



**HAL**  
open science

# A robust whitens test for the identification of discrete-time linear models: Use of orthonormal transfer functions

Bernard Vau, Henri Bourlès

## ► To cite this version:

Bernard Vau, Henri Bourlès. A robust whitens test for the identification of discrete-time linear models: Use of orthonormal transfer functions. *Automatica*, 2022, 139, pp.110174. 10.1016/j.automatica.2022.110174 . hal-03525813

**HAL Id: hal-03525813**

**<https://hal.science/hal-03525813>**

Submitted on 14 Jan 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.

# A robust whiteness test for the identification of discrete-time linear models: Use of orthonormal transfer functions

Bernard Vau, Henri Bourlès

*SATIE, Ecole normale supérieure de Paris-Saclay, 4 avenue des sciences, 91190 Gif-sur-Yvette, France*

---

## Abstract

A novel whiteness test of residuals is proposed, which makes use of generalized bases of orthonormal transfer functions. It can be viewed as a robustified version of the classical whiteness test in the sense that it reduces the risk of type II errors, by introducing a frequency weighting in the assessment of the flatness in the residual power spectrum density. This frequency weighting, which depends on the basis poles, can be employed for the validation of reduced order models, when the flatness of the residual power spectrum density is evaluated over a limited frequency band.

*Key words:* Identification, validation methods for identification, whiteness test

---

## 1 Introduction

In the field of automatic control and signal processing, the necessity of assessing the whiteness of a signal is often encountered in practice. In particular, this is the case when one wants to validate a model resulting from identification (see [11], [7]), by checking the whiteness of the residual between the true system output and the predicted one. The issue of determining residual whiteness has been of the utmost importance in the identification field for decades, since it is related to the question of model order selection<sup>1</sup>. A simple way consists in evaluating the residual autocorrelation functions which are asymptotically null in case of whiteness (for a time lag different from 0). The squares sum of these functions obey to a chi-square distribution, and the null hypothesis  $H_0$  asserting that the residual is white can be rejected with a risk  $\alpha$ , if this sum is larger than a threshold depending on the degrees of freedom. This is the principle of the Box-Pierce test [3]. In any statistical test two kinds of errors can occur: The type I error corresponds to the risk of rejecting  $H_0$  when it is true (and is measured by  $\alpha$ ), the type II error is the risk of accepting  $H_0$  when it is false (which depends on the

statistical power of the test and is unknown in general). However, from a practical point of view this second risk is far more crucial in many situations.

In Section 2 it is shown by a simple example, that the classical whiteness test can entail a large proportion of type II errors, if the flatness defect in the residual power density spectrum occurs in low frequency. That leads to propose in Section 3 a more robust whiteness test by assessing the cross-correlation of outputs resulting from generalized orthonormal basis functions (GOBF) (introduced by Heuberger et al., see [6]), fed with the residual. The statistical properties of these cross-correlation functions are detailed, allowing for a chi-square based whiteness evaluation. The use of GOBF induces a frequency weighting depending on the basis pole selection. As shown in [12], this selection yields a dilatation or a compression of the frequency scale which can be used for a robust evaluation of the residual whiteness over a large frequency band, or on the contrary to validate reduced-order models, as shown in Section 4.

## 2 The classical whiteness test and its limitation

At first, let us recall briefly the principle of the classical whiteness test of a residual sequence  $\{\varepsilon(t)\}$ . Define  $r_{\varepsilon\varepsilon}^{(N)}(0) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t)$ , an estimation of the

---

*Email address:* `bernard.vau@satie.ens-cachan.fr`  
(Bernard Vau, Henri Bourlès).

<sup>1</sup> However, in more recent developments stemming from the machine learning community, the model order is no longer the key tuning parameter of identification algorithms, see [9].

residual variance over  $N$  samples, and  $r_{\varepsilon\varepsilon}^{(N)}(k) = \frac{1}{N} \sum_{t=1}^N \varepsilon(t)\varepsilon(t-k)$  an estimation of the cross-correlation function with a time lag  $k$ , where  $k = 1, 2, \dots, k_{max}$ . Let  $rn_{\varepsilon\varepsilon}^{(N)}(k) = \frac{r_{\varepsilon\varepsilon}^{(N)}(k)}{r_{\varepsilon\varepsilon}^{(N)}(0)}$  be the corresponding normalized functions. It is well known that if  $\{\varepsilon(t)\}$  is a white noise sequence, one has  $\mathbf{E}[rn_{\varepsilon\varepsilon}^{(N)}(k)] = 0$  and  $\mathbf{E}[(rn_{\varepsilon\varepsilon}^{(N)}(k))^2] = \frac{1}{N}$  (see [7] p. 512). Moreover,  $rn_{\varepsilon\varepsilon}^{(N)}(k)$  converges in distribution towards a Gaussian stochastic variable as  $N \rightarrow \infty$ , and thus the quantity  $\mathcal{T}(N, k_{max}) = N \sum_{k=1}^{k_{max}} (rn_{\varepsilon\varepsilon}^{(N)}(k))^2$  converges towards a chi-square distributed variable. Therefore a threshold  $\mathbf{K}_\alpha(k_{max})$  as a function of the risk  $\alpha$  and the degree of freedom  $k_{max}$  can be defined following a chi-square table, leading to the rejection of  $H_0$  if  $\mathcal{T}(N, k_{max}) > \mathbf{K}_\alpha(k_{max})$ . As said in the introduction, the risk of type II error is not known and depends on how much the residual differs from a white noise [11], p. 425. If the purpose of the test is to assess the whiteness of the residual over a large frequency band (for example in order to validate a fast sampled system), the classical whiteness test can entail considerable type II errors, as shown by the following example: Let us assume that  $\{\varepsilon(t)\}$  results from the filtering of a Gaussian white noise sequence  $\{e(t)\}$  by a transfer operator  $G(q)$  such that  $\varepsilon(t) = G(q)e(t)$ , the system  $G$  having two complex poles at  $0.002Hz$  with a damping of 0.5, and two zeros at  $0.001Hz$  with a damping of 0.5 too. The sample time is set to  $Te = 1s$  and the corresponding Power Spectrum Density (PSD) of  $\{\varepsilon(t)\}$  is displayed in Fig. 1

