



# Comments on the bias distribution of PLR algorithms

Bernard Vau, Henri Bourlès

## ► To cite this version:

Bernard Vau, Henri Bourlès. Comments on the bias distribution of PLR algorithms. 2022. hal-03788883

**HAL Id: hal-03788883**

**<https://hal.science/hal-03788883>**

Preprint submitted on 27 Sep 2022

**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.

# Comments on the bias distribution of PLR algorithms

Bernard Vau\*\* Henri Bourlès\*

\* SATIE, Ecole normale supérieure de Paris-Saclay, 91130 Gif sur Yvette, France (e-mail: henri.bourles@satie.ens-cachan.fr)

\*\* IXBLUE, 94380 Bonneuil-sur-Marne, France (e-mail: bernard.vau@ixblue.com)

## 1. INTRODUCTION

## 3. ANALYSIS

The purpose of this technical note is to discuss some aspects of the bias distribution in Pseudo-Linear Regression (PLR) identification algorithms. PLR algorithms have been widely used in adaptive control (3), and a recent contribution make them able to identify multi-frequency scale systems in discrete-time (7). In (6) the concept of equivalent prediction error was introduced: it is a non-measurable signal whose variance is asymptotically minimized by some PLR schemes (output error and ARMAX models, either in open- or closed-loop operations), and the bias distribution of these algorithms is inferred from the expressions of this equivalent prediction error (whose expression, specific to each algorithm is given in Table 2 of (6)). In Lemma 2 of (6), it is claimed that the stationarity condition of the PLR schemes (namely the orthogonality between the prediction error and the regressor) entails the orthogonality between the equivalent prediction error and its gradient (independent of the estimated parameters) called the equivalent regressor.

In a personal communication to the authors, Pr. Xavier Bombois pointed out that this assertion is correct if the true system is in the model set, but that Lemma 2 of (6) does not prove that it is still exact if the true systems is no longer in the model set (which is the situation when a bias occurs). This question is worthwhile since the bias analysis proposed in (6) depends on this property. This paper sheds light on this matter, and it is proven under some assumptions that the effective stationarity condition is very close to the one claimed in (6).

## 2. NOTATION

In the sequel, the following notation related to signal, transfer functions and transfer functions operators is used:  $q$  the forward shift operator,

$z$  the symbol of the  $z$ -transform

$G : z \mapsto G(z)$  a system, for example  $G(z) = \frac{b_1 z + b_2}{z^2 + a_1 z + a_2}$

$G^* : z \mapsto G^*(z)$  the reciprocal system of  $G$ , for example  $G^*(z) = \frac{b_1 + b_2 z}{1 + a_1 z + a_2 z^2}$

$\check{G} : z \mapsto \check{G}(z^{-1})$  for example  $\check{G}(z) = \frac{b_1 z^{-1} + b_2}{z^{-2} + a_1 z^{-1} + a_2}$  and

$\check{G}^*(z^{-1}) = \frac{b_1 + b_2 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}}$

$\mathbf{1} : z \mapsto 1$ .

In this paper all the signal are assumed to be quasi-stationary signals, and

$\mathbf{E}[x(t)] = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{\tau=0}^N \mathbf{E}[\tau]$

Three assumptions are made

- (1) **H1** The excitation sequences are persistently exciting signals, meaning that  $\Phi_u(z) > 0 \quad \forall z$  and  $\Phi_{r_u}(z) > 0 \quad \forall |z| = 1$ , where  $\Phi_u(z)$  and  $\Phi_{r_u}(z)$  are the Power Spectrum Density (PSD) of respectively  $\{u(t)\}$  (the input excitation signal in open-loop) and  $\{r_u(t)\}$  (the additive excitation signal on the system input in closed-loop).
- (2) **H2** One defines  $n_a = \deg(A(q))$ ,  $n_b = \deg(B(q))$ ,  $n_c = \deg(C(q))$  the degrees of  $A, B, C$  respectively, and for sake of simplicity (but without any restriction) one chooses in this article  $n_b = n_a - 1$ , and  $n_c = n_a$ .
- (3) **H3** The controller  $K(q)$  is stable ( $S(q)$  has all its zeros strictly inside de unit circle)

