8.1 Introduction

Consider, as in the previous chapters, that

$$d(n) = \frac{B(q)}{A(q)}x(n) + \nu(n) = y(n) + \nu(n)$$

To estimate the parameter associated to $A(q)$ and $B(q)$, the following function was proposed

$$J(\boldsymbol{\theta}(n+1)) = E\left\{\left(A_{n+1}(q)\frac{d(n)}{A_n(q)} - \hat{B}_{n+1}(q)\frac{x(n)}{A_n(q)}\right)^2\right\} \tag{88}$$

In order to obtain $\boldsymbol{\theta}(n+1)$, (88) is minimized assuming known $\boldsymbol{\theta}(n)$, i.e., a LS problem at the $(n+1)$-th iteration.

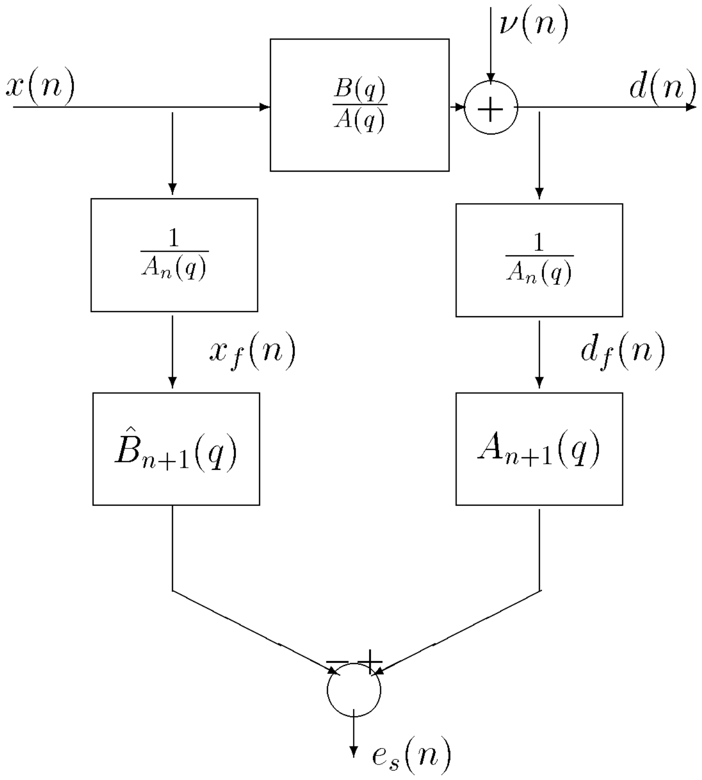

Figure 53: The Steiglitz-McBride method

Related to equation (88), the following definitions are useful

- $$\boldsymbol{\theta}_o = [a_1^o, ...a_{n_a}^o, b_0^o, ..., b_{n_b}^o]^T$$

  the ideal coefficients,

- $$\boldsymbol{\theta}(n) = [a_1(n), ...a_{n_a}(n), b_0(n), ..., b_{n_b}(n)]^T$$

  the estimated or filter coefficients,

- $$\boldsymbol{\phi}(n) = [\frac{d(n-1)}{A_n(q)}, ...\frac{d(n-N)}{A_n(q)}, \frac{x(n)}{A_n(q)}, ...\frac{x(n-M)}{A_n(q)}]^T$$

  the regressor related to the SM method.

Then, asymptotically, the Least Squared solution of equation (88) can be written as follows

$$\begin{aligned}\boldsymbol{\theta}(n+1) &= \left[E\left\{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\right\}\right]^{-1} E\left\{\boldsymbol{\phi}(n)\frac{d(n)}{A_n(q)}\right\} \\ &= \boldsymbol{\theta}(n) + \left[E\left\{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\right\}\right]^{-1} E\left\{\boldsymbol{\phi}(n)e_s(n)\right\}\end{aligned}$$

where $e_s(n) = \frac{d(n)}{A_n(q)} - \boldsymbol{\phi}^T(n)\boldsymbol{\theta}(n)$.

8.1.1 Generic stability

As a direct extension of the EE method properties, at a stationary point (88) is

$$\begin{aligned} J(\boldsymbol{\theta}) &= [-\boldsymbol{a}^T \; -\boldsymbol{b}^T] E\{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\} \begin{bmatrix} -\boldsymbol{a} \\ -\boldsymbol{b} \end{bmatrix} \\ &= [-\boldsymbol{a}^T \; -\boldsymbol{b}^T] \begin{bmatrix} \boldsymbol{R}_d^f & {\boldsymbol{R}_{xd}^f}^T \\ {\boldsymbol{R}_{xd}^f}^T & \boldsymbol{R}_x^f \end{bmatrix} \begin{bmatrix} -\boldsymbol{a} \\ -\boldsymbol{b} \end{bmatrix} \end{aligned}$$

where $\boldsymbol{R}_d^f = E\{\frac{d(n-i)}{A_n(z)}\frac{d(n-j)}{A_n(z)}\}$, $\boldsymbol{R}_{xd}^f = E\{\frac{x(n-i)}{A_n(z)}\frac{d(n-i)}{A_n(z)}\}$ and $\boldsymbol{R}_x^f = E\{\frac{x(n-i)}{A_n(z)}\frac{x(n-j)}{A_n(z)}\}$.
The factorization of the covariance matrix is useful

$$\begin{bmatrix} \boldsymbol{R}_d^f & {\boldsymbol{R}_{xd}^f}^T \\ \boldsymbol{R}_{xd}^f & \boldsymbol{R}_x^f \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{N+1} & {\boldsymbol{R}_x^f}^{-1}\boldsymbol{R}_{xd}^f \\ & \boldsymbol{I}_{N+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_d^f - {\boldsymbol{R}_{xd}^f}^T {\boldsymbol{R}_x^f}^{-1}\boldsymbol{R}_{xd}^f & \\ & \boldsymbol{R}_x^f \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{N+1} & \\ {\boldsymbol{R}_{xd}^f}^T {\boldsymbol{R}_x^f}^{-1} & \boldsymbol{I}_{N+1} \end{bmatrix}$$

and by defining $\boldsymbol{R}_{d/x}^f = \boldsymbol{R}_d^f - {\boldsymbol{R}_{xd}^f}^T {\boldsymbol{R}_x^f}^{-1}\boldsymbol{R}_{xd}^f$, and pre and post-multiplying by the parameter vector

$$J(\boldsymbol{\theta}) = \boldsymbol{a}^T \boldsymbol{R}_{d/x}^f \boldsymbol{a} + [\boldsymbol{b} - {\boldsymbol{R}_x^f}^{-1}\boldsymbol{R}_{xd}^f\boldsymbol{a}]^T \boldsymbol{R}_x^f [\boldsymbol{b} - {\boldsymbol{R}_x^f}^{-1}\boldsymbol{R}_{xd}^f\boldsymbol{a}]$$

Minimizing with respect to $\boldsymbol{b}$

$$E\{e_e^2(n)\} = \boldsymbol{a}^T \boldsymbol{R}_{d/x}^f \boldsymbol{a}$$

Now, with the monic constraint on $\boldsymbol{a}$ ($\boldsymbol{a}\boldsymbol{\psi} = 1$, with $\boldsymbol{\psi} = [1, 0, ..., 0]^T$), this last equation leads to

$$\boldsymbol{R}_{d/x}^f \boldsymbol{a} = \sigma_e^2 \boldsymbol{\psi}$$

where σ_e is the value of $J(\boldsymbol{\theta})$ at the minimum.

Theorem

- *Sufficient order case.* With $deg\ H(z) = N$ and for a signal-to-noise ratio given by $S = \frac{E\left\{\left(\frac{B(q)}{A(q)}x(n)\right)^2\right\}}{E\{\nu^2(n)\}}$, $A_n(q)$ has zeros inside the stability region if some of the following conditions is satisfied
  - S is sufficiently high.
  - S is sufficiently low.
- *Insufficient order case.* With $deg\ H(z) > N$, then $A_n(q)$ has zeros inside the stability region if $x(n)$ is white noise.

The second part of the theorem is a direct extension of the result discussed for the EE method, and follows from

Lemma: Let the sequences $x^f(n)$ and $d^f(n)$, related by

$$d^f(n) = H(z)u^f(n) + \nu^f(n)$$

where $H(z)$ is stable and causal, and where the disturbance $\nu^f(n)$ is statistically independent of $x^f(n)$. If $x^f(n)$ is an AR process of order not exceding N, then $A_n(q)$ obtained with the monic constraint in the SM method has zeros inside the stability region.

Now if $x(n)$ is white, $x^f(n)$ is an N order AR process and the previous lemma guarantee the stability of the estimate.

