

ON HILBERT SPECTRAL REPRESENTATION: A TRUE TIME-FREQUENCY REPRESENTATION FOR NONLINEAR AND NONSTATIONARY DATA

NORDEN E. HUANG*, XIANYAO CHEN[†], MEN-TZUNG LO*
and ZHAOHUA WU[‡]

**Research Center for Adaptive Data Analysis,
National Central University, Zhongli, Taiwan 32001*

[†]*The First Institute of Oceanography,
State Ocean Administration, Qingdao, China*

[‡]*Department of Earth, Ocean and Atmospheric Science,
& Center for Ocean-Atmospheric Prediction Studies,
Florida State University,
Tallahassee, FL 32306, USA*

As the original definition on Hilbert spectrum was given in terms of total energy and amplitude, there is a mismatch between the Hilbert spectrum and the traditional Fourier spectrum, which is defined in terms of energy density. Rigorous definitions of Hilbert energy and amplitude spectra are given in terms of energy and amplitude density in the time-frequency space. Unlike Fourier spectral analysis, where the resolution is fixed once the data length and sampling rate is given, the time-frequency resolution could be arbitrarily assigned in Hilbert spectral analysis (HSA). Furthermore, HSA could also provide zooming ability for detailed examination of the data in a specific frequency range with all the resolution power. These complications have made the conversion between Hilbert and Fourier spectral results difficult and the conversion formula is elusive until now. We have derived a simple relationship between them in this paper. The conversion factor turns out to be simply the sampling rate for the full resolution cases. In case of zooming, there is another additional multiplicative factor. The conversion factors have been tested in various cases including white noise, delta function, and signals from natural phenomena. With the introduction of this conversion, we can compare HSA and Fourier spectral analysis results quantitatively.

Keywords: Hilbert spectrum; marginal Hilbert spectrum; Fourier spectrum; instantaneous frequency; spectral resolution; sampling rate; Nyquist frequency; uncertainty principle.

1. Introduction

When Huang *et al.* (1998) first introduced the empirical mode decomposition (EMD) and Hilbert spectral analysis (HSA) as a new adaptive data analysis tool, it was clearly stated that the ultimate goal of the whole endeavor was to have a truly time-frequency representation for nonlinear and nonstationary data, the Hilbert

spectral representation. Consequently, the HSA and EMD together were designated as the Hilbert–Huang transform (HHT) by the National Aeronautics and Space Administration (NASA), where EMD was the pre-processing step to prepare the data so that the Hilbert transform could be used to define the instantaneous frequency for the time-frequency representation. Ever since the introduction of this new approach, overwhelming efforts were directed toward the EMD part, for, by itself, EMD had already served many useful applications in data analysis in a wide range of studies and applications that include physical, engineering, biomedical, and financial areas. As a result, most of the theoretical studies are also concentrated in the EMD area. New advances and innovations in EMD are numerous, such as the ensemble EMD (EEMD, Wu and Huang, 2009) and complementary EEMD (CEEMD, Yeh *et al.*, 2010), and smoothing EMD (SEMD, Pao *et al.*, 2010), and also some alternatives approaches by Olhede and Walden (2004) with wavelet packets, Hou *et al.* (2009) with multi-scale separation, and Daubechies *et al.* (2010) with synchrosqueezing wavelets. All the alternatives were primarily in the pre-processing decomposition step. The HSA by comparison was left on the side untreated and with no true advancements.

When the Hilbert spectral representation, $H(\omega, t)$, was first introduced by Huang *et al.* (1998), it was simply defined as

$$H(\omega, t) = \sum_{j=1}^n a_j(t) \exp\left(i \int \omega_j(t) dt\right) \quad (1)$$

where $\omega_j(t)$ is the instantaneous frequency, which is a function of time as the derivative of the phase function from Hilbert transform of the j th intrinsic mode function (IMF), $a_j(t)$ is the corresponding amplitude that is also a function of time.

In addition to the Hilbert spectrum, Huang *et al.* (1998) also introduced the marginal spectrum, $h(\omega)$ as

$$h(\omega) = \int_0^T H(\omega, t) dt \quad (2)$$

in which T is the total length of the data span.