In the sequel  $\mathcal{H}_2$  is the set of square summable transfer functions analytic outside the unit circle including infinity, and  $\mathcal{H}_{2-}$  is the subset of  $\mathcal{H}_2$  including only the functions having a relative degree higher or equal to 1. Moreover, the inner product of two transfer functions  $H_1$  and  $H_2$ , denoted  $\langle H_1, H_2 \rangle$  is equal to

$$\frac{1}{2\pi i} \oint_{\mathbf{T}} H_1(z^{-1}) H_2(z) \frac{dz}{z}$$

where  $\mathbf{T}$  is the unit circle, and one has  $\langle H_1, H_2 \rangle = \langle \mathbf{1}, \check{H}_1 H_2 \rangle$ .

Since the publication of (6), it has been proposed more recently in (7) some improved versions of PLR algorithms where the predictor is parametrized on generalized orthonormal bases transfer function (GOBF) in closed-loop, the schemes being called H-CLOE and H-XCLOE, and as mentioned in this latter article the corresponding algorithms in an open-loop context can directly be inferred by taking a controller null (where  $R^*(q^{-1}) = 0, S^*(q^{-1}) = 1$ ). The corresponding output-error algorithm has been detailed in (8) and is called H-OLOE, and a similar simplification of H-XCLOE to the open-loop case (for ARMAX models) will be presented in a further publication.

It is thus necessary to recall (very quickly) the properties of GOBF, the reader interested in this topic can report to (2). These functions stem from a balanced realization of an

all-pass function  $G_b(z) = \prod_{j=0}^{n_p-1} \frac{-z \bar{p}_j + 1}{z - p_j}$  where  $p_j$  ( $|p_j| < 1$ ) are the basis poles and  $n_p$  the basis poles number. There exists a balanced state space realization such that  $G_b(z) = D_b + C_b(zI - A_b)^{-1}B_b$  (for the construction of the state space matrices see chap. 2 of (2)). The orthonormal transfer function vectors  $V_k(z)$  with  $k = 1, 2, \dots$  and of size  $(n_p, 1)$  are given by the relation  $V_k(z) = (zI - A_b)^{-1}B_b G_b^{k-1}(z)$ . Because of the orthonormal state space realisation of  $G_b(z)$ , orthonormality between these functions holds. Particular configurations of  $n_p$  and  $p_k$  correspond to well known cases:  $n_p = 1, p_0 = 0$  is the classical  $z^{-1}, z^{-2}, \dots$  basis, and  $n_p = 1, |p_0| < 1$  correspond to the Laguerre basis. The following result related to GOBF is employed in the sequel:

*Theorem 1.* Consider any function  $F : z \mapsto F(z)$ . One has

$$\langle V_k, F \rangle = 0 \iff F(z) = G_b^k(z)L(z)$$

where  $L \in \mathbf{H}_{2-}$ .

Proof:

1) Sufficient part : Let us assume  $F(z) = G_b^k(z)L(z)$  with  $L \in \mathbf{H}_{2-}$ . For any integer  $k \in [1, \infty[$ , one can write  $\langle V_k, F \rangle = \langle V_k G_b^{-k}, L \rangle = \langle V_1 G_b^{-1}, L \rangle$ . But according to Corollary 3.6 of (1),  $V_1(z)G_b^{-1}(z)$  belongs to  $\mathbf{H}_{2-}^\perp$  the orthogonal complement of  $\mathbf{H}_{2-}$ . Therefore  $\langle V_1 G_b^{-1}, L \rangle = \langle V_k, F \rangle = 0$ .

