Instantaneous frequency estimation using weighted least squares Tikhonov regularization with box constraints

Ahmed H. Sigiuk[1] , Matthew J. Yedlin[2]

ABSTRACT

In this report, we begin with a brief introduction into the concept of instantaneous frequency, where we use the complex analytical signal to obtain the time dependent signal phase, and differentiate the phase with respect to time to obtain the instantaneous frequency. We reformulate our problem into a weighted least squares Tikhonov regularization with box constraints, and explain the advantages of the added terms in helping to smooth and stabilize our results. We conclude with some synthetic and real data simulation results and compare the proposed algorithm results with a well-established time-frequency distribution method.

INTRODUCTION

The importance of instantaneous frequency (IF) in spectral signal analysis stems from the need to analyze signals that are both nonlinear and nonstationary. In many applications such as seismic, sonar and biomedical engineering, the instantaneous frequency is a time-varying parameter, which defines the location of the signal spectral peak as a function of time. This only gives physical meaning if the signal is a monocomponent signal with only one frequency or a narrow band of frequencies at each time instant. For multicomponent signals, the single value instantaneous frequency gives only an average frequency value over time, which has very little physical meaning. Hence, a decomposition of the signal is required to extract the different monocomponent modes present in the signal, prior to calculating the instantaneous frequency of each mode.

In this report we start with the classical instantaneous frequency definition introduced by Gabor (Gabor,1946), which we will expand on by reformulating the problem into a weighted least squares Tikhonov regularization with box constraints inverse problem and compare our results with an alternative time-frequency energy distribution method based on the first frequency moment of the squares modulus of the STFT using a Gaussian window.

This report is an extension of previous work (Yedlin et al. 2015) using Tikhonov regularization for instantaneous frequency estimation.

[1] Department of Electrical and Computer Engineering, University of British Columbia
[2] Department of Electrical and Computer Engineering, University of British Columbia

THEORY

Classical Instantaneous Frequency Definition

Gabor proposed a method for generating a unique complex signal "Analytical signal (AS)" from a real one by using Hilbert transform. His method for obtaining the AS was achieved by first finding the FFT of the real signal, then by suppressing the negative frequency amplitude and multiplying the positive frequency amplitudes by two, which he then showed was equivalent to the following time domain representation.

$$z(t) = f(t) + i\,\mathcal{H}[f(t)], \tag{1}$$

$$z(t) = f(t) + ig(t). \tag{2}$$

where $z(t)$ is the analytical signal (AS), $f(t)$ is the real signal, and $\mathcal{H}$ is the Hilbert transform (HT) defined as

$$\mathcal{H}[f(t)] = \int_{-\infty}^{+\infty} \frac{f(\tau)}{\pi(t-\tau)}\, d\tau. \tag{3}$$

The analytical signal can also be written in polar form with $a(t)$ representing the magnitude with respect to time, and $\phi(t)$ representing the phase with respect to time. We note that the Hilbert transform $\mathcal{H}[f(t)]$ is not a 90° phase shift of the original signal $f(t)$ unless the frequency spectrum of the amplitude $a(t)$ is less than the frequency spectrum of the phase signal $\phi(t)$ with no overlap (Boashash, 1992a,b). Under these conditions the AS using the Hilbert transform and the quadrature representation (7) and (8) are equivalent.

$$z(t) = a(t)\, e^{i\phi(t)} = f(t) + ig(t), \tag{4}$$

$$a(t) = \sqrt{\big(f(t)\big)^2 + \big(g(t)\big)^2}, \tag{5}$$

$$\phi(t) = \tan^{-1}\left(\frac{g(t)}{f(t)}\right). \tag{6}$$

These can be written in quadrature form as follows

$$f(t) = a(t)\cos[\phi(t)], \tag{7}$$

$$f(t) = a(t)\sin[\phi(t)]. \tag{8}$$

where $\phi(t)$ is the instantaneous phase of the signal with respect to time, and by differentiating the (6), we obtain the instantaneous angular frequency (IAF) as follows

$$\omega(t) = \frac{d\,\phi(t)}{dt} = \frac{f(t)\,g'(t) + g(t)\,f'(t)}{(f(t))^2 + (g(t))^2} \tag{9}$$

$$\omega(t) = \frac{f(t)\,g'(t) + g(t)\,f'(t)}{(a(t))^2}. \tag{10}$$

We can also find the instantaneous frequency by dividing the (IAF) by 2π.

$$IF = \frac{\omega(t)}{2\pi}. \tag{11}$$

IF Estimation Using Weighted Least squares with Tikhonov regularization and box constraints

By reformulating (10) into matrix format, we obtain the following linear equations.

$$A\omega = b, \tag{12}$$

$$where, \qquad A = \| \mathrm{f} \|^2 + \| \mathrm{g} \|^2 \ \ Envelope, \tag{13}$$

$$b = f \circ \left(\frac{\Delta g}{\Delta t}\right) - g \circ \left(\frac{\Delta f}{\Delta t}\right). \tag{14}$$

where A is a diagonal matrix with squared envelope value, and the vector b is computed from the first order derivative of the signal and its Hilbert transform. It is apparent that for most real signals the matrix A is ill-conditioned (high ratio between smallest and highest eigenvalue), and the measurement vector b is highly contaminated with noise.

