Connections between Reassigned Spectrum and Least Squares Estimation for Sinusoidal Models

Yannis Pantazis
Institute of Applied and Computational Mathematics, *FORTH*
Heraklion, Greece
pantazis@iacm.forth.gr

Vassilis Tsiaras
Department of Computer Science
University of Crete
Heraklion, Greece
tsiaras@csd.uoc.gr

Yannis Stylianou
Department of Computer Science
University of Crete
Heraklion, Greece
yannis@csd.uoc.gr

***Abstract*—The parameter estimation of sinusoidal signals, especially the frequency estimation is for decades a very challenging problem. Among the various frequency estimation methods, this paper compares and connects the reassigned spectrum and an iterative, nonlinear Least Squares method referred to as iQHM (iterative Quasi Harmonic Model). Interestingly, there are subtle connections between these two –seemingly different– iterative methods both in frequency as well in time domain. Moreover, inspired by the optimal performance of reassigned spectrum for mono-component sinusoidal signals, a variant of iQHM is proposed. The new method improves the performance of the original iQHM approach in frequency estimation by increasing the region of convergence by 40% on average.**

***Index Terms*—Sinusoidal models, Reassigned spectrum, Quasi-harmonic model, Nonlinear least squares**

I. Introduction

Sinusoidal representation is a powerful tool for analysis, synthesis, decomposition and transformation of various periodic signals such as speech, audio, medical signals, animal sounds, etc. Non-parametric approaches [1]–[4] aim to estimate the sinusoidal parameters based on various spectral transformations. On the other hand, parametric methods [5]–[7] have also been proposed for the estimation of the time-varying amplitudes and frequencies in multi-component sinusoidal signals. It is well known that the hardest task in sinusoidal parameter estimation is the inference of the frequency values due to their nonlinear nature. Additionally, the attraction region for unbiased frequency estimation is inversely proportional to the utilized window length which results in a trade-off between accuracy and convergence.

In [5], [8], authors developed a model, which is referred to as Quasi-Harmonic Model (QHM), based on Least Squares (LS) method for the representation of sinusoidal signals which accurately estimates the frequency parameters. In a correspondence with P. Babu and P. Stoica [9], [10], an iterative version of QHM (iQHM) was shown to be (almost) equivalent to the nonlinear Least Squares frequency estimation approach using Newton-Gauss (NG) method. That correspondence motivated us to search for connections and relevance between iterative LS method (either formulated as iQHM or NG method) with other iterative frequency estimation methods.

Another iterative frequency estimation method is Reassigned Spectrum (RS) which is a general post-processing method applied on time-frequency representations [1], [11]. In RS, time and frequency relocation is performed based on the center of gravity of the representation. Hence, the overall resolution of the representation is improved which is crucial especially for representations such as spectrogram where there is a trade-off between time resolution and frequency resolution. Restricting the space of all signals to the sinusoidal subspace, we observe that both time and frequency relocation of RS have similar structure with iQHM. Qualitatively, in both iterative methods, frequency estimation is performed by weighting the sinusoidal signal with an odd function. While the two methods do not coincide in general for the frequency re-estimation/relocation procedure, we show here that when that weighting function is a Gaussian window, then the two methods provide the same formula for the frequency re-estimation/relocation procedure. On the other hand, time relocation restricted to sinusoidal signals has exactly the same formulation for both methods.

The application of RS to a mono-component sinusoidal signal reveals that the frequency relocation provides almost perfect frequency estimation in the noiseless case. Moreover, since RS is a method based on the center of the gravity of the time-frequency representation, we expect that for multi-component sinusoidal signals with different amplitudes, the frequency estimation to be biased against the weaker sinusoids. Based on these observations as well on the fact that the Least Squares method decouples the sinusoids in the multi-component case when the sample size is sufficiently large, we developed a variant of iQHM inspired by the frequency relocation of RS which is named reassigned spectrogram QHM (rsQHM). To our knowledge, this is the first time that RS is used for *parametric modeling*. Simulations on multi-component sinusoidal signals showed that rsQHM enjoyed on average 40% larger area of convergence as well it attained the Cramer-Rao lower bound for both amplitudes and frequencies.

