



# Closed-loop output error identification algorithms with predictors based on generalized orthonormal transfer functions: Convergence conditions and bias distribution

Bernard Vau, Henri Bourlès

## ► To cite this version:

Bernard Vau, Henri Bourlès. Closed-loop output error identification algorithms with predictors based on generalized orthonormal transfer functions: Convergence conditions and bias distribution. *Automatica*, 2021, 125, pp.109377. 10.1016/j.automatica.2020.109377 . hal-03035258

**HAL Id: hal-03035258**

**<https://hal.science/hal-03035258>**

Submitted on 2 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Closed-loop output error identification algorithms with predictors based on generalized orthonormal transfer functions: Convergence conditions and bias distribution

Bernard Vau<sup>a</sup>, Henri Bourlès<sup>a</sup>

<sup>a</sup>*SATIE Ecole normale supérieure de Paris-Saclay, 94230 CACHAN France.*

---

## Abstract

This paper proposes an improved version of closed-loop output-error identification algorithms, where the predictor is established on a generalized basis of orthonormal transfer functions. It is shown that the selection of the basis poles impacts the convergence conditions and the bias distribution of the schemes. These algorithms present several advantages: They are able to identify in closed-loop fast sampled systems, stiff systems (with modes spread over three decades or more), and reduced order models. Moreover, they are suitable for unstable systems or controllers. A simulation example shows the effectiveness of this approach. These algorithms can be employed in an open-loop context by using a straightforward simplification.

*Key words:* Closed-loop identification, stiff systems identification, orthonormal transfer functions bases

---

## Notation

$\mathbb{N}$	set of natural integers $\{1, 2, \dots\}$
$q, q^{-1}$	forward/backward shift operator
$z$	$z$ transform variable
$\lambda$	Hambo operator
$\mathbb{D}$	exterior of the open unit disc $\cup \infty$
$\mathbb{T}$	unit circle
$\mathcal{H}_2(\mathbb{D})$	Hardy space of analytic functions in $\mathbb{D}$ which are square integrable over $\mathbb{T}$
$\mathcal{H}_{2-}$	the set of all functions in $\mathcal{H}_2(\mathbb{D})$ that are null at infinity
$\star$	transpose conjugate symbol
$l_2$	space of square summable sequences on the time interval $\mathbb{N}$
$\ M\ _2$	Largest singular value of matrix $M$
$L(q)$	Polynomial in $q$ i.e. with $q$ substituted to the indeterminate ( $l_0 \neq 0$ if $L \neq 0$ ): $L(q) = l_0 q^{n_l} + l_1 q^{n_l-1} + \dots + l_{n_l}$
$L^*(q^{-1})$	reciprocal polynomial of $L(q)$ $L^*(q^{-1}) = l_0 + l_1 q^{-1} + \dots + l_{n_l} q^{-n_l}$
$L^{*\dagger}(q^{-1})$	$L^*(q^{-1})$ minus its constant term: $L^{*\dagger}(q^{-1}) = l_1 q^{-1} + \dots + l_{n_l} q^{-n_l}$
$d^o(L)$	degree of polynomial $L$ : $d^o(L) = n_l$

---

*Email addresses:* [bernard.vau@satie.ens-cachan.fr](mailto:bernard.vau@satie.ens-cachan.fr)  
(Bernard Vau), [henri.bourles@satie.ens-cachan.fr](mailto:henri.bourles@satie.ens-cachan.fr)  
(Henri Bourlès).

## 1 Introduction

Identification in closed-loop is unavoidable whenever a system includes marginally stable or unstable poles, or if for practical reasons an open-loop identification cannot be carried out (e.g. because of safety constraints), which is a common situation, in particular in industrial applications. The main specific issue related to closed-loop identification lies in the fact that the system input is necessarily correlated with the noise, entailing a biased estimation if operated by means of an open-loop algorithm involving an erroneous noise model or a prefilter ([12] pp. 433, 436). Several identification methods in closed-loop operation have been existing for a long time [16]. Nevertheless, a significant interest for this topic has been triggered in the nineties in the perspective of identification for control [6], [18], entailing several research teams to work simultaneously on this subject: see for example [5], [4], [10]. The schemes proposed in the latter reference by I.D. Landau and A. Karimi, called CLOE and X-CLOE (CLOE being dedicated to an output error model, and X-CLOE to an ARMAX model structure), present a special interest especially in the context of adaptive control, since their recursive structure makes them usable in a real-time situation. Their philosophy is based on the concept of Model Reference Adaptive Systems (MRAS), in which the true closed-loop works in parallel with a simulated closed-loop. These algorithms belong to the pseudo-linear regression (PLR) class, and

their convergence conditions are well-known (see [10]). However, it has been shown only recently in [20], that their bias distributions differs from those of prediction error methods (PEM), and that the native CLOE and X-CLOE structures strongly penalizes the model misfit in high frequency. In general, most of discrete-time identification algorithms are not reliable in a fast sampling situation [11], and the specific bias distribution of CLOE and X-CLOE algorithms make them even more unsuited in case of oversampling. As a consequence, identification of systems having poles separated from several decades (stiff systems) is totally intractable using this algorithms family.

This contribution is twofold:

- (1) Complete CLOE and X-CLOE algorithms in order to make them able to cope with unstable controllers (which is not the case of the initial versions of these schemes), a situation that can occur when, for example, a double integrator is introduced in the control law. For this, a specific closed-loop formulation of the predictor is used (see propositions 1 and 2),
- (2) Make these algorithms suitable for the identification of stiff (or multi-frequency scale) systems.

For purpose (2), a novel parametrization is used, which is established on generalized bases of orthonormal transfer functions (GBOF) introduced by Heuberger et al. in [7] and [8], that are associated with the Hambo transform [9]. The models employed in the present article can be considered as a generalization of infinite impulse response (IIR) filters. In contrast, the existing identification method based on GBOF as described in section 4.3 of [8] (in open-loop operation), rely on a predictor being, roughly speaking, a generalization of finite impulse response (FIR) filters: the past outputs (whether true or predicted) are not used in the predictor. For this latter method, the selection of the basis poles is crucial: if this selection is not relevant, a large amount of estimated parameters is required, leading to increased uncertainties. This is the reason why several authors have proposed methodologies about data-driven pole selection: see for example [1], [15], [8] (chap. 11), or [17]. On the contrary, in the schemes proposed here, an asymptotically unbiased model of minimal order can be obtained without any stringent condition on the basis poles, provided only that these latter ensure the algorithm convergence. Nevertheless, these poles have a strong impact on the bias distribution. Limit models in the frequency domain are obtained by expressing the equivalent prediction error (a non measurable signal) as defined in [20]. The variance of this equivalent prediction error is the convex function minimized by the PLR algorithms for a full order model. These limit models are good indicators of the bias distribution and show that the bias depends on the predictor parametrization, which is a situation that differs from PEM for which

this parametrization has no impact on the model fit (see [12] p. 437). The interest of the proposed developments is to allow for a tuning of the identification schemes in function of the requirements that the user may have. We show through the simulations that one can either identify a wide band model over, say, 3 decades (corresponding to a stiff system), or a local model in the frequency domain that can be useful for control, or even an almost "punctual" model (only valid in a small neighbourhood of one frequency) with highly reduced order. The basis poles play the role of tuning parameters, and an indicator of their effect on the fit is used.

The article is organized as follows: Section 2 contains a short review of generalized orthonormal basis transfer functions. In section 3, these functions are employed in a specific parametrization of output error and ARMAX predictors in closed-loop. In section 4, these predictors are introduced in recursive closed-loop identification schemes belonging to the pseudo-linear regression class. These algorithms are detailed and their convergence conditions are provided. Section 5 presents the limit models of the algorithms in the frequency domain. In section 6, an indicator of the basis poles effect on these limit models is introduced (function  $\chi$ ), and the properties of this indicator are given. Last, section 7 presents some relevant simulations showing the interest of these algorithms, and of the function  $\chi$  as an heuristic indicator of the basis poles effect on the bias distribution. Finally, section 8 contains four appendices.

## 2 Orthogonal transfer functions bases and Hambo transforms

At first, some definitions useful for the understanding of this paper are recalled. They are related to orthonormal transfer functions issued from a balanced realization of an all-pass transfer function, as proposed in [7], and to the associated Hambo signal and operator transforms. These notions are presented in [8] (chap. 3 and 12), and [9]: the reader can refer to these references for more information. In the rest of this section, we recall some basic facts for the reader's convenience. Let us consider the Blaschke product  $G_b(z)$ , with  $G_b(z^{-1})G_b(z) = 1$ , given by

$$G_b(z) = \prod_{k=0}^{n_p-1} \frac{-p_k z + 1}{z - p_k} \quad (1)$$

where  $p_k$  ( $|p_k| < 1$ ) are the basis poles (being either real or complex conjugate), and  $n_p$  the number of poles.

This transfer function can be represented by means of a balanced state-space realization  $G_b(z) = D_b + C_b(zI - A_b)^{-1}B_b$ , which satisfies

$$\begin{bmatrix} A_b & B_b \\ C_b & D_b \end{bmatrix}^* \begin{bmatrix} A_b & B_b \\ C_b & D_b \end{bmatrix} = I \quad (2)$$

The orthonormal transfer functions basis proposed by Heuberger and al. [7] consist of the vectors of functions  $V_k$  with size  $(n_p, 1)$  and where  $k \in \mathbb{N}$ , defined as

$$V_k(z) = (zI - A_b)^{-1} B_b G_b^{k-1}(z) \quad (3)$$

One can observe that the basis poles are repeated  $k$  times for each entry of  $V_k(z)$  and these transfer functions form a Hilbert basis of strictly proper transfer functions in  $\mathcal{H}_2(\mathbb{D})$ . Particular configurations of  $n_p$  and  $p_k$  correspond to well known cases:  $n_p = 1, p_0 = 0$  is nothing but the classical  $z^{-1}, z^{-2}, \dots$  basis, and  $n_p = 1, |p_0| < 1$  ( $p_0$  being real) corresponds to the Laguerre basis. Define

$$v_1(t) = A_b^{t-1} B_b \quad v_{k+1}(t) = G_b(q) v_k(t) \quad k \in \mathbb{N}$$

The Hambo signal transform of the sequence  $\{x(t)\}$  belonging to  $l_2$  is  $\{\check{x}(k)\}$  with  $\check{x}(k) = \sum_{t \in \mathbb{N}} v_k(t) x(t)$ . The Hambo operator  $\lambda$  is defined according to  $\lambda^{-1} = G_b(z)$ . If  $n_p > 1$ , the mapping  $\lambda \mapsto z$  is one to  $n_p$  (i.e.  $n_p$ -valued) except at branch points which cannot lie on the unit circle. Define  $N(\lambda) = A_b + B_b(\lambda - D_b)^{-1} C_b$ . It is shown that the eigenvalues  $\{z_j, j = 1, \dots, n_p\}$  of  $N(1/\lambda)$  satisfy the equation  $\lambda^{-1} = G_b(z_j)$ . The  $\lambda$  transform of  $\{x(t)\}$ , denoted as  $\check{X} : \lambda \mapsto \check{X}(\lambda)$ , is given by

$$\check{X}(\lambda) = \sum_{k \in \mathbb{N}} \check{x}(k) \lambda^{-k} \quad (4)$$

