



ESTIMATION OF DAMPED & DELAYED SINUSOIDS : ALGORITHM AND CRAMER-RAO BOUND.

Rémy Boyer and Karim Abed-Meraim

ENST, Department of Signal and Image Processing
46, rue Barrault, 75634 Paris Cedex 13
{boyer, abed} @tsi.enst.fr

ABSTRACT

In this work, we present an iterative Fourier-type algorithm which is introduced for the DDS (Damped and Delayed Sinusoid) parameter estimation using a sub-band processing approach. This algorithm is shown to improve the existing ones with respect to the computational cost and the estimation accuracy. Moreover, we derive the Cramer-Rao Bound (CRB) expression for the DDS model and we perform a simulation-based performance analysis of noisy fast time-varying synthetic signal and in the audio transient signal modeling context.

1. INTRODUCTION

Parametric models, like the constant-amplitude sinusoidal or EDS (Exponentially Damped Sinusoidal) models are popular and efficient tools in many area of interest including spectral-line [1] or pole estimation [2], source localization [3], biomedical signal processing [4] and audio signal compression [5].

In this paper, we use a generalization of these models, named the Damped & Delayed Sinusoidal (DDS) model. The latter adds a time-delay parameter which allows time-shifting of each waveform. Note that this paper goes further into the work initiated in [6]. Properties of this model are studied and we show that it allows to obtain compact representations $\hat{x}(n)$ of fast time-varying or "transient" signals $x(n)$ i.e. $M \ll N$ and $\sum_{0 \leq n \leq N-1} |x(n) - \hat{x}(n)|^2 \ll \sum_{0 \leq n \leq N-1} |x(n)|^2$ where N is the analysis duration and M is the number of sinusoids. Finally, we derive the Cramer-Rao bound for the DDS process.

2. THE DDS MODEL

The real M -DDS model definition is given by [6] :

$$\begin{aligned} \hat{x}(n) &\stackrel{\Delta}{=} \sum_{m=1}^M a_m e^{d_m(n-t_m)} \cdot \cos(\omega_m(n-t_m) + \phi_m) \cdot \psi(n-t_m) \\ &\stackrel{\Delta}{=} \sum_{p=1}^M p_m(n), \quad n = 0, \dots, N-1 \end{aligned} \quad (1)$$

where $\{a_m, \phi_m, d_m, \omega_m, t_m\}_{1 \leq m \leq M}$ are the $5M$ real amplitude, phase, damping-factor, angular-frequency and time-delay parameters and $\psi(n)$ is the Heaviside function, i.e., $\psi(n) = 1$ for $n \leq 0$ and zero otherwise. Note that if we choose $t_m = 0$ for all m , we obtain the M -EDS model [5]. The M -DDS model can be understood as a generalization of this parametric model.

3. SKETCH OF THE SOLUTION

Given a (real-valued) signal $x(n)$, the global non-linear criterion to be solved is the 2-norm of the signal $x(n) - \hat{x}(n)$ where $\hat{x}(n)$ is the M -DDS signal given by (2).

Let \mathbf{p}_m be the N -sample 1-DDS component, i.e., $\mathbf{p}_m = (p_m(0), \dots, p_m(N-1))^T$. We can consider two cases. The components are quasi-orthogonal. In other words, for $t_j \neq t_m$, we have $\langle \mathbf{p}_m, \mathbf{p}_j \rangle \approx 0$ where $\langle \cdot, \cdot \rangle$ defines the inner product. This definition can be seen as a separation constraint on the component time-supports. Indeed, if we fix $t_m < t_j$, the component \mathbf{p}_m has a sharp decreasing part (large damping-factor) in such a way that the component \mathbf{p}_j is practically not disrupted. This approach is studied in [7] where we propose several high-resolution algorithms well adapted to the audio signals. The second case is a more difficult situation since we have for $t_m \neq t_j$, $|\langle \mathbf{p}_m, \mathbf{p}_j \rangle| \gg 0$. In this case, the components are non-orthogonal and, thus, the j -th component is not clearly separated from the m -th. Consequently, a direct estimation of the time delay is a difficult task. However, the angular-frequencies estimation by means of Fourier-type [8] or subspace [2] methods, directly applied to the observed signal, remains relatively robust while a direct damping-factor estimation, on the 1-DDS signal, is systematically biased [6]. In this context, we propose to solve this problem by performing a narrow band-pass filtering around each component in conjunction with a component deflation procedure in a view to decrease the influence of the other components. Afterwards, in each sub-band, we estimate the filtered 1-DDS model parameters.

