

TFARMA MODELS: ORDER ESTIMATION AND STABILIZATION

Michael Jachan, Gerald Matz*, and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
 Gusshausstrasse 25/389, A-1040 Vienna, Austria
 phone: +43 1 58801 38914, fax: +43 1 58801 38999, email: michael.jachan@tuwien.ac.at

*On leave with Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Electricité
 3 Rue Joliot-Curie, F-91190 Gif-sur-Yvette, France

ABSTRACT

The *time-frequency ARMA (TFARMA) model* is introduced as a time-varying ARMA model for nonstationary random processes that is formulated in terms of time shifts and frequency (Doppler) shifts. We present Akaike and minimum description length information criteria for the practically important task of selecting the TFARMA model orders. Because the estimated inverse filter used by the resulting order selection procedures is not guaranteed to be stable, we propose an iterative stabilization algorithm that is based on the concepts of instantaneous roots and root reflection/shrinkage. The performance of the proposed order selection and stabilization techniques is assessed through simulation.

1. INTRODUCTION

Time-varying AR, MA, and ARMA models are powerful tools for a wide range of signal processing applications involving nonstationary random processes (e.g. [1]). In [2, 3], we introduced the *time-frequency AR (TFAR)* and *time-frequency MA (TFMA)* models whose new feature is their formulation in terms of time-frequency (TF) shifts, i.e., time delays and Doppler frequency shifts. We also proposed efficient estimators for the TFAR and TFMA parameters and demonstrated the advantage of TFAR and TFMA models and the corresponding parameter estimators over previously proposed time-varying AR and MA models and estimators [2, 3].

Here, we discuss the practically important task of estimating the TFAR and TFMA model orders. For efficiency of exposition, we consider the general framework of *TFARMA models*, which are obtained by straightforward combination of the TFAR and TFMA models. A problem in the context of order estimation is the stabilization of the estimated inverse innovations filter. The main contributions of this paper are the adaptation of two information criteria for model order selection to the novel TFARMA models and the formulation of an iterative stabilization algorithm.

The paper is organized as follows. The TFARMA model is introduced in Section 2. In Section 3, we present the Akaike information criterion (AIC) and the minimum description length (MDL) criterion for TFARMA models. The concept of λ -hyperstability and an algorithm for stabilization are proposed in Section 4. Finally, simulation results are presented in Section 5.

2. TFARMA MODELS

A nonstationary random process $x[n]$ corresponding to a TFARMA model is defined by the equation

$$(\mathbb{A}x)[n] = (\mathbb{B}e)[n], \quad n = 0, \dots, N-1, \quad (1)$$

where $e[n]$ is the innovations noise (assumed zero-mean, white, and stationary) and \mathbb{A} and \mathbb{B} denote the TFAR and TFMA operators, respectively. These operators are linear, time-varying systems defined as [2, 3]

$$\mathbb{A} \triangleq \sum_{m=0}^{M_A} \sum_{l=-L_A}^{L_A} a_{m,l} \mathbb{S}_{m,l}, \quad \mathbb{B} \triangleq \sum_{m=0}^{M_B} \sum_{l=-L_B}^{L_B} b_{m,l} \mathbb{S}_{m,l}, \quad (2)$$

Funding by FWF grants P15156 and J2302.

where $a_{m,l}$ and $b_{m,l}$ denote the (complex) TFAR and TFMA parameters, respectively, and $\mathbb{S}_{m,l}$ is the TF shift operator defined by $(\mathbb{S}_{m,l}x)[n] = e^{j\frac{2\pi}{N}ln}x[n-m]$. Assuming a monic TFAR operator (i.e., $a_{0,l} = \delta[l]$), the TFARMA process $x[n]$ can be rewritten as

$$x[n] = - \sum_{m=1}^{M_A} \sum_{l=-L_A}^{L_A} a_{m,l} e^{j\frac{2\pi}{N}ln} x[n-m] + \sum_{m=0}^{M_B} \sum_{l=-L_B}^{L_B} b_{m,l} e^{j\frac{2\pi}{N}ln} e[n-m]. \quad (3)$$

Formally, we can also write $x[n] = (\mathbb{H}e)[n]$ with the time-varying innovations filter $\mathbb{H} = \mathbb{A}^{-1}\mathbb{B}$. Without loss of generality, we assume that $b_0[n] \triangleq \sum_{l=-L_B}^{L_B} b_{0,l} e^{j\frac{2\pi}{N}nl}$ is positive (i.e., $b_{0,l}$ is a correlation sequence). The TFAR and TFMA models are special cases obtained for $b_{m,l} = b_{0,l}\delta[m]$ and $a_{m,l} = \delta[m]\delta[l]$, respectively.

