DEPARTAMENTO DE INGENIERIA ELECTRICA

UNIVERSIDAD NACIONAL DEL SUR

Adaptive Signal Processing: IIR Filtering

Juan Cousseau

1998

1 Introduction

Some common applications of adaptive filters

Applications contemplated

a. Echo cancelling.

b. Voice coding.

c. Inverse filtering (equalization).

d. Interference cancelling (active noise control, detection).

Aspects related to identification and control

1. Persistent excitation is important for robustness. Lack of persistent excitation leads to *drift* and *bursting* in feedback adaptive systems.

2. General algorithms using prediction error correlate a (filtered) version of the prediction error with a (filtered) version of the regressor.

3. A constant convergence (or gain) factor related to the updating algorithm leads to a bounded, but finite, asymptotic variance in the parameter estimation (misadjustement).

1.0.1 Basic recursive identifier

Consider the model $y(n)$ and the identifier $\hat{y}(n)$ (FIR! only to introduce) as follows

$$\begin{aligned} y(n) &= \sum_{i=0}^{N} b_i x(n-i) \\ \hat{y}(n) &= \sum_{i=0}^{N} \hat{b}_i(n) x(n-i) \end{aligned}$$

then the *prediction error* can be written as

$$\begin{aligned} e(n) &= y(n) + \nu(n) - \hat{y}(n) \\ &= (\boldsymbol{b}(n) - \hat{\boldsymbol{b}}(n))^T \boldsymbol{x}(n) + \nu(n) \\ &= \tilde{\boldsymbol{b}}^T(n)\boldsymbol{x}(n) + \nu(n) \end{aligned}$$

such that the LMS algorithm is defined by

$$\hat{\boldsymbol{b}}(n+1) = \hat{\boldsymbol{b}}(n) + \mu \boldsymbol{x}(n) e(n)$$

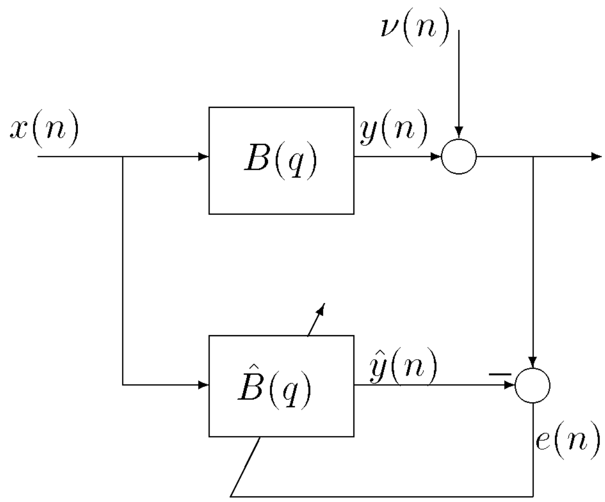

Figure 1: Basic recursive identifier

1.1 System identification: Echo Cancelling

Relevant aspects of the application

- Useful in typical long distance telephone loops. Essential in full duplex DSL.
- The hybrid design can not achieve echo attenuation lower than 6 dB.
- *Double talk* situation need to be detected. This can be interpreted in the figure by $f(n)$ (the far-end signal) similar to $\nu(n)$ (the near-end signal) in order that the identifier works suitably (this happens in practice if $x(n) - (y(n) + \nu(n)) < 6$ dB).

Formulation similar to the *basic recursive identifier* (except when feedback exist!).