The organization of the paper is as follows. Section II introduces iQHM and RS methods for frequency and time estimation/relocation and continues by exploring their similarities and dissimilarities. Then, a variant of iQHM inspired by RS is introduced in Section III. We then validate the new modelling approach in a series of experiments. Section IV concludes the paper.

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

II. Sinusoidal Models, Representations and their Connection

Lets consider a complex-valued discrete-time signal, $x(n)$, whose short time Fourier transform (STFT) is defined by

$$X(m,\omega) = \sum_{n=-N}^{N} w(n)x(n+m)e^{-j\omega(n+m)} \tag{1}$$

where $w(n)x(n+m)$ is the windowed frame of the signal centered at time instant m while $w(n)$ is a real symmetric window function with support in $\{-N,\ldots,N\}$.

The common model-basis for comparison between iQHM and RS is the multi-component sinusoidal signal representation, however, it is intractable to study them analytically. On the other hand, for the case of mono-component sinusoidal signals, the analytic calculations can be performed. Thus, in the following, we assume that the frame of $x(n)$, centered at time instant m, is given by

$$x(n+m) = c_1 e^{j\omega_1(n+m)}, \quad n=-N,\ldots,N \tag{2}$$

where $c_1 = c_1(m)$ is the complex amplitude while $\omega_1 = \omega_1(m)$ is the angular frequency.

A. Iterative Least Squares (iQHM)

In iQHM formulation [5], a frame of the signal centered at time instant m is modeled by

$$\tilde{x}(n+m) = (a_1 + nb_1)e^{j\tilde{\omega}_1 n}, \quad n=-N,\ldots,N \tag{3}$$

where $\tilde{\omega}_1 = \tilde{\omega}_1(m)$ is an initial estimate of the frequency, $a_1 = a_1(m)$ is the complex amplitude and $b_1 = b_1(m)$ is the complex slope. Assuming $\tilde{\omega}_1$ is known, a_1 and b_1 are computed by minimizing the weighted sum of squared error

$$\epsilon = \sum_{n=-N}^{N} w(n)(x(n+m) - \tilde{x}(n+m))^2 \tag{4}$$

Minimization of ϵ provides the LS estimates

$$\begin{aligned} a_1 &= \frac{1}{W_0}\sum_{n=-N}^{N} w(n)x(n+m)e^{-j\tilde{\omega}_1 n} \\ b_1 &= \frac{1}{W_2}\sum_{n=-N}^{N} nw(n)x(n+m)e^{-j\tilde{\omega}_1 n} \end{aligned} \tag{5}$$

with W_k being the kth moment of the window (i.e., $W_k = \sum_{n=-N}^{N} n^k w(n),\ k=0,1,\ldots$).

Assuming that the complex amplitude, a_1, is non-zero, the decomposition (or projection) of the complex slope, b_1, into a parallel and an orthogonal component of the complex amplitude defined by

$$b_1 = \rho_{1,1}a_1 + \rho_{1,2}ja_1 \tag{6}$$

provides a mean to estimate the frequency mismatch error. Indeed, it has been shown in [8] that frequency mismatch error defined as the difference between the true frequency and the initially provided (i.e., $\omega_1 - \tilde{\omega}_1$), is estimated from $\rho_{1,2}$. Therefore, an updating procedure for the frequency estimation problem is given by

$$\hat{\omega}_1 = \tilde{\omega}_1 + \rho_{1,2} \tag{7}$$

This iterative procedure was shown to be equivalent to the nonlinear LS frequency estimation problem based on NG method [9], [10].