2) Necessary part: Let us assume that  $\langle V_k, F \rangle = \langle V_1 G_b^{k-1}, F \rangle = 0$ . Let  $l$  be an integer such that  $l \in [k + 1, \infty[$ , one can write  $\langle V_1 G_b^{k-1} G_b^{-l+1}, G_b^{-l+1} F \rangle = 0$ , and let us define  $r = k - l$ . One has  $r \in ]-\infty, -1]$ , and  $\langle V_1 G_b^r, G_b^{-l+1} F \rangle = 0$ . But according to the corollary 3.6 of (1),  $\{V_1(z)G_b^r(z)\}$  is a basis for  $\mathbf{H}_{2-}^\perp$ . Therefore a necessary condition so that this last equality holds is that  $G^{-l+1}F \in \mathbf{H}_{2-}$  for  $l = k + 1$  (since  $l \in [k + 1, \infty[$ ), or equivalently  $G_b^{-k}F \in \mathbf{H}_{2-}$  which means that  $G_b^{-k}F$  must be stable and which implies that  $F = G_b^k L$  where  $L \in \mathbf{H}_{2-}$ . This ends the proof.  $\square$

For sake of homogeneity with a majority of papers published in the identification field, some notation differ from the notation employed in (6) as in (4) :  $\theta_0$  is the true parameter vector of size  $n_f$  (previously denoted as  $\theta$  in (6)),  $\theta$  is the parameter vector of the model of size  $n \leq n_f$  (denoted as  $\hat{\theta}$  in (6)). The estimated parameter vector at time  $t$  is denoted  $\hat{\theta}(t)$ , and  $\theta^* = \lim_{t \rightarrow +\infty} \hat{\theta}(t)$ .

The transfer function operators  $G_0(q) = \frac{B_0(q)}{A_0(q)}$  and  $W_0(q) = \frac{C_0(q)}{A_0(q)}$  are respectively the deterministic and stochastic parts of the true system, and  $G(q) = \frac{B(q)}{A(q)}$  and  $W(q) = \frac{C(q)}{A(q)}$  the deterministic and stochastic parts of the identified model. The excitation signal sequence is  $\{u(t)\}$  in open-loop and  $\{r_u(t)\}$  in closed-loop with  $u(t) = -K(q)y(t) + r_u(t)$ , and where  $K(q) = \frac{R(q)}{S(q)}$  is the R-S controller. Three classes of models are addressed in (6) and the present document.

- The open-loop output error model where  $y(t) = G_0(q)u(t) + v(t)$ , the sequence  $\{v(t)\}$  being a centred,

gaussian (non necessarily white) stochastic disturbance, independent of  $\{u(t)\}$ , the input sequence of the model.

- The closed-loop output error model where  $y(t) = S_{yv0}(q)r_u(t) + S_{yp0}(q)v(t)$ ,  $S_{yp0}(q)$  being the closed-loop sensitivity function  $S_{yp0}(q) = \frac{1}{1+G_0(q)K(q)}$ , and  $S_{yv0}(q) = \frac{G_0(q)}{1+G_0(q)K(q)}$
- The ARMAX model where  $y(t) = G_0(q)u(t) + W_0(q)e(t)$ , where  $\{e(t)\}$  is centred, gaussian white sequence.

For PLR algorithms, according to (6) the predicted output  $\hat{y}(t+1, \theta)$  is given by  $\hat{y}(t+1, \theta) = \theta^T \phi(t, \theta)$ , where  $\phi(t, \theta)$ , is the regressor of the predictor, and  $\varepsilon(t+1, \theta) = y(t+1) - \hat{y}(t+1, \theta)$  is the prediction error, where  $y(t+1)$  is the true system output at time  $t+1$ .

In Theorem 1 of (6) the transfer function operator  $Q$  has been introduced. This transfer function is by definition the one that satisfies the relation

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

and it has been shown that it satisfies <sup>1</sup>

$$Q^*(q^{-1}, \theta) = 1 + \theta^T \frac{\partial \phi}{\partial q \varepsilon} \quad (2)$$

Note that in any case  $Q^*(q^{-1})$  is causal and is the ratio of two monic polynomials in  $q^{-1}$ . The expressions of  $Q^*(q^{-1})$  (for classical PLR algorithms) are given in Table 2 of (6) (denoted  $Q(q^{-1})$  in this last reference).

