ADAPTIVE IIR FILTERING: CONVERGENCE SPEED PROPERTIES IN THE UNDERMODELLED CASE

Phillip M. S. Burt *

Dept. of Telecomm. and Control Eng.
Escola Politécnica
Universidade de São Paulo
CEP 05508-900 São Paulo SP Brazil

Phillip A. Regalia

Dept. of Electrical Eng. and Computer Sci.
Catholic University of America
Washington, D.C., 20064

ABSTRACT

Previous results based on balanced realization theory and concerning the local convergence speed of adaptive IIR filters apply to the sufficient order case. In the undermodelled case, situations of greater practical interest are those in which the order chosen for the adaptive filter provides a good approximation of the system being modelled. A relevant question is then if the existence of a good approximation of the system implies a good approximation of the sufficient order convergence speed properties. We address this problem here, based on the same balanced realization theory framework. Our results suggest a positive answer to the question.

1. INTRODUCTION

Several aspects of adaptive IIR filters have been tackled over the years, leading to different adaptive algorithms and realization structures. We can mention, for instance, problems related to local minima, stability and the effect of poles close to the unit circle. Many references can be found in [1]−[3]. In [4, 5], we presented an analysis of the local convergence speed of adaptive IIR filters, based on balanced realization theory. From it, followed a new adaptive IIR algorithm. These results applied mostly to the sufficient order case.

In practice, however, the undermodelled case is more common. Not because the order of the adaptive filter is wronlgy chosen, but because the underlying physical system being modelled has a very high order, possibly infinite (meaning its transfer function is not rational). It is also true, though, that situations of greater practical interest are those in which the order chosen for the adaptive filter provides a good approximation of the system being modelled. A relevant question in these cases, then, is whether convergence speed properties derived for sufficient order can be applied with any confidence. In other words, does the existence of a good approximation of the system (according to a suitable criterion) imply a good approximation of the sufficient order convergence speed properties ? We address this problem here, based on the framework presented in [4, 5].

2. INITIAL ASSUMPTIONS

We consider an adaptive IIR filtering identification problem: a rational function $\widehat{H}(z)$ is adapted so as to minimize the mean square error between the output $\widehat{y}(n) = \widehat{H}(z)u(n)$ produced for a known white input $u(n)$ and the noisy output of a system $H(z)$ to the same input, $y(n) = H(z)u(n) + \eta(n)$. In this mixed notation, z is the unit-delay operator, with $zu(n) = u(n-1)$. Assuming additive noise $\eta(n)$ is independent of $u(n)$ makes the problem equivalent to the minimization of the L_2 norm $||H(z) - \widehat{H}(z)||$. We consider that in $\widehat{H}(z)$ we can vary M zeros, M poles and a gain. We say, therefore, that $\widehat{H}(z)$ is "of order M" .

The system being identified $H(z)$ is assumed to have order $N \geq M$ and to have the form

$$H(z) = H_m(z) + \delta H_d(z), \tag{1}$$

where $H_m(z)$ has order M. Undermodelling is due to the discrepancy $\delta H_d(z)$, for which we assume

$$||[H_d(z)]_+||_\infty = 1,$$

with $||[H_d(z)]_+||_\infty = \sup_\omega |\sum_{n=1}^{\infty} h_d(n)e^{-j\omega n}|$. These assumptions imply that $\min_{\widehat{H}(z)} ||H(z) - \widehat{H}(z)||$ is upper bounded, as we now verify.

A norm inequality true for any $F(z)$ is

$$||[F(z)]_+|| \leq ||\Gamma_F|| \leq ||[F(z)]_+||_\infty, \tag{2}$$

*Supported by FAPESP and CAPES.

where $||\Gamma_F||$ is the Hankel norm of $F(z)$. With $F(z) = H(z) - H_m(z)$, it follows that

$$||[F(z)]_+|| \leq ||[F(z)]_+||_\infty = \delta||[H_d(z)]_+||_\infty = \delta.$$

Now, the minimization of $||H(z) - \widehat{H}(z)||$ always leads to $\widehat{h}(0) = h(0)$ [3]. Making then $\widehat{h}(0) = h(0)$ and $\widehat{h}(n) = h_m(n)$, $n > 0$, we obtain $||H(z) - \widehat{H}(z)|| \leq ||H(z) - \widehat{H}(z)||_\infty = \delta$. Therefore, under the adopted assumptions, $\min_{\widehat{H}(z)} ||H(z) - \widehat{H}(z)|| \leq \min_{\widehat{H}(z)} ||H(z) - \widehat{H}(z)||_\infty \leq \delta$.

