2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

On consistent estimation of farthest NMP zeros of stable LTI systems

Lirong Huang, Håkan Hjalmarsson, Cristian R. Rojas

Abstract— **Recently, identification of real-valued NMP zeros of discrete-time LTI systems has been studied in a few works. Among the key results, an adaptive input design approach is presented for consistent estimation of NMP zeros of stable LTI systems. However, this result is not applicable if the sign of the first impulse response coefficient is unknown. This paper studies the problem of estimation of the farthest NMP zeros without such prior knowledge.**

I. INTRODUCTION

A model is often used in control design for both analysis and synthesis purposes. Consequently, system identification with focus on control design has been a research area with a lot of activity. The overall objective of identification for control is to deliver models suitable for control design (see, e.g., [5], [10] and the references therein). Non-minimum-phase (NMP) zeros play important roles in many control applications since they limit closed loop performance. Recently, identification of real-valued NMP zeros of discrete-time LTI systems has been studied in a few works (see, e.g., [5], [10] and [12]). Particularly, [12] proposes an adaptive approach to consistent estimation of real-valued NMP zeros of stable LTI systems, which shows that it is possible to estimate a real-valued NMP zero with multiplicity one outside the unit circle consistently using a simple two parameter FIR model if the input can be manipulated and some prior information is available. In [12], it is assumed that not only some prior knowledge about the location of the NMP zero of interest but also the sign of the first impulse response coefficient are known. In fact, the assumption on the sign of the frequency gain is frequently used in adaptive control (see [1], [11], [12] and the references therein). However, prior knowledge on the sign of the first impulse response coefficient is unavailable in many practical cases. Based on the result in [12], this work studies the problem of consistent estimation of the farthest NMP zeros of stable LTI systems when the sign of the zero is known while that of the first impulse response coefficient is unknown. In this paper, a real-valued NMP zero z of a system is called a farthest NMP zero if there is no zero of the system larger (resp., less) than z when $z > 0$ (resp., $z < 0$).

II. REVIEW OF ADAPTIVE NMP ZERO ESTIMATION

In this section, we review the development of an adaptive method for consistent estimation of real-valued NMP zeros in [12]. The assumptions used in [12] are cited as follows:

The authors are with Automatic Control Lab, School of Electrical Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden. Email addresses: lirong.huang@ee.kth.se (L. Huang), hakan.hjalmarsson@ee.kth.se (H. Hjalmarsson), cristian.rojas@ee.kth.se (C. R. Rojas).

Assumption 2.1 (System): The system has a state-space representation of the form

$$\begin{aligned} \xi_{n+1} &= A^o\xi_n + B^o u_n + K^o e_n^o \\ y_n &= C^o\xi_n + e_n^o \end{aligned} \tag{1}$$

where $u_n \in \mathbb{R}$ and $y_n \in \mathbb{R}$ represent the input and measured output at time n, respectively, where $\xi \in \mathbb{R}^m$, for some positive integer m, is the state vector, and e_n^o represents noise acting on the system.
The transition matrix A^o has all its eigenvalues strictly inside the unit circle, i.e., the system is internally stable.
The input-output relationship of the system is given by

$$y_n = G^o(q)u_n + w_n^o, \tag{2}$$

where

$$G^o(q) = C^o(qI - A^o)^{-1}B^o = \sum_{j=1}^{\infty} g_j^* q^{-j}, \tag{3}$$

and $w_n^o = H^o(q)e_n^o$ with $H^o(q) = C^o(qI - A^o)^{-1}K^o + 1$.
The system has one pure time delay, i.e., $g_1^* \neq 0$.
The system has a real-valued NMP zero of multiplicity 1 at an unknown location z_*, i.e., $G^o(z_*) = 0$, where $z_* \in \mathbb{R}$ and $|z_*| > 1$.

Assumption 2.2 (Prior system knowledge): The following prior knowledge is assumed:

i) *A compact interval*

$$\mathcal{G} = \{g_1 : \underline{g}_1 \leq g_1 \leq \bar{g}_1\} \;\; with \;\; 0 \notin \mathcal{G}, \; g_1^* \in \mathcal{G}; \tag{4}$$

ii) *A compact interval $\mathcal{Z} \subset \mathbb{R}$ with the following properties*

$$\begin{aligned} G^o(z) = 0, \; z \in \mathcal{Z} \;&\Rightarrow\; z = z_* \\ z \in \mathcal{Z} \;&\Rightarrow\; |z| > 1; \end{aligned} \tag{5}$$

iii) *The parity of the number of system zeros on the ray $\{\alpha z_*, \alpha > 1\}$ is known.*

Assumption 2.3 (Noise): The noise $\{e_n^o\}$ is a sequence of independent random variables of zero mean and variance λ_o for which

$$\sup_n \mathbb{E}\big[e^{\varepsilon (e_n^o)^2}\big] < \infty \tag{6}$$

holds for some $\varepsilon > 0$.

Assumption 2.4 (Input): The input is generated by

$$u_{n+1} = \rho_n^{-1} u_n + \sqrt{\lambda_u}\sqrt{1 - \rho_n^{-2}}\; r_n \tag{7}$$

where λ_u is a user-defined positive constant and $\{r_n\}$ is a sequence of independent random variables of zero mean and unit variance. Furthermore, $\{r_n\}$ is independent of $\{e_n^o\}$ and subject to the condition $\sup_n \mathbb{E}\big[e^{\varepsilon (r_n)^2}\big] < \infty$ for some $\varepsilon > 0$.