Weighted Least Squares Approach

Let's first start by reformulating the problem into a weighted least squares minimization problem with the following general solution

$$min \ \| \ C_b^{\frac{1}{2}} (A\omega - b) \ \|_2^2 \tag{15}$$

$$A^T C_b A\widehat{\omega} = A^T C_b b \tag{16}$$

From (16) we can show that to minimize the expected error in the estimated $\widehat{\omega}$ value, we select the matrix C_b to be the covariance matrix to our error vector e ($e = \omega - \widehat{\omega}$). In the simplest case of WGN, C_b is a constant diagonal variance matrix (Strang, 2007).

Least Squares Tikhonov Regularization Approach

First, we introduce the general minimization of the sum of two squares.

$$min \ \| \ A\omega - b \ \|_2^2 + \ \alpha \ \| \ B\omega - d \ \|_2^2 \tag{17}$$

with the following solution,

$$(A^T A + \alpha B^T B)\widehat{\omega} = A^T b + \alpha B^T d \tag{18}$$

The general solution for $\widehat{\omega}$ depends on the value of the weight α. Our original problem (12) can be very "ill-posed." Which often arises in inverse problems (Mead and Renaut, 2010). Our solution in (16) is unreliable either because $A^T A$ has a high ratio between smallest and largest eigenvalues (worse case $A^T A$ might be singular), or because b is very noisy which is amplified through differentiation, giving a wrong value for $\widehat{\omega}$.

A simpler form of least squares regularization is Tikhonov regularization, which takes the following form,

$$min \parallel A\omega - b \parallel^2 + \alpha \parallel \omega \parallel^2 \tag{19}$$

$$(A^T A + \alpha I)\widehat{\omega}_\alpha^e = A^T b \tag{20}$$

If we set $\alpha = 0$, the error e will be greatly amplified at the output $\widehat{\omega}_0^e$ of the least squares. The role of αI is to stabilize the least squares solution $\widehat{\omega}_\alpha^e$, which uses the noisy data $b - e$. We compensate for e by a good choice of α.

Using an upper and lower bound on the error of our estimated output value can be calculated (21) and (22) that is dependent on the value of α (Strang, 2007) .

$$\parallel \widehat{\omega}_0^0 - \widehat{\omega}_\alpha^0 \parallel \leq C\alpha \parallel \mathrm{b} \parallel \tag{21}$$

$$\parallel \widehat{\omega}_\alpha^0 - \widehat{\omega}_\alpha^e \parallel \leq \frac{\parallel \mathrm{e} \parallel}{2\sqrt{\alpha}} \tag{22}$$

A possible choice for α would be to equalize both parts of the overall error, yielding.

$$\alpha = \left(\frac{\parallel \mathrm{e} \parallel}{2C \parallel \mathrm{b} \parallel}\right)^{\frac{2}{3}} \approx (\parallel \mathrm{e} \parallel^2)^{\frac{1}{3}} \tag{23}$$

The matrix C depends on A^{-1} and the value of α is dependent on the variance of the error vector e (or noise).

Box constraints to quadratic inequality constraints

In finding the solution for (12), we most likely know the range of frequencies of the target solution. This can be added as a quadratic inequality constraint to our minimization problem as follows.

$$min \parallel A\omega - b \parallel_2^2 \;\; s.t\, (\omega_i - \bar{\omega}_i)^2 \leq \sigma^2 \tag{24}$$

$$\gamma < \omega < \beta, \qquad \sigma^2 = \frac{(\beta - \gamma)^2}{4}, \bar{\omega} = \frac{(\beta + \gamma)}{2} \tag{25}$$

By finding the Lagrange dual function for (24), and applying a penalty C_ϵ (where $C_\epsilon = diag(\sigma^2)$). We solve the following equation as follows.

$$min \parallel A\omega - b \parallel^2 + \parallel C_\epsilon^{\frac{1}{2}} (\omega - \bar{\omega}) \parallel^2, \tag{26}$$

$$(A^T A + C_\epsilon)\widehat{\omega} = A^T b + C_\epsilon \bar{\omega}. \tag{27}$$

In practice to find the optimal value for the penalty matrix C_ϵ, we run the solution through a loop until the constraints are satisfied.