Concerning the different physical nature of the various spectra defined, Huang *et al.* (1998) made the following statements:

The time-frequency distribution of the amplitude is designated as the Hilbert amplitude spectrum, $H(\omega, t)$, or simply Hilbert spectrum. If amplitude squared is more desirable, which is used commonly to represent energy density, the squared values of amplitude can be substituted to produce the Hilbert energy spectrum as well.

It represents the cumulated amplitude over the entire data span in a probabilistic sense.

The frequency in either $H(\omega, t)$ or $h(\omega)$ has a total different meaning from the Fourier spectral analysis.

With these cautionary notes, the true meaning and the properties of the Hilbert spectral representation had never been fully explored and understood. For example, the default designation of amplitude spectrum as the Hilbert spectrum had led Huang *et al.* (2001) to compare the marginal spectrum with the Fourier power spectrum. The only study other than the original one seemed to be the one by Wen and Gu (2009), who pointed that the integration of the Hilbert energy spectrum gave the total energy rather than the energy density as in the Fourier spectral representation. Such discrepancy had made the meaning of the Hilbert spectrum representation ambiguous and quantitative comparisons among Hilbert spectral result with Fourier and wavelet analyses difficult, if not impossible.

Over the last few years, advances in different fronts of HHT had been made: as stated above, the EMD had been extended to EEMD and CEEMD. The computation of instantaneous frequency had been extensively discussed and rigorously defined by Huang *et al.* (2009). Further, EMD under general conditions and instantaneous frequency through more precise and direct quadrature computations has been explored by Hou and Shi (2011). All these research studies are trying to eschew the traditional Hilbert transform, which is found to give only approximation frequency values under highly restrictive conditions. It seems that the condition is ripe for a renewal examination of the Hilbert spectral representation to make its mathematical definition clear and physical meaning lucid, so that its results could be compared with the traditional Fourier and wavelet analyses, and its applications could be made quantitatively.

In the paper, we start from the basic definition of the Hilbert spectral representation in terms of energy density as in the traditional spectral representations. Then, we proceed to the detailed comparisons with other existing method of time-frequency analysis methods. The emphasis, however, is on the prowess and versatility of the Hilbert spectrum as a truly time-frequency representation based on the example of various types of data ranging from delta function, white noise, and real data from natural phenomena. We then discuss the various limitations of spectral representations based on frequency obtained through integral transformations, and contrast them with the Hilbert spectral representation. Finally, we draw a conclusion based on our study. Now, let us start with the discussion on the definition of the Hilbert spectrum.

2. The Definition of the Hilbert Spectrum

According to Huang *et al.* (1998), after decomposing the data, $x(t)$, into IMF components, $c_j(t)$, we should have

$$x(t; \omega) = \sum_{j=1}^n c_j(t) = \sum_{j=1}^n a_j(t) \cos \theta_j(t) \Rightarrow \sum_{j=1}^n a_j(t) \cos \left(\int_0^t \omega_j(\tau) d\tau \right) \quad (3)$$

where $\theta_j(t)$ is the phase function with its derivative as the instantaneous frequency, and the arrow indicating a proper transform of the phase function into the frequency

space. Therefore, the squared signal should be

$$\begin{aligned} x^2(t; \omega) &= \left(\sum_{j=1}^n c_j(t) \right)^2 = \sum_{j=1}^n c_j^2(t) + \sum_{k \neq j} \sum_{j=1}^n c_j(t)c_k(t) \\ &\Rightarrow \sum_{j=1}^n a_j^2(t) \cos^2 \left(\int_0^t \omega_j(\tau) d\tau \right) \end{aligned} \quad (4)$$

in which the double summation term in the first line was set to zero by virtue of orthogonality of the IMFs.

According to the original definition given by Huang *et al.* (1998), the Hilbert spectrum is only defined schematically as given in Fig. 1. This is only a qualitative definition in terms of total energy. To facilitate future comparison, we define the Hilbert spectrum quantitatively in terms of energy density as:

The Hilbert energy spectrum is defined as the energy density distribution in a time-frequency space divided into equal-sized bins of $\Delta t \times \Delta \omega$ with the value in each bin summed and designated as $a^2(t)$ at the proper time, t , and proper instantaneous frequency, ω . With this definition, one can see that the resolution of the Hilbert spectrum is determined by the bin size selected but not by the total data length and sampling rate as in the Fourier spectral analysis.