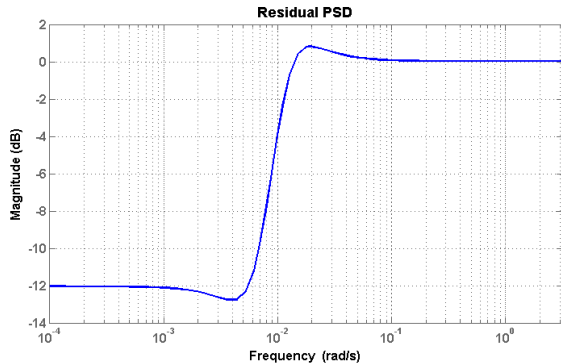


Fig. 1. Power spectrum density of  $\varepsilon(t)$

The residual sequence  $\{\varepsilon(t)\}$  is obviously not white, since a patent default of flatness of this PSD occurs around  $10^{-2}rad/s$ . However, as shown in Table. 1, the percentage of type II errors for  $\alpha = 0.05$  (a value commonly employed) is considerable, unless  $N$  is very large (at least 200 or 400 times the filter dominant modes period). In order to account for this result, it is useful to consider the asymptotic expression of  $rn_{\varepsilon\varepsilon}^{(N)}(k)$  in the frequency

domain. One has immediately

$$\lim_{N \rightarrow \infty} rn_{\varepsilon\varepsilon}^{(N)}(k) = \frac{\int_{-\pi}^{+\pi} \Phi_{\varepsilon\varepsilon}(\omega) \cos(k\omega) d\omega}{\int_{-\pi}^{+\pi} \Phi_{\varepsilon\varepsilon}(\omega) d\omega} \quad (1)$$

And since the frequency scale of Fig. 1 is logarithmic, one must express  $rn_{\varepsilon\varepsilon}^{(N)}(k)$  in the same scale with  $\bar{\omega} = \log(\omega)$ , and one obtains<sup>2</sup>

$$\lim_{N \rightarrow \infty} rn_{\varepsilon\varepsilon}^{(N)}(k) = \frac{\int_{-\infty}^{\log(\pi)} \Phi_{\varepsilon\varepsilon}(\bar{\omega}) \cos(ke^{\bar{\omega}}) e^{\bar{\omega}} d\bar{\omega}}{\int_{-\infty}^{\log(\pi)} \Phi_{\varepsilon\varepsilon}(\bar{\omega}) e^{\bar{\omega}} d\bar{\omega}} \quad (2)$$

Owing to the weighting term  $e^{\bar{\omega}}$ , the flatness default of the PSD that appears clearly in Fig. 1 is severely underweighted in (2): if one considers in this example that even for  $N = 15000$  the limit expression of (2) is a good approximation of  $rn_{\varepsilon\varepsilon}^{(N)}(k)$ , that can explain why the whiteness test is so insensitive to this low frequency default, unless a very large amount of data is available such that the test threshold depending on  $\mathbf{E}[(rn_{\varepsilon\varepsilon}^{(N)}(k))^2] = \frac{1}{N}$  becomes very low.

N \ k <sub>max</sub>	30	50	100	200	500
15000	93%	93%	86%	80%	77%
30000	94%	87%	49%	40%	58%
50000	90%	72%	23%	8%	36%
100000	65%	27%	2%	0%	0%
200000	22%	0%	0%	0%	0%

Table 1

Percentage of type II errors, in function of the sample number and the max. time lags of the auto-correlation function (classical whiteness test with  $\alpha = 0.05$  -100 realizations)

### 3 A whiteness test established on a generalized orthogonal functions basis

The novel test proposed here uses the generalized bases of orthonormal transfer functions (GOBF) proposed by Heuberger et al. in [6], that stem from a balanced realization of an all-pass function  $G_b(z) = \prod_{j=0}^{n_p-1} \frac{-z \cdot \bar{p}_j + 1}{z - p_j}$  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 [6]). 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

<sup>2</sup> The bar of  $\bar{\omega}$  must not be confounded with the complex conjugate symbol employed elsewhere

of the orthonormal state space realisation of  $G_b(z)$ , orthonormality between these functions holds (meaning that  $\oint_{\mathbf{T}} V_k^T(z)V_k'(z^{-1})\frac{dz}{z} = \delta_{k,k'}$ , where  $\mathbf{T}$  is the unit circle, and  $\delta = 1$  if  $k = k'$ ,  $\delta = 0$  otherwise). 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.

A transform is associated to the basis functions: The Hambo transform. The Hambo operator  $\lambda$  is given by  $\lambda^{-1} = G_b(z)$ . The mapping  $\lambda \mapsto z$  is multi-valued in a domain including the unit circle, and the  $z_j$  ( $j = 1, \dots, n_p$ ) such that  $G_b(z_j) \in \lambda^{-1}$  are the eigenvalues of  $\mathcal{N}(1/\lambda)$  with  $\mathcal{N}(\lambda) = A_b + (\lambda - D_b)^{-1}C_b$  (see [6] Section 3.3.3). The Hambo frequency  $\omega_\lambda \in [-n_p\pi; n_p\pi]$ , with  $\omega_\lambda = \beta(\omega)$  and  $e^{-i\omega_\lambda} = G_b(e^{i\omega})$  is such that  $\omega_\lambda = \sum_{k=0}^{n_p-1} \eta_k + 2 \arctan\left(\left(\frac{1+\rho_k}{1-\rho_k}\right) \tan\left(\frac{\omega-\eta_k}{2}\right)\right)$ , where  $p_k = \rho_k e^{i\eta_k}$ . Furthermore, one has  $d\omega_\lambda = \beta'(\omega)d\omega$ , with  $\beta'(\omega) = V_1^T(e^{i\omega})V_1(e^{-i\omega})$  (see [10]). Now define