In brief, the proposed parameter estimation approach proceeds in the following steps :

- Angular-frequencies estimation using an iterative Fourier-type method.
- Sub-band filtering and deflation to 'separate' the sinusoidal components and mitigate the inter-components interferences.
- In each sub-band, estimate the damping factor, the phase and amplitude (eventually refine the frequency estimation) of the considered component.

4. DDS-D ALGORITHM : "DEFLATION APPROACH"

In [6], we have presented a new algorithm named DDS-B (B stands for Block) for the estimation of the M -DDS model parameters. This algorithm is based on the use of subspace methods and exploit a filter-bank architecture. In this paper, we introduce an improved estimation algorithm that uses a deflation approach (to en-

force the 1-DDS separation) in conjunction with a tight band-pass filter procedure. The latter, named DDS-D (where D stands for deflation), is shown to improve both the estimation accuracy and the computational cost of the DDS-B as it uses essentially FFT-based estimation procedures.

4.1. Primary angular-frequency estimation by a Fourier-type iterative algorithm

Consider the m -th residual signal defined by the recurrent equation : $x_m(n) \triangleq x_{m-1}(n) - p_m(n) = x(n) - \sum_{j=1}^m p_j(n)$ with $x_0(n) = x(n)$. Contrary to the DDS-B algorithm, we determine a primary angular-frequency $\omega_m^{(1)}$ estimation by simply maximizing the modulus of the Fourier Transform (FT) of the m -th residual signal $x_m(n)$. We denote the m -th synthetic signal by $\hat{x}_m(n) \triangleq \sum_{j=1}^m p_j(n)$ and from the previous expression, we have :

$$\varepsilon_m \triangleq \sum_{n=0}^{N-1} |x_m(n)|^2 = \sum_{n=0}^{N-1} |x(n) - \hat{x}_m(n)|^2. \quad (2)$$

This process can be stopped when the 2-norm of the residual is lower than a threshold, according to $\varepsilon_m \leq \epsilon \cdot \sum_{n=0}^{N-1} |x(n)|^2$ where $\epsilon > 0$ is a chosen threshold.

4.2. Filtering stage and sub-band processing

In order to process a tight band-pass filtering of the signal centered on the sinusoidal frequency $\omega_m^{(1)}$, we use the modulated raised cosine filter $\{h_m(n)\}$ [3]. In the context of DDS-D algorithm, the m -th residual signal is filtered such as $y_m(n) = h_m(n) * x_m(n)$. This approach enforces the extraction of the m -th component and the mitigation on the inter-components interferences.

After that, we perform the angular-frequency back-estimation and damping-factor estimation within the sub-band as described in the following section.

4.3. Sub-band parameters estimation

In each sub-band indexed by m , we find the filtered 1-DDS component which best matches the m -th N -sample sub-band signal \mathbf{y}_m , *i.e.*, we resolve the following criterion :

$$\arg \min_{\alpha_m, \mathbf{z}_m, t_m} \|\mathbf{y}_m - \hat{\mathbf{y}}_m\|_2^2 \text{ where } \hat{\mathbf{y}}_m = \mathbf{H}_m \mathbf{p}_m \quad (3)$$

where \mathbf{H}_m is the Filtering matrix computed from the filter $\{h_m(n)\}_{0 \leq n \leq P-1}$ (P being the filter length) and $\mathbf{p}_m = \mathbf{G}(\mathbf{z}_m, t_m) \alpha_m$ where $\alpha_m = (a_m e^{i\phi_m} / 2 \ a_m e^{-i\phi_m} / 2)^T$ is the complex amplitude vector, $\mathbf{z}_m = (z_m \ z_m^*)^T$ is the pole vector where $z_m = e^{d_m + i\omega_m}$ and $\mathbf{G}(\mathbf{z}_m, t_m) = [\xi_m \ \xi_m^*]$ is the $N \times 2$ Vandermonde matrix where $\xi_m = (0_{t_m}^T \ 1 \ z_m^{N-t_m-1})^T$.