Similarly, the Hilbert amplitude spectrum is defined as the amplitude density distribution in a time-frequency space divided into equal-sized bins of $\Delta t \times \Delta \omega$ with the value in each bin summed and designated as $a(t)$ at the proper time, t , and

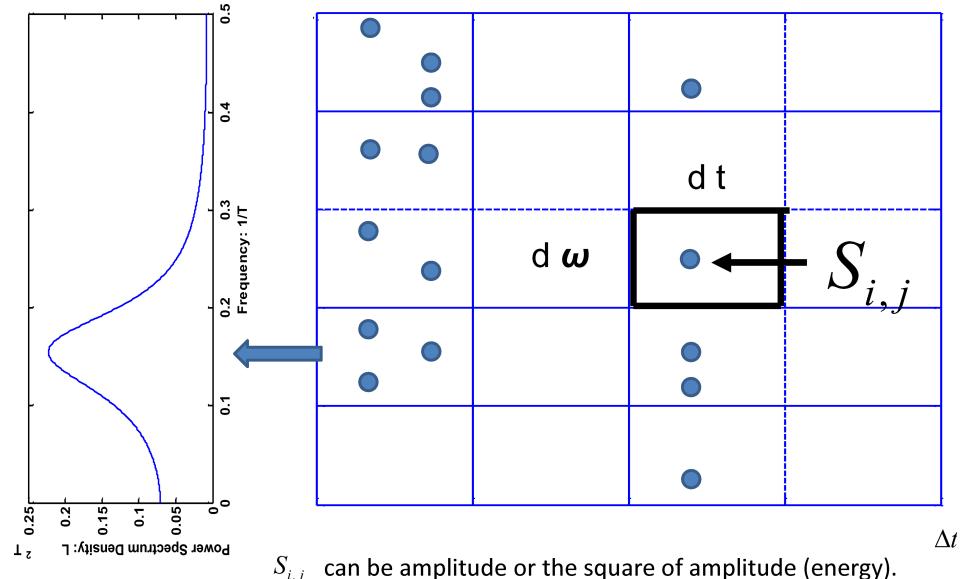


Fig. 1. The schematic of Hilbert energy distribution spectrum. The time-frequency is subdivided into equal-sized bins. Energy density values fall in the bin would be summed. Therefore, the bin size determines the spectral resolution.

proper instantaneous frequency, ω . Again, the resolution of the amplitude spectrum is determined by the bin size selected.

To implement this definition, the time-frequency space is divided as

$$\begin{aligned} t_0, t_0 + \Delta t, \dots, t_0 + i\Delta t, \dots, t_1. \\ \omega_0, \omega_0 + \Delta\omega, \dots, \omega_0 + j\Delta\omega, \dots, \omega_1. \end{aligned}$$

Thus

$$t_i = t_0 + i\Delta t \quad \text{and} \quad \omega_j = \omega_0 + j\Delta\omega$$

Here, the values and sizes of the time and frequency bins could be selected independently and arbitrarily to fit our need subjecting only to the following constrictions:

- (1) t_0 and t_1 have to reside within the interval of the data span, $[0, T]$.
- (2) Δt cannot be smaller than the sampling rate steps.
- (3) ω_1 should be less than a Nyquist frequency to be defined below.

With this division of time and frequency space, we have the Hilbert energy density spectrum, $S_{i,j}$, defined as

$$S_{i,j} = H(t_i, \omega_j)$$

$$= \frac{1}{\Delta t \times \Delta\omega} H \left[\sum_{k=1}^n a_k^2(t) : t \in \left(t_i - \frac{\Delta t}{2}, t_i + \frac{\Delta t}{2} \right), \omega \in \left(\omega_j - \frac{\Delta\omega}{2}, \omega_j + \frac{\Delta\omega}{2} \right) \right] \quad (5)$$

It should be noted that in the traditional Fourier spectral representation, the time scale does not appear in the formulation at all, due to the nature of the Fourier transform. The frequency scale is determined jointly by the sampling rate, Δt , and the data span, T , through the frequency resolution and the Nyquist frequency:

$$\begin{aligned} \Delta\omega = \frac{1}{T} &: \text{frequency resolution;} \\ \frac{1}{2\Delta t} &= \text{the Nyquist frequency, } \omega_q. \end{aligned} \quad (6)$$

Here, the Nyquist frequency is the highest frequency value one can ever get in the Fourier spectral analysis. The only condition that the time would appear explicitly in the Fourier spectral analysis is through spectrogram. There, the window location and window width would both involve time, as discussed later.

