**ABSTRACT**

The most common sinusoidal models for non-stationary analysis represent either complex amplitude modulated exponentials with exponential damping (cPACED) or log-amplitude/frequency modulated exponentials (generalised sinusoids), by far the most commonly used modulation function being polynomials for both signal families. Attempts to tackle a hybrid sinusoidal model, i.e. a generalised sinusoid with complex amplitude modulation were relying on approximations and iterative improvement due to absence of a tractable analytical expression for their Fourier Transform. In this work a simple, direct solution for the aforementioned model is presented.

1. INTRODUCTION

Sinusoidal analysis algorithms’ vast area of application, ranging from sound and music analysis [1, 2], medical data analysis [3], imaging [4], Doppler radar [5] and sonar applications, seismic signal analysis, hydrogen atom spectrum analysis [6], laser technology [7] and financial data analysis [8] had fuelled the research field for decades. Rapidly modulated sinusoids found in most of the aforementioned applications have sparked the interest in non-stationary sinusoidal analysis.

Recent development in this area, have provided efficient and accurate methods for either cPACED or generalised sinusoid model. A somehow hybrid model was attempted with real polynomial amplitude modulation and frequency modulation, however to authors’ knowledge only approximate, iterative-improvement type algorithm were developed to date [9][10].

A very old idea of TF energy reassignment [11] has been a focus of much research lately [12,13,14,15,16]. The vast variety of modifications of the original reassignment, be it merely a generalisation for higher modulations [17] or redefined for an entirely different model [18], has called for a more general name for this family of algorithms. Recently a reallocation [19] of TF energy was proposed and will be used in this work.

The paper is organised as follows: in section 2 the hybrid signal model is outlined, while section 3 derives the non-linear multivariate polynomial system and its solution, inspired by the distribution derivative method. Section 4 compares the accuracy and computational complexity of the proposed algorithm to the high-resolution method based on rotational invariance.

2. HYBRID SIGNAL MODEL

The hybrid sinusoidal model will be defined as follows:

\( s(t) = a(t)e^{r(t)} \)

\( a(t) = \sum_{k=0}^{\infty} a_k m_k(t), a_k \in \mathbb{C} \)

\( r(t) = \sum_{l=0}^{\infty} r_l n_l(t), r_l \in \mathbb{C} \)

where \( m_k, n_k \) are the complex amplitude and log-AM/FM model functions respectively. To accommodate for the static amplitude and phase, \( m_0 = n_0 = 1 \) is assumed. A most common, but by no means mandatory selection for the model functions are polynomials: \( m_k = n_k = t^k \).

The above model is ambiguous with respect to parameters \( r_0 \) and \( a_0 \). To show this, the following derivation is considered:

\( s(t) = a(t)e^{r(t)} = a_0 \tilde{a}(t)e^{r_0 + \tilde{r}(t)} \)

\( = \tilde{a}(t) \exp(\log(|a_0|) + j\angle(a_0) + r_0 + \tilde{r}(t)) \)

Clearly \( a_0 \) and \( r_0 \) are in fact the same parameter in either Cartesian or polar coordinates. The decision seems irrelevant, however as will be shown in section 3 using the Cartesian form would result in a rank-deficient system, therefore the model will be constrained to \( a_0 = 1 \).

It is important to note that since modulation functions are complex they both contribute to overall AM/FM. If the same model functions are used (\( m_k = n_k \)) that can lead to some ambiguity, especially when the energy of \( m_k, n_k \) declines fast with \( k \). Such ambiguity can be demonstrated when using polynomials for the modulation functions \( m_k = n_k = t^k \):

\( a(t)e^{r(t)} = \exp(\log(a(t)) + r(t)) \)

\( \approx \exp(a_1 t + (2a_2 - a_1^2)t^2 + r(t)) \)

using the 2nd degree truncated Taylor expansion. It is expected that an estimator for the model in (7) could be inaccurate when separate parameter estimates are considered, but generally much more flexible due to twin AM/FM functions. In practice however one is mostly concerned with algorithm’s overall ability to fit to the signal under investigation, rather than individual per-parameter accuracy. It would be feasible to devise a disambiguation procedure but this is considered to be outside the scope of this document.