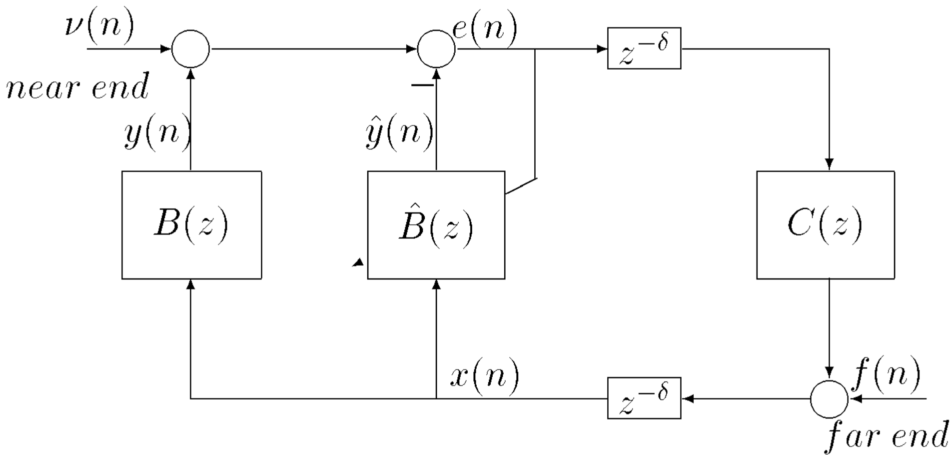

Figure 2: Echo cancelling

1.2 Prediction: Speech coding

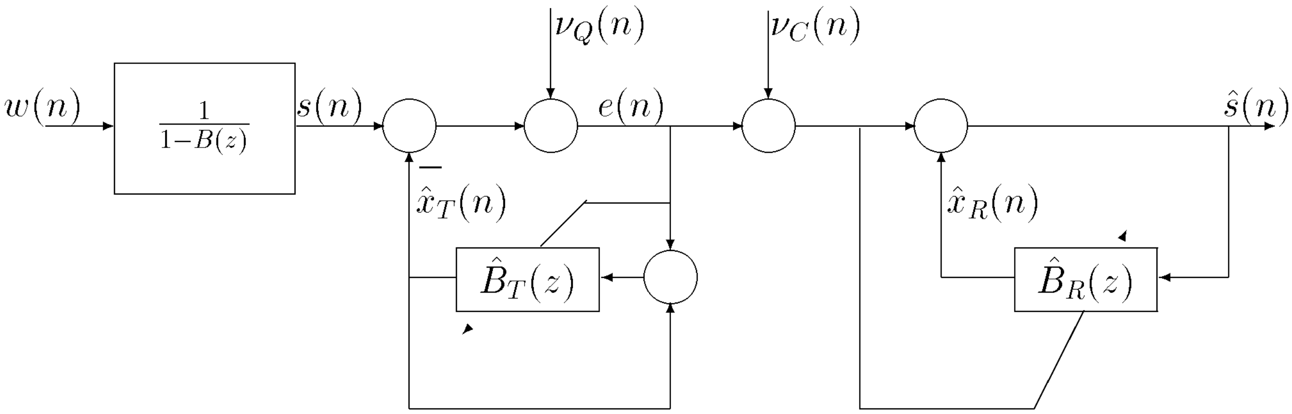

Figure 3: Speech coding application

- The signal model is: $s(n) = w(n) + \sum_{i=1}^{N} b_i s(n-i)$.
- The predictor is: $\hat{B}_T(q) = \sum_{i=1}^{N} \hat{b}_{iT} q^{-i}$.
- The transfer function between $s(n)$ and $e(n)$ will be: $1 - \hat{B}_T(q)$.
- The transfer function between $e(n)$ and $\hat{s}(n)$ will be: $\frac{1}{1-\hat{B}_R(q)}$.

In the transmitter:

$$\hat{x}_T(n) = \sum_{i=1}^{N} \hat{b}_{iT}(n)[\hat{x}_T(n-i) + e(n-i)] \tag{1}$$

