

## Research Article

# Stability Analysis of a Variant of the Prony Method

**Rodney Jaramillo and Marianela Lentini**

*Escuela de Matemáticas, Universidad Nacional de Colombia, Sede Medellín, Medellín, Colombia*

Correspondence should be addressed to Rodney Jaramillo, [rjjustic@unal.edu.co](mailto:rjjustic@unal.edu.co)

Received 15 June 2012; Accepted 12 October 2012

Academic Editor: Bin Liu

Copyright © 2012 R. Jaramillo and M. Lentini. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prony type methods are used in many engineering applications to determine the exponential fit corresponding to a dataset. In this paper we study a variant of Prony's method that was used by Martín-Landrove et al., in a process of segmentation of  $T_2$ -weighted MRI brain images. We show the equivalence between that method and the classical Prony method and study the stability of the computed solutions with respect to noise in the data set. In particular, we show that the relative error in the calculation of the exponential fit parameters is linear with respect to noise in the data. Our analysis is based on classical results from linear algebra, matrix computation theory, and the theory of stability for roots of polynomials.

## 1. Introduction

In this paper we consider the problem of recovering the parameters  $b$ ,  $\lambda_j$ , and  $C_j$   $j = 1, \dots, k$ , from a set of measurements  $y_i$ ,  $i = 1, \dots, n$  with  $n \geq 2k + 1$  and a given exponential model

$$\mu(t) = b + \sum_{j=1}^k C_j e^{-\lambda_j t}. \quad (1.1)$$

The measurement  $y_i$  corresponds to the value of expression (1.1) for  $t = i\Delta t$ ,  $\Delta t > 0$ . Hence we get the set of equations

$$y_i = b + \sum_{j=1}^k C_j e^{-i\lambda_j \Delta t} \quad i = 1, \dots, n. \quad (1.2)$$

This kind of exponential fit appears in the case of  $T_2$ -weighted magnetic resonance images of the brain.  $T_2$  is the transverse relaxation time in the process of magnetic resonance; it measures

for how long the transverse magnetization lasts in a uniform external magnetic field. Two parameters, TR (repetition time) and TE (echo time) characterize the acquisition of the signal;  $T_2$  weighted refers to a signal produced with long TR and TE.

The exponents  $\lambda_j$  are the nonlinear parameters and correspond to the relaxation rates associated with the different tissues present in the images. The coefficients  $b, C_1, \dots, C_k$  are the linear parameters and are related to the noise, and the proportions of those tissues in the given images. In previous works Martín-Landrove et al. [1, 2], it is considered a variant of the Prony method to recover the relaxation rates, the noise and the proportions. Paluszny et al. [3] compared that variant of the Prony method with the Variable Projection Method of Golub and Pereyra, [4–6]. The method proved to be reliable for low noise levels and  $k = 1, 2, 3$ . The Variable Projection Method proved to be more robust in the presence of larger noise levels but required ten times more computational time to get results comparable to those of the variant of the Prony method. In this paper we first show that the method proposed by Martín-Landrove is equivalent to the Prony method described by Kahn et al. in [7], and then we study the behavior of the linear systems and the conditioning of the polynomial roots that have to be computed to obtain the model parameters using the method.

## 2. Prony Method

In the model formulation (1.1) set  $b = C_0$  and  $\lambda_0 = 0$ , and write

$$\mu(t) = \sum_{j=0}^k C_j e^{-\lambda_j t}. \quad (2.1)$$

The Prony type methods, also known as polynomial methods, use the fact that  $\mu(t)$  satisfies a difference equation of the form

$$\left( \delta_{k+2} E^{k+1} + \dots + \delta_2 E + \delta_1 \right) \mu(t) = 0, \quad (2.2)$$

where  $E$  is the forward shift operator

$$(E\mu)(t) = \mu(t + \Delta t), \quad (2.3)$$

and the values  $\beta_j = e^{-\lambda_j \Delta t}$  are the roots of the polynomial

$$P(z) = \delta_{k+2} z^{k+1} + \dots + \delta_2 z + \delta_1 = 0, \quad (2.4)$$

which is called the characteristic polynomial of the associated difference equation (2.2). Evaluating (2.2) for  $t_i = i\Delta t$ ,  $i = 1, \dots, n$ , we get the following set of linear equations:

$$\begin{aligned} \delta_{k+2}\mu(t_{k+2}) + \dots + \delta_1\mu(t_1) &= 0 \\ \vdots & \\ \delta_{k+2}\mu(t_n) + \dots + \delta_1\mu(t_{n-k-1}) &= 0. \end{aligned} \quad (2.5)$$

Let  $\delta = [\delta_1, \dots, \delta_{k+2}]^T$ ,  $v = [v_1, \dots, v_n]^T = [\mu(t_1), \dots, \mu(t_n)]^T$  and

$$X_\delta = \begin{bmatrix} \delta_1 & & & & & \\ & \ddots & & & & \\ & & \ddots & & & \\ & & & \delta_1 & & \\ \delta_{k+2} & & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \delta_{k+2} & \end{bmatrix}. \quad (2.6)$$

Then the above system of linear equations can be rewritten in matrix form as

$$X_\delta^T v = 0. \quad (2.7)$$

Alternatively, if  $W_v$  is the  $(n - k - 1) \times (k + 2)$  Hankel matrix

$$W_v = \begin{bmatrix} v_1 & v_2 & \cdots & v_{k+2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ v_{n-k-1} & \cdot & \cdots & v_n \end{bmatrix} \quad (2.8)$$

then

$$W_v \delta = X_\delta^T v. \quad (2.9)$$

If there is no noise in the observation data, then  $y_i = \mu(t_i)$  and the coefficients  $\delta_i$  can be determined from the equivalent system of equations

$$X_\delta^T y = 0, \quad (2.10)$$

where  $y = [y_1, \dots, y_n]^T$ . Then the  $\beta_j$  are computed as the roots of  $P(z)$  and finally

$$\lambda_j = \frac{-\log(\beta_j)}{\Delta t} \quad j = 0, \dots, k. \quad (2.11)$$

In the presence of noisy data instead of solving the system (2.10), we consider the nonlinear optimization problem

$$\min_{\delta} y^T X_\delta X_\delta^T y. \quad (2.12)$$

In order to obtain a nontrivial solution, it is necessary to impose restrictions over the parameters  $\delta_j$ ; each choice of restrictions characterizes a particular version of the Prony method. For example the modified Prony method described in Osborne and Smyth [8, 9], uses

$$\min_{\delta} y^T X_\delta X_\delta^+ y \quad \text{subject to } \|\delta\|_2^2 = 1, \quad (2.13)$$

where  $X_\delta^+$  is the Moore Penrose generalized inverse of  $X_\delta$ . Osborne et al. considered  $X_\delta$  to be full rank, hence  $X_\delta^+ = (X_\delta^T X_\delta)^{-1} X_\delta^T$ .

We consider a different optimization problem, which appears in the literature, to compare with the Martín-Landrove method:

$$\min_{\delta} y^T X_\delta X_\delta^T y \quad \text{subject to } \delta_{k+2} = 1. \quad (2.14)$$

The above methods and others, such as classical Prony method, Pisarenko's method, and the linear predictor method, are described in [7–11]. Once the nonlinear parameters have been found, the linear ones are computed as the least squares solution of the linear system obtained by replacing the nonlinear parameters in (1.2), a separation of variables approach.