An example is shown in figure 1, where 2 hybrid model sinusoids with significantly different \( a_1, r_1 \) reach a signal-to-residual ratio (SRR) of 24dB. To reach higher SRR the parameters would...
Figure 1: Two sinusoids with significantly different parameters obtain a similar SRR of 24 dB.

have to eventually match exactly, however in noisy conditions a relatively high SRR is achievable with substantial error in parameter estimates. From figure[1] is also evident that the proposed model includes sinusoids with negative amplitude, suggesting good coding abilities for sinusoid pairs with close frequencies (i.e.: beating partials). The negative amplitude can occur when $\Im[a(t)] = 0$, since the overall amplitude corresponds to: $\sqrt{a(t)(t)e^{j\Im[r(t)]}}$. The notion of negative amplitude is purely artificial ($\sqrt{a(t)(t)e^{j\Im[r(t)]}}$ cannot be negative), as it does not have a natural physical meaning, however it comes handy as a mathematical generalisation. If negative amplitude is not allowed and $\Im[a(t)] = 0$, the derivative of $\sqrt{a(t)(t)e^{j\Im[r(t)]}} = |a(t)|e^{j\Im[r(t)]}$ is not continuous for all roots of $\mathcal{a}(t)$ and thus the model function could not be considered a holomorphic function. In such cases the absolute value can be easily dropped and negative amplitude introduced, leading to a mathematically sound model in the context of holomorphic functions.

3. NON-LINEAR MULTIVARIATE POLYNOMIAL SYSTEM OF EQUATIONS

The non-linear system can be derived by considering the signal time derivative manipulated in the following way:

\[ s'(t) = a'(t)e^r(t) + r'(t)a(t)e^{r(t)}, \quad (8) \]
\[ a(t)s'(t) = a'(t)s(t) + a(t)r'(t)s(t). \quad (9) \]

The last row can be rewritten in a more verbose form that reveals the non-linearity of the system:

\[ m_0s' + \sum_{k=1}^{K} a_k m_k s' = \sum_{k=1}^{K} a_k m_k s + \left(m_0 + \sum_{k=1}^{K} a_k m_k \right) \sum_{l=0}^{L} r_l n_l s, \quad (10) \]

where time variable $t$ was omitted for compactness. The only non-linear terms arise from the last - double sum expression. Multiplying both sides with a window function $w(t)$ and taking a Fourier Transform (FT) at frequency $\omega$ yields:

\[ S_{w_m}^f(\omega) + \sum_{k=1}^{K} a_k S_{w_m}^f(\omega) = \sum_{k=1}^{K} a_k S_{w_m}^f(\omega) + \sum_{l=0}^{L} r_l S_{w_m}^f n_l^f(\omega) + \sum_{k=1}^{K} a_k \sum_{l=0}^{L} r_l n_l^f S_{w_m}^f(\omega), \]

where $S_{\mathcal{F}}(\omega) = \langle s(t)f(t), e^{j\omega t} \rangle$ is the FT of the signal $s$ multiplied by function $f$, and $S_{\mathcal{F}}^f(\omega) = \langle s(t)g(t), e^{j\omega t} \rangle$ is the FT of the signal derivative multiplied by function $g$ at frequency $\omega$. Note that $n_0^f = n_0^g = 0$ and thus the sums on the right-hand side start at index 1 rather than 0. Above equation can be viewed as a (non-linear) multivariate polynomial with respect to parameters $a_k : k = 1 \ldots K, r_l : l = 1 \ldots L.$ The expressions $S_{\mathcal{F}}^f, S_{\mathcal{F}}^g$ can be considered constants for any $f, g$ as they can be directly computed from the signal. To calculate $S_{\mathcal{F}}^f$ accurately, sample difference in time domain should be avoided [12]. A common approach is the use of distribution derivative rule $(x', y') = -(x, y')$ and a real window function $w$ as a part of the kernel $y$:

\[ S_{g_u}^f(\omega) = \langle s g, w \psi_\omega \rangle = \langle s, g w \psi_\omega \rangle = \langle s, g w \psi_\omega \rangle + 2 \langle s g', w \psi_\omega \rangle \]
\[ = \langle s g', w \psi_\omega \rangle + 2 \langle s g', w \psi_\omega \rangle + \langle s g, w \psi_\omega \rangle + j \omega(s, g w \psi_\omega) \]
\[ = -S_{g_u}^f(\omega) - S_{g_u}^f(\omega) + j \omega(S_{g_u}^f(\omega) + S_{g_u}^f(\omega)), \]

where $\psi_\omega$ is generally a kernel function with FT centred around frequency $\omega$. For the last equality to hold the kernel is set simply to the Fourier kernel: $\psi_\omega(t) = e^{j\omega t}$. Higher time derivatives can accurately be computed by chaining the above expression. Rearranging the equation and collecting together the model parameters yields:

\[ S_{w_m}^f(\omega) = \sum_{k=1}^{K} a_k (S_{w_m}^f(\omega) - S_{w_m}^f(\omega)) + \sum_{l=0}^{L} r_l S_{w_m}^f n_l^f(\omega) \quad (16) \]

Taking the FT at different frequencies close to the peak provides as many equations as necessary. Assuming polynomial modulation functions, the following system can be derived:

\[ S_{w}^f(\omega) + \sum_{k=1}^{K} a_k S_{e_k}^f(\omega) = \sum_{l=0}^{L} k a_k S_{e_{k-1}}^f(\omega) + \sum_{l=0}^{L} \sum_{l=0}^{L} (l + 1) r_{l+1} S_{e_{l+1}}^f(\omega) \quad (17) \]

For a cPACED sinusoids with polynomial amplitude of degree 3 (ie: $K=3, L=1$) the following case can be deduced:

\[ S_{w}^f(\omega) + a_1 S_{w}^f(\omega) + a_2 S_{e_2}^f(\omega) + a_3 S_{e_3}^f(\omega) = \]
\[ a_1 S_w(\omega) + 2a_2 S_{e_2}^f(\omega) + a_3 S_{e_3}^f(\omega) + a_1 r_1 S_{e_1}^f(\omega) + a_2 r_2 S_{e_2}^f(\omega) + a_3 r_1 S_{e_3}^f(\omega) + r_1 S_{w}^f(\omega) \quad (18) \]
Grouping the linear and non-linear terms in respect to $a_t, r_1$:

$$S'_w(\omega) = a_1(S_w(\omega) - S'_{tw}(\omega)) + a_2(2S_w(\omega) - S'_{\tau w}(\omega)) + a_3(3S_{\tau w}(\omega) - S'_{\tau w}(\omega)) + r_1S_w(\omega) + a_1r_1S_{tw}(\omega) + a_2r_1S_{\tau w}(\omega) + a_3r_1S'_{\tau w}(\omega).$$  \hspace{1cm} (19)

The distribution derivative rule can be applied to the $S'$ terms:

$$S'_{w'}(\omega) = -kS_{w'-1}(\omega) - S'_{w'}(\omega) + j\omega S'_{w'}(\omega),$$  \hspace{1cm} \text{for } k > 0 \hspace{1cm} (20)

$$S'_w(\omega) = -S'_w(\omega) + j\omega S_w(\omega),$$  \hspace{1cm} \text{for } k = 0 \hspace{1cm} (21)