and $e(n) = s(n) - \hat{x}_T(n) + \nu_Q(n)$ or

$$s(n) = e(n) - \nu_Q(n) + \hat{x}_T(n) \tag{2}$$

By replacing the signal model $s(n)$ in (2) and using (1),

$$\begin{aligned} e(n) &= w(n) + \sum b_i[e(n-i) - \nu_Q(n-i) + \hat{x}_T(n-i)] + \nu_Q(n) + \hat{x}_T(n) \\ &= w(n) + \sum b_i[e(n-i) - \nu_Q(n-i) \\ &\quad + \hat{x}_T(n-i)] + \nu_Q(n) + \sum \hat{b}_{iT}[e(n-i) + \hat{x}_T(n-i)] \\ &= \sum [b_i - \hat{b}_{iT}(n)][e(n-i) + \hat{x}_T(n-i)] \\ &\quad + w(n) + \nu_Q(n) - \sum \hat{b}_{iT} n_Q(n-i) \end{aligned}$$

$x(n)$	$e(n) + \hat{x}_T$
$y(n) + \nu(n)$	$s(n) + \nu_Q(n)$
$\hat{y}(n)$	$\hat{x}_T(n)$
$\nu(n)$	$w(n) + \nu_Q(n)[1 - \hat{B}_{iT}(q)]$

Then, the LMS algorithm related to this problem will be:

$$\hat{b}_i(n+1) = \hat{b}_i(n) + \mu e(n)[e(n-i) + \hat{x}(n-i)]$$

When $\nu_C(n) = 0$, this equation is useful for both transmitter and receiver. Since in general $\nu_C(n) \neq 0$, *leakage* is introduced in the receiver, i.e.,

$$\hat{b}_i(n+1) = (1-\lambda)\hat{b}_i(n) + \mu e(n)[e(n-i) + \hat{x}(n-i)]$$

1.3 Inverse filtering: Linear and Decision feedback equalization

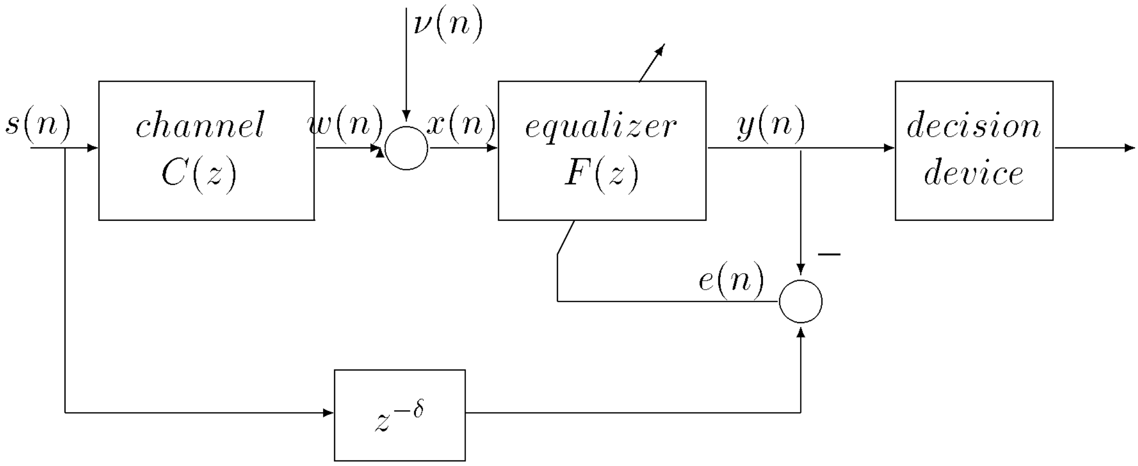

Figure 4: Linear equalization

In this case

$$y(n) = \sum_{i=0}^{N} f_i(n)x(n-i) \qquad x(n) = w(n) + \nu(n)$$

With an AR model of the channel $C(q) = \frac{1}{B(q)}$ such that $s(n) = \sum_{i=0}^{N} b_i w(n-i)$, i.e., a **minimum phase stable channel**, then

$$w(n) = \frac{s(n)}{b_0} - \sum_{i=1}^{N} \frac{b_i}{b_0} w(n-i)$$

Using training

$$\begin{aligned} e(n) &= s(n) - y(n) = \sum_{i=0}^{N} b_i w(n-i) - \sum_{i=1}^{N} f_i(n)x(n-i) \\ &= \sum_{i=0}^{N} (b_i - f_i(n))x(n-i) - \sum_{i=0}^{N} b_i \nu(n-i) \\ &= \tilde{\boldsymbol{b}}^T(n)\boldsymbol{x}(n) - \sum_{i=0}^{N} b_i \nu(n-i) \end{aligned}$$