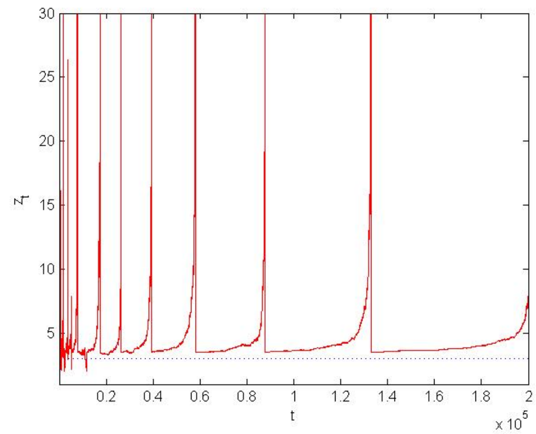

Fig. 1. Solid (red) line: Typical realization of (26) in [12] applied to (47) with initial values $g_1 = -0.5$ and $z = 3.5$. Dotted (blue) line: True location of the zero.

It is assumed that the input is stationary and generated by (7) with $\rho_n = \rho$ and $|\rho| > 1$, which is emphasized by indexing the signals with ρ, e.g.,

$$u_n(\rho) = \rho^{-1} u_{n-1}(\rho) + \sqrt{\lambda_u}\,\sqrt{1-\rho^{-2}}\; r_{n-1}, \quad (8)$$

$$y_n(\rho) = G^o(\rho) u_n(\rho) + w_n^o. \quad (9)$$

The model parameters used in [12] are the first impulse response coefficient g_1 and the zero of interest z. Correspondingly, the model predictor and the normal equations for the least squares critirion are given by (15) and (16) in [12], respectively. It is verified that $g_1 = g_1^*$ and $z = z_*$ is the unique solution to the normal equations with $\rho = z_*$, i.e., (22) in [12], and an adaptive algorithm is proposed by (26) in [12] for estimation of NMP zeros of the system (2). It should be pointed out that, according to Assumption 2.2, the sign of the first impulse response coefficient g_1^* has to be known in advance since $0 \notin \mathcal{G}$. This is also explicitly stated in [12]. In fact, without prior knowledge on the sign of g_1^*, the proposed method in [12] is not applicable. For example, with prior knowledge of the sign of g_1^*, the algorithm presented by [12] managed to estimate the farthest NMP zero $z_* = 3$ of the system (47) by choosing initial values $g_1 = 0.5$ and $z = 3.5$ (see Numerical example (61) in [12]). But it does not work when the initial values are given by $g_1 = -0.5$ and $z = 3.5$, which represents the case in which the sign of g_1^* is unknown in advance, see Fig. 1. To cope with this problem, this paper proposes an improved algorithm for estimation of the farthest NMP zeros of the system (2) when prior knowledge on the sign of g_1^* is unavailable.

III. An improved algorithm

Let z_* be the farthest NMP zero of interest and $s_* = sgn(z_*)$. Since there is no prior knowledge on the sign of g_1^*, we replace Assumption 2.2 with the following one, in which the origin can be contained in the set $\mathcal{G}$.

Assumption $2.2'$: *The following prior knowledge is assumed:*

i) *A real convex compact set*

$$\mathcal{G} = \{g_1 : \underline{g}_1 \le g_1 \le \bar{g}_1\} \;\; with \;\; g_1^* \in \mathcal{G}; \quad (10)$$

ii) *A compact interval* $\mathcal{Z} \subset \mathbb{R}$ *with the following properties*

$$\begin{array}{c} z_* \in \mathcal{Z} \\ G^o(z) = 0,\, z \in \mathcal{Z} \;\Rightarrow\; z = z_* \\ z \in \mathcal{Z} \;\Rightarrow\; |z| > 1. \end{array} \quad (11)$$

So there is no zero of the system (2) larger (resp. less) than z_* if $s_* = 1$ (resp. $s_* = -1$). Since $\mathcal{Z}$ is a compact interval, there is a pair of positive numbers $l_{\mathcal{Z}}$ and $h_{\mathcal{Z}}$ such that $1 < l_{\mathcal{Z}} \le |z| \le h_{\mathcal{Z}}$ for all $z \in \mathcal{Z}$. Given $h_M > h_{\mathcal{Z}}$, define a compact interval $\bar{\mathcal{Z}} \subset \mathbb{R}$ by

$$\bar{\mathcal{Z}} = \{z \in \mathbb{R} : sgn(z) = s_*, l_{\mathcal{Z}} \le |z| \le h_M\}. \quad (12)$$

Clearly, $\mathcal{Z} \subset \bar{\mathcal{Z}}$ and, for any arbitrarily large h_M, there is no zero of the system other than z_* contained in $\bar{\mathcal{Z}}$. Moreover, based on the prior knowledge of the system, i.e., the compact sets $\mathcal{G}$ and $\mathcal{Z}$, it is easy to find a compact set

$$\mathcal{K} = \{k_1 : \underline{k}_1 \le k_1 \le \bar{k}_1\} \quad (13)$$

such that $k_1^* \in \mathcal{K}$ with $k_1^* = g_1^*/z_*$, e.g., $\mathcal{K} = \{k_1 : k_1 = g_1/z,\; g_1 \in \mathcal{G},\; z \in \mathcal{Z}\}$.