$$w_k(t) = V_k(q)\varepsilon(t)$$

If  $\{\varepsilon(t)\}$  is a white noise with variance  $\lambda^2$ , one has immediately  $\mathbf{E}[w_1^T(t)w_k(t)] = \lambda^2 n_p \delta_{1,k}$ , with  $\delta_{1,k} = 1$  for  $k = 1$ , and  $\delta_{1,k} = 0$ , otherwise. In the following, we evaluate the whiteness of  $\{\varepsilon(t)\}$  by assessing the cross correlation between  $\{w_1\}$  and  $\{w_k\}$ . For this purpose, one defines the normalized cross-correlation functions

$$rn_{ww}^{(N)}(k) = \frac{r_{ww}^{(N)}(k)}{r_{ww}^{(N)}(1)} = \frac{\frac{1}{N} \sum_{t=1}^N w_1^T(t)w_k(t)}{\frac{1}{N} \sum_{t=1}^N w_1^T(t)w_1(t)} \quad (3)$$

which should be asymptotically null if  $\{\varepsilon(t)\}$  is white for  $k = 2, 3, \dots, k_{max}$ . By shifting in the frequency domain, one has immediately

$$\lim_{N \rightarrow \infty} rn_{ww}^{(N)}(k) = \frac{\int_{-\pi}^{+\pi} \Phi_{\varepsilon\varepsilon}(\omega) G_b^{k-1}(e^{i\omega}) \beta'(\omega) d\omega}{\int_{-\pi}^{+\pi} \Phi_{\varepsilon\varepsilon}(\omega) \beta'(\omega) d\omega} \quad (4)$$

This expression can be rewritten in the Hambo frequency domain

$$\lim_{N \rightarrow \infty} rn_{ww}^{(N)}(k) = \frac{\int_{-n_p\pi}^{+n_p\pi} \tilde{\Phi}_{\varepsilon\varepsilon}(\omega_\lambda) \cos(k\omega_\lambda) d\omega_\lambda}{\int_{-n_p\pi}^{+n_p\pi} \tilde{\Phi}_{\varepsilon\varepsilon}(\omega_\lambda) d\omega_\lambda} \quad (5)$$

where  $\tilde{\Phi}_{\varepsilon\varepsilon}(\omega_\lambda) = \Phi_{\varepsilon\varepsilon}(\omega)|_{\omega=\beta^{-1}(\omega_\lambda)}$ . One finds the same expression as (1), but now in the distorted Hambo frequency scale.

Let us consider again the example in Section 2, and let us express  $\tilde{\Phi}_{\varepsilon\varepsilon}(\omega_\lambda)$  in the Hambo scale  $\omega_\lambda$  for various

Laguerre bases: For  $p = 0$  (the  $\omega$  scale) the residual appears very slightly coloured, and this is no longer the case for  $p = 0.9$  or  $p = 0.95$  as shown in Fig. 2.

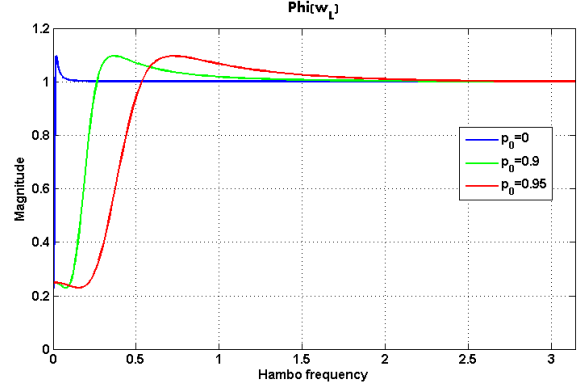


Fig. 2. Expression of  $\tilde{\Phi}_{\varepsilon\varepsilon}(\omega_\lambda)$  in various Hambo frequency scales, in function of the basis poles ( $p = 0$  blue,  $p = 0.9$  green,  $p = 0.95$  red)

Another manner to consider the problem consists in expressing (4) in the logarithmic frequency scale  $\bar{\omega} = \log(\omega)$  (the scale associated to the Bode diagram of Fig. 1) which yields

$$\lim_{N \rightarrow \infty} rn_{ww}^{(N)}(k) = \frac{\int_{-\infty}^{\log(\pi)} \Phi_{\varepsilon\varepsilon}(e^{\bar{\omega}}) \cos(k\beta(e^{\bar{\omega}})) \chi(e^{\bar{\omega}}) d\bar{\omega}}{\int_{-\infty}^{\log(\pi)} \Phi_{\varepsilon\varepsilon}(e^{\bar{\omega}}) \chi(e^{\bar{\omega}}) d\bar{\omega}} \quad (6)$$

where, according to [6], p. 222

$$\chi(e^{\bar{\omega}}) = \frac{1}{\pi} e^{\bar{\omega}} \bar{V}_1^T(e^{ie^{\bar{\omega}}}) V_1(e^{-ie^{\bar{\omega}}}) = \frac{1}{\pi} e^{\bar{\omega}} \sum_{j=0}^{n_j-1} \frac{1 - |p_j|^2}{|1 - \bar{p}_j e^{ie^{\bar{\omega}}}|^2} \quad (7)$$

The function  $\chi$ <sup>3</sup> is now the weighting term associated with the proposed whiteness test established on GOBF, and it substitutes the term  $e^{\bar{\omega}}$  appearing in (2). This function  $\chi$  has been studied in [12] and is directly related to the reproducing kernel of the Hilbert space associated with the GOBF. In particular, the following nice properties exist (the proof can be found in [12]).