Coefficient $\rho_{1,1}$, which in the context of iQHM accounts for the linear amplitude slope of the signal, is given by [8]

$$\rho_{1,1} = \Re\left\{\frac{b_1}{a_1}\right\} = \frac{W_0}{W_2}\Re\left\{\frac{\sum_{n=-N}^{N} nw(n)x(n+m)e^{-j\tilde{\omega}_1 n}}{\sum_{n=-N}^{N} w(n)x(n+m)e^{-j\tilde{\omega}_1 n}}\right\} \tag{8}$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex number, respectively. Coefficient $\rho_{1,2}$, which, as already mentioned above, provides an estimate of the frequency mismatch, is given by

$$\rho_{1,2} = \Im\left\{\frac{b_1}{a_1}\right\} = \frac{W_0}{W_2}\Im\left\{\frac{\sum_{n=-N}^{N} nw(n)x(n+m)e^{-j\tilde{\omega}_1 n}}{\sum_{n=-N}^{N} w(n)x(n+n)e^{-j\tilde{\omega}_1 n}}\right\} \tag{9}$$

B. Reassigned Spectrum

In the reassignment method, a time-frequency representation of the signal is sharpened by appropriately post-processing the time-frequency representation. Two central operations, namely, time relocation and frequency relocation are performed. Time relocation is a displacement of the energy in time which depends on the center of gravity of the windowed signal in time-domain. When the time-frequency representation is the spectrogram, time relocation at coordinates (m,ω) corresponds to the estimate of the local group delay and it is given by [1]

$$\hat{m} = m - \Re\left(\frac{X_{\mathcal{T}w}(m,\omega)}{X(m,\omega)}\right) \tag{10}$$

where $X_{\mathcal{T}w}(n,\omega) = -\sum_{n=-N}^{N} nw(n)x(n+m)e^{-j\omega(n+m)}$ is the STFT of $x(n)$ with tilted "window" function $nw(n)$. Thus,

$$\hat{m} = m + \Re\left(\frac{\sum_{n=-N}^{N} nw(n)x(n+m)e^{-j\omega n}}{\sum_{n=-N}^{N} w(n)x(n+m)e^{-j\omega n}}\right). \tag{11}$$

Frequency relocation is a displacement of the energy in frequency which again depends on the center of gravity of the windowed signal but now the operation is performed in frequency-domain. When the time-frequency representation is the spectrogram, frequency relocation at coordinates (m,ω) corresponds to local estimate of instantaneous frequency and it is given by [1]

$$\hat{\omega} = \omega + \Im\left(\frac{X_{\mathcal{D}w}(m,\omega)}{X(m,\omega)}\right) \tag{12}$$

where $X_{\mathcal{D}w}(m,\omega) = -\sum_{n=-N}^{N} w'(n)x(n+m)e^{-j\omega(n+m)}$ is the STFT of $x(n)$ with "window" function the discretized version of the derivative of the respective continuous-time window, $w(t)$. Therefore,

$$\hat{\omega} = \omega - \Im\left(\frac{\sum_{n=-N}^{N} w'(n)x(n+m)e^{-j\omega n}}{\sum_{n=-N}^{N} w(n)x(n+m)e^{-j\omega n}}\right). \tag{13}$$

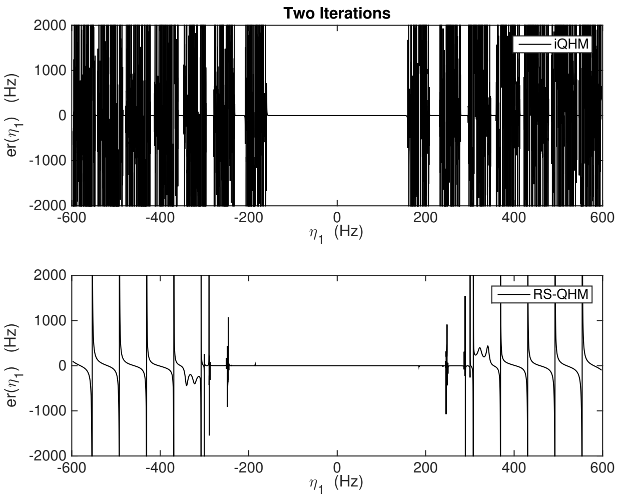

Fig. 1. Frequency estimation error for iQHM (upper panel) and rsQHM (lower panel) with $\eta_1 := \omega_1 - \tilde{\omega}_1$ being the initial frequency mismatch. Hann window was used and two iterations were performed.

