ALGEBRAIC PARAMETER ESTIMATION OF DAMPED EXPONENTIALS

Aline Neves, Maria D. Miranda**, Mamadou Mboup****

*CECS - Federal University of ABC - Rua da Catequese, 242, Santo André, Brazil;
phone: +55 11 44371600 ramal 456; aline.panazio@ufabc.edu.br
** Department of Telecomunication and Control - University of São Paulo, Av. Prof Luciano Gualberto, trav. 3, n. 158, Brazil; maria@lcs.poli.usp.br
*** UFR Math-Info, Université Réné Descartes and ALIEN INRIA-Futurs, 45 rue des Saints-Pères
75270 Paris Cedex 06, France; mboup@math-info.univ-paris5.fr

ABSTRACT

The parameter estimation of a sum of exponentials or the exponential fitting of data is a well known problem with a rich history. It is a nonlinear problem which presents several difficulties as the ill-conditioning when roots have close values and the order of the estimated parameters, among others. One of the best existing methods is the modified Prony algorithm [1] which suffers in the presence of noise. In this paper we propose an algebraic method for the parameter estimation. The method, differently from the modified Prony method, is considerably robust to noise. The comparison of both through simulations confirm the good performance of the algebraic method.

1. INTRODUCTION

The parameter estimation of exponentially damped sinusoids is a problem that arises in many data applications. For example, models that describe heat and chemical components diffusion, or time series related to biomedical signals are represented by a sum of exponentials. The problem is specially interesting since frequency estimation is fundamental in signal processing. In addition, this kind of signal is transient due to the damping factor, what makes the problem challenging.

Usually, these signals are represented by:

$$x(t) = \sum_{j=1}^{p} a_j e^{\alpha_j t} \qquad (1)$$

where a_j are amplitudes and α_j are constants that may be complex and usually have a negative real part.

Estimation of a_j and α_j is well known to be numerically difficult [1]. Since, in practice, $x(t)$ can never be exactly observed due to noise and measurement errors, a first solution would be to use a minimum square criteria where the error is $e(t) = x(t) - y(t)$ with $y(t) = x(t) + \eta(t)$ being the observed signal and $\eta(t)$ a random Gaussian noise. The estimation of a_j and α_j through this method, however, results in a nonlinear problem.

The study of this kind of parameter estimation has a long and rich history. The classical technique of Prony, for example, extracts sinusoid or exponential signals from time series data by solving a set of linear equations for the coefficients of the recurrence equation satisfied by these signals [1]. It is a technique closely related to Pisarenko's method which finds the smallest eigenvalue of an estimated covariance matrix [2]. Prony's method, however, is well known to perform poorly in the presence of noise. On the other hand, Pisarenko's method is consistent but inefficient for estimating sinusoid signals and inconsistent for estimating damped sinusoids or exponential signals. In addition, choosing initial values, ill-conditioning when two or more α_j are close and the order of the estimated values are just some of the encountered difficulties in solving the problem.

Searching for a better performance, Osborne [3] proposed a modified Prony algorithm equivalent to maximum likelihood estimation. The method was generalized to estimate any function which satisfies a difference equation with constant coefficients. In [1], the authors apply the method to fitting sums of exponential functions. The method is shown to be relatively insensitive to starting values and also solves the ill-conditioning problem as far as the convergence of the algorithm is concerned, but may return a pair of damped sinusoids in place of two exponentials which are coalescing. Moreover, it suffers in the presence of noise even though it performs better then other existing methods.

In this paper, we propose an algebraic estimation method for the exponential fitting of the observed signal that results from an identification/estimation theory based on differential algebra and operational calculus [4, 5]. Application of this approach in other contexts as the deconvolution of BPSK (*Binary Phase Shift Keying*) and QPSK (*Quadrature Phase Shift Keying*) signals [6] and the demodulation of CPM (*Continuous Phase Modulation*) signals [7], among others [8], all associated with the estimation of different parameters, showed that the resulting methods are specially robust to noise, enabling a good estimation of the desired parameters even at very low signal to noise ratios [9]. Since noise always largely degrades the performance of existing methods for the parameter estimation of exponentially damped sinusoids, we expect the algebraic approach to enhance the results obtained under these conditions.