- A conservation principle holds:  $\int_{-\infty}^{\log(\pi)} \chi(e^{\bar{\omega}}) d\bar{\omega} = 1$
- For a basis pole  $p_j$  sufficiently close to the unit circle, the function  $\bar{\omega} \mapsto \frac{1}{\pi} e^{\bar{\omega}} \frac{1 - |p_j|^2}{|1 - \bar{p}_j e^{ie^{\bar{\omega}}}|^2}$  has a local maximum roughly at the frequency of the pole  $p_j$ .

Consequently, the weighting function  $\chi$  can be modelled by the selection of the basis poles  $p_j$ . As an example,

<sup>3</sup> This variable should not be confused with that of the chi-square test

Fig. 2 shows various functions  $\chi$ , and are compared with the weighting term of the classical whiteness test where  $p_0 = 0$ .

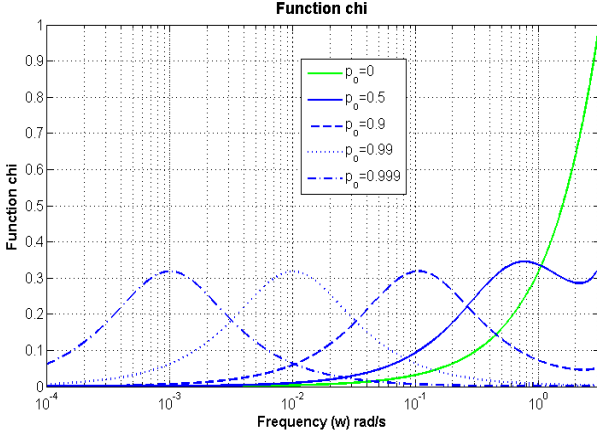


Fig. 3. Functions  $\chi(e^{\omega})$  for Laguerre bases

The following results hold:

**Lemma 1** If  $\{\varepsilon(t)\}$  is a centered Gaussian white noise such that  $\mathbf{E}[\varepsilon^2(t)] = \lambda^2$ , then for any integer  $k > 1$ , one has  $\mathbf{E}[r_{ww}^{(N)}(k)] = 0$ , and  $r_{ww}^{(N)}(k)$  converges in distribution towards a Gaussian variable as  $N$  tends to infinity.

*Proof:* see Appendix ?? .  $\square$

**Lemma 2** Consider a transfer function basis having only one (possibly complex) pole, such that

$$V_1(z) = \frac{\sqrt{1-|p_0|^2}z^{-1}}{1-p_0z^{-1}}.$$

Set  $r_{1k}(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \bar{w}_1(t)w_k(t+\tau)$  where  $\bar{w}_1(t)$  is the complex conjugate of  $w_1(t)$ . If the residual  $\{\varepsilon(t)\}$  is a Gaussian centered white noise, for any  $\tau \leq 0$  the following relation is satisfied

$$r_{1k}(\tau) = 0$$

*Proof:* One has  $w_k(t) = \frac{\sqrt{1-|p_0|^2}q^{-1}}{1-p_0q^{-1}} \left( \frac{-p_0+q^{-1}}{1-p_0q^{-1}} \right)^{k-1} \varepsilon(t)$ .

Therefore  $w_k(t) = p_0w_k(t-1) + v_1(t-1)$  where  $v_1(t) = \sqrt{1-|p_0|^2} \left( \frac{-p_0+q^{-1}}{1-p_0q^{-1}} \right)^{k-1} \varepsilon(t)$ . Thus one obtains

$r_{1k}(\tau) = p_0r_{1k}(\tau-1) + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \bar{w}_1(t)v_1(t-1+\tau)$ . Let  $\Phi_{\bar{w}_1v_1}(z)$  be the Cross Power Spectral Density (CPSD) associated to  $\bar{w}_1(t)$  and  $v_1(t)$ .

Since  $\bar{w}_1(t) = \frac{\sqrt{1-|p_0|^2}q^{-1}}{1-p_0q^{-1}} \varepsilon(t)$ , by the interference formula<sup>4</sup> on one has the relation  $\Phi_{\bar{w}_1v_1}(z) = (1 - |p_0|^2) \frac{z}{1-p_0z} \left( \frac{-p_0+z^{-1}}{1-p_0z^{-1}} \right)^{k-1} \Phi_{\varepsilon\varepsilon}(z)$ . Therefore

<sup>4</sup> If  $x_1$  and  $x_2$  are two signals,  $H_1(q)$ ,  $H_2(q)$  are two stable causal operators with complex coefficients, and  $y_j = H_j(q)x_j$  ( $j = 1, 2$ ), then according to the interference formula  $\Phi_{y_1, y_2}(z) = \bar{H}_1(z^{-1})H_2(z)\Phi_{x_1, x_2}(z)$ , see [4] p. 326.

$\Phi_{\bar{w}_1v_1}(z) = (1 - |p_0|^2) \frac{(-p_0+z^{-1})^{k-2}}{(1-p_0z^{-1})^{k-1}} \Phi_{\varepsilon\varepsilon}(z)$ . Now set  $v'_1(t) = (1 - |p_0|^2) \frac{(-p_0+q^{-1})^{k-2}}{(1-p_0q^{-1})^{k-1}} \varepsilon(t)$ . Consequently  $\Phi_{\bar{w}_1v_1}(z) = \Phi_{\varepsilon v'_1}(z)$ . By shifting in the time domain one gets  $r_{1k}(\tau) = p_0r_{1k}(\tau-1) + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \varepsilon(t)v'_1(t-1+\tau)$ . Thus if  $\{\varepsilon(t)\}$  is white,  $r_{1k}(0) = 0$  by the orthogonality of functions  $V_k$ , and for  $\tau \leq 0$   $\lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{t=-N}^N \varepsilon(t)v'_1(t-1+\tau) = 0$ . Combining this last result with  $r_{1k}(0) = 0$ , one concludes that  $r_{1k}(-1) = 0$ . Now let us assume that for  $\tau \leq 0$  one has  $r_{1k}(\tau) = 0$ . Since  $\lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{t=-N}^N \varepsilon(t)v'_1(t-1+\tau) = 0$  and owing to the expression of  $r_{1k}(\tau)$  above, one has necessarily  $r_{1k}(\tau-1) = 0$ , and therefore by induction  $r_{1k}(\tau) = 0 \forall \tau \leq 0$ .

