NOTES ON MODEL-BASED NON-STATIONARY SINUSOID ESTIMATION METHODS USING DERIVATIVES

Wen Xue and Mark Sandler

Centre for Digital Music, School of EE & CS,
Queen Mary, University of London
London, United Kingdom
{xue.wen;mark.sandler}@elec.qmul.ac.uk

ABSTRACT

This paper reviews the derivative method and explores its capacity for estimating time-varying sinusoids of complicated parameter variations. The method is reformulated on a generalized signal model. We show that under certain arrangements the estimation task becomes solving a linear system, whose coefficients can be computed from discrete samples using an integration-by-parts technique. Previous derivative and reassignment methods are shown to be special cases of this generic method. We include a discussion on the continuity criterion of window design for the derivative method. The effectiveness of the method and the window design criterion are confirmed by test results. We also show that, thanks to the generalization, off-model sinusoids can be approximated by the derivative method with a sufficiently flexible model setting.

1. INTRODUCTION

Estimation of parameters of slow-varying sinusoids has received much attention in the digital audio researches, and especially in the area of sinusoid modelling of audio and speech [1]. A slow-varying sinusoid is described by its instantaneous phase and amplitude functions, i.e.

$$s(t) = a(t)e^{j\varphi(t)} . \tag{1}$$

Unlike the case of constant sinusoids, which only have three parameters, for slow-varying sinusoids the parameter set has a higher degree of freedom than the signal itself. Extra information regarding the behaviour of parameters is therefore necessary to reach at any unique set of parameters from the signal.

One widely adopted way to handle this issue is to assume that the sinusoid parameters obey certain parametric model within a short interval from which they are estimated. Early methods assumes short-time parameter stationarity and estimate three basic sinusoid parameters, i.e. frequency, amplitude and phase angle, as if they are constant [1][2]. Later methods developed this assumption by involving extra parameters that describe parameter variations. For example, [3] and [4] assume linear frequency and amplitude, [5] and [6] assume linear frequency and log-amplitude, while [7] assumes linear frequency and constant amplitude. Although not all sinusoids obey these models, the methods are generally helpful in reducing estimation errors.

Several parameter estimation methods involve taking derivatives or differences regarding time. For example, in [7] the derivatives of the window function are used for calculating spectra, in [8] the instantaneous frequency is estimated by taking the difference of phase angles, while in [9] it is estimated by comparing the spectra of the signal and its derivative. If we look at the simplest case $s = ae^{\varphi_0 + j\omega_0 t}$, then

$$s' = j\omega_0 s \tag{2a}$$

By taking the short-time Fourier transform (STFT) of (2a) we get

$$S'_w(t,\omega) = j\omega_0 S_w(t,\omega) , \tag{2b}$$

where the STFT is defined as ω

$$S_w(t,\omega) = \int s(t+\tau)w(\tau)e^{-j\omega\tau}d\tau \tag{3}$$

The key feature of (2a) is that by taking the derivative, the parameter ω is singled out as a coefficient that applies to s, and remains at a similar position in the frequency-domain equivalent (2b). Once $S'_w(t,\omega)$ and $S_w(t,\omega)$ are computed for a proper pair of t and ω, the angular frequency ω_0 can be solved from (2b). This is the derivative method for constant sinusoids [9]. The conversion into frequency domain serves to keep wide-band or distant-frequency noises from partaking in the calculation.

In this paper we explore the possibilities of the derivative framework in a more general sense, including the existing non-stationary derivative and reassignment methods as special cases. A systematic approach is taken that allows highly complicated signal models, while examples are given to illustrate technical details as appear in various special cases. Section 2 formulates the derivative method for a general signal model as well as several specific ones; section 3 discusses the numeric solution of the general model along with window continuity considerations; section 4 tests both the effectiveness of the method and the continuity criterion regarding window design.

2. THE DERIVATIVE METHOD

We give the generalized signal model of the derivative method as

$$s(t) = e^{R(t)} ,\ R(t) = \sum_{m=0}^{M-1} r_m h_m(t) , \tag{4}$$

where $h_m(t)$, m=0, 1, …, M-1, are fixed real functions of t with sufficient order of continuity, and r_m is a flexible complex coefficient of h_m. If we separate the real and imaginary parts of R as

$$R(t) = P(t) + jQ(t) ,\ P(t), Q(t) \in \mathbf{R}, \tag{5a}$$

then $P(t) = \sum_{m=0}^{M-1} (\mathrm{Re}\, r_m) h_m(t)$, $Q(t) = \sum_{m=0}^{M-1} (\mathrm{Im}\, r_m) h_m(t)$, and

$$s(t) = a(t)e^{j\varphi(t)},\ a(t) = e^{P(t)},\ \varphi(t) = Q(t), \tag{5b}$$

i.e. $P(t)$ and $Q(t)$ respectively define the amplitude and frequency variation laws of s. We always let $h_0(t)=1$, so that r_0 is interpreted as a global amplification and phase-shift factor. Once the functions $h_m(t)$ are selected, the signal is parameterized by the coefficients r_m, $m=0, \ldots, M-1$. In practice the real and imaginary parts of r_m are often treated separately. For convenience we write $p_m=\mathrm{Re}\, r_m$ and $q_m=\mathrm{Im}\, r_m$.

Examples 1. Constant sinusoids are modelled by $M=2$, $h_0=1$, $h_1(t)=t$, $p_1=0$; exponentially enveloped constant sinusoids are modelled by $M=2$, $h_0=1$, $h_1(t)=t$; linear chirps are modelled by $M=3$, $h_0=1$, $h_1(t)=t$, $h_2(t)=0.5t^2$, $p_1=p_2=0$; exponentially enveloped linear chirps are modelled by $M=3$, $h_0=1$, $h_1(t)=t$, $h_2(t)=0.5t^2$, $p_2=0$; a constant sinusoid frequency-modulated by a sinusoidal modulator of angular frequency ω_M is modelled by $M=4$, $h_0=1$, $h_1(t)=t$, $h_2(t)=\cos\omega_M t$, $h_3(t)=\sin\omega_M t$, $p_1=p_2=p_3=0$.■

The choice of M, h_m and constraints on p_m, q_m, such as in the examples above, determine the nature of a signal model. We call this a *model setting*. Given signal s and a model setting, the derivative method evaluates r_m from the spectra of s and its derivatives against t. The number of derivatives to take depends on the number of unknown values in the model. As will be shown later, by taking each derivative we obtain one complex equation, which reduces the degree of freedom by 2 in terms of real unknown values.