The paper is organized as follows. Section 2 briefly reviews the modified Prony method. Section 3 proposes an algebraic method for the estimation of the desired parameters. Considering the same framework used in [1], we study the case of a signal given by the sum of two exponentials and briefly show how to generalize the resulting method. Simulations and discussions are presented in section 4, where both methods are compared. The paper concludes with section 5.

2. THE MODIFIED PRONY METHOD

The modified Prony algorithm is equivalent to maximum likelihood estimation for Gaussian noise. It was generalized in [10] to estimate any function that satisfies a difference equation with coefficients that are linear and homogeneous in the parameters. In [1], the algorithm was applied to exponential fitting. The authors start considering a signal as shown in (1), which satisfies a constant coefficient differential equation. In the discrete time case, it is known that this signal can be written as an AR (Auto-Regressive) signal which satisfies the difference equation:

$$\sum_{k=1}^{p+1} \vartheta_k y(n-1-k) = 0 \tag{2}$$

for some suitable choice of ϑ_k with $y(n)$ being equally spaced samples of $y(t)$. From (2), setting $\boldsymbol{y} = [y(n)\ y(n-1)\ ...\ y(n-N)]^T$, we can write

$$X^T(\vartheta)\boldsymbol{y} = 0 \tag{3}$$

where $X(\vartheta)$ is the convolution matrix

$$X(\vartheta) = \begin{pmatrix} \vartheta_1 & & \\ \vdots & \ddots & \\ & & \vartheta_1 \\ \vartheta_{p+1} & & \\ & \ddots & \vdots \\ & & \vartheta_{p+1} \end{pmatrix}$$

The algorithm results from the minimization of the square error $\phi(\boldsymbol{a},\boldsymbol{\alpha}) = (\boldsymbol{y}-\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})$ subject to the constraint $\vartheta^T\vartheta = 1$. A necessary condition for it to be attained is:

$$(B(\vartheta) - \lambda I)\,\vartheta = 0 \tag{4}$$

where λ is the Lagrange multiplier and

$$\begin{aligned} B(\vartheta)_{ij} &= \boldsymbol{y}^T X_i \left(X^T X\right)^{-1} X_j^T \boldsymbol{y} + \\ & -\boldsymbol{y}^T X \left(X^T X\right)^{-1} X_i^T X_j \left(X^T X\right)^{-1} X^T \boldsymbol{y} \end{aligned}$$

with $X_j = \partial X(\vartheta)/\partial\vartheta_j$. The problem is then a nonlinear eigenvalue problem. The modified Prony algorithm solves (4) using a succession of linear problems converging to $\lambda = 0$. Given an estimate ϑ^k of the solution $\hat{\vartheta}$, solve:

$$\begin{aligned} \left(B(\vartheta^k) - \lambda^{k+1} I\right)\vartheta^{k+1} &= 0 \\ {\vartheta^{k+1}}^T \vartheta^{k+1} &= 1 \end{aligned} \tag{5}$$

with λ^{k+1} the nearest to zero of such solutions. Convergence is accepted when λ^{k+1} is small compared to $\|B(\vartheta)\|$.

In order to recover α_j having estimated ϑ, we need to obtain the roots of the characteristic polynomial $p_\vartheta = \sum_{k=1}^{p+1} \vartheta_k z^{k-1}$ which will result in $\zeta_j = n(1-e^{-\alpha_j/n})$. The final step is then

$$\alpha_j = -nlog(1-\zeta_j/n) = n\sum_{j=1}^{\infty} j^{-1}(\zeta_j/n)^j$$

The modified Prony method is relatively insensitive to starting values. It may return a pair of damped sinusoids in place of two exponentials which are coalescing. More details about this method can be found in [10, 1].

3. ALGEBRAIC ESTIMATION METHOD

Let us consider a signal given by the sum of two complex exponentials in noise:

$$y(t) = a_1 e^{\alpha_1 t} + a_2 e^{\alpha_2 t} + \gamma + \eta(t) \tag{6}$$

where α_1 and α_2 may be complex, γ is a constant bias perturbation and $\eta(t)$ is an additive noise.