**Theorem 1** Consider a transfer function basis having only one (possibly complex) pole, such that  $V_1(z) = \frac{\sqrt{1-|p_0|^2}z^{-1}}{1-p_0z^{-1}}$ . If the residual  $\{\varepsilon(t)\}$  is a Gaussian centered white noise such that  $\mathbf{E}[\varepsilon^2(t)] = \lambda^2$ , then for any integer  $k > 1$ , the following results hold

(1)

$$\mathbf{E}\left[\left(r_{ww}^{(N)}\right)^2(k)\right] = \mathbf{E}[g_0^2(t)] \frac{\lambda^4}{N} \quad (8)$$

where  $g_0(t)$  is given by  $g_0(t) = \left( \frac{\sqrt{1-|p_0|^2}}{1-|p_0|q^{-1}} \right)^2 e_n(t)$ ,

in function of a centered white noise  $\{e_n(t)\}$  with variance equal to 1.

(2)

$$\mathbf{E}\left[\left(r_{ww}^{(N)}\right)^2(k)\right] = \frac{\lambda^4}{N} \frac{1+|p_0|^2}{1-|p_0|^2} \quad (9)$$

*Proof:* Proof of 1. One has  $\lim_{N \rightarrow \infty} N \mathbf{E}\left[\left(r_{ww}^{(N)}\right)^2(k)\right] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=1}^N \mathbf{E}[\bar{w}_1(t)w_k(t)\bar{w}_1(s)w_k(s)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=1}^{t-1} \mathbf{E}[\bar{w}_1(t)w_k(t)\bar{w}_1(s)w_k(s)] + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=t+1}^N \mathbf{E}[\bar{w}_1(t)w_k(t)\bar{w}_1(s)w_k(s)] + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{E}[\bar{w}_1(t)w_k(t)\bar{w}_1(t)w_k(t)]$ .

Now, owing to Lemma 2 one obtains  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=1}^{t-1} \mathbf{E}[\bar{w}_1(t)w_k(t)\bar{w}_1(s)w_k(s)] = 0$

Similarly  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=t+1}^N \mathbf{E}[\bar{w}_1(t)w_k(t)\bar{w}_1(s)w_k(s)] = 0$ .

Moreover,  $\mathbf{E}[\bar{w}_k(t)w_k(t)] = \mathbf{E}[\bar{w}_1(t)w_1(t)]$ . Therefore, by shifting in the frequency domain and by considering

$p_0 = |p_0|e^{i\varphi}$ , one obtains  $\lim_{N \rightarrow \infty} N \mathbf{E}\left[\left(r_{ww}^{(N)}\right)^2(k)\right] = \lim_{N \rightarrow \infty} N \mathbf{E}[(\bar{w}_1(t)w_1(t))^2] =$

$$\frac{\lambda^4}{2\pi} \int_{-\pi}^{\pi} \left( \frac{1-|p_0|^2}{1+|p_0|^2-2|p_0|\cos(\omega+\varphi)} \right)^2 d\omega =$$

$$\frac{\lambda^4}{2\pi} \int_{-\pi}^{\pi} \left( \frac{1-|p_0|^2}{1+|p_0|^2-2|p_0|\cos(\omega)} \right)^2 d\omega.$$

But  $\left( \frac{1-|p_0|^2}{1+|p_0|^2-2|p_0|\cos(\omega)} \right)^2$  is the spectral density of a

signal that is the output of the filter  $\left(\frac{\sqrt{(1-|p_0|^2)}}{1-|p_0|q^{-1}}\right)^2$  fed by a white noise  $\{e(t)\}$  with variance 1.

*Proof of 2.* It is known (see for example [2]) that the covariance associated to the output of the ARMA process  $\frac{1}{(1-|p_0|z^{-1})^2}$  driven by a white noise with variance equal to 1 is  $\frac{1+|p_0|^2}{(1-|p_0|^2)^3}$ , therefore by combining with the result of 1, one gets  $\mathbf{E}[(r_{ww}^{(N)})^2(k)] = \frac{\lambda^4}{N} \frac{1+|p_0|^2}{1-|p_0|^2}$ .  $\square$

Let us now consider the general case of basis functions having several (possibly complex) poles  $p_j$  ( $j = 1, \dots, n_p$ )

**Theorem 2** *If  $\{\varepsilon(t)\}$  is a centered Gaussian white noise, then for any integer  $k > 1$ , one has:*

$$\mathbf{E}[(r_{ww}^{(N)})^2(k)] = \frac{\lambda^4}{N} \sum_{j=0}^{n_p-1} \frac{1+|p_j|^2}{1-|p_j|^2} \quad (10)$$