This can be related to the basic identifier as shown in the in the following table

$\hat{\boldsymbol{b}}(n)$	$\boldsymbol{f}(n)$
$\nu(n)$	$-\sum_{i=0}^{N} b_i \nu(n-i)$

Note: since in general the noise term is not white (correlated by the channel parameters), the LMS parameter estimation is biased.

With a FIR model (**a truncated version**) of the channel, i.e., $C(q)$ such that

$$w(n) = \sum_{i=0}^{N} c_i s(n-i)$$

then $x(n) = \sum_{i=0}^{N} c_i s(n-i) + \nu(n)$.

The equalizer is described by

$$y(n) = \sum_{i=0}^{N} f_i x(n-i)$$

The channel - equalizer combination given by

$$\boldsymbol{h} = [h_0, ..., h_{2N}]^T$$

(convolution of $C(q)$ and $F(q)$) can be written as

$$\boldsymbol{h} = \boldsymbol{\Delta f}$$

where $\boldsymbol{\Delta} = \begin{bmatrix} c_0 & 0 & 0 \\ c_1 & c_0 & 0 \\ . & c_1 & c_0 \\ . & . & . \\ c_{2N+1} & . & . \\ 0 & . & . \\ & c_{2N+1} & . \\ & 0 & c_{2N+1} \end{bmatrix}$ $(2N+1) \times (N+1)$ and $\boldsymbol{f} = [f_0, ..., f_N]^T$.

With a delay of δ units, $\boldsymbol{h}^{opt} = [0...010...0]^T$.

Given $\boldsymbol{\Delta}$, the equation above can be solved in the mean square sense (pseudoinverse).

Problem exists in the extreme values (maximums) of the frequency response of the equalizer where the channel has a transfer function with small values.

Then, if channel noise exist, the noise power is amplified at the equalizer output.

A possible reduction of noise sensitivity is obtained considering the re-use of past detected symbols, i.e., **Decision Feedback Equalization**