Consider a rational transfer function  $H : z \mapsto H(z)$  belonging to  $\mathcal{H}_{2-}$ . Its corresponding Hambo signal transform, denoted as  $\check{H} : \lambda \mapsto \check{H}(\lambda)$ , is given by

$$\check{H}(\lambda) = \sum_{k \in \mathbb{N}} \check{h}(k) \lambda^{-k} \quad (5)$$

where

$$\check{h}(k) = \frac{1}{2\pi i} \oint_{\mathbb{T}} V_k(z) H(z^{-1}) \frac{dz}{z} \quad (6)$$

The Hambo operator transform of  $H$ , which is denoted as  $\tilde{H} : \lambda \mapsto \tilde{H}(\lambda)$  is

$$\tilde{H}(\lambda) = \sum_{\tau=0}^{\infty} h(\tau) N^{\tau}(\lambda) \quad (7)$$

where  $\{h(\tau)\}$  is the impulse response of  $H$ . The following notation is commonly used:

$$\tilde{H}(\lambda) = H(z)|_{z^{-1}=N(\lambda)} \quad (8)$$

and as shown in ([8], Sect. 3.3.10),

$$\tilde{H}(\lambda) = \sum_{j=1}^{n_p} H(z_j) \frac{V_1(1/z_j) V_1^T(z_j)}{V_1^T(z_j) V_1(1/z_j)} \quad (9)$$

The relation from  $\check{H}(\lambda)$  to  $\tilde{H}(\lambda)$  is given by ([8], prop. 12.5)

$$\check{H}(\lambda) = \tilde{H}(\lambda) \frac{\lambda^{-1} C_b^T}{\lambda^{-1} - D_b} \quad (10)$$

Consider now  $\{u(t)\}, \{y(t)\}$ , two sequences in  $l_2$ , and  $H \in \mathcal{H}_{2-}$ , such that  $y(t) = H(q)u(t)$ . Then one has  $\check{y}(\lambda) = \check{H}(\lambda)\check{u}(\lambda)$ . Moreover, the transformation  $\widetilde{(\cdot)} : H \mapsto \tilde{H}$  is an isomorphism of associative and commutative  $\mathbb{R}$ -algebras ([8], sect. 12.4). Explicitly, for any transfers functions  $H_1, H_2 \in \mathcal{H}_{2-}$ , the followings relations hold:

$$\widetilde{(H_1 H_2)}(\lambda) = \tilde{H}_1(\lambda) \tilde{H}_2(\lambda) = \tilde{H}_2(\lambda) \tilde{H}_1(\lambda) \quad (11)$$

and for any  $a, b \in \mathbb{R}$

$$\widetilde{(aH_1 + bH_2)}(\lambda) = a\tilde{H}_1(\lambda) + b\tilde{H}_2(\lambda) \quad (12)$$

### 3 A parameterization for closed-loop identification predictors of output error and ARMAX models

The purpose of this section is to derive expressions of closed-loop predictors for output error and ARMAX models by using the generalized orthonormal basis of transfer functions presented in section 2. Let us consider  $\{u(t)\}, \{y(t)\}$ , the input and output sequences of the system to be identified (true system), defined as

$$y(t) = G_0(q)u(t) + W_0(q)e(t) \quad (13)$$

or

$$y(t) = G_0(q)u(t) + v(t) \quad (14)$$

The white noise sequence  $\{e(t)\}$  is centered and gaussian and  $\{v(t)\}$  is a centred gaussian colored noise independent of  $\{u(t)\}$ . One assumes that  $G_0(q) = \frac{B_0(q)}{A_0(q)}$ , and  $W_0(q) = \frac{C_0(q)}{A_0(q)}$ , where  $A_0, B_0, C_0$  are polynomials. The polynomials  $A_0$  and  $C_0$  are monic with the same degree. The deterministic and stochastic parts of the model are denoted respectively as  $G(q) = \frac{B(q)}{A(q)}$ , and  $W(q) = \frac{C(q)}{A(q)}$ , where  $A, B, C$  are polynomials,  $A$  and  $C$  being monic, with  $d^o(A) = d^o(C) = n_a$  and  $d^o(B) \leq n_a - 1$ . The system is driven by an R-S (causal) controller, such that

$$u(t) = -\frac{R(q)}{S(q)}y(t) \quad (15)$$

where  $S, R$  are coprime polynomials,  $S$  being monic, and it is assumed that  $d^o(R) \leq d^o(S)$ . Let us define the relative degrees  $\delta$  and  $\kappa$  respectively of  $G(q)$  and of the controller, such that  $\delta = d^o(A) - d^o(B) \geq 1$  and  $\kappa = d^o(S) - d^o(R) \geq 0$ .

According to ([10] p. 296), the predicted output at time  $t + 1$ , denoted as  $\hat{y}(t + 1)$ , of the output error model in closed-loop operation, is given by

$$\hat{y}(t+1) = -A^{*\dagger}(q^{-1})\hat{y}(t+1) + q^{-\delta}B^*(q^{-1})\hat{u}(t+1) \quad (16)$$

with

$$S^*(q^{-1})\hat{u}(t) = -q^{-\kappa}R^*(q^{-1})\hat{y}(t) \quad (17)$$

The ARMAX predictor in closed-loop operation is given in ([10], p. 300):

$$\begin{aligned} \hat{y}(t+1) = & -A^{*\dagger}(q^{-1})\hat{y}(t+1) + q^{-\delta}B^*(q^{-1})\hat{u}(t+1) \\ & \dots + \frac{H^{*\dagger}(q^{-1})}{S^*(q^{-1})}\varepsilon(t+1) \end{aligned} \quad (18)$$

where  $H^* = 1 + C^*S^* - A^*S^* - q^{-\delta-\kappa}B^*R^*$ . Moreover, the closed-loop prediction error is  $\varepsilon(t+1) = y(t+1) - \hat{y}(t+1)$ .

Let  $\bar{S}$  be a monic schur polynomial with the same degree as  $S$ , and

$$A_p(q) = \prod_{k=0}^{n_p-1} (q - p_k)^n \quad (19)$$

where  $n = n_a/n_p$  is chosen to be an integer by a suitable choice of  $n_p$  (the number of basis poles), and  $n_a = d^o(A)$ .

**Proposition 1** *The closed-loop output predictor can be written*

$$\hat{y}(t+1) = -\frac{A_I^{*\dagger}(q^{-1})}{A_p^*(q^{-1})}\hat{y}_f(t+1) + \frac{q^{-\delta}B_I^*(q^{-1})}{A_p^*(q^{-1})}\hat{u}_f(t+1) \quad (20)$$

with

$$\bar{S}^*(q^{-1})\hat{y}_f(t+1) = S^*(q^{-1})\hat{y}(t+1) \quad (21a)$$

$$\bar{S}^*(q^{-1})\hat{u}_f(t+1) = -q^{-\kappa}R^*(q^{-1})\hat{y}(t+1) \quad (21b)$$

and

$$A^{*\dagger}(q^{-1}) = A_f^{*\dagger}(q^{-1}) + A_I^{*\dagger}(q^{-1}) \quad (22a)$$

$$B^*(q^{-1}) = B_f^*(q^{-1}) + B_I^*(q^{-1}) \quad (22b)$$

where  $A_f, B_f$  have the same degree as  $A$  and  $B$  respectively, and satisfy the Bézout equation

$$A_f S + B_f R = A_p \bar{S} \quad (23)$$

*Proof:* By (21), and (23), (20), can be written  $-A_I^{*\dagger}\hat{y}(t+1) + q^{-\delta}B_I^*\hat{u}(t+1) = \frac{A_f^*S^* + q^{-\delta-\kappa}B_f^*R^*}{\bar{S}^*}\hat{y}(t+1)$ , which is equivalent to (16) owing to (17).  $\square$

Set

$$A_p'(q) = \prod_{k=0}^{n_p-1} (q - p_k)^{n'} \quad (24)$$

where  $n' = \frac{n_h - n_a}{n_p}$ ,  $n'$  is an integer by a suitable choice of  $n_p$  and  $n_h = d^o(H)$ .

**Proposition 2** *The closed-loop ARMAX predictor can be written*

$$\begin{aligned} \hat{y}(t+1) = & -\frac{A_I^{*\dagger}(q^{-1})}{A_p^*(q^{-1})}\hat{y}_f(t+1) + \dots \\ & \dots + \frac{q^{-\delta}B_I^*(q^{-1})}{A_p^*(q^{-1})}\hat{u}_f(t+1) + \frac{H^{*\dagger}(q^{-1})}{A_p^*(q^{-1})A_p'^*(q^{-1})}\varepsilon_f(t+1) \end{aligned} \quad (25)$$

where

$$\bar{S}^*(q^{-1})\varepsilon_f(t+1) = A_p'^*(q^{-1})\varepsilon(t+1) \quad (26)$$

*Proof:* Similar to the proof of proposition 1.  $\square$

Thus, according to (21), if  $\{\hat{y}(t)\}$  is bounded,  $\{\hat{y}_f(t)\}$  and  $\{\hat{u}_f(t)\}$  are bounded too. Consequently, an unstable controller can be employed, contrary to the classical versions of CLOE and X-CLOE (see [10], pp. 299-301). In order to solve the equation (23) (in which  $A_f$  and  $B_f$  are the unknown terms), it is necessary to specify  $\bar{S}$  (the specification of  $A_p$  results from the basis poles selection). The zeros of  $\bar{S}$  are chosen according to the principle of  $\rho$ -stability developed in [3]. Let  $\rho \geq 1$ ,  $\bar{S}(z/\rho)$  is the stable spectral factor of  $S(z/\rho)$ , i.e

$$S(z/\rho)S(\rho/z) = \bar{S}(z/\rho)\bar{S}(\rho/z) \quad (27)$$

where  $z \mapsto \bar{S}(z/\rho)$  has all its zeros inside the unit circle, assuming that  $\rho$  has been chosen so that  $z \mapsto \bar{S}(z/\rho)$  has no zero  $\zeta$  such that  $|\zeta| = 1$ . In term of zeros placement, the choice of  $\bar{S}$  is equivalent to leaving unchanged the zeros of  $S$  that are inside the circle with radius  $1/\rho$ , and to reflect the other zeros with respect to this circle. In any case,  $\rho$  should be chosen close to 1, and can be exactly equal to 1 if the controller has no pole on the unit circle.