4.3.1. Models Equivalence and filtering effects

By supposing that the sub-band signal is well isolated by the filtering process, we introduce the following time offset $\rho_m = \arg \max_{0 \leq n \leq N-1} |y_m(n)|$. After that, we define the truncated sub-band signal $\bar{\mathbf{y}}_m(n) = y_m(n+\rho_m)$ for $n = 0, \dots, N-\rho_m-1$ and \mathbf{J}_m which is a selection matrix such as $\bar{\mathbf{y}}_m = \mathbf{J}_m \mathbf{y}_m$. The truncated sub-band signal is efficiently approximated by the real 1-EDS model. This assumption is based on the Model Equivalence

(ME) property between the 1-EDS model and the 1-DDS with a reduced time support and modified complex amplitudes. Indeed, due to the filtering properties, we have $P_m \triangleq \rho_m - t_m \approx P$ where P is the time-delay introduced by the filter and using the fact that $p_m(n)$ is causal¹, the estimated truncated sub-band signal admits the following expression :

$$\begin{aligned} \hat{y}_m(n + \rho_m) &= \sum_{k=0}^{P_m-1} h_m(k) p_m(n - k + \rho_m) \\ &= \alpha_m H(z_m) z_m^n + \alpha_m^* H(z_m^*) z_m^{*n} \end{aligned} \quad (4)$$

with $H(z) = z^{P_m} \sum_k h_m(k) z^{-k}$. Consequently, we can see that only, the complex amplitudes are modified by the filter $h_m(n)$. Let $\mathbf{S}_m \triangleq \text{diag}\{H(z_m), H(z_m^*)\}$ then, knowing an estimate of the time-delay, criterion (3) reduces to the following one :

$$\arg \min_{\alpha_m, \mathbf{z}_m} \|\bar{\mathbf{y}}_m - \mathbf{J}_m \mathbf{G}(\mathbf{z}_m, \rho_m) \mathbf{S}_m \alpha_m\|_2^2 \quad (5)$$

4.3.2. Angular-frequency and damping-factor estimation

We perform the angular-frequency $\omega_m^{(2)}$ estimation by maximizing the modulus of the FT of the truncated sub-band signal. This is a refining of the first estimate $\omega_m^{(1)}$ of the angular frequency. After that, we estimate the damping-factor d_m by the shifted-FT method [6] where we add a windowing step to improve the performances of this approach [8]. We choose here a Blackman's window.

4.3.3. Time delay estimation

The delay parameter is estimated via a 'model-data' matching criterion. Therefore, in each sub-band m , we resolve criterion (3) with respect to the time-delay. Optimizing first over the complex amplitude then over t leads to :

$$t_m = \arg \min_{t \in \mathcal{V}(\rho_m)} f(d_m^{(1)}, t) \text{ where } f(d, t) = \|\Pi_{\mathcal{G}}^\perp(d, t) \mathbf{y}_m\|_2^2 \quad (6)$$

with $\Pi_{\mathcal{G}}^\perp(d, t) = \mathbf{I}_N - \mathcal{G} \mathcal{G}^\dagger$ (\mathcal{G}^\dagger being the pseudo-inverse of \mathcal{G}) the orthogonal projector where we have omitted the arguments for the simplicity of the notation and $\mathcal{G}(\mathbf{z}_m, t_m) = \mathbf{H}_m \mathbf{G}(\mathbf{z}_m, t_m)$ is the filtered matrix of the m -th signal pole ξ_m and its conjugate. $\mathcal{V}(\rho_m)$ is a given time interval centered at $\rho_m - P_m$. We solve (6) by a simple enumeration of the possible values in $\mathcal{V}(\rho_m)$, so as to reduce the search cost.

4.3.4. Back-estimation of the damping-factor

Once we estimate the delay t_m we can sharpen the damping-factor estimation using a non-linear optimization algorithm such as Newton's [10]. The back-estimation (using Newton method) of the damping-factor corresponds to²

$$d_m^{(2)} = d_m^{(1)} - \left(\frac{\partial^2 f}{\partial d^2}(d_m^{(1)}, t_m) \right)^{-1} \frac{\partial f}{\partial d}(d_m^{(1)}, t_m) \quad (7)$$

¹*i.e.*, $\psi(n+a) = \psi(n)$ for $n \geq 0$.

²Obviously, we can iterate equation (9) to further improve the estimation of the damping parameter.