*Proof:* The Transfer function  $V_1(z)$  can be chosen such that:

$$V_1(z) = \begin{bmatrix} \frac{\sqrt{1-|p_0|^2}z^{-1}}{1-p_0z^{-1}} \\ \frac{\sqrt{1-|p_1|^2}z^{-1}}{1-p_1z^{-1}} \frac{-\bar{p}_0+z^{-1}}{1-p_0z^{-1}} \\ \vdots \\ \frac{\sqrt{1-|p_{n_p-1}|^2}z^{-1}}{1-p_{n_p-1}z^{-1}} \prod_{k=0}^{n_p-2} \frac{-\bar{p}_k+z^{-1}}{1-p_kz^{-1}} \end{bmatrix} \quad (11)$$

the entries of which are Takenaka-Malmquist transfer functions (see [6], chap. 2, p. 18). If  $\{\varepsilon(t)\}$  is a Gaussian white noise,  $\{w_1(t)\}$  and  $\{w_k(t)\}$  are Gaussian signal vectors and one has:  $\mathbf{E}[(\bar{w}_1^T(t)w_k(t))^2] = \sum_{l=1}^{n_p} \mathbf{E}[(\bar{w}_{1(l)}(t)w_{k(l)}(t))^2]$  for  $l \in [1, n_p]$  ( $w_{k(l)}$  is the  $l^{\text{th}}$  entry of  $w_k$ , and  $V_{k(l)}$  is the  $l^{\text{th}}$  entry of  $V_k$ ). And from Result 1 of Theorem 1, one has  $\mathbf{E}[(\bar{w}_{1(l)}(t)w_{k(l)}(t))^2] = \frac{\lambda^4}{N} \mathbf{E}[g_i^2(t)]$ , with  $g_i(t) = (V_{1(l)}(z))^2 e_n(t)$ ,  $\{e_n(t)\}$  being a centered Gaussian white noise with variance 1. But according to the choice of  $V_1(z)$  that has been done above, and to result 2 of Theorem 1, one gets  $\mathbf{E}[g_i^2(t)] = \frac{1+|p_i|^2}{1-|p_i|^2}$ , thus for this choice of  $V_1(z)$ , one obtains  $\mathbf{E}[(\bar{w}_1^T(t)w_k(t))^2] = \frac{\lambda^4}{N} \sum_{j=0}^{n_p-1} \frac{1+|p_j|^2}{1-|p_j|^2}$ . If another basis vector  $V_1'(z)$  is chosen instead of  $V_1(z)$ , as pointed out in [5],  $V_1'(z)$  can be obtained from  $V_1(z)$  by a premultiplication of  $V_1(z)$  with a square complex unitary matrix  $M_u$  so that  $V_1'(z) = M_u V_1(z)$ . Define  $w_1'(t) = V_1'(z)\varepsilon(t)$ ,  $w_k'(t) = V_k'(z)\varepsilon(t)$ . We get  $\mathbf{E}[(\bar{w}_1'(t)w_k'(t))^2] = \mathbf{E}[(\bar{w}_1(t)M_u^T M_u w_k(t))^2] = \mathbf{E}[(\bar{w}_1 w_k(t))^2]$ .  $\square$

Consequently if  $\{\varepsilon(t)\}$  is white, one has for  $k > 1$

$$\mathbf{E}[(rn_{ww}^{(N)})^2(k)] = \frac{1}{Nn_p} \sum_{j=0}^{n_p-1} \frac{1+|p_j|^2}{1-|p_j|^2} \quad (12)$$

Now a chi-square test with  $k_{max}$  degrees of freedom can be performed on the normalized quantity

$$\mathcal{T}(N, k_{max}) = \frac{Nn_p}{\sum_{j=0}^{n_p-1} \frac{1+|p_j|^2}{1-|p_j|^2}} \sum_{k=1}^{k_{max}} \left( rn_{ww}^{(N)}(k+1) \right)^2 \quad (13)$$

Let us go back to the example of section 2: The whiteness test is now performed with a Laguerre basis where  $p_0 = 0.9$ . Table. 2 shows that the percentage of type II errors is dramatically reduced, even for a low amount of data ( $N = 15000$  or  $N = 30000$ ). This can be explained by the frequency weighting induced by the test (see the corresponding function  $\chi$  for  $p_0 = 0.9$  in Fig. 3), in the frequency area where the defect of flatness occurs in the PSD of  $\{\varepsilon(t)\}$ . In low frequency, the all-pass function  $\frac{-p_0z+1}{z-p_0}$  can be considered approximately as a delay (called the Laguerre shift, see [6], chap. 3), equal to  $\frac{1+p_0}{1-p_0}$ , and the present test remains reliable provided  $\frac{1+p_0}{1-p_0} k_{max} \ll N$ , which is a generalization for a Laguerre basis of the condition  $k_{max} \ll N$  in the classical whiteness test. In this example, one has  $\frac{1+p_0}{1-p_0} k_{max} = 950$  for  $k_{max} = 50$ , and  $\frac{1+p_0}{1-p_0} k_{max} = 1900$  for  $k_{max} = 100$ .

N	$k_{max}$		
	30	50	100
15000	0%	3%	23%
30000	0%	0%	0%
50000	0%	0%	0%
100000	0%	0%	0%
200000	0%	0%	0%

Table 2

Percentage of type II errors (100 realizations,  $\alpha = 0.05$ ), in function of the sample number  $N$  and  $k_{max}$  (whiteness test established on a Laguerre basis and  $p_0 = 0.9$ )

#### 4 An application of the frequency weighting to the validation of reduced order models

Another interest of the proposed test is to allow for the evaluation of a reduced order model. For controller synthesis purposes, it is quite common to employ reduced order models valid only in low frequency, the high frequency modes being considered as unstructured uncertainties for which the designed controller must be sufficiently robust. We propose now to show on an example the interest of the whiteness test with frequency weighting of section 3. In this example, the true system has an order equal to 4, its static gain is 1 and its poles and its zeros are given in Table. 3, the sample time being 1 second.