Order-recursive parameter estimators for TFARMA models can be obtained by suitable combination of the TFAR and TFMA parameter estimators described in [2, 3]; they will be presented in a future paper [4]. Given an observed signal (process realization) $x[n]$, these parameter estimators yield estimates $\hat{a}_{m,l}$ and $\hat{b}_{m,l}$ of the TFARMA parameters in (2) or (3) for prescribed model orders M_A, L_A, M_B, L_B . This leaves the model orders to be estimated.

3. TFARMA ORDER ESTIMATION

Estimation of the TFARMA model orders M_A, L_A, M_B, L_B in the Gaussian case can be accomplished by minimizing one of the “information criteria” (IC) proposed e.g. in [5–7]. Here, we consider the adaptation of AIC and MDL to the case of TFARMA models. Without loss of generality, we assume that $b_{0,0} = 1$.

For convenience, we combine all model orders into the *model order vector* $\mathbf{s} \triangleq [M_A \ L_A \ M_B \ L_B]^T$. The number $N_{\mathbb{H}}$ of complex TFARMA parameters $a_{m,l}, b_{m,l}$ depends on \mathbf{s} according to

$$N_{\mathbb{H}}(\mathbf{s}) = M_A(2L_A+1) + (M_B+1)(2L_B+1) - 1$$

(recall that $a_{0,l} = \delta[l]$ and $b_{0,0} = 1$ are known). Usually $N_{\mathbb{H}}(\mathbf{s}) \ll N$. We also arrange the real and imaginary parts of all $a_{m,l}, b_{m,l}$ into the real *TFARMA parameter vector* $\boldsymbol{\theta}$ of length $2N_{\mathbb{H}}(\mathbf{s})$.

3.1. Fundamentals

Most order estimation procedures use inverse filtering to obtain an estimate $\hat{e}[n]$ of the innovations noise $e[n]$ in (1) or (3), i.e., $\hat{e}[n] \triangleq (\hat{\mathbb{H}}^{-1}x)[n]$ with $\hat{\mathbb{H}} = \hat{\mathbb{A}}^{-1}\hat{\mathbb{B}}$. Here, $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ are given by (2) with $a_{m,l}$ and $b_{m,l}$ replaced by estimates $\hat{a}_{m,l}$ and $\hat{b}_{m,l}$, respectively. This inverse filtering requires that $\hat{\mathbb{H}}^{-1}$ is stable (cf. Section 4). We model the random variables $\hat{e}[n]$ as i.i.d., zero-mean, circular complex Gaussian with variance $v \triangleq \text{var}\{\hat{e}[n]\}$. Thus, the probability density function (pdf) of a single sample of $\hat{e}[n]$ is

$$p(\hat{\varepsilon}[n]; v(\mathbf{s})) = \frac{1}{\pi v(\mathbf{s})} \exp\left(-\frac{|\hat{\varepsilon}[n]|^2}{v(\mathbf{s})}\right), \quad (4)$$

where the notation $v(\mathbf{s})$ indicates the dependence of the variance v on the model order vector \mathbf{s} (due to the dependence of $\hat{\mathbb{H}}$ on \mathbf{s}).

The IC build upon the estimated variance of $\hat{\varepsilon}[n]$,

$$\hat{v} \triangleq \frac{1}{N} \sum_{n=0}^{N-1} |\hat{\varepsilon}[n]|^2, \quad (5)$$

which also depends on \mathbf{s} , i.e., $\hat{v} = \hat{v}(\mathbf{s})$. They include a penalty term that increases with the number of parameters $N_{\mathbb{H}}(\mathbf{s})$ to prevent overfitting. The search set to be used for determining \mathbf{s} is

$$\mathcal{S} = \{0, \dots, M_A^{\max}\} \times \{0, \dots, L_A^{\max}\} \times \{0, \dots, M_B^{\max}\} \times \{0, \dots, L_B^{\max}\},$$

where M_A^{\max} etc. are suitably chosen maximum model orders. All models with $\mathbf{s} \in \mathcal{S}$ can be estimated in an order-recursive manner using the methods described in [2–4], yielding an estimated TFARMA parameter vector $\hat{\boldsymbol{\theta}}$ for each order vector \mathbf{s} .