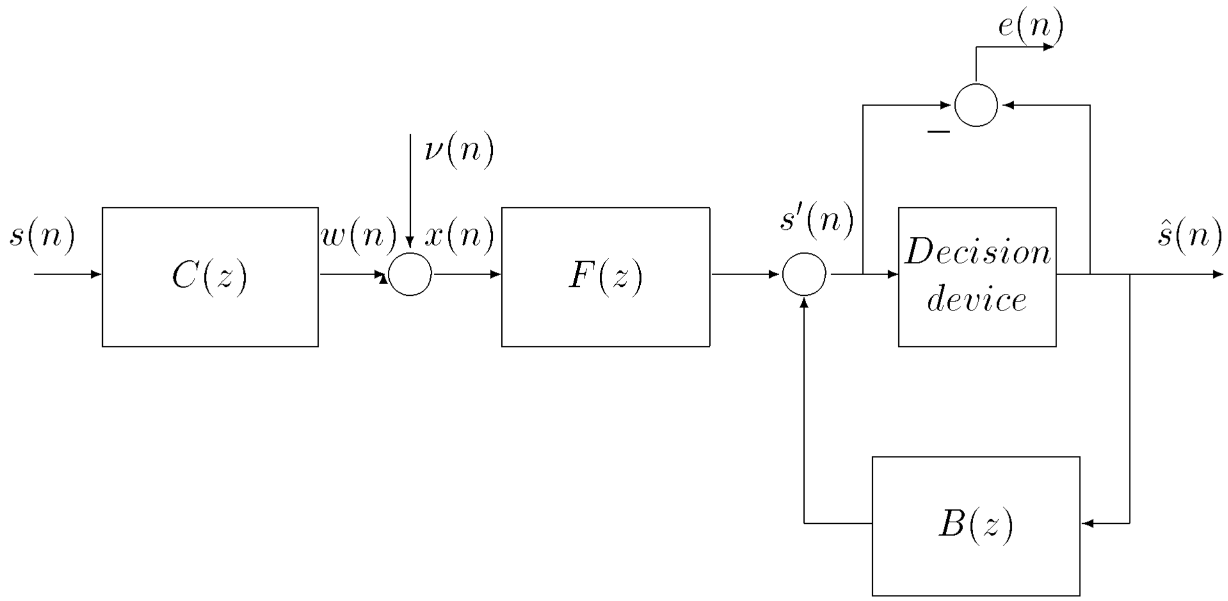

Figure 5: Decision feedback equalization

Consider the channel divided in

- a *precursor* response: $\sum_{i=-(N-1)}^{0} c_i(n)$ (equalized by $F(z)$) the interference to be equalized in future symbols,
- a *poscursor* response: $\sum_{i=1}^{M} c_i(n)$ (equalized by $B(z)$) the interference remaining in correct detected symbols.

In such a way that

$$s'(n) = \sum_{i=-(N-1)}^{0} f_i(n)x(n-i) - \sum_{i=1}^{M} b_i(n)\hat{s}(n-i) = F(q)x(n) - B(q)\hat{s}(n)$$

If the detected symbols are correct (or if a training period exist): $\hat{s}(n) = s(n)$, is easy to see that:

$$\begin{aligned} s'(n) &= F(q)(C(q)s(n) + \nu(n)) - B(q)s(n) \\ &= (F(q)C(q) - B(q))s(n) + F(q))\nu(n) \end{aligned}$$

Note that, for a given channel transfer function, noise not intervenes in the first term. Then the poscursor filter $B(q)$ can be obtained as a function of the precursor filter.

Because we work with a truncated version of the channel $C(q)$ we can not obtain a straightforward relationship to the basic recursive identifier.
But, considering that the error can be written as

$$e(n) = \hat{s}(n) - s'(n) = \hat{s}(n) - [(F(q)C(q) - B(q))s(n)] + F(q)\nu(n)]$$

is not hard to see that the associated LMS algorithm has the form

$$\begin{aligned} \hat{f}_i(n+1) &= \hat{f}_i(n) + \mu e(n)[\sum_{k=-(N-1)}^{0} f_k(n)x(n-k-i) - \sum_{k=1}^{M} b_k(n)\hat{s}(n-k-i)] \\ \hat{b}_j(n+1) &= \hat{b}_j(n) + \mu e(n)[\sum_{k=-(N-1)}^{0} f_k(n)x(n-k-j) - \sum_{k=1}^{M} b_k(n)\hat{s}(n-k-j)] \end{aligned}$$

for $i = -(N-1), ...0$ and $j = 1, ..., M$.

Another alternative is the independent channel identification and their utilization in equalization.