We suppose that the amplitudes a_1 and a_2 are irrelevant for our purpose and consider them equal to one. Our objective is, thus, to estimate α_1 and α_2.

Considering that we will observe the signal $y(t)$ during a finite time interval $T \le t \le T+\lambda$, the noise $\eta(t)$ can be decomposed as:

$$\eta(t) = \gamma_T^\lambda + \eta_T^\lambda(t) \tag{7}$$

where the constant γ_T^λ represents its mean (average) value and $\eta_T^\lambda(t)$ is a zero-mean term. Returning to (6), we can thus consider, without any loss of generality, that in the perturbation $\gamma + \eta(t)$, $\eta(t)$ is zero-mean.

The algebraic method proposed is based on differential algebra and operational calculus. Therefore, the first step is to obtain a differential equation satisfied by (6). Ignoring the zero-mean noise term $\eta(t)$ for the moment, the signal

$$y(t) = a_1 e^{\alpha_1 t} + a_2 e^{\alpha_2 t} + \gamma \tag{8}$$

satisfies the following differential equation

$$y^{(3)} = \theta_1 y^{(2)} + \theta_2 y^{(1)} \tag{9}$$

where the super-script (i) denotes de order of differentiation of $y(t)$ with respect to time, $\theta_1 = \alpha_1 + \alpha_2$ and $\theta_2 = -(\alpha_1\alpha_2)$.

Observing (9), it is important to note that the differential equation has, as parameters, functions of the desired α_1 and α_2. In addition, (9) does not depend on the structured perturbation γ. Since it is a constant, it can easily be eliminated through differentiation.

The next step is to obtain the Laplace transform of (9):

$$s^3\hat{y} - \left(s^2 y_0 + s\dot{y}_0 + \ddot{y}_0\right) = \theta_1(s^2\hat{y} - (sy_0 + \dot{y}_0)) + \theta_2(s\hat{y} - y_0) \tag{10}$$

where y_0, $\dot{y}_0$ and $\ddot{y}_0$ are the initial conditions of $y(t)$ in the time interval being considered. These unknown constants can also be viewed as structured perturbations that can be eliminated by taking the derivative of (10) recursively with respect to the variable s three times:

$$(s^3\hat{y})^{(3)} = \theta_1(s^2\hat{y})^{(3)} + \theta_2(s\hat{y})^{(3)} \tag{11}$$

Now, we can use (11) to generate a system of equations that will allow us to estimate the desired parameters. Since we have only two unknown parameters, it suffices to differentiate (11) once more to obtain:

$$\begin{bmatrix} (s^3\hat{y})^{(3)} \\ (s^3\hat{y})^{(4)} \end{bmatrix} = \begin{bmatrix} (s^2\hat{y})^{(3)} & (s\hat{y})^{(3)} \\ (s^2\hat{y})^{(4)} & (s\hat{y})^{(4)} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} \tag{12}$$

Solving the system of equations (12) will give us an estimation of α_1 and α_2. It is clear, however, that the signal $y(t)$ is only known in the time-domain. Applying the inverse Laplace transform to (12), we observe that the multiplication by s in the operational domain will result in a differentiation with respect to time, in the time domain. Since differentiations are not robust operations to be calculated numerically, we can avoid them by dividing both sides of equation (12) by s^ν. Developing the $(.)^k$ for $k = 3,4$ and dividing by s^ν we obtain:

$$\begin{bmatrix} \frac{y^{(3)}}{s^{\nu-3}} + 9\frac{y^{(2)}}{s^{\nu-2}} + 18\frac{y^{(1)}}{s^{\nu-1}} + 6\frac{y}{s^\nu} \\ \frac{y^{(4)}}{s^{\nu-3}} + 12\frac{y^{(3)}}{s^{\nu-2}} + 36\frac{y^{(2)}}{s^{\nu-1}} + 24\frac{y^{(1)}}{s^\nu} \end{bmatrix} =$$

$$\begin{bmatrix} \frac{y^{(3)}}{s^{\nu-2}} + 6\frac{y^{(2)}}{s^{\nu-1}} + 6\frac{y^{(1)}}{s^\nu} & \frac{y^{(3)}}{s^{\nu-1}} + 3\frac{y^{(2)}}{s^\nu} \\ \frac{y^{(4)}}{s^{\nu-2}} + 8\frac{y^{(3)}}{s^{\nu-1}} + 12\frac{y^{(2)}}{s^\nu} & \frac{y^{(4)}}{s^{\nu-1}} + 4\frac{y^{(3)}}{s^\nu} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad (13)$$

where ν is larger than the largest power of s appearing in the system. In the considered case, ν has to be larger than 3.