Following the approach of the distribution derivative method and considering the FT at 4 different frequencies, we obtain the following non-linear multivariate system:

$$Ax = b \hspace{1cm} (22)$$

$$A = \begin{pmatrix}
S_w(\omega_1) - S'_{tw}(\omega_1) & \cdots & S_w(\omega_N) - S'_{tw}(\omega_N) \\
2S_w(\omega_1) - S'_{\tau w}(\omega_1) & \cdots & 2S_w(\omega_N) - S'_{\tau w}(\omega_N) \\
3S_{\tau w}(\omega_1) - S'_{\tau w}(\omega_1) & \cdots & 3S_{\tau w}(\omega_N) - S'_{\tau w}(\omega_N) \\
S_w(\omega_1) & \cdots & S_w(\omega_N) \\
S_{tw}(\omega_1) & \cdots & S_{tw}(\omega_N) \\
S_{\tau w}(\omega_1) & \cdots & S_{\tau w}(\omega_N) \\
S'_{w}(\omega_1) & \cdots & S'_{w}(\omega_N)
\end{pmatrix}$$

$$x = \begin{pmatrix}
a_1 \\ a_2 \\ a_3 \\ r_1 \\ r_{1a_1} \\ r_{1a_2} \\ r_{1a_3}
\end{pmatrix}, \quad b = \begin{pmatrix}
S'_w(\omega_1) \\ S'_w(\omega_2) \\ S'_w(\omega_3) \\ S'_w(\omega_4)
\end{pmatrix}. \hspace{1cm} (23)$$

Note that for high parameter values, the frequency spread of the signal might be large - a small number of frequency bins (in the above case 4) might not suffice to cover enough information in the Fourier domain. In such cases more frequency bins can be considered.

### 4. Tests and Results

The proposed method was implemented in a Matlab script for an arbitrary degree of amplitude and exponential complex polynomials. The resulting multivariate non-linear systems were solved by the Matlab function `fsolve`. This function only accepts real variables and coefficients as parameters, although internally can use complex variables and solve the system if the complex equations are split into real and imaginary parts.

A polynomial amplitude of degree 3 was studied and the polynomial denoted as: $[a_3, a_2, a_1, 1] = [p_3 + jq_3, p_2 + jq_2, p_1 + jq_1, 1]$. The test values for $p_3, p_2, p_1$ were chosen so all the terms of the amplitude polynomial have equal impact on the final value:

$$p_3 \in \left[ -\left( \frac{f_s}{8T_1} \right)^3, \left( \frac{f_s}{8T_1} \right)^3 \right] \hspace{1cm} (24)$$

$$p_2 \in \left[ -\left( \frac{f_s}{8T_1} \right)^2, \left( \frac{f_s}{8T_1} \right)^2 \right] \hspace{1cm} (25)$$

$$p_1 \in \left[ -\frac{f_s}{8T_1}, \frac{f_s}{8T_1} \right] \hspace{1cm} (26)$$

The exact same value sets were used for the imaginary part of the polynomial $q(t)$. A Hann window function of length 511 samples was used for pole estimation and Hann window for the complex polynomial coefficients estimation. The damping factor was set to [-150,0,150] and only one frequency of 10000Hz was considered to match tests performed in [12]. $r_0$ was set to 0, since gain has theoretically no effect when the snr is fixed. For $p_1, p_2, p_3, q_1, q_2$ and $q_3$ parameters, only 5 linearly distributed values have been tested in order to keep the computational time reasonable. The comparison to a 31st degree (i.e. 4 poles and amplitudes) simple high-resolution method (HRM) implementation from DESAM Toolbox [20] (section 5.1.2.) without whitening and the cPACED reassignment method (cPACED-RM) [13] was conducted. The signal tested is the real part of the complex cPACED signal, reflecting the real world scenario when analytical signal isn’t available.

To measure accuracy, the commonly used Signal-to-Residual-Ratio (SRR) metric was used,

$$SRR = \frac{(s, w(s))}{(s - \hat{s}, w(s - \hat{s}))} \hspace{1cm} (27)$$

where $s, \hat{s}$ are the original signal (without noise) and the estimated signal respectively, and $w$ is the Hann window. The Signal-to-Noise-Ratio (SNR) range from -20 to 50dB with steps of 10dB was studied. The total computation times for both methods follow:

<table>
<thead>
<tr>
<th>Method</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>cPACED-RM</td>
<td>7 min</td>
</tr>
<tr>
<td>cPACED-DDM</td>
<td>300 min</td>
</tr>
<tr>
<td>High-resolution RM</td>
<td>880 min</td>
</tr>
</tbody>
</table>

Since HRM involves singular value decomposition of correlation matrix of size $N/2 \times N/2$ the computation cost is significantly the highest among the tested methods. cPACED-RM method requires $K - 1$ FFTs for the pole and $K$ DFTs for the complex polynomial estimates to build a linear system. In contrast, the proposed method requires only $K$ DFTs to build a non-linear multivariate system. However, solving such system by iterative methods requires a significant computation cost. cPACED-RM method performs much faster since it requires solving a linear system.