The vector of model parameters that will be used is $\theta = [k_1 \; z]^T$, where z is the farthest NMP zero of interest and k_1 represents the numerical relationship between the NMP zero z and the non-zero first impulse response coefficient g_1 such that $g_1 = k_1 z$. The true value of the vector of parameters is denoted by $\theta_* = [k_1^* \; z_*]^T$. Substition of $g_1 = k_1 z$ into (15) in [12] gives the corresponding model predictor as follows

$$\hat{y}_n(\theta, \rho) = k_1 z \big(u_{n-1}(\rho) - z u_{n-2}(\rho)\big). \quad (14)$$

In the limit $n \to \infty$, the prediction-error estimator of k_1 and z are defined by (see, e.g., [9] and [7])

$$\mathbb{E}\big[\psi_n(\theta(\rho), \rho)\big(y_n(\rho) - \hat{y}(\theta, \rho)\big)\big] = 0, \quad (15)$$

where $\psi_n(\theta, \rho)$ is the gradient of the predictor (14) with respect to the model parameters, i.e.,

$$\psi_n(\theta, \rho) := \frac{\partial \hat{y}_n(\theta, \rho)}{\partial \theta} = \begin{bmatrix} z(u_{n-1}(\theta) - z u_{n-2}(\theta)) \\ k_1(u_{n-1}(\theta) - 2 z u_{n-2}(\theta)) \end{bmatrix}. \quad (16)$$

As in [12], we observe

$$\begin{array}{l} \mathbb{E}[u_n(\rho) u_{n-k}(\rho)] = \lambda_u \rho^{-|k|}, \\ \mathbb{E}[u_{n-k}(\rho) y_n(\rho)] = \lambda_u \sum_{j=1}^{\infty} g_j^* \rho^{-|j-k|}, \\ \mathbb{E}[u_{n-1}(\rho) y_n(\rho)] = \lambda_u \rho G^o(\rho), \\ \mathbb{E}[u_{n-2}(\rho) y_n(\rho)] = \lambda_u [\rho^2 G^o(\rho) + g_1^* \rho(\rho^{-2} - 1)], \end{array} \quad (17)$$

which are useful for the development of our proposed result.

Let $\Theta \subset \mathbb{R}^2$ be a set such that $\theta \in \Theta$ implies $|k_1| < \infty$ and $|z| > 1$. Since $g_1 = k_1 z$ and $|z| > 1$, $g_1 = 0$ if and only if $k_1 = 0$. Recall that $|g_1^*| > 0$ is taken for granted by Assumption 2.1 and therefore $|k_1^*| > 0$. Moreover, assume that $\theta_* \in \Theta$ and there is no zero of the system other than z_* contained in Θ, that is, $\theta = [k_1 \; z]^T \in \Theta$ and $G^o(z) = 0$

imply $z = z_*$. The following result is presented as the basis for the algorithm.

Lemma 3.1: On the set Θ, $\theta_* = [k_1^* \ z_*]^T$ is the unique solution to

$$\mathbb{E}\big[\psi_n(\theta, z)\big(y_n(z) - \hat{y}(\theta, z)\big)\big] = 0, \tag{18}$$

where $\theta = [k_1 \ z]^T$.

Proof: By computation with (17), we have

$$\begin{aligned}&\mathbb{E}\big[\psi_n(\theta, z)\big(y_n(z) - \hat{y}(\theta, z)\big)\big] \\ &= \lambda_u \begin{bmatrix} z(1-z^2)\big(zG^o(z) - (g_1^* - g_1)\big) \\ k_1\big(z(1-2z^2)G^o(z) - 2(1-z^2)(g_1^* z^{-1} g_1)\big)\end{bmatrix} \\ &= 0. \end{aligned} \tag{19}$$

It is observed that the first element of (19) is equal to zero if and only if there are z and $g_1 = k_1 z$ with $\theta \in \Theta$ such that

$$g_1^* - g_1 = zG^o(z). \tag{20}$$

Substitution of (20) into the second element of (19) gives

$$k_1 z G^o(z) = g_1 G^o(z) = 0 \ \Rightarrow \ G^o(z) = 0 \ \Rightarrow \ z = z_*$$

since z_* is the only zero of the system contained in Θ. Now (20) implies $g_1 = g_1^*$ and hence $k_1 = g_1^*/z_* = k_1^*$. □

By Lemma 3.1, solving normal equation (18) yields consistent estimates of k_1^* and z_* (and hence that of g_1^*). The model (14) with the gradient (16) immediately suggests the following recursive prediction error method scheme [9]:

$$\begin{aligned}\begin{bmatrix}\hat{k}_{1,n+1} \\ \hat{z}_{n+1}\end{bmatrix} &= \begin{bmatrix}\hat{k}_{1,n} \\ \hat{z}_n\end{bmatrix} + \frac{1}{n+1} R_n^{-1}\psi_n(y_{n+1} - \hat{y}_{n+1}) \\ \hat{y}_n &= \hat{k}_{n-1}\hat{z}_{n-1}(u_{n-1} - \hat{z}_{n-1}u_{n-2}) \\ \psi_n &= \begin{bmatrix}\hat{z}_{n-1}(u_{n-1} - \hat{z}_{n-1}u_{n-2}) \\ \hat{k}_{1,n-1}(u_{n-1} - 2\hat{z}_{n-1}u_{n-2})\end{bmatrix} \\ R_{n+1} &= R_n + \frac{1}{n+1}(\psi_{n+1}\psi_{n+1}^T - R_n). \end{aligned} \tag{21}$$

Let ρ_n in the input filter (7) be fixed on ρ with $|\rho| > 1$, then the system is described by (8)-(9) and the model with fixed parameters $\theta = [k_1 \ z]^T$ is given by (14), (16) and

$$\begin{aligned}&R_{n+1}(\theta, \rho) = R_n(\theta, \rho) \\ &\quad + \frac{1}{n+1}\big(\psi_{n+1}(\theta, \rho)\psi_{n+1}^T(\theta, \rho) - R_n(\theta, \rho)\big). \end{aligned} \tag{22}$$