3.2. The AIC

The AIC was originally derived by approximative minimization of the Kullback-Leibler distance between the pdf of the data and the pdf of the model [5, 7]. It is given by twice the negative log-likelihood of the i.i.d. random vector $[\hat{\varepsilon}[0] \cdots \hat{\varepsilon}[N-1]]^T$, with the innovations variance $v(\mathbf{s})$ replaced by its estimate $\hat{v}(\mathbf{s})$ in (5), plus twice the number $2N_{\mathbb{H}}(\mathbf{s})$ of real-valued TFARMA parameters, i.e., $-2 \sum_{n=0}^{N-1} \log p(\hat{\varepsilon}[n]; \hat{v}(\mathbf{s})) + 4N_{\mathbb{H}}(\mathbf{s})$. Inserting (4) and suppressing irrelevant constant terms, we obtain

$$\text{AIC}(\mathbf{s}) = \log \hat{v}(\mathbf{s}) + 2 \frac{N_{\mathbb{H}}(\mathbf{s})}{N}.$$

The model orders are estimated by minimizing $\text{AIC}(\mathbf{s})$ over \mathcal{S} ,

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{S}} \text{AIC}(\mathbf{s}).$$

The AIC is biased (even asymptotically for $N \rightarrow \infty$), and it tends to overfit (i.e., produce order estimates that are too large) [7].

3.3. The MDL

The MDL criterion was introduced in [6] for dynamic innovations models and time-invariant ARMA models with real-valued parameters. The underlying principle is to find the model for which the description length required to represent the observation $x[n]$ using the information contained in the estimated model parameters $\hat{a}_{m,l}$, $\hat{b}_{m,l}$, and $\hat{\varepsilon}[n]$ is shortest. For TFARMA models with complex parameters, the MDL is shown in Appendix A to be given by

$$\text{MDL}(\mathbf{s}) = \log \hat{v}(\mathbf{s}) + [\log(N+1) + \rho] \frac{N_{\mathbb{H}}(\mathbf{s}) + 1/2}{N}, \quad (6)$$

with $\rho = 1 - \log 12 \approx -1.5$ (the ρ term was suppressed in [6] but is included here because we observed it to improve the MDL's performance). The model order estimates are again obtained by minimizing $\text{MDL}(\mathbf{s})$ over \mathcal{S} .

4. STABILIZATION OF TFARMA MODELS

The inverse of the estimated TFAR operator $\hat{\mathbb{A}}$ occurs in the estimated innovations filter $\hat{\mathbb{H}} = \hat{\mathbb{A}}^{-1} \hat{\mathbb{B}}$. Similarly, the inverse of the estimated TFMA operator $\hat{\mathbb{B}}$ occurs in the inverse estimated innovations filter $\hat{\mathbb{H}}^{-1} = \hat{\mathbb{B}}^{-1} \hat{\mathbb{A}}$ that is required for the inverse filtering $\hat{\varepsilon}[n] = (\hat{\mathbb{H}}^{-1} x)[n]$. To obtain reasonable modeling results, it is essential that the inverses $\hat{\mathbb{A}}^{-1}$ and $\hat{\mathbb{B}}^{-1}$ exist and are *stable* in an

appropriate sense. In this section, we discuss a stability concept based on *instantaneous roots* and a stabilization procedure that has produced satisfactory results in our simulations. We will use the generic symbols \mathbb{C} , $c_{m,l}$, M , L to denote either \mathbb{A} , $a_{m,l}$, M_A , L_A (TFAR part) or \mathbb{B} , $b_{m,l}$, M_B , L_B (TFMA part).

