# Enhanced Steiglitz-McBride Procedure for Minimax IIR Digital Filters

Wu-Sheng Lu
University of Victoria
Victoria, Canada

Takao Hinamoto
Hiroshima University
Higashi-Hiroshima, Japan

May 30, 2018

# Outline

- Early Work
- Problem Formulation
- The Steiglitz-McBride (SM) Procedure
- An Enhanced SM Procedure
- The Stability Issue
- Design Examples

# 1. Early Work

- K. Steiglitz and L. E. McBride, 1965. (System identification)
- P. Stoica and T. Soderstrom, 1981. (System identification)
- W.-S. Lu, S.-C. Pei, and C.-C. Tseng, 1998. (IIR filters)
- B. Dumitrescu and R. Niemisto, 2004. (IIR digital filters)
- X. Lai and Z. Lin, 2010. (IIR filters)

## 2. Problem Formulation

- Consider designing an IIR filter of order $(n, m)$ with transfer function

$$H(z) = \frac{b(z)}{a(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_m z^{-m}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}}$$

such that $H(e^{j\omega})$ best approximates a desired frequency response $H_d(\omega)$ in minimax sense over a frequency region of interest $\Omega$ subject to stability.

- The design problem can be formulated as finding parameter vectors $\boldsymbol{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^T$ and $\boldsymbol{b} = \begin{bmatrix} b_0 & b_1 & \cdots & b_m \end{bmatrix}^T$ by solving the constrained problem

$$\begin{aligned} &\underset{\boldsymbol{a},\, \boldsymbol{b}}{\text{minimize}} \quad \max_{\omega \in \Omega} \; W(\omega) \, | H(e^{j\omega}) - H_d(\omega) | \\ &\text{subject to: } \; H(z) \text{ stable} \end{aligned}$$

where $W(\omega) \geq 0$ is a known weighting function over $\Omega$.

## 3. The Steiglitz-McBride (SM) Procedure

- First, convert the constrained problem to

$$\underset{\boldsymbol{a},\,\boldsymbol{b},\,\beta}{\text{minimize}} \qquad \beta$$

$$\text{subject to:} \quad W(\omega)\,|\,H(e^{j\omega}) - H_d(\omega)\,| \le \beta \quad \text{for} \quad \omega \in \Omega$$

$$H(z) \text{ stable}$$

- Next, rewrite the first set of constraints as

$$W(\omega)\,|\,b(e^{j\omega}) - H_d(\omega)a(e^{j\omega})\,| \le \beta\,|\,a(e^{j\omega})\,| \qquad (1)$$

- In the $(k+1)$th SM iteration, the polynomial $a(e^{j\omega})$ on the right-hand side is set to the ***known*** $a_k(e^{j\omega})$ while optimizing the polynomials $a(e^{j\omega})$ and $b(e^{j\omega})$ on the left-hand side. In this way, the $(k+1)$th SM iteration solves the convex constrained problem

$$\underset{\boldsymbol{a}_{k+1},\,\boldsymbol{b}_{k+1},\,\beta_{k+1}}{\text{minimize}} \qquad \beta_{k+1}$$

$$\text{subject to:} \quad W(\omega)\,|\,b_{k+1}(e^{j\omega}) - H_d(\omega)a_{k+1}(e^{j\omega})\,| \le \beta_{k+1}\,|\,a_k(e^{j\omega})\,|$$

$$H(z) \text{ stable}$$

# 4. An Enhanced SM Procedure

- Suppose we don't set the polynomial $a(e^{j\omega})$ on the right-hand side of (1) to the ***known*** $a_k(e^{j\omega})$ but let it be an unknown $a_{k+1}(e^{j\omega})$, then in $(k+1)$th iteration (1) would become

$$W(\omega)\,|\,b_{k+1}(e^{j\omega}) - H_d(\omega)a_{k+1}(e^{j\omega})\,| \le \beta_{k+1}\,|\,a_{k+1}(e^{j\omega})\,| \qquad (2)$$

which is not convex because of the presence of term $|\,a_{k+1}(e^{j\omega})\,|$ on its right-hand side.