Note $\lim_{n\to\infty} R_n(\theta, \rho) = R(\theta, \rho) := \mathbb{E}[\psi_n(\theta, \rho)\psi_n^T(\theta, \rho)]$ and let the input filter (7) use the model zero. Using (17), we obtain

$$\begin{aligned}R(\theta) = R(\theta, z) &= \mathbb{E}\big[\psi_n(\theta, z)\psi_n^T(\theta, z)\big] \\ &= \lambda_u \begin{bmatrix} z^2(z^2-1) & 2k_1 z(z^2-1) \\ 2k_1 z(z^2-1) & k_1^2(4z^2-3)\end{bmatrix}\end{aligned}$$

and hence $R^{-1}(\theta) = \frac{1}{\lambda_u}\begin{bmatrix}\frac{4z^2-3}{z^2(z^2-1)} & -\frac{2}{k_1 z} \\ -\frac{2}{k_1 z} & \frac{1}{k_1^2}\end{bmatrix}$. Furthermore, we have

$$R^{-1}(\theta)\psi_n(\theta, z) = \frac{1}{\lambda_u}\begin{bmatrix}\frac{2z^2-1}{z(z^2-1)}u_{n-1}(z) - \frac{1}{z^2-1}u_{n-2}(z) \\ -\frac{1}{k_1}u_{n-1}(z)\end{bmatrix}. \tag{23}$$

Substitution of (23) with $\theta = [k_1 \ z]^T$ replaced by $\hat{\theta}_n = (\hat{k}_{1,n} \ \hat{z}_n)^T$ into the recursive scheme (21) yields our proposed algorithm as follows

$$\begin{aligned}&\hat{y}_{n+1} = \hat{k}_{1,n}\hat{z}_n(u_n - \hat{z}_n u_{n-1}) \\ &\begin{bmatrix}\hat{k}_{1,n+1} \\ \hat{z}_{n+1}\end{bmatrix} = \begin{bmatrix}\hat{k}_{1,n} \\ \hat{z}_n\end{bmatrix} + \frac{1}{\lambda_u(n+1)} \\ &\qquad \times \begin{bmatrix}\frac{2\hat{z}_n^2-1}{\hat{z}_n(\hat{z}_n^2-1)}u_{n-1} - \frac{1}{\hat{z}_n^2-1}u_{n-2} \\ -\frac{1}{\hat{k}_{1,n}}u_{n-1}\end{bmatrix}(y_{n+1} - \hat{y}_{n+1}). \end{aligned} \tag{24}$$

In the sequel, the entire adaptive system is presented by (25)-(29) for analysis, where $\eta_n = [e_{n+1}^o \ r_n]^T$, $\theta_n = [\hat{k}_{1,n} \ \hat{z}_n]^T$, $\Phi_n = [\Phi_n^1 \ \Phi_n^2 \ \cdots \ \Phi_n^7]^T$ with $\Phi_n^1 = u_n$, $\Phi_n^2 = \xi_n$, $\Phi_n^3 = e_n^o$, $\Phi_n^4 = u_{n-1}$, $\Phi_n^5 = \frac{2\hat{z}_n^2-1}{\lambda_u \hat{z}_n(\hat{z}_n^2-1)}u_{n-1} - \frac{1}{\lambda_u(\hat{z}_n^2-1)}u_{n-2}$, $\Phi_n^6 = -\frac{1}{\lambda_u \hat{k}_{1,n}}u_{n-1}$, $\Phi_n^7 = \hat{y}_n$, $Q_{1,n+1} = \Phi_{n+1}^5(C^o\Phi_{n+1}^2 + \Phi_{n+1}^3 - \Phi_{n+1}^7)$, $Q_{2,n+1} = \Phi_{n+1}^6(C^o\Phi_{n+1}^2 + \Phi_{n+1}^3 - \Phi_{n+1}^7)$,

$$A(\theta) = \begin{bmatrix} z^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ B^o & A^o & K^o & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{2z^2-1}{\lambda_u z(z^2-1)} & 0 & 0 & -\frac{1}{\lambda_u(z^2-1)} & 0 & 0 & 0 \\ -\frac{1}{\lambda_u k_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ k_1 z & 0 & 0 & -k_1 z^2 & 0 & 0 & 0 \end{bmatrix},$$

$$B(\theta) = \begin{bmatrix} 0 & \sqrt{\lambda_u}\sqrt{1-z^{-2}} \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad Q(\Phi_{n+1}) = \begin{bmatrix} Q_{1,n+1} \\ Q_{2,n+1}\end{bmatrix}.$$

Adaptive system

$$D_0 := \{[k_1 \ z]^T : k_1 \in \mathcal{K}, z \in \tilde{\mathcal{Z}}\} \tag{25}$$

$$\theta_0 = [k_{1,0} \ z_0]^T \in \mathrm{int}D_0, \ k_{1,0} \neq 0, \ z_0 = s_* h_{\mathcal{Z}} \tag{26}$$

$$\Phi_{n+1} = A(\theta_n)\Phi_n + B(\theta_n)\eta_n \tag{27}$$

$$\theta_{n+1-} = \theta_n + \tfrac{1}{n+1}Q(\Phi_{n+1}) \tag{28}$$

$$\theta_{n+1} = \begin{cases} \theta_{n+1-} & \theta_{n+1-} \in D_0 \\ \theta_0 & \text{otherwise}\end{cases} \tag{29}$$

It is reasonable to choose $\theta_0 \in \mathrm{int}D_0$ with $|k_{1,0}| > 0$ since $|z_0| = h_{\mathcal{Z}}$ and $|g_1^*| > 0$. In practice, $k_{1,0}$ should be chosen with small $|k_{1,0}|$ while h_M is set to be large.