2.1. First derivative

Taking the derivative of (4) we get

$$s'(t) = s(t)R'(t) = \sum_m r_m s(t)h'_m(t) \tag{6a}$$

Applying STFT on (6a) we get

$$S'_w(t,\omega) = \sum_m r_m \int s(t+\tau)h'_m(t+\tau)w(\tau)e^{-j\omega\tau}d\tau \tag{6b}$$

(6b) is a linear equation of the coefficients r_m, $m=1, \ldots, M-1$. Notice that r_0 is not covered by this equation, since $h'_0=0$. When evaluated for $t=0$ (6b) is simplified as

$$S'_w(\omega) = \sum_m r_m \int s(\tau)h'_m(\tau)w(\tau)e^{-j\omega\tau}d\tau, \tag{6c}$$

in which case the coefficient of r_m can be regarded as the STFT of s calculated with the window function $h'_m w$. By (6c) alone we can solve models with up to 2 real parameters apart from r_0.

Example 2. Applying the exponentially enveloped constant sinusoid model ($M=2$, $h_0=1$, $h_1(t)=t$) in (6c) we get

$$S'_w(\omega) = r_1\int s(\tau)w(\tau)e^{-j\omega\tau}d\tau = r_1 S_w(\omega) \tag{7a}$$

(7a) is the "damped" version of (2b). The parameters p_1 (decay rate) and q_1 (angular frequency) can be estimated as

$$p_1 = \mathrm{Re}\frac{S'_w(\omega)}{S_w(\omega)},\ q_1 = \mathrm{Im}\frac{S'_w(\omega)}{S_w(\omega)}\ .■ \tag{7b}$$

Example 3. Applying the linear chirp model ($M=3$, $h_0=1$, $h_1(t)=t$, $h_2(t)=t^2$, $p_1=p_2=0$) in (6c) we get

$$\begin{aligned} S'_w(\omega) &= j\left(q_1\int s(\tau)w(\tau)e^{-j\omega\tau}d\tau + q_2\int s(\tau)\tau w(\tau)e^{-j\omega\tau}d\tau\right) \\ &= j\left(q_1 S_w(\omega) + q_2 S_{tw}(\omega)\right) \end{aligned} \tag{8a}$$

where $S_{tw}(\omega)$ is the STFT of $s(t)$ calculated with window $tw(t)$. (8a) is the "chirped" version of (2b). The parameters q_1 (angular frequency) and q_2 (angular chirp rate) can be estimated as

$$q_1 = -\frac{\mathrm{Re}\, S'_w(\omega)S^*_{tw}(\omega)}{\mathrm{Im}\, S_w(\omega)S^*_{tw}(\omega)},\ q_2 = -\frac{\mathrm{Re}\, S'_w(\omega)S^*_w(\omega)}{\mathrm{Im}\, S_{tw}(\omega)S^*_w(\omega)} \tag{8b}$$

Unlike the reassignment method in [7], by (8b) we are able to jointly estimate the frequency and frequency slope without taking any 2nd-order derivative.■

2.2. Second derivative: direct form

Taking the derivative of (6a) we get

$$\begin{aligned} s''(t) &= s(t)\left((R'(t))^2 + R''(t)\right) \\ &= \sum_{m=0}^{M-1}\sum_{n=0}^{M-1} r_m r_n s(t)h'_m(t)h'_n(t) + \sum_{m=0}^{M-1} r_m s(t)h''_m(t) \end{aligned} \tag{9a}$$

Applying STFT on (9a) we get

$$\begin{aligned} S''_w(t,\omega) &= \sum_{m=0}^{M-1} r_m \int s(t+\tau)h''_m(t+\tau)w(\tau)e^{-j\omega\tau}d\tau \\ &+ \sum_{m=0}^{M-1}\sum_{n=0}^{M-1} r_m r_n \int s(t+\tau)h'_m(t+\tau)h'_n(t+\tau)w(\tau)e^{-j\omega\tau}d\tau \end{aligned} \tag{9b}$$

which, when evaluated at $t=0$, becomes

$$\begin{aligned} S''_w(\omega) &= \sum_{m=0}^{M-1} r_m \int s(\tau)h''_m(\tau)w(\tau)e^{-j\omega\tau}d\tau \\ &+ \sum_{m=0}^{M-1}\sum_{n=0}^{M-1} r_m r_n \int s(\tau)h'_m(\tau)h'_n(\tau)w(\tau)e^{-j\omega\tau}d\tau \end{aligned} \tag{9c}$$

(9c) is a quadratic equation of the coefficients r_m, $m=1, \ldots, M-1$. Again r_0 is not covered. Using (9c) in conjunction with (6c) we can solve models with up to 4 real parameters apart from r_0.

Example 4. Applying the exponentially enveloped linear chirp model ($M=3$, $h_0=1$, $h_1(t)=t$, $h_2(t)=0.5t^2$, $p_2=0$) in (6c) and (9c) we get

$$S'_w(\omega) = r_1 S_w(\omega) + jq_2 S_{tw}(\omega) \tag{10a}$$

$$S''_w(\omega) = jq_2 S_w(\omega) + r_1^2 S_w(\omega) + j2r_1 q_2 S_{tw}(\omega) - q_2^2 S_{t^2w}(\omega) \tag{10b}$$

The parameters are solved by

$$q_2 = \frac{-\mathrm{Re}\,B \pm \left((\mathrm{Re}\,B)^2 - 4\,\mathrm{Re}\,A\,\mathrm{Re}\,C\right)^{1/2}}{2\,\mathrm{Re}\,A}, \tag{10c}$$

$$r_1 = \frac{S'_w(\omega) - jq_2 S_{tw}(\omega)}{S_w(\omega)}, \tag{10d}$$

where $A=S_{tw}(\omega)^2 - S_{t^2w}(\omega)S_w(\omega)$, $B=jS_w(\omega)^2$, and $C=S'_w(\omega)^2 - S''_w(\omega)S_w(\omega)$. Notice that (10c) yields two values of q_2. To choose the right one we test that

$$q_2^2\,\mathrm{Im}\,A + q_2\,\mathrm{Im}\,B + \mathrm{Im}\,C = 0\ .■ \tag{10e}$$

We notice that in the direct form of the derivative method, high-order differentiations introduce high powers to the coefficients, so that the problem quickly becomes overcomplicated with the expansion of signal model. Fortunately, by deriving derivatives recursively, we are able to maintain the coefficients within a linear system. This we describe as follows.