8.1.2 Stationary points

The basic equations related to the SM iteration are

$$\boldsymbol{R}^f_{d/x}\boldsymbol{a} = \sigma_e^2\boldsymbol{\psi} \tag{89}$$

$$\boldsymbol{R}^f_x\boldsymbol{b} = \boldsymbol{R}_{xd}\boldsymbol{a} \tag{90}$$

Based on the definitions of the variables involved, the second equation can be rewritten as

$$\mathbf{0}_{N+1} = \left\langle \begin{bmatrix} 1/A_n(z) \\ z^{-1}/A_n(z) \\ \vdots \\ z^{-N}/A_n(z) \end{bmatrix}, \mathcal{S}_x(z)\left[\frac{A_{n+1}(z)}{A_n(z)}H(z) - \frac{B_{n+1}(z)}{A_n(z)}\right]\right\rangle$$

Then at any stationary point, i.e., $A_{n+1}(z) = A_n(z)$,

$$\mathbf{0}_{N+1} = \left\langle \begin{bmatrix} 1/A(z) \\ z^{-1}/A(z) \\ \vdots \\ z^{-N}/A(z) \end{bmatrix}, \mathcal{S}_x(z)\left[H(z) - \hat{H}(z)\right]\right\rangle$$

Then, assuming $\mathcal{S}_x(z)\left[H(z) - \hat{H}(z)\right] = f(z) = \sum_{k=-\infty}^{\infty} f_k z^{-k}$, and using the decomposition theorem,

$$[f(z)]_+ = z^{-1}V(z)g(z) \quad g(z) \in \mathcal{H}_2 \quad and \quad f_0 = 0$$

where $V(z) = \frac{z^{-N}A(z^{-1})}{A(z)}$.

On the other hand, (89) can be rewritten as

$$\begin{aligned}\mathbf{0}_N &= \left\langle \begin{bmatrix} z^{-1}/A_n(z) \\ \vdots \\ z^{-N}/A_n(z) \end{bmatrix} H(z), \mathcal{S}_x(z)\left[\frac{A_{n+1}(z)}{A_n(z)}H(z) - \frac{B_{n+1}(z)}{A_n(z)}\right]\right\rangle \\ &\quad + E\left\{ \begin{bmatrix} \nu^f(n-1) \\ \vdots \\ \nu^f(n-N) \end{bmatrix} A_{n+1}\nu^f(n) \right\}\end{aligned}$$

and at a stationary point this reads as

$$\mathbf{0}_N = \left\langle \begin{bmatrix} z^{-1}/A(z) \\ \vdots \\ z^{-N}/A(z) \end{bmatrix} H(z), \mathcal{S}_x(z)\left[H(z) - \hat{H}(z)\right]\right\rangle + E\left\{ \begin{bmatrix} \nu^f(n-1) \\ \vdots \\ \nu^f(n-N) \end{bmatrix} \nu(n) \right\}$$

Assuming for a moment that $\nu(n) = 0$

$$\begin{aligned}\mathbf{0}_N &= \left\langle \begin{bmatrix} z^{-1}/A(z) \\ \vdots \\ z^{-N}/A(z) \end{bmatrix} H(z), \mathcal{S}_x(z)\left[H(z) - \hat{H}(z)\right]\right\rangle \\ &= \left\langle \begin{bmatrix} z^{-1}/A(z) \\ \vdots \\ z^{-N}/A(z) \end{bmatrix} H(z), z^{-1}V(z)g(z)\right\rangle\end{aligned}$$

After some reordering, using the $V(z)$ definition,

$$\mathbf{0}_N = \left\langle \begin{bmatrix} 1/A(z) \\ z^{-1}/A(z) \\ \vdots \\ z^{-(N-1)}/A(z) \end{bmatrix}, z[H(z)g(z^{-1})]_+ \right\rangle$$

Then, by the decomposition theorem, this is satisfied if $z[H(z)g(z^{-1})]_+ = V(z)Q(z)$.

A complementary result here is the following

$$\begin{aligned}
\left\langle H(z)-\hat{H}(z), \mathcal{S}_x(z)[H(z)-\hat{H}(z)]\right\rangle &= \left\langle H(z)-\hat{H}(z), f(z)\right\rangle \\
&= \left\langle H(z)-\hat{H}(z), z^{-1}V(z)g(z)\right\rangle \\
&= \left\langle H(z), z^{-1}V(z)g(z)\right\rangle - \left\langle \hat{H}(z), z^{-1}V(z)g(z)\right\rangle \\
&= \left\langle H(z), z^{-1}V(z)g(z)\right\rangle \\
&= \left\langle H(z)g(z^{-1}), z^{-1}V(z)\right\rangle \\
&= \left\langle [H(z)g(z^{-1})]_+, z^{-1}V(z)\right\rangle \\
&= \left\langle zV(z)Q(z), z^{-1}V(z)\right\rangle = \langle Q(z), 1\rangle = Q(0)
\end{aligned}$$

(the second term of the third equation vanish because $\hat{H}(z)V(z^{-1})$ is anti-causal).

Theorem: (noise free case) Let $f(z) = \mathcal{S}_x(z)\left[H(z)-\hat{H}(z)\right]$. $\hat{H}(z) \in \mathcal{H}_2$ is a stationary point of the SM iteration if and only if:

- $f_0 = 0$ and $[f(z)]_+ = z^{-1}V(z)g(z)$, $g(z) \in \mathcal{H}_2$, that satisfies $[H(z)g(z^{-1})]_+ = z^{-1}V(z)Q(z)$, for some $Q(z) \in \mathcal{H}_2$.
- Also, the error in $\mathbf{L}_2$ norm is: $\left\langle H(z)-\hat{H}(z), \mathcal{S}_x(z)[H(z)-\hat{H}(z)]\right\rangle = Q(0)$.

- If *deg* $H(z) = N$ (sufficient order) and $R(z) = [H(z)g(z^{-1})]_+$.
- But *deg* $R(z) \leq N$, then if $R(z)$ is not zero it can have at most N zeros. Since it is strictly causal, it must have a zero at $z = 0$, i.e., $R(0) = 0$.
- Then $zR(z) = z[H(z)g(z^{-1})]_+$, a causal function now has $N-1$ zeros.
- But, by the theorem $z[H(z)g(z^{-1})]_+ = V(z)Q(z)$, N zeros outside the unit circle (i.e., those of $V(z)$), but this is impossible, at least than $R(z) = 0$, or $Q(z) = 0$.
- Then, $\left\langle H(z)-\hat{H}(z), \mathcal{S}_x(z)[H(z)-\hat{H}(z)]\right\rangle = 0$.
- For a general input $x(n)$, this is not the case for the MSOE minimization.

- Note that $R(z) = [H(z)g^{-1}]_+$ in matrix form is

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots \\ h_2 & h_3 & h_4 & \cdots \\ h_3 & h_4 & h_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \end{bmatrix}$$

  or $\boldsymbol{r} = \boldsymbol{\Gamma}_H \boldsymbol{g}$. For the sufficient order case, $\boldsymbol{r} = 0$, and the relationship with Hankel form is evident.

- For the undermodelled case, $R(z) = z^{-1}V(z)Q(z)$ appears as

$$\begin{bmatrix} v_0 & 0 & 0 & \cdots \\ v_1 & v_0 & 0 & \cdots \\ v_2 & v_1 & v_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots \\ h_2 & h_3 & h_4 & \cdots \\ h_3 & h_4 & h_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \end{bmatrix}$$

  or $\boldsymbol{V}\boldsymbol{q} = \boldsymbol{\Gamma}_H \boldsymbol{g}$. As will be discussed, this will give an approximate solution to $\boldsymbol{\Gamma}_H$.

If we consider the disturbance term, the previous result generalizes to the following

Theorem: Let $f(z) = \mathcal{S}_x(z)\left[H(z) - \hat{H}(z)\right]$. $\hat{H}(z) \in \mathcal{H}_2$ is a stationary point of the SM iteration if and only if:

- $f_0 = 0$ and $[f(z)]_+ = z^{-1}V(z)g(z)$, $g(z) \in \mathcal{H}_2$, that satisfies $[H(z)g(z^{-1})]_+ + [\mathcal{S}_x(z)]_+ = z^{-1}V(z)Q(z)$, for some $Q(z) \in \mathcal{H}_2$.
- Also, the error in $\mathbf{L}_2$ norm is: $\left\langle H(z) - \hat{H}(z), \mathcal{S}_x(z)[H(z) - \hat{H}(z)]\right\rangle = Q(0) - < \mathcal{S}_x(z), z^{-1}V(z) >$.

this can leads to an interpolation interpretation that extend the results disposable for the EE method,

Corollary: If $x(n)$ is white noise,

- $H(z) - \hat{H}(z) = z^{-1}V(z)g(z)$, for some $g(z) \in \mathcal{H}_2$.
- $[H(z)H(z^{-1}]_+ - [\hat{H}(z^{-1})\hat{H}(z)]_+ = z^{-1}V(z)Q(z) - [\mathcal{S}_x(z)]_+$.

- The first condition reads as $\hat{H}(\alpha_k) = H(\alpha_k)$ and $\hat{H}(0) = H(0)$, $k = 1, ..., N$.

- Defining $r_k = \sum_{l=0}^{\infty} h_l h_{l+k}$ and $\hat{r}_k = \sum_{l=0}^{\infty} \hat{h}_l \hat{h}_{l+k}$, the second part can be written as $\sum_{l=1}^{\infty} \hat{r}_l \alpha_k^{-l} = \sum_{l=1}^{\infty} r_l \alpha_k^{-l} + [\mathcal{S}(z)]_+|_{z=\alpha_k^{-1}}$, $k = 1, .., N$.

  this leads to interpolation constraints similar to those obtained with the unit norm equation error, but now at the reciprocal of the poles of $\hat{H}(z)$ not at $z = 0$ as in the EE case.