IV. Stability and Sensitivity Analysis of the Associated ODE

The asymptotic behaviour of the adaptive system (25)-(29) is determined by the so called associated ODE (see [9]). For rigorous analysis for recursive estimation schemes with resetting, readers are referred to [2] and [3]. The associated ODE for system (27)-(29) is given by

$$\dot{\theta}_t = \frac{1}{t}F(\theta_t) \tag{30}$$

where $F(\theta) := \mathbb{E}[Q(\bar{\Phi}_n(\theta)]$ and in turn $\bar{\Phi}_n(\theta)$ corresponds to (27) with frozen parameters θ, i.e.,

$$\bar{\Phi}_{n+1}(\theta) = A(\theta)\bar{\Phi}_n(\theta) + B(\theta)\eta_n. \tag{31}$$

This implies

$$F(\theta) = \frac{1}{\lambda_u}\mathbb{E}\Big\{ \begin{bmatrix} \frac{2z^2-1}{z(z^2-1)}u_{n-1}(z) - \frac{1}{z^2-1}u_{n-2}(z) \\ -\frac{1}{k_1}u_{n-1}(z) \end{bmatrix} \times \big(y_n(z) - \hat{y}_n(\theta, z)\big)\Big\}. \tag{32}$$

Computing (32) with (17), we obtain

$$F(\theta) = \begin{bmatrix} G^o(z) + \big(\frac{g_1^*}{z} - k_1\big) \\ -\frac{z}{k_1}G^o(z) \end{bmatrix}. \tag{33}$$

Substitution of (33) into (30) and introduction of $t = e^v$ give the associated ODE for the adaptive system (27)-(29) as follows

$$\dot{k}_1(v) = G^o(z(v)) + \frac{g_1^*}{z(v)} - k_1(v), \tag{34}$$

$$\dot{z}(v) = -\frac{z(v)}{k_1(v)}G^o(z(v)). \tag{35}$$

Recall that $g_1(v) = k_1(v)z(v)$. Multiplying by $z(v)$ and $k_1(v)$ both sides of (34) and (35) respectively, we have

$$z(v)\dot{k}_1(v) = z(v)G^o(z(v)) + g_1^* - g_1(v), \tag{36}$$

$$k_1(v)\dot{z}(v) = -z(v)G^o(z(v)). \tag{37}$$

Since $\dot{g_1}(v) = z(v)\dot{k}_1(v) + k_1(v)\dot{z}(v)$, (36) and (37) give

$$\dot{g_1}(v) = g_1^* - g_1(v), \tag{38}$$

which immediately yields

$$g_1(v) - g_1^* = (g_1(0) - g_1^*)e^{-v}. \tag{39}$$

This means that the equilibrium g_1^* of system (38) is exponentially stable. If $k_1(0)$ has the same sign as k_1^*, then $g_1(0)$ has the same sign as g_1^* and the analysis of stability and sensitivity has been discussed in [12]. Let us now consider the cases when $k_1(0)$ has has different sign from that of k_1^*, i.e., $g_1(0)$ has a different sign from that of g_1^*. According to (39), there exists a number $v_1 \geq 0$ such that $g_1(v_1) = 0$ and hence $k_1(v_1) = 0$. Define $v_* = \inf\{v \geq 0 : z(v) = z_*\}$. If $v_* \leq v_1$, then the Lebesgue measure of set $\{v \geq v_* : \dot{z}(v) \neq 0\}$ is zero. This implies $z(v) = z_*$ for all $v \geq v_*$. In this case, $k_1(v) \to k_1^*$ as $v \to \infty$ since $g_1(v) \to g_1^*$ as $v \to \infty$. So we only need to consider the cases with $v_* > v_1$.

First, let us consider the case when $s_* = 1$, $k_1^* > 0$ and $k_1(0) < 0$, which implies $g_1^* > 0$ and $g_1(0) < 0$. Since $k_1(v) < 0$, $g_1(v) < 0$ for $v < v_1$ and $k_1(v) > 0$, $g_1(v) > 0$ for $v > v_1$, there are numbers $a_1 > 0$ and $0 < a_2 < v_* - v_1$ such that $\dot{k}_1(v) > 0$ and $\dot{g_1}(v) > 0$ on $v \in [v_1 - a_1, v_1 + a_2]$. Obviously, $z(v)$ is continuous and bounded on $[0, v_1 - a_1]$ and $[v_1 + a_2, v_*)$. We claim that $z(v)$ is essentially bounded on $[v_1 - a_1, v_1 + a_2]$. Since $z(0) > z_*$ and $v_* > v_1 + a_2$, $-z(v)G^o(z(v))$ does not change its sign on $v \in [v_1 - a_1, v_1 + a_2]$. So we have

$$g_1(v_1 + a_2) - g_1(v_1 - a_1) = \int_{v_1-a_1}^{v_1+a_2} \dot{g_1}(v)\mathrm{d}v = \int_{v_1-a_1}^{v_1+a_2} \big[\dot{k}_1(v)z(v) - z(v)G^o(z(v))\big]\mathrm{d}v \tag{40}$$

If $-z(v)G^o(z(v))$ is positive on $[v_1 - a_1, v_1 + a_2]$, then, by the mean value theorem, (40) gives

$$g_1(v_1 + a_2) - g_1(v_1 - a_1) \geq \int_{v_1-a_1}^{v_1+a_2} \dot{k}_1(v)z(v)\mathrm{d}v = \dot{k}_1(v_a)\int_{v_1-a_1}^{v_1+a_2} z(v)\mathrm{d}v \tag{41}$$

where $v_a \in (v_1 - a_1, v_1 + a_2)$. This implies that $z(v)$ is essentially bounded on $[v_1 - a_1, v_1 + a_2]$. If $-z(v)G^o(z(v))$ is negative on $[v_1 - a_1, v_1 + a_2]$, then (40) gives