Properties of the convergence speed of IIR adaptive filters were obtained in [4, 5], mostly for the case of sufficient modelling ($N = M$). In the setting described above, this is achieved with $\delta = 0$, leading to $H(z) = H_m(z)$. Our aim here is to analyze how these properties extend to the undemodelled case for small values of δ.

3. REVIEW OF CONVERGENCE SPEED PROPERTIES

We review here the convergence speed properties of adaptive IIR filters obtained in [4, 5]. A new parameterization of the adaptive filter's poles was introduced: $\alpha_k \doteq \langle \zeta_k(z), V(z) \rangle$, $k = 1, 2, \ldots, N$ where $\langle .,. \rangle$ denotes the standard inner product, $\zeta_k(z)$ is the normalized transfer function between the input $u(n)$ and the $k-$th state variable $x_k(n+1)$ in a balanced realization of $H(z)$ and $V(z)$ is the M-order all-pass function with the same poles as the adaptive filter $\widehat{H}(z)$. The adaptation process of a chosen set of pole parameters w_j, $j = 1, 2, \ldots, M$ (for instance, direct form parameters or lattice parameters) can be locally described in terms of these new parameters as

$$\boldsymbol{\alpha}(n+1) \approx \left[\mathbf{I} - \mu \mathbf{J}\mathbf{J}^t \boldsymbol{\Sigma}^2\right] \boldsymbol{\alpha}(n),$$

where $\boldsymbol{\alpha}(n) = [\alpha_1 \ \ldots \ \alpha_N]^t$, $\boldsymbol{\Sigma}$ is the diagonal matrix of Hankel singular values of $H(z)$ and the elements of sensitivity matrix $\mathbf{J}$ are $[\mathbf{J}]_{kj} = \partial \alpha_k / \partial w_j$. We note that $\mathbf{J}$ depends on $\boldsymbol{\alpha}(n)$, which is not explicitely indicated for greater simplicity of the notation.

Local convergence speed depends on the eigenvalue spread $\chi(\mathbf{J}\mathbf{J}\boldsymbol{\Sigma}^2)$ of the M non-null eigenvalues of the $N \times N$ matrix $\mathbf{J}\mathbf{J}\boldsymbol{\Sigma}^2$. In summary its properties are:

1) When $H(z)$ is far from being all-pass ($\chi(\boldsymbol{\Sigma}) \gg 1$), convergence tends to be slow, irrespective of parameterization.

2.1) When $H(z)$ is close to all-pass ($\chi(\boldsymbol{\Sigma}) \approx 1$) and $\widehat{H}(z) \approx H(z)$ (sufficient order assumed), convergence for direct form parameters is fast if the poles of $H(z)$ are uniformly distributed and slow if they are concentrated.

2.2) Under the same conditions, convergence of lattice parameters is fast and less dependent on the pole distribution.

4. UNDERMODELLED CASE

On the assumptions made in Section 2 and based on the convergence speed properties discussed in Section 3, our aim is to analyze how the eigenvalue spread $\chi(\mathbf{J}\mathbf{J^t}\boldsymbol{\Sigma}^2)$ for the undermodelled case ($\delta > 0$ in (1)) relates to the eigenvalue spread $\chi(\mathbf{J}_m\mathbf{J}_m^t\boldsymbol{\Sigma}_m^2)$ for the associated sufficient order case ($\delta = 0$). We will in fact analyze the eigenvalues of the $M \times M$ matrix $\mathbf{J^t}\boldsymbol{\Sigma}^2\mathbf{J}$, which are the same as the non-null eigenvalues of $\chi(\mathbf{J}\mathbf{J^t}\boldsymbol{\Sigma}^2)$. We note that for any $\mathbf{X}$ and $\mathbf{Y}$ such that $\mathbf{XY}$ is square, for the non-null eigenvalues we have $\lambda(\mathbf{XY}) = \lambda(\mathbf{YX})$. This matrix property is repeatedly used in the following.