In HSA, we have both time and frequency variables appear explicitly. The limitation on Δt is the sampling rate as stated above. The frequency resolution in Fourier spectral representation is defined as an integer times the frequency resolution size, $\Delta\omega$; therefore, the frequency can only have values at $n \cdot \Delta\omega$. The frequency value in Hilbert spectral representation can assume any value on a continuous scale. There, however, still is a limitation on the highest frequency value, or the Nyquist frequency, in HSA. It can be shown that the Nyquist frequency in HSA is exactly the same as in Fourier analysis: given the sampling rate at Δt , the largest phase change

between the neighboring data points is from a maximum at one point changes to a minimum at the next point, or vice versa. The total phase change is therefore π , and the highest frequency should be

$$\omega_q = \frac{\pi \text{ radian}}{\Delta t \text{ second}} = \left(\frac{1}{2\pi} \frac{\pi}{\Delta t} \text{ Hz} \right) = \frac{1}{2\Delta t} \text{ Hz} \quad (7)$$

This is the same as given in Eq. (6). It should be pointed out that the instantaneous frequency could have singularity or erroneous values higher than the Nyquist value, if one uses the Hilbert transform approach. For example, there would be a singular point when the amplitude is precisely zero. Otherwise, erroneous values could also occur when the local amplitude is small. None of these conditions are physically meaningful and should be excluded.

This new freedom of arbitrary frequency and time bins made the data density a variable. As a result, the comparison between Fourier spectral analysis and HSA becomes very difficult and the conversion formula has been elusive until now. However, when the spectral unit is defined in terms of energy density, there should be a simple conversion factor. The goal of this paper is to define such a factor under the general condition for the Hilbert spectra as well as for the marginal spectra defined in the next section.

With the Hilbert spectral representations given, the marginal energy spectral representation is defined as

$$h(\omega_j) = \sum_{i=1}^N H(t_i, \omega_j) \times \Delta t = \left(\frac{1}{\Delta t \times \Delta \omega} \sum_{k=1}^m a_k^2(t) \right) \times \Delta t = \frac{1}{\Delta \omega} \sum_{k=1}^m a_k^2(t) \quad (8)$$

Integration of the marginal energy spectrum would be in terms of the total energy:

$$\int h(\omega_j) d\omega = \left(\frac{1}{\Delta \omega} \sum_{k=1}^m a_k^2(t) \right) \times \Delta \omega = \sum_{k=1}^m a_k^2(t) = 2 \sum_{i=1}^N x^2(t_i) = 2E \quad (9)$$

Thus, the integral of the Hilbert energy spectrum is the twice the total energy, for

$$E = \sum_{i=1}^n x^2(t_i) = \sum_{i=1}^N \sum_{j=1}^n c_j^2(t_i) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^n a_j^2(t_i) \quad (10)$$

By definition of Fourier spectral representation, we have

$$\overline{x^2} = \int S(\omega) d\omega, \quad (11)$$

where $S(\omega)$ is the Fourier spectrum for $x(t)$, and that

$$\overline{x^2} = \frac{\sum_{i=1}^N x^2(t_i)}{N} \quad (12)$$

with N as the total number of data point. Therefore, the Fourier spectrum gives the mean rather than the total energy. The different definitions of spectral function make the comparison difficult. By combining Eqs. (10) and (12), we have

$$\sum_{i=1}^N x^2(t_i) = N \cdot \overline{x^2} = N \cdot \sum_{j=1}^n S(\omega_j) \Delta\omega = \frac{N}{T} \sum_{j=1}^n S(\omega_j) = \frac{1}{2} \sum_{k=1}^m a_k^2 \quad (13)$$

Therefore, we have

$$\frac{1}{2} \sum_{k=1}^m a_k^2 = \frac{N}{T} \sum_{i=1}^M S(\omega_i) \quad (14)$$