The equivalent prediction error  $\varepsilon_E$  was defined in (6)

$$\varepsilon_E(t+1, \theta) = Q^*(q^{-1}, \theta) \varepsilon(t+1, \theta) + (1 - Q^*(q^{-1}, \theta)) w(t+1) \quad (3)$$

where  $w(t+1) = v(t+1)$  for the open-loop output error model,  $w(t+1) = e(t+1)$  for the open and closed-loop ARMAX model, and  $w(t+1) = S_{yp}(q)v(t+1)$  for the closed-loop output error model. From the first part of Lemma 2 in (6) one has

$$\varepsilon_E(t+1, \theta) = (\theta_0 - \theta)^T \phi_E(t) + w(t+1) \quad (4)$$

where  $\phi_E$  is by definition the equivalent regressor and is independent of  $\theta$ , and one has

$$\phi_E(t) = -\frac{\partial \varepsilon_E(t+1, \theta)}{\partial \theta} \quad (5)$$

If the true system is in the model set, and if  $\theta = \theta_0$ , one verifies that

$$\varepsilon(t+1, \theta_0) = \varepsilon_E(t+1, \theta_0) = w(t+1) \quad (6a)$$

$$\phi(t, \theta_0) = \phi_E(t) \quad (6b)$$

Therefore the stationarity condition

$$\bar{\mathbf{E}}[\varepsilon(t+1, \theta_0) \phi(t, \theta_0)] = 0 \quad (7)$$

entails the orthogonality between the equivalent prediction error and the equivalent regressor

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta_0) \phi_E(t)] = 0 \quad (8)$$

and expressions in the frequency domain can be given like in Table 3 of (6). However in this last reference it

<sup>1</sup> The use of partial derivative relative to signals is justified in the appendix of (6).

is not proven that these expressions remain exact if the true system is no longer in the model set, which occurs necessarily in the case of bias. The following result can now be presented

*Theorem 2.* Under the assumptions H1,H2 and H3, for the four PLR algorithms, namely

- a) Output error in open-loop
- b) Output error in closed-loop
- c) Extended least squares in open-loop
- d) Extended closed-loop output error

the stationnarity condition

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

entails

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta)\phi(t, \theta)] = 0$$

even if the true system is not in the model set.

Proof:

a) Output error in open-loop. For this algorithm,  $w(t+1) = v(t+1)$  and from (7), and (2),  $Q^*(q^{-1}) = \frac{A^*(q^{-1})}{A_p^*(q^{-1})}$ , where  $A_p(q) = \prod_{j=0}^{n_p-1} (q-p_j)^n$  ( $p_j$  being the basis poles of the predictor -which are null for the classical output error scheme of PLR and  $n = \frac{n_a}{n_p}$ ), the stationarity condition entails for  $1 \leq k \leq n$

$$\bar{\mathbf{E}}[\{V_k(q)\hat{y}(t+1, \theta)\}(\varepsilon(t+1, \theta) - w(t+1))] = 0 \quad (9a)$$

$$\bar{\mathbf{E}}[\{V_k(q)u(t+1, \theta)\}(\varepsilon(t+1, \theta) - w(t+1))] = 0 \quad (9b)$$

since

$$\bar{\mathbf{E}}[w(t+1)\hat{y}(t+1-l, \theta)] = 0 \quad (10a)$$

$$\bar{\mathbf{E}}[w(t+1)u(t+1-l, \theta)] = 0 \quad (10b)$$

for any  $l > 0$ .

But one has  $\varepsilon(t, \theta) = (G_0(q) - G(q))u(t) + v(t)$  and  $\hat{y}(t, \theta) = G(q)u(t)$ . Therefore from equations (9) one can write,

$$\langle V_k, \check{G}(G_0 - G)\Phi_u \rangle = 0 \quad (11a)$$

$$\langle V_k, (G_0 - G)\Phi_u \rangle = 0 \quad (11b)$$

And Theorem 1 leads to

$$\check{G}(G_0 - G)\Phi_u = G_b^k L_{\hat{y}} \quad (12a)$$

$$(G_0 - G)\Phi_u = G_b^k L_u \quad (12b)$$

where  $L_{\hat{y}}$  and  $L_u$  belong to  $\mathbf{H}_{2-}$ .