4.1. Discrepancy of controllability functions

As seen in Section 3, let $\zeta_k(z) = \sum_n \zeta_{kn} z^n$, $k = 1, \ldots, N$ be the normalized controllability functions of a balanced realization of $H(z)$. Let now $\zeta_{m,k}(z)$ be the ones related to $H_m(z)$, with $k = 1, 2, \ldots, M$. We define then the discrepancy functions $\Delta_k(z) \doteq \zeta_k(z) - \zeta_{m,k}(z)$, $k = 1, 2, \ldots, M$.

The norms $||\Delta_k(z)||$ and $||[H(z) - H_m(z)]_+||_\infty$ can be related, as described in the following. Defining the normalized infinite horizon controllability matrix $\mathcal{C}_M \doteq [\boldsymbol{\zeta}_1 \ \boldsymbol{\zeta}_2 \ \ldots \ \boldsymbol{\zeta}_M]^t$, where $\boldsymbol{\zeta}_k \doteq [\zeta_{k0} \ \zeta_{k1} \ \ldots]^t$, and likewise $\mathcal{C}_m \doteq [\boldsymbol{\zeta}_{m,1} \ \boldsymbol{\zeta}_{m,2} \ \ldots \ \boldsymbol{\zeta}_{m,M}]^t$, we can write

$$\begin{aligned} \sum_{k=1}^{M} \sigma_{m,k}^2 ||\Delta_k(z)||^2 \quad &= \quad \sum_{k=1}^{M} \sigma_{m,k}^2 ||\zeta_k(z) - \zeta_{m,k}(z)||^2 \\ &= \text{trace } [\boldsymbol{\Sigma}_m(\mathcal{C}_M - \mathcal{C}_m)(\mathcal{C}_M - \mathcal{C}_m)^t\boldsymbol{\Sigma}_m] \\ &\leq M\lambda_{\max}[(\mathcal{C}_M - \mathcal{C}_m)(\mathcal{C}_M - \mathcal{C}_m)^t\boldsymbol{\Sigma}_m^2], \qquad (3) \end{aligned}$$

where $\boldsymbol{\Sigma}_m = \text{diag}[\sigma_{m,1} \ \sigma_{m,2} \ \ldots \ \sigma_{m,M}]$ contains the Hankel singular values of $H_m(z)$ and $\lambda_{\max}[.]$ denotes the maximum eigenvalue of the argument.

Now, the Hankel form of $H(z)$ is given by $\Gamma_H = \mathcal{O}\boldsymbol{\Sigma}\mathcal{C}$ [3, p. 150], where $\mathcal{C} \doteq [\boldsymbol{\zeta}_1 \ \boldsymbol{\zeta}_2 \ \ldots \ \boldsymbol{\zeta}_N]^t$ and $\mathcal{O}$ is an infinite horizon observability matrix. Likewise, $\Gamma_{H_m} = \mathcal{O}_m\boldsymbol{\Sigma}_m\mathcal{C}_m$. At this point we make a necessary approximation, which is good for small values of δ:

$$\Gamma_H \approx \mathcal{O}_m\boldsymbol{\Sigma}_m\mathcal{C}_M.$$

From this, $\Gamma_{H-H_m} \approx \mathcal{O}_m\boldsymbol{\Sigma}_m(\mathcal{C}_M - \mathcal{C}_m)$. We have then

$$\begin{aligned} ||\Gamma_{H-H_m}||^2 \quad &= \quad \lambda_{\max}(\Gamma_{H-H_m}\Gamma_{H-H_m}^t) \\ &\approx \lambda_{\max}(\mathcal{O}_m\boldsymbol{\Sigma}_m[\mathcal{C}_M - \mathcal{C}_m][\mathcal{C}_M - \mathcal{C}_m]^t\boldsymbol{\Sigma}_m\mathcal{O}_m^t) \end{aligned}$$