- Let $\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \boldsymbol{\delta}_a$ and $\boldsymbol{b}_{k+1} = \boldsymbol{b}_k + \boldsymbol{\delta}_b$. If we replace the right-hand side with its first-order approximation:

$$\beta_{k+1}\,|\,a_{k+1}(e^{j\omega})\,| \approx |\,a_k(e^{j\omega})\,|\,\beta_{k+1} + \beta_k \nabla_a^T\,|\,a_k(e^{j\omega})\,|\,\boldsymbol{\delta}_a$$

then a convex relaxation of (2) is obtained as

$$W(\omega)\,|\,b_{k+1}(e^{j\omega}) - H_d(\omega)a_{k+1}(e^{j\omega})\,| \le |\,a_k(e^{j\omega})\,|\,\beta_{k+1} + \beta_k \nabla_a^T\,|\,a_k(e^{j\omega})\,|\,\boldsymbol{\delta}_a \qquad (3)$$

- Note that if the conventional SM procedure were applied to (2), the right-hand side of its convex relaxation would be $|\,a_k(e^{j\omega})\,|\,\beta_{k+1}$. Since the second term on the right-side of (3) involves gradient $\nabla_a\,|\,a_k(e^{j\omega})\,|$, (3) may be regarded as a *first-order* SM procedure while the conventional SM is regarded as a *zeroth-order* procedure.

## 5. The Stability Issue

- For convenience of stability analysis, consider polynomial

$$\tilde{a}(z) = z^n + a_1 z^{n-1} + \cdots + a_n$$

which is equivalent to polynomial $a(z)$. $\tilde{a}(z)$ is called a Schur polynomial if its zeros are strictly inside the unit circle $C = \{z : |z| = 1\}$.

- Suppose $\tilde{a}(z)$ is a Schur polynomial and we want to examine the stability of polynomial $f(z) = \tilde{a}(z) + \tilde{\delta}(z)$ with $\tilde{\delta}(z)$ a perturbing polynomial.

**Rouche's Theorem**

$f(z)$ remains Schur if $\tilde{\delta}(z)$ obeys $|\tilde{\delta}(z)| < |\tilde{a}(z)|$ for all $z$ on $C$.

Rouche's theorem is well known in complex analysis and finds application for the design of stable IIR filters (Lang, 2000). What is not so well known is an enhanced variant of the theorem stated below.

**Theorem 1** If functions $f(z)$ and $\tilde{a}(z)$ are analytic inside and on $C$ and

$$|f(z) - \tilde{a}(z)| < |f(z)| + |\tilde{a}(z)| \qquad \text{for all } z \text{ on } C \qquad (4)$$

then $f(z)$ and $\tilde{a}(z)$ have the same number of zeros inside $C$.

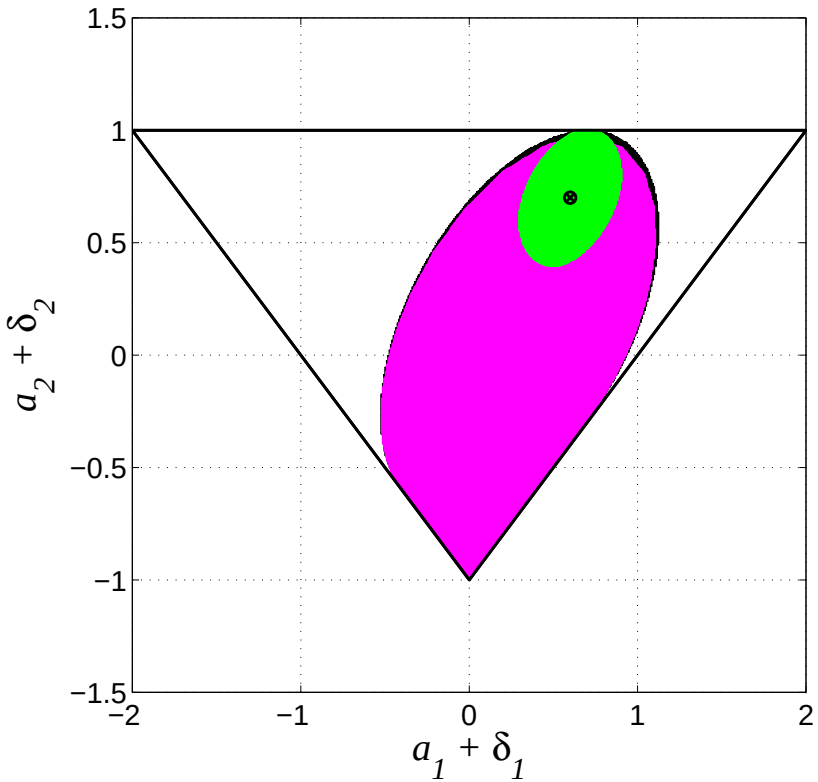
- Interpreting Theorem 1 in the context of IIR filters, let $f(z) = \tilde{a}(z) + \tilde{\delta}(z)$, then (4) becomes