Consequently, the conversion factor between the Hilbert and the Fourier spectra is a key, but simple, factor of N/T . As T/N is the sampling rate, the conversion factor is simply

$$\frac{1}{\text{sampling rate}} \quad (15)$$

Based on Eq. (14), if one would like to convert the Fourier spectral density to the Hilbert energy density distribution unit, the Fourier spectral values need to be multiplied by N/T . Conversely, if one would like to have the spectral representation to be expressed in the Fourier energy density unit, then the Hilbert energy distribution values need to be multiplied by T/N , the sampling rate. This is certainly true if the frequency range is all naturally defined to include all the possible values up to the Nyquist frequency. Although this sounds so simple, the actual determination of the energy density in Hilbert spectral representation is much more complicate, which is caused primarily by the freedom in defining the bin size and also the zooming capability in the HSA. Unlike Fourier analysis, in Hilbert spectral representation, we can select arbitrary frequency resolution bin size and arbitrary frequency range to zoom in. This freedom would introduce additional complications, or an additional modification of the conversion factor. Suppose that we will put all our resolution in frequency at a special range and designate this special frequency resolution as $\Delta\omega_a$ with the associated Hilbert and marginal energy spectra, $S_{i,aj}$ and $h_a(\omega)$, we would have

$$h_a(\omega) = \frac{1}{\Delta\omega_a \times \Delta t} \sum_{i,aj} S_{i,aj} \Delta t = \frac{1}{\Delta\omega_a} \sum_{k=1}^m a_k^2 = \frac{1}{\Delta\omega_a} (\Delta\omega h(\omega)) \quad (16)$$

Then, according to Eq. (15), the new conversion factor would have an extra term:

$$\frac{\Delta\omega}{\Delta\omega_a} \quad \text{for } h_a(\omega) = \frac{\Delta\omega}{\Delta\omega_a} h(\omega) \quad (17)$$

With the above derivation, we have established the firm conversion between Fourier spectra and Hilbert marginal spectra. The complications actually came from the power endowed in the Hilbert spectral representations with total freedom in selecting an arbitrary combination for the time and frequency resolution and zooming

ability. As we routinely evoke the zooming feature to examine minuate and subtle changes in special frequency ranges, the new freedom made the quantitative definition of the spectral density elusive ever since the introduction of the HSA. The above analysis has established for the first time a definitive conversion between the Hilbert spectral density with the Fourier counterpart. Such a conversion would enable us to make quantitative comparison in the future.

Before proceeding any further, there are several observations worthy of our attention. First, in the Fourier spectral analysis, the frequency resolution is determined totally by factors such as the total data length and sampling rate. However, these factors are unrelated to the physical properties we are looking for. In the HSA, we can separate the time and frequency and select them to the degree of accuracy to fit our needs for physical understanding. We can also select the frequency resolution independent of the data length and the sampling rate up to the Nyquist value. The resolution could be as fine as we want inspite of the fact that there should be a upper Nyquist limit on how many independent frequency values. Yet, as these values are changing on a continuous real positive number axis, we could select the resolution as fine as we want to isolate the true values of the instantaneous frequency.

Second, we have concentrated exclusively on the conversion of the Hilbert energy spectrum to Fourier, and did not pay any attention on the amplitude spectrum. We made this decision because the Fourier spectral is in terms of energy density (or power spectrum as some other authors prefer). The original definition of Hilbert spectrum in terms of amplitude is somewhat a misnomer, for at that time we do not have the conversion to spectral density in consideration. Now, as we study the conversion between Hilbert and Fourier spectra, it is only natural to consider the spectral density in terms of energy rather than amplitude.

Third, we have concentrated on the effects on frequency resolution here without any attention on time bin size and resolution. Although the time variable does not appear explicitly in Fourier spectral analysis, in the spectrogram the time and frequency resolution is linked in the so-called uncertainty principle: i.e., one could not determine the frequency and time resolution to an arbitrary precision independently. As time variable does not appear in Fourier analysis, its effect in HSA is implicit as far as comparison with Fourier analysis is concerned. If the instantaneous energy is of interest, then time resolution would be more important. There, however, is no need for neither conversion nor comparison between Hilbert and Fourier analyses along the time axis. Therefore, time resolution is not emphasized here.