2.3. Second derivative: recursive form

We reconsider the derivative of (6a) in the arrangement of

$$\begin{aligned} s''(t) &= s(t)R''(t) + s'(t)R'(t) \\ &= \sum_{m=0}^{M-1} r_m s(t) h''_m(t) + \sum_{m=0}^{M-1} r_m s'(t) h'_m(t) \end{aligned} \tag{11a}$$

which leads to

$$S''_w(\omega) = \sum_{m=0}^{M-1} r_m \int \left(s(\tau)h''_m(\tau) + s'(\tau)h'_m(\tau) \right) w(\tau) e^{-j\omega\tau} d\tau \tag{11b}$$

(11b) is the linear version of (9c). Compared to the latter, it not only avoids quadratic terms, but also has fewer summands.

Example 5. Applying the exponentially enveloped linear chirp model (M=3, h_0=1, $h_1(t)$=t, $h_2(t)$=0.5t^2, p_2=0) in (11b) we get

$$S''_w(\omega) = r_1 S'_w(\omega) + jq_2 \left(S_w(\omega) + S'_{tw}(\omega) \right), \tag{12a}$$

then q_2 can be solved from (10a) and (12a) as

$$q_2 = \mathrm{Im} \frac{S''_w(\omega) S_w(\omega) - S'_w(\omega)^2}{S_w(\omega)^2 + S'_{tw}(\omega) S_w(\omega) - S_w(\omega) S'_w(\omega)}. \tag{12b}$$

With (12b) there is no need to test against extraneous roots. Once q_2 is evaluated r_1 can be calculated by (10d).■

Example 6. Applying the frequency modulated sinusoid model (M=4, h_0=1, $h_1(t)$=t, $h_2(t)$=cos $\omega_M t$, $h_3(t)$=sin $\omega_M t$, p_1=p_2=p_3=0) in (6c) and (11b) we get

$$-jS'_w(\omega) = q_1 S_w(\omega) - q_2 \omega_M S_{ws}(\omega) + q_3 \omega_M S_{wc}(\omega) \tag{13a}$$

$$\begin{aligned} -jS''_w(\omega) &= q_1 S'_{tw}(\omega) - q_2 \left(\omega_M^2 S_{wc}(\omega) + \omega_M S'_{ws}(\omega) \right) \\ &\quad - q_3 \left(\omega_M^2 S_{ws}(\omega) - \omega_M S'_{wc}(\omega) \right) \end{aligned} \tag{13b}$$

where S_{wc}/S_{ws} (S'_{wc} / S'_{ws}) are the STFT of s (s') calculated with window functions $w(t)$cos $\omega_M t$ / $w(t)$sin $\omega_M t$, respectively. (13a) and (13b) provide 4 real linear equations, from which parameters q_1~q_3 can be easily solved.■

2.4. Higher-order derivatives

For signal models with more than 4 real parameters apart from r_0 we need to take derivatives of orders higher than 2. For an arbitrary order k, we write

$$s^{(k)}(t) = \left(s(t)R'(t) \right)^{(k-1)} = \sum_m r_m \left(s(t) h'_m(t) \right)^{(k-1)},\ k\geq 1, \tag{14a}$$

by taking the STFT of which we arrive at

$$S_w^{(k)}(t,\omega) = \sum_m r_m \int \left(s(t+\tau) h'_m(t+\tau) \right)^{(k-1)} w(\tau) e^{-j\omega\tau} d\tau \tag{14b}$$

which, when evaluated at t=0, is reduced to

$$S_w^{(k)}(\omega) = \sum_m r_m \int \left(s(\tau) h'_m(\tau) \right)^{(k-1)} w(\tau) e^{-j\omega\tau} d\tau\ . \tag{14c}$$

(14c) is the order-k version of (6c) and (11b). Obviously it is a linear equation of the coefficients r_m, or equivalently, two linear equations of p_m and q_m. (14b) and (14c) are also the general formulations of the (recursive) derivative method. By substituting k=1, 2, …, K in (14c) we obtain a system of K complex (2K real) linear equations, from which up to 2K real parameters can be solved.

3. COMPUTATION ISSUES

In this section we explain the numerical implementation of the derivative method proposed in section 2, which can not be applied directly due to the general non-availability of s and its derivatives in their continuous form. In practice we only have the signal s in its sampled form

$$s_n = s(n),\ n \in \mathbf{Z}, \tag{15}$$

from which we need to calculate the coefficients of r_m in (14b) or (14c). All these coefficients have the form of

$$X_w^{(k)} = \int x^{(k)}(\tau) w(\tau) e^{-j\omega\tau} d\tau\ . \tag{16a}$$

where $x(\tau)$ is known at $\tau\in\mathbf{Z}$, and $x^{(k)}$ stands for $(d^k/dt^k)\cdot x$. In equation (14b) $x(\tau)$ corresponds to $s(t+\tau)h'_m(t+\tau)$, in (14c) to $s(\tau)h'_m(\tau)$. For numerical computation we assume that w is compactly supported on [-T, T], $T\in\mathbf{Z}^+$.

3.1. Derivatives

In the literature of derivative methods attempts have been reported to calculate signal derivatives by interpolating s, e.g. [6]. This, however, risks breaching the signal model, and may require extra data outside the duration of the window function. In this paper we take the approach of integration by parts, which transfers the differentiation operators from the signal s to the window function w. We extend (16a) by allowing differentiation of the window function w:

$$X_{w^{(l)}}^{(k)} = \int x^{(k)}(\tau) w^{(l)}(\tau) e^{-j\omega\tau} d\tau \tag{16b}$$

Integrating the right side of (16b) by parts we get

$$\begin{aligned} X_{w^{(l)}}^{(k)} &= \int w^{(l)} e^{-j\omega\tau} dx^{(k-1)} \\ &= w^{(l)} e^{-j\omega\tau} x^{(k-1)} \Big|_{-T}^{T} - \int x^{(k-1)} \left(w^{(l+1)} e^{-j\omega\tau} - j\omega w^{(l)} e^{-j\omega\tau} \right) d\tau \\ &= w^{(l)} e^{-j\omega\tau} x^{(k-1)} \Big|_{-T}^{T} - X_{w^{(l+1)}}^{(k-1)} + j\omega X_{w^{(l)}}^{(k-1)} \end{aligned} \tag{17a}$$

If we let $w^{(l)}(-T)$= $w^{(l)}(T)$=0 then (17a) is simplified as

$$X_{w^{(l)}}^{(k)} = -X_{w^{(l+1)}}^{(k-1)} + j\omega X_{w^{(l)}}^{(k-1)} \tag{17b}$$

(17b) reduces the differentiation order of x by 1. By repeating (17b) recursively we are able to, eventually, calculate (16a) without differentiating x.