**Proposition 3** *Let  $\mu_k, \nu_k, k \in [1, n]$  be some vectors of size  $(n_p, 1)$  with real entries. The closed-loop output error predictor (20) that we call the H-CLOE predictor*

( $H$  stands for Hambo) can be expressed as

$$\hat{y}(t+1) = - \sum_{k=1}^n \mu_k^T V_k(q) \hat{y}_f(t+1) + \sum_{k=1}^n \nu_k^T V_k(q) \hat{u}_f(t+1) \quad (28)$$

*Proof:* Let  $P_{n_a}$  be the space of all polynomials of degree inferior or equal to  $n_a - 1$ . By (3), the entries of the column  $A_p V_k$  belong to  $P_{n_a}$  and form a basis of this space by the orthogonality condition of the transfer functions  $V_k$ , with  $k = 1, \dots, n$ . Since  $A_p$  is the denominator of  $V_n$ , there exist vectors  $\mu_k, \nu_k$  with real entries and size  $(n_p, 1)$  such that  $\frac{A_L^*(q^{-1})}{A_p^*(q^{-1})} = \sum_{k=1}^n \mu_k^T V_k(q)$  and  $q^{-\delta} \frac{B_L^*(q^{-1})}{A_p^*(q^{-1})} = \sum_{k=1}^n \nu_k^T V_k(q)$ , hence the result.  $\square$

**Proposition 4** Let  $\gamma_k, k \in [1, n + n']$  be some vectors with real entries and size  $(n_p, 1)$ . The closed-loop AR-MAX predictor called in the sequel H-XCLOE can be put under the form

$$\hat{y}(t+1) = - \sum_{k=1}^n \mu_k^T V_k(q) \hat{y}_f(t+1) + \sum_{k=1}^n \nu_k^T V_k(q) \hat{u}_f(t+1) + \sum_{k=1}^{n_h} \gamma_k^T V_k(q) \varepsilon_f(t+1) \quad (29)$$

*Proof:* Similar to the proof of proposition 3 since one can write  $\frac{H^*(q^{-1})}{A_p^*(q^{-1})} = \sum_{k=1}^{n+n'} \gamma_k^T V_k(q)$ .  $\square$

Remark 1: Note that causality is respected in (28), (29), since  $V_k$  is a strictly proper transfer operator.

Remark 2: The Bézout equation (23) has solutions provided that  $R$  and  $S$  are coprime, which means that the controller is expressed in a minimal form. It is preferable to solve this equation with the Hambo operator (in order to have a well conditioned problem). The corresponding algebra is explained in Appendix A.

Remark 3: Open-loop predictors of output error and AR-MAX models can trivially be obtained by taking  $R = 0$ ,  $S = \bar{S} = 1$ . The convergence conditions and the limit models of the corresponding open-loop algorithms can be deduced from the general case treated in sections 4 and 5 below.

## 4 Algorithms and their convergence conditions

### 4.1 Recursive structure of the algorithms

The predictors presented in section 2 can now be used in identification schemes belonging to the pseudo-linear class. In the following,  $\theta_0$  is the parameters vector of the

true system,  $\theta$  is the parameters vector of the predictor,  $\hat{\theta}(t)$  is the estimated parameters vector at time  $t$ ,  $\theta^*$  is the limit estimated parameters vector (if it exists), and  $\phi(t, \theta)$  is the regressor of the predictor. According to (28) and (29), the predicted output of H-CLOE and of H-XCLOE can be put in the form (which corresponds to a pseudo-linear regression structure)

$$\hat{y}(t+1) = \theta^T \phi(t, \theta) \quad (30)$$

The expressions of  $\phi(t, \theta)$  and  $\theta$  are given in sections 4.2 and 4.3. In the context of pseudo-linear regression,  $\theta$  is generally computed in a recursive manner by means of the parameters adaptation algorithm ([10], p. 101-102):

$$\hat{\theta}(t+1) = \hat{\theta}(t) + F(t) \phi(t, \theta) \varepsilon(t+1) \quad (31a)$$

$$F^{-1}(t+1) = \lambda_1 F^{-1}(t) + \lambda_2 \phi(t, \theta) \phi^T(t, \theta) \quad (31b)$$

where  $\varepsilon(t+1) = y(t+1) - \hat{y}(t+1)$  is the (a-posteriori) prediction error,  $F(t)$  the adaptation gain (positive definite matrix),  $0 < \lambda_1 \leq 1$ ,  $0 \leq \lambda_2 < 2$  are the forgetting factors.

It will be shown that the expressions of the prediction error of H-CLOE and H-XCLOE schemes can be put in the form

$$\varepsilon(t+1) = \mathcal{H}(q^{-1}) (\theta_0 - \theta)^T \phi(t, \theta) + w(t+1) \quad (32)$$

where  $\mathcal{H} : z^{-1} \mapsto 1 + \sum_{k=1}^{\infty} h_k z^{-k}$  is a rational function, and  $\{w(t+1)\}$  is a centred stochastic disturbance such that  $\mathbf{E}[w(t+1) \phi(t, \theta)] = 0$ . According to ([12] p. 396), and Theorem 4.1 of [10], if one assumes that  $\lambda_1 = 1$ , and that  $\hat{\theta}(t)$  generated by the algorithm always belongs to the domain  $\mathcal{D}_S$  in which the stationary processes  $\{\phi(t, \theta)\}$  and  $\{\varepsilon(t+1, \theta)\}$  can be defined, and if the transfer function  $\mathcal{H}(z^{-1}) - \frac{\lambda_2}{2}$  is strictly positive real, one has

$$\text{Prob} \quad \left\{ \lim_{t \rightarrow \infty} \hat{\theta}(t) \in D_C \right\} = 1$$

where  $D_C = \{\theta : \phi^T(t, \theta)[\theta_0 - \theta] = 0\}$  is the convergence domain.

Therefore it is crucial to make explicit the expression of  $\mathcal{H}$  in H-CLOE and H-XCLOE algorithms : This is the purpose of the next subsections. Note that the strictly real positivity of  $\mathcal{H}(z^{-1}) - \frac{\lambda_2}{2}$  is a sufficient convergence condition.

### 4.2 Output error predictor and the H-CLOE algorithm

In the H-CLOE algorithm, the regressor  $\phi(t, \theta)$ , the parameters vector of the model  $\theta$  and the true system  $\theta_0$  are respectively

$$\phi^T(t, \theta) = [-V_1^T(q) \hat{y}_f(t+1) - V_2^T(q) \hat{y}_f(t+1) \cdots \cdots V_1^T(q) \hat{u}_f(t+1) V_2^T(q) \hat{u}_f(t+1) \cdots], \quad (33)$$

$$\theta^T = [\mu_1^T \mu_2^T \cdots \nu_1^T \nu_2^T \cdots] \quad (34a)$$

$$\theta_0^T = [\mu_{01}^T \mu_{02}^T \cdots \nu_{01}^T \nu_{02}^T \cdots] \quad (34b)$$

Moreover, let us define  $S_{yp0}$  as the sensitivity function of the closed-loop, where

$$S_{yp0}(z^{-1}) = \frac{A_0^*(z^{-1})S^*(z^{-1})}{A_0^*(z^{-1})S^*(z^{-1}) + z^{-\delta-\kappa}B_0^*(z^{-1})R^*(z^{-1})} \quad (35)$$

**Lemma 1** A sufficient convergence condition of the H-CLOE algorithm is that

$$\frac{A_p^*(z^{-1})\bar{S}^*(z^{-1})}{A_0^*(z^{-1})S^*(z^{-1}) + z^{-\delta-\kappa}B_0^*(z^{-1})R^*(z^{-1})} - \frac{\lambda_2}{2} \quad (36)$$

be a strictly positive real transfer function.

*Proof:* Let  $y_f(t)$ ,  $u_f(t)$  be such that  $\bar{S}^*(q^{-1})y_f(t) = S^*(q^{-1})y(t)$  and  $\bar{S}^*(q^{-1})u_f(t) = -q^{-\kappa}R^*(q^{-1})y(t)$ . Let  $A_{0I}$  be a monic polynomial with  $d^o(A_{0I}) = n_a$ , and  $B_{0I}$  a polynomial with  $d^o(B_{0I}) = d^o(B_0)$ , these polynomials satisfying  $A_{0I}^{\dagger}(q^{-1}) = A_0^{\dagger}(q^{-1}) + A_f^{\dagger}(q^{-1})$  and  $B_0^*(q^{-1}) = B_{0I}^*(q^{-1}) + B_f^*(q^{-1})$ . The parameters vectors of the true system  $\mu_{0k}$ ,  $\nu_{0k}$  are such that  $\frac{A_{0I}^{\dagger}(q^{-1})}{A_p^*(q^{-1})} = \sum_{k=1}^n \mu_{0k}^T V_k(q)$  and  $q^{-\delta} \frac{B_{0I}^*(q^{-1})}{A_p^*(q^{-1})} = \sum_{k=1}^n \nu_{0k}^T V_k(q)$ . On the other hand, by using the same procedure as that leading to (20), one can write from (14),  $y(t+1) = -\frac{A_{0I}^{\dagger}(q^{-1})}{A_p^*(q^{-1})}y_f(t+1) + \frac{B_{0I}^*(q)}{A_p^*(q)}u_f(t+1-\delta) + \frac{A_0^*(q^{-1})S^*(q^{-1})}{A_p^*(q^{-1})\bar{S}^*(q^{-1})}v(t+1)$ . By combining this latter equation with (20) one obtains  $\varepsilon(t+1) = \mathcal{H}(q^{-1})(\theta_0 - \theta)^T \phi(t, \theta) + S_{yp0}(q^{-1})v(t+1)$ , with  $\mathcal{H}(z^{-1}) = \frac{A_p^*(z^{-1})\bar{S}^*(z^{-1})}{A_0^*(z^{-1})S^*(z^{-1}) + z^{-\delta-\kappa}B_0^*(z^{-1})R^*(z^{-1})}$ . Owing to this expression of  $\mathcal{H}$ , the sufficient convergence condition is obtained from subsection 4.1.  $\square$

Note that (36) is a generalization of the original CLOE algorithm which correspond to the situation:  $A_p^* = 1$  and  $\bar{S}^* = S^*$

### 4.3 ARMAX predictor and the H-XCLOE algorithm

The regressor of the H-XCLOE algorithm, the model parameters vector and the true parameters vector are given by:

$$\begin{aligned} \phi^T(t, \theta) = & [-V_1^T(q)\hat{y}_f(t+1) - V_2^T(q)\hat{y}_f(t+1) \cdots \\ & \cdots V_1^T(q)\hat{u}_f(t+1) V_2^T(q)\hat{u}_f(t+1) \cdots \\ & \cdots V_1^T(q)\varepsilon_f(t+1) V_2^T(q)\varepsilon_f(t+1) \cdots] \end{aligned}$$

$$\theta^T = [\mu_1^T \mu_2^T \cdots \nu_1^T \nu_2^T \cdots \gamma_1^T \gamma_2^T \cdots] \quad (37a)$$

$$\theta_0^T = [\mu_{01}^T \mu_{02}^T \cdots \nu_{01}^T \nu_{02}^T \cdots \gamma_{01}^T \gamma_{02}^T \cdots] \quad (37b)$$

**Lemma 2** A sufficient convergence condition of H-XCLOE is that

$$\frac{A_p^*(z^{-1})\bar{S}^*(z^{-1})}{C^*(z^{-1})S^*(z^{-1})} - \frac{\lambda_2}{2} \quad (38)$$

be a strictly positive real transfer function.