4.1. Instantaneous Roots

In the time-invariant case, stability of the inverse system means that the original system is minimum-phase, i.e., the roots (zeros) of the transfer function of the original system in the complex frequency plane (z -plane) are within the unit circle. For a time-varying system occurring in the TFARMA context, $\mathbb{C} = \sum_{m=0}^M \sum_{l=-L}^L c_{m,l} \mathbb{S}_{m,l}$, a *time-varying transfer function* can be defined as

$$H_{\mathbb{C}}(n, z) \triangleq \sum_{m=0}^M c_m[n] z^{-m},$$

with the time-varying impulse response

$$c_m[n] \triangleq \sum_{l=-L}^L c_{m,l} e^{j \frac{2\pi}{N} nl}, \quad m = 0, \dots, M.$$

For later use, we define the *monic version* of $c_m[n]$ as

$$d_m[n] \triangleq \frac{c_m[n]}{c_0[n]}.$$

We will also use $H_{\mathbb{C}}(n, z)$ evaluated on the unit circle $z = e^{j \frac{2\pi}{N} k}$,

$$H_{\mathbb{C}}[n, k] \triangleq H_{\mathbb{C}}(n, e^{j \frac{2\pi}{N} k}) = \sum_{m=0}^M \sum_{l=-L}^L c_{m,l} e^{-j \frac{2\pi}{N} (km - nl)},$$

where $k \in [-N/2, N/2 - 1]$ is a frequency index.

The *instantaneous roots* (or *frozen roots*) $r_m[n]$ are defined as the roots of $H_{\mathbb{C}}(n, z)$ for the respective time instant n [8–10]. It follows that $H_{\mathbb{C}}(n, z)$ can be factored in terms of the $r_m[n]$ as

$$H_{\mathbb{C}}(n, z) = c_0[n] \prod_{m=1}^M (1 - r_m[n] z^{-1}).$$

There exists a nonlinear one-to-one mapping between the (monic) time-varying impulse response $\{d_m[n]\}_{m=0, \dots, M}$ and the (unordered) instantaneous roots $\{r_m[n]\}_{m=1, \dots, M}$ at the same time n .

4.2. λ -Hyperstability

Similar to the time-invariant case, it is undesirable that the instantaneous roots $r_m[n]$ have magnitude close to or larger than one. In particular, in the underspread case [11] the TF transfer functions of a TFARMA model and of its inverse can be approximated as

$$H_{\mathbb{H}}[n, k] \approx \frac{H_{\mathbb{B}}[n, k]}{H_{\mathbb{A}}[n, k]}, \quad H_{\mathbb{H}^{-1}}[n, k] \approx \frac{H_{\mathbb{A}}[n, k]}{H_{\mathbb{B}}[n, k]}.$$

Consider e.g. the first relation, in which $H_{\mathbb{A}}[n, k] = H_{\mathbb{A}}(n, e^{j \frac{2\pi}{N} k})$ is the denominator. If some instantaneous roots of \mathbb{A} fall on or outside the unit circle during a certain time interval, this will cause excessive values of $H_{\mathbb{H}}[n, k]$ during that time interval, which can be considered a temporary instability. Conversely, if during a certain time interval $H_{\mathbb{H}}[n, k]$ is very large at a certain frequency k_0 , it can be expected that during that time interval at least one of the instantaneous roots of \mathbb{A} lies close to the unit circle about $z = e^{j \frac{2\pi}{N} k_0}$.

Motivated by these reflections, we will call a time-varying system \mathbb{C} *λ -hyperstable* (cf. [9]) if all its instantaneous roots have magnitudes less than λ for all n , i.e.,

$$|r_m[n]| < \lambda \quad \text{for all } m, n.$$

Typically, λ is chosen slightly below one.

4.3. Model Stabilization

We next present an iterative procedure for stabilizing a TF model \mathbb{C} via nonlinear modification of the estimated parameters $\hat{c}_{m,l}$. This procedure uses *root reflection* as known from the time-invariant case [8] as well as the fact that slight modifications of the instantaneous roots do not drastically change the parameters $\hat{c}_{m,l}$ (cf. the continuous dependence theorem [12]). The procedure is initialized by $r_m^{(0)}[n] \triangleq \hat{r}_m[n]$, i.e., the estimated instantaneous roots derived from $\hat{d}_m[n] = \hat{c}_m[n]/\hat{c}_0[n] = \sum_{l=-L}^L \hat{c}_{m,l} e^{j\frac{2\pi}{N}nl} / \sum_{l=-L}^L \hat{c}_{0,l} e^{j\frac{2\pi}{N}nl}$.