The classic Cramer-Rao bounds (CRBs) parameter-by-parameter comparison would total to 10 plots, overcomplicating the results and obscuring the overall accuracy. A more intuitive approach involves only one SRR/SNR plot, although a different upper accuracy bound is required. For each test case the CRBs for each parameter were computed. Denoting a CRB for parameter $\hat{\epsilon}_{a_0}$ as $\epsilon_{a_0}$, the minimum SRR for the specific CRB set can be defined:

$$\min SRR(\hat{s}(a_3 \pm \epsilon_{a_3}, a_2 \pm \epsilon_{a_2}, a_1 \pm \epsilon_{a_1}, r_0 \pm \epsilon_{r_0}, r_1 \pm \epsilon_{r_1}) \hspace{1cm} (29)$$

The mean and variance of the minimum SRR represents a good upper SRR bound. Figure 1 depicts the mean and variance of the upper SRR bound, the proposed method (cPACED-DDM), cPACED-RM and HRM. At low SNR, cPACED-DDM and cPACED-RM
perform roughly the same, up to 5dB below the upper bound, while cPACED-RM performs ~3dB better. For mid-high SNRs, all methods perform roughly the same, about ~5dB below the upper bound. In general cPACED-DDM performs ~1dB below cPACED-RM.

The main advantage of the proposed algorithm is that it offers the flexibility to analyze generalised sinusoids with complex amplitude of any polynomial degree combination. With the purpose of having an initial exploration of its general performance, we have computed the mean SRR obtained for several combinations of amplitude and exponential complex polynomial degrees. In particular, amplitude polynomial degrees were set to [0,1,2,3] and exponential polynomial degrees to [1,2,3]. Table 1 shows the results obtained. In this experiment we used the same polynomial coefficient value sets as for the previous cPaced case. For each SNR, the results show a general tendency to decrease the SRR as we increase the complexity of the signal by increasing the degrees of the polynomials. Figure 3 compares the mean and variance of the SRR obtained for all degree combinations with the SNR. It shows a general trend of reaching an SRR ~24dB above the SNR value, although the difference decreases significantly to ~13dB for SNR~20dB, and to ~14dB for SNR=50dB. This seems to indicate that the proposed method reaches a plateau for high SNRs above 50dB.

5. DISCUSSION AND FUTURE WORK

In this work the currently most flexible sinusoidal method for TF energy reassignment analysis has been described. The concept used in the distribution derivative method is used to generate a nonlinear multivariate system of polynomials obtained by the first signal derivative. It is important to note that higher signal derivatives would provide enough equations for a solution to exist, however a significantly more complex system would be obtained. Even if solution could eventually be obtained, it is desirable to avoid higher signal derivatives due to ill conditioning.

The method showed a similar performance than the high resolution and the reassignment methods for the cPACED signal model, however in theory the proposed method is much more flexible since it can be applied to generalised sinusoids with complex amplitude of any polynomial degree combination. The initial exploration of the performance obtained for several degree combinations was promising, although a more in depth evaluation was left for future work. Further, the proposed sinusoidal model seems promising for the analysis of overlapping partials, as the beating function corresponds to real value amplitude/frequency modulated sinusoids - a subfamily of signals described by the proposed model.

On the other hand high-resolution methods’ intrinsic frequency resolution of 1 frequency bin [12] for damped sinusoids has not been surpassed, as common window function mainlobe width (several bins) and significant sidelobe amplitude both reduce the frequency resolution.