### 3. An Alternative Formulation of the Prony Method

To simplify the explanation let us consider the case  $k = 3$  and  $n \geq 7$ . Then, the system of equations (1.2) can be written as

$$\begin{aligned} y_1 &= b + C_1\beta_1 + C_2\beta_2 + C_3\beta_3 \\ y_2 &= b + C_1\beta_1^2 + C_2\beta_2^2 + C_3\beta_3^2 \\ &\vdots \\ y_n &= b + C_1\beta_1^n + C_2\beta_2^n + C_3\beta_3^n. \end{aligned} \quad (3.1)$$

From (3.1) and defining  $q_i = y_i - y_{i+1}$  we get

$$\begin{aligned} q_1 &= C_1(1 - \beta_1)\beta_1 + C_2(1 - \beta_2)\beta_2 + C_3(1 - \beta_3)\beta_3 \\ &\vdots \\ q_{n-1} &= C_1(1 - \beta_1)\beta_1^{n-1} + C_2(1 - \beta_2)\beta_2^{n-1} + C_3(1 - \beta_3)\beta_3^{n-1}. \end{aligned} \quad (3.2)$$

The dimension of the system can be reduced by using the transformation  $q_{j+1} - q_j\beta_1$ , to get

$$\begin{aligned} q_2 - q_1\beta_1 &= C_2(1 - \beta_2)\beta_2(\beta_2 - \beta_1) + C_3(1 - \beta_3)\beta_3(\beta_3 - \beta_1) \\ &\vdots \\ q_{n-1} - q_{n-2}\beta_1 &= C_2(1 - \beta_2)\beta_2^{n-2}(\beta_2 - \beta_1) + C_3(1 - \beta_3)\beta_3^{n-2}(\beta_3 - \beta_1). \end{aligned} \quad (3.3)$$

Then

$$\begin{aligned} q_2 &= q_1\beta_1 + C_2(1 - \beta_2)\beta_2(\beta_2 - \beta_1) + C_3(1 - \beta_3)\beta_3(\beta_3 - \beta_1) \\ &\vdots \\ q_{n-1} &= q_{n-2}\beta_1 + C_2(1 - \beta_2)\beta_2^{n-2}(\beta_2 - \beta_1) + C_3(1 - \beta_3)\beta_3^{n-2}(\beta_3 - \beta_1). \end{aligned} \quad (3.4)$$

We now apply the transformation  $q_{j+1} - q_j\beta_2$ , and the new set can be written as

$$\begin{aligned} q_3 - q_2\beta_2 &= q_2\beta_1 - q_1\beta_1\beta_2 + C_3(1 - \beta_3)\beta_3(\beta_3 - \beta_1)(\beta_3 - \beta_2) \\ &\vdots \\ q_{n-1} - q_{n-2}\beta_2 &= q_{n-2}\beta_1 - q_{n-3}\beta_1\beta_2 + C_3(1 - \beta_3)\beta_3^{n-3}(\beta_3 - \beta_1)(\beta_3 - \beta_2). \end{aligned} \quad (3.5)$$

The equations above are equivalent to

$$\begin{aligned} q_3 &= q_2(\beta_1 + \beta_2) - q_1\beta_1\beta_2 + C_3(1 - \beta_3)\beta_3(\beta_3 - \beta_1)(\beta_3 - \beta_2) \\ q_4 &= q_3(\beta_1 + \beta_2) - q_2\beta_1\beta_2 + C_3(1 - \beta_3)\beta_3^2(\beta_3 - \beta_1)(\beta_3 - \beta_2) \\ &\vdots \\ q_{n-1} &= q_{n-2}(\beta_1 + \beta_2) - q_{n-3}\beta_1\beta_2 + C_3(1 - \beta_3)\beta_3^{n-3}(\beta_3 - \beta_1)(\beta_3 - \beta_2). \end{aligned} \quad (3.6)$$

Finally we use the transformation  $q_{j+1} - q_j\beta_3$ . At this final step we get the following system:

$$\begin{aligned} q_4 &= q_3(\beta_1 + \beta_2 + \beta_3) - q_2(\beta_1\beta_2 + \beta_1\beta_3 + \beta_2\beta_3) + q_1\beta_1\beta_2\beta_3 \\ &\vdots \\ q_{n-1} &= q_{n-2}(\beta_1 + \beta_2 + \beta_3) - q_{n-3}(\beta_1\beta_2 + \beta_1\beta_3 + \beta_2\beta_3) + q_{n-4}\beta_1\beta_2\beta_3. \end{aligned} \quad (3.7)$$

In matrix notations we have

$$Mw = Q, \quad (3.8)$$

where  $[w_1, w_2, w_3] = [\beta_1 + \beta_2 + \beta_3, -\beta_1\beta_2 - \beta_1\beta_3 - \beta_2\beta_3, \beta_1\beta_2\beta_3]$ ,

$$M = \begin{bmatrix} q_3 & q_2 & q_1 \\ q_4 & q_3 & q_2 \\ \vdots & \vdots & \vdots \\ q_{n-2} & q_{n-3} & q_{n-4} \end{bmatrix}, \quad Q = \begin{bmatrix} q_4 \\ q_5 \\ \vdots \\ q_{n-1} \end{bmatrix}. \quad (3.9)$$

If the data values  $y_i$  are noiseless, the values  $w_1$ ,  $w_2$ , and  $w_3$  may be obtained from the previous system of equations and the  $\beta_j$  are the roots of the polynomial

$$\alpha(z) = z^3 - w_1z^2 - w_2z - w_3. \quad (3.10)$$

Once the roots have been computed the nonlinear parameters can be calculated using (2.11), and likewise the linear parameters, as stated before.

In general, for an arbitrary  $k$ , let  $M^{(k)}$  be the  $(n - k - 1) \times k$  matrix

$$M^{(k)} = \begin{bmatrix} q_k & \cdot & \cdot & q_1 \\ q_{k+1} & \cdot & \cdot & q_2 \\ \vdots & \vdots & \vdots & \\ q_{n-2} & \cdot & \cdot & q_{n-1-k} \end{bmatrix}, \quad Q^{(k)} = \begin{bmatrix} q_{k+1} \\ q_{k+2} \\ \vdots \\ q_{n-1} \end{bmatrix}, \quad w^{(k)} = \begin{bmatrix} w_1^{(k)} \\ \vdots \\ w_k^{(k)} \end{bmatrix}. \quad (3.11)$$

In this case the coefficients of the polynomial  $\alpha(z)$  are the symmetric functions of  $\beta_1, \dots, \beta_k$  defined as

$$\begin{aligned} w_1^{(k)} &= \beta_1 + \dots + \beta_k, \\ w_2^{(k)} &= -\sum_{l \neq r} \beta_l \beta_r, \\ w_3^{(k)} &= \sum_{l \neq r, l \neq s, s \neq r} \beta_l \beta_r \beta_s, \\ &\vdots \\ w_{k-1}^{(k)} &= (-1)^k \sum_{j=1}^k \left( \prod_{l \neq j} \beta_l \right), \\ w_k^{(k)} &= (-1)^{k+1} \prod_{j=1}^k \beta_j. \end{aligned} \quad (3.12)$$

These coefficients are determined by the solution of the system

$$M^{(k)} w^{(k)} = Q^{(k)}. \quad (3.13)$$

Finally, the  $\beta_j$  are the roots of the polynomial

$$\alpha^{(k)}(z) = z^k - \sum_{j=1}^k w_j^{(k)} z^{k-j}. \quad (3.14)$$

Next we will study the relationship between the solution obtained by the procedure described above and the Prony method described in Section 2.