Note that conditions (12) entails that the set of zeros of  $(G_0 - G)\Phi_u$  include the zeros of  $G_b^n$  (the inverse of the basis poles with a multiplicity  $n$ ) and the inverse of the poles of  $G$  (so that  $\check{G}(G_0 - G)\Phi_u$  belongs to  $\mathbf{H}_{2-}$ ), and this is nothing but the result of Theorem 2 in (5).

Now, since  $Q^*$  is a ratio of two monic polynomials and is stable by construction, one has

$$\check{G}Q^*(G_0 - G)\Phi_u = G_b^k Q^* L_{\hat{y}} = G_b^k L'_{\hat{y}} \quad (13a)$$

$$Q^*(G_0 - G)\Phi_u = G_b^k Q^* L_u = G_b^k L'_u \quad (13b)$$

where  $L'_{\hat{y}}$  and  $L'_u$  also belong to  $\mathbf{H}_{2-}$ .

By applying again Theorem 1, one gets

$$\langle V_k, \check{G}Q^*(G_0 - G)\Phi_u \rangle = 0 \quad (14a)$$

$$\langle V_k, Q^*(G_0 - G)\Phi_u \rangle = 0 \quad (14b)$$

Now from the definition of the equivalent prediction error (3), one has

$$Q^*(q^{-1}, \theta) (\varepsilon(t+1, \theta) - w(t+1)) = \varepsilon_E(t+1, \theta) - w(t+1) \quad (15)$$

By combining (14) and (15) and shifting in the time domain one has

$$\bar{\mathbf{E}}[\{V_k(q)\hat{y}(t+1, \theta)\}(\varepsilon_E(t+1, \theta) - w(t+1))] = 0 \quad (16a)$$

$$\bar{\mathbf{E}}[\{V_k(q)u(t+1, \theta)\}(\varepsilon_E(t+1, \theta) - w(t+1))] = 0 \quad (16b)$$

but owing to (10), one finally obtains

$$\bar{\mathbf{E}}[\{V_k(q)\hat{y}(t+1, \theta)\}\varepsilon_E(t+1, \theta)] = 0 \quad (17a)$$

$$\bar{\mathbf{E}}[\{V_k(q)u(t+1, \theta)\}\varepsilon_E(t+1, \theta)] = 0 \quad (17b)$$

b) Output error in closed-loop. For this algorithm,  $w(t+1) = S_{yp0}v(t+1)$ , and  $\hat{u}(t) = S_{yp}(q)r_u(t)$ , moreover  $\hat{y}(t) = \frac{B(q)}{A(q)}S_{yp}(q)r_u(t)$ . As in (7) one defines  $\hat{u}_f(t) = \frac{S(q)}{\bar{S}(q)}\hat{u}(t)$  and  $\hat{y}_f(t) = \frac{S(q)}{\bar{S}(q)}\hat{y}(t)$ , where  $\bar{S}(q)$  is a stability polynomial defined from  $S(q)$  whose roots outside the unit circle are projected inside and become the roots of  $\bar{S}(q)$  (enabling the use of an unstable corrector in this algorithm -see (7)). Under assumption H3, one has  $\bar{S}(q) = S(q)$ . The prediction error is given by  $\varepsilon(t+1) = S_{yp}(q)(G_0(q) - G(q))S_{yp0}(q)r_u(t) + w(t+1)$ , and according to (7)  $Q^*(q^{-1}) = \frac{A^*(q^{-1})S^*(q^{-1}) + q^{-1}B^*(q^{-1})R^*(q^{-1})}{A_p^*(q^{-1})S^*(q^{-1})}$ .