8.2 On-line algorithm

Considering signal processing applications an on-line version of the SM method is important. Some basic approach are the following

- Ljung propose the following error

$$\begin{aligned} e_s(n) &= \frac{d(n)}{A_n(q)} - \boldsymbol{\phi}^T(n)\boldsymbol{\theta}(n) \\ &= \frac{A_n(q)}{A_{n-1}(q)} e_o(n) \end{aligned}$$

where $\boldsymbol{\phi}(n)$ is defined as previously, and

$$e_o(n) = d(n) - \hat{y}(n) = d(n) - \frac{A_n(q)}{B_n(q)} x(n)$$

is the output error. On the other side, Fan consider that $e_s(n) \approx e_o(n)$.

- The basic difference between the *independent filtering* algorithm proposed by Fan, and the OE method studied is that in the latter the regressor is composed of filtered versions of the adaptive filter output and in the former the regressor is composed of filtered versions of $d(n)$.

Then, a suitable realization for on-line adaptive filtering of the SM method is the following

$$\begin{aligned} \boldsymbol{P}(n+1) &= \frac{1}{1-\alpha}\left[\boldsymbol{P}(n) - \frac{\boldsymbol{P}(n)\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\boldsymbol{P}(n)}{\frac{1-\alpha}{\alpha} + \boldsymbol{\phi}^T(n)\boldsymbol{P}(n)\boldsymbol{\phi}(n)}\right] \\ \boldsymbol{\theta}(n+1) &= \boldsymbol{\theta}(n) + \alpha \boldsymbol{P}(n+1)\boldsymbol{\phi}(n)e_s(n) \end{aligned}$$

where α is the convergence factor. Making $\boldsymbol{P}(n+1) = \boldsymbol{I}$, we obtain the stochastic gradient version algorithm, given by

$$\boldsymbol{\theta}(n+1) = \boldsymbol{\theta}(n) + \alpha\boldsymbol{\phi}(n)e_s(n)$$

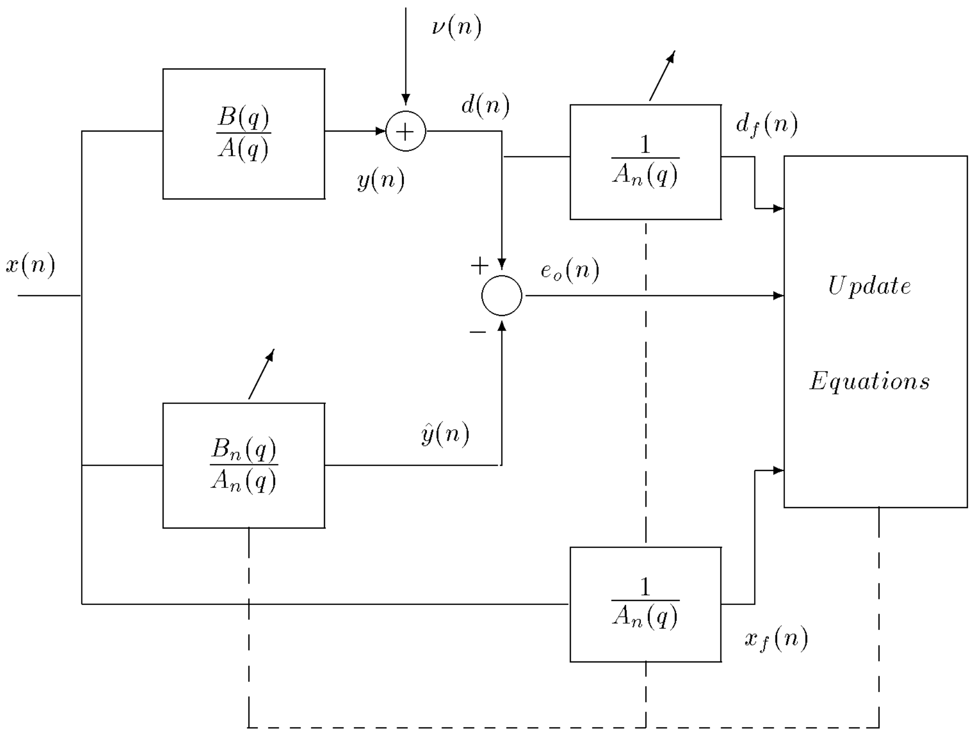

Figure 54: Block diagram of the on-line Steiglitz-McBride method.

8.2.1 ODE associated

The ODE associated to the SM method in the on-line version, assuming white measurement noise and a stability check mechanism, is the following

$$\begin{aligned}\frac{\partial \boldsymbol{\vartheta}(t)}{\partial t} &= E\{\boldsymbol{\phi}(n)e'(n)\} \\ &= -\mathbf{G}(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o)\end{aligned}$$

where $e'(n) = e_o(n)$ and $e_o(n)$ is the output error,

$$\boldsymbol{\phi}(n) = [\frac{d(n-1)}{A_n(q)}, \ldots, \frac{d(n-N)}{A_n(q)}, \frac{x(n)}{A_n(q)}, \ldots, \frac{x(n-N)}{A_n(q)}]^T \tag{91}$$

and

$$\mathbf{G} = E\{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\}$$

and for the Gauss-Newton version

$$\begin{aligned}\frac{\partial \boldsymbol{\vartheta}(t)}{\partial t} &= \boldsymbol{\varrho}^{-1}(t)\mathbf{G}\,(\boldsymbol{\theta_o} - \boldsymbol{\vartheta}(t)) \\ \frac{\partial \boldsymbol{\varrho}(t)}{\partial t} &= \mathbf{G} - \boldsymbol{\varrho}(t)\end{aligned}$$

The Liapunov function related to the ODE analysis is given by

$$V(\boldsymbol{\vartheta}(t)) = \frac{1}{2}(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o)^T(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o)$$

such that

$$\begin{aligned}(dV/dt) &= \frac{1}{2}(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o)^T\frac{\partial\boldsymbol{\theta}(t)}{\partial t} + \frac{1}{2}\left(\frac{\partial\boldsymbol{\theta}(t)}{\partial t}\right)^T(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o) \\ &= -(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o)^T\mathbf{G}(\boldsymbol{\theta}(t)-\boldsymbol{\theta}_o) \\ &\leq 0\end{aligned}$$

As in the OE method, the stability check for this algorithm is not trivial, specially for a direct form adaptive filter realization.

8.3 Relationship between the MSOE and MSSME

8.3.1 Structural interpretation and stationary points

Consider $\boldsymbol{x}^f(n)$ and $\boldsymbol{d}^f(n)$ rewritten as

$$\begin{bmatrix} x^f(n) \\ x^f(n-1) \\ \vdots \\ x^f(n-(N-1)) \\ x^f(n-N) \end{bmatrix} = \begin{bmatrix} a_1(n) & a_2(n) & \cdots & a_N(n) & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x^f(n-1) \\ x^f(n-2) \\ \vdots \\ x^f(n-N) \\ x(n) \end{bmatrix}$$

$$= \boldsymbol{Q}_d(n) \begin{bmatrix} x^f(n-1) \\ x^f(n-2) \\ \vdots \\ x^f(n-N) \\ x(n) \end{bmatrix}$$

$$\begin{bmatrix} d^f(n) \\ d^f(n-1) \\ \vdots \\ d^f(n-(N-1)) \\ d^f(n-N) \end{bmatrix} = \boldsymbol{Q}_d(n) \begin{bmatrix} d^f(n-1) \\ d^f(n-2) \\ \vdots \\ d^f(n-N) \\ d(n) \end{bmatrix}$$

in particular $\boldsymbol{Q}_d(n)$ is invertible, i.e, $[\boldsymbol{Q}_d(n)]^{-1}\boldsymbol{Q}_d(n) = \boldsymbol{I}_{N+1}$. Also, if $\boldsymbol{q}_d^T(n)$ denote the final row of $[\boldsymbol{Q}_d(n)]^{-1}$, then

$$\begin{aligned} \boldsymbol{q}_d^T(n) &= [1\ a_1(n)\ \cdots\ a_N(n)] = \boldsymbol{a}^T(n) \\ \boldsymbol{q}_d^T(n)\boldsymbol{Q}_d(n) &= [0\ \cdots\ 0\ 1] \end{aligned}$$

Then the SM error can be written as

$$\begin{aligned} e(n+1) &= \boldsymbol{q}_d^T(n+1) \begin{bmatrix} d^f(n) \\ \vdots \\ d^f(n-N) \end{bmatrix} - \boldsymbol{b}^T(n+1) \begin{bmatrix} x^f(n) \\ \vdots \\ x^f(n-N) \end{bmatrix} \\ &= \boldsymbol{q}_d^T(n+1)\boldsymbol{Q}_d(n) \begin{bmatrix} d^f(n-1) \\ \vdots \\ d^f(n-N) \\ d(n) \end{bmatrix} - \boldsymbol{b}^T(n+1) \begin{bmatrix} x^f(n) \\ \vdots \\ x^f(n-N) \end{bmatrix} \end{aligned}$$

Then, the Steiglitz-McBride method can be realized based on the following properties of $\boldsymbol{Q}_d$ matrix

- $\boldsymbol{Q}_d$ is invertible.
- $\boldsymbol{Q}_d$ is uniquely determined from N free parameters.
- This parameters can be uniquely deduced from the last row of $\boldsymbol{Q}_d^{-1}(= \boldsymbol{Q}_d^T)$.

Remembering chapter 4, this can be obtained, for example, for the lattice realization and the general orthonormal realization.