$$= \lambda_{\max}([\mathcal{C}_M - \mathcal{C}_m][\mathcal{C}_M - \mathcal{C}_m]^t \mathbf{\Sigma}_m^2), \quad (4)$$

where we used $\mathcal{O}_m^t \mathcal{O}_m = \mathbf{I}$. Combining (2), (3) and (4) we arrive at

$$\sum_{k=1}^{M} \sigma_{m,k}^2 ||\Delta_k(z)||^2 \lesssim M||\Gamma_{H-H_m}||^2$$
$$\leq M||[H(z) - H_m(z)]_+||_\infty^2, \quad (5)$$

where $\lesssim$ indicates an approximate upper bound.

4.2. Sensitivity matrix partition

We make the partition $\mathbf{J}^t = [\mathbf{J}_M^t \ \mathbf{J}_R^t]$, where $\mathbf{J}_M$ is $M \times M$ and $\mathbf{J}_R$ is $R \times M$, $R = N - M$. We have, then $\mathbf{J}^t \mathbf{\Sigma}^2 \mathbf{J} = \mathbf{J}_M^t \mathbf{\Sigma}_M^2 \mathbf{J}_M + \mathbf{J}_R^t \mathbf{\Sigma}_R^2 \mathbf{J}_R$. For $\mathbf{J}_M$ we also have

$$[\mathbf{J}_M]_{k,j} = \langle \frac{\partial}{\partial w_j} V(z), \zeta_k(z) \rangle = [\mathbf{J}_m]_{k,j} + [\mathbf{J}_\Delta]_{k,j} \quad (6)$$

where $[\mathbf{J}_m]_{k,j} \doteq \langle \frac{\partial}{\partial w_j} V(z), \zeta_{m,k}(z) \rangle$ and $[\mathbf{J}_\Delta]_{k,j} \doteq \langle \frac{\partial}{\partial w_j} V(z), \Delta_k(z) \rangle$.

For greater notational simplicity, we define $\mathbf{K}_m \doteq \mathbf{\Sigma}_M \mathbf{J}_m$ and $\mathbf{K}_\Delta \doteq \mathbf{\Sigma}_M \mathbf{J}_\Delta$. An approximation for the upper bound of the norm of $\mathbf{K}_\Delta$ is obtained as follows. For a unit-norm $\mathbf{x}$ we have

$$\begin{aligned}
\mathbf{x}^t \mathbf{K}_\Delta \mathbf{K}_\Delta^t \mathbf{x} &= \sum_{j=1}^{M} \langle \frac{\partial}{\partial w_j} V(z), \sum_{k=1}^{M} x_k \sigma_k \Delta_k(z) \rangle^2 \\
&\leq \sum_{j=1}^{M} ||\frac{\partial}{\partial w_j} V(z)||^2 \, ||\sum_{k=1}^{M} x_k \sigma_k \Delta_k(z)||^2 \\
&\leq \sum_{j=1}^{M} ||\frac{\partial}{\partial w_j} V(z)||^2 \, \sum_{k=1}^{M} \sigma_k^2 ||\Delta_k(z)||^2,
\end{aligned}$$

where we used

$$||\sum_{k=1}^{M} x_k \Delta_k(z)||^2 = ||\sum_{k=1}^{M} x_k \sum_l \Delta_{k,l} z^l||^2 =$$
$$= ||\sum_l (\sum_{k=1}^{M} x_k \Delta_{k,l}) z^l||^2 = \sum_l (\sum_{k=1}^{M} x_k \Delta_{k,l})^2 \leq$$
$$\leq \sum_l (\sum_{k=1}^{M} x_k^2)(\sum_{k=1}^{M} \Delta_{k,l})^2 = \sum_l \sum_{k=1}^{M} \Delta_{k,l}^2 =$$
$$= \sum_{k=1}^{M} \sum_l \Delta_{k,l}^2 = \sum_{k=1}^{M} ||\Delta_k(z)||^2.$$