We give the expressions of the first and the second order derivative with respect to the damping-factor as

$$\begin{aligned}\frac{\partial f}{\partial d}(d, t_m) &= 2\Re\{y_m^T \Pi_G^\perp \mathcal{G}' \mathcal{G}^\dagger y_m\} \\ \frac{\partial^2 f}{\partial d^2}(d, t_m) &= 2\Re\{y_m^T (\Pi_G^\perp (\mathcal{G}'' \mathcal{G}^\dagger + \mathcal{G}' \mathcal{G}^{\dagger\prime}) - (\mathcal{G}' \mathcal{G}^\dagger + \mathcal{G} \mathcal{G}^{\dagger\prime}) \mathcal{G}' \mathcal{G}^\dagger) y_m\}\end{aligned}\quad (8)$$

where $\mathcal{G}^{\dagger\prime} = (\mathcal{G}^H \mathcal{G})^{-1} (\mathcal{G}^{H\prime} - (\mathcal{G}^{H\prime} \mathcal{G} + \mathcal{G}^H \mathcal{G}') (\mathcal{G}^H \mathcal{G})^{-1})$. \mathcal{G}' and \mathcal{G}'' denote the first and second order derivative of \mathcal{G} , respectively.

In order to simplify the above re-estimation procedure, we did use in our simulation a Newton implementation based on the real-valued (instead of complex) vectors which lead to vector instead of matrix manipulations according to:

$$d_m^{(2)} = d_m^{(1)} - \frac{\tilde{f}'(d_m^{(1)})}{\tilde{f}''(d_m^{(1)})} \quad (9)$$

where $\tilde{f}(d) \triangleq (\tilde{g}_d^T y_m)^2$ with $\tilde{g}_d = (\tilde{g}_d(0), \dots, \tilde{g}_d(N-1))^T$ and $\tilde{g}_d(n) \triangleq e^{d(n-t_m)} \cos(\omega_m^{(2)}(n-t_m) + \hat{\phi}_m) \psi(n-t_m)$. $\hat{\phi}_m$ represents an estimate of the phase parameter given by $\hat{\phi}_m = (\angle(\alpha_m(1)) - \angle(\alpha_m(2))) / 2$ where $\alpha_m = (\alpha_m(1) \ \alpha_m(2))^T$ is estimated as in (10) using $\hat{z}_m = e^{d_m^{(1)} + i\omega_m^{(2)}}$.

4.3.5. Complex amplitude estimation in the sub-band

We have to estimate the complex amplitude (amplitude and phase parameter) of the sub-band signal by minimizing the least squares fitting criterion (5). Given the frequency and damping factor estimates, *i.e.*, $\omega_m^{(2)}$ and $d_m^{(2)}$, the solution of (5) is expressed as :

$$\alpha_m = S_m^{-1} (\mathbf{J}_m \mathbf{G}(\hat{z}_m, \rho_m))^\dagger \bar{y}_m \quad (10)$$

where $\hat{z}_m = e^{d_m^{(2)} + i\omega_m^{(2)}}$.

5. CRAMER-RAO BOUND FOR THE DDS MODEL

The CRB for the parameter estimations of a DDS process is derived in this section. The CRB is useful as a touchstone against which the efficiency of the considered estimators can be tested.

The CRB has been investigated in [9] for a damped sinusoidal process. We derive, here, the conditional CRB for the more general DDS case. More precisely, the CRB is computed conditionally to the exact knowledge of the discrete-valued delay parameters. Let consider a real-valued M -DDS process corrupted by zero-mean, unit-variance, white gaussian noise $x(n) = \hat{x}(n) + \sigma w(n)$ where $\hat{x}(n)$ is given by (2). Let $\gamma = [\mathbf{d}^T, \omega^T, \phi^T, \mathbf{a}^T]^T$ be the vector of desired damping-factor, angular-frequency, phase and amplitude parameters where $\mathbf{d} = (d_1, \dots, d_M)^T$ (ω, ϕ and \mathbf{a} are defined in the same way³). We have the following results (proofs are omitted due to space limitation):

Lemma : Under the above assumptions, the elements of the Fisher Information matrix $\mathcal{F}_t(\gamma, \sigma^2)$ corresponding to the cross terms of γ_k and σ^2 are zero.

This lemma allows to "ignore" the noise parameter and compute the "sub-matrix" of the Fisher Information matrix corresponding to the desired parameters γ .

³The time-delay parameter vector $\mathbf{t} = (t_1, \dots, t_M)^T$ is omitted here as it is assumed perfectly known.