For the lattice realization, $\boldsymbol{Q}_l = \boldsymbol{Q}_1\boldsymbol{Q}_2...\boldsymbol{Q}_N$ and

$$\boldsymbol{Q}_k = \begin{bmatrix} \boldsymbol{I}_{k-1} & & & \\ & -\sin\theta_k & \cos\theta_k & \\ & \cos\theta_k & \sin\theta_k & \\ & & & \boldsymbol{I}_{N-k} \end{bmatrix}$$

then by chosen the respective filtered variables,

$$\begin{aligned} \begin{bmatrix} \boldsymbol{r}(n+1) \\ w(n) \end{bmatrix} &= \boldsymbol{Q}_l(\theta(n)) \begin{bmatrix} \boldsymbol{r}(n) \\ x(n) \end{bmatrix} \\ \begin{bmatrix} \boldsymbol{u}(n+1) \\ s(n) \end{bmatrix} &= \boldsymbol{Q}_l(\theta(n)) \begin{bmatrix} \boldsymbol{u}(n) \\ d(n) \end{bmatrix} \end{aligned}$$

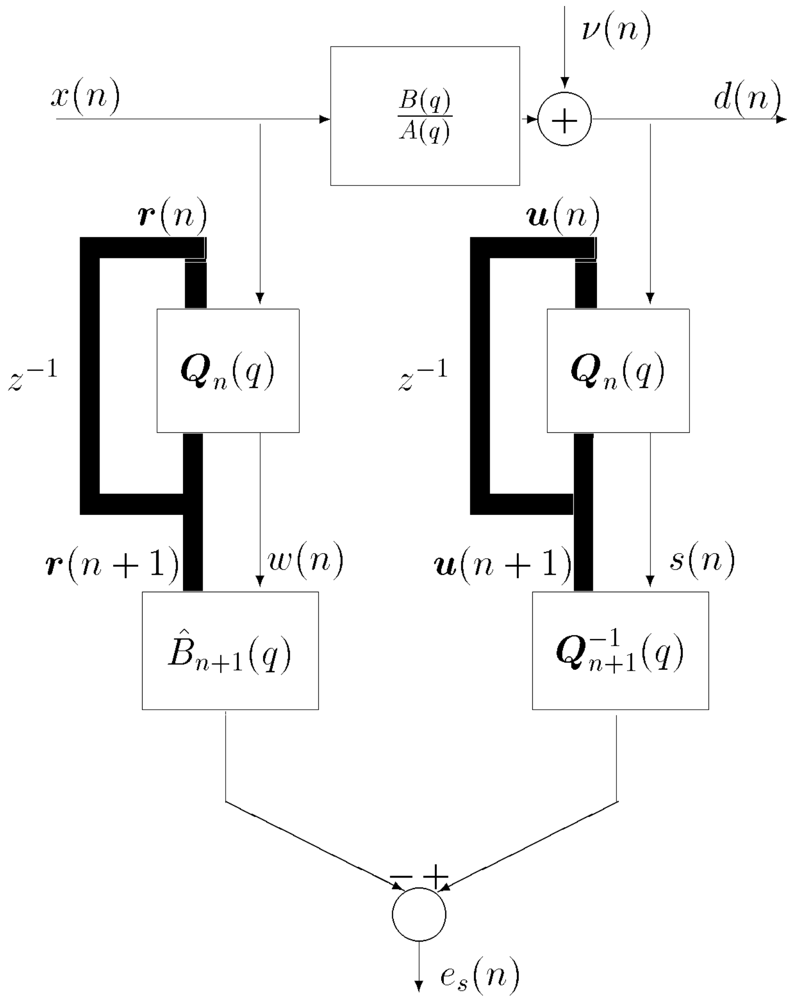

Figure 55: Structural interpretation of the SM method.

The error is

$$
\begin{aligned}
e(n+1) &= \boldsymbol{q}_l^T(n+1)\begin{bmatrix} \boldsymbol{u}(n+1) \\ s(n) \end{bmatrix} - \boldsymbol{\nu}^T(n+1)\begin{bmatrix} \boldsymbol{r}(n+1) \\ x(n) \end{bmatrix} \\
&= \boldsymbol{q}_l^T(n+1)\boldsymbol{Q}_l(n)\begin{bmatrix} \boldsymbol{u}(n) \\ s(n) \end{bmatrix} - \boldsymbol{\nu}^T(n+1)\begin{bmatrix} \boldsymbol{r}(n+1) \\ x(n) \end{bmatrix}
\end{aligned}
$$

where

$$\boldsymbol{q}(\theta) = \boldsymbol{Q}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \Pi_{i=1}^{M}\cos\theta_i \\ \sin\theta_1 \Pi_{i=2}^{M}\cos\theta_i \\ \vdots \\ \sin theta_{N-1}\cos\theta_N \\ \sin\theta_N \end{bmatrix}$$

then, in terms of the lattice parameters the SM method reads as

$$J = [\boldsymbol{q}^T(n+1) \quad -\boldsymbol{\nu}^T(n+1)]\begin{bmatrix} \boldsymbol{R}_u(\theta(n)) & \boldsymbol{R}_{ru}{}^T(\theta(n)) \\ \boldsymbol{R}_{ru}(\theta(n)) & \boldsymbol{R}_r(\theta(n)) \end{bmatrix}\begin{bmatrix} \boldsymbol{q}(n+1) \\ -\boldsymbol{\nu}(n+1) \end{bmatrix}$$

Minimizing J with respect to $\boldsymbol{\nu}$ results in $\boldsymbol{R}_r\boldsymbol{\nu}(n+1) = \boldsymbol{R}_{ru}\boldsymbol{q}(n+1)$, that substituted in the previous equation gives

$$\begin{aligned} J_r &= \boldsymbol{q}^T(n+1)\left[\boldsymbol{R}_u(\theta(n)) - \boldsymbol{R}_{ru}{}^T(\theta(n))\boldsymbol{R}_r^{-1}(\theta(n))\boldsymbol{R}_{ru}(\theta(n))\right]\boldsymbol{q}(n+1) \\ &= \boldsymbol{q}^T(n+1)\boldsymbol{R}_{u/r}(\theta(n))\boldsymbol{q}(n+1) \end{aligned}$$

Since, $\boldsymbol{q}^T\boldsymbol{q} = 1$, this is equivalent to

$$J_r = \frac{\boldsymbol{q}^T(n+1)\boldsymbol{R}_{u/r}(\theta(n))\boldsymbol{q}(n+1)}{\boldsymbol{q}^T(n+1)\boldsymbol{q}(n+1)}$$

that represents the *Rayleigh quotient* related to $\boldsymbol{R}_{u/r}$, and in particular

$$\frac{\boldsymbol{q}^T(n+1)\boldsymbol{R}_{u/r}(\theta(n))\boldsymbol{q}(n+1)}{\boldsymbol{q}^T(n+1)\boldsymbol{q}(n+1)} \geq \lambda_{min}(\boldsymbol{R}_{u/r})$$

Theorem: Convergence points of the SM method. Let $\theta(n+1) = \theta(n)$ denote any convergent point (if one exists). Then

$$\boldsymbol{R}_{u/r}(\theta(n))\boldsymbol{q}(n+1) = \lambda_{min}(\boldsymbol{R}_{u/r})\boldsymbol{q}(n+1)$$

The resulting output error has variance

$$E\{(d(n) - \hat{H}(z)x(n))^2\} = \lambda_{min}(\boldsymbol{R}_{u/r})$$

To relate this result with the direct-form realization

Theorem: Let $\hat{H}(z)$ be a stationary point of the direct-form SM.

1. If $\boldsymbol{\nu}$ and $\boldsymbol{\theta}$ are the lattice parameters of $\hat{H}(z)$, then

$$\begin{aligned} \boldsymbol{R}_{u/r}(\boldsymbol{\theta})\boldsymbol{q}(\boldsymbol{\theta}) &= \lambda\boldsymbol{q}(\boldsymbol{\theta}) \\ \boldsymbol{R}_r(\boldsymbol{\theta})\boldsymbol{\nu} &= \boldsymbol{R}_{ru}\boldsymbol{q}(\boldsymbol{\theta}) \end{aligned}$$

   for some eigenvalue λ of $\boldsymbol{R}_{u/r}$. Conversely, if $\boldsymbol{\theta}$ and $\boldsymbol{\nu}$ are compatible with these equations, then the resulting $\hat{H}(z)$ is a stationary point of the direct-form iteration as well.

2. If the input $x(n)$ is white noise, the eigenvalue λ from the first part is necessarily an extremal eigenvalue of $\boldsymbol{R}_{u/r}$.

Remarks

- λ is not an extremal eigenvalue in the first part of the theorem. Then any stationary points of the direct-form realization is not translated to a stationary point of the lattice realization.
- By contrast in the second part, and only for white noise input, the unique eigenvalue is an extremal one, an the mapping conversion leads to coincident stationary points for both realizations.

8.3.2 An a priori error bound

If $\boldsymbol{P}_1 > 0$ and $\boldsymbol{P}_2 > 0$, then $\lambda_{m+n+1}[\boldsymbol{P}_1 + \boldsymbol{P}_2] \leq \lambda_m[\boldsymbol{P}_1] + \lambda_n[\boldsymbol{P}_2]$.

Then, with $\boldsymbol{P}_1 = \boldsymbol{R}_{u/r}|_{d(n)=H(z)x(n)}$ and $\boldsymbol{P}_2 = \boldsymbol{R}_u|_{d(n)=\nu(n)}$, $\boldsymbol{R}_{u/r} = \boldsymbol{P}_1 + \boldsymbol{P}_2$ and

$$\lambda_{min}[\boldsymbol{R}_{u/r}] \;=\; \lambda_{N+1}[\boldsymbol{R}_{u/r}] \leq \lambda_{N+1}[\boldsymbol{P}_1] + \lambda_1[\boldsymbol{P}_2]$$

To find an eigenvalue bound of $\lambda_{min}[\boldsymbol{R}_{u/r}]$ it is equivalent to find an eigenvalue bound for the signal induced part and for the noise induced part.