Now, since $\lambda_{\max}(\mathbf{K}_\Delta \mathbf{K}_\Delta^t) = \max_{||x||=1}(\mathbf{x}^t \mathbf{K}_\Delta \mathbf{K}_\Delta^t \mathbf{x})$, with (5) and assuming $\sigma_k \approx \sigma_{m,k}$, $k = 1, \ldots, M$, we arrive at

$$\lambda_{\max}(\mathbf{K}_\Delta \mathbf{K}_\Delta^t) \lesssim M \sum_{j=1}^{M} ||\frac{\partial V(z)}{\partial w_j}||^2 \, ||[H(z) - H_m(z)]_+||_\infty^2. \quad (7)$$

4.3. Eigenvalue bounds

From the partition of the sensitivity matrix follows

$$\lambda_{\max}(\mathbf{J}^t \mathbf{\Sigma}^2 \mathbf{J}) \leq \lambda_{\max}(\mathbf{J}_M^t \mathbf{\Sigma}_M^2 \mathbf{J}_M) +$$
$$+ \lambda_{\max}(\mathbf{J}_R^t \mathbf{\Sigma}_R^2 \mathbf{J}_R). \quad (8)$$

The second term of this sum is

$$\lambda_{\max}(\mathbf{J}_R^t \mathbf{\Sigma}_R^2 \mathbf{J}_R) = \lambda_{\max}(\mathbf{J}_R \mathbf{J}_R^t \mathbf{\Sigma}_R^2)$$
$$\leq \lambda_{\max}(\mathbf{J}_R \mathbf{J}_R^t) \sigma_{M+1}^2.$$

AAK theory gives $\sigma_{M+1}^2 \leq ||\Gamma_{H-H_m}||^2$. With (2), then,

$$\lambda_{\max}(\mathbf{J}_R^t \mathbf{\Sigma}_R^2 \mathbf{J}_R) \leq \lambda_{\max}(\mathbf{J}_R \mathbf{J}_R^t) \times$$
$$||[H(z) - H_m(z)]_+||_\infty^2. \quad (9)$$

In [5] it is shown that $\lambda(\mathbf{J}\mathbf{J}^t)$ depends only on the poles of $H(z)$. The same can be shown for an upper bound for $\lambda(\mathbf{J}_R \mathbf{J}_R^t)$, bound which is therefore constant with δ. Due to its growth with the square of the norm, when δ is small we can then discard $\lambda_{\max}(\mathbf{J}_R^t \mathbf{\Sigma}_R^2 \mathbf{J}_R)$.

Using (6) now in the first term of (8) gives

$$\lambda_{\max}(\mathbf{J}_M^t \mathbf{\Sigma}_M^2 \mathbf{J}_M) =$$

$$\begin{aligned}
&= \max_{||\mathbf{x}||=1} (\mathbf{x}^t \mathbf{K}_m^t \mathbf{K}_m \mathbf{x} + 2\mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_m \mathbf{x} + \mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_\Delta \mathbf{x}) \\
&\leq \max_{||\mathbf{x}||=1} (\mathbf{x}^t \mathbf{K}_m^t \mathbf{K}_m \mathbf{x} + 2|\mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_m \mathbf{x}| + \mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_\Delta \mathbf{x}) \\
&\leq \max_{||\mathbf{x}||=1} (||\mathbf{K}_m \mathbf{x}|| + ||\mathbf{K}_\Delta \mathbf{x}||)^2 \\
&\leq \left[\sqrt{\lambda_{\max}(\mathbf{K}_m \mathbf{K}_m^t)} + \sqrt{\lambda_{\max}(\mathbf{K}_\Delta \mathbf{K}_\Delta^t)} \right]^2.
\end{aligned}$$