*Proof:* Set  $H_0^* = 1 + C_0^*S^* - A_0^*S^* - q^{-\delta-\kappa}B_0^*R_0^*$ , and define  $\gamma_{0k}$  such that  $\frac{H_0^{*\dagger}(q^{-1})}{A_p^*(q^{-1})A_p^*(q^{-1})} = \sum_{k=1}^{n+n'} \gamma_{0k}^T V_k(q)$ . Since  $\bar{S}^*(q^{-1})y_f(t) = S^*(q^{-1})y(t)$  and  $\bar{S}^*(q^{-1})u_f(t) = -q^{-\kappa}R^*(q^{-1})y(t)$ , by using (13) one has  $y(t+1) = -\frac{A_{0I}^{\dagger}(q^{-1})}{A_p^*(q^{-1})}y_f(t+1) + \frac{q^{-\delta}B_{0I}^*(q^{-1})}{A_p^*(q^{-1})}u_f(t+1) + \cdots + \frac{C_0^*(q^{-1})S^*(q^{-1})}{A_p^*(q^{-1})\bar{S}^*(q^{-1})}e(t+1)$ . Combining this expression with (25) one obtains  $\varepsilon(t+1) = \mathcal{H}(q^{-1})(\theta_0 - \theta)^T \phi(t, \theta) + e(t+1)$  with  $\mathcal{H}(q^{-1}) = \frac{A_p^*(q^{-1})\bar{S}^*(q^{-1})}{C^*(q^{-1})S^*(q^{-1})}$ , and one can directly use the result of subsection 4.1.  $\square$

The expression (38) shows that H-XCLOE is again a generalization of the classical X-CLOE where  $A_p^* = 1$  and  $\bar{S}^* = S^*$ .

### 4.4 Inclusion of excitation signals

The excitation is an additional signal either to the system input or to the reference. In case of an input system excitation  $r_u(t)$ , or a reference excitation  $r_r(t)$ , one has respectively

$$S^*(q^{-1})\hat{u}(t) = -q^{-\kappa}R^*(q^{-1})\hat{y}(t) + S^*(q^{-1})r_u(t) \quad (39a)$$

$$S^*(q^{-1})\hat{u}(t) = -q^{-\kappa}R^*(q^{-1})\hat{y}(t) + q^{-\kappa}R^*(q^{-1})r_r(t) \quad (39b)$$

And the predicted output expression of the H-CLOE algorithm with excitation signals is:

$$\begin{aligned} \hat{y}(t+1) = & -\sum_{k=1}^n \mu_k^T V_k(q)\hat{y}_f(t+1) + \\ & \cdots + \sum_{k=1}^n \nu_k^T V_k(q)\hat{u}_f(t+1) + r_{uf}(t) + r_{rf}(t) \end{aligned} \quad (40)$$

where  $r_{uf}(t)$ ,  $r_{rf}(t)$  satisfy the relations

$$A_p^*(q^{-1})\bar{S}^*(q^{-1})r_{uf}(t) = q^{-\delta}B_f^*(q^{-1})S^*(q^{-1})r_u(t) \quad (41a)$$

$$A_p^*(q^{-1})\bar{S}^*(q^{-1})r_{rf}(t) = q^{-\delta-\kappa}B_f^*(q^{-1})R^*(q^{-1})r_r(t) \quad (41b)$$

The expression of the (a-posteriori) predicted output:  $\hat{y}(t+1) = \hat{\theta}(t+1)\phi(t, \theta) + r_{uf}(t) + r_{rf}(t)$  is not of the form required by the pseudo-linear regression if  $r_{uf}(t) + r_{rf}(t) \neq 0$ . Defining the modified measured output  $\bar{y}(t+1) = y(t+1) - r_{uf}(t) - r_{rf}(t)$ , we get

$$\varepsilon^0(t+1) = \bar{y}(t+1) - \hat{\theta}^T(t)\phi(t, \theta) \quad (42)$$

which has the usual expression and is called the a-priori prediction error. The a-posteriori prediction error  $\varepsilon(t+1)$  is given by  $\varepsilon(t+1) = \frac{\varepsilon^0(t+1)}{1+\phi^T(t)F(t)\phi(t)}$  (see [10], p. 101). This substitution applies *mutatis mutandis* for the H-XCLOE algorithm.

## 5 Limit models in the frequency domain

The stationary condition of the parameter adaptation algorithm is (see [12], p. 222):

$$\mathbf{E}[\varepsilon(t+1)\phi(t, \theta)] = 0 \quad (43)$$

The regressor  $\phi(t, \theta)$  depends on the estimated parameters. As shown in [20], Condition (43) does not imply in general the minimization of  $\mathbf{E}[\varepsilon^2(t, \theta)]$ . This is the major difference with prediction error methods (PEM) that aim directly at minimizing this latter expression. For a full order model, the signal (generally non measurable) whose variance is effectively minimized if Condition (43) is satisfied is what we call the equivalent prediction error (see [20]) denoted by  $\varepsilon_E(t+1)$  such that

$$\theta^* = \text{Argmin} \mathbf{E}[\varepsilon_E^2(t, \theta)] \quad (44)$$

The signal  $t \mapsto \mathbf{E}[\varepsilon_E^2(t, \theta)]$  is ergodic, thus the mapping  $\theta \mapsto \mathbf{E}[\varepsilon_E^2(t, \theta)]$  is a strictly convex and coercive function independent of  $t$ , and therefore the minimum of this function exists and is achieved for a unique value  $\theta^*$  of  $\theta$ . As shown in [20] this equivalent prediction error is related to the prediction error  $\varepsilon(t)$ . Let  $\mathcal{Q}(q^{-1}, \theta)$  be a transfer function operator, ratio of two polynomials in  $q^{-1}$  with constant term equal to 1, that satisfies

$$\mathcal{Q}(q^{-1}, \theta) \frac{\partial \varepsilon(t+1, \theta)}{\partial \theta} = -\phi(t, \theta) \quad (45)$$

For the output error model (in closed-loop) one has (see [20])

$$\varepsilon_E(t+1, \theta) = \mathcal{Q}(q^{-1}, \theta)\varepsilon(t+1) + (1 - \mathcal{Q}(q^{-1}, \theta))S_{yp}(q^{-1})v(t+1) \quad (46)$$

and for the ARMAX model

$$\varepsilon_E(t+1, \theta) = \mathcal{Q}(q^{-1}, \theta)\varepsilon(t+1) + (1 - \mathcal{Q}(q^{-1}, \theta))e(t+1) \quad (47)$$

Define the sensitivity function operator  $S_{up0}(q^{-1})$  such that

$S_{up0}(q^{-1}) = \frac{-A_0^*(q^{-1})R^*(q^{-1})}{A_0^*(q^{-1})S^*(q^{-1})+q^{-\delta-\kappa}B_0^*(q^{-1})R^*(q^{-1})}$ , and let  $T_{ur0}(q^{-1})$  be the transfer function operator from the excitation signal to the system input. One has  $T_{ur0} = S_{yp0}$  in case of an additive excitation to the system input, and  $T_{ur0} = -S_{up0}$  in case of an excitation on the reference. In the following,  $\Phi_r$  is the excitation power spectrum density (PSD) (of  $r_u$  or  $r_r$  depending on the location of this excitation), and  $\Phi_e$  the PSD of  $\{e(t)\}$ . The two following results hold:

**Proposition 5** *The limit model of the H-CLOE algorithm is given by*

$$\theta^* = \text{Argmin} \int_{-\pi}^{+\pi} \left| \frac{A(e^{i\omega})S(e^{i\omega})}{A_p(e^{i\omega})\bar{S}(e^{i\omega})} \right|^2 \times |G_0(e^{i\omega}) - G(e^{i\omega})|^2 |T_{ur0}(e^{i\omega})|^2 \Phi_r(\omega) d\omega \quad (48)$$

*Proof:* One has to calculate  $\mathcal{Q}(q^{-1}, \theta)$ . For the H-CLOE algorithm, one verifies easily that:

$\mathcal{Q}(q^{-1}, \theta) = \frac{A^*(q^{-1})S^*(q^{-1})+q^{-\delta-\kappa}B^*(q^{-1})R^*(q^{-1})}{A_p^*(q^{-1})\bar{S}^*(q^{-1})}$ . On the other hand, the H-CLOE predictor expression is exactly the same as that of the classical CLOE algorithm (see [10], p. 308):

$\varepsilon(t) = S_{yp}(G_0 - G)[S_{yp0}r_u(t) - S_{up0}r_r(t)] + S_{yp0}v(t)$ , where  $S_{yp}(q^{-1}) = \frac{A^*(q^{-1})S^*(q^{-1})}{A^*(q^{-1})S^*(q^{-1})+q^{-\delta-\kappa}B^*(q^{-1})R^*(q^{-1})}$ . By combining with (46), one obtains the expression of the equivalent prediction error

$\varepsilon_E(t) = \frac{A^*S^*}{A_p^*\bar{S}^*}(G_0 - G)[S_{yp0}r_u(t) - S_{up0}r_r(t)] + S_{yp0}v(t)$ , hence the equation (48), since  $d^o(A) = d^o(A_p)$ , and  $d^o(S) = d^o(\bar{S})$ .  $\square$

**Proposition 6** *The limit model of the H-XCLOE algo-*



ritms is given by

$$\theta^* = \underset{\omega}{\operatorname{Argmin}} \int_{-\pi}^{+\pi} \left| \frac{A(e^{i\omega})S(e^{i\omega})}{A_p(e^{i\omega})\tilde{S}(e^{i\omega})} \right|^2 \left[ |G_0(e^{i\omega}) - G(e^{i\omega})|^2 |T_{ur0}(e^{i\omega})|^2 \Phi_r(\omega) + |W_0(e^{i\omega}) \frac{S_{yp0}(e^{i\omega})}{S_{yp}(e^{i\omega})} - W(e^{i\omega})|^2 \Phi_e(\omega) \right] d\omega \quad (49)$$

*Proof:* One verifies that:  $C^*(q^{-1})S^*(q^{-1})\frac{\partial \varepsilon(t+1)}{\partial \theta} = -A_p^*(q^{-1})\tilde{S}^*(q^{-1})\phi(t)$ , thus:  $\mathcal{Q}(q^{-1}, \theta) = \frac{C^*(q^{-1})S^*(q^{-1})}{A_p^*(q^{-1})\tilde{S}^*(q^{-1})}$ . On the other hand, the prediction error of the classical X-CLOE algorithm is  $\varepsilon(t) = \frac{A^*}{C^*} [(G_0 - G)(S_{yp0}r_u(t) - S_{up0}r_r(t)) + (W_0 \frac{S_{yp0}}{S_{yp}} - W)e(t)] + e(t)$ . By combining with (47),  $\varepsilon_E(t) = \frac{A^*S^*}{A_p^*\tilde{S}^*} [(G_0 - G)(S_{yp0}r_u(t) - S_{up0}r_r(t))] + \frac{A^*S^* + q^{-\delta-\kappa}B^*R^*}{A_p^*\tilde{S}^*} (W_0S_{yp0} - WS_{yp})e(t) + e(t)$ , and the result in the frequency domain follows.  $\square$

## 6 Effect of the basis poles on the limit models: The function $\chi$

There exists a bijective and strictly increasing function  $\beta : ]-\pi, \pi] \rightarrow ]-n_p\pi, n_p\pi]$  such that  $G_b(e^{i\omega}) = e^{-i\beta(\omega)}$  for all  $\omega \in ]-\pi, \pi]$ , [13], ([8] chap. 9). Putting  $\omega_\lambda := \beta(\omega)$ ,  $\omega_\lambda$  is called the Hambo frequency and the mapping  $\mathbb{T} \rightarrow ]-n_p\pi, n_p\pi] : \lambda \mapsto \omega_\lambda$  is  $n_p$ -valued. The derivative of this function plays a major role, owing to the property (see [8], chap. 9)

$$\beta'(\omega) = V_1^T(e^{i\omega})V_1(e^{-i\omega}) \quad (50)$$

The function  $\beta'$  is a particular expression of the reproducing kernel of the Hilbert space associated with the orthogonal transfer functions basis (see [8], chap.2). The Hambo operator transform of  $G_0(z)$  is denoted  $\tilde{G}_0(\lambda)$ . The Hambo operator transform of  $G(z) = \frac{B(z)/A_p(z)}{A(z)/A_p(z)}$ , denoted as  $\tilde{G}(\lambda)$ , is given by the expression  $\tilde{G}(\lambda) = \tilde{\mathcal{A}}^{-1}(\lambda)\tilde{\mathcal{B}}(\lambda)$ , where  $\tilde{\mathcal{A}}(\lambda)$ , and  $\tilde{\mathcal{B}}(\lambda)$  are polynomial matrices in  $\lambda$ , and are the Hambo operator transforms of  $A(z)/A_p(z)$  and  $B(z)/A_p(z)$  respectively, see proposition 8 below.