Weighted least squares Tikhonov regularization with box constraints

By combining (15), (19) and (26) we obtain the following equation with the optimal solution as follows,

$$min \parallel C_b^{\frac{1}{2}} (A\omega - b) \parallel^2 + \alpha \parallel W\omega \parallel^2 + \parallel C_\epsilon^{\frac{1}{2}} (\omega - \bar{\omega}) \parallel^2, \tag{28}$$

With the general solution,

$$(A^T C_b A + \alpha W^T W + C_\epsilon)\hat{\omega} = A^T C_b b + C_\epsilon \bar{\omega}. \tag{29}$$

The Tikhonov regularization term in (28) was modified, by taking the first order derivative of the output parameter ω we assumes that our instantaneous frequency function is smooth with time and no sudden frequency jumps occur. We use the value α that equalize both parts of the overall error in \(30).

$$\alpha \approx \left(\frac{\sigma_s^2 + \sigma_e^2}{(\Delta t)^2} \right)^{\frac{1}{3}}. \tag{30}$$

NUMERICAL EVALUATION

We present a series of representative results from two datasets.

1. First, a synthetic data that is corrupted by zero-mean Gaussian noise.

2. Second, seismic data from a quarry blast in Jordan.

We will compare the results from two different algorithms.

A. Time-frequency energy distribution method based on the first frequency moment of the squares modulus of the STFT using a Gaussian window. The algorithm uses a 40-sample window size with a 20-sample overlap .We will refer to this method throughout the report as Algorithm 1.

B. Weighted least squares Tikhonov regularized with box constraints, using Equation 29. The algorithm uses the upper and lower frequency band of our input signal to calculate C_ϵ. We will refer to this method as Algorithm 2 throughout the rest of the report.

We also note that for both datasets the sampling frequency f_s is $40\ Hz$, and the first $25\ sec$ of the signal will only contain noise with no signal data; the algorithm uses this to evaluate the noise characteristics of the signal. As for the multicomponent signal, we empirical mode decomposition (EMD) (Huang et al. 2009) to obtain the monocomponent intrinsic mode function (IMF) prior running the IF algorithm on each IMF.

Synthetic signal

A. A signal with constant frequency modulation, a fixed frequency sinusoidal signal corrupted by zero-mean AWGN. Where $f_1 = 10\ Hz$ and the $SNR = 15dB$.

$$f_1(t) = \sin(2\pi f_1 t) + Noise \tag{31}$$

B. A signal with hyperbolic frequency modulation, the frequency-time plot follows the quadratic equation Eq 33. The signal is also corrupted by the same level of noise as in A, with $f_1 = 2\ Hz$, $f_2 = 12\ Hz$ and $t_1 = 30\ sec$

$$f_2(t) = chirpquad(f_1, f_2, t_1, t) + Noise \tag{32}$$

$$ideal\ IF\ for\ chirpquad = f_1 + \frac{(f_2 - f_1)}{(t_1)^2} t^2 \tag{33}$$

C. A multicomponent signal, consisting of a constant frequency signal and a chirp signal. The signal is also corrupted by the same level of noise as in A, with $f_1 = 2\ Hz$, $f_2 = 7\ Hz$, $f_3 = 15\ Hz$ and $t_1 = 30\ sec$.