For the minimum eigenvalue,

$$\lambda_{\min}(\mathbf{J}^t \mathbf{\Sigma}^2 \mathbf{J}) \geq \lambda_{\min}(\mathbf{J}_M^t \mathbf{\Sigma}_M^2 \mathbf{J}_M) =$$

$$\begin{aligned}
&= \min_{||\mathbf{x}||=1} (\mathbf{x}^t \mathbf{K}_m^t \mathbf{K}_m \mathbf{x} + 2\mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_m \mathbf{x} + \mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_\Delta \mathbf{x}) \\
&\geq \min_{||\mathbf{x}||=1} (\mathbf{x}^t \mathbf{K}_m^t \mathbf{K}_m \mathbf{x} - 2|\mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_m \mathbf{x}| + \mathbf{x}^t \mathbf{K}_\Delta^t \mathbf{K}_\Delta \mathbf{x}) \\
&\geq \min_{||\mathbf{x}||=1} (||\mathbf{K}_m \mathbf{x}|| - ||\mathbf{K}_\Delta \mathbf{x}||)^2.
\end{aligned}$$

Assuming $\lambda_{\max}(\mathbf{K}_\Delta \mathbf{K}_\Delta^t) \leq \lambda_{\min}(\mathbf{K}_m \mathbf{K}_m^t)$ gives then

$$\lambda_{\min}(\mathbf{J}_M^t \mathbf{\Sigma}_M^2 \mathbf{J}_M) \geq$$
$$\geq \left[\sqrt{\lambda_{\min}(\mathbf{K}_m \mathbf{K}_m^t)} - \sqrt{\lambda_{\max}(\mathbf{K}_\Delta \mathbf{K}_\Delta^t)} \right]^2.$$

Proceeding similarly, we obtain also the bounds

$$\lambda_{\max}(\mathbf{J}^t \mathbf{\Sigma}^2 \mathbf{J}) \geq \lambda_{\max}(\mathbf{J}_M^t \mathbf{\Sigma}_M^2 \mathbf{J}_M) \geq$$

$$\geq \left[\sqrt{\lambda_{\max}(\mathbf{K}_m\mathbf{K}_m^t)} - \sqrt{\lambda_{\max}(\mathbf{K}_\Delta\mathbf{K}_\Delta^t)}\right]^2$$

and $\lambda_{\min}(\mathbf{J}^t\mathbf{\Sigma}^2\mathbf{J}) \leq \lambda_{\min}(\mathbf{J}_M^t\mathbf{\Sigma}_M^2\mathbf{J}_M)+\lambda_{\max}(\mathbf{J}_R^t\mathbf{\Sigma}_R^2\mathbf{J}_R) \lesssim \lambda_{\min}(\mathbf{J}_M^t\mathbf{\Sigma}_M^2\mathbf{J}_M)$, with

$$\lambda_{\min}(\mathbf{J}_M^t\mathbf{\Sigma}_M^2\mathbf{J}_M) \leq$$

$$\leq \left[\sqrt{\lambda_{\min}(\mathbf{K}_m\mathbf{K}_m^t)} + \sqrt{\lambda_{\max}(\mathbf{K}_\Delta\mathbf{K}_\Delta^t)}\right]^2.$$

Combining these expressions with (7) and using $G = \sqrt{M\sum_{j=1}^{M}||\frac{\partial}{\partial w_j}V(z)||^2}$, we obtain the following bounds for the eigenvalue spread of $\mathbf{J}^t\Sigma^2\mathbf{J}$:

$$\chi(\mathbf{J}^t\Sigma^2\mathbf{J}) = \frac{\lambda_{\max}(\mathbf{J}^t\Sigma^2\mathbf{J})}{\lambda_{\min}(\mathbf{J}^t\Sigma^2\mathbf{J})} \lesssim$$

$$\lesssim \left[\frac{\sqrt{\lambda_{\max}(\mathbf{K}_m\mathbf{K}_m^t)} + G||[H(z)-H_m(z)]_+||_\infty}{\sqrt{\lambda_{\min}(\mathbf{K}_m\mathbf{K}_m^t)} - G||[H(z)-H_m(z)]_+||_\infty}\right]^2 \tag{10}$$

and

$$\chi(\mathbf{J}^t\Sigma^2\mathbf{J}) \gtrsim$$

$$\gtrsim \left[\frac{\sqrt{\lambda_{\max}(\mathbf{K}_m\mathbf{K}_m^t)} - G||[H(z)-H_m(z)]_+||_\infty}{\sqrt{\lambda_{\min}(\mathbf{K}_m\mathbf{K}_m^t)} + G||[H(z)-H_m(z)]_+||_\infty}\right]^2. \tag{11}$$