For the signal induced part and $x(n)$ white.

$\boldsymbol{u}(n+1)$ and $\boldsymbol{r}(n+1)$ in terms of the orthogonal lattice filter transfer functions $F_k(z)$,

$$\begin{bmatrix} \boldsymbol{u}(n+1) \\ s(n) \end{bmatrix} = \begin{bmatrix} F_0(z) \\ \vdots \\ F_{N-1}(z) \\ F_N(z) \end{bmatrix} H(z)x(n)$$

$$\begin{bmatrix} \boldsymbol{r}(n+1) \\ w(n) \end{bmatrix} = \begin{bmatrix} F_0(z) \\ \vdots \\ F_{N-1}(z) \\ F_N(z) \end{bmatrix} x(n)$$

and define the following orthogonal expansion in $\mathcal{H}_2$,

$$F_l(z)H(z) = \sum_k^{\infty} \tilde{h}_{l,k} F_k(z) \qquad \tilde{h}_{l,k} =< F_l(z)H(z), F_k(z) >$$

in such a way that: $< F_l(z)H(z), F_i(z)H(z) >= \sum_{k=0}^{\infty} \tilde{h}_{l,k}\tilde{h}_{i,k}$.

Then

$$\begin{aligned}
\boldsymbol{R}_u &= E\left\{\left[\begin{array}{c}\boldsymbol{u}(n+1)\\ s(n)\end{array}\right]\left[\begin{array}{c}\boldsymbol{u}(n+1)\\ s(n)\end{array}\right]^T\right\}\\
&= \left[\begin{array}{ccc} < F_0(z)H(z), F_0(z)H(z) > & \cdots & < F_0(z)H(z), F_N(z)H(z) > \\ \vdots & \ddots & \vdots \\ < F_N(z)H(z), F_0(z)H(z) > & \cdots & < F_N(z)H(z), F_N(z)H(z) > \end{array}\right]\\
&= \left[\begin{array}{cccccc} \tilde{h}_{0,0} & \tilde{h}_{0,1} & \cdots & \tilde{h}_{0,N} & \tilde{h}_{0,N+1} & \cdots \\ \tilde{h}_{1,0} & \tilde{h}_{1,1} & \cdots & \tilde{h}_{1,N} & \tilde{h}_{1,N+1} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \tilde{h}_{N,0} & \tilde{h}_{N,1} & \cdots & \tilde{h}_{N,N} & \tilde{h}_{N,N+1} & \cdots \end{array}\right][.]^T
\end{aligned}$$

similarly

$$\begin{aligned}
\boldsymbol{R}_{ru}^T &= E\left\{\left[\begin{array}{c}\boldsymbol{u}(n+1)\\ s(n)\end{array}\right]\left[\begin{array}{c}\boldsymbol{r}(n+1)\\ w(n)\end{array}\right]^T\right\}\\
&= \left[\begin{array}{ccc} < F_0(z)H(z), F_0(z) > & \cdots & < F_0(z)H(z), F_N(z) > \\ \vdots & \ddots & \vdots \\ < F_N(z)H(z), F_0(z) > & \cdots & < F_N(z)H(z), F_N(z) > \end{array}\right]\\
&= \left[\begin{array}{cccc} \tilde{h}_{0,0} & \tilde{h}_{0,1} & \cdots & \tilde{h}_{0,N} \\ \tilde{h}_{1,0} & \tilde{h}_{1,1} & \cdots & \tilde{h}_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{h}_{N,0} & \tilde{h}_{N,1} & \cdots & \tilde{h}_{N,N} \end{array}\right][.]^T
\end{aligned}$$

and $\boldsymbol{R}_r = \boldsymbol{I}_{N+1}$. This results in

$$\begin{aligned}
\boldsymbol{R}_{u/r} &= \boldsymbol{R}_u - \boldsymbol{R}_{ru}^T\boldsymbol{R}_r^{-1}\boldsymbol{R}_{ru}\\
&= \left[\begin{array}{cccc} \tilde{h}_{0,N+1} & \tilde{h}_{0,N+2} & \tilde{h}_{0,N+1} & \cdots \\ \tilde{h}_{1,N+1} & \tilde{h}_{1,N+2} & \tilde{h}_{1,N+1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \tilde{h}_{N,N+1} & \tilde{h}_{N,N+2} & \tilde{h}_{N,N+1} & \cdots \end{array}\right][.]^T
\end{aligned}$$

It is interesting now to rewrite the previous equation pursuing a form similar to a Hankel matrix. To do this, consider that

$$\begin{aligned}\tilde{h}_{l,N+k} &= < F_l(z)H(z), F_{N+k}(z) > \\ &= < F_l(z)H(z), z^{-k}V(z) > \\ &= < H(z), z^{-k}F_l(z^{-1})V(z) > \\ &= < H(z), z^{-k}\hat{F}_l(z) >\end{aligned}$$

where in particular $\hat{F}_l(z) = \frac{\overline{D}_l(z^{-1})}{D_N(z^{-1})}\frac{z^{-N}D_N(z^{-1})}{D_N(z)} = \frac{z^{-N}\overline{D}_l(z^{-1})}{D_N(z)}$. Since these are causal functions (with the first $N-l$ coefficients of $z^{-N}\overline{D}_l(z^{-1})$ equal to zero), it can be expanded as

$$\hat{F}_l(z) = \sum_k^{\infty} \hat{f}_{l,k} z^{-k}$$

in such a way that we can write: $\tilde{h}_{l,N+k} =< H(z), z^{-k}\hat{F}_l(z) >= h_k\hat{f}_{l,0} + h_{k+1}\hat{f}_{l,1} + h_{k+2}\hat{f}_{l,2} + \cdots$. Using this

$$\begin{aligned}\boldsymbol{R}_{u/r} &= \boldsymbol{R}_u - \boldsymbol{R}_{ru}^T\boldsymbol{R}_r^{-1}\boldsymbol{R}_{ru} \\ &= \begin{bmatrix} \hat{f}_{0,0} & \hat{f}_{0,1} & \hat{f}_{0,2} & \cdots \\ \hat{f}_{1,0} & \hat{f}_{1,1} & \hat{f}_{1,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \hat{f}_{N,0} & \hat{f}_{N,1} & \hat{f}_{N,2} & \cdots \end{bmatrix} \begin{bmatrix} h_1 & h_2 & h_3 & \cdots \\ h_2 & h_3 & h_4 & \cdots \\ h_3 & h_4 & h_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \boldsymbol{\Gamma}_H \boldsymbol{C}^T \\ &= \boldsymbol{C}\boldsymbol{\Gamma}_H^2\boldsymbol{C}^T\end{aligned}$$

considering that $\boldsymbol{C}$ has orthonormal rows and using the fact that for a symmetric positive definite $\boldsymbol{R}$, then $\boldsymbol{P} = \boldsymbol{C}\boldsymbol{R}\boldsymbol{C}^T$, where $\boldsymbol{C}$ has orthonormal rows then

$$\lambda_k[\boldsymbol{P}] \leq \sigma_k[\boldsymbol{R}]$$

we can arrive to the following result

$$\lambda_{min}[\boldsymbol{R}_{u/r}] \leq \sigma_{N+1}(\boldsymbol{\Gamma}_H)$$

Considering in general the noise induced term and the signal induced term we have the following

Lemma (eingenvalue bounds). Suppose $\boldsymbol{\theta}$ such that $|\theta_k| < \pi/2$, then

- The noise induced component to $\boldsymbol{R}_{u/r}$ satisfies: $\lambda_1[\boldsymbol{R}_r] \leq sup_w \mathcal{S}_\nu(e^{jw})$.
- If the input is white noise, the signal induced component to $\boldsymbol{R}_{u/r}$ satisfies: $\lambda_{min}[\boldsymbol{R}_{u/r}] \leq E\{x^2(n)\}\sigma_{N+1}(\boldsymbol{\Gamma}_H)$

Based on this result

Theorem: Suppose the input $x(n)$ is white noise, and let $\hat{H}(z)$ be the Nth-order transfer function obtained at any stationary point of the SM method. Then

$$\|H(z) - \hat{H}(z)\|_2^2 \;\leq\; \sigma_{N+1} + \frac{\max_w \mathcal{S}_\nu(e^{jw}) - E\{\nu^2(n)\}}{E\{x^2(n)\}}$$

Note that, in the noise free case for example, this reduces to $\|H(z) - \hat{H}(z)\|_2 \leq \sigma_{N+1}$. From chapter 5 we know that for MSOE minimization, the global minimum satisfies $\|H(z) - \hat{H}_{L_2}(z)\|_2 \leq \sigma_{N+1}$, then

$$\begin{aligned}
\|\hat{H}_{L_2}(z) - \hat{H}(z)\|_2 &= \|[H(z) - \hat{H}(z)] - [H(z) - \hat{H}_{L_2}(z)]\|_2 \\
&\leq \|[H(z) - \hat{H}(z)]\|_2 + \|[H(z) - \hat{H}_{L_2}(z)]\|_2 \leq 2\sigma_{N+1}
\end{aligned}$$