$$f_3(t) = \sin(2\pi f_1 t) + chirp(f_2, f_3, t_1, t) + Noise \tag{34}$$

$$ideal\ \ chirp\ IF = f_1 + \frac{(f_2 - f_1)}{t_1} t \tag{35}$$

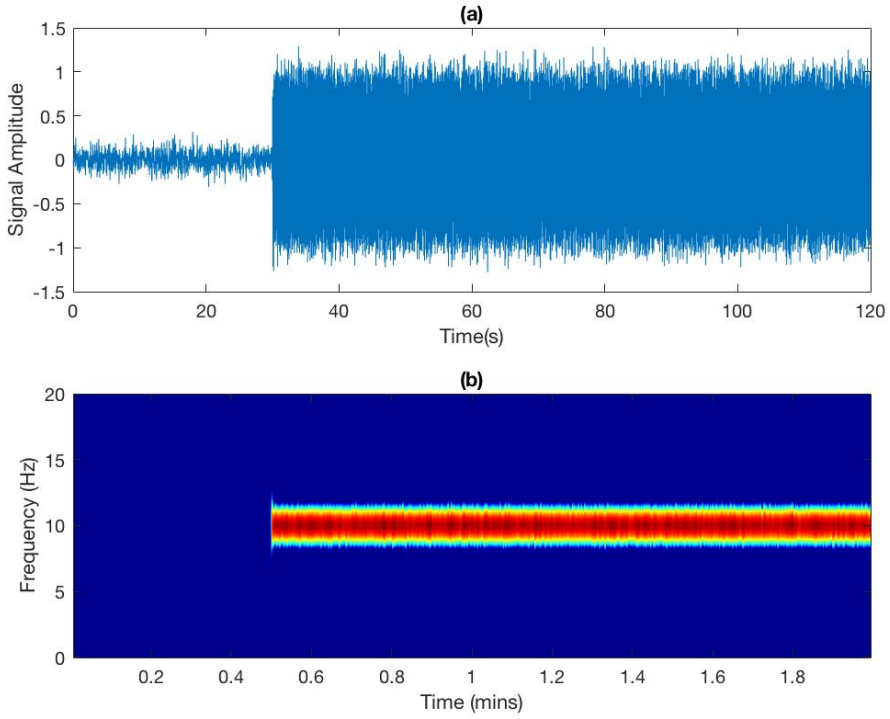

FIG. 1. (a) $f_1(t)$ signal time-amplitude plot. (b) $f_1(t)$ Spectrogram.

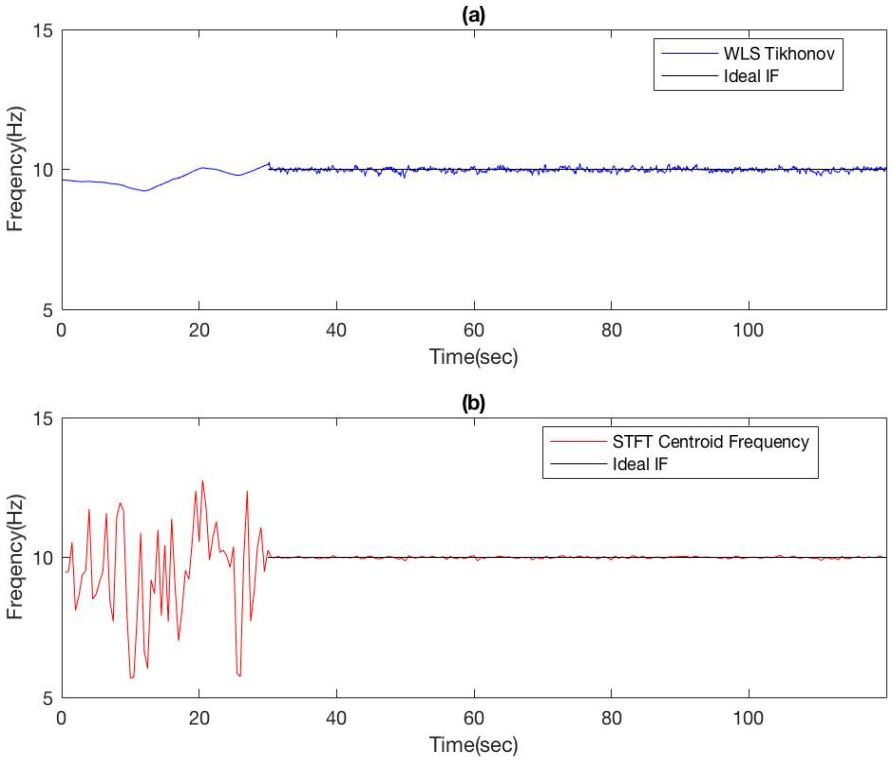

FIG. 2. (a) $f_1(t)$ IF estimation using Algorithm 2. (b) $f_1(t)$ IF estimation using Algorithm 1.