The proof is similar to the open-loop case where (9) is replaced by

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\hat{y}_f(t+1, \theta)\}] = 0 \quad (18a)$$

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\hat{u}_f(t+1, \theta)\}] = 0 \quad (18b)$$

Equation (10) is replaced by

$$\bar{\mathbf{E}}[w(t+1)\hat{y}_f(t+1-l, \theta)] = 0 \quad (19a)$$

$$\bar{\mathbf{E}}[w(t+1)\hat{u}_f(t+1-l, \theta)] = 0 \quad (19b)$$

Equation (11) is replaced by

$$\langle V_n, \check{G}|S_{yp}|^2(G_0 - G)S_{yp0}\Phi_{ru} \rangle = 0 \quad (20a)$$

$$\langle V_n, |S_{yp}|^2(G_0 - G)S_{yp0}\Phi_{ru} \rangle = 0 \quad (20b)$$

The application of Theorem 1, similarly to (14) entails

$$\langle V_n, \check{G}Q^*|S_{yp}|^2(G_0 - G)S_{yp0}\Phi_{ru} \rangle = 0 \quad (21a)$$

$$\langle V_n, Q^*|S_{yp}|^2(G_0 - G)S_{yp0}\Phi_{ru} \rangle = 0 \quad (21b)$$

leading like in (17) to

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta)\{V_k(q)\hat{y}_f(t+1, \theta)\}] = 0 \quad (22a)$$

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta)\{V_k(q)\hat{u}_f(t+1, \theta)\}] = 0 \quad (22b)$$

c) Extended least squares in open-loop. For this algorithm, which corresponds to the extended least-squares, where the predictor developed on GBOF (this is a simplification to the open-loop case of H-XCLOE in (7)), one has

$w(t+1) = e(t+1)$ . The prediction error is given by  $\varepsilon(t) = \frac{A}{C} [(G_0 - G)u(t) + (W_0 - \frac{C}{A})e(t)]$ , and  $Q^*(q^{-1}) = \frac{C^*(q^{-1})}{A_p^*(q^{-1})}$ . The proof is again similar to the detailed one for the output error in open-loop. Equation (9) is replaced by

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)y(t+1, \theta)\}] = 0 \quad (23a)$$

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)u(t+1, \theta)\}] = 0 \quad (23b)$$

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\varepsilon(t+1, \theta)\}] = 0 \quad (23c)$$

Since  $\varepsilon(t) = y(t) - \hat{y}(t)$ , the first and third condition lead to

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\hat{y}(t+1, \theta)\}] = 0$$

On the other hand, from (3), the predicted output of this algorithm can be written

$$\hat{y}(t) = G(q)u(t) + \frac{C(q) - A(q)}{A(q)}\varepsilon(t)$$

Therefore like in (11) one can write

$$\langle V_n, \check{G}[G_0 - G]\Phi_u \rangle + \langle V_n, \frac{\check{C} - \check{A}}{\check{A}}[G_0 - G]\Phi_{\varepsilon u} \rangle = 0 \quad (24a)$$

$$\langle V_n, [G_0 - G]\Phi_u \rangle = 0 \quad (24b)$$

$$\langle V_n, \left| \frac{A}{C} \right|^2 \left( [G_0 - G]\Phi_u + [W_0 - \frac{C}{A}]\Phi_e \right) \rangle = 0 \quad (24c)$$

and the application of Theorem 1, similarly to (14), leads to

$$\langle V_n, \check{G}Q^*[G_0 - G]\Phi_u \rangle + \langle V_n, \frac{\check{C} - \check{A}}{\check{A}}Q^*[G_0 - G]\Phi_{\varepsilon u} \rangle = 0 \quad (25a)$$

$$\langle V_n, Q^*[G_0 - G]\Phi_u \rangle = 0 \quad (25b)$$

$$\langle V_n, Q^* \left| \frac{A}{C} \right|^2 \left( |G_0 - G|^2 \Phi_u + |W_0 - \frac{C}{A}|^2 \Phi_e \right) \rangle = 0 \quad (25c)$$

and by shifting in the time domain, like in (17) one obtains finally

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta)\{V_k(q)y(t+1, \theta)\}] = 0 \quad (26a)$$