8.4 On-line realizations

8.4.1 Lattice realization

Based on the error

$$e(n+1) = \boldsymbol{q}^T(n+1)\begin{bmatrix}\boldsymbol{u}(n+1)\\ s(n)\end{bmatrix} - \boldsymbol{\nu}^T(n+1)\begin{bmatrix}\boldsymbol{r}(n+1)\\ x(n)\end{bmatrix}$$

then an algorithm for the coefficients updating reads as

$$\begin{aligned}\nu_k(n+1) &= \nu_k(n) + \mu e(n)\nabla_{\nu_k}(n)\\ \theta_k(n+1) &= \theta_k(n) + \mu e(n)\nabla_{\theta_k}(n)\end{aligned}$$

Consider $\begin{bmatrix}\nabla_{\theta_1}(n)\\ \vdots\\ \nabla_{\theta_N}(n)\end{bmatrix} = \begin{bmatrix}\delta_1\boldsymbol{q}^T\\ \vdots\\ \delta_N\boldsymbol{q}^T\end{bmatrix}\begin{bmatrix}\boldsymbol{u}(n+1)\\ s(n)\end{bmatrix}$
and the fact that the orthogonal Hessemberg $\boldsymbol{Q}^T(\theta) = [\boldsymbol{q}_1^T, ..., \boldsymbol{q}_{N+1}^T]^T$ matrix verifies

$$\boldsymbol{Q}^T(\theta) = \begin{bmatrix}\delta_1\boldsymbol{q}_{N+1}^T(\theta)/\gamma_1\\ \vdots\\ \delta_N\boldsymbol{q}_{N+1}^T(\theta)/\gamma_N\\ \boldsymbol{q}_{N+1}^T(\theta)\end{bmatrix}$$

where $\gamma_k = \prod_{l=k+1}^{N}\cos\theta_l$, $\gamma_N = 1$, then

$$\begin{bmatrix}r_0\\ \vdots\\ r_N\\ r_{N+1}\end{bmatrix} = \boldsymbol{Q}^T\begin{bmatrix}\boldsymbol{u}(n+1)\\ s(n)\end{bmatrix}$$

with $e(n) = r_{N+1} - \boldsymbol{\nu}^T\begin{bmatrix}\boldsymbol{r}(n+1)\\ w(n)\end{bmatrix}$ and

$$\begin{bmatrix}\nabla_{\theta_1}(n)\\ \vdots\\ \nabla_{\theta_N}(n)\end{bmatrix} = \begin{bmatrix}\gamma_1 & & & \\ & \gamma_2 & & \\ & & \ddots & \\ & & & \gamma_N\end{bmatrix}\begin{bmatrix}r_0\\ \vdots\\ r_N\end{bmatrix}$$

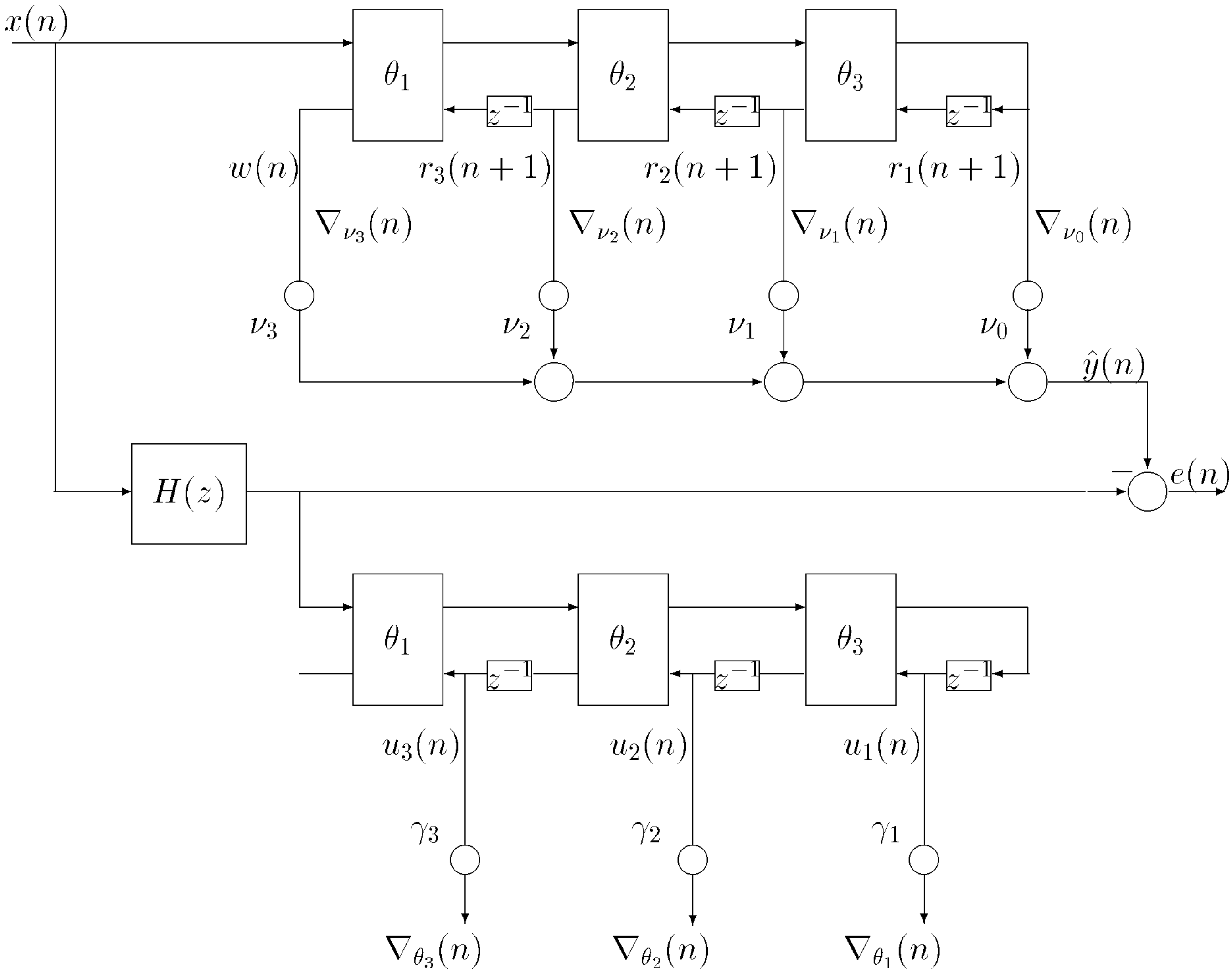

Figure 56: SM lattice realization

8.4.2 Orthonormal realization

For the particular case of real poles, the basis functions is

$$F_k(q) = \frac{\alpha_k}{1 - a_k q^{-1}} \prod_{n=1}^{k-1} \frac{q^{-1} - a_n}{1 - a_n q^{-1}} \tag{92}$$

where a_k is the k-th pole and $\alpha_k = \sqrt{1 - a_k^2}$ is a normalization constant.

- Each function can be represented in a recursive form

$$\begin{pmatrix} F_i(n) \\ r_i(n) \end{pmatrix} = \boldsymbol{Q}_i \begin{pmatrix} F_i(n-1) \\ u_i(n) \end{pmatrix}$$
$$\boldsymbol{Q}_i = \begin{pmatrix} \boldsymbol{A}_i & \boldsymbol{b}_i \\ \boldsymbol{g}_i & \rho_i \end{pmatrix} = \begin{pmatrix} a_i & \alpha_i \\ \alpha_i & -a_i \end{pmatrix}$$

  where $u_i(n)$ is the input to the i-th section and $r_i(n)$ is an auxiliary variable chosen to be orthonormal to $F_i(n)$, so that $\boldsymbol{Q}_i$ will be orthogonal, i.e. $\boldsymbol{Q}_i \boldsymbol{Q}_i^T = \boldsymbol{I}$.

- Also, with only one orthogonal matrix $\boldsymbol{Q}_o$

$$\begin{pmatrix} \boldsymbol{F}(n) \\ r(n) \end{pmatrix} = \boldsymbol{Q}_o \begin{pmatrix} \boldsymbol{F}(n) \\ u(n) \end{pmatrix}$$
$$\boldsymbol{Q}_o = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{g} & \rho \end{pmatrix}$$

  where

$$\boldsymbol{F}(n) = [F_1(n), F_2(n), ..., F_N(n)]^T \tag{93}$$

  and $r(n)$ is an auxiliary variable chosen to be orthonormal to the $F_i(n)$, $i = 1, ..., N$.

The matrix $\boldsymbol{Q}_o$ can be constructed in a recursive way from the matrices $\boldsymbol{Q}_i$, starting from $\boldsymbol{Q}_1$ up to $\boldsymbol{Q}_N$. This construction is based in the following lemma.

Lemma The matrix $\boldsymbol{Q}_o$ obtained from the previous arrangement is orthogonal.

Then, the structural interpretation of the SM will be useful also in this case.

Consider the a priori error is $e(n) = d(n) - \hat{y}(n)$, where

$$\begin{aligned} e(n) &= \boldsymbol{q}_o^T(n) \begin{bmatrix} \boldsymbol{u}(n) \\ s(n) \end{bmatrix} - \boldsymbol{\nu}^T(n) \begin{bmatrix} \boldsymbol{r}(n) \\ x(n) \end{bmatrix} \\ &= \boldsymbol{q}_o^T(n) \boldsymbol{Q}_o(n) \begin{bmatrix} \boldsymbol{u}(n-1) \\ d(n) \end{bmatrix} - \boldsymbol{\nu}^T(n) \begin{bmatrix} \boldsymbol{r}(n) \\ x(n) \end{bmatrix} \end{aligned}$$

where, $\boldsymbol{Q}_o = [\boldsymbol{q}_1^T(n)\, \boldsymbol{q}_2^T(n) \, \cdots \, \boldsymbol{q}_N^T(n)]^T$.