**Theorem 3.1.** *Let  $R$  be the  $k \times k$  matrix defined as follows: for  $k = 1$  set  $R = 1$  and for  $k > 1$*

$$R(i, j) = \begin{cases} 1, & \text{if } i = j, \\ -1, & \text{if } j = i + 1, \\ 0, & \text{otherwise.} \end{cases} \quad (3.15)$$

In addition, let  $P(z)$  and  $\alpha^{(k)}(z)$  be the polynomials defined in (2.4) and (3.14), respectively. If  $\delta = [\delta_1, \dots, \delta_{k+1}, 1]$  is the solution of the optimization problem (2.14), then  $w^{(k)} = R^{-1}[\delta_k, \dots, \delta_1]^T$  satisfies

$$M^{(k)}w^{(k)} - Q^{(k)} = X_\delta^T y. \quad (3.16)$$

Moreover,

$$P(z) = (z - 1)\alpha^{(k)}(z). \quad (3.17)$$

*Proof.* The solution  $\delta = [\delta_1, \dots, \delta_{k+1}, \delta_{k+2}]$  of (2.14) satisfies  $\delta_{k+2} = 1$ . In the case we are considering  $\beta_0 = 1$  is a root of  $P(z)$ , which implies that

$$-\sum_{j=1}^{k+1} \delta_j = 1. \quad (3.18)$$

Then we have

$$\begin{aligned} M^{(k)}w^{(k)} - Q^{(k)} &= M^{(k)}R^{-1}Rw^{(k)} - Q^{(k)} \\ &= M^{(k)}R^{-1} \begin{bmatrix} \delta_k \\ \vdots \\ \delta_1 \end{bmatrix} - \begin{bmatrix} y_{k+1} - y_{k+2} \\ \vdots \\ y_{n-1} - y_n \end{bmatrix} \\ &= M^{(k)}R^{-1} \begin{bmatrix} \delta_k \\ \vdots \\ \delta_1 \end{bmatrix} - \begin{bmatrix} -\left[\sum_{j=1}^{k+1} \delta_j\right] y_{k+1} - y_{k+2} \\ \vdots \\ -\left[\sum_{j=1}^{k+1} \delta_j\right] y_{n-1} - y_n \end{bmatrix} \\ &= M^{(k)}R^{-1} \begin{bmatrix} \delta_k \\ \vdots \\ \delta_1 \end{bmatrix} + \begin{bmatrix} y_{k+1} & \cdots & y_{k+1} \\ \vdots & \vdots & \vdots \\ y_{n-1} & \cdots & y_{n-1} \end{bmatrix} \begin{bmatrix} \delta_k \\ \vdots \\ \delta_1 \end{bmatrix} \\ &\quad + \cdots \delta_{k+2} \begin{bmatrix} y_{k+2} \\ \vdots \\ y_n \end{bmatrix} + \delta_{k+1} \begin{bmatrix} y_{k+1} \\ \vdots \\ y_{n-1} \end{bmatrix} \\ &= \begin{bmatrix} y_1 & y_2 & \cdots & y_{k+2} \\ y_2 & y_3 & \cdots & y_{k+3} \\ y_3 & y_4 & \cdots & y_{k+4} \\ \cdot & \cdot & \cdots & \cdot \\ y_{n-k-1} & y_{n-k} & \cdots & y_n \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{k+1} \\ \delta_{k+2} \end{bmatrix} = W_y \delta. \end{aligned} \quad (3.19)$$

Then, using (2.9), it follows that

$$M^{(k)}w^{(k)} - Q^{(k)} = X_\delta^T y. \quad (3.20)$$

Now, for the polynomial  $P(z)$  we have

$$\begin{aligned}
P(z) &= \delta_{k+2}z^{k+1} + \dots + \delta_2z + \delta_1 \\
&= (z-1) \left( z^k - \left( \sum_{j=1}^k \delta_j \right) z^{k-1} - \left( \sum_{j=1}^{k-1} \delta_j \right) z^{k-2} - \dots - (\delta_1 + \delta_2)z - \delta_1 \right) \\
&= (z-1) \left( z^k - w_1^{(k)} z^{k-1} - w_2^{(k)} z^{k-2} - \dots - w_{k-1}^{(k)} z - w_k^{(k)} \right) \\
&= (z-1) \left( z^k - \sum_{j=1}^k w_j^{(k)} z^{k-j} \right) \\
&= (z-1) \alpha^{(k)}(z).
\end{aligned} \tag{3.21}$$

□

**Theorem 3.2.** *Let us suppose that there exists a unique solution to the optimization problem (2.14). A vector  $\delta \in \mathbb{R}^{k+2}$  is the solution of the optimization problem (2.14) if and only if  $R^{-1}[\delta_k, \dots, \delta_1]^T$  is the least squares solution of the linear equation (3.13).*

*Proof.* Let  $\delta \in \mathbb{R}^{k+2}$  be the solution of the optimization problem (2.14),  $\zeta = R^{-1}[\delta_k, \dots, \delta_1]^T$ , and let  $\psi$  be the least squares solution of the linear system (3.13). From Theorem 3.1 it follows that

$$\begin{aligned}
\|X_\delta^T \mathbf{y}\| &= \|M^{(k)} \zeta - Q^{(k)}\| \\
&\geq \min_z \|M^{(k)} z - Q^{(k)}\| \\
&= \|M^{(k)} \psi - Q^{(k)}\|.
\end{aligned} \tag{3.22}$$

Let us consider  $\xi \in \mathbb{R}^k$  given by

$$[\xi_k, \dots, \xi_1]^T = R\psi \tag{3.23}$$

and  $\gamma \in \mathbb{R}^{k+2}$  defined as

$$\gamma = \left[ \xi_1, \dots, \xi_k, -1 - \sum_{j=1}^k \xi_j, 1 \right]^T. \tag{3.24}$$

By Theorem 3.1 we have

$$\|M^{(k)} \psi - Q^{(k)}\| = \|X_\gamma^T \mathbf{y}\|, \tag{3.25}$$

and therefore  $\|X_\delta^T \mathbf{y}\| \geq \|X_\gamma^T \mathbf{y}\|$ . By hypothesis  $\delta$  is the only solution of the optimization problem, thus  $\delta = \gamma$ , which implies that  $[\delta_k, \dots, \delta_1] = [\xi_k, \dots, \xi_1]$ . Now we have  $R\zeta = R\psi$  and thus  $\zeta = \psi$ .