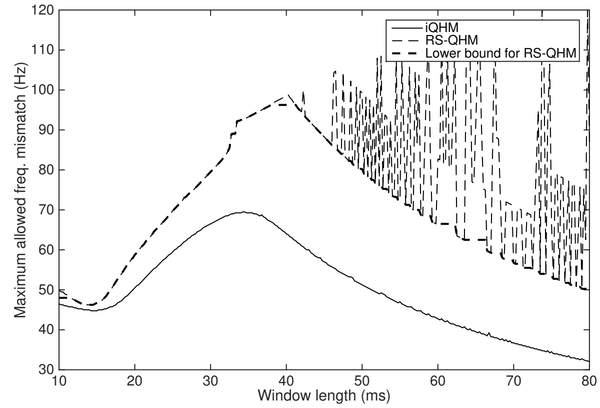

Fig. 2. Maximum allowed frequency mismatch in multi-component signals for iQHM (solid line) and rsQHM (dashed line) utilizing Hann window.

Remark: In equations (8), (9), (11) and (13), the denominator of the fraction, which equals to $X(m,\omega)$, may take the value 0 resulting in ill-posed formulation. This erroneous behavior is evident at Fig. 1 where the frequency estimation diverges at these particular points. However, this issue has been tackled in RS by not performing relocation when $|X(m,\omega)|$ is below a small threshold value while in iQHM it was tackled by considering only a region (of attraction) where no zeros of $X(m,\omega)$ are present.

C. Connection between iQHM and RS

Comparing (8) and (11), we observe that the time relocation of the reassignment method can be written as

$$\hat{m} = m + \frac{W_2}{W_0}\rho_{1,1} \tag{14}$$

Notice that this result is independent of the underlying signal, $x(n)$. Thus, the term $\frac{W_2}{W_0}\rho_{1,1}$ is equal to the local group delay. The group delay has been used, among others, for spectrum estimation [12] and for determining the instants of significant excitation in speech signals [13].

The comparison between (7), (9) and (13) is more involved. Equations (7) and (13) are identical only when

$$w'(n) = -\kappa n w(n) \Rightarrow w(n) = e^{-\frac{\kappa}{2}n^2} \tag{15}$$

where $\kappa = \frac{W_0}{W_2}$. Hence, for Gaussian windows iQHM and RS have similar frequency update rules and the frequency relocation of the reassignment method can be written as

$$\hat{\omega} = \omega + \frac{W_2}{W_0}\rho_{1,2} \tag{16}$$

Therefore, for Gaussian windows one iteration is enough for the iQHM method in order to converge. However, for any other window type the frequency update of iQHM and RS are different.

III. RSQHM

A. Motivation

RS is a general nonparametric method for improving the localization of the spectrogram. If we restrict RS to the case of (mono-component) noise-free, sinusoidal model, frequency relocation is perfect. Indeed, the spectrogram for the sinusoidal signal is given by

$$X(m,\omega) = c_1 e^{j(\omega_1-\omega)m}\sum_{n=-N}^{N} w(n)e^{j(\omega_1-\omega)n} \tag{17}$$

while the spectrogram of the signal with window function the discretized derivative of the window is given by

$$\begin{aligned} X_{\mathcal{D}w}(m,\omega) &= \sum_{n=-N}^{N} w'(n)e^{j(\omega_1-\omega)(n+m)} \\ &\approx j(\omega-\omega_1)c_1 e^{j(\omega_1-\omega)m}\sum_{n=-N}^{N} w(n)e^{j(\omega_1-\omega)n} \end{aligned} \tag{18}$$

Substituting (17) and (18) into (13), we obtain that

$$\hat{\omega} \approx \omega - \Im(j(\omega-\omega_1)) = \omega + \omega_1 - \omega = \omega_1$$

Therefore the reassignment method achieves almost perfect relocation of the frequency in one step. The induced numerical error is inverse proportional to the sampling frequency while it is noteworthy that the frequency relocation is perfect when the analysis is performed in continuous time.