In the following, it is assumed that  $\{r(t)\}$  and  $\{v(t)\}$  have rational spectra, and that the power spectrum density (PSD) of  $\{\varepsilon_E(t)\}$ , denoted as  $\Phi_{\varepsilon_E}(\omega)$ , is rational too. Additionally, the PSD of  $\{e(t)\}$  is denoted as  $\Phi_e(\omega)$ .

**Proposition 7** *Under the assumption that  $\{r(t)\}$  is independent of  $\{v(t)\}$ , one has*

$$\theta^* = \underset{\omega}{\operatorname{Argmin}} \int_{-n_p\pi}^{n_p\pi} \Psi_{\varepsilon_E}(\omega_\lambda) d\omega_\lambda \quad (51)$$

where

• For H-CLOE

$$\Psi_{\varepsilon_E}(\omega_\lambda) = \|\tilde{\mathcal{A}}(e^{i\omega_\lambda})\|_2^2 \|\tilde{\mathcal{S}}(e^{i\omega_\lambda})\tilde{\mathcal{S}}^{-1}(e^{i\omega_\lambda})\|_2^2 \times \dots \|\tilde{G}_0(e^{i\omega_\lambda}) - \tilde{G}(e^{i\omega_\lambda})\|_2^2 \|\tilde{T}_{ur0}(e^{i\omega_\lambda})\|_2^2 \Psi_r(\omega_\lambda) \quad (52)$$

• For H-XCLOE

$$\Psi_{\varepsilon_E}(\omega_\lambda) = \|\tilde{\mathcal{A}}(e^{i\omega_\lambda})\|_2^2 \|\tilde{\mathcal{S}}(e^{i\omega_\lambda})\tilde{\mathcal{S}}^{-1}(e^{i\omega_\lambda})\|_2^2 \times \dots \left( \|\tilde{G}_0(e^{i\omega_\lambda}) - \tilde{G}(e^{i\omega_\lambda})\|_2^2 \|\tilde{T}_{ur0}(e^{i\omega_\lambda})\|_2^2 \Psi_r(\omega_\lambda) + \dots \|\tilde{S}_{yp0}(e^{i\omega_\lambda})\tilde{S}_{yp}^{-1}(e^{i\omega_\lambda})\tilde{W}_0(e^{i\omega_\lambda}) - \tilde{W}(e^{i\omega_\lambda})\|_2^2 \Psi_e(\omega_\lambda) \right) \quad (53)$$

with  $\Psi_r(\omega_\lambda) = \Phi_r(\omega) \frac{1}{\beta'(\omega)} \Big|_{\omega=\beta^{-1}(\omega_\lambda)}$ ,

and  $\Psi_e(\omega_\lambda) = \Phi_e(\omega) \frac{1}{\beta'(\omega)} \Big|_{\omega=\beta^{-1}(\omega_\lambda)}$

and where  $\tilde{(\cdot)}$  is the Hambo operator transform of the transfer matrix in parentheses, with  $\|\tilde{(\cdot)}\|_2^2 = |(\cdot)|_{\omega=\beta^{-1}(\omega_\lambda)}^2$ .

*Proof:* The proof is given in appendix B.  $\square$

Equations (51), (52), (53) provides some intrinsic expressions, independent of the pole basis, showing how the discrepancy between the true system and the model is weighted in the Hambo frequency domain. These equations are similar to those of native PLR closed-loop algorithms (corresponding to equations (48) and (49) in which all basis poles are set to 0 or  $A_p(e^{i\omega}) = 1$ ), since  $\|\tilde{\mathcal{S}}(e^{i\omega_\lambda})\tilde{\mathcal{S}}^{-1}(e^{i\omega_\lambda})\|_2^2 \approx 1$  at almost every frequency (see Section 3 above). Instead of considering the effect of the basis poles induced by the weighting function  $|1/A_p(e^{i\omega})|^2$  that appears in the limit models (48) and (49) of section 5, it is more convenient to analyse the generic expressions (52), (53) with respect to the frequency distortion from  $\omega$  to  $\omega_\lambda$ . The function  $\beta'$  characterizes this distortion, which is well known for the Laguerre basis (see [14]). From (50), one shows easily that [8] (p. 222), [13]

$$\beta'(\omega) = \sum_{k=0}^{n_p-1} \beta'_k(\omega) \quad \beta'_k(\omega) = \frac{1 - |p_k|^2}{|1 - \bar{p}_k e^{i\omega}|^2} \quad (54)$$

Since most identified systems are represented in Bode diagrams with a logarithmic scale  $\bar{\omega} = \log(\omega)$ , one should study the distortion from this logarithmic frequency scale  $\bar{\omega}$  to the Hambo frequency scale  $\omega_\lambda$ . Let us denote as  $\chi$  the function which expresses the dilatation rate from the  $\bar{\omega}$  scale to  $\omega_\lambda$  scale. Note that

$$d\omega_\lambda = e^{\bar{\omega}} \beta'(e^{\bar{\omega}}) d\bar{\omega} \quad (55)$$

This function  $\chi$  is given by

$$\chi(\omega) = \frac{1}{\pi} \omega \beta'(\omega) = \frac{1}{\pi} e^{\bar{\omega}} \beta'(e^{\bar{\omega}}) \quad (56)$$

where the term  $\frac{1}{\pi}$  is introduced for normalization.

**Lemma 3** *The following conservation principle holds:*

$$\int_{-\infty}^{\log(\pi)} \chi(e^{\bar{\omega}}) d\bar{\omega} = 1 \quad (57)$$

*Proof:* From (50)  $\beta'(\omega) = V_1^T(e^{i\omega})V_1(e^{-i\omega})$ , and because of the orthonormality of  $V_1(e^{i\omega})$  the result is obtained immediately, see [8], p. 89.  $\square$

**Theorem 1** Consider  $\chi_k(\omega) = \frac{1}{\pi} \omega \beta'_k(\omega)$  and define the  $k$ -th basis pole from its natural frequency  $\omega_{ok}$  and its damping  $\xi_k$  as follows:  $p_k = \rho_k e^{i\sigma_k}$  with  $\rho_k = e^{-\xi_k \omega_{ok}}$  and  $\sigma_k = \sqrt{1 - \xi_k^2} \omega_{ok}$ . Assume that  $\xi_k^2 \geq 1 - \frac{\pi^2}{4\omega_{ok}^2}$ . One has the following results:

- (1) If  $\cosh(\xi_k \omega_{ok}) - \sqrt{1 - \xi_k^2} \omega_{ok} < \frac{\pi}{2}$ ,  $\chi_k$  has a unique maximum on  $[0, \pi]$ . Additionally if:  $\cosh(\xi_k \omega_{ok}) + \cos(\sqrt{1 - \xi_k^2} \omega_{ok}) - \pi \sin(\sqrt{1 - \xi_k^2} \omega_{ok}) > 0$ ,  $\chi_k$  has a local minimum on  $]0, \pi[$ .
- (2) If  $p_k$  is real ( $\xi_k = 1$ ), and if  $\frac{\pi - \sqrt{\pi^2 - 4}}{2} < p_k < 1$ ,  $\chi_k$  has a unique maximum on  $[0, \pi]$ , and a local minimum on  $]0, \pi[$ . If  $p_k \leq \frac{\pi - \sqrt{\pi^2 - 4}}{2}$ ,  $\chi_k$  is an increasing function on  $[0, \pi]$ , and has its maximum at  $\omega = \pi$ .

*Proof:* Given in appendix C.  $\square$

This result is illustrated in Fig. 1 below, in case (2) with  $n_p = 1$ .

**Corollary 1** Let  $\omega_{max}$  the frequency for which  $\chi_k(\omega)$  is maximum. If  $\omega_{ok} \rightarrow 0$ , one has:

$$\omega_{max} = \omega_{ok} + o(|\omega_{ok}|) \quad (58)$$

*Proof:* Provided in appendix D.  $\square$

The frequencies in the  $\bar{\omega}$  scale for which the dilatation rate (function  $\chi$ ) is maximum are over-weighted in the minimization problem, and a better model fit around these frequencies can be expected, for given excitation and noise spectra. On the contrary an area in the  $\bar{\omega}$  domain with small values of  $\chi$  will be under-weighted. Thus it can be inferred that the function  $\chi$  plays the role of a heuristic indicator of the pole basis effect on the model adjustment in the frequency domain, and this is confirmed in section 7. Figure 1 displays the frequency distortion rate  $\chi = \chi(\omega, p_0)$  corresponding to Laguerre bases for many values of the Laguerre poles. One can observe the conservation principle of Lemma 3, and the

frequency of the maximum equal roughly to the frequency of the pole (for low frequency poles). Figure 2 shows three examples of  $\chi$ , for various basis poles corresponding to 1) one pole basis  $p_0 = 0.99$ , 2) two poles basis with  $p_0 = 0.9, p_1 = 0.999$ , 3) three poles basis with  $p_0 = 0.9, p_1 = 0.99, p_2 = 0.999$ .