Having established the conversions between Hilbert and Fourier spectral representations, we should also introduce a variation of the Fourier representation in the form of a spectrogram in which the Fourier spectrum is defined in a subdivision of time space known as the window, τ_w . Thus, the spectrogram is defined as

$$F(\omega_j, t_i) = S(\omega_j; t_i) : t_i \in \left(t_i - \frac{\tau_w}{2}, t_i + \frac{\tau_w}{2} \right) \quad (18)$$

In this form, the spectrogram retained some indication of the neighborhood of its validity. This gain of temporal mark is paid in the degradation of frequency resolution. As the Fourier transform is performed within the window, the frequency resolution is only

$$\Delta\omega_w = \frac{1}{\tau_w} \quad (19)$$

The Nyquist frequency is the same as given in Eq. (6). By comparison between Eqs. (6) and (19), one can immediately see that the frequency degradation is proportional to the size of windows selected: the finer the time resolution one wants, the cruder the resulting frequency resolution would be. This is precisely the so-called “uncertainty principle” in data analysis. It should be pointed out that the uncertainty principle in data analysis is not a fundamental law as in physics. It is a consequence of the integral transform (in this case, the Fourier transform) used in converting the data from temporal to frequency space. By virtue of the integral, the time mark is obliterated within the integral limit. Crude as the resolution of a spectrogram is, the spectrogram was the only time-frequency representation available before the introduction of wavelet and Wigner–Ville distribution. As a result, it is still being used heavily and almost exclusively in speech analysis. Now, let us demonstrate the conversion formula between Hilbert and Fourier spectral representations derived above in some concrete examples.

3. Examples of Conversion Between Hilbert and Fourier Spectral Representations

To illustrate the conversion between the Hilbert and Fourier spectra, we use some concrete examples here. Of particular interest are the white noise and delta function, for these are the function having given many user trouble. Additionally, we also use the conversion on a set of earthquake data.

3.1. White noise

The first example we use is the white noise, partly because it has been studied extensively by Flandrin and Gonçalves (2004) and Wu and Huang (2004). Based on those studies, Wu and Huang (2009) also used the statistical properties of white noise as the foundation for the most important advances in HHT, the ensemble EMD. This is, partly, because Gledhill (2003) had also studied the white noise and found that the marginal spectrum is not flat. He had embarked on a heroic effort to make the spectrum flat by introducing multiplicative factors but to no avail. Let us take a white noise signal of 1,000 data points as shown in Fig. 2. Assuming the sampling rate is 1 Hz, we have a Nyquist frequency at 0.5 Hz and the Fourier spectrum would have 500 spectral lines. Using straightforward EMD, we have 10 IMFs as given in Fig. 3. The corresponding Hilbert spectra with different frequency bin numbers at 10, 20, 30, 50, 100, 300, 500, 600, and 800 are given in Fig. 4. A casual

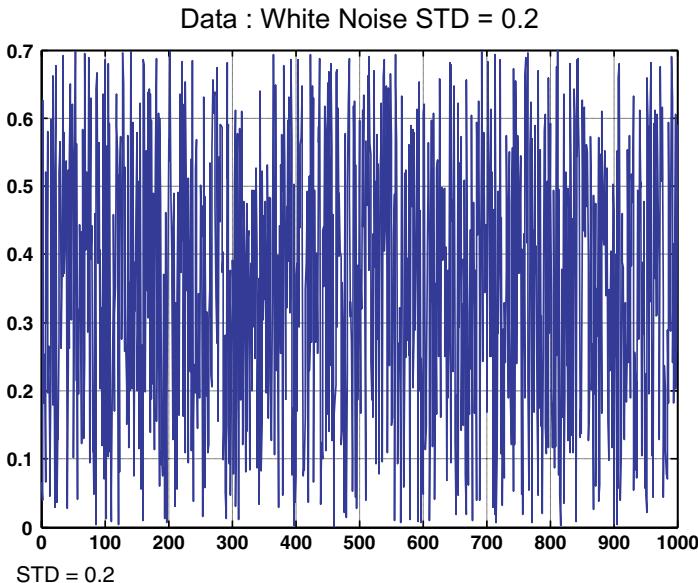


Fig. 2. The uniformly distributed 1,000 data points.