Returning to the time domain, we will have integral operations instead of differentiations. In general, (13) is composed by terms of the form $\frac{y^{(i)}}{s^{\nu-j}}$, which inverse Laplace transform is given by the $\nu - j$ order iterated integral

$$\begin{aligned} \mathscr{L}^{-1}\left\{\frac{y^{(i)}}{s^{\nu-j}}\right\} &= \int_0^\lambda d\lambda_{\nu-j-1}\int_0^{\lambda_{\nu-j-1}}\cdots\int_0^{\lambda_1}\tau^i y(\tau)d\tau \\ &= \frac{(-1)^i}{(\nu-j-1)!}\int_0^\lambda (\lambda-\tau)^{\nu-j-1}\tau^i y(\tau)d\tau \end{aligned} \quad (14)$$

where λ is the integration interval.

Finally, the parameters can be estimated by solving

$$P\boldsymbol{\theta} = Q \quad (15)$$

where $\boldsymbol{\theta} = [\theta_1 \ \ \theta_2]^T$,

$$Q = \int_0^\lambda \begin{bmatrix} 1 & 0 \\ 0 & \tau \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \kappa^{\nu-4} y(\tau)d\tau$$

and

$$P = \int_0^\lambda \begin{bmatrix} 1 & 0 \\ 0 & \tau \end{bmatrix} \begin{bmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{bmatrix} \begin{bmatrix} \kappa^{\nu-3} & 0 \\ 0 & \tau\kappa^{\nu-2} \end{bmatrix} y(\tau)\tau d\tau$$

with $\kappa = \lambda - \tau$ and

$$\begin{aligned} p_{1,1} &= -\tau^2 b_{1,2} + 6\tau\kappa b_{1,1} - 6\kappa^2 \\ p_{2,1} &= \tau^2 b_{1,2} - 8\tau\kappa b_{1,1} + 12\kappa^2 \\ p_{1,2} &= -\tau b_{1,1} + 3\kappa \\ p_{2,2} &= \tau b_{1,1} - 4\kappa \\ q_1 &= -\tau^3 b_{1,3} + 9\tau^2\kappa b_{1,2} - 18\tau\kappa^2 b_{1,1} + 6\kappa^3 \\ q_2 &= \tau^3 b_{1,3} - 12\tau^2\kappa b_{1,2} + 36\tau\kappa^2 b_{1,1} - 24\kappa^3 \\ b_{1,j} &= \prod_{i=1}^{j}(\nu - i) \end{aligned}$$

Note that the estimation time λ, which depends on the observation signal y(t), has to be sufficient to the integrals to converge. In general this interval can be small, which translates into fast estimation.

In addition, when the zero-mean noise term $\eta(t)$ is added to the signal, the iterated integration procedure is going to largely reduce its effect at the output. For this reason the method is significantly robust to noise.

Having estimated $\boldsymbol{\theta}$, we still have to recover α_1 and α_2. From (9), α_1 and α_2 are the roots of the characteristic equation:

$$p_x = z^2 - \theta_1 z - \theta_2 = (z - \alpha_1)(z - \alpha_2) \quad (16)$$

The generalization of the procedure shown is quite direct. Consider a signal given by the sum of p exponentials:

$$y(t) = \sum_{i=1}^{p} a_i e^{\alpha_i t} + \gamma \quad (17)$$

It will satisfy the following differential equation:

$$\frac{d^{p+1}y}{dt^{p+1}} = \theta_1\frac{d^p y}{dt^p} + \theta_2\frac{d^{p-1}y}{dt^{p-1}} + \ldots + \theta_p\frac{dy}{dt} \quad (18)$$

where the coefficients θ_k are algebraic functions of the desired unknown parameters:

$$\theta_k = (-1)^{k-1} \sum_{\substack{i_1,\ldots,i_k = 1 \\ i_1 < \ldots < i_k}}^{p} \alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_k}$$