In a similar way we can prove that  $\gamma$  in (3.24) is the solution of the optimization problem (2.14), whenever  $\varphi$  is the least squares solution of the linear equation (3.13).  $\square$

#### 4. Stability of the Alternative Prony Method

In (3.13), arrays  $M^{(k)}$  and  $Q^{(k)}$  depend on the vector of measurements  $y$ . In the presence of noisy data the equation that really holds is

$$\left(M^{(k)} + \epsilon_1\right)w^{(k)} = Q^{(k)} + \epsilon_2, \quad (4.1)$$

where  $\epsilon_1$  and  $\epsilon_2$  depend on the level of noise. Then it is necessary to determine the condition number of the matrix in order to see the accuracy in the coefficients given by  $w^{(k)}$ . Let us consider (1.2) for  $n = 2k + 1$ . In this situation  $M^{(k)}$  is a nonsingular square matrix of dimension  $k \times k$ . In the following lemma we establish a factorization of the matrix  $M^{(k)}$  that will be used in the stability analysis of the system (3.13). Our analysis is similar to the one described in [12].

**Lemma 4.1.** *Let us consider  $n = 2k + 1$ . Let  $G$  and  $H$  be the  $k \times k$  matrices defined by*

$$g_{i,j} = \beta_j^{i-1}, \quad h_{i,j} = \beta_i^{k-j}, \quad (4.2)$$

and let  $D$  be the  $k \times k$  diagonal matrix given by  $d_{i,i} = C_i(1 - \beta_i)\beta_i$ . Then

$$M^{(k)} = GDH. \quad (4.3)$$

*Proof.* By definition  $q_i = y_i - y_{i+1}$ ,  $i = 1, \dots, 2k$ . Then

$$q_i = \sum_{j=1}^k C_j \beta_j^i (1 - \beta_j). \quad (4.4)$$

Thus

$$\begin{aligned} M^{(k)} &= \begin{bmatrix} q_k & \cdots & q_1 \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ q_{2k-1} & \cdots & q_k \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^k C_j \beta_j^k (1 - \beta_j) & \cdots & \sum_{j=1}^k C_j \beta_j (1 - \beta_j) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \sum_{j=1}^k C_j \beta_j^{2k-1} (1 - \beta_j) & \cdots & \sum_{j=1}^k C_j \beta_j^k (1 - \beta_j) \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} 1 & \cdots & 1 \\ \beta_1 & \cdots & \beta_k \\ \vdots & \cdots & \vdots \\ \beta_1^{k-1} & \cdots & \beta_k^{k-1} \end{bmatrix} \begin{bmatrix} d_{1,1} & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & d_{k,k} \end{bmatrix} \begin{bmatrix} \beta_1^{k-1} & \cdots & \beta_1 & 1 \\ \beta_2^{k-1} & \cdots & \beta_2 & 1 \\ \cdot & \cdots & \cdot & \cdot \\ \beta_k^{k-1} & \cdots & \beta_k & 1 \end{bmatrix} \\
&= GDH.
\end{aligned} \tag{4.5}$$

□

**Lemma 4.2.** *Let us consider  $n = 2k + 1$ . Let  $\eta_{M^{(k)}}$ ,  $\eta_G$  and  $\eta_D$  be the condition numbers in the infinite norm of the matrices  $M^{(k)}$ ,  $G$ , and  $D$ , respectively. Then*

$$\eta_{M^{(k)}} \leq k^2 \eta_G^2 \eta_D. \tag{4.6}$$

*Proof.* Let  $\eta_H$  be the condition number in the infinite norm of the matrix  $H$ . Using Lemma 4.1 it follows that

$$\eta_{M^{(k)}} \leq \eta_G \eta_D \eta_H. \tag{4.7}$$

To get (4.6) it is enough to observe that

$$\begin{aligned}
\eta_H &= \|H\|_\infty \|H^{-1}\|_\infty = \|H^T\|_1 \|(H^{-1})^T\|_1 \\
&= \|H^T\|_1 \|(H^T)^{-1}\|_1 = \|G\|_1 \|G^{-1}\|_1 \\
&\leq k^2 \|G\|_\infty \|G^{-1}\|_\infty = k^2 \eta_G.
\end{aligned} \tag{4.8}$$

□

The following theorem establishes an estimate for the variation  $\Delta w^{(k)}$  in the vector components  $w^{(k)}$  in (3.13), as it depends on the noise level in the vector of the measurements  $y = [y_1, \dots, y_n]^T$ .

**Theorem 4.3.** *Let  $n = 2k + 1$ ,  $y = [y_1, \dots, y_n]^T$ ,  $\Delta y = [\Delta y_1, \dots, \Delta y_n]^T$ , and*

$$\Lambda_1 = \min \left\{ \|M^{(k)}\|_\infty, \|q^{(k)}\|_\infty \right\}. \tag{4.9}$$

Let  $\delta > 0$  such that

$$2k^3 \eta_G^2 \eta_D \delta < \Lambda_1. \tag{4.10}$$

If  $\|\Delta y\|_\infty < \delta$ , then the variation  $\Delta w^{(k)}$  in the solution of the perturbed system

$$\left(M^{(k)} + \Delta M^{(k)}\right)\left(w^{(k)} + \Delta w^{(k)}\right) = q^{(k)} + \Delta q^{(k)} \quad (4.11)$$

satisfies

$$\frac{\|\Delta w^{(k)}\|_\infty}{\|w^{(k)}\|_\infty} \leq 4k^3 \eta_G^2 \eta_D \frac{\delta}{\Lambda_1} + O(\delta^2). \quad (4.12)$$

*Proof.* Let  $\epsilon = 2\delta k / \Lambda_1$ . If  $\|\Delta y\|_\infty < \delta$ , then  $\|\Delta M^{(k)}\|_\infty \leq 2k\delta$  and  $\|\Delta q^{(k)}\|_\infty \leq 2k\delta$ . Therefore

$$\|\Delta q^{(k)}\|_\infty \leq 2k\delta \frac{\|q^{(k)}\|_\infty}{\Lambda_1} = \epsilon \|q^{(k)}\|_\infty. \quad (4.13)$$

In a similar way we get

$$\|\Delta M^{(k)}\|_\infty \leq \epsilon \|M^{(k)}\|_\infty. \quad (4.14)$$

By using (4.6) we see that

$$\epsilon \eta_{M^{(k)}} \leq \epsilon k^2 \eta_G^2 \eta_D < 1. \quad (4.15)$$

From (4.13), (4.14), and (4.15), and using Theorem 2.7.2 in [13], it follows that

$$\begin{aligned} \frac{\|\Delta w^{(k)}\|_\infty}{\|w^{(k)}\|_\infty} &\leq 2 \frac{\epsilon \eta_{M^{(k)}}}{1 - \epsilon \eta_{M^{(k)}}} \\ &\leq 2\epsilon \eta_{M^{(k)}} + O\left(\epsilon^2 \eta_{M^{(k)}}^2\right) \\ &\leq 4k^3 \eta_G^2 \eta_D \frac{\delta}{\Lambda_1} + O(\delta^2). \end{aligned} \quad (4.16)$$