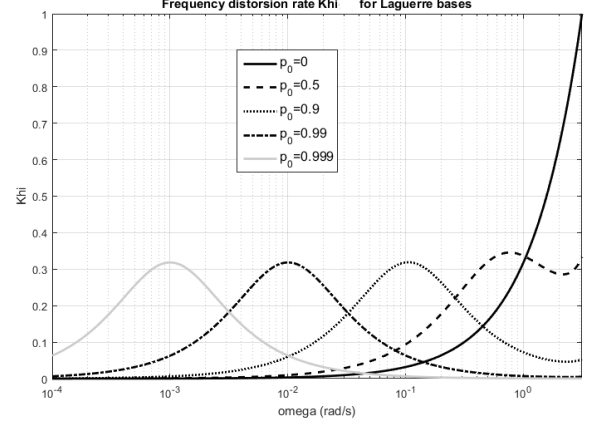


Fig. 1. Frequency distortion rate  $\chi$  for Laguerre bases

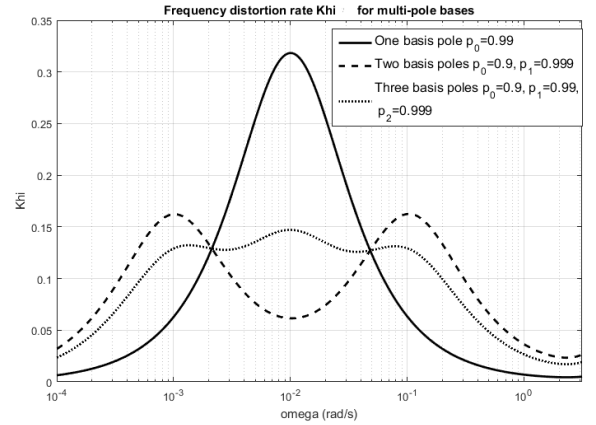


Fig. 2. Frequency distortion rate  $\chi$  for multi-poles bases

## 7 Simulation results

### 7.1 Comparison of CLOE and H-CLOE for a simple system

The aim of this first simulation is to compare H-CLOE to CLOE and AF-CLOE (an adaptive filtered version of CLOE, see [10], p. 299-300) for the identification of a simple multi-scale system having two resonating modes at frequencies  $10^{-2}$  rad/s and  $0.26$  rad/s (sample frequency  $1$  s). The poles of this system are  $0.9575 \pm 0.0097i$  and  $0.8318 \pm 0.2211i$ , its static gain is  $1$ , and it has three zeros  $[0.9712; 0.9696; 1.3578]$ . The controller is a gain equal to  $1$ . The system is excited by a wide spectrum PRBS (added on the input), and is disturbed by a white noise. In closed-loop, the signal/noise (variance) ratio

is equal to 12 dB. The sample number is  $N = 10^4$ . When identified with CLOE (Fig. 3 (top)), the resulting models are correct only in high frequency, whereas the low frequency mode is never captured (over 20 realizations): This is the typical drawback of native CLOE family schemes (see [20]). This could be imputed to the fact that the (only sufficient) convergence condition is not satisfied here. However, the same simulation carried out without any stochastic disturbance with X-CLOE (which has no convergence condition in a deterministic context) leads to similar identified models. If the AF-CLOE is employed (which aims at relaxing the convergence condition of CLOE), the performances are not better (Fig. 3 (middle)). On the contrary, if one uses H-CLOE, with a two poles basis in order to encompass the frequencies of interest of this system (the poles frequencies are set respectively to  $0.03 \text{ rad/s}$  and  $0.3 \text{ rad/s}$ , corresponding to  $p = [0.97; 0.74]$ ), one can observe in (Fig. 3 (bottom)) that the identified models are hardly distinguishable from the true system. The pole selection results from an initial guess, and the algorithm convergence depends on it.

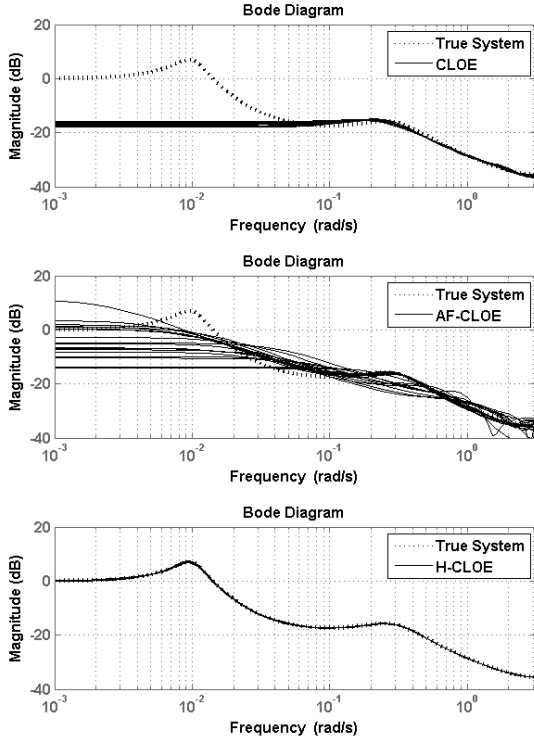


Fig. 3. Identification with CLOE (top) and AF-CLOE (middle), and with H-CLOE (bottom), 20 realizations

## 7.2 Use of a double integrator in the controller, and bias management

The purpose of this second set of simulations is to show that the algorithms proposed in this paper are able to include a double integrator, and to provide reduced order models in a prescribed frequency range, according to the poles of the orthonormal basis. Moreover, these simulations show the interest of the function  $\chi$  as an indicator of the basis poles effect on the bias distribution. The academic system to be identified has an order equal to 12, with modes frequencies lying from  $1.25 \cdot 10^{-3} \text{ rad/s}$  to  $1.26 \text{ rad/s}$  (three decades), the sampling period being equal to 1s. The poles, zeros of this system are displayed in Table 1 (the static gain being equal to 1):

Poles		Zeros	
Frequency (rad/s)	Damping	Frequency (rad/s)	Damping
$1.25 \cdot 10^{-3}$	0.25	$10^{-3}$	1
$1.25 \cdot 10^{-3}$	0.25	$2.02 \cdot 10^{-2}$	0.03
$2.24 \cdot 10^{-2}$	0.03	$2.02 \cdot 10^{-2}$	0.03
$2.24 \cdot 10^{-2}$	0.03	$1.68 \cdot 10^{-1}$	0.098
$1.88 \cdot 10^{-1}$	0.01	$1.68 \cdot 10^{-1}$	0.098
$1.88 \cdot 10^{-1}$	0.01	$5.42 \cdot 10^{-1}$	-0.431
$8.32 \cdot 10^{-1}$	0.04	$5.42 \cdot 10^{-1}$	-0.431
$8.32 \cdot 10^{-1}$	0.04	$6.51 \cdot 10^{-1}$	0.23
$8.66 \cdot 10^{-1}$	0.86	$6.51 \cdot 10^{-1}$	0.23
$8.66 \cdot 10^{-1}$	0.86	$6.54 \cdot 10^{-1}$	0.61
1.26	0.5	$6.54 \cdot 10^{-1}$	0.61
1.26	0.5		

Table 1

Poles and zeros of the true system

This system is controlled with a three order R-S controller such that  $\frac{R(z)}{S(z)} = k_c \frac{\prod_{l=1}^3 (z - z_c(l))}{\prod_{l=1}^3 (z - p_c(l))}$  where  $z_c(l), p_c(l)$ ,  $l = \{1, 2, 3\}$  are the controller zeros and poles given in table 2, and one has  $k_c = 10^{-3}$ . The controller includes a double integrator, and for this reason the classical CLOE or X-CLOE schemes cannot be employed, since their regressors diverge. A white noise is added to the system output (signal/noise -variance-ratio of 10 dB). The system input is excited with a 17 registers PRBS ( $2^{17} - 1$  samples). Four simulations are carried out with different system orders and various basis poles configurations. In all these simulations, as recommended for the identification of stationary systems in [10], p. 69, the forgetting factor  $\lambda_1$  has an initial value strictly inferior to 1 and tends exponentially towards 1 during the simulation: That leads to an increased convergence rate of the algorithm.

Poles		Zeros	
Frequency (rad/s)	Damping	Frequency (rad/s)	Damping
0	1	$1.0 \cdot 10^{-1}$	1
0	1	$1.0 \cdot 10^{-1}$	1
$1.0 \cdot 10^{-2}$	1	3.14	0

Table 2

Poles and zeros of the controller

**Simulation 1:** The overall system is identified by choosing a 12th order model ( $n_a = 12$ ), and the basis poles are set in order that the function  $\chi$  is roughly constant over almost three decades. Their values are 0.5, 0.95, 0.995. Figure 4 provides a comparison of the true system versus the estimated one on a Bode diagram (magnitude and phase) and the bottom graph displays the function  $\chi$ . We notice immediately that the estimated model coincides with the true system.

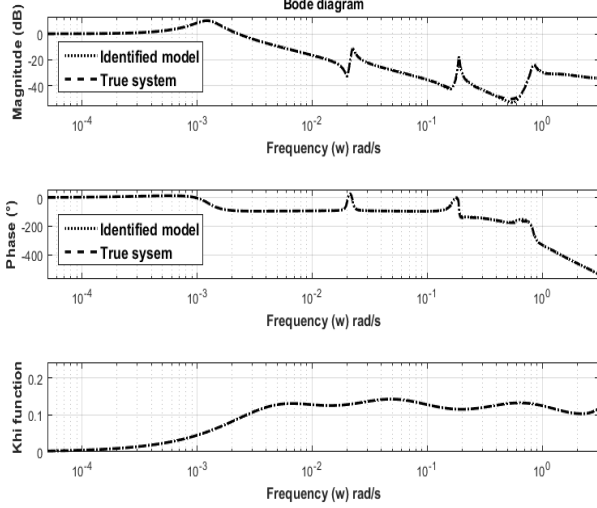


Fig. 4. Identification with a 12th order model and three basis poles equal to 0.5, 0.95, 0.995, respectively.

**Simulation 2:** A 3rd order model is chosen, with a single basis poles equal to  $p = 0.997$ , such that the frequency for which  $\chi$  is maximum is roughly equal to 0.003 rad/s. Figure 5 shows that the first modes of the system are perfectly captured, while the high frequency modes are sheerly ignored.

**Simulation 3:** A single basis pole is employed, equal to 0.5. The function  $\chi$  has its largest value within one decade below the Nyquist frequency. The model order is equal to 6, and Figure 6 shows clearly that the system modes situated in the decade below this Nyquist frequency are the only ones that are properly identified.

**Simulation 4:** The purpose of this identification is now to obtain a reduced order model valid locally around 0.02 rad/s. For this purpose we choose a function  $\chi$  having a narrow peak around this frequency. This is obtained by selecting a basis including 3 poles  $0.98, 0.9968 + 0.01971i, 0.9968 - 0.01971i$  respectively. The use of complex poles at a frequency equal to 0.02 rad/s and a relatively small damping (0.15) allows for this narrow peak. The model order is equal to 3. Owing to the function  $\chi$  aspect, we can expect a good fit around 0.02 rad/s: This is exactly what can be observed in Figure 7.

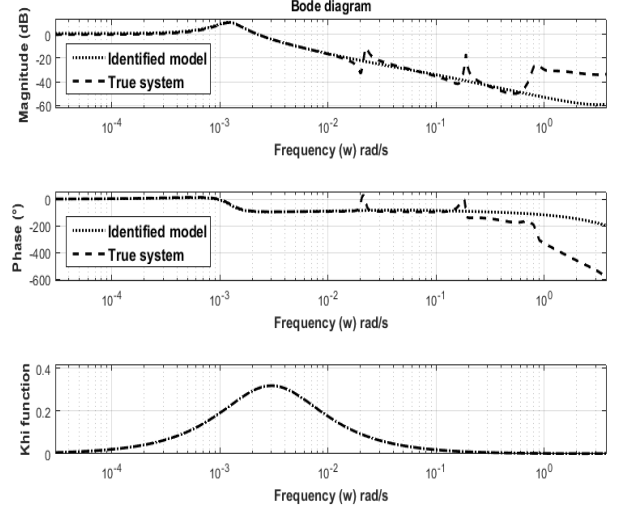


Fig. 5. Identification with a 3th order model and a single basis poles equal to 0.997.