Having found the respective differential equation, we can resume the necessary steps for the development of the method as follows:

1. Obtain the Laplace transform of (18), *i.e.*, a differential equation satisfied by the observed signal, with constant coefficients that depend on the desired parameters.
2. Differentiate the resulting equation with respect to the variable s recursively $p+1$ times in order to eliminate the dependence on initial conditions of $y(t)$.
3. Continue to differentiate the resulting expression recursively in order to generate a system of equations with a number of equations equal to the number of unknown coefficients. We will then have a system with p equations.
4. Divide all equations by a term s^ν, with $\nu > p+1$, to obtain proper operators.
5. Return to the time-domain, calculating the necessary integrals as defined in (14).
6. Solve the resulting linear system. Once we have found the estimates of θ_i, $i = 1,2$, the coefficients α_i can be found as the roots of the characteristic polynomial of (18).

4. SIMULATION RESULTS AND DISCUSSION

Comparing the proposed algebraic method with the modified Prony's method, it is interesting to observe that both start considering the differential equation that is satisfied by $y(t)$. The modified Prony's method, however, treats the problem in discrete time, using the corresponding difference equation, while the algebraic method treats it directly in continuous time. In addition, both methods are insensitive to initial conditions.

Let us start with the case where $y(t)$ is given by only one exponential:

$$y(t) = e^{\alpha t}$$

where α is complex. The development of the algebraic method for this case follows exactly the same steps presented in section 3. Here, $y(t)$ satisfies the following differential equation:

$$\ddot{y}(t) = \alpha\dot{y}(t)$$

where the structured perturbation γ was already eliminated through differentiation.

Figures 1 and 2 show the convergence of both methods with respect to time, for the estimation of the real and imaginary components of α respectively. The value of α was set to $-0.2+0.9i$, ν was considered equal to 4 and white gaussian noise was added to $y(t)$. The presented curves are the resulting mean of 100 Monte Carlo simulations. The signal $y(t)$ was observed during 10 seconds and we considered 1000 samples for the algebraic method and 15 for the modified Prony method. This difference comes from the fact that the algebraic method treats the signals in continuous time and needs a sufficient number of samples for the integrals to converge, while the modified Prony method uses equally spaced samples of the signal. For the later, increasing the number of points degrades the performance.

As we can see, the performances of the algebraic method are similar for SNR (Signal to Noise Ratio) of 15 dB and 0 dB. However, this is not the case for the modified Prony method. In the estimation of α at 0 dB presented here, the obtained value is out of the graphic scale.