Provided G is itself bounded, as $||[H(z)-H_m(z)]_+||_\infty$ tends to zero these bounds tend uniformly to the eigenvalue spread of the sufficient modelling case. In relation to our initial problem, then, this means that a good approximation of the sufficient order convergence speed properties is obtained for sufficiently small values of $||[H(z)-H_m(z)]_+||_\infty$. We note that, in these expressions, G is a constant, since it depends only on the poles of $\widehat{H}(z)$. An analytical assessment of this dependency is a matter for future work.

5. NUMERICAL EXAMPLE

In (1) we consider that $H_m(z^{-1})$ is all-pass with uniformly distributed poles at $0.8\angle\pm 45$ and $0.8\angle\pm 135$. In this case, for direct form parameters, $\chi(\mathbf{J}_m^t\mathbf{\Sigma}_m^2\mathbf{J}_m) = 1$ at $\widehat{H}(z) = H_m(z)$ and convergence is fast for sufficient order. Function $H_d(z^{-1})$ is also all-pass but with more concentrated poles, at $0.8\angle\pm 5$ and $0,8\angle\pm 20$. For a sufficiently large δ, therefore, convergence would be slow. The true value of the eigenvalue spread $\chi(\mathbf{J}^t\mathbf{\Sigma}^2\mathbf{J})$ at $\widehat{H}(z) = H_m(z)$ as well as the bounds (10) and (11) were calculated for different values of δ. The results are in Figure 1, where the fact that $||[H(z)-H_m(z)]_+||_\infty = \delta$ is used.

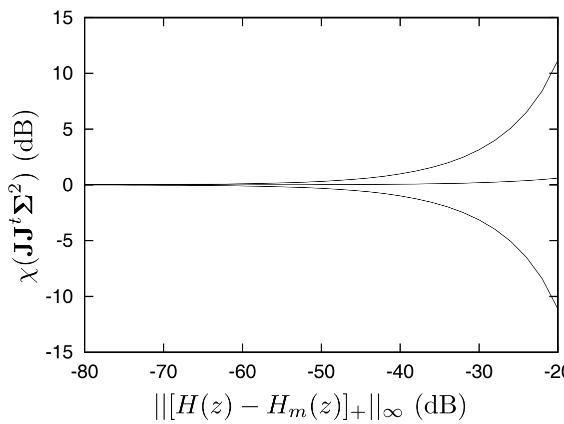

Fig. 1. True eigenvalue spread, lower and upper bounds.

6. CONCLUSION

Approximate upper and lower bounds were obtained for the eigenvalue spread that determines local convergence speed of adaptive IIR filters in the identification of a system $H(z)$, assuming an undermodelled setting. These bounds tend uniformly to the eigenvalue spread of the sufficient order identification of a system $H_m(z)$, as the norm $||[H(z)-H_m(z)]_+||_\infty$ tends to zero. This result suggests that when there is a good undemodelled approximation of a system, as is the case in situations of greater practical interest, then convergence speed properties are close to those derived under the sufficient order assumption.

7. REFERENCES

[1] J. Treichler, C. R. Johnson, Jr., and M. G. Larimore, *Theory and design of adaptive filters*, Prentice-Hall, Upper Saddle River, 2001.

[2] J. J. Shynk, "Adaptive IIR filtering," *IEEE ASSP Magazine*, vol. 6, no. 2, pp. 4–21, Apr. 1989.

[3] P. A. Regalia, *Adaptive IIR filtering in signal processing and control*, Marcel Dekker, New York, 1995.

[4] P. M. S. Burt and P. A. Regalia, "A new framework for convergence analysis and algorithm development of adaptive IIR filters," in *Proc. IEEE ICASSP*, Montreal, 2004, vol. 2, pp. 441–444.

[5] P. M. S. Burt and P. A. Regalia, "A new framework for convergence analysis and algorithm development of adaptive IIR filters," *Accepted for publication,IEEE Transactions on Signal Processing.*