$$g_1(v_1 + a_2) - g_1(v_1 - a_1) + \int_{v_1-a_1}^{v_1+a_2} z(v)G^o(z(v))\mathrm{d}v = \int_{v_1-a_1}^{v_1+a_2} \dot{k}_1(v)z(v)\mathrm{d}v = \dot{k}_1(v_b)\int_{v_1-a_1}^{v_1+a_2} z(v)\mathrm{d}v \tag{42}$$

where $v_b \in (v_1 - a_1, v_1 + a_2)$. Since $z(v) > 1$ for $v \in [v_1 - a_1, v_1 + a_2]$ and the system is internally stable, (3) implies that $0 < z(v)G^o(z(v)) < \infty$ on $[v_1 - a_1, v_1 + a_2]$. So we have

$$\int_{v_1-a_1}^{v_1+a_2} z(v)G^o(z(v))\mathrm{d}v < \infty.$$

But, by (42), this implies $z(v)$ is essentially bounded on $[v_1 - a_1, v_1 + a_2]$. Therefore, we have

$$\|z(v)\|_\infty := \operatorname*{ess\,sup}_{v \geq 0} |z(v)| < \infty. \tag{43}$$

in this case. It is also observed that $\dot{z}(v)$ is well defined and hence $z(v)$ is continuous for $v \in [v_1 - a_1, v_1)$ and $v \in (v_1, v_1 + a_2]$. Similarly, we can find these properties in the other cases when $k_1(0)$ (resp. $g_1(0)$) has a different sign from that of k_1^* (resp. g_1^*). When h_M is chosen sufficiently large with $\|z(v)\|_\infty < h_M$, $k_1(v)$ and $z(v)$ are in the interior of D_0 almost everywhere on $[0, v_1 + a_2]$ (see Lemma 5.1 [12]), which implies that $k_1(v)$ (resp. $g_1(v)$) computed by (24) change its sign from the one of $k_1(0)$ to that of k_1^* (resp. g_1^*) almost surely as the step number n increases from 0 to larger than $v_1 + a_2$. Let $v_0 \geq v_1 + a_2$. Then we have $sgn(k_1(v_0)) = sgn(k_1^*)$, i.e., $sgn(g_1(v_0)) = sgn(g_1^*)$ and $\theta(v_0) = [k_1(v_0)\ z(v_0)]^T \in D_0$. The rest of analysis of stability and sensitivity on $[v_0, \infty)$ is similar to that in [12] and hence omitted.

According to the analysis given above, we have the following result.

Lemma 4.1: Let $[k_1(v, v_0, \gamma)\ z(v, v_0, \gamma)]^T$ with $\gamma = [\gamma_1\ \gamma_2]^T \in D_0$ and $v \geq v_0$ be the solution to (34)-(35) with $k_1(v_0) = \gamma_1$ and $z(v_0) = \gamma_2$. Then there are positive constants C_1, C_2, a and b such that

$$\left|\frac{\partial k_1(v, v_0, \gamma)}{\partial \gamma}\right| \leq C_1 e^{-b(v-v_0)}, \left|\frac{\partial z(v, v_0, \gamma)}{\partial \gamma}\right| \leq C_2 e^{-a(v-v_0)}$$

for all $v \geq v_0$.

V. Convergence analysis

In this section, we consider the convergence of our algorithm (24) based on the above analysis. For the readers' convenience, we cite the following definition (see [2] and [3]).

Definition 5.1: A random process $\{s_n\}$ is said to be M-bounded, which is denoted by $s_n = O_M(1)$, if $M_q(s) = \sup_{n\geq 0} \mathbb{E}^{1/q}\left[|s_n|^q\right] < \infty$ for all $1 \leq q < \infty$.

Suppose that $\{c_n\}$ is a sequence of positive numbers. We also write $s_n = O_M(c_n)$ if $s_n/c_n = O_M(1)$.

We can verify that Conditions 1-3 in [12, Appendix A] (see also [2]) are satisfied, which is done in the same way as the proof of [12, Theorem 6.1] (see [12, Appendix E]). Therefore, by [12, Theorem A.1] (see also [2, Theorem 4.1]), we have

Theorem 5.1: Let the above assumptions with Assumption 2.2 replaced by Assumption 2.2$'$ hold. If $[\hat{k}_{1,n}\ \hat{z}_n]^T$ is computed by recursive scheme (24), then

$$\hat{k}_{1,n} - k_1^* = O_M(n^{-\beta}),\ \hat{z}_n - z_* = O_M(n^{-\beta}) \tag{44}$$

where $\beta = \min\{1/2, |z_*\tilde{G}^o(z_*)/k_1^*|\}$ with $\tilde{G}^o$ given by $\tilde{G}^o(z)(z - z_*) = G^o(z)$.

By the well-known Borel-Cantelli lemma, convergence in L_q for all $q \geq 1$ with rate $O_M(n^{-\beta})$ and $\beta > 0$ implies almost sure convergence. Thus, Theorem 5.1 implies that $[\hat{k}_{1,n}\ \hat{z}_n]^T$ converges to $\theta_* = [k_1^*\ z_*]^T$ almost surely.

VI. Numerical examples

In this section, we show the effectiveness of our proposed result. To compare with the result in [12], we consider the estimation of the farthest NMP zeros of the following examples, i.e., the systems (59)-(61) in [12], for which $s_* = 1$ and $\mathcal{Z} = [2, 10]$ (cf. [12]). Consequently, $z_0 = 10$. It is noticed that, for any arbitrarily large h_M, z_* is the only zero of the system in $\bar{\mathcal{Z}}$, where $\bar{\mathcal{Z}}$ is defined by (12). Since we do not have much prior knowledge on $\|z(v)\|_\infty$, we should choose sufficiently large h_M in practice. Moreover, for $g_{1,0} = 0.5$ in [12], we choose $|k_{1,0}| = |g_{1,0}/z_0| = 0.05$. In practice, since we do not know the sign of g_1^*, we should choose $k_{1,0}$ such that $|g_{1,0}|$ and hence $|k_{1,0}|$ are small so that the sign of g_1 (resp., k_1) will change to the same as g_1^* (resp., k_1^*) within a few steps in case that $g_{1,0}$ (resp., $k_{1,0}$) has different sign from that of g_1^* (resp., k_1^*).