Once we have estimated the vector  $\tilde{w}^{(k)} = w^{(k)} + \Delta w^{(k)}$ , it remains to consider the corresponding impact in the calculation of the roots of the polynomial

$$\tilde{\alpha}^{(k)}(z) = z^k - \sum_{j=1}^k \tilde{w}_j^{(k)} z^{k-j}. \quad (4.17)$$

Let  $\beta_j$  be one of the roots of the polynomial  $\alpha^{(k)}(z)$  and suppose that  $\tilde{\beta}_j = \beta_j + \Delta\beta_j$  is the root of the polynomial  $\tilde{\alpha}^{(k)}(z)$  closest to  $\beta_j$ . The following is an estimate for  $\Delta\beta_j$  which follows from the theory of stability for simple roots of polynomials, see [14].  $\square$

**Lemma 4.4.** *Let  $0 < \theta \ll 1$  such that*

$$\frac{\|\Delta w^{(k)}\|_\infty}{\|w^{(k)}\|_\infty} < \theta. \quad (4.18)$$

Then there exists  $\Lambda_2 > 0$  such that

$$|\Delta\beta_j| \leq \Lambda_2 \frac{\|\Delta\boldsymbol{w}^{(k)}\|_\infty}{\|\boldsymbol{w}^{(k)}\|_\infty}. \quad (4.19)$$

From (4.12) we see that there is a constant  $\Lambda_3$  such that

$$\frac{\|\Delta\boldsymbol{w}^{(k)}\|_\infty}{\|\boldsymbol{w}^{(k)}\|_\infty} \leq \left( \frac{4k^3\eta_G^2\eta_D}{\Lambda_1} + \Lambda_3\delta \right) \delta. \quad (4.20)$$

Thus, if  $\delta$  in Theorem 4.3 is sufficiently small such that

$$\delta \left( \frac{4k^3\eta_G^2\eta_D}{\Lambda_1} + \Lambda_3\delta \right) = \theta < 1, \quad (4.21)$$

then the conditions of Lemma 4.4 are satisfied and therefore

$$\|\Delta\boldsymbol{\beta}\|_\infty \leq \Lambda_2 \frac{\|\Delta\boldsymbol{w}^{(k)}\|_\infty}{\|\boldsymbol{w}^{(k)}\|_\infty}. \quad (4.22)$$

With the computed  $\tilde{\beta}_j$ ,  $1 \leq j \leq k$ , we should analyze the least squares solution of the perturbed system

$$\begin{bmatrix} 1 & \tilde{\beta}_1 & \cdots & \tilde{\beta}_k \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 1 & \tilde{\beta}_1^{2k+1} & \cdots & \tilde{\beta}_k^{2k+1} \end{bmatrix} \begin{bmatrix} \tilde{b} \\ \tilde{C}_1 \\ \cdot \\ \tilde{C}_k \end{bmatrix} = \boldsymbol{y} + \Delta\boldsymbol{y}, \quad (4.23)$$

which will be written as  $\tilde{B}\tilde{C} = \tilde{\boldsymbol{y}}$ . The following theorem provides an estimate of  $\|\Delta C\|_2$ , being  $\tilde{C} = C + \Delta C$  the solution of the perturbed system. Note that the following result is the first one in this paper where we require the nonlinear parameters in the model (1.1) to be positive.

**Theorem 4.5.** *Let  $n$ ,  $\boldsymbol{y}$ ,  $\Delta\boldsymbol{y}$ , and  $\delta$  be as in the statement of Theorem 4.3, with  $\|\Delta\boldsymbol{y}\|_\infty < \delta$ . Let  $\hat{\eta}_\beta$  be the condition number of matrix  $B$  in the Euclidean norm and consider*

$$\Lambda_4 = \min\{\|B\|_2, \|\boldsymbol{y}\|_2\}. \quad (4.24)$$

Let us suppose that  $\delta$  satisfies (4.10) and (4.21) and let

$$\rho_\delta = \delta\sqrt{2k+1} \max\left\{k(2k+1) \left( \frac{4k^3\eta_G^2\eta_D}{\Lambda_1} + \Lambda_3\delta \right) \Lambda_2, 1\right\} < \min\left\{1, \frac{\Lambda_4}{\hat{\eta}_\beta}\right\}. \quad (4.25)$$

Then

$$\frac{\|\Delta C\|_2}{\|C\|_2} \leq \frac{\rho_\delta}{\Lambda_4} \left( \frac{\hat{\eta}_\beta}{\cos \sigma} + \hat{\eta}_\beta^2 \tan \sigma \right) + O(\delta^2), \quad (4.26)$$

where

$$\sin \sigma = \frac{\|y - y^*\|_2}{\|y\|_2}, \quad (4.27)$$

and  $y^*$  is the orthogonal projection of the vector  $y$  on the subspace spanned by the columns of  $B$ .

*Proof.* Let  $\epsilon = \rho_\delta / \Lambda_4$ . Using the mean value theorem we see that

$$\left| (\beta_l + \Delta\beta_l)^j - \beta_l^j \right| \leq j(\beta_l^*)^{j-1} \Delta\beta_l, \quad (4.28)$$

with  $\beta_l^*$  between  $\beta_l$  and  $\beta_l + \Delta\beta_l$  and  $0 < \beta_l < 1$ . Then

$$\begin{aligned} \|\Delta B\|_\infty &= \max_{1 \leq i \leq 2k+1} \left( \sum_{l=1}^k \left| (\beta_l + \Delta\beta_l)^j - \beta_l^j \right| \right) \\ &\leq k(2k+1) \|\Delta\beta\|_\infty. \end{aligned} \quad (4.29)$$

From (4.20) and (4.22) it follows that

$$\begin{aligned} \|\Delta B\|_2 &\leq \sqrt{2k+1} \|\Delta B\|_\infty \\ &\leq k(2k+1)^{3/2} \|\Delta\beta\|_\infty \\ &\leq k(2k+1)^{3/2} \Lambda_2 \delta \left( \frac{4k^3 \eta_G^2 \eta_D}{\Lambda_1} + \Lambda_3 \delta \right) \\ &\leq \frac{\rho_\delta}{\Lambda_4} \|B\|_2. \end{aligned} \quad (4.30)$$

On the other hand

$$\begin{aligned} \|\Delta y\|_2 &\leq \sqrt{2k+1} \|\Delta y\|_\infty \\ &\leq \sqrt{2k+1} \delta \leq \frac{\rho_\delta}{\Lambda_4} \|y\|_2. \end{aligned} \quad (4.31)$$

By definition

$$\frac{\rho_\delta}{\Lambda_4} < \frac{1}{\hat{\eta}_\beta}, \quad (4.32)$$

and (4.26) follows from (4.30), (4.31), and (4.32), by Theorem 5.3.1 in [13].  $\square$

If the value of  $\rho_\delta$  in (4.25) is given by

$$\rho_\delta = k(2k+1)^{3/2} \left( \frac{4k^3 \eta_G^2 \eta_D}{\Lambda_1} + \Lambda_3 \delta \right) \Lambda_2 \delta \quad (4.33)$$

and  $\sin \sigma = 0$ , then

$$\frac{\|\Delta C\|_2}{\|C\|_2} \leq \frac{4k^4(2k+1)^{3/2} \Lambda_2}{\Lambda_1 \Lambda_4} \eta_G^2 \eta_D \hat{\eta}_\beta \delta + O(\delta^2). \quad (4.34)$$

This estimate is similar to the is studied one obtained in reference [12], in which it is studied the stability of a confluent Prony system. Let us suppose, without loss of generality, that  $\beta_1 < \beta_2 \cdots < \beta_k$ . From [15] we see that

$$\eta_G = \max\{k, g\} \max_{1 \leq j \leq k} g_j, \quad (4.35)$$

where

$$g = \sum_{j=1}^k \beta_j^{k-1}, \quad g_j = \prod_{l \neq j} \frac{1 + \beta_l}{|\beta_j - \beta_l|}. \quad (4.36)$$

$D$  is a diagonal matrix, therefore. We have

$$\eta_D = \frac{\max_j |C_j \beta_j (1 - \beta_j)|}{\min_j |C_j \beta_j (1 - \beta_j)|}. \quad (4.37)$$

We do not know a reasonable upper bound for the condition number  $\hat{\eta}_B$ .