Poles		Zeros	
Frequency (rad/s)	Damping	Frequency (rad/s)	Damping
$1.0 \cdot 10^{-2}$	0.5	$1.10 \cdot 10^{-2}$	1
$1.0 \cdot 10^{-2}$	0.5	$2.1 \cdot 10^{-1}$	0.273
$2.4 \cdot 10^{-1}$	0.115	$2.1 \cdot 10^{-1}$	0.273
$2.4 \cdot 10^{-1}$	0.115		

Table 3

Poles and zeros of the system to be identified

This system is disturbed by an output (centered and Gaussian) white noise, such that the signal/noise ratio (variance) is equal to 10 dB, and this system is excited with a wide spectrum Pseudo Random Binary Sequence (PRBS). One assumes that a reduced order model (second order) issued from identification has been obtained which fits well with the true system in low frequency, as shown in Fig. 4 a). The test described in Section 3 is performed for various sets of basis poles:

- Test 1:  $p = 0$  Classical whiteness test
- Test 2:  $p = 0.99$  Frequency weighting centred around the low frequency modes
- Test 3:  $p = 0.995 \pm 0.0086i$  Selective frequency weighting centred around the low frequency modes (poles damping: 0.5)
- Test 4:  $p = 0.996 \pm 0.0091i$  Very selective frequency weighting centred around the low frequency modes (poles damping 0.4)

The functions  $\chi$  associated with each test are represented in Fig. 4 b) and Table. 4 displays the percentage of acceptance of  $H_0$  (in simulation) of each test for 100 realizations, and for various sample data  $N$  (the ratio of  $N$  by the period of low frequency modes of the true system denoted  $T_m$  is displayed -here one has  $T_m = 1/0.01 = 100s$ ).

Test number \ $N/T_m$	52	104	209	417
1	0%	0%	0%	0%
2	97%	63%	5%	0%
3	100%	100%	98%	98%
4	100%	100%	100%	100%

Table 4

Percentage of acceptance of  $H_0$  for the reduced order model (over 100 realizations), in function of the sample number  $N$  and the period of the low frequency modes  $T_m$

As expected, the classical whiteness test (test 1) leads systematically to the rejection of  $H_0$  because of the associated function  $\chi$ , and the discrepancy between the true system and the model that occurs only in high frequency. On the contrary, Table 4 shows that a sufficiently frequency weighted test (as in tests 3 and 4), leads to the quasi-systematic acceptance of  $H_0$ . Other simulations, where a model misfit occurs in low frequency (the static

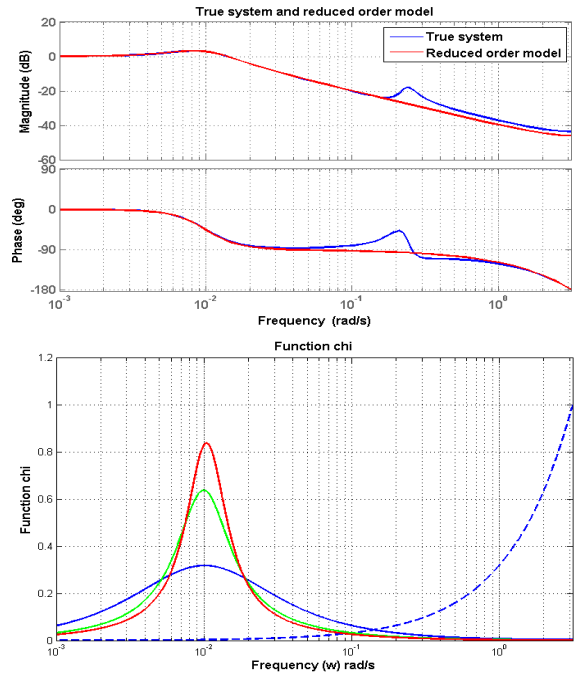


Fig. 4. a): True system and reduced order model, b): Function  $\chi$  for various basis poles

gain of the reduced order model is 1.1 instead of 1) lead to 100% rejection of  $H_0$  whatever the basis poles. This example shows that the frequency weighted test proposed here is able to discriminate reduced order models.

## 5 Concluding remarks

This paper has shown that the classical whiteness test of residuals is not robust if the whiteness default occurs in low frequency, and this can be accounted for if one considers the asymptotic expression (2) of the residual autocorrelation functions in the frequency domain. That led to propose a novel whiteness test based on the signals cross-correlation (3) resulting from orthonormal transfer functions fed by the residual. The asymptotic expression of these cross-correlation functions (5) in the distorted Hambo frequency scale reveals to have a form similar to (2). The frequency distortion from the classical logarithmic frequency scale to the Hambo frequency scale is expressed by the function  $\chi$  given in (7), and depends on the basis pole selection. The covariance of these cross-correlation functions is provided in Theorem 2, and it is shown that a proper basis pole selection overcomes the lack of reliability of the classical whiteness test, for the detection of whiteness defaults in low frequency. Another interest of the test, as shown in Section 4, is its ability to evaluate the fit of an identified model over a restricted frequency band, by selecting the basis poles such that the corresponding function  $\chi$  presents a narrow peak around the frequencies of interest. In a more general way, the simulation examples of this article

show that the presented test can be employed advantageously in a perspective of identification for control, especially in case of fast sampling, a situation that occurs more and more frequently in practice.