Also,

$$d(n) = [00 \cdots 1] \boldsymbol{Q}_o^{-1} \begin{pmatrix} \boldsymbol{u}(n) \\ s(n) \end{pmatrix} \tag{94}$$

The coefficient update is given by

$$\boldsymbol{\theta}(n+1) = \boldsymbol{\theta}(n) - 2\mu e(n) \boldsymbol{\nabla}(n) \tag{95}$$

where $\boldsymbol{\theta}(n) = [\nu_1(n), \ldots, \nu_N(n), a_1(n), \ldots, a_N(n)]^T$, $\boldsymbol{\nabla}(n) = [\nabla_{\nu_1}, \ldots, \nabla_{\nu_N}, \nabla_{a_1}, \ldots, \nabla_{a_N}]^T$

$$\nabla_{\nu_i} = \frac{\partial e(n)}{\partial \nu_i(n)} = -F_i(z) x(n) \qquad \nabla_{a_i} = \frac{\partial e(n)}{\partial a_i(n)}$$

To obtain the gradient with respect to a_i is necessary to consider

$$\nabla_{a_i}(n) \;=\; [\delta_i \boldsymbol{q}_N^T(n)] \begin{pmatrix} \boldsymbol{u}(n) \\ s(n) \end{pmatrix}$$

where $\delta_i = \frac{\partial}{\partial a_i}$. Since $\boldsymbol{Q}_o$ can be expressed as

$$\boldsymbol{Q}_o^T \;=\; \begin{pmatrix} \lambda_1 \delta_1 \boldsymbol{q}_N^T(n) \\ \vdots \\ \lambda_{N-1} \delta_{N-1} \boldsymbol{q}_N^T(n) \\ \boldsymbol{q}_N^T(n) \end{pmatrix}$$

where $\lambda_i = (-1)^i \frac{\sqrt{1-a_i^2}}{\prod_{l=1}^{i-1} a_l}$, then using (94) and the previous equation the gradients with respect to the a_i coefficients can be written as

$$\begin{pmatrix} \nabla_{a_1}(n) \\ \vdots \\ \nabla_{a_{N-1}}(n) \end{pmatrix} \;=\; \begin{pmatrix} \delta_1 \boldsymbol{q}_N^T(n) \\ \vdots \\ \delta_{N-1} \boldsymbol{q}_N^T(n) \end{pmatrix} \begin{pmatrix} \boldsymbol{u}(n) \\ s(n) \end{pmatrix} \tag{96}$$

this implies that

$$\begin{pmatrix} \lambda_1 \nabla_{a_1}(n) \\ \vdots \\ \lambda_{N-1} \nabla_{a_{N-1}}(n) \\ d(n) \end{pmatrix} = \boldsymbol{Q}_o^T \begin{pmatrix} \boldsymbol{u}(n) \\ s(n) \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}(n-1) \\ d(n) \end{pmatrix}$$

From this equation we can conclude that the gradient of the a priori error with respect to each denominator coefficient is given by

$$\begin{aligned} -\frac{\partial e(n)}{\partial a_k} &= \lambda_k u_k(n-1) \\ &= \lambda_k F_k(z) d(n-1) \end{aligned}$$

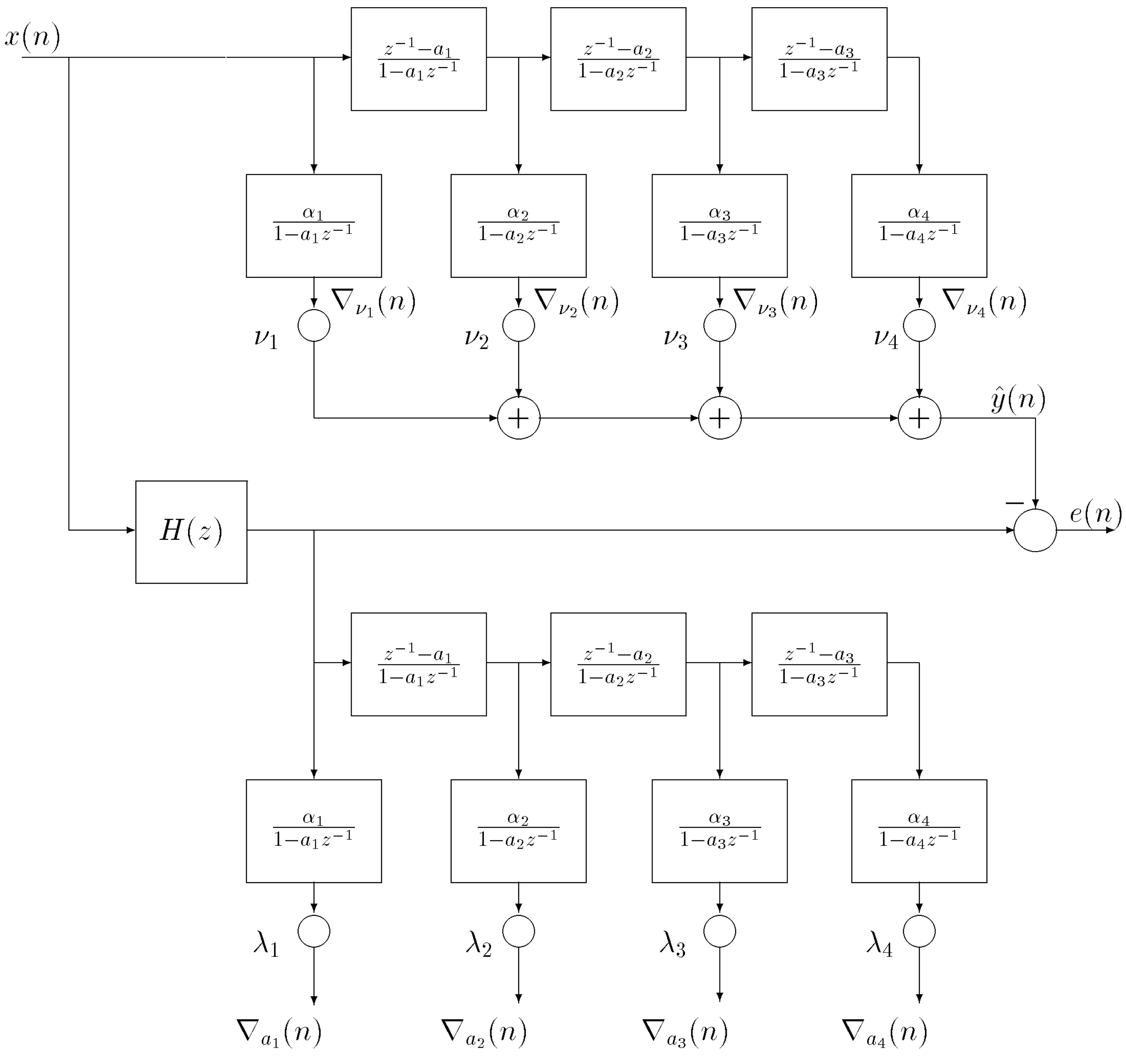

Figure 57: SM orthonormal realization with first-order sections

8.5 Some examples

Example 1:

$$d(n) = \frac{0.05 - 0.4q^{-1}}{1 - 1.1314q^{-1} + 0.25q^{-2}} x(n)$$

This is an example of *global convergence* of the SM method, that gives estimates that approximate closely the global minima obtained by the OE method.

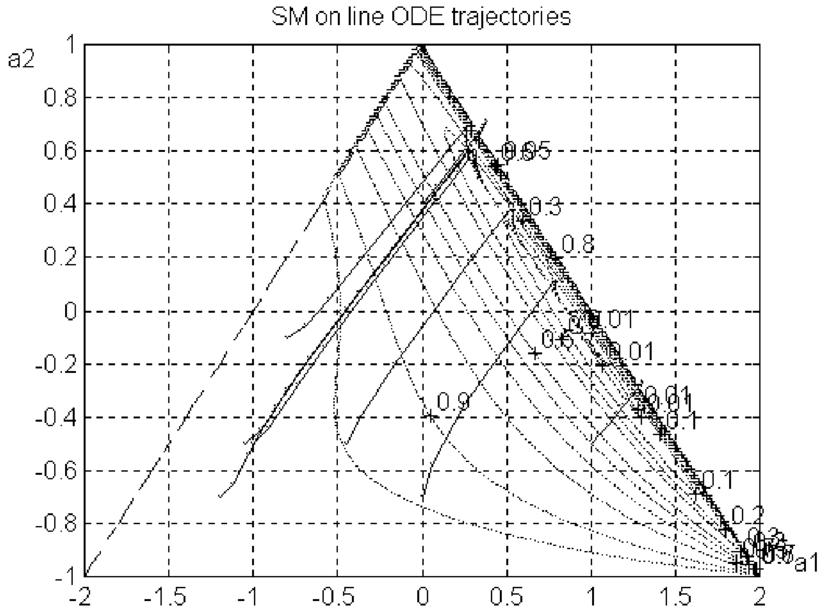

Figure 58: ODE trajectories of the SM method with sufficient order.