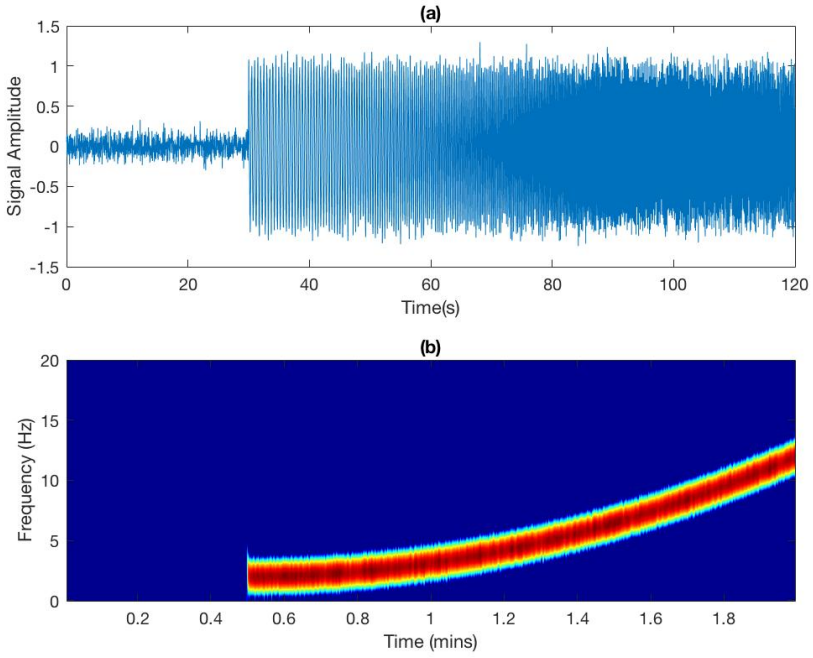

FIG. 3. (a) $f_2(t)$ signal time-amplitude plot. Panel (b) $f_2(t)$ Spectrogram.

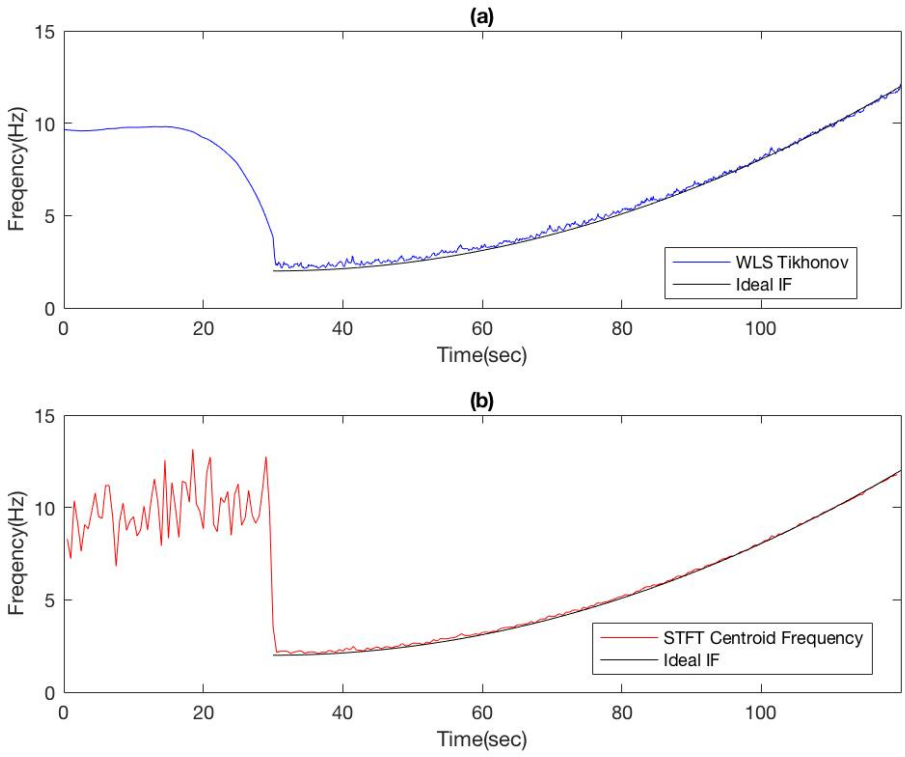

FIG. 4. (a) $f_2(t)$ IF estimation using Algorithm 2. (b) $f_2(t)$ IF estimation using Algorithm 1.

We note that both Algorithm 1 and Algorithm 2 provide good frequency estimation for the constant frequency modulation signal $f_1(t)$, and the hyperbolic frequency modulation signal $f_2(t)$. The average frequency variance for Algorithm 2 was 0.0118 from the ideal IF compared with a frequency variance of 0.0031 from ideal IF for Algorithm 1.