Example 7. Using (17b) in (8a) for linear chirps we get

$$-S_{w'}(\omega) + j\omega S_w(\omega) = j\left(q_1 S_w(\omega) + q_2 S_{tw}(\omega) \right), \tag{18a}$$

from which we have

$$q_1 + q_2 \,\mathrm{Re} \frac{S_{tw}(\omega)}{S_w(\omega)} = \omega - \mathrm{Im} \frac{S_{w'}(\omega)}{S_w(\omega)}. \tag{18b}$$

(18b) can be regarded as a reassignment equation [10], where $\mathrm{Re}\left(S_{tw}(\omega)/S_w(\omega) \right)$ and $-\mathrm{Im}\left(S_{w'}(\omega)/S_w(\omega) \right)$ are the amounts of time and frequency reassignments, respectively. It provides an alternative proof that reassignment is perfect for linear chirps.■

Example 8. Using (17b) in (10a) and (12a) for exponentially enveloped linear chirps we get

$$-S_{w'}(\omega)+j\omega S_w(\omega)=r_1 S_w(\omega)+jq_2 S_{tw}(\omega) \tag{19a}$$

$$\begin{aligned} &S_{w''}(\omega)-j2\omega S_{w'}(\omega)-\omega^2 S_w(\omega) \\ &=r_1(-S_{w'}(\omega)+j\omega S_w(\omega))-q_2(jS_{tw'}(\omega)+\omega S_{tw}(\omega)) \end{aligned} \tag{19b}$$

From which we eliminate r_1 and get

$$q_2=\frac{\operatorname{Im}\dfrac{S_{w''}(\omega)}{S_w(\omega)}-\operatorname{Im}\dfrac{S_{w'}(\omega)S_{w'}(\omega)}{S_w^2}}{\operatorname{Re}\dfrac{S_{tw}(\omega)S_{w'}(\omega)}{S_w^2(\omega)}-\operatorname{Re}\dfrac{S_{tw'}(\omega)}{S_w(\omega)}} \tag{19c}$$

(19c) is the reassignment method [7] for calculating linear chirp rates, which requires 1st and 2nd derivatives. By this example we have shown that the same method applies to linear chirps with exponential amplitudes too. However, the methods in [7] for calculating instantaneous frequency can not be directly applied. One solves (19a) instead.■

To implement (14c) for k=1, …, K we need to calculate $S_w^{(k)}(\omega)$ for k=1, …, K and $X_{m\,w}^{(k)}(\omega)$ for k=0, …, K-1, m=0, …, M-1, where $x_m = s\cdot h'_m$. These are evaluated at the same ω, which can be selected at the DFT peak corresponding to the sinusoid. The following chart gives the data flow for calculating $S_w^{(k)}$, k=1, …, K:

$$\begin{array}{ccccccccccc}
S_w & & S_{w'} & & S_{w''} & \cdots & \cdots & & S_{w^{(K-1)}} & & S_{w^{(K)}} \\
\downarrow & \swarrow & \downarrow & \swarrow & \downarrow & & & \swarrow & \downarrow & \swarrow & \\
S'_w & & S'_{w'} & & S'_{w''} & \cdots & \cdots & & S'_{w^{(K-1)}} & & \\
\downarrow & \swarrow & \downarrow & \swarrow & \downarrow & & & \swarrow & & & \\
S''_w & & S''_{w'} & & S''_{w''} & \cdots & \iddots & & & & \\
\vdots & & \vdots & & \vdots & \iddots & & & & & \\
\vdots & & \vdots & & \iddots & & & & & & \\
\downarrow & \swarrow & \downarrow & \swarrow & & & & & & & \\
S_w^{(K-1)} & & S_{w'}^{(K-1)} & & & & & & & & \\
\downarrow & \swarrow & & & & & & & & & \\
S_w^{(K)} & & & & & & & & & &
\end{array}$$

Figure 1: *Data flow for calculating* $S_w^{(k)}$, k=1, …, K.

The first row of the chart is evaluated directly by Fourier transform (see 3.2). The second row can be evaluated either by (17a) or (17b). The rest K-1 rows can only be evaluated by (17b). w and its derivatives must be 0 at $-T$ and T up to order K-2 (or K-1 if (17a) is to be avoided). The computation of $X_{m\,w}^{(k)}$ is very similar to the above, with one less differentiation.

3.2. Discrete Fourier transform

The first row in Figure 1 contains Fourier transforms in the form of

$$X^c(\omega)=\int_{-T}^{T} x(\tau)e^{-j\omega\tau}d\tau \tag{20a}$$

where x has the form $sw^{(k)}$ or $sh'_m w^{(k)}$. Here the superscript "c" stands for "continuous". x is only known in its sampled form $x_n=x(n)$, $n\in\mathbf{Z}$. We approximately calculate (20a) by

$$X(\omega)=\sum_{n=-T}^{T} x_n e^{-j\omega n} \tag{20b}$$

(20b) is evaluated only at the selected angular frequency ω. The relation between $X^c(\omega)$ and $X(\omega)$ is given by the sampling theorem as

$$X(\omega)-X^c(\omega)=\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} X^c(\omega-2k\pi) \tag{21}$$

According to (21), for $X(\omega)$ to be a good approximation of $X^c(\omega)$, X^c must decay fast enough to be absolutely integrable. We rewrite (17b) as

$$X_{w^{(l)}}^{(k)}+X_{w^{(l+1)}}^{(k-1)}=j\omega X_{w^{(l)}}^{(k-1)}\,. \tag{22}$$

(22) shows that the decay rate of $X_{w^{(l)}}^{(k)}$ is roughly outlined by l+k. With each increase in l+k, the decay becomes slower by order of 1. To guarantee that $X_{w^{(l)}}^{(k)}$ is absolutely integrable, X_w needs to have a decay rate faster than $\omega^{-(l+k+1)}$. Accordingly xw (and therefore w) needs to be l+k times continuously differentiable. In the context of Figure 1, w and its derivatives up to order K are required to remain zero at $-T$ and T.