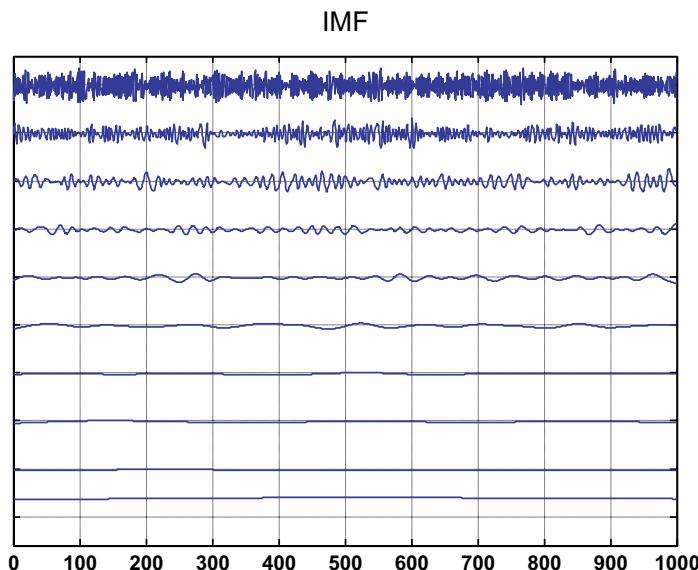


Fig. 3. The IMF of the white noise data given in Fig. 2.

inspection suggests that the details and quality of the spectra do not bear remarkable resemblance. If we use the 500 bin Hilbert spectrum to compute the marginal spectrum, the conversion is simply a factor of 2, and we obtain the result as given in Fig. 5. Both the spectra are flat over most of the frequency range except the

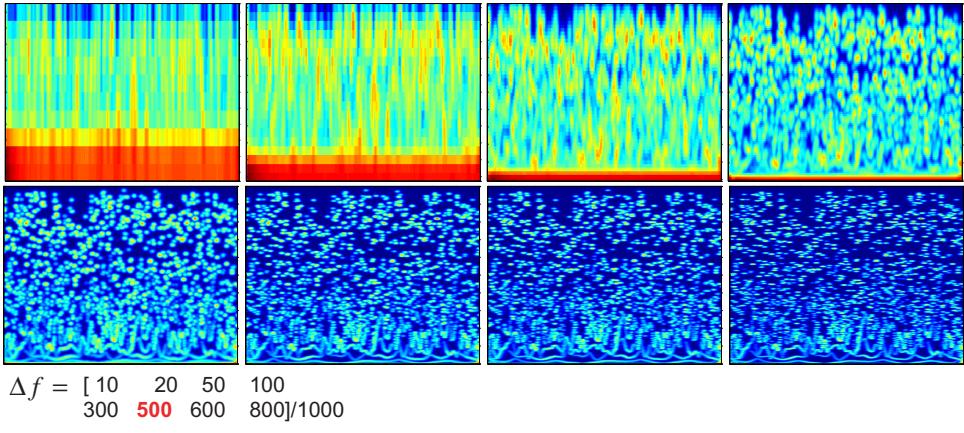


Fig. 4. The Hilbert spectra for bin numbers equal to 10, 20, 50, 100, 300, 500, 600 and 800, respectively. The spectra do not have any obvious similarity although they were all from the same data but with different bin sizes.