Example 6.1 Consider a system described by

$$y_n = (q^{-1} - 3q^{-2})u_n + e_n^o \tag{45}$$

where $\{e_n^o\}$ is Gaussian white noise of variance 0.01. Note that system (45) has exactly one NMP zero at $z_* = 3$. Let $\mathcal{Z} = [2, 10]$ and hence $z_0 = 10$. Moreover, $\bar{\mathcal{Z}} = [2, 10^6]$. Since there is no prior knowledge on the sign of g_1^*, $k_{1,0}$ can be a positive or negative number. Typical realizations of the algorithm presented in Section III with $\theta_0 = [k_{1,0}\ z_0]^T = [\pm 0.05\ 10]^T$ are given in Fig. 2 and Fig. 3, respectively.

Example 6.2 Consider a system that has a pole and an additional zero described by

$$y_n = \frac{(q-3)(q-0.1)}{q^2(q-0.5)}u_n + e_n^o. \tag{46}$$

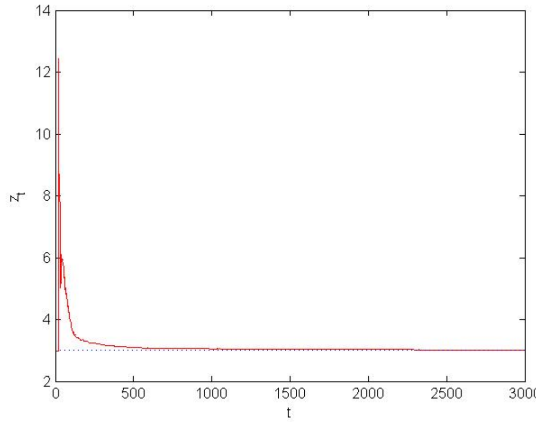

Fig. 2. Solid (red) line: Typical realization of (24) applied to (45) with $\theta_0 = [0.05\ 10]^T$. Dotted (blue) line: True location of the zero.

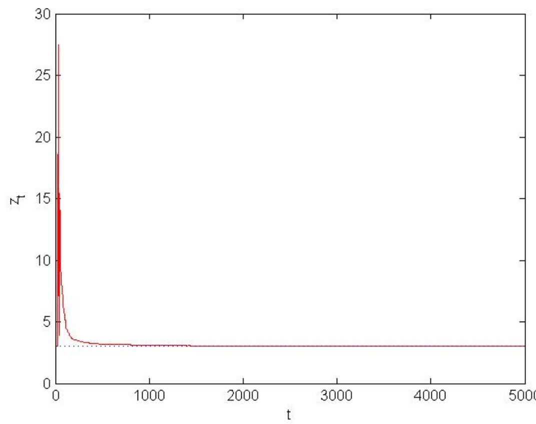

Fig. 3. Solid (red) line: Typical realization of (24) applied to (45) with $\theta_0 = [-0.05\ 10]^T$. Dotted (blue) line: True location of the zero.

System (46) has one NMP zero at $z_* = 3$. Let $\bar{\mathcal{Z}} = [2, 10^6]$. Typical realizations of the algorithm presented in Section III with $\theta_0 = [k_{1,0}\ z_0]^T = [\pm 0.05\ 10]^T$ are given in Fig. 4 and Fig. 5, respectively.

Example 6.3 To study the behavior of algorithm (24) applied to a system of higher complexity (i.e., where G and H have more poles and zeros), let us consider the following system

$$y_n = \frac{(q-3)(q-1.5)(q-0.2)(q+0.3)}{q^4(q-0.5)}u_n + \frac{q}{q-0.8}e_n^o. \tag{47}$$

The farthest NMP zero of system (47) is at $z_* = 3$. It is noticed that the result in [12] does not work when the initial values are given by $g_{1,0} = -0.5$ and $z_0 = 3.5$, see Fig. 1. Let us now turn to the algorithm (24) with $\bar{\mathcal{Z}} = [2, 10^6]$. Typical realizations of the algorithm presented in Section III with $\theta_0 = [k_{1,0}\ z_0]^T = [\pm 0.05\ 10]^T$ are given in Fig. 6 and Fig. 7, respectively.

VII. Conclusion

In this paper, a modified version of an adaptive technique for the estimation of NMP zeros has been presented, which

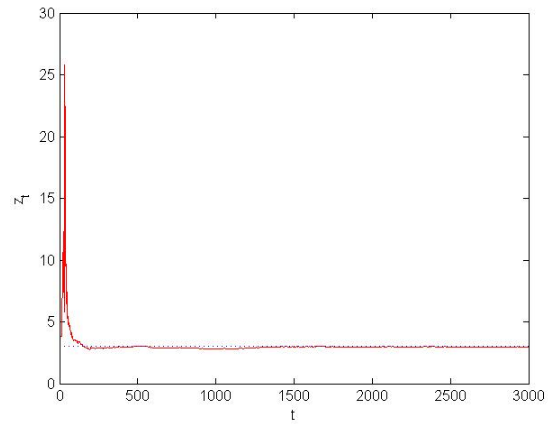

Fig. 4. Solid (red) line: Typical realization of (24) applied to (46) with $\theta_0 = [0.05\ 10]^T$. Dotted (blue) line: True location of the zero.