Corollary : The CRB on the variance of any unbiased estimate of γ (conditionally to the perfect knowledge of the time-delay parameter vector \mathbf{t}) is given by :

$$[\text{CRB}_t(\gamma)]_{i,j} = \sigma^2 \left[\frac{\partial \hat{\mathbf{x}}^T}{\partial \gamma_i} \cdot \frac{\partial \hat{\mathbf{x}}}{\partial \gamma_j} \right]^{-1} \quad (11)$$

Remarks : 1) We choose here to compute a bound conditionally to the exact knowledge of delay parameters because the latter are discrete-valued and consequently the computation of a (non-conditional) bound leads to inextricable derivations. On the other hand, choosing time-delay parameters with continuous real-valued lead to the following model indeterminacy : for $n = 0, \dots, N-1$, we have

$$\begin{aligned}ae^{d(n-t)} \cos(\omega(n-t) + \phi) \psi(n-t) &= \\ ae^{d(n-(t+\tau))} \cos(\omega(n-(t+\tau)) + \phi + \tau\omega) \psi(n-(t+\tau))\end{aligned}\quad (12)$$

for any τ such that $\lceil t + \tau \rceil = \lceil t \rceil$ where $\lceil \cdot \rceil$ is the integer part.

2) We observed in our simulation a relatively small distance (especially when the damping factor is low) between our estimation methods performances and the CRB for low and moderate values of the SNR. However, the gap becomes significant for high SNRs. The reason is due to the fact that the delay parameters are assumed perfectly known in the derivation of the CRB. Indeed, based on this a priori knowledge, we can obtain much better frequency and damping factor estimates than those given by the DDS-D algorithm (for example, in the noiseless case, we can perform an exact estimation of these parameters from a finite-length observation). An exact CRB (that takes into account the estimation of delay parameters) can be easily obtained for a *Soft* DDS model where the Heaviside function used in the signal modeling is replaced by an appropriate continuous function. This point is still under investigation and will be the focus of future works.

6. SIMULATIONS

6.1. Synthetic signal

We choose a 2-DDS non-orthogonal components, *i.e.*, $|t_2 - t_1|$ is small. In this case, a time-delay estimation/detection based on the variation of the signal envelope is inefficient (see [7]).

The algorithms (DDS-D and DDS-B) are compared, in terms of parameter estimation accuracy through the normalized Mean Square Error (MSE), evaluated for several Signal to Noise Ratios (SNR) using 100 Monte-Carlo trials. The MSE is defined by the ratio of the square difference between the true parameter value and its estimated value over the square value of the true parameter. In relation to figures 1 and 2, we can say that the DDS-D algorithm, outperforms the DDS-B algorithm in this simulation context. This conclusion can be explained by the fact that the deflation character of the DDS-D algorithm enforces the separation of the components. Finally, we can note that the performances of these two algorithms are quite far from the ideal performances of the CRB. We can improve the efficiency of these algorithms by considering a joint Newton algorithm in ω, d, ϕ . However, this improvement has a price : the computational complexity. Moreover, we obtain approximatively, the same performances on the real percussive audio signals of the next section. Consequently, we have kept a "simple" (and cheaper) Newton on the parameter d since this approach represents a good trade-off between computational complexity and performance.

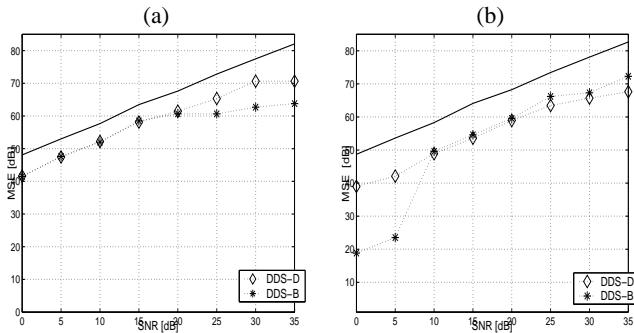


Fig. 1. Angular-frequency estimation performance, (a) first component, (b) second component, CRB (solid line).

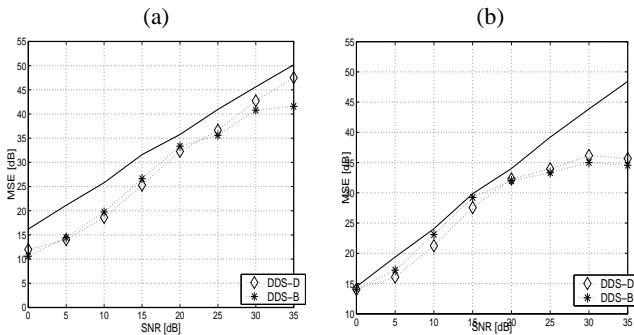


Fig. 2. Damping-factors estimation performance, (a) first component, (b) second component, CRB (solid line).