Apart from global integrability, we also require that X^c has significant decay at $\omega\pm2\pi$. For slow-varying complex sinusoids this is automatically satisfied if ω is selected near the instantaneous frequency, so that $\omega\pm2\pi$ are about twice the Nyquist frequency from the sinusoid. For real sinusoids it is desirable that ω be not too close to 0 or π to avoid the conjugate partial being picked up in (21).

3.3. More on window functions

Both in 3.1 and 3.2 we have raised the continuity issue of the window function w, requiring w be continuously differentiable up to a certain order. While in 3.1 it appears as a technical convenience for adopting the integration-by-parts method, in 3.2 it concerns the accuracy of numerical computation. The requirement raised in 3.2 applies to the derivative methods in general as long as the spectrum is calculated from discrete samples, no matter how the differentiation is implemented. Since a lower level of continuity is demanded in 3.1 than in 3.2, we see that the integration by parts does not introduce extra constraint to w.

When the order of differentiation is high the continuity requirement of w rules out all the commonly used window functions. Rectangular, Hamming, Bartlett, Gaussian and Kaiser windows are either discontinuous or C^0 functions, therefore should not be used in the derivative method. Hann and Blackman windows are C^1 functions, so that they should not be used in derivative methods involving 2nd-order derivatives. Window functions of higher continuity level are rarely seen in the literature.

Highly continuous window functions can be constructed by multiplying or convolving less continuous windows. The multiplication method aims at directly eliminating discontinuity at the vanishing ends of w, i.e. $-T$ and T. In fact, if w_1 and w_2 are window functions that vanish at $-T^+$ with vanishing moments k_1 and k_2, i.e.

$$w_1(-T+\tau)=a_1\tau^{k_1}+O(\tau^{k_1+1}),\ w_2(-T+\tau)=a_2\tau^{k_2}+O(\tau^{k_2+1}), \tag{23a}$$

then for w=w_1w_2 we have

$$w(-T+\tau)=a_1a_2\tau^{k_1+k_2}+O(\tau^{k_1+k_2+1})\ , \tag{23b}$$

i.e. w vanishes at $-T^+$ with vanishing moment k_1+k_2. The multiplication method does not improve continuity inside ($-T$, T), so both w_1 and w_2 ought to have sufficient interior continuity desired of w. This method preserves window length and increases time concentration, at the cost of a wider bandwidth. On the other hand, the convolution method aims at accelerating decay of the window spectrum as $\omega\to\pm\infty$. If the spectra of w_1 and w_2 decay like $|\omega|^{-k_1}$ and $|\omega|^{-k_2}$, respectively, then the convolution theorem states that the spectrum of w=w_1*w_2 decays like $|\omega|^{-(k_1+k_2)}$. The convolution method improves continuity on the whole window support. It yields a longer window but increases frequency concentration.

Example 9. Multiplying a window function on [-1, 1] with the half cosine window (a C^0 function)

$$w_{\cos}(t)=\begin{cases}\cos\pi(t/2), & -1\le t\le 1,\\ 0, & otherwise\end{cases} \tag{24}$$

increases the order of continuity by 1. Multiplying $w_{\cos}$ with itself yields the Hann window, which is a C^1 function. By repetitively multiplying the Hann window with itself we get a series of window functions we denote as Hann^k, k=0, 1, …, Hann^0 being the rectangular window. This $\{\text{Hann}^k\}_k$ series make up the basis of the cosine window family on [-1, 1], in which Hann^k is the simplest function to offer C^{2k-1} continuity.■

Example 10. Convolving a window function with the rectangular window (a discontinuous but bounded function) increases the order of continuity by 1. Convolving the rectangular window on [-1, 1] with itself yields the triangular window on [-2, 2], which is a C^0 function. Convolving the latter with the same rectangular window yields a piecewise parabolic C^1 function

$$w(t)=\begin{cases}(3-t^2)/2, & |t|\le 1,\\ (3-|t|)^2/4, & 1<|t|\le 3,\ \blacksquare\\ 0, & otherwise.\end{cases} \tag{25}$$

3.4. Extra equations

Given the general derivative method in (14b) or (14c), we notice that the number of real equations is always even. Therefore if the number of real unknown parameters to solve is odd, we need to drop one equation from the system. In this paper we only drop from the two real equations obtained by the highest order of derivative taken, and of these two we always choose to drop the one that yields the smaller determinant of the coefficient matrix of the linear system. This can be conveniently incorporated into the Gaussian elimination or LU factorization method for solving the linear system without significantly adding to the computation load.

An alternative to dropping extra equation is solving the linear system in the least square sense: i.e. instead of the overdetermined linear system **Ar**=**b**, we solve $\mathbf{A}^T\mathbf{A}\mathbf{r}=\mathbf{A}^T\mathbf{b}$. In the ideal case that the signal strictly obeys (4) and the numerical computation incurs no error, the least square solution is accurate (i.e. the square error is 0), and is equivalent to the solution obtained by dropping the extra equation.

3.5. Amplitude and phase angle

The coefficient r_0, which represents global amplification and phase shift, is not involved in the linear system derived from (14b) or (14c). Once r_1, …, r_{M-1} have been estimated, various methods can be engaged in evaluating r_0, one of which is by comparing the spectra of s with and without the contribution from r_0:

$$e^{r_0}=\frac{S_w(\omega)}{\int e^{\sum_{m=1}^{M-1}r_mh_m(\tau)}w(\tau)e^{-j\omega\tau}d\tau} \tag{26}$$

Again the integral is evaluated from discrete data with ω selected near the instantaneous angular frequency.

3.6. Extension to multiple frames

The derivative method discussed above focuses on one data frame only. If the signal model can be assumed to be stable over the span of more than one frame, then it is possible to avoid taking high-order derivatives by applying (14b) to multiple frames with the same unknown parameters.

4. TEST RESULTS

We test the model-based derivative method on synthesized real sinusoids. In the first part of the tests the sinusoids are synthesized on signal model (4), and the derivative method evaluates the parameters using the true model setting. In the second part the parameter variations are not known to the estimator, in which case the derivative method yields an approximate result from within the modelled signal space. For all tests the window length of 1024 is used.