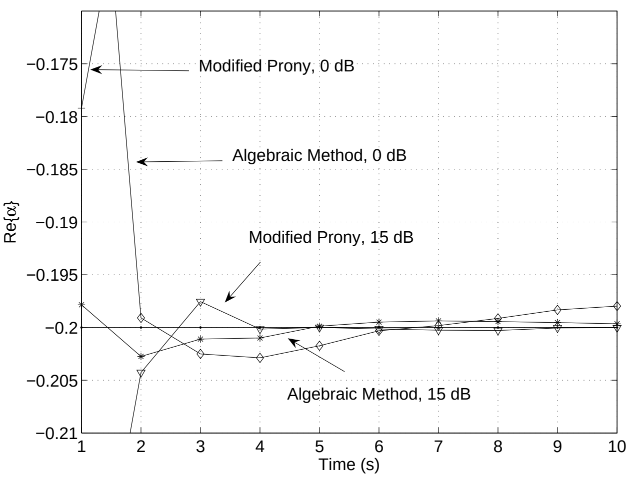

Figure 1: Estimation of Re$\{\alpha\}$

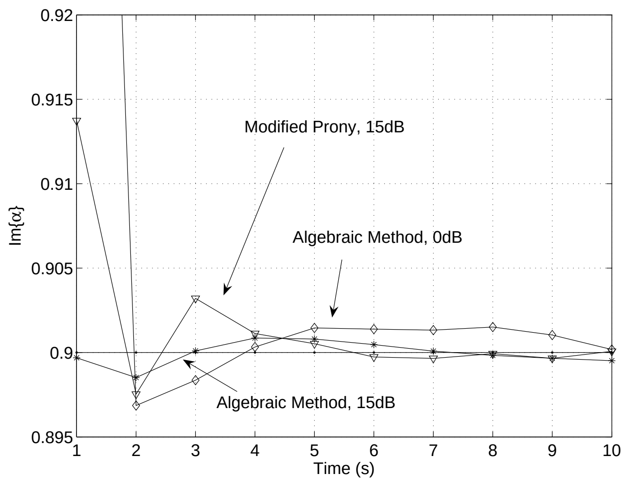

Figure 2: Estimation of Im$\{\alpha\}$

In the sequel, we will consider $y(t)$ given by (6), *i.e.*, the sum of two exponentials. Figure 3 shows the result obtained using real valued α: $\alpha_1 = -0.25$ and $\alpha_2 = -0.8$. The figure shows the mean square error (MSE) obtained comparing the estimated values of α with the correct ones for 1000 Monte Carlo simulations. The signal was observed during an interval of 20 seconds. The Algebraic Method (AM) used 5000 samples while the Modified Prony method (MP) used 25. The parameter ν was set equal to 5.

It is clear that the modified Prony method suffers considerably with the addition of noise, differently from the algebraic method. As mentioned in [1], the former may give a complex value as a result of the estimation of real parameters. The error shown in figure 3 was calculated considering only the real part of the estimated parameter. The number of inaccurate estimations increases as the SNR decreases. The algebraic method does not present such a behavior, *i.e.*, if the desired parameters are real it always returns a real valued estimation. For comparison, the figure also shows the Cramer Rao Bound (CRB).

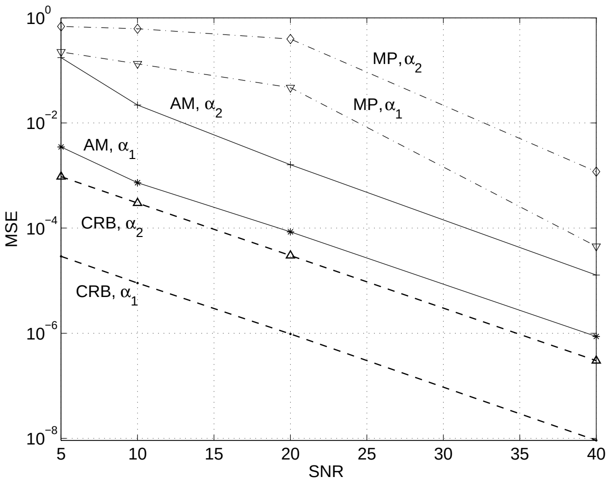

Figure 3: Comparison of the performances of the algebraic method and the modified Prony method in identifying $\alpha_1 = -0.25$ and $\alpha_2 = -0.8$

Figures 4 and 5 show the error obtained in the estimation of complex valued α: $\alpha_1 = -0.25-0.9i$ and $\alpha_2 = -1.2+i$. Once again we can see the good performance of the algebraic method when compared to the modified Prony method. The simulations in this case used the same scenario above: $\nu = 5$, 5000 samples for MA and 25 for MP, 1000 Monte Carlo simulations.

It should be noted that, in the case without noise, both methods coincide estimating the desired parameters with an error that tends to zero.

5. CONCLUSION

We have proposed an algebraic method for the parameter estimation of a sum of exponentials. This method was compared to the modified Prony method, one of the best existing methods in the literature. The starting point of both methods is the same: a linear homogeneous differential/difference equation with constant coefficients that are algebraic functions of the desired parameters. In both cases, the estima-

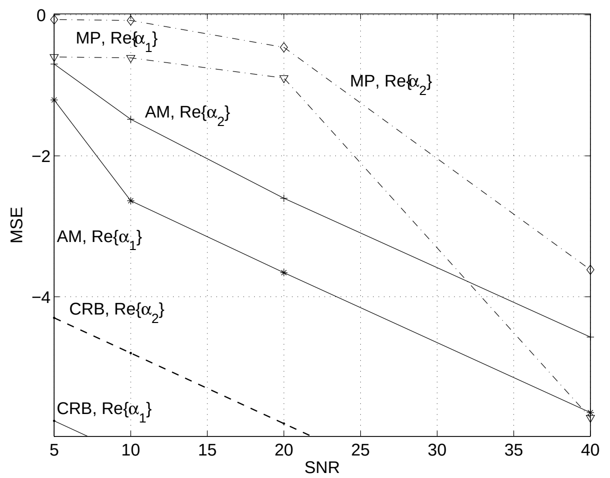

Figure 4: Comparison of the performances of the algebraic method and the modified Prony method in identifying $Re\{\alpha\}$