At the i th iteration, which is only performed if the model was not λ -hyperstable at the end of the $(i-1)$ th iteration, the instantaneous roots $r_m^{(i-1)}[n]$ obtained at the $(i-1)$ th iteration are ‘‘compressed.’’ First, all instantaneous roots with magnitude larger than one are reflected on the unit circle:

$$\rho_m^{(i)}[n] \triangleq \begin{cases} r_m^{(i-1)}[n], & |r_m^{(i-1)}[n]| \leq 1, \\ \frac{1}{(r_m^{(i-1)}[n])^*}, & |r_m^{(i-1)}[n]| > 1. \end{cases}$$

This root reflection leaves the magnitude of $H_{\mathbb{C}}(n, z)$ or $H_{\mathbb{C}}[n, k]$ unchanged. Next, a ‘‘root shrinkage’’ is performed according to

$$r_m^{(i)'}[n] \triangleq \begin{cases} \lambda \rho_m^{(i)}[n], & |\rho_m^{(i)}[n]| < \lambda, \\ \lambda^2 \rho_m^{(i)}[n], & \lambda \leq |\rho_m^{(i)}[n]| \leq 1. \end{cases}$$

We use the factor λ^2 rather than λ in the second case because we observed that this modification results in faster convergence.

Let $d_m^{(i)'}[n]$ denote the monic time-varying impulse response corresponding to the modified instantaneous roots $r_m^{(i)'}[n]$. Furthermore let $c_{m,l}^{(i)'} = \frac{1}{N} \sum_{n=0}^{N-1} d_m^{(i)'}[n] \hat{c}_0[n] e^{-j\frac{2\pi}{N}ln}$ be the associated model parameters. To enforce the given Doppler order L , we set

$$c_{m,l}^{(i)} \triangleq \begin{cases} c_{m,l}^{(i)'}, & |l| \leq L, \\ 0, & |l| > L. \end{cases}$$

This corresponds to the following modified (lowpass-filtered) impulse response:

$$c_m^{(i)}[n] \triangleq \sum_{l=-L}^L c_{m,l}^{(i)} e^{j\frac{2\pi}{N}nl} = \sum_{l=-L}^L c_{m,l}^{(i)'} e^{j\frac{2\pi}{N}nl}.$$

Finally, we calculate the instantaneous roots $r_m^{(i)}[n]$ associated with $d_m^{(i)}[n] = c_m^{(i)}[n]/\hat{c}_0[n]$ and check if the λ -hyperstability property is satisfied: If $|r_m^{(i)}[n]| < \lambda$ for all m and n , the stabilization procedure is terminated, otherwise the next $((i+1)$ th) iteration is performed. An example of an iteration is illustrated in Fig. 1.

Typically, a few iterations suffice to achieve λ -hyperstability. In the example of Fig. 1, only two iterations were required. Our simulations showed that reasonable results are obtained as long as the initial root magnitudes $|\hat{r}_m[n]|$ are not larger than about 1.2.

5. SIMULATION RESULTS

In our first experiment, we generated 100 realizations of a synthetic TFAR process with orders $M_A = 4$ and $L_A = 2$. From each realization, the parameters of all TFAR models up to orders $M_A^{\max} = 2M_A$, $L_A^{\max} = 2L_A$ were estimated using the estimator proposed in [2]. These parameter estimates were then stabilized as described in Section 4.3. After inverse filtering of the process realizations using all stabilized estimated models, the model orders were estimated from each filtered realization using the AIC and MDL criteria. Fig. 2 shows marginal histograms of the estimated model orders \hat{M}_A and \hat{L}_A . It is seen that both order estimation criteria are quite reliable. The MDL outperforms the AIC in terms of correctly estimated model orders, and the AIC tends to overfit the model orders.