4.1. Estimation with the correct model setting

The performance of sinusoid estimation is evaluated by signal-to-residue ratio (SRR), where the residue is calculated by subtracting a sinusoid frame constructed from the estimated signal model from the original sinusoid. The SRR is formulated as

$$SRR=\frac{\sum_{n=-T}^{T}w_ns_n^2}{\sum_{n=-T}^{T}w_n(s_n-\hat{s}_n)^2}\ ,\ \hat{s}_n=e^{\hat{R}(n)}\ ,\ \hat{R}(t)=\sum_{m=0}^{M-1}\hat{r}_mh_m(t)\ , \tag{27}$$

where $\hat{r}_m$, m=0, …, M-1, are the model parameter estimates, and w is a Hann window used to emphasize the contributions of both signal and residue at the frame centre.

The first test set contains exponentially enveloped linear chirps (signal model M=3, h_0=1, $h_1(t)$=t, $h_2(t)$=$0.5t^2$, p_2=0). A total number of 3000 frames are included in this set, with p_0=0, 10 q_0 values uniformly selected between 0 and 0.45π, 10 p_1 values between 0 and 0.0045, 10 q_1 values between 255 and 255.9 bins (1 bin=$\pi/2^9$) , and 10 q_2 values between 0 and 27 bin/frame (1bin/frame=$\pi/2^{19}$). To limit the number of frames, p_1 and q_2 do not vary independently, but appear in three groups: an AM-only group in which q_2=0, an FM-only group in which p_1=0, and an AM-FM group in which p_1 and q_2 are paired up in order. This test set is summarized in Table 1. 0dB Gaussian white noise is applied to test performance under noisy environment.

Table 1: *Test set 1*

Group	i_0, i_1=0~9	p_0= 0, q_0=0.05$\pi \cdot i_0$, q_1=255+0.1$\cdot i_1$(bin)
1	i=0~9	p_1=0.0005i, q_2=0
2		p_1=0, q_2=3i (bin/frame)
3		p_1=0.0005i, q_2=3i (bin/frame)

(M=3, h_0=1, $h_1(t)=t$, $h_2(t)=0.5t^2$, p_2=0.)

The result (marked "D") is compared against the Abe-Smith QIFFT (quadratically interpolated fast Fourier transform) method [5] which was designed for the same signal model. In this paper the QIFFT method is implemented in both its original version in [5] ("QI") and an enhanced version in [11] ("EQI"). A Gaussian window is required for both QIFFT methods. Theoretically the derivative method yields very accurate result if the proper window is used, while the QIFFT methods may suffer moderately from the truncation of the Gaussian window at both ends. In the test we use Hann^2 window in the derivative method. Test results for test set 1 are given in Figure 2. For clean signals both the derivative and enhanced QIFFT methods work very well, with the worst SRR above 70dB. There is a gap between 70~100dB between the two methods, which can be attributed to the truncation of Gaussian window in the QIFFT method. Both methods appear to be susceptible to noise. The derivative and original QIFFT methods show similar results with 0dB noise, while the enhanced QIFFT method does not yield consistent results, partially due to random errors of p_1 being accumulated in the enhancement step.

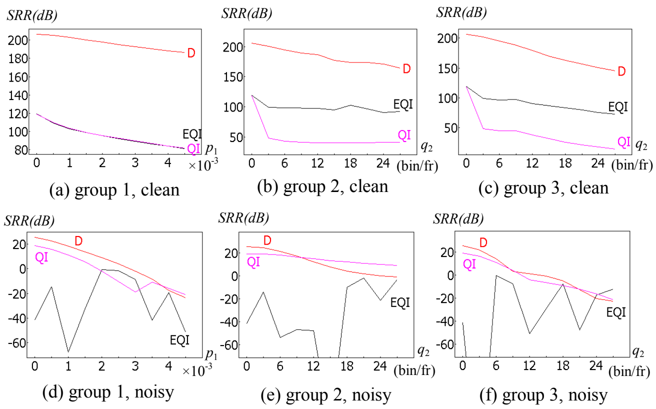


Figure 2: *Results for test set 1.*

We run a separate test on the same test set to illustrate the importance of window continuity to the derivative method. We apply the method on clean and noisy data with 4 window functions with different order of continuity, namely Hamming window the discontinuous, Hann window of C^1, Hann$^{1.5}$ window of C^2, and Hann^2 window of C^3. Results on clean data are given in Figures 3 (a)~(c), which show Hann^2>Hann$^{1.5}$>Hann>Hamming consistently. Results on noisy data are given in (d)~(f), showing very close performance for all window functions (so much so that we are not always able to mark them out in the figures). This indicates that the difference between window functions, as a cause of bias, is overridden by the noise.

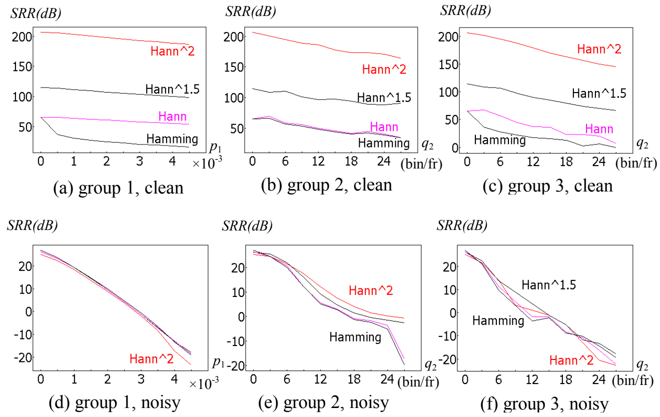


Figure 3: *Comparing windows for test set 1.*

The second test set contains sinusoids frequency-modulated by a sinusoidal modulator with frequency ω_M (signal model M=4, h_0=1, $h_1(t)=t$, $h_2(t)=\cos \omega_M t$, $h_3(t)=\sin \omega_M t$, $p_1=p_2=p_3=0$). A total number of 6000 frames are included in this set, with ω_M=0.001, p_0=0, 10 q_0 values uniformly selected between 0 and 0.45π, 5 q_1 values between 255 and 255.8 bins (1 bin=$\pi/2^9$), 10 modulator angular frequency ω_M between 0.333×10^{-3} and 6.333×10^{-3}, 10 modulator amplitudes a_M from 0 to 27 bins (1bin=$\pi/2^9$), and 6 modulator phase angles φ_M between 0 and π/2, $q_2=a_M\cos\varphi_M$, $q_3=a_M\sin\varphi_M$. To limit the number of frames, ω_M and a_M do not vary independently, but appear in two groups: in the first group ω_M is fixed at 10^{-3}, in the second group a_M is fixed at 9 bins. This test set is summarized in Table 2.