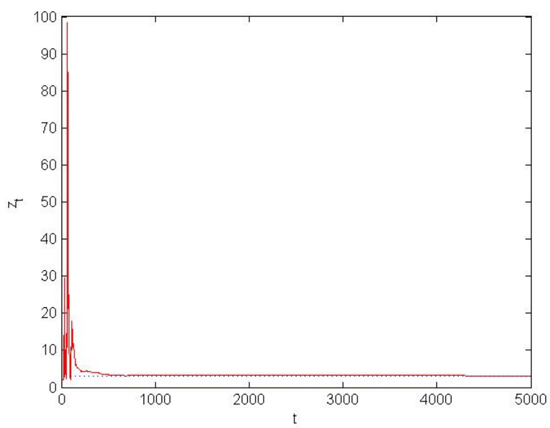

Fig. 5. Solid (red) line: Typical realization of (24) applied to (46) with $\theta_0 = [-0.05\ 10]^T$. Dotted (blue) line: True location of the zero.

relaxes one important assumption on the prior knowledge available to the user, namely, that of the sign of the high frequency gain. This modification improves the applicability of the method, and its performance has been shown both theoretically and via simulation examples.

Our proposed method introduces and employs a parameter k_1 instead of g_1, which represents the numerical relationship between the NMP zero z and the non-zero first impulse response coefficient g_1 such that $g_1 = k_1 z$. By exploiting such numerical relationship, we propose an improved algorithm and find that the solution to the associated ODE is essentially bounded and hence the proposed algorithm can go through the singular point almost surely in a case when $g_{1,0}$ (resp., $k_{1,0}$) has different sign from that of g_1^* (resp., k_1^*). Therefore, our proposed method applies to the cases when there is no prior information on the sign of the first impulse response coefficient g_1^*.

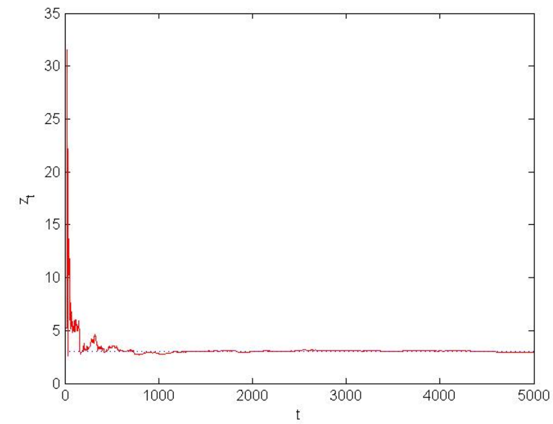

Fig. 6. Solid (red) line: Typical realization of (24) applied to (47) with $\theta_0 = [0.05\ 10]^T$. Dotted (blue) line: True location of the zero.

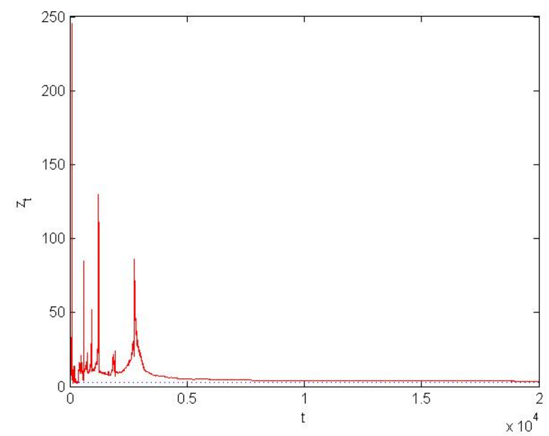

Fig. 7. Solid (red) line: Typical realization of (24) applied to (47) with $\theta_0 = [-0.05\ 10]^T$. Dotted (blue) line: True location of the zero.

REFERENCES

[1] K. Åström and B. Witttenmark, *Adaptive control*, 2nd ed. Massachusetts, USA: Addison-Wesley, 1995.

[2] L. Gerencsér, Rate of convergence of recursive estimators, *SIAM J. Control and Optimization*, vol.30, pp.1200-1227, 1992.

[3] L. Gerencsér, A representation theorem for the error of recursive estimators, *SIAM J. Control and Optimization*, vol.44, pp.2123-2188, 2006.

[4] L. Gerencsér and H. Hjalmarsson, Adaptive input design in system identification, *Proc. of the 44th IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, pp.4988-4993.

[5] H. Jansson, *Experiment design with applications in identification for control*, PhD thesis, Stockholm, Sweden: KTH Royal Institute of Technology, 2004.

[6] L. Ljung, Analysis of recursive stochastic algorithms, *IEEE Trans. Automatic Control*, vol.22, pp.551-575, 1977.

[7] L. Ljung, *System identification: theory for the user*, 2nd ed. New Jersey, USA: Prentice Hall, 1999.

[8] L. Ljung, Recursive identification algorithms, *Circuits Systems Signal Processing*, vol.21, pp.57-68, 2002.

[9] L. Ljung and T. Söderström, *Theory and practice of recursive identification*, Massachusetts, USA: MIT Press, 1983.

[10] J. Mårtensson and H. Hjalmarsson, "Input design for identification of zeros", *16th World Congress on Automatic Control*, IFAC, Prague, Czech Republik, 2005.

[11] K. S. Narendra and A. M. Annaswamy, *Stable adaptive systems*, New Jersey, USA: Prentice Hall, 1989.

[12] C. R. Rojas, H. Hjalmarsson, L. Gerencsér and J. Mårtensson, An adaptive method for consistent estimation of real-valued non-minimum phase zeros in stable LTI systems, *Automatica*, vol.47, pp.1388-1398, 2011.