$$|\tilde{\delta}(z)| < |\tilde{a}(z) + \tilde{\delta}(z)| + |\tilde{a}(z)| \quad \text{for all } z \text{ on } C \tag{5}$$

• The stability region obtained based on (5) is shown to be identical to that based on the concept of strictly positive real (SPR) (Dumitrescu and Niemisto, 2004).

• A set of $K$ linear stability constraints has been derived based on (5) and is shown to be identical to that obtained by Lai and Lin (2010) derived from SPR.

**<u>Example 1</u>** Consider 2nd-order Schur polynomial $\tilde{a}(z) = z^2 + 0.6z + 0.7$. Shown below are the stability regions constructed based on Rouche theorem (in green), on SPR (with boundary in black), and on (5) with $K = 100$ (in magenta).

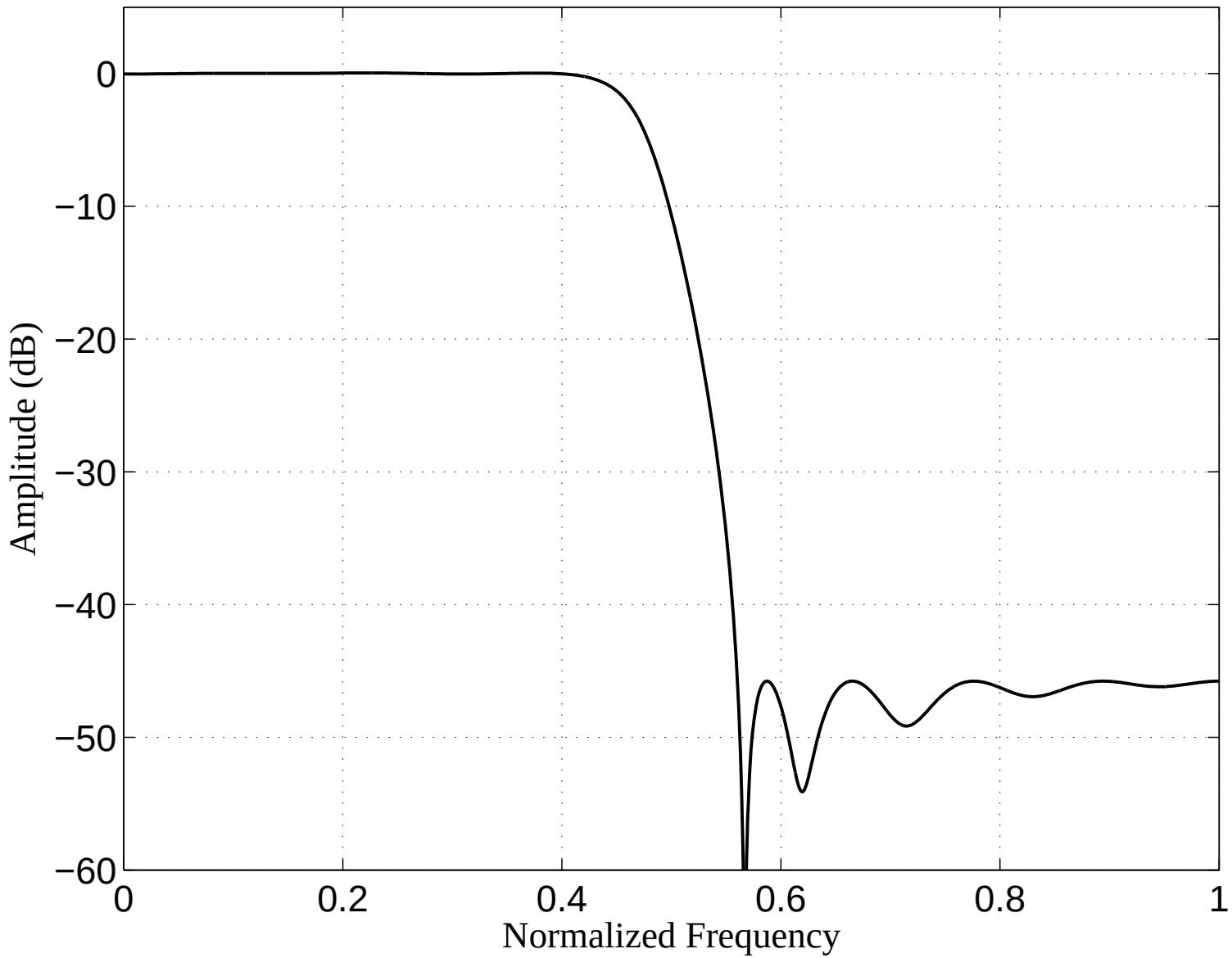


# 6. Design Examples

**<u>Example 2</u>** To design a lowpass filter of order $(n, m) = (4, 15)$ with normalized passband edge $\omega_p = 0.4\pi$, stopband edge $\omega_a = 0.56\pi$, and passband group delay $\kappa = 12$. The initial design was set to $\boldsymbol{a}_0 = \mathbf{0}$ and $\boldsymbol{b}_0$ = the impulse response of a least-squares linear-phase FIR filter of order $15$. The weight $W(\omega)$ was a piecewise constant function being unity over the passband and constant $w$ over the stopband, where the value of $w$ was determined by trial-and-error to equalize the approximation errors in passband and stopband. With $L = 400$, $K = 1000$, $r = 0.9$, $w = 1.01$, $\varepsilon = 0.002$, and $\varepsilon_c = 2 \times 10^{-9}$, it took the algorithm 23 iterations to converge to a solution.

- The maximum amplitude of the poles was $0.8596$. The amplitude response of the IIR filter obtained is depicted below.