Table 2: *Test set 2*

i_0=0~9, i_1=0~4		p_0= 0, q_0=0.05$\pi \cdot i_0$, q_1=255+0.3$\cdot i_1$(bin)
i_2=1~10	group 1	ω_M=10^{-3}, a_M=3$\cdot i_2$ (bin)
	group 2	ω_M=10$^{-3}\cdot$(2$\cdot i_2$-1)/3, a_M=9 (bin)
i_3=0~5		φ_M=0.1$\pi \cdot i_3$, $q_2=a_M\cos\varphi_M$, $q_3=a_M\sin\varphi_M$

(M=4, h_0=1, $h_1(t)=t$, $h_2(t)=\cos 10^{-3}t$, $h_3(t)=\sin 10^{-3}t$, $p_1=p_2=p_3=0$.)

Results for the second test set are not compared against another estimator, as there has been none reported to support this signal model. Figure 4 compares test results obtained using the four window functions listed above, with the modulation extent and angular rate as the x axes, respectively. Again we have accurate result when the window function has sufficient order of continuity.

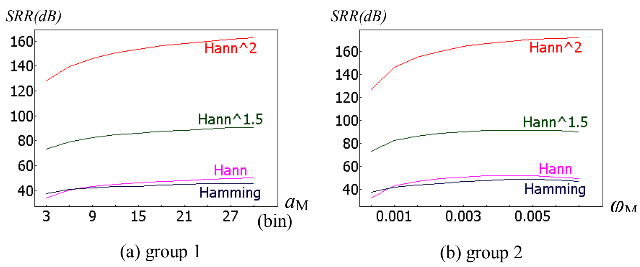


Figure 4: *Comparing windows for test set 2.*

4.2. Approximation of unknown model settings

In the second part of the test we show how the derivative method behaves when applied to sinusoids whose $R(t)$ model is not known to the estimator. In this case it is generally impossible to select a limited number M of functions h_0, …, h_{M-1} so that R is their linear combination. However, if R is close enough to the linear space with basis $\{h_0, …, h_{M-1}\}$, then it is possible to approximate $R(t)$ by a linear combination of the basis functions. In the following we use a simple example to illustrate this.

In this example we let $h_m(t)=t^m$, so that the afore-mentioned function space is that of polynomials of orders up to M-1. The ability of polynomials to locally approximate arbitrary smooth functions is proved by Taylor's theorem. For test signals we use a frequency modulated sinusoid with optional accompanying amplitude modulation:

$$s_1(t) = \cos(\omega_0 t + \delta_c \cos\omega_M t + \delta_s \sin\omega_M t) \tag{28a}$$

$$s_2(t) = \left(1.5 + \cos(\omega_M t + 1)\right)\cos(\omega_0 t + \delta_c \cos\omega_M t + \delta_s \sin\omega_M t) \tag{28b}$$

where ω_M=0.005, ω_0=255 bin, δ_c=6 bin and δ_s=8 bin. In (28a) and (28b) both frequency and amplitude modulators are sinusoids with angular frequency ω_M. Obviously neither signal fits into any polynomial frequency-and-log-amplitude model.

With the above settings we run the derivative method with M=2, …, 7. A Hann^2 window is used in this test. To avoid taking derivatives above 3rd order, we apply the derivative method at two frames (see 3.6), with the second frame being 256 point (i.e. 25%) shifted from the first, therefore 1280 data points are needed here instead of 1024. Equal numbers of derivatives are taken from both frames. Possible extra equations are handled by the least square approach (see 3.4). For s_1 we solve the system assuming p_m=0, $\forall m$; for s_2 no such restriction is applied.

Figure 5 shows progressively the approximation of the instantaneous frequency of s_1 by polynomials, achieved by the derivative method. The dotted curve (R) is the true frequency, while the solid curves (Rx) are the estimated approximations. The value x in "Rx" is the degree of the polynomial used for approximating the instantaneous frequency. The time span in these figures is -512~512. In general, the more polynomials we use, the closer the frequency estimate is to the true frequency. We observed that the approximation is better in right half. This can be attributed to the use of a second frame centred at 256.

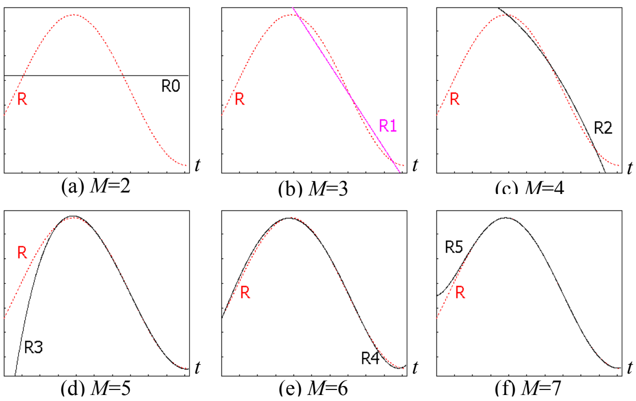


Figure 5: *Approximating frequency modulation*

Figure 6 shows the approximation progress of the instantaneous frequency (a~f) and amplitude (g~l) of signal s_2. For amplitude results the x in "Rx" is the degree of polynomial, used for approximating the logarithmic amplitude, minus 1. The results are similar to those observed in Figure 5, with moderately higher bias due to the doubling of the number of unknown variables.