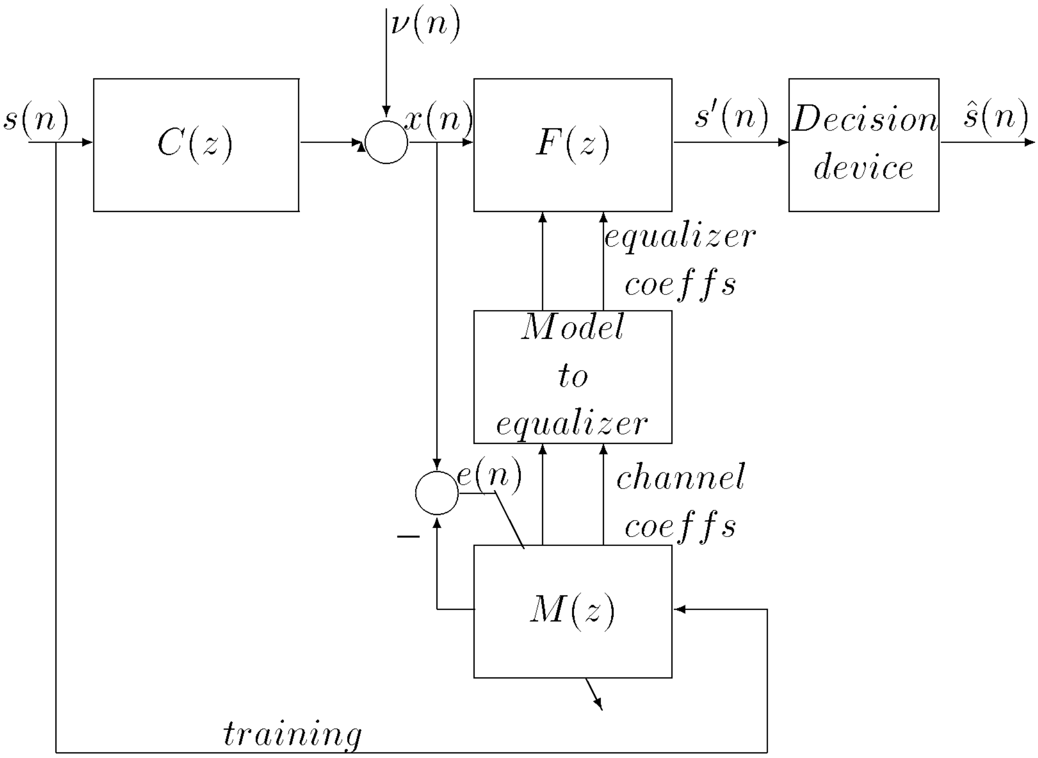

Figure 6: Independent channel identification

1.4 Interference cancelling

1.4.1 Active noise control

Many different configuration exist for Active noise control. A suitable one is the following

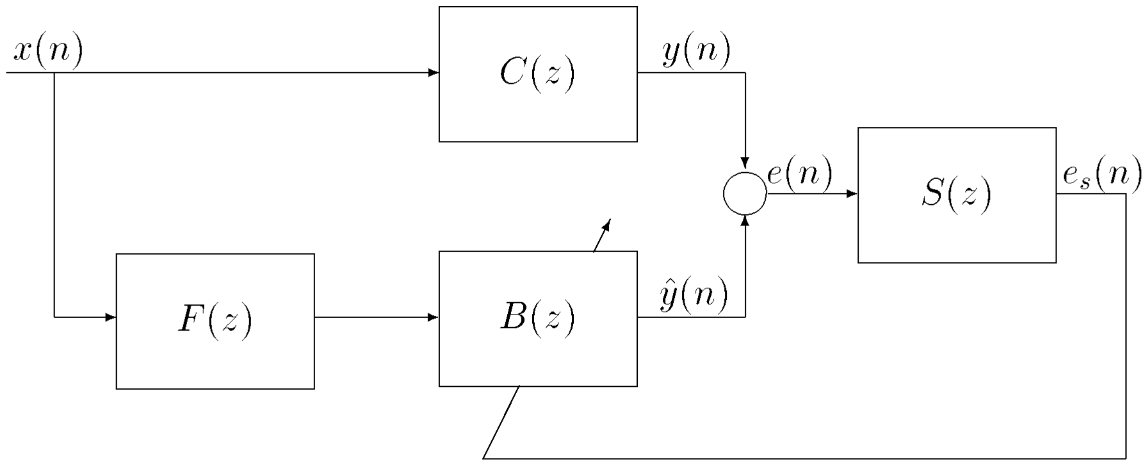

Figure 7: Active noise control: feedfoward (broadband) system

- $x(n)$: noise source (primary source, microphone).
- $\hat{y}(n)$ (secondary source, loudspeaker).
- $e_f(n)$ error source (monitor microphone).
- $F(z)$: noise transfer function.
- $C(z)$ acoustic system (i.e.: a duct).
- $B(z)$ adaptive noise controller.
- $S(z)$ transducer transfer function.

