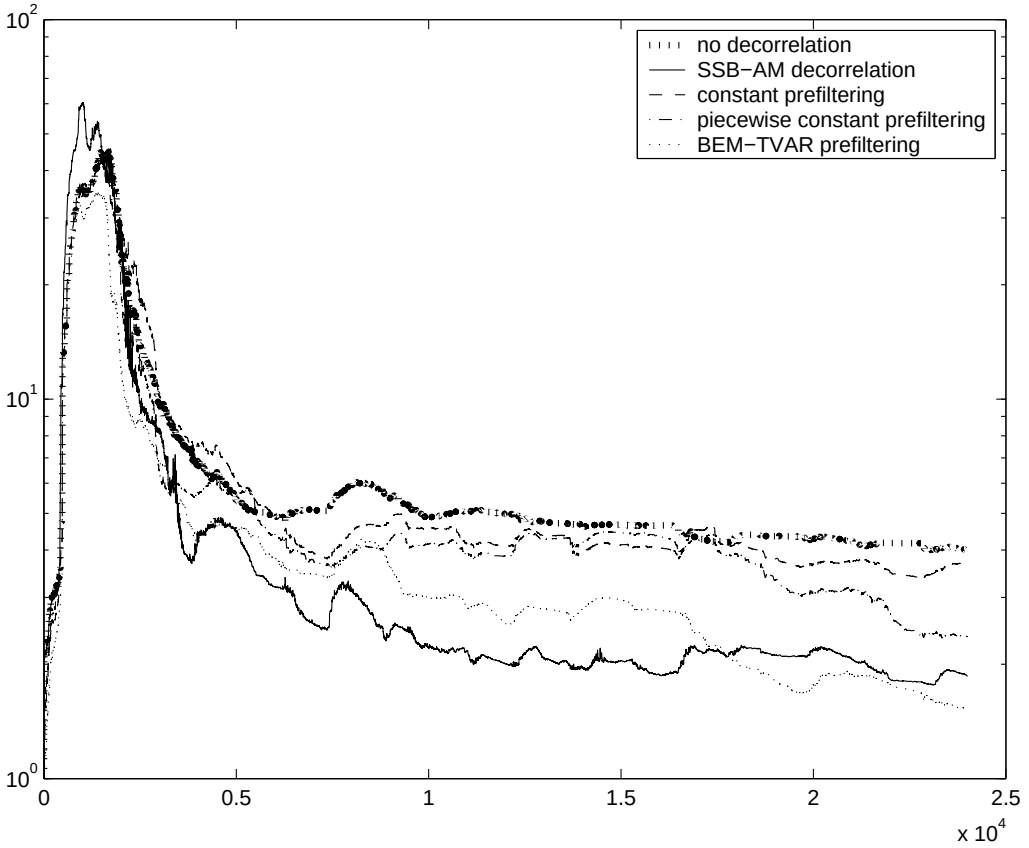
Simulations: TVAR prefiltering (2)

Source signal $v(t)$ = stationary AR sequence



Simulations: TVAR prefiltering (3)

Source signal $v(t)$ = true speech signal



Conclusions

- Consistent PEM identification of the acoustic feedback path $F(q)$ in adaptive feedback cancellation using the direct method, can be achieved by including a noise model $\hat{H}(q)$ in the identification procedure.
- The noise model $\hat{H}(q)$ should be a consistent estimate of the source signal model $H(q)$, which is generally time-varying due to the non-stationary nature of audio signals.
- Several approaches to time-varying autoregressive (TVAR) modelling exist. A BEM-TVAR technique that expands the TVAR coefficients onto a set of predefined basis functions was implemented under the assumption that the source signal was known.
- Under the latter assumption the BEM-TVAR prefiltering technique performs similar to a state-of-the-art reference algorithm in which decorrelation by frequency shifting is applied. However, the BEM-TVAR prefiltering technique does not introduce any signal distortion in the forward path, whereas the frequency shifting technique does.
- Concurrent estimation of the acoustic feedback path $F(q)$ and the inverse source signal model $H^{-1}(q)$ will be the subject of further work.