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta)\{V_k(q)u(t+1, \theta)\}] = 0 \quad (26b)$$

$$\bar{\mathbf{E}}[\varepsilon_E(t+1, \theta)\{V_k(q)\varepsilon(t+1, \theta)\}] = 0 \quad (26c)$$

d) Extended closed-loop output error. For this algorithm  $w(t+1) = e(t+1)$ . The prediction error is given by  $\varepsilon(t) = \frac{A(q)}{C(q)} [(G_0(q) - G(q))S_{yp0}(q)r_u(t) + \dots + (W(q)\frac{S_{yp0}(q)}{S_{yp}(q)} - \frac{C(q)}{A(q)})e(t)]$  and one has  $Q^*(q^{-1}) = \frac{C^*(q^{-1})}{A_p^*(q^{-1})}$ . According to assumptions H1 to H3, and owing to relation (9.32) of (3), the predicted output is given by

$$\hat{y}(t) = G(q)\hat{u}(t) + \frac{C(q)S(q) - P(q)}{A(q)S(q)}\varepsilon(t)$$

Moreover, since  $\hat{u}(t) = -\frac{R(q)}{S(q)}\hat{y}(t) + r_u(t)$  one gets

$$\hat{y}(t) = \frac{BS}{P}r_u(t) + \left( \frac{CS}{P} - 1 \right) \varepsilon(t)$$

$$\hat{u}(t) = \frac{AS}{P}r_u(t) - \left( \frac{CS}{P} - 1 \right) \varepsilon(t)$$

where  $P(q) = A(q)S(q) + B(q)R(q)$ . Set  $\varepsilon_f(t) = \frac{1}{S^*(q^{-1})}\varepsilon(t)$ . The proof follows the same procedure as for the other algorithms, and equations (9) is replaced by

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\hat{y}(t+1, \theta)\}] = 0 \quad (27a)$$

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\hat{u}(t+1, \theta)\}] = 0 \quad (27b)$$

$$\bar{\mathbf{E}}[(\varepsilon(t+1, \theta) - w(t+1))\{V_k(q)\varepsilon_f(t+1, \theta)\}] = 0 \quad (27c)$$

where  $1 \leq k \leq n = \frac{n_a}{n_p}$  for the two first equations and  $1 \leq k \leq n = \frac{2n_a}{n_p}$  for the third one. Equations (11) are replaced by

$$\begin{aligned} & \langle V_k, \frac{\check{B}\check{S}}{\check{P}} \frac{A}{C} (G_0 - G)S_{yp0}\Phi_{ru} \rangle + \langle V_k, \left( \frac{\check{C}\check{S}}{\check{P}} - 1 \right) \dots \\ & \dots \left| \frac{A}{C} \right|^2 \left[ |G_0 - G|^2 |S_{yp0}|^2 \Phi_{ru} + \left| \frac{WS_{yp0}}{S_{yp}} - \frac{C}{A} \right|^2 \Phi_e \right] \rangle = 0 \end{aligned} \quad (28)$$

$$\begin{aligned} & \langle V_k, \frac{\check{A}\check{S}}{\check{P}} \frac{A}{C} (G_0 - G)S_{yp0}\Phi_{ru} \rangle - \langle V_k, \left( \frac{\check{C}\check{S}}{\check{P}} - 1 \right) \dots \\ & \dots \left| \frac{A}{C} \right|^2 \left[ |G_0 - G|^2 |S_{yp0}|^2 \Phi_{ru} + \left| \frac{WS_{yp0}}{S_{yp}} - \frac{C}{A} \right|^2 \Phi_e \right] \rangle = 0 \end{aligned} \quad (29)$$

$$\langle V_k, \frac{1}{\check{S}} \left| \frac{A}{C} \right|^2 \left( |G_0 - G|^2 \Phi_u + \left| W \frac{S_{yp0}}{S_{yp}} - \frac{C}{A} \right|^2 \Phi_e \right) \rangle = 0 \quad (30)$$

and by the same machinery, Theorem 1 leads to

$$\bar{\mathbf{E}}[(\varepsilon_E(t+1, \theta))\{V_k(q)\hat{y}(t+1, \theta)\}] = 0$$

$$\bar{\mathbf{E}}[(\varepsilon_E(t+1, \theta))\{V_k(q)\hat{u}(t+1, \theta)\}] = 0$$

$$\bar{\mathbf{E}}[(\varepsilon_E(t+1, \theta))\{V_k(q)\varepsilon_f(t+1, \theta)\}] = 0$$