Assuming $F(z) = 1$, and in order to cope with the unobservable error $e(n)$, it is necessary to work with a **Filtered x-LMS algorithm**,

$$\hat{b}_i(n+1) = \hat{b}_i(n) + \mu e_s(n) x_f(n-i)]$$

where $x_f(n) = \hat{S}(q)x(n)$, where $\hat{S}(q)$ is an estimate of $S(q)$.

Since in this case,

$$\begin{aligned} e_s(n) &= S(q)e(n) = S(q)[C(q)x(n) - B(q)x(n)] \\ &= S(q)[\sum_{i=1}^{N}(c_i - b_i(n))x(n-i)] \\ &\cong \tilde{\boldsymbol{b}}^T(n)\boldsymbol{x}_f(n) \end{aligned}$$

where $\tilde{\boldsymbol{b}}(n) = (\boldsymbol{c} - \boldsymbol{b}(n))$. Then, the basic recursive identifier is related by

$\nu(n)$	0 (noise is the primary source)
$x(n)$	$x_f(n)$
$e(n)$	$e_s(n)$

1.4.2 Adaptive notch filters

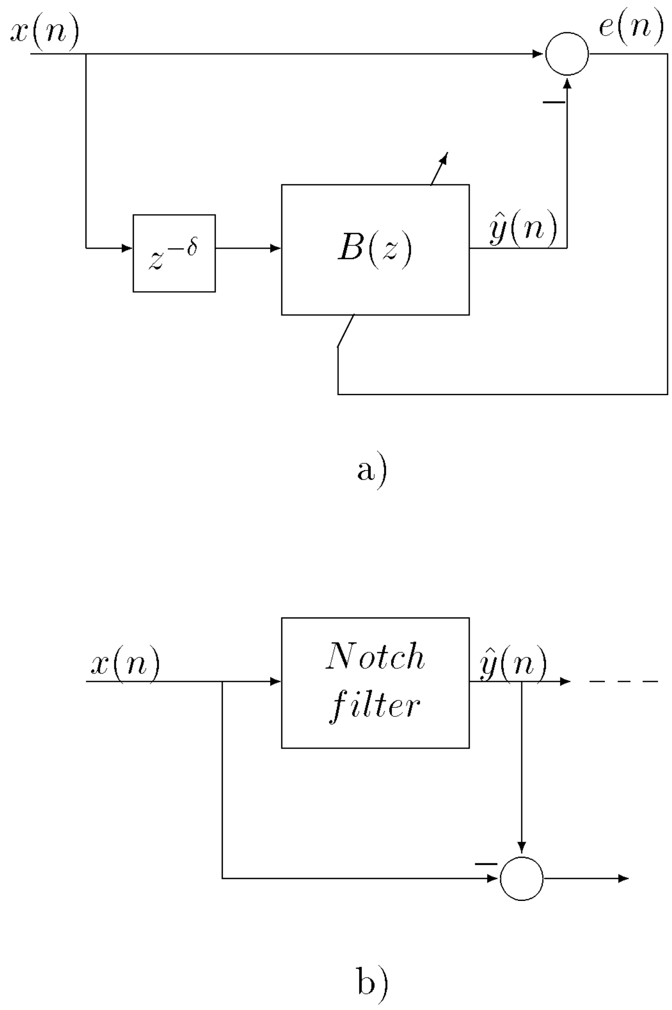

Figure 8: Adaptive line enhancer: a) FIR filter, b) IIR notch filter.

In this case the requirement is to improve the signal-to noise ratio for the input signal $x(n)$, described by

$$x(n) = c_o \, sin(w_o \, n + \phi_o) + \nu(n)$$

where c_o is the constant amplitude of the sinusoid of unknown frequency w_o, and ϕ_o is its phase.