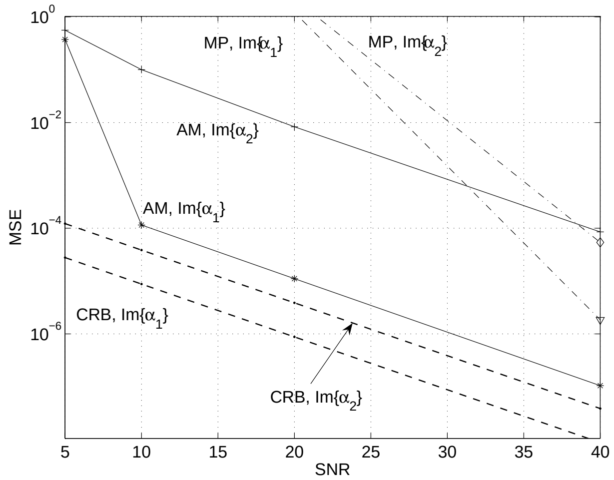

Figure 5: Comparison of the performances of the algebraic method and the modified Prony method in identifying $Im\{\alpha\}$

tions are obtained by finding the roots of a polynomial. One of the main differences is that the modified Prony is based on a nonlinear optimization while the algebraic method solves a linear system of equations.

Simulations have shown that the algebraic method is more robust to noise than the modified Prony method. The problem of yielding complex results when estimating real parameters was not observed in the simulations with the algebraic method, while for the modified Prony method it occurred in simulations where SNR ≤ 20. The number of inaccurate estimations increased with the decrease of the SNR value.

REFERENCES

[1] M. R. Osborne and G. K. Smyth, "A modified Prony algorithm for exponential function fitting," *SIAM Journal of Scientific Computing*, vol. 16, pp. 119–138, 1995.

[2] H. Ouibrahim, "Prony, Pisarenko, and the matrix pencil: a unified presentation," *IEEE Trans. Acoustic, Speech and Signal Processing*, vol. 37, pp. 133–134, 1989.

[3] M. R. Osborne, "Some special nonlinear least squares problems," *SIAM Journal of Numerical Analysis*, vol. 12, pp. 571–592, 1975.

[4] M. Fliess and H. Sira-Ramírez, "An algebraic framework for linear identification," *ESAIM Control Optim. Calculus Variations*, vol. 9, pp. 151–168, 2003.

[5] M. Fliess, M. Mboup, H. Mounier, and H. Sira-Ramírez, "Questioning some paradigms of signal processing via concrete examples," *Proc 1st Workshop Flatness, Signal Processing and State Estimate, CINVESTAV-IPN, Mexico, D.F.*, 2003.

[6] A. Neves, M. Mboup, and M. Fliess, "An algebraic identification method for the demodulation of QPSK signal through a convolutive channel," *12th European Signal Processing Conference (EUSIPCO), Viena, Austria*, 2004.

[7] A. Neves, M. Mboup, and M. Fliess, "An algebraic receiver for full response CPM demodulation," *VI International Telecommunication Symposium (ITS), Fortaleza, Brazil*, pp. 925–930, 2006.

[8] M. Mboup, "Parameter estimation via differential algebra and operational calculus," *in preparation*, 2007.

[9] A. Neves, *Identification Algébrique et Déterministe de Signaux et Systèmes à Temps Continu: Application à des Problèmes de Communication Numérique*, Ph.D. thesis, Université René Descartes - Paris V, 2005.

[10] M. R. Osborne and G. K. Smyth, "A modified Prony algorithm for fitting functions defined by difference equations," *SIAM Journal of Scientific Computing*, vol. 12, pp. 362–382, 1991.