B. Formulation of rsQHM

Motivated by the fact that RS for sinusoidal signals provides perfect frequency relocation, we suggest varying QHM so as to approach closer to the reassignment method. Thus, the variant of QHM is defined by

$$\tilde{x}(n) = \left(a_1 w(n) - b_1\tilde{\kappa}w'(n)\right)e^{j\tilde{\omega}_1 n}, \quad n = -N,\ldots,N \tag{19}$$

where $\tilde{\kappa} = \frac{\sum_n |n\sqrt{w(n)}|}{\sum_n \left|w'(n)/\left(2\sqrt{w(n)}\right)\right|}$ is the appropriate normalization constant. Similar to iQHM, the LS method is used for

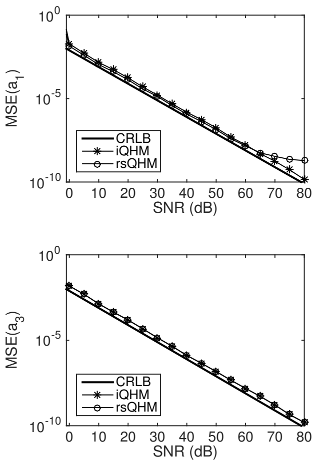

Fig. 3. Mean squared error (MSE) of the amplitudes for iQHM and rsQHM after two iterations utilizing Hann window.

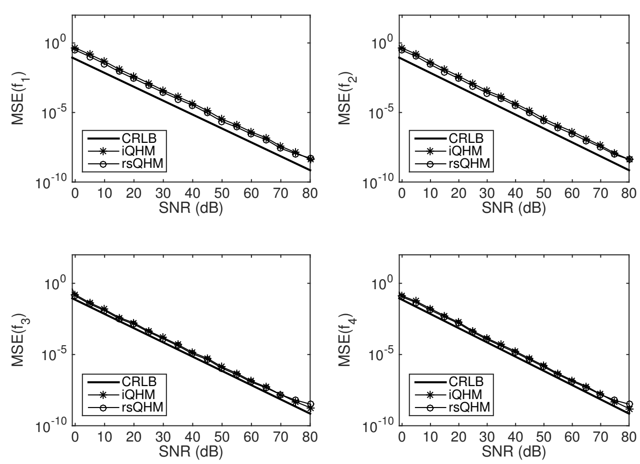

Fig. 4. Mean squared error (MSE) of the frequencies for iQHM and rsQHM after two iterations utilizing Hann window.

the estimation of the parameters and the same decomposition/projection is performed for rsQHM.

C. Numerical Validation

The comparison between iQHM (i.e. Newton-Gauss sinusoidal parameter estimation) and its variant (rsQHM) is threefold. All the presented figures are obtained utilizing a squared Hann window. The reason of using the square of the window is that the normalization constant is now given by $\tilde{\kappa} = \frac{\sum_n |nw(n)|}{\sum_n |w'(n)|}$ which is well-posed even when the window function takes zero values. Fig. 1 show the error between the true frequency mismatch and the estimated one as it is given by $\rho_{1,2}$ after two iterations. Upper panels shows the frequency mismatch error for the iQHM while the lower panels for the rsQHM. Note also that the analyzed signal is monocomponent and the window length was $16ms$.

It is evident that rsQHM estimates correctly the frequency of the sinusoid for a wider range of values as it is expected from Section III-A. However, as we mentioned in the previous Section, there are frequency mismatch values which cannot be correctly estimated because the amplitude becomes zero at these frequencies. Yet, rsQHM seems to have a larger region of convergence compared to iQHM. Indeed, Fig. 2 depicts the maximum allowed frequency mismatch (i.e., the region of convergence) for iQHM and rsQHM as a function of the window length. In this example, a three-component sinusoidal signal was considered and the figure shows the maximum allowed frequency mismatch (MAFM) for the middle sinusoid. Thus, the expected behavior of MAFM is initially increasing as the window length is increased because the frequency resolution is increased, but after a point MAFM starts to decrease due to the larger window length that reduces the area of guaranteed convergence. The dark dashed-line shows the lower bound of MAFM for rsQHM. On average, MAFM for the lower bound of rsQHM is 40% larger providing an advantage over iQHM. The oscillatory behavior of rsQHM for window length larger than $40ms$ is explained by the following fact. When the frequency mismatch falls between the diverging points (see lower panel of Fig. 1) then rsQHM correctly estimates the frequency error thus the MAFM is computed to be larger. Nevertheless, this is not always the case therefore we observe oscillations in MAFM.