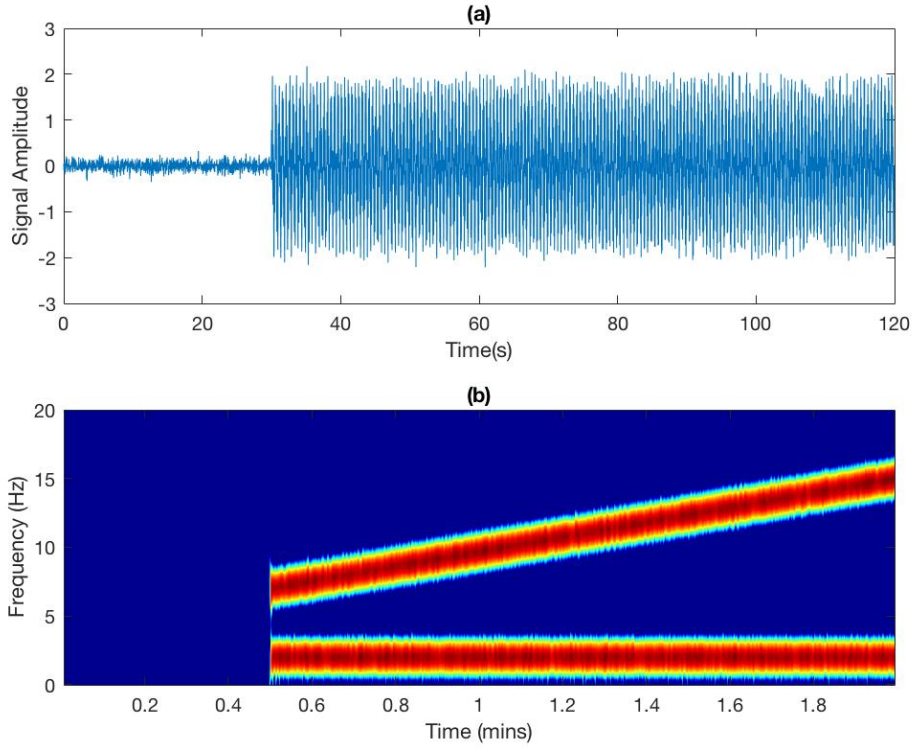

FIG. 5. (a) Synthetic $f_3(t)$ signal time-amplitude plot. (b) Spectrogram plot of $f_3(t)$ signal.

Before we can compute the instantaneous frequency for $f_3(t)$, first we must decompose the signal into two IMF signals, and then each IMF signal is processed separately, using both Algorithms 1 and 2. Fig. 6 panel (a) shows the IF plot for both IMF signals using Algorithm 2, and panel (b) shows the IF plot results from Algorithm 1. Both algorithms track the frequency fairly well at low frequency with some separation apparent at higher frequencies, Algorithm 2 having average frequency variance of 0.063 and Algorithm 1 doing slightly worse with an average frequency variance of 0.2.

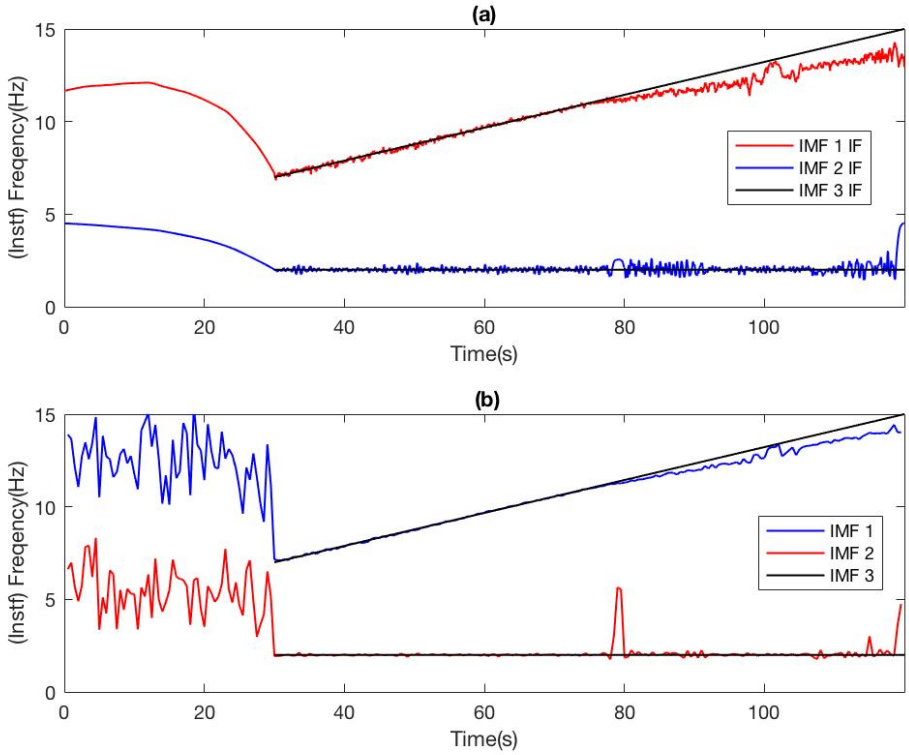

FIG. 6. (a) $f_3(t)$ IF estimation using Algorithm 2. (b) $f_3(t)$ IF estimation using Algorithm 1.

Seismic data from quarry blast in Jordan

The seismic data we present here was recorded from the HRFI station; located in southern Israel, latitude 30:04◦ N and longitude 35:04◦ E. The station recorded a three-component sequence of 1805 quarry blasts shot from by the Jordan Phosphate Mines Company, from March 2002 to January 2015.

We present here one seismic trace as an example, and evaluate the performance of the algorithm. Since the seismic data is also considered a multicomponent signal we first decompose the signal into three IMF signals (IMF 1, IMF 2 and IMF 3) before analysis the signal.