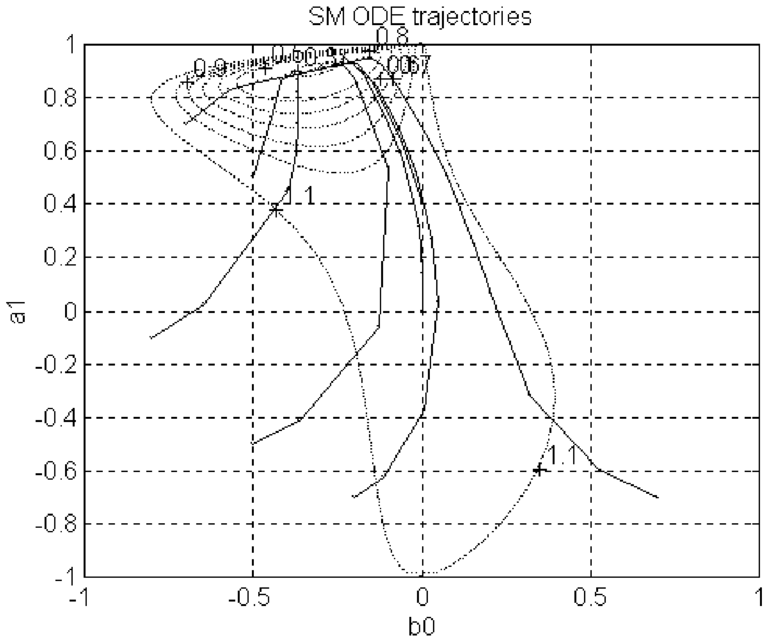

Figure 59: ODE trajectories of the SM method with $N = 1$, $M = 0$

Example 2:

$$d(n) = \frac{0.05 - 0.4\, q^{-1}}{(1 - 0.8303\, q^{-1})(1 + 0.83\, q^{-1})} x(n)$$

This example shows that multiple solutions can exist for the SM method in insufficient order cases.

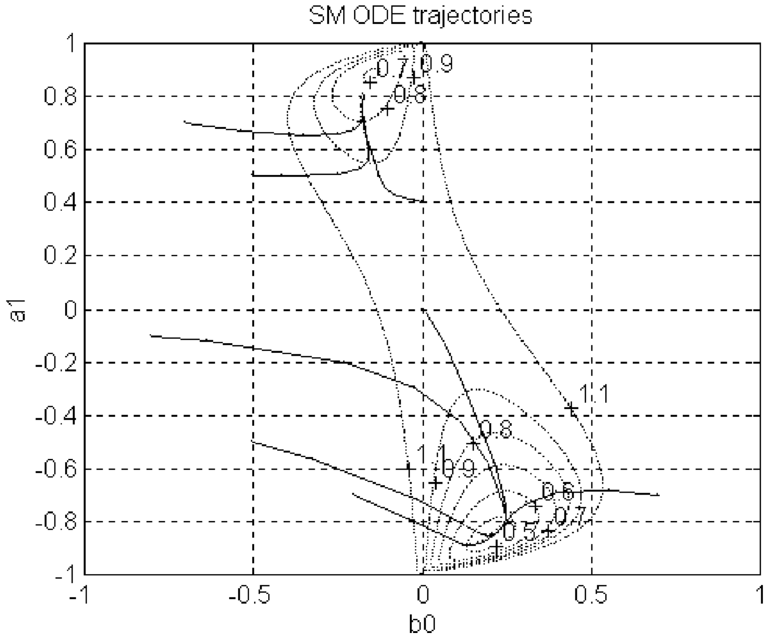

Figure 60: ODE trajectories of the SM method with $N = 1$, $M = 0$

With the purpose of study the *global behavior* of the SM method for insufficient order cases, some stationary points comparisons were performed in the following examples (with an IIR adaptive filter of orders $N = 1$, $M = 0$, i.e., when the stability can be guaranteed).

Example 3:

$$d(n) = \frac{0.05 - 0.4q^{-1}}{(1 - 0.8303\, q^{-1})(1 - p_2 q^{-1})} x(n)$$

for $-1.0 \geq p_2 \geq 1.0$.

For $p_2 = 0.3011$ or $p_2 = 0.83$, this example corresponds to examples 1 and 2.

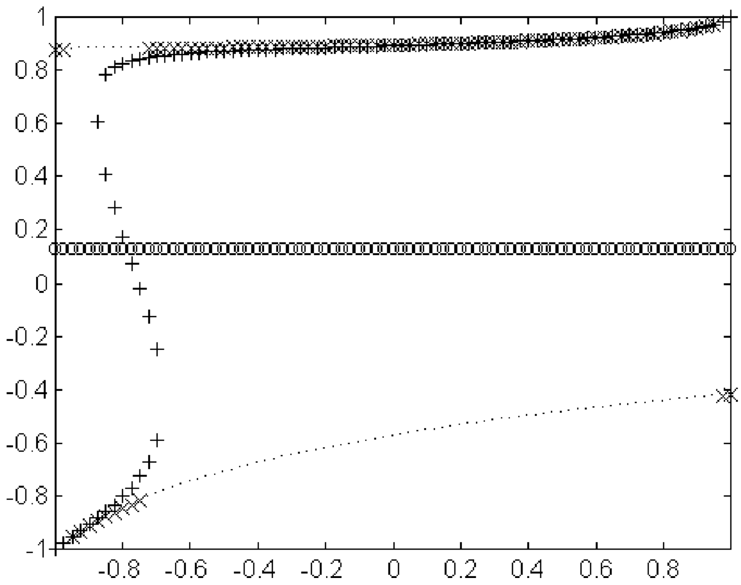

Figure 61: Stationary points for the SM and OE methods. $\times$ MSOE global minima, $\bullet$ MSOE local minima, $\circ$ degenerated point, $+$ stationary point of the SM method.

ODE of different realizations

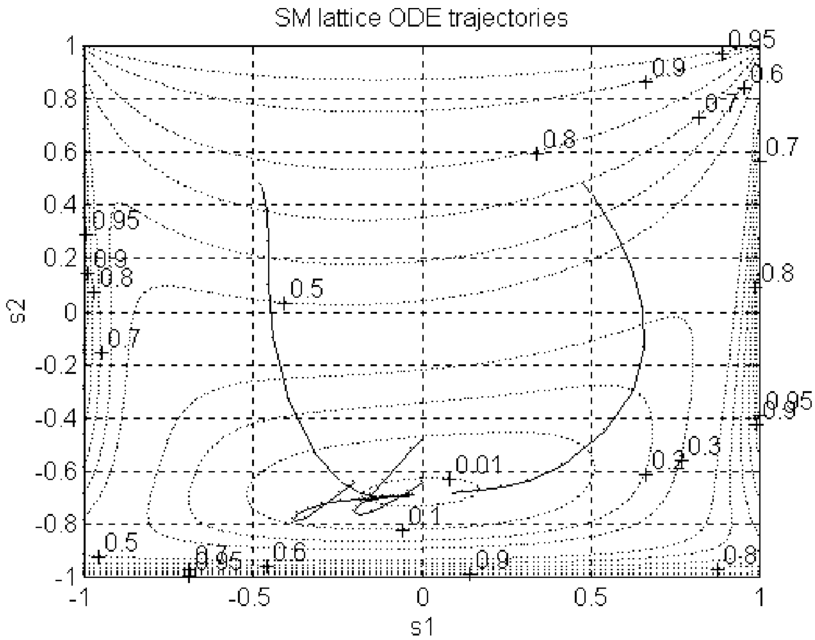

Figure 62: ODE of the Lattice SM algorithm (sufficient order), example 2.

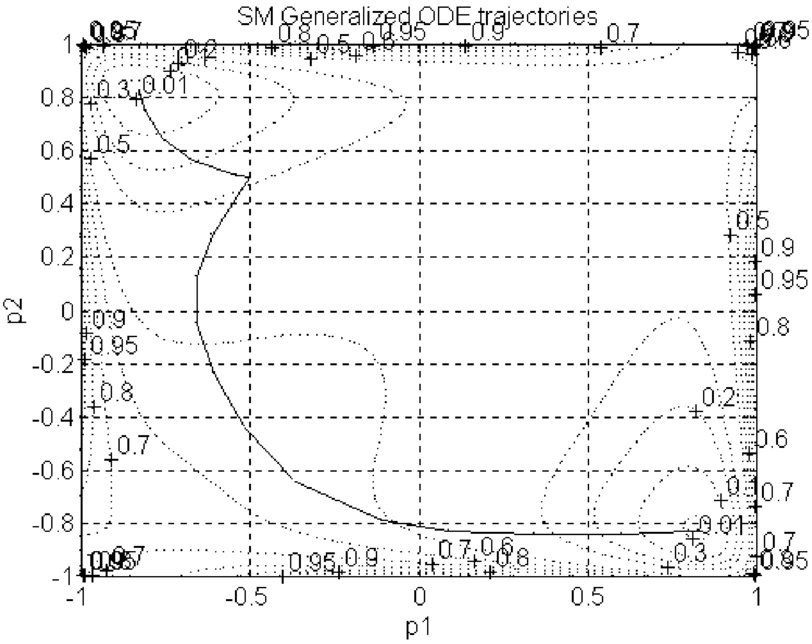

Figure 63: ODE of the Orthonormal SM algorithm (sufficient order), example 2.