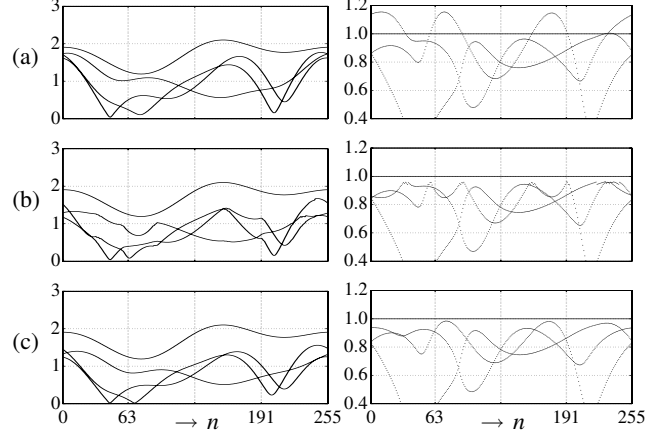


Figure 1: Illustration of the first stabilization iteration ($\lambda = 0.98$) for a model of orders $M = 3$, $L = 2$ with $|\hat{r}_m[n]|_{\max} = 1.1545$: (a) initial time-varying impulse response $\hat{c}_m[n]$, $m = 0, \dots, 3$ (left) and corresponding instantaneous roots $\hat{r}_m[n]$ (right), (b) impulse response and roots after reflection and shrinkage, (c) impulse response and roots after lowpass filtering. At the end of this first iteration, λ -hyperstability is nearly achieved ($|r_m^{(1)}[n]|_{\max} = 0.984$).

We carried out similar experiments for TFARMA processes with $M_A \in \{2, \dots, 5\}$, $M_B = M_A - 1$, and $L_A = L_B = 2$, as well as for TFARMA processes with $M_A = 2$, $M_B = M_A - 1 = 1$, and $L_A = L_B \in \{0, \dots, 3\}$. The percentage of correctly estimated model orders $M_A = M_B + 1$ for the first case and correctly estimated model orders $L_A = L_B$ for the second case is shown in Fig. 3. The results of both methods degrade with increasing model order and tend to be poorer than for the TFAR model considered previously. It is also seen that the AIC is better than the MDL for estimating the M order whereas the MDL is better than the AIC for estimating the L order. Both estimators are perfect in the limiting case of stationary processes ($L = 0$). Our simulations also showed that the MDL tends to underfit the M order (but not the L order). For TFMA models (not shown here), our simulations produced results that were similar to those obtained for TFARMA models.

6. CONCLUSION

We proposed estimators for the delay and Doppler orders of time-frequency ARMA (TFARMA) models, including TFAR and TFMA models as special cases. The classical AIC and MDL information criteria were adapted to TFARMA models, and a novel technique for stabilization of the inverse time-varying innovations filter was developed. Our simulations showed that MDL-based order estimation outperforms AIC-based order estimation in most cases.

ACKNOWLEDGMENT

The authors would like to thank Prof. B. M. Pötscher for helpful discussions and suggestions.

A. DERIVATION OF THE MDL

Our derivation of the MDL in (6) is largely analogous to that of [6]. We nevertheless include a sketch of this derivation because there are some subtle differences, mostly due to the fact that the model parameters and innovations process are complex.

The idea is to code (describe) \mathbf{s} as well as quantized versions of $\hat{\varepsilon}[n]$, $v = \text{var}\{\hat{\varepsilon}[n]\}$, and $\hat{\theta}_i = (\hat{\theta})_i$, and choose the model orders with minimal overall description length (DL). The quantization accuracies are $\pm\delta_e/2$, $\pm\delta_v/2$, and $\pm\delta_i/2$ ($i = 1, \dots, 2N_{\text{H}}$), respectively.

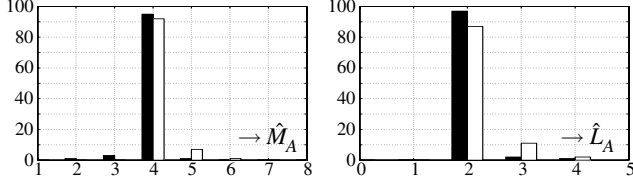


Figure 2: Histograms of order estimates \hat{M}_A and \hat{L}_A for a TFAR process with $M_A = 4$, $L_A = 2$ (white bins: AIC, black bins: MDL).