In Fig. 7 and Fig. 8 panel (a) is the z-component full signal recording and panels (b), (c) and (d) are the three major signal component or IMF after decomposition algorithm is applied to the full signal. We note that in addition to the three IMF presented here there are seven more IMF components which have negligible power contribution to the full signal.

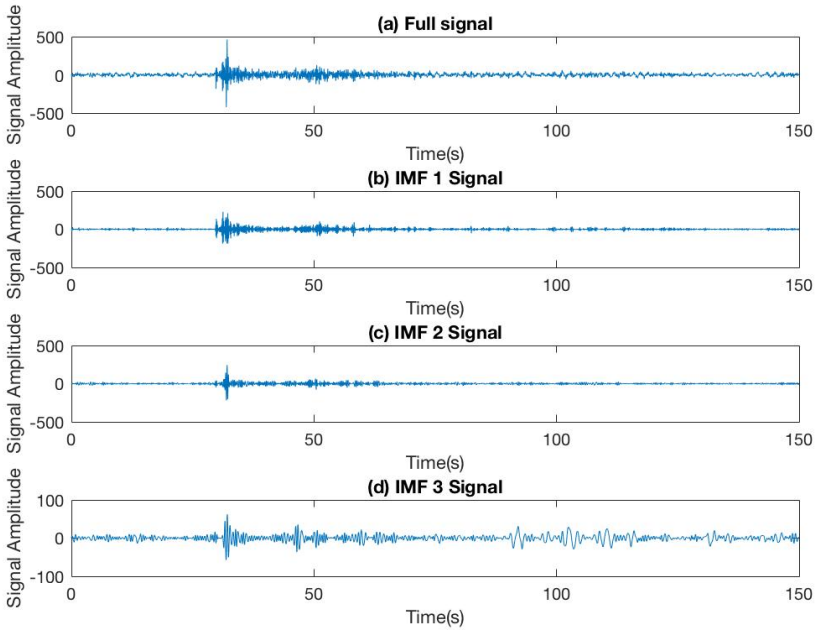

FIG. 7. Empirical mode decomposition (EMD) of the seismic signal. (a) z-component of the full seismic trace. (b) First IMF of the signal. (b) Second IMF of the signal. (c) Third IMF of the signal.

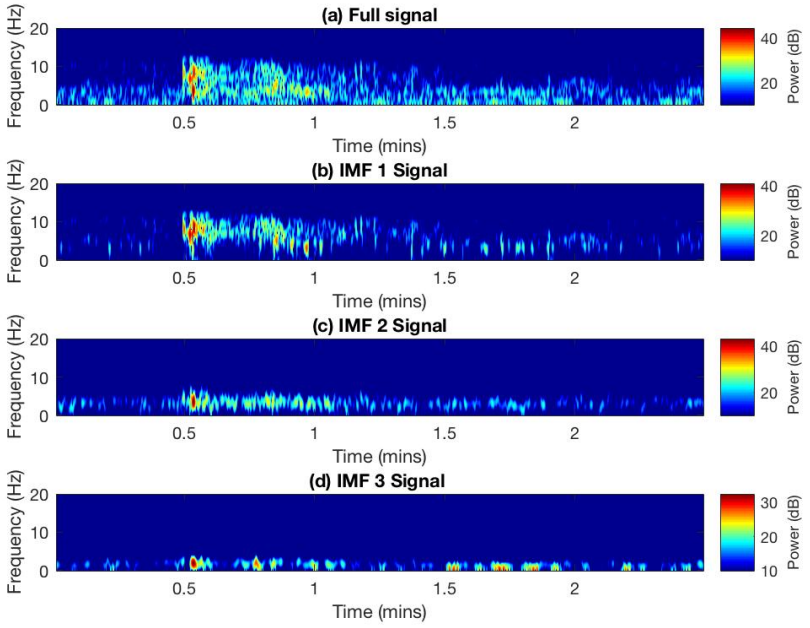

FIG. 8. Spectrogram. (a) z-component full signal spectrogram. (b) First IMF spectrogram. (c) Second IMF spectrogram. (c) Third IMF spectrogram.