6. ACKNOWLEDGMENTS

This research is partially funded by the SIGMUS project (TIN2012-36650).

7. REFERENCES

Figure 2: SRR mean and variance for cPaced signals.

Figure 3: mean and variance of SRR means for all combinations of \([0,1,2,3]\) amplitude polynomial degrees and \([1,2,3]\) exponential polynomial degrees sets. Details are given in Table 1.

Table 1: SRR mean for several polynomial degree combinations. Column numbers in top row indicate amplitude and exponential complex polynomial degrees respectively (e.g. a1r3 means amplitude polynomial degree 1 and exponential polynomial degree 3).

<table>
<thead>
<tr>
<th>SNR</th>
<th>a0r1</th>
<th>a0r2</th>
<th>a0r3</th>
<th>a1r1</th>
<th>a1r2</th>
<th>a1r3</th>
<th>a2r1</th>
<th>a2r2</th>
<th>a2r3</th>
<th>a3r1</th>
<th>a3r2</th>
<th>a3r3</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>72.2</td>
<td>66.7</td>
<td>64.6</td>
<td>62.8</td>
<td>63.9</td>
<td>62.3</td>
<td>64.2</td>
<td>61.2</td>
<td>60.5</td>
<td>62.3</td>
<td>60.8</td>
<td>58.2</td>
</tr>
<tr>
<td>40</td>
<td>67.0</td>
<td>63.1</td>
<td>61.7</td>
<td>59.2</td>
<td>61.3</td>
<td>60.0</td>
<td>61.7</td>
<td>59.2</td>
<td>58.7</td>
<td>60.1</td>
<td>58.9</td>
<td>56.6</td>
</tr>
<tr>
<td>30</td>
<td>56.6</td>
<td>56.5</td>
<td>54.2</td>
<td>54.3</td>
<td>54.6</td>
<td>53.2</td>
<td>54.5</td>
<td>53.0</td>
<td>52.4</td>
<td>53.3</td>
<td>52.4</td>
<td>50.5</td>
</tr>
<tr>
<td>20</td>
<td>50.0</td>
<td>46.4</td>
<td>44.7</td>
<td>44.9</td>
<td>45.2</td>
<td>43.8</td>
<td>44.8</td>
<td>43.6</td>
<td>42.9</td>
<td>44.0</td>
<td>43.0</td>
<td>41.1</td>
</tr>
<tr>
<td>10</td>
<td>36.2</td>
<td>35.2</td>
<td>34.7</td>
<td>36.8</td>
<td>34.8</td>
<td>33.6</td>
<td>35.0</td>
<td>33.5</td>
<td>32.8</td>
<td>33.9</td>
<td>32.7</td>
<td>29.4</td>
</tr>
<tr>
<td>0</td>
<td>25.9</td>
<td>25.7</td>
<td>24.9</td>
<td>26.4</td>
<td>25.2</td>
<td>23.2</td>
<td>25.2</td>
<td>23.1</td>
<td>22.1</td>
<td>23.8</td>
<td>21.7</td>
<td>15.1</td>
</tr>
<tr>
<td>-10</td>
<td>13.7</td>
<td>16.8</td>
<td>14.9</td>
<td>15.4</td>
<td>14.6</td>
<td>10.3</td>
<td>14.7</td>
<td>10.9</td>
<td>8.4</td>
<td>13.3</td>
<td>7.8</td>
<td>-1.8</td>
</tr>
<tr>
<td>-20</td>
<td>3.2</td>
<td>1.4</td>
<td>-3.0</td>
<td>-1.5</td>
<td>-2.4</td>
<td>-7.4</td>
<td>-1.6</td>
<td>-6.9</td>
<td>-13.1</td>
<td>-2.5</td>
<td>-10.4</td>
<td>-19.8</td>
</tr>
</tbody>
</table>

Table 1: SRR mean for several polynomial degree combinations. Column numbers in top row indicate amplitude and exponential complex polynomial degrees respectively (e.g. a1r3 means amplitude polynomial degree 1 and exponential polynomial degree 3).