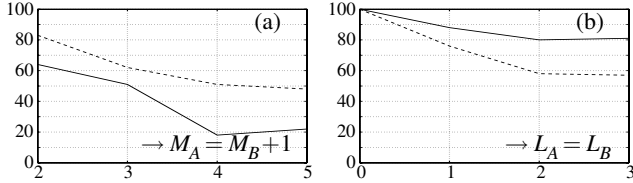


Figure 3: (a) Percentage of correctly estimated model order $M_A = M_B + 1$ for TFARMA processes with $M_A \in \{2, \dots, 5\}$, $M_B = M_A - 1$, $L_A = L_B = 2$; (b) percentage of correctly estimated model order $L_A = L_B$ for TFARMA processes with $M_A = 2$, $M_B = 1$, $L_A = L_B \in \{0, \dots, 3\}$ (dashed lines: AIC, solid lines: MDL).

The probability mass function of a single quantized sample $\hat{e}[n]$ is approximately given by $\delta_e^2 p(\hat{e}_R[n]) p(\hat{e}_I[n])$. Here, $\hat{e}_R[n]$ and $\hat{e}_I[n]$ are respectively the real and imaginary part of $\hat{e}[n]$, which are independent due to circular symmetry and distributed as $\mathcal{N}(0, v/2)$, i.e., $p(\hat{e}_{R/I}[n]) = (1/\sqrt{\pi v}) \exp(-\hat{e}_{R/I}^2[n]/v)$. The average DL (in nats) for $\hat{e}[n]$ using separate coding of $\hat{e}_R[n]$ and $\hat{e}_I[n]$ is given by

$$L(\hat{e}[n]) = -\log(\delta_e^2 p(\hat{e}_R[n]) p(\hat{e}_I[n])) = \log \frac{\pi v}{\delta_e^2} + \frac{|\hat{e}[n]|^2}{v}.$$

The average DL for the i.i.d. vector $\hat{e} = [\hat{e}[0] \dots \hat{e}[N-1]]^T$ follows as

$$L(\hat{e}) = \sum_{n=0}^{N-1} L(\hat{e}[n]) = N \log \frac{\pi v}{\delta_e^2} + \frac{1}{v} \sum_{n=0}^{N-1} |\hat{e}[n]|^2.$$

The DLs for the quantized versions of v and $\hat{\theta}$ are respectively given by (the signs of the parameters $\hat{\theta}_i$ are not coded [6])

$$L(v) = \log \frac{v}{\delta_v}, \quad L(\hat{\theta}) = \sum_{i=1}^{2N_{\text{H}}} \log \frac{|\hat{\theta}_i|}{\delta_i}.$$

The total DL needed for coding \hat{e} , v , $\hat{\theta}$, and s is thus obtained as

$$\begin{aligned} L(\hat{e}, v, \hat{\theta}, s) &= L(\hat{e}) + L(v) + L(\hat{\theta}) + L(s) \\ &= N \log \frac{\pi v}{\delta_e^2} + \frac{1}{v} \sum_{n=0}^{N-1} |\hat{e}[n]|^2 + \log \frac{v}{\delta_v} + \sum_{i=1}^{2N_{\text{H}}} \log \frac{|\hat{\theta}_i|}{\delta_i} + L(s), \end{aligned}$$

where $L(s) = \log(M_A + 1) + \log(L_A + 1) + \log(M_B + 1) + \log(L_B + 1)$. Minimizing $L(\hat{e}, v, \hat{\theta}, s)$ with respect to v yields the estimate

$$\hat{v}_{\text{MDL}} = \frac{1}{N+1} \sum_{n=0}^{N-1} |\hat{e}[n]|^2 = \frac{N}{N+1} \hat{v}, \quad (7)$$

with \hat{v} from (5). Substituting \hat{v}_{MDL} for v in $L(\hat{e}, v, \hat{\theta}, s)$ yields

$$L(\hat{e}, \hat{v}_{\text{MDL}}, \hat{\theta}, s) = N \log \frac{\pi}{\delta_e^2} + (N+1)(1 + \log \hat{v}_{\text{MDL}})$$

$$+ \log \frac{1}{\delta_v} + \sum_{i=1}^{2N_{\text{H}}} \log \frac{|\hat{\theta}_i|}{\delta_i} + L(s).$$