It follows that the method is well conditioned if each one of the following statements is satisfied: there is  $\sigma > 0$  such that  $\beta_j \geq \sigma$  for  $1 \leq j \leq k$ , the powers  $\beta_j$  are widely separated, and the number of nonlinear parameters remains small. The proximity between two  $\beta_j$  values leads to a large value of the condition number of  $G$ ,  $\eta_G$ , and hence to the ill-conditioning of the algorithm. This characteristic is consistent with the observations of Varah, [16], who remarked the ill conditioning behaviour for any exponential fitting algorithm, provided that there are two parameters  $\beta_j$  very close.

## 5. Numerical Results

In this section we present some numerical experiments to see the actual behavior of our implementation of the alternative Prony method described above. We have results for models with two and three exponentials.

Since the application which motivated our work is related to  $T_2$  magnetic resonance brain images, we will take the linear parameters  $C_j$  to be positive and such that  $\sum_{j=1}^k C_j = 255$ , and the nonlinear ones to lie in the interval  $[0.65, 26]$ , see [17].

Given a particular model  $y(t)$ , the data values  $y_i$  are generated evaluating  $y(t_i)$  with  $\Delta t = 0.04$ ,  $t_i = i\Delta t$  and adding noise. We have used two different kinds of noise: Gaussian and Rician.

In MRI, data are intrinsically complex valued and corrupted with zero mean Gaussian distributed noise with equal variance [18]. MR magnitude images are formed by simply taking the square root of the two independent Gaussian random variables (real and imaginary parts) pixel by pixel. For an MR magnitude image, the Rician probability density function of the measure pixel intensity  $x$  is given by

$$p(x | y) = \frac{x}{\delta^2} e^{-(x^2+y^2)/2\delta^2} I_0\left(\frac{yx}{\delta^2}\right), \quad (5.1)$$

where  $I_0$  is the modified zeroth-order Bessel function of the first kind,  $y$  is the underlying noise-free signal amplitude, and  $\delta$  denotes the standard deviation of the Gaussian noise in the real and imaginary parts.

The figures are log-log graphs, the horizontal axis corresponds to the noise, and the vertical axis corresponds to the relative errors in the computations of the linear and nonlinear parameters

$$db = \left| \frac{\Delta b}{b} \right|, \quad dC_j = \left| \frac{\Delta C_j}{C_j} \right|, \quad d\lambda_j = \left| \frac{\Delta \lambda_j}{\lambda_j} \right|. \quad (5.2)$$

For each example, we first show the results for one run corresponding to a particular value of the noise and then we show the average value of the relative errors for one hundred runs. The Gaussian noise has zero mean, and the variance,  $\delta$ , varies between  $10^{-10}$  and 1. For the Rice noise we use  $(y_i, \delta)$ , with  $\delta$  varying between  $10^{-10}$  and 1 as the parameters for each simulation.

For Figures 1, 2, 3 and 4 we used the model

$$y(t) = 10 + 80e^{-11t} + 165e^{-20t}. \quad (5.3)$$

Note the linear dependence of the errors with respect to level of noise. We also noticed that, in general, the errors in the nonlinear parameters are slightly smaller than the errors in the linear parameters. In Figures 1 and 2 we have the results using Gaussian noise. Figures 3 and 4 show the results for the same model and Rice noise; as the behaviour is similar for Gaussian and Rice noises, we will only consider Rice noise for the rest of the experiments.

Next we present the results for the model

$$y(t) = 5 + 100e^{-11t} + 150e^{-12t}. \quad (5.4)$$

In Figures 5 and 6 the graphs show a deterioration, as the theory suggests, because the nonlinear parameters are closer than in the previous example. In this case the errors in the linear parameters are greater, for more than one order of magnitude, than the errors on the nonlinear parameters.

Finally, we use the model

$$y(t) = 5 + 70e^{-2t} + 85e^{-10t} + 95e^{-18t}. \quad (5.5)$$

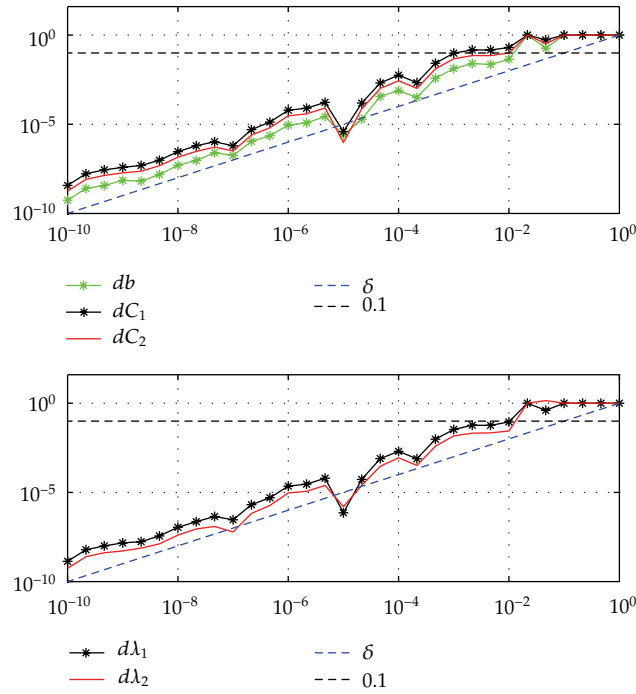


Figure 1: Relative errors for the first model corresponding to Gaussian noise.