This ends the proof.  $\square$

#### 4. SIMULATION EXAMPLE

In this example, the true system is

$G_0(q) = \frac{q^{-2} + 0.5q^{-3}}{(1-1.5q^{-1} + 0.7q^{-2})(1-0.5q^{-1})}$  and we consider a deterministic context. This system is excited with a white noise sequence. Three algorithms are compared

- The output error scheme of PEM where

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega})|^2 \Psi_u(\omega) d\omega$$

( $\Psi_u(\omega)$  being the Power spectrum density (PSD) of the input)

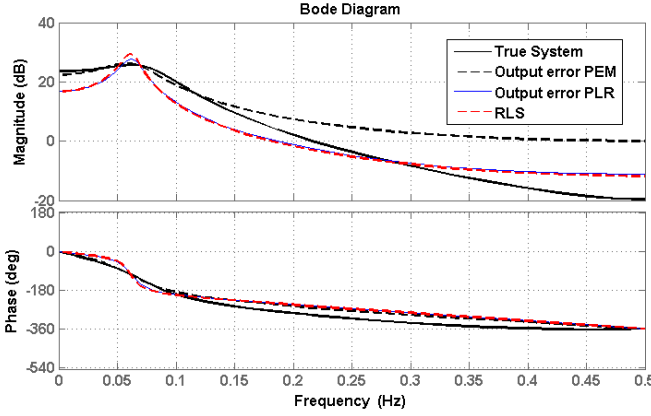


Fig. 1. Comparison of the limit models

- The classical output error scheme of PLR (the basis poles of  $V_k$  are all null) where according (6) and the present developments

$$\theta^* \approx \underset{\theta}{\operatorname{argmin}} \int_{-\pi}^{\pi} |A(e^{i\omega})|^2 |G_0(e^{i\omega}) - G(e^{i\omega})|^2 \Psi_u(\omega) d\omega$$

- The recursive least squares algorithm (RLS) where

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \int_{-\pi}^{\pi} |A(e^{i\omega})|^2 |G_0(e^{i\omega}) - G(e^{i\omega})|^2 \Psi_u(\omega) d\omega$$

The simulation results are displayed in Fig. 1: they confirm the bias analysis proposed here: The limit models of RLS and output error of PLR are very close (only a small discrepancy appears around  $0.06 \text{ rad/s}$  between them). By comparison, in accordance with the limit models expressions in the frequency domain displayed above, the fit of the output error scheme of PEM is better in low frequency and worse in high frequency.

## REFERENCES

- [1] P.S.C Heuberger, T. 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 Optimisation, vol. 42(4), pp 1347-1373, 2003.
- [2] P.S.C Heuberger, P.M.J. Van den Hof, B. Wahlberg (eds), *Modelling and Identification with rational orthonormal basis functions*, Springer, 2005.
- [3] I.D. Landau, R. Lozano, M. M'Saad, A. Karimi, *Adaptive Control*, Springer 2011.
- [4] L. Ljung, *System Identification, Theory for the User, second edition*, Upper Saddle River, Prentice Hall, 1999.
- [5] Ph. Regalia, M. Mboup, M. Ashari, "Existence of stationary points for Pseudo-linear regression identification algorithms", IEEE Trans on Automatic Control, vol. 44(5), 1999.
- [6] B. Vau, H. Bourlès, "Some remarks on the bias distribution analysis of discrete-time identification algorithms based on pseudo-linear regressions", Systems and Control Letters, vol(119), pp. 46-51, 2018.
- [7] B. Vau, H. Bourlès, "Closed-loop output error identification algorithms whit predictors based on generalized orthonormal transfer functions: Convergence conditions and bias distribution", Automatica, vol. 125, 2021.

- [8] B. Vau, H. Bourlès, "A pseudo-linear regression algorithm in discrete-time for the efficient identification of stiff systems", Proc. of the 19th IFAC symposium on system identification (Sysid), 2021.