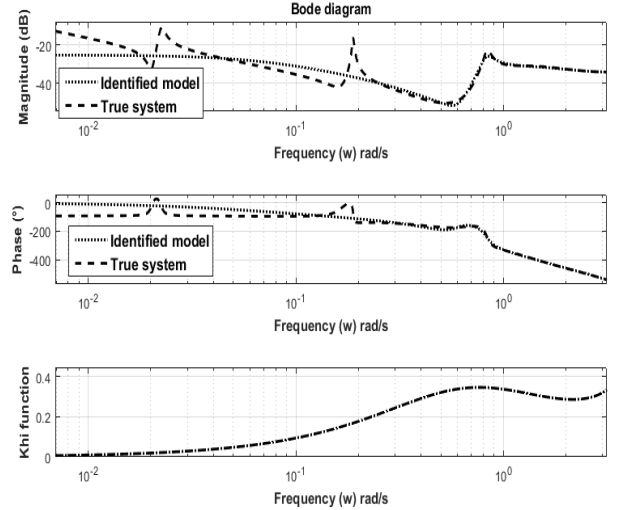


Fig. 6. Identification with a 6th order model and one basis pole equal to 0.5.

These simulations show the tuning effect of the basis poles on the bias distribution: by the use of a predictor established on a generalized basis of orthogonal transfer functions, the bias distribution can be managed, which is not the case when using a direct identification procedure in a closed-loop context, as mentioned in [12], p. 436. Moreover, these results confirm that H-XCLOE can cope with an unstable controller, contrary to the classical X-CLOE scheme.

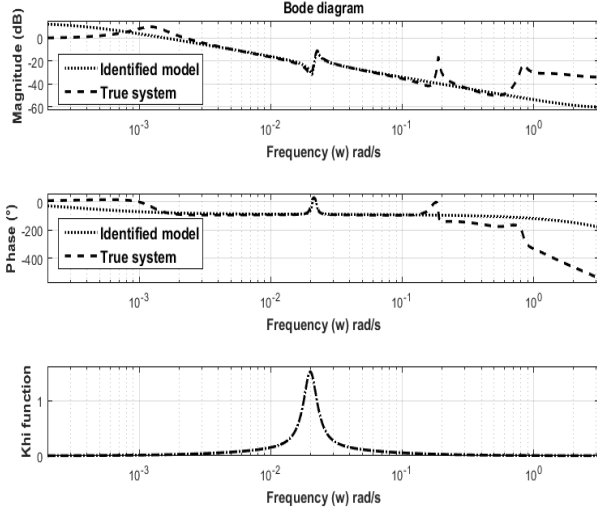


Fig. 7. Identification with a 3th order model and three basis poles equal to  $0.9968+0.01971i$ ,  $0.9968-0.01971i$  and  $0.98$ .

## 8 Acknowledgement

The authors express their gratitude to two anonymous reviewers for their valuable advice that helped improving the manuscript substantially.

## 9 Appendix

### 9.1 Appendix A: Some algebra to solve the Bézout equation using the Hambo transform

Owing to the expressions of  $A_p$  in (19), and  $\bar{S}$  as defined in Section 3, there exists a unique solution to the Bézout equation (23), where  $d^o(A_f) = n_a$ , and  $d^o(B_f) \leq n_a - 1$ . This is the solution considered in the sequel. In case of fast sampling or a high model/controller order, the resolution of equation (23) may become difficult due to an ill-conditioning of the associated Sylvester system. This is the reason why it is strongly recommended to solve it with the Hambo operator. At first, consider the commutative ring  $\mathbf{R}$  of polynomials in  $z$  with real coefficients. The set of polynomials  $\mathcal{S}_a = \{A_p^k : k \in \mathbb{N}\}$ , is a multiplicative subset of  $\mathbf{R}$ , and the set of fractions  $\mathbf{R}\mathcal{S}_a^{-1} = \{\frac{L}{A_p^k} : L \in \mathbf{R}, A_p^k \in \mathcal{S}_a\}$  is, like  $\mathbf{R}$ , a principal ideal domain. Therefore (23) can be written

$$\frac{A_f}{A_p} \frac{S}{A_p''} + \frac{B_f}{A_p} \frac{R}{A_p''} = \frac{\bar{S}}{A_p''} \quad (59)$$

where  $A_p'' = (A_p)^k$  is such that  $d^o(A_p'') \geq d^o(S)$ , and (59) is again a Bézout equation over  $\mathbf{R}\mathcal{S}_a^{-1}$ . Consider  $\tilde{A}_f(\lambda, \xi)$ ,  $\tilde{S}(\lambda)$ ,  $\tilde{B}_f(\lambda, \kappa)$ ,  $\tilde{R}(\lambda)$ ,  $\tilde{\bar{S}}(\lambda)$ , the Hambo operator transforms of respectively  $\frac{A_f}{A_p}$ ,  $\frac{S}{A_p''}$ ,  $\frac{B_f}{A_p}$ ,  $\frac{R}{A_p''}$ ,  $\frac{\bar{S}}{A_p''}$ , and

let  $\xi$  and  $\kappa$  the coefficients vectors of  $\tilde{A}_f$  and  $\tilde{B}_f$  respectively. Define  $\widetilde{\mathbf{R}\mathcal{S}_a^{-1}}$  to be the set of those  $\tilde{H}$  which are the image of a proper and stable transfer function  $H$  of  $\mathbf{R}\mathcal{S}_a^{-1}$ , by the Hambo transform  $\mathcal{H}$ .

**Proposition 8** *The mapping  $\mathcal{H} : H \mapsto \tilde{H}$*

$$\mathbf{R}\mathcal{S}_a^{-1} \xrightarrow{\mathcal{H}} \widetilde{\mathbf{R}\mathcal{S}_a^{-1}}$$

*is a ring isomorphism, and its codomain is a ring of polynomial matrices.*

*Proof:*

*At first the mapping  $\mathcal{H}$  is a homomorphism: For any transfer function  $H_1, H_2$  one has  $\widetilde{H_1 + H_2} = \tilde{H}_1 + \tilde{H}_2$  (from (12)), and  $\widetilde{H_1 H_2} = \tilde{H}_1 \tilde{H}_2$  (from (11)). Moreover, by its very definition  $\mathcal{H}$  is surjective. Additionally we show in the sequel that its kernel is 0: The relation  $\tilde{H}(\lambda) = 0$  implies necessarily*

$$\sum_{k=1}^{n_p} \frac{H(z_k) V_1(z_k) V_1^T(1/z_k)}{V_1^T(z_k) V_1(1/z_k)} = 0$$

*for an infinite number of values of  $z_k$ . Since  $H$  has a finite number of zeros if it is nonzero,  $\tilde{H} = 0$  implies  $H = 0$ . Thus  $\mathcal{H}$  is injective, and therefore this homomorphism is bijective, its codomain is a ring of polynomial matrices, owing to remarks of section 12.5.1 in [8].*

Therefore (59) is equivalent to

$$\tilde{A}_f(\lambda, \xi) \tilde{S}(\lambda) + \tilde{B}_f(\lambda, \eta) \tilde{R}(\lambda) = \tilde{\bar{S}}(\lambda) \quad (60)$$

If  $\frac{S}{A_p''}$  and  $\frac{R}{A_p''}$ , are coprime in  $\mathbf{R}\mathcal{S}_a^{-1}$ , i.e. if the only common divisors of  $R$  and  $S$ , if any, belong to  $\mathcal{S}_a$ , (59) has a solution, and (60) has a solution too,  $\xi$  and  $\eta$  being the unknown terms.

### 9.2 Appendix B: Proof of Proposition 7

Let  $H_{\varepsilon_E}(e^{i\omega})$  be the stable spectral factor of  $\Phi_{\varepsilon_E}(\omega)$ . Its Hambo signal transform, denoted as  $\check{H}_{\varepsilon_E}(e^{i\omega_\lambda})$ , is  $\check{H}_{\varepsilon_E}(e^{i\omega_\lambda}) = H_{\varepsilon_E}(e^{i\omega}) \frac{V_1(e^{-i\omega}) V_1^T(e^{i\omega})}{V_1^T(e^{i\omega}) V_1(e^{-i\omega})} \Big|_{\omega=\beta^{-1}(\omega_\lambda)} \frac{e^{-i\omega_\lambda} C_b^T}{e^{-i\omega_\lambda} - D_b}$  (owing to (10)). From equation (3.10) of [8] (chap.3), one gets  $\frac{C_b^T C_b}{1 + D_b^2 - 2\cos(\omega_\lambda)} V_1(e^{-i\omega}) V_1^T(e^{i\omega}) \Big|_{\omega=\beta^{-1}(\omega_\lambda)} = I_{n_p}$ . Thus  $\check{H}_{\varepsilon_E}^T(e^{i\omega_\lambda}) \check{H}_{\varepsilon_E}(e^{-i\omega_\lambda}) = \Phi_{\varepsilon_E}(\omega) \frac{1}{\beta'(\omega)} \Big|_{\omega=\beta^{-1}(\omega_\lambda)} = \Psi_{\varepsilon_E}(\omega_\lambda)$ . Similarly,  $\Psi_r(\omega_\lambda) = \check{H}_r^T(e^{i\omega_\lambda}) \check{H}_r(e^{-i\omega_\lambda})$ , and  $\Psi_{\varepsilon_E}(\omega_\lambda) = \check{H}_e^T(e^{i\omega_\lambda}) \check{H}_e(e^{-i\omega_\lambda})$ , where  $\check{H}_r(e^{i\omega_\lambda})$  and  $\check{H}_e(e^{i\omega_\lambda})$  are the Hambo signal transforms of the stable spectral factors of  $\Phi_r(\omega)$  and  $\Phi_e(\omega)$  respectively. Now one has  $\tilde{A}(e^{i\omega_\lambda}) = \frac{A(e^{i\omega})}{A_p(e^{i\omega})} \frac{V_1(e^{-i\omega}) V_1^T(e^{i\omega})}{V_1^T(e^{i\omega}) V_1(e^{-i\omega})} \Big|_{\omega=\beta^{-1}(\omega_\lambda)}$ .

Therefore  $\left| \frac{A(e^{i\omega})}{A_p(e^{i\omega})} \right|_{\omega=\beta^{-1}(\omega_\lambda)}^2 = \|\tilde{A}(e^{i\omega_\lambda})\|_2^2$ , and similar relations hold for  $|G_0(e^{i\omega}) - G(e^{i\omega})|^2$ ,  $|T_{ur0}(e^{i\omega})|^2$ ,  $|\frac{\tilde{S}(e^{i\omega})}{\tilde{S}(e^{i\omega})}|^2$ ,  $|W_0(e^{i\omega})|^2$ ,  $|W(e^{i\omega})|^2$ ,  $|\frac{S_{yp0}(e^{i\omega})}{S_{yp}(e^{i\omega})}|^2$ . Thus, by using equations of the limit model in the  $\omega$  frequency domain (48) and (49), and by considering that from [19] (section 5)  $\mathbf{E}[\varepsilon_E^2(t)] = \frac{1}{2n_p\pi} \int_{-n_p\pi}^{+n_p\pi} \Psi_{\varepsilon_E}(\omega_\lambda) d\omega_\lambda$ , one obtains the claimed results.