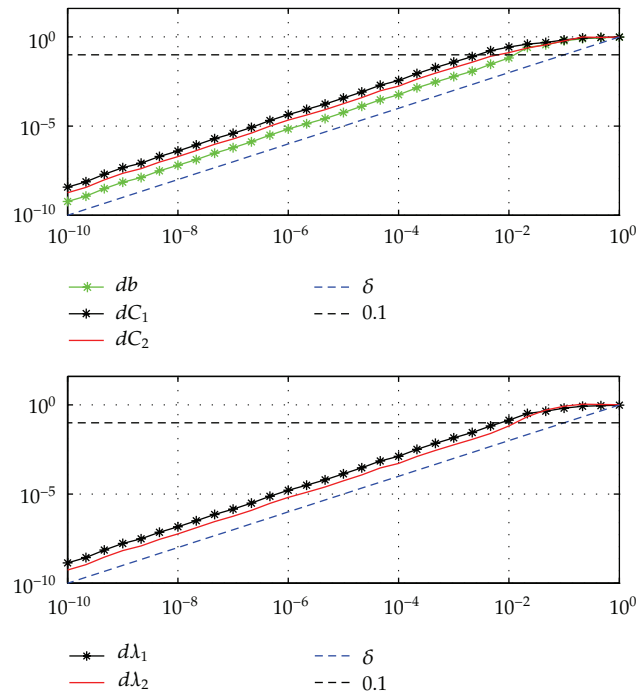


Figure 2: Relative error average, after 100 runs, for the first model corresponding to Gaussian noise.



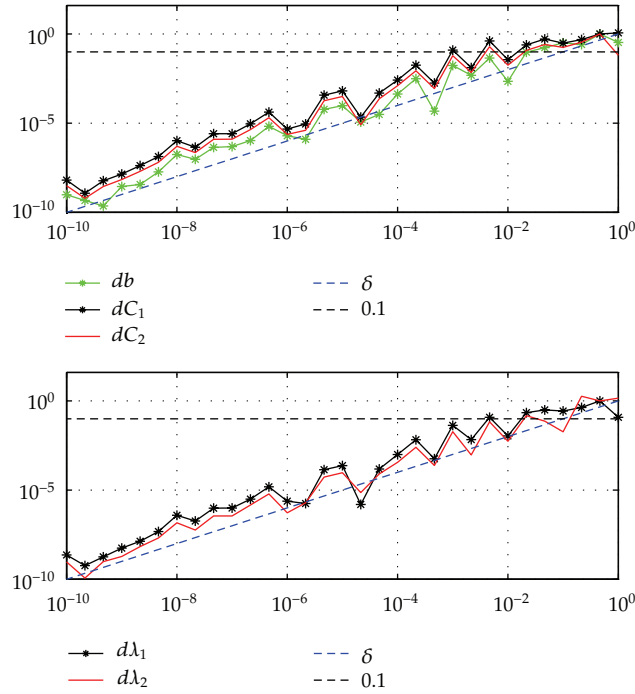


Figure 3: Relative errors for the first model corresponding to Rice noise.

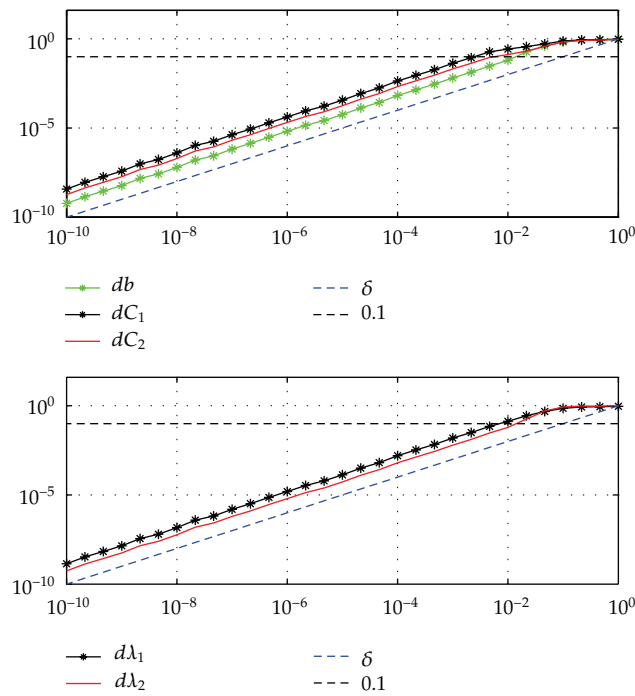


Figure 4: Relative error average, after 100 runs, for the first model corresponding to Rice noise.

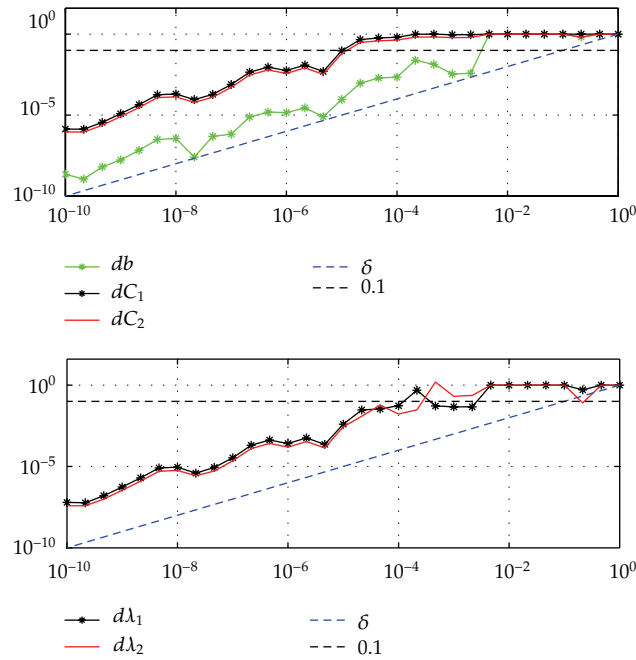


Figure 5: Relative errors for the second model corresponding to Rice noise.

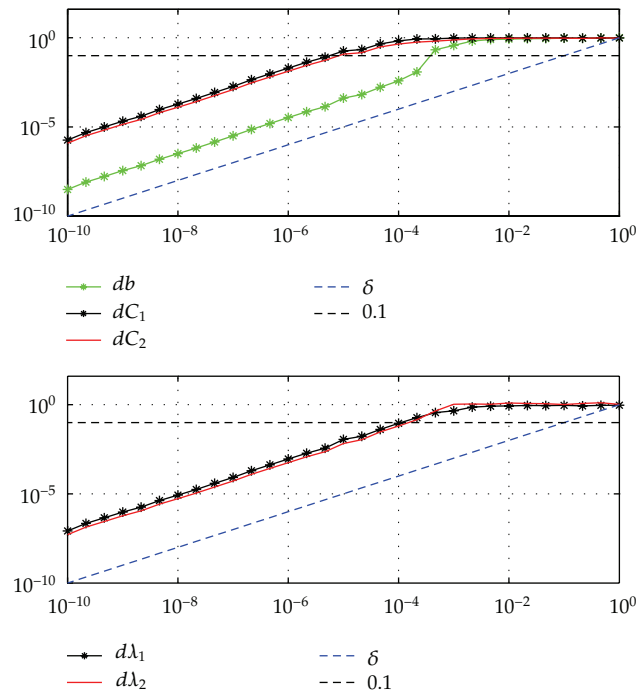


Figure 6: Relative error average, after 100 runs, for the second model corresponding to Rice noise.