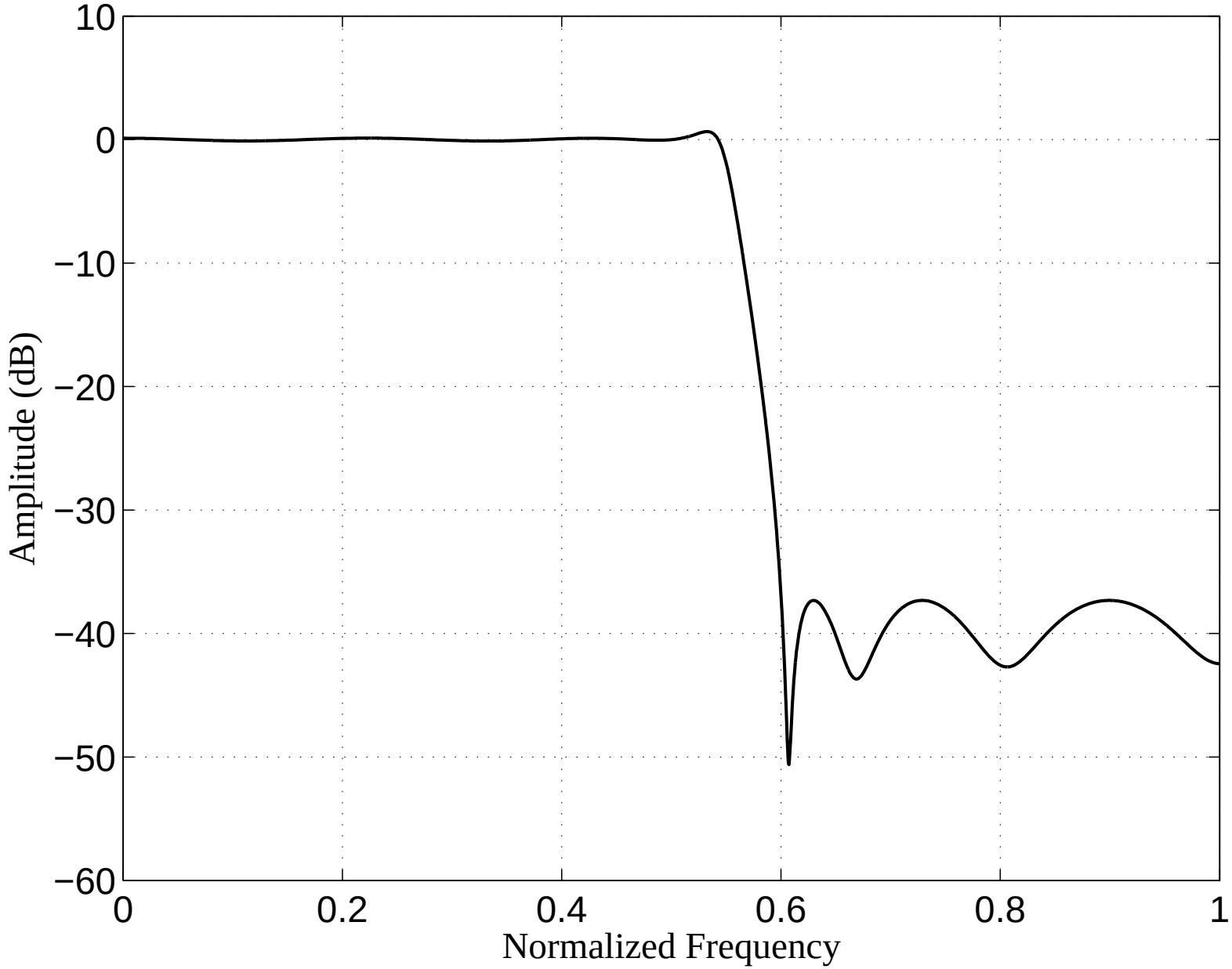
| | Largest deviation in passband amplitude | Minimum stopband attenuation |
|---|---|---|
| Method of this paper | 0.005121 | −45.7723 dB |
| Method of [15] | 0.005177 | –45.7117 dB |
| Method of [14] | 0.005127 | –45.6616 dB |


0
−10
−20
−30
−40
−50
−60
Amplitude (dB)
0
0.2
0.4
0.6
0.8
1
Normalized Frequency

**<u>Example 3</u>** To design a lowpass filter of order $(n, m) = (6, 12)$ with normalized passband edge $\omega_p = 0.5\pi$, stopband edge $\omega_a = 0.6\pi$, and passband group delay $\kappa = 9$. The initial design was set to $\boldsymbol{a}_0 = \mathbf{0}$ and $\boldsymbol{b}_0$ = the impulse response of a least-squares linear-phase FIR filter of order $12$. With $L = 400$, $K = 1000$, $r = 0.975$, $w = 1.112$, $\varepsilon = 0.002$, and $\varepsilon_c = 8 \times 10^{-5}$, it took the algorithm $29$ iterations to converge to a solution.

- The maximum amplitude of the poles was $0.9422$. The amplitude response of the IIR filter obtained is depicted below.

| | Largest deviation in passband amplitude | Minimum stopband attenuation |
|---|---|---|
| Method of this paper | 0.01363 | −37.3101 dB |
| Method of [15] | 0.01470 | –36.5363 dB |
| Method of [10] | 0.01550 | –36.1934 dB |


10
0
−10
−20
−30
−40
−50
−60
Amplitude (dB)
0
0.2
0.4
0.6
0.8
1
Normalized Frequency

Thank you.

Q & A