Next, we find optimal quantization levels δ_v and δ_i . Following [6], the dependence of $L(\hat{e}, \hat{v}_{\text{MDL}}, \hat{\theta}, s)$ on \hat{v}_{MDL} and $\hat{\theta}$ is approximated by a second-order Taylor series about the nonquantized versions of \hat{v}_{MDL} and $\hat{\theta}$. The corresponding quantization errors (differences between the quantized and nonquantized versions) are modeled as random variables that are independent and uniformly distributed within the respective elementary quantization intervals. Taking the expectation of the Taylor series approximation of $L(\hat{e}, \hat{v}_{\text{MDL}}, \hat{\theta}, s)$ with respect to the quantization errors and minimizing it with respect to δ_v and δ_i yields the optimal values (cf. [6])

$$\hat{\delta}_v = \hat{v}_{\text{MDL}} \sqrt{\frac{12}{N+1}}, \quad \hat{\delta}_i = \sqrt{\frac{12}{(N+1) \frac{\partial^2 \log \hat{v}_{\text{MDL}}}{\partial \hat{\theta}_i^2}}}.$$

Substituting $\hat{\delta}_v$ for δ_v and the $\hat{\delta}_i$'s for the δ_i 's in the expectation of the Taylor series approximation for $L(\hat{e}, \hat{v}_{\text{MDL}}, \hat{\theta}, s)$ yields

$$\begin{aligned} N \log \frac{\pi}{\delta_e^2} + N + 1 + N \log \hat{v}_{\text{MDL}} + \frac{(2N_{\text{H}} + 1)[\log(N+1) + \rho]}{2} \\ + \frac{1}{2} \sum_{i=1}^{2N_{\text{H}}} \log \left(\hat{\theta}_i^2 \frac{\partial^2 \log \hat{v}_{\text{MDL}}}{\partial \hat{\theta}_i^2} \right) + L(s), \end{aligned}$$

with $\rho = 1 - \log 12$. Dividing by N , suppressing terms that do not depend on s , neglecting the last two terms [6], and replacing \hat{v}_{MDL} by \hat{v} (cf. (7)) finally yields the MDL (6).

REFERENCES

- [1] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 899–911, Aug. 1983.
- [2] M. Jachan, G. Matz, and F. Hlawatsch, "Time-frequency-autoregressive random processes: Modeling and fast parameter estimation," in *Proc. IEEE ICASSP-2003*, vol. VI, (Hong Kong), pp. 125–128, April 2003.
- [3] M. Jachan, G. Matz, and F. Hlawatsch, "Time-frequency-moving-average random processes: Principles and cepstral methods for parameter estimation," in *Proc. IEEE ICASSP-2004*, vol. II, (Montreal, Canada), pp. 757–760, May 2004.
- [4] M. Jachan, G. Matz, and F. Hlawatsch, "Time-frequency ARMA processes: Model formulation and parameter estimation," in preparation.
- [5] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Contr.*, vol. 19, pp. 716–723, Dec. 1974.
- [6] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 6, pp. 465–471, 1978.
- [7] B. S. Choi, *ARMA Model Identification*. Springer, 1992.
- [8] Y. Grenier and M.-C. Omnes-Chevalier, "Autoregressive models with time-dependent log area ratios," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1602–1612, Oct. 1988.
- [9] M. Juntunen, J. Tervo, and J. P. Kaipio, "Stabilization of Subba Rao-Liporace models," *Circuits, Systems, Signal Processing*, vol. 18, pp. 395–406, 1999.
- [10] S. Mukhopadhyay and P. Sircar, "Parametric modeling of nonstationary signals: A unified approach," *Signal Processing*, vol. 60, pp. 135–152, 1997.
- [11] G. Matz and F. Hlawatsch, "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory," *J. Math. Phys., Special Issue on Wavelet and Time-Frequency Analysis*, vol. 39, pp. 4041–4071, Aug. 1998.
- [12] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge (UK): Cambridge Univ. Press, 1999.