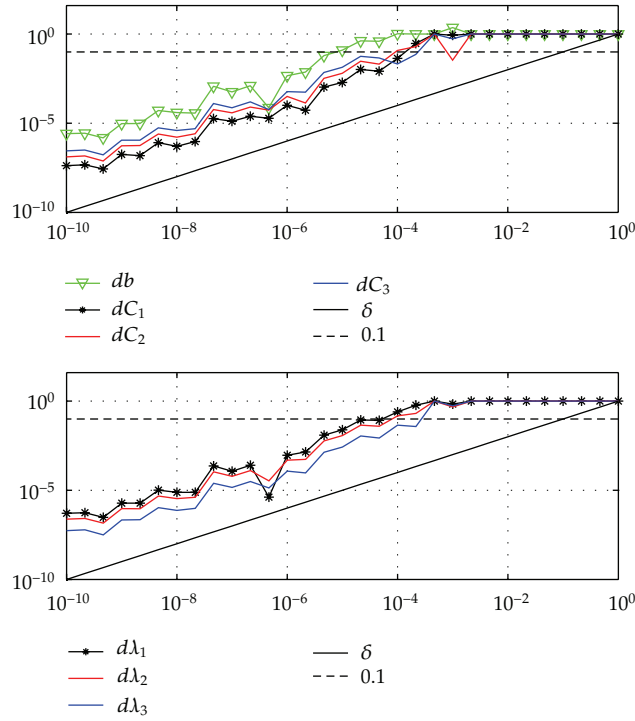


Figure 7: Relative errors for the third model corresponding to Rice noise.

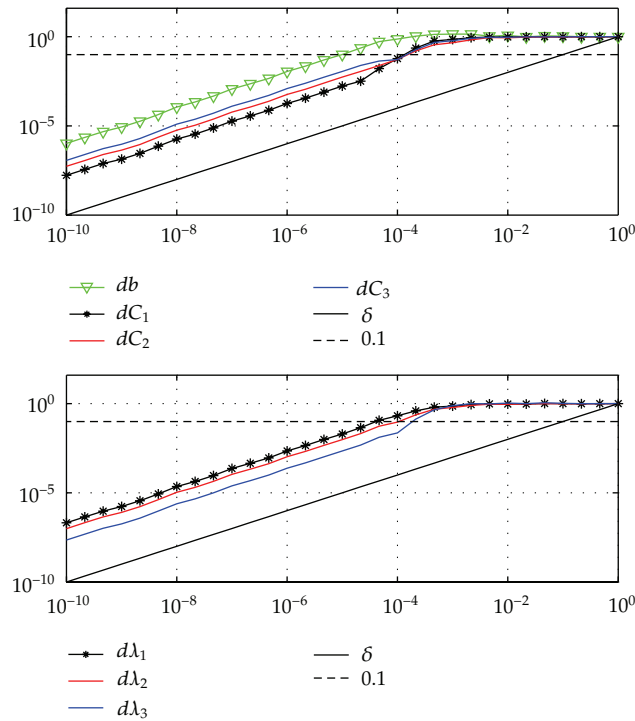


Figure 8: Relative error average, after 100 runs, for the third model corresponding to Rice noise.

**Table 1:** Condition numbers.

	$\eta_G$	$\eta_D$	$\hat{\eta}_B$
Example 1	16.88724660783	2.22604845378	53.04339947209
Example 2	130.20508573840	1.54342570227	397.12140292138
Example 3	184.8409648621	4.7771637988	269.0372904019

The results are shown in Figures 7 and 8. Again, there is a deterioration, with respect to the first example, which is caused by the addition of one term to the exponential model.

Now we present Table 1 with the computed condition numbers for the different relevant matrices for each one of the three examples.

## 6. Conclusions

In this work we have developed a stability analysis for the alternative formulation of the Prony method presented by Martín-Landrove et al. [1, 2]. The analysis shows the linear dependence between the perturbation of the data and the relative errors in the computed values for the model linear and nonlinear parameters. It is also shown that the errors in the linear parameters depend upon both the number and the closeness of the nonlinear parameters.

## Acknowledgment

This research was partially supported by Vicerrectoría de Investigación de la Universidad Nacional de Colombia, grant numbers 12427 and 16084.

## References

- [1] M. Martín-Landrove, G. Figueroa, M. Paluszny, and W. Torres, "A quasi-analytical method for relaxation rate distribution determination of T<sub>2</sub>-weighted MRI in brain," in *Proceedings of the 29th Annual International Conference of IEEE-EMBS, Engineering in Medicine and Biology Society (EMBC '07)*, pp. 1318–1321, Lyon, France, August 2007.
- [2] M. Martín-Landrove, G. Figueroa, W. Torres, and M. Paluszny, "Boosting the inverse interpolation problem by a sum of decaying exponentials using an algebraic approach," *Electronic Transactions on Numerical Analysis*, vol. 34, pp. 163–169, 2009.
- [3] M. Paluszny, M. Lentini, M. Martín-Landrove, W. Torres, and R. Martín, "Recovery of relaxation rates in MRI T<sub>2</sub>-weighted brain images via exponential fitting," in *Exponential Data Fitting and Its Applications*, pp. 69–88, Bentham Science Publishers, 2010.
- [4] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, pp. 413–432, 1973.
- [5] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Problems*, vol. 19, no. 2, pp. R1–R26, 2003.
- [6] D. P. O'Leary and B. W. Rust, "Variable projection for nonlinear least squares problems," *Computational Optimization and Applications*. In press, <http://www.cs.umd.edu/~oleary/software/varpro.pdf>.
- [7] M. H. Kahn, M. S. Mackisack, M. R. Osborne, and G. K. Smyth, "On the consistency of Prony's method and related algorithms," *Journal of Computational and Graphical Statistics*, vol. 1, no. 4, pp. 329–349, 1992.
- [8] M. R. Osborne and G. K. Smyth, "A modified Prony algorithm for fitting functions defined by difference equations," *Society for Industrial and Applied Mathematics*, vol. 12, no. 2, pp. 362–382, 1991.

- [9] M. R. Osborne and G. K. Smyth, "A modified Prony algorithm for exponential function fitting," *SIAM Journal on Scientific Computing*, vol. 16, no. 1, pp. 119–138, 1995.
- [10] M. R. Osborne, "Some special nonlinear least squares problems," *SIAM Journal on Numerical Analysis*, vol. 12, no. 4, pp. 571–592, 1975.
- [11] V. Pereyra and G. Scherer, "Exponential data fitting," in *Exponential Data Fitting and Its Applications* Bentham Science Publishers, pp. 15–41, 2010.
- [12] D. Batenkov and Y. Yomdin, "On the accuracy of solving confluent Prony system," <http://arxiv.org/pdf/1106.1137v1.pdf>.
- [13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.
- [14] J. H. Wilkinson, *Rounding Errors in Algebraic Processes*, Dover Publications, 1994.
- [15] W. Gautschi, "Optimally conditioned Vandermonde matrices," *Numerische Mathematik*, vol. 24, pp. 1–12, 1975.
- [16] J. M. Varah, "On fitting exponentials by nonlinear least squares," *Society for Industrial and Applied Mathematics*, vol. 6, no. 1, pp. 30–44, 1985.
- [17] M. Martín-Landrove, F. Mayobre, I. Bautista, and R. Villalta, "Brain tumor evaluation and segmentation by in vivo proton spectroscopy and relaxometry," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 18, no. 6, pp. 316–331, 2005.
- [18] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, 1995.