The final comparison is the sensitivity of the models under noisy conditions. It is well known that Newton-Gauss method (as well as iQHM) is robust under mild noise conditions (see [10]) and we test if the same holds for rsQHM. Figs. 3 and 4 show the mean squared error (MSE) for the amplitudes and the frequencies of a multi-component sinusoidal signal after two iterations. The signal is exactly the same as those used in [10]. For comparison purposes, the Cramer-Rao lower bound (CRLB) is also provided. Presumably, rsQHM is robust under a wide range of noise conditions as it is measured by the signal-to-noise ratio (SNR) and two iterations are more than adequate for converging to the optimal value. We observe that rsQHM performs as good as iQHM in terms of MSE. Finally, we remark that similar results are obtained when other window types such as Gaussian, Hamming or Blackman are utilized.

IV. Conclusion

This paper highlights the connections between iterative LS estimation (Newton-Gauss method) and RS for sinusoidal models. Moreover, a variant of iQHM based on RS (rsQHM) was shown to perform better in terms of increasing the region of convergence of iQHM. We presented results for mono-component as well multi-component signals. Numerical evaluations on multi-component signals showed that rsQHM achieved the CRLB.

References

[1] F. Auger and P. Flandrin. Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method. *IEEE Trans. on Signal Processing*, 43:1068–1089, 1995.

[2] D. Iatsenko, P.V.E. McClintock, and A. Stefanovska. Extraction of instantaneous frequencies from ridges in time-frequency representations of signals. *Signal Processing*, 125:290 – 303, 2016.

[3] I. Daubechies, Y Wang, and H-T. Wu. ConceFT: concentration of frequency and time via a multitapered synchrosqueezed transform. *Phil. Trans. R. Soc. A*, 374:20150193, 2016.

[4] J. Harmouche, D. Fourer, F. Auger, P. Borgnat, and P. Flandrin. The sliding singular spectrum analysis: a data-driven non-stationary signal decomposition. *IEEE Transactions on Signal Processing*, 66:251–263, 2017.

[5] Y. Pantazis, O. Rosec, and Y. Stylianou. Adaptive AM-FM Signal Decomposition with Application to Speech Analysis. *IEEE Trans. on Audio, Speech and Language Processing*, 19(2):290–300, 2011.

[6] F. Ge, Q. Wan, L. Guo, and B. Sun. Frequency estimation of sinusoidal signals in multiplicative and additive noise. *IEEE Journal of Oceanic Engineering*, 41:810–819, 2016.

[7] X. Meng, A. Jakobsson, X.Li, and Y. Lei. Estimation of chirp signals with time-varying amplitudes. *Signal Processing*, 147:1 – 10, 2018.

[8] Y. Pantazis, O. Rosec, and Y. Stylianou. Iterative Estimation of Sinusoidal Signal Parameters. *IEEE Signal Processing Letters*, 17(5):461–464, 2010.

[9] P. Babu and P. Stoica. Comments on 'Iterative Estimation of Sinusoidal Signal Parameters. *IEEE Signal Processing Letters*, 17(12):1022–1023, 2010.

[10] Y. Pantazis, O. Rosec, and Y. Stylianou. Reply to "comments on 'iterative estimation of sinusoidal signal parameters"'. *IEEE Signal Processing Letters*, 17(12):1024–1026, 2010.

[11] K. Fitz and S. Fulop. A unified theory of time-frequency reassignment. arXiv, 2009. cs.SD (arXiv:0903.3080v1).

[12] B. Yegnanarayana and H. A. Murthy. Significance of group delay functions in spectrum estimation. *IEEE Trans. on Signal Processing*, 40:2281–2289, Sep 1992.

[13] P. Satyanarayana Murthy and B. Yegnanarayana. Robustness of group-delay-based method for extraction of significant instants of excitation from speech signals. *IEEE Trans. on Speech and Audio Processing*, 7:609–619, Nov 1999.