This is a classical detection problem whose optimum solution (maximum signal-to-noise ratio) to recover a real signal $h(n)$ is given by the associated matched filter, $h(-n)$. In this case a discrete sinusoid.

- The adaptive FIR filter solution, $B(q) = \sum_{i=1}^{N} b_i(n) q^{-i}$, is obviously only a finite memory approximation of the matched filter. The delay $z^{-\delta}$ allows to work with several cycles.
- The specific sinusoid frequency w_o can be obtained using an FFT on the estimated impulse response coefficients, $b_i(n)$.
- As could be expected, high N improves the signal-to-noise ratio of the estimate.
- The main problem, without regarding computational complexity, is the occurrence of false noise-induced peaks in the frequency response.

- Consider now the variance $E[y^2(n)]$ of the output signal $y(n)$ in the second configuration, that can be written as

$$E[y^2(n)] = c_o^2\,|H(e^{jw})|^2 + E[\nu^2(n)]$$

if $H(e^{jw}) = \begin{cases} 0 & w = w_o \; and \; w = -w_o \\ 1 & \forall w \end{cases}$, i.e., an ideal notch filter, we obtain

$$E[y^2(n)] = \begin{cases} E[\nu^2(n)] & w = w_o \\ c_o^2 + E[\nu^2(n)] & \forall w \end{cases}$$

- then we can recover the sinusoid using an ideal bandpass filter given by

$$|G(e^{jw})|^2 = 1 - |H(e^{jw})|^2 = \begin{cases} 1 & w = w_o \; and \; w = -w_o \\ 0 & \forall w \end{cases}$$

- The adaptive IIR filter solution contemplates the use of a practical notch filter with, for example, the following transfer function:

$$H(z) = \frac{1 + a\, z^{-1} + z^{-2}}{1 + a\, r\, z^{-1} + r^2 z^{-2}}$$

where: $0 < r < 1$ is a constant, and a is related to the sinusoid frequency w_o by $w_o = arcos(-a/2)$, provided $-2 < a < 2$.

- It is not hard to see that the frequency response of $H(z)$ is a notch filter with notch bandwidth decreasing when $r \rightarrow 1$.

- A basic recursive identifier is not trivial in this case, mainly because the adaptive filter is IIR in this case. Anyway, the algorithm to be considered has the familiar form

$$a(n+1) = a(n) - \mu y(n) \nabla_a(n)$$

where $\nabla_a = \frac{\partial y^2(n)}{\partial a(n)}$ and obviously $a(n)$ is the parameter to be updated.

- Straightforward calculations show that

$$\nabla_a(n) = (1-r)\left(\frac{1 - r\, q^{-2}}{(1 + a\, r\, q^{-1} + r^2 q^{-2})^2}\right) y(n-1)$$

- Finally the relationship with the basic recursive identifier is given by

$\nu(n)$	0 (noise exist at input)
$x(n)$	$\nabla_a(n)$
$e(n)$	$y(n)$

1.5 Overview and objectives

An outline of the proposed contents is the following:

1. Adaptive FIR filters. Some algorithms and their limitations.

2. Adaptive IIR filters. Motivation from system identification theory.

3. Some useful tools. Concepts on Approximation and Stability theory.
   (a) Considerations on time variant linear systems.
   (b) ODE, conditions for the association, Liapunov function. Stationary points: theorems.
   (c) Approximation concepts: Orthonormal space decomposition of $\mathcal{L}_2$ (interpolation), relationship with Hankel norm.
   (d) Stability concepts: Stability of a quasi-time-invariant linear system. Stability of a particular non linear system: passivity and hyperstability.

4. MSOE minimization and related algorithms.

5. The Equation Error perspective. An IIR extension of the FIR adaptive filter.

6. Alternative criteria I: HARF, an stable but incomplete solution.

7. Alternative criteria II: Steiglitz-McBride, the closest approximation to the global minimum.

8. A brief discussion of adaptive IIR filters.