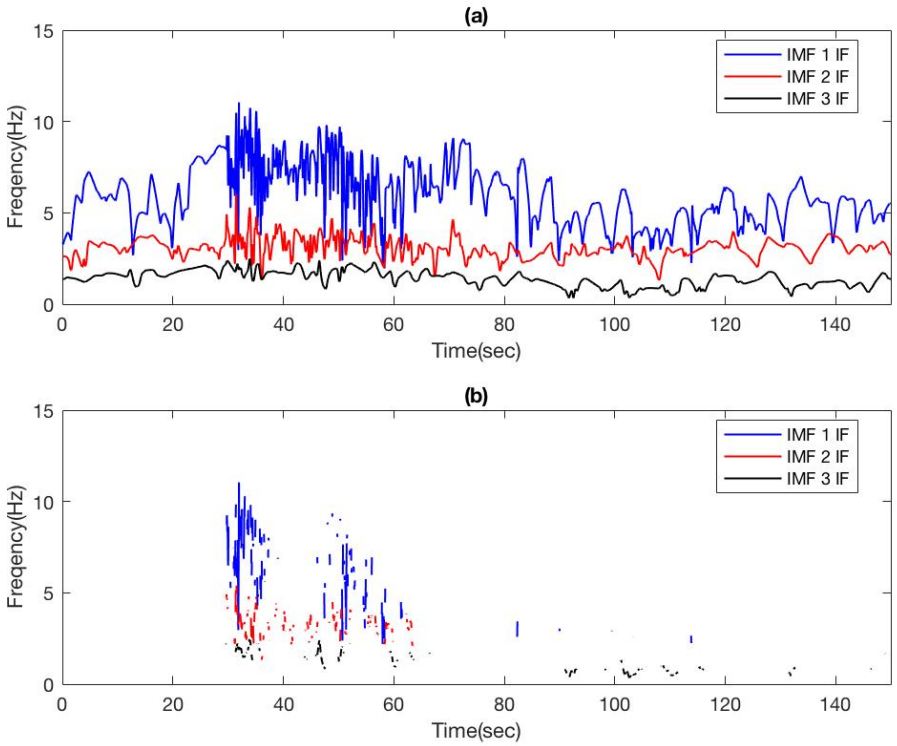

FIG. 9. Panel (a) Algorithm 2 Instantaneous frequency plot for each IMF mode. (b) Algorithm 2 Instantaneous frequency plot after applying minimum power threshold cutoff.

The instantaneous frequency output from Algorithm 2 is shown in Fig.9, where all three of the IMF are plotted. From the IMF 1 curve, we can clearly identify the P-wave and S-wave arrivals as well as the maximum frequency content. Also as shown in panel (b) we can apply a power threshold cutoff to the signal to filter out some of the noise. Application of this threshold allows us to measure the P and S wave arrivals from each IMF signal, and by taking the average time measurements from all three IMF components, we find that the P-wave onset is at 29.9 sec. and the S-wave onset is at 45 sec.

CONCLUSIONS

In this report, we toke a simple least squares minimization problem and enhanced it by adding three additional terms

1. Weighted least squares,
2. Tikhonov regularization,
3. Box constrain.

We generated three synthetic signals that have well-defined frequency properties and compared the results we obtained from the proposed algorithm to the central time-frequency distribution algorithm. The results obtained from both algorithms were very

close to each other as well as to the ideal frequency in the synthetic cases. In the quarry data example, we were able to use the Algorithm 2 , to pick S-wave onset in very noisy data.

Further work will focus on performing windowed decomposition prior to solving the least squares Tikhonov regularized problem, this will help speed up the algorithm execution time as well as improve the frequency smoothing output.

ACKNOWLEDGEMENTS

This research was supported by the CBIE-LNASP (SIGIUK)

REFERENCES

Boashash, B., 1992a, Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals: Proceedings of the IEEE, 80 , No. 4, 520–538.

Boashash, B., 1992b, Estimating and interpreting the instantaneous frequency of a signal. ii. algorithms and applications: Proceedings of the IEEE, 80 , No. 4, 540–568.

Mead, J. L., & Renaut, R. A., 2010, Least squares problems with inequality constraints as quadratic constraints: Linear Algebra and its Applications, 432 ,No. 8, 1936-1949.

Strang, G., 2007, Computational Science and Engineering, vol. 1: Wellesley: Wellesley-Cambridge Press.

Yedlin, M., Margrave, G. F., Vorst, D. V., and Horin, Y. B., 2014, Tikhonov regularization of instantaneous frequency attribute comparisons: CREWES Research Report, 26 , No. 80.1–80.14.

Gabor, D., 1946, Theory of communication. part 1: The analysis of information, part 2: The analysis of hearing, part 3: Frequency compression and expansion: Electrical Engineers - Part III: Radio and Communication Engineering, Journal of the Institution of, 93 , No. 26, 429–457.

Huang, Norden E., Zhaohua Wu, Steven R. Long, Kenneth C. Arnold, Xianyao Chen And Karin Blank., 2009a, On Instantaneous Frequency: Advances in Adaptive Data Analysis, 01, No. 02, 177-229.

Huang, Norden E., Shen, Samuel S P., 2005, Hilbert-Huang Transform and its Applications, vol. 5: World, River Edge, US.