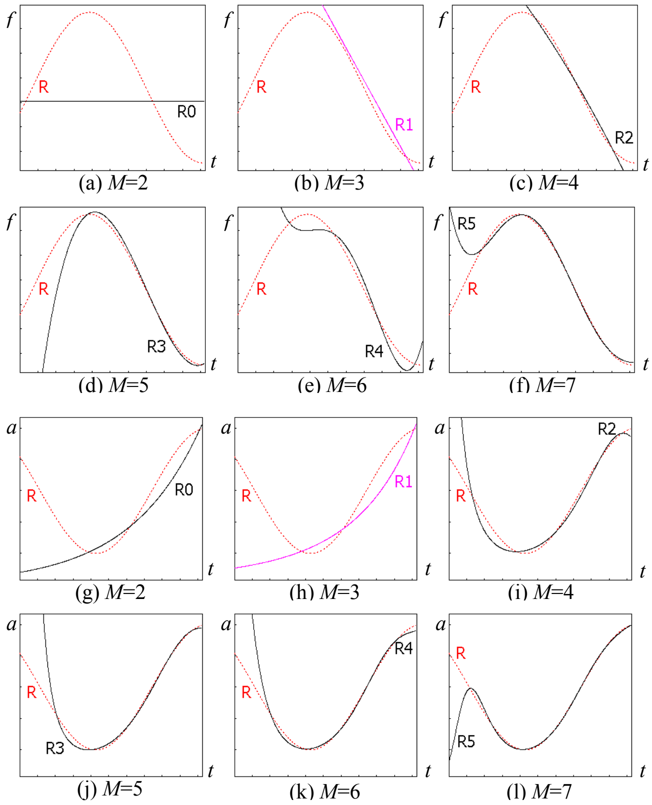


Figure 6: *Approximating amplitude and frequency modulation*

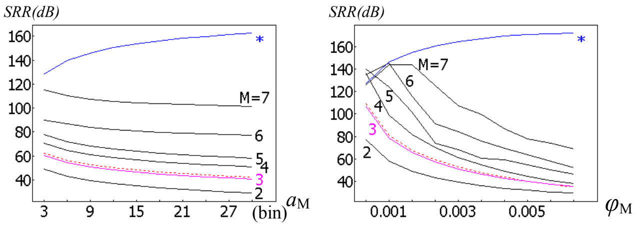
We run a numerical experiment on frequency-modulated sinusoids to test the approximation performance using the polynomial basis. A test set 3, based on test set 2 but with a higher modulation rate in group 1, is used for this purpose. Again the derivative method is implemented to take 2 frames with 25% overlap, and the least-square approach in 3.4 is used to handle possible extra equations. The Hann^2 window is used for all values of M. Results are given in SRR, calculated from the first frames (1024 samples) only.

The approximation results are depicted in Figure 7, where we have marked the curves by the different values of M, from 2 to 7. For comparison we also run the enhanced Abe-Smith method, whose results we plot in dotted curves, and the derivative method using the exact model setting, whose results we mark by "*". For both test groups the SRR improves consistently as M becomes larger. The performance of Abe-Smith method is very close to the derivative method with M=3, which is very reasonable as for M=3 the two methods actually have the same signal model. The derivative method with exact model setting, unsurprisingly, has the best performance on both test groups.

Table 3: *Test set 3*

i_0=0~9, i_1=0~4		p_0= 0, $q_0=0.05\pi\cdot i_0$, $q_1=255+0.3\cdot i_1$(bin)
i_2=1~10	group 1	$\omega_M=3\cdot10^{-3}$, $a_M=3\cdot i_2$ (bin)
	group 2	$\omega_M=10^{-3}\cdot(2\cdot i_2-1)/3$, a_M=9 (bin)
i_3=0~5		$\varphi_M=0.1\pi\cdot i_3$, $q_2=a_M\cos\varphi_M$, $q_3=a_M\sin\varphi_M$

(M=4, h_0=1, $h_1(t)=t$, $h_2(t)=\cos10^{-3}t$, $h_3(t)=\sin10^{-3}t$, $p_1=p_2=p_3$=0.)



(a) group 1 (b) group 2

Figure 7: *Results for approximating vibratos*

5. CONCLUSIONS

In this paper we have reviewed the derivative method for non-stationary sinusoid estimation in a generalized signal model (4). We have shown that by modelling the complex logarithm of the signal s as the linear combination of a limited number of basis functions, it is possible to obtain a linear system of the combination weights by taking derivatives. The coefficients of the linear system are short-time Fourier transforms involving s, the basis functions and their derivatives. For the evaluation of these coefficients an integration-by-parts method is applied to transfer the differentiation from s to the window function, whose derivatives are known by design. The derivative method requires the window function be continuously differentiable up to a certain order, for which we have discussed various ways to obtain highly continuous windows. Tests show that the derivative method yields accurate results for clean sinusoids with the correct model setting, and is able to approximate sinusoids of unknown model type with a sufficiently large polynomial basis.

On the other hand, despite the improved flexibility of (4), the derivative method is still model-based, and therefore suffers from deficiencies typical to all model-based estimators, such as potential overfitting for sinusoids outside the modelled signal space. Modelling error due to these deficiencies can be compensated by engaging non-parametric error control methods, such as the one in [11], which is designed to work with arbitrary estimators.

6. ACKNOWLEDGMENTS

This work was supported by the EPSRC EP/E017614/1 project OMRAS2 (Online Music Recognition and Searching).

7. REFERENCES

[1] R.J. McAulay and T. F. Quatieri, “Speech analysis/synthesis based on a sinusoidal representation,” *IEEE Tran. Acous. Sp. Sig. Proc.*, vol.34, no.4, 1986, pp. 744-754.

[2] F. Keiler and S. Marchand, “Survey on extraction of sinusoids in stationary sounds,” in *Proc. DAFx’02*, Hamburg, 2002.

[3] P. Masri and N. Canagarajah, “Extracting more detail from the spectrum with phase distortion analysis,” in *Proc. DAFx’98*, Barcelona, 1998.

[4] A. Röbel, “Frequency slope estimation and its application for non-stationary sinusoidal parameter estimation,” in *Proc. DAFx’07*, Bordeaux, 2007.

[5] M. Abe and J. O. Smith III, “AM/FM rate estimation for time-varying sinusoidal modeling,” in *Proc. ICASSP’05*, Philadelphia, 2005.

[6] S. Marchand and P. Depalle, “Generalization of the derivative analysis method to non-stationary sinusoidal modeling,” in *Proc. DAFx’08*, Espoo, 2008.

[7] A. Röbel, “Estimating partial frequency and frequency slope using reassignment operators,” in *Proc. ICMC’02*. Göteborg. 2002.

[8] J. C. Brown and M. S. Puckette, “A high resolution fundamental frequency determination based on phase changes of the Fourier transform,” *J. Acous. Soc. Am.*, vol.94, no.2, 1993, pp.662-667.

[9] M. Desainte-Catherine and S. Marchand, “High precision Fourier analysis of sounds using signal derivatives,” *J. AES*, vol.48 no.7/8, 2000, pp.654-667.

[10] F. Auger and P. Flandrin, “Improving the readibility of time-frequency and time-scale representations by the reassignment method,” *IEEE Tran. Sig. Proc.*, vol.43(5), 1995, pp.1068-1089.

[11] Wen X. and M. Sandler, “Evaluating parameters of time-varying sinusoids by demodulation,” in *Proc. DAFx’08*, Espoo, 2008.