### 9.3 Appendix C: Proof of Theorem 1

One has

$$\chi_k(\omega) = \frac{1}{\pi} \frac{1 - |p_k|^2}{|1 - \bar{p}_k e^{i\omega}|^2} \omega = \frac{1}{\pi} \frac{(1 - \rho_k^2)}{1 + \rho_k^2 - 2\rho_k \cos(\omega - \sigma_k)} \omega$$

$$\frac{\partial \chi_k(\omega)}{\partial \omega} = \frac{1}{\pi} \frac{(1 - \rho_k^2)}{2\rho_k} \frac{\frac{1+\rho_k^2}{2\rho_k} - \cos(\omega - \sigma_k) - \omega \sin(\omega - \sigma_k)}{\left( \frac{1+\rho_k^2}{2\rho_k} - \cos(\omega - \sigma_k) \right)^2}$$

The sign of  $\frac{\partial \chi_k(\omega)}{\partial \omega}$  depends upon the sign of

$g(\omega) = \frac{1+\rho_k^2}{2\rho_k} - \cos(\omega - \sigma_k) - \omega \sin(\omega - \sigma_k)$ . One has  $\frac{\partial g(\omega)}{\partial \omega} = -\omega \cos(\omega - \sigma_k) = -\omega \cos(\omega - \sqrt{1 - \xi_k^2} \omega_{ok})$ ,  $\frac{\partial g(\omega)}{\partial \omega} \leq 0$  if and only if  $\omega \leq \frac{\pi}{2} + \sqrt{1 - \xi_k^2} \omega_{ok}$ , and  $\frac{\partial g(\omega)}{\partial \omega} > 0$  otherwise. Let  $\tilde{\omega}$  be the frequency for which  $g(\omega)$  is minimum. One has  $g(\tilde{\omega}) = \cos(\xi_k \omega_{ok}) - \sqrt{1 - \xi_k^2} \omega_{ok} - \frac{\pi}{2}$ . This quantity is strictly negative if and only if

$$\cosh(\xi_k \omega_{ok}) - \sqrt{1 - \xi_k^2} \omega_{ok} < \frac{\pi}{2} \quad (61)$$

Additionally  $g(0) = \cosh(\xi_k \omega_{ok}) - \cos(\sqrt{1 - \xi_k^2} \omega_{ok}) > 0$  for any  $\omega_{ok} > 0$ ,  $g(\omega)$  is a positive decreasing function for  $\omega$  close to 0, and has a minimum at  $\omega = \frac{\pi}{2} + \sqrt{1 - \xi_k^2} \omega_{ok}$  only if  $\omega_{ok} < \frac{\pi}{2\sqrt{1 - \xi_k^2}}$ , i.e

$$\xi_k^2 \geq 1 - \frac{\pi^2}{4\omega_{ok}^2} \quad (62)$$

Now  $g(\pi) = \cosh(\xi_k \omega_{ok}) + \cos(\sigma_k) - \pi \sin(\sigma_k)$ , therefore  $g(\pi) > 0$  is equivalent to

$$\cosh(\xi_k \omega_{ok}) + \cos(\sqrt{1 - \xi_k^2} \omega_{ok}) - \pi \sin(\sqrt{1 - \xi_k^2} \omega_{ok}) > 0 \quad (63)$$

Therefore, if we assume that (62) and (61) are satisfied,  $\chi_k$  has a unique maximum on  $[0, \pi[$ . Furthermore if (63) is satisfied  $\chi_k$  has a local minimum on  $]0, \pi[$ . If condition (62) is satisfied and (61) is not,  $\chi_k$  is an increasing function on this interval and admits a unique maximum at  $\omega = \pi$ . If the pole  $p_k$  is real, condition (63) is necessarily fulfilled, and (61) reduces to:  $1 + p_k^2 - \pi p_k < 0$ . Since we consider only stable poles, this is equivalent to  $p_k > \frac{\pi - \sqrt{\pi^2 - 4}}{2}$ .

### 9.4 Appendix D: Proof of Corollary 1

According to Theorem 1,  $\omega_{max}$  is the smallest frequency such that  $g(\omega) = 0$ . This frequency is such that

$$h(\omega_{ok}, \omega) = \cosh(\xi_k \omega_{ok}) - \cos\left(\omega - \sqrt{1 - \xi_k^2} \omega_{ok}\right) \cdots$$

$$\cdots - \omega \cos\left(\omega - \sqrt{1 - \xi_k^2} \omega_{ok}\right) = 0.$$

Let us consider  $\omega_{ok}$  as the function variable and  $\omega$  as a parameter. One has  $h(\omega_{ok}, \omega) = -\left(\cos(\omega) \cos\left(\sqrt{1 - \xi_k^2} \omega_{ok}\right) + \sin(\omega) \sin\left(\sqrt{1 - \xi_k^2} \omega_{ok}\right)\right) \cdots$

$$- \omega \left(\sin(\omega) \cos\left(\sqrt{1 - \xi_k^2} \omega_{ok}\right) - \cos(\omega) \sin\left(\sqrt{1 - \xi_k^2} \omega_{ok}\right)\right) \cdots$$

$$+ \cosh(\xi_k \omega_{ok}).$$

A first order Taylor-Young approximation yields

$$h(\omega_{ok}, \omega) = 1 - \cos(\omega) + \sin(\omega) \sqrt{1 - \xi_k^2} \omega_{ok} - \omega \sin(\omega) + \omega \cos(\omega) \sqrt{1 - \xi_k^2} \omega_{ok} + o(\omega_{ok}).$$

This quantity can be null only if  $1 - \cos(\omega) - \omega \sin(\omega) = 0$ , implying  $\omega = 0$ . If we perform a second order Taylor-Young expansion near 0, we get

$$h(\omega_{ok}, \omega) = 1 + \frac{\xi_k^2 \omega_{ok}^2}{2} - \left(1 - \frac{\omega^2}{2}\right) \left(1 - \frac{(1 - \xi_k^2) \omega_{ok}^2}{2}\right) -$$

$$\omega \sqrt{1 - \xi_k^2} \omega_{ok} - \omega \left(\omega - \sqrt{1 - \xi_k^2} \omega_{ok}\right) + o(\|(\omega_{ok}, \omega)\|^2)$$

$$= 1 + \frac{\xi_k^2 \omega_{ok}^2}{2} - 1 + \frac{(1 - \xi_k^2) \omega_{ok}^2}{2} + \frac{\omega^2}{2} - \omega \sqrt{1 - \xi_k^2} \omega_{ok} - \omega^2 +$$

$$\omega \sqrt{1 - \xi_k^2} \omega_{ok} + o(\|(\omega_{ok}, \Omega(\omega_{ok}))\|^2)$$

$$= \frac{1}{2} (\omega^2 - \omega_{ok}^2) + o(\|(\omega_{ok}, \Omega(\omega_{ok}))\|^2).$$

Consequently,  $h(\omega_{ok}, \omega) = o(\|(\omega_{ok}, \omega)\|^2)$  if and only if  $\omega^2 = \omega_{ok}^2$ . The relation  $h(\omega_{ok}, \omega) = 0$  entails an implicit function  $\omega_{max} = \Omega(\omega_{ok})$ , and one has:  $h(\omega_{ok}, \Omega(\omega_{ok})) = o(\|(\omega_{ok}, \Omega(\omega_{ok}))\|^2)$ . Hence the result.

### References

- [1] P. Bodin, L.F. Villermoes, B. Wahlberg, "An algorithm for pole selection of best orthonormal basis", Proc. of the 36th IEEE Conference on Decision and Control, San Diego, 1997.
- [2] H. Boursès, *Linear Systems*, ISTE Wiley, 2010.
- [3] H. Boursès, Y. Joannic, O. Mercier, "Rho-stability and robustness: discrete-time case", International Journal of Control, 52(5), pp. 1217-1239, 1990.
- [4] F. de Bruyne, B. Anderson, N. Linard, "The Hansen scheme revisited", Proc. of the 37th IEEE Conference on Decision and Control, Tampa, Florida, 1998.
- [5] U. Forsell, I. Ljung, "Closed-loop identification revisited", Automatica, 35(7), pp. 1215-1241, 1999.
- [6] M. Gevers, "Towards a joint design of identification and control?", In H.L. Trentelma, and J.C. Willems (Eds), *Essays on control: Perspectives in the theory and its applications*, Birhäuser, Boston, pp. 111-151, 1993.
- [7] P.S.C. Heuberger, P.M.J. Van den Hof, O.H. Bosgra, "A generalized orthonormal basis for linear dynamical systems", IEEE, Trans. on Automatic Control, vol. 40, pp. 451-465, 1995.
- [8] P.S.C. Heuberger, P.M.J. Van den Hof, B. Wahlberg, *Modelling and Identification with Rational Orthogonal Basis Functions*, Springer Verlag, 2005.
- [9] P.S.C. Heuberger, T.J. de Hoog, P.M.J. Van den Hof, B. Wahlberg, "Orthonormal basis functions in time and

- frequency domain: Hambo transform theory", SIAM, J. Control and Opt., vol. 42(4), pp. 1347-1373, 2003.
- [10] I.D. Landau, R. Lozano, M. M'Saad, A. Karimi, *Adaptive Control, second edition*, Springer Verlag, 2011.
- [11] Ph. de Larminat, *Automatique Appliquée*, Hermès Sciences, 2009.
- [12] L. Ljung, *System Identification, Theory for the User, second edition*, Upper Saddle River, Prentice Hall, 1999.
- [13] F. Shipp, L. Gianone, J. Bokor, Z. Szabo, "Identification in generalized orthogonal basis - a frequency domain approach", preprints of the 13th IFAC World congress, vol. 1, pp. 387-392, Elsevier, 1996.
- [14] T. Oliveira e Silva, "Laguerre filters - An introduction", Revista do Detua, vol.1(3), pp. 237-248, 1995.
- [15] T. Oliveira e Silva "On the determination of the optimal pole position of Laguerre filters", IEEE Trans. on Signal Processing, vol. 43(9), pp. 2079-2087, 1995.
- [16] T. Söderström, P. Stoïca, *System Identification*, Upper Saddle River, Prentice Hall, 1989.
- [17] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, Springer Verlag, 2010.
- [18] P.M.J. Van den Hof, R.J.P. Schrama, "Identification and control - Closed-loop issues", Automatica, vol. 31, pp. 1571-1770, 1995.
- [19] P.M.J Van den Hof, P.S.C. Heuberger, J. Bokor, "System Identification with generalized orthonormal basis functions", Automatica, vol. 31(12), pp. 1821-1834, 1995.
- [20] B. Vau, H. Bourlès, "Some remarks on the bias distribution analysis of pseudo-linear regression discrete-time identification algorithms", Systems and Control Letters, vol. 119, pp. 46-51, 2018.