## A Appendix: Proof of Lemma 1

The proof is inspired from the first part of the proof of Theorem 9.1 in [7], p. 309. Set  $S_{N,l} = \frac{1}{\sqrt{N}} \sum_{t=1}^N w_{1(l)}(t)w_{k(l)}(t)$ , where  $l \in [1, n_p]$ . One has  $r_{ww}^{(N)}(k) = \frac{1}{\sqrt{N}} S_{N,l}$ . We are going to demonstrate that  $r_{ww}^{(N)}(k)$  is Gaussian. One can write  $w_{1(l)}(t) = \frac{1}{\sqrt{N}} \sum_{l=0}^{\infty} d^{(1,l)} \varepsilon(t-l)$  and  $w_{k(l)}(t) = \frac{1}{\sqrt{N}} \sum_{l=0}^{\infty} d^{(k,l)} \varepsilon(t-l)$ . Let  $M$  be an integer and define  $w_{1(l)}^M(t) = \frac{1}{\sqrt{N}} \sum_{l=0}^M d^{(1,l)} \varepsilon(t-l)$ ,  $w_{k(l)}^M(t) = \frac{1}{\sqrt{N}} \sum_{l=0}^M d^{(k,l)} \varepsilon(t-l)$ ,  $\tilde{w}_{1(l)}^M(t) = \frac{1}{\sqrt{N}} \sum_{l=M+1}^{\infty} d^{(1,l)} \varepsilon(t-l)$  and  $\tilde{w}_{k(l)}^M(t) = \frac{1}{\sqrt{N}} \sum_{l=M+1}^{\infty} d^{(k,l)} \varepsilon(t-l)$ , where  $d^{(p,l)}$  is real valued,  $|d^{(x,l)}| \leq \beta_l$  for  $x = 1$  or  $x = k$ , and  $\sum_{l=1}^{\infty} \beta_l < \infty$ , since the transfer functions  $V_1(z)$  and  $V_k(z)$  are stable. Now set  $S_{N,l}(N) = Z_{M,l}(N) + T_{M,l}(N)$ , with  $Z_{M,l}(N) = \frac{1}{\sqrt{N}} \sum_{t=1}^N w_{1(l)}^M(t)w_{k(l)}^M(t)$ , and  $T_{M,l}(N) = \frac{1}{\sqrt{N}} \sum_{t=1}^N \tilde{w}_{1(l)}^M(t)w_{k(l)}^M(t) + w_{1(l)}^M(t)\tilde{w}_{k(l)}^M(t) + \tilde{w}_{1(l)}^M(t)\tilde{w}_{k(l)}^M(t)$ . One has  $\mathbf{E}[|w_{1(l)}^M(t)w_{k(l)}^M(t)|^{2+\delta}] \leq \frac{1}{N} N^{-\delta/2} \sqrt{\mathbf{E}|w_{1(l)}^M(t)|^{4+2\delta} \mathbf{E}|w_{k(l)}^M(t)|^{4+2\delta}} \leq \frac{1}{N} N^{-\delta/2} C$ , where  $C$  is a constant. The terms  $Z_{M,l}$  are zero mean and M-dependent in the sense of [8]. One has  $\lim_{N \rightarrow \infty} \sup \sum_{t=1}^N \mathbf{E}|w_{1(l)}^M(t)w_{k(l)}^M(t)|^2 < \infty$  and  $\lim_{N \rightarrow \infty} \sum_{t=1}^N \mathbf{E}|w_{1(l)}^M(t)w_{k(l)}^M(t)|^{2+\delta} = 0$ . Therefore from lemma 9.A1 of [7], derived from [8], one obtains  $Z_{M,l}(N) \rightarrow A_s \mathcal{N}(0, Q)$  with  $Q = \lim_{N \rightarrow \infty} \mathbf{E}[Z_{M,l}(N)Z_{M,l}(N)^T]$ . On the other hand, one has  $\mathbf{E}|T_{M,l}(N)|^2 \leq C [\sum_{k=M+1}^{\infty} \beta_k]^2$ , and one gets  $\lim_{M \rightarrow \infty} \mathbf{E}|T_{M,l}(N)|^2 = 0$ . From lemma 9.A.2 of [7] derived from [4], and [1], it follows that the asymptotic distribution of  $\sqrt{N}r_{ww}^{(N)}(k)$  is the same as the distribution of  $Z_{M,l}(N)$ , and is therefore Gaussian. Since  $r_{ww}^{(N)}(k)$  is the sum of the terms  $r_{ww}^{(N)}(k)$  for  $l = [1, n_p]$ , one concludes that  $r_{ww}^{(N)}(k)$  converges in distribution towards a Gaussian variable as N tends to infinity.

## References

- [1] T.W. Anderson, "On asymptotic distributions of estimated parameters of stochastic difference equations", Ann. Math. Stat. vol. 30, pp. 676-687, 1959.
- [2] G.N. Boshnakov, "Bartlett's formulae-closed form and recurrent equations", Annals of the institute of statistical mathematics, vol. 48(1), pp. 49-59, 1996.

- [4] H. Boursès, *Linear systems*, ISTE-Wiley, 2010.
- [3] G. Box, D. Pierce, "Distribution of residual correlations in autoregressive-integrated moving average time series models", Journal of the american statistical association, vol(65), p.1509-1526,1970.
- [4] P.H. Diananda, "Some probability limit theorems with statistical applications", Proceedings of Cambridge Philos. Society, vol(49), pp.239-246, 1953.
- [5] P.S.C. Heuberger, P.M.J Van den Hof, O.H. Bosgra, "A generalized orthonormal basis for linear dynamical systems", IEEE Transaction on Automatic control vol. 40(3), pp.451-465, 1995.
- [6] P.S.C. Heuberger, P.M.J Van den Hof, B. Wahlberg, *Modelling and identification with rational orthogonal basis functions*, Springer Verlag, 2005.
- [7] L. Ljung, *System identification, theory for the user (second edition)*, Upper Saddle River, Prentice Hall, 1999.
- [8] S. Orey, "A central limit theorem for m-dependent random variables", Duke Math. journal, vol. 25, pp. 543-546, 1958.
- [9] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey", Automatica vol. 50(3), March 2014.
- [10] F. Shipp, L. Gianone, J. Bokor, Z. Szabó, "Identification in generalized orthogonal basis- a frequency domain approach", preprints of the 13th IFAC World congress, vol(1), p.387-392, San Francisco, CA, 1996, Elsevier.
- [11] T. Söderström, P. Stoica, *System identification*, Prentice Hall, 1989.
- [12] B.Vau, H. Boursès, "Closed-loop error identification algorithms with predictors based on generalized orthonormal transfer functions: Convergence conditions and bias distribution", Automatica vol. 125, March 2021. <https://doi.org/10.1016/j.automat.2020.109377>.