6.2. Typical audio transient signal

In the context of percussive audio modeling, we choose to apply the proposed algorithms on a castanet onset which is a typical audio transient signal (see the top of figure 3-a). On the same figure, we have represented on the middle (respectively bottom), the 20-order modeling by the DDS-B (respectively DDS-D) algorithm. The chosen criterion is the SMNR (Signal to Modeling Noise Ratio) which is a time matching criterion between the synthesized waveform and the original signal. Then, we obtain 11.2 dB for the DDS-B algorithm and 12.7 dB for the DDS-D algorithm. This result is confirmed by the observation of figure 3-b. Indeed, we can see that the DDS-B algorithm estimates several time-delay parameters lower than 223 samples. Consequently, we observe on figure 3-a (middle), a small pre-echo (distortion before the sound onset). Inversely, the DDS-D modeling presents a total absence of pre-echo and a good reproduction of the onset dynamic.

7. CONCLUSION

In this article, we have presented a Fourier-type iterative model parameter estimation algorithm for DDS signals. We compare this approach with the subspace-type method DDS-B, introduced in [6] and we show through simulations on synthetic and real transient signals that the DDS-D algorithm is more efficient than the DDS-B approach since its deflation scheme enforces the separation of the

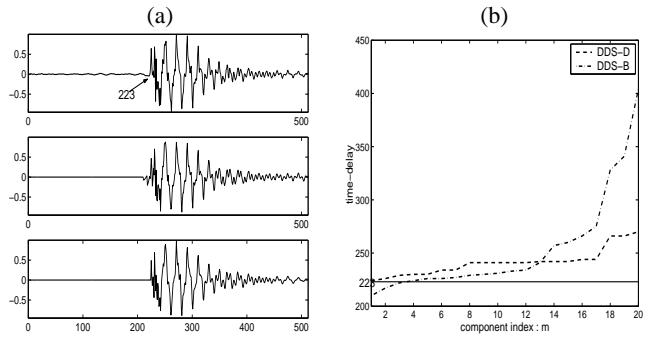


Fig. 3. (a), (top) original castanet onset (normalized amplitude), (middle) 20-order modeling by the DDS-B algorithm, (bottom) 20-order modeling by the DDS-D algorithm, (b) time-delay estimation with respect to the index component.

components, for a lower computational cost. Finally, we derive the expression of the Cramer-Rao Bound for the DDS process.

8. REFERENCES

- [1] P. Stoica and R. Moses, *Introduction to spectral analysis*, Prentice-Hall, 1997.
- [2] Y. Hua and T.K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 38 Issue: 5 , May 1990.
- [3] J.G. Proakis, *Digital Communications*, McGraw-Hill Companies, 1995.
- [4] S. Van Huffel, C. Decanniere, H. Chen and P. Van Hecke, "Algorithm for time-domain NMR data fitting based on total least squares", *Journal of Magnetic Resonance A*, Vol. 110, 1994, pp. 228-237.
- [5] R. Boyer, S. Essid and Nicolas Moreau, "Dynamic temporal segmentation in parametric non-stationary modeling for percussive musical signals", *Proc. of IEEE Int. Conf. on Multimedia and Expo*, August 2002.
- [6] R. Boyer and K. Abed-Meraim, "Audio transients modeling by Damped & Delayed Sinusoids (DDS)", *Proc. of IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, May 2002.
- [7] R. Boyer and K. Abed-Meraim "Efficient parametric modeling for audio transients", *Proc. of 5th Int. Conf. on Digital Audio Effects (DAFx-02)*, Hamburg, Germany, September 2002.
- [8] P. O'Shea, "The use of sliding spectral windows for parameter estimation in power system disturbance monitoring", *IEEE Trans. on PES*, 2000.
- [9] T. Wigren, A. Nehorai, "Asymptotic Cramer-Rao bounds for estimation of the parameters of damped sine waves in noise", *IEEE Trans. on SP*, April 1991
- [10] J.E. Dennis, JR & R.B. Schnabel, *Numerical methods for unconstrained optimization and non-linear equations*, Prentice-Hall, 1983.