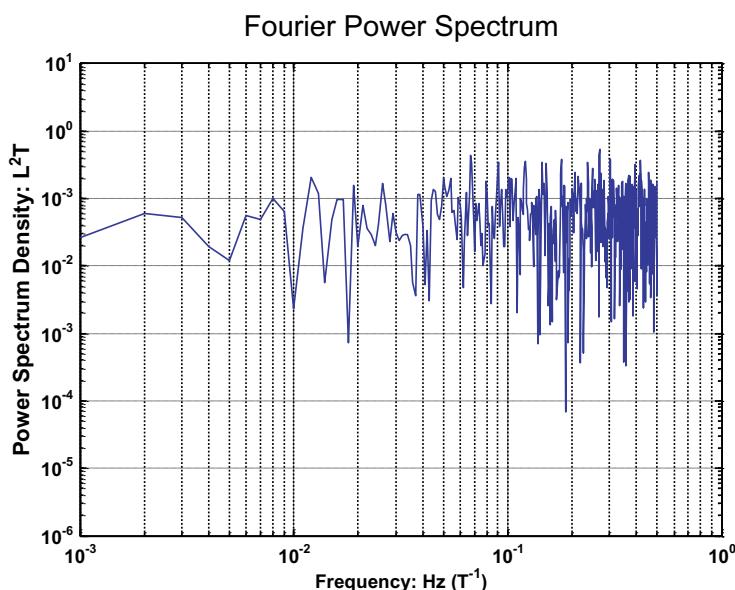


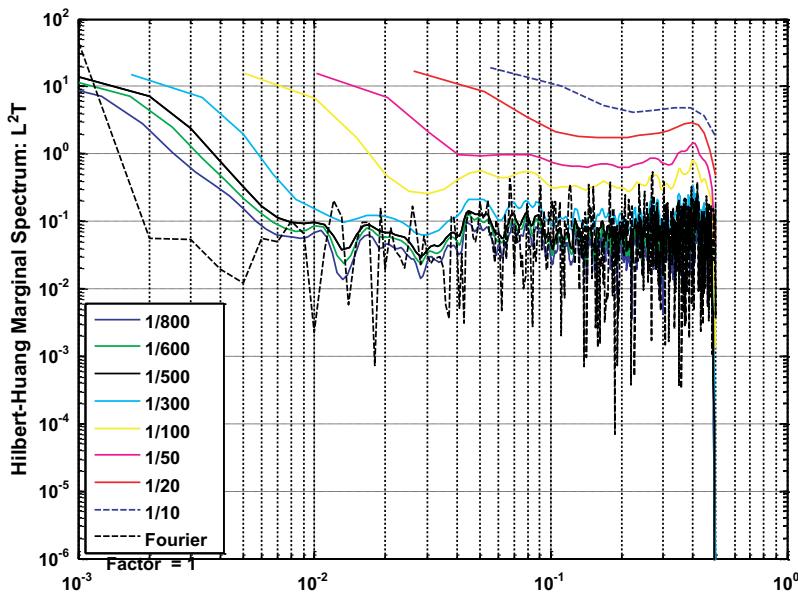
Fig. 5. Fourier spectrum with 500 frequency bins.

region near the lowest frequency range. As we demonstrate in this study, this must be a coincidence of this particular realization of the white noise data set as other trials shown later would not have this abundance of energy at the low-frequency region. Furthermore, this abundance of energy is very different in characteristics from the one reported by Huang *et al.* (1998) or Gledhill (2003).

Now, we also show the various marginal spectra with different frequency bin sizes in Figs. 6(a) and 6(b), for the cases before and after the conversion. In the unconverted cases, we have also used some 5×5 Laplacian filter to smooth the Hilbert spectra before computing the marginal values. This exercise serves to demonstrate that we could smooth the full time frequency spectrum and produce a smoother marginal spectral form. Both sets of spectra show flat spectral form comparable in both Hilbert and Fourier representations. Based on these tests, we conclude that the Hilbert marginal spectra for the white noise are indeed flat and therefore also white.

It should be pointed out that the selection of bin size could influence the resolution of the spectrum drastically. Let us take the cases with 50 and 500 bins for example. As the Nyquist frequency is 0.5 Hz, the 50 and 500 bin size cases would have frequency resolution of 0.01 and 0.001 Hz, respectively. Therefore, the first spectral value for the 50 bin size case is 0.01 Hz and the corresponding case for the 500 bin size case is 0.001 Hz. Thus, the coarse bin size case has not been explored for a whole decade in the spectral representation as shown in Figs. 6(a) and 6(b). This trend could be easily identified from Fig. 6(b).

To demonstrate that we could indeed select frequency resolution finer than the number in the discrete Fourier spectral lines, we have selected the bin size at 600



(a) Effects on Frequency Resolution MHS

Fig. 6. The unnormalized (a) and normalized (b) Hilbert marginal spectra for the various bin numbers given in Fig. 4. Notice that we could select bin number (600, 800, for example) greater than the fixed value (500 spectral lines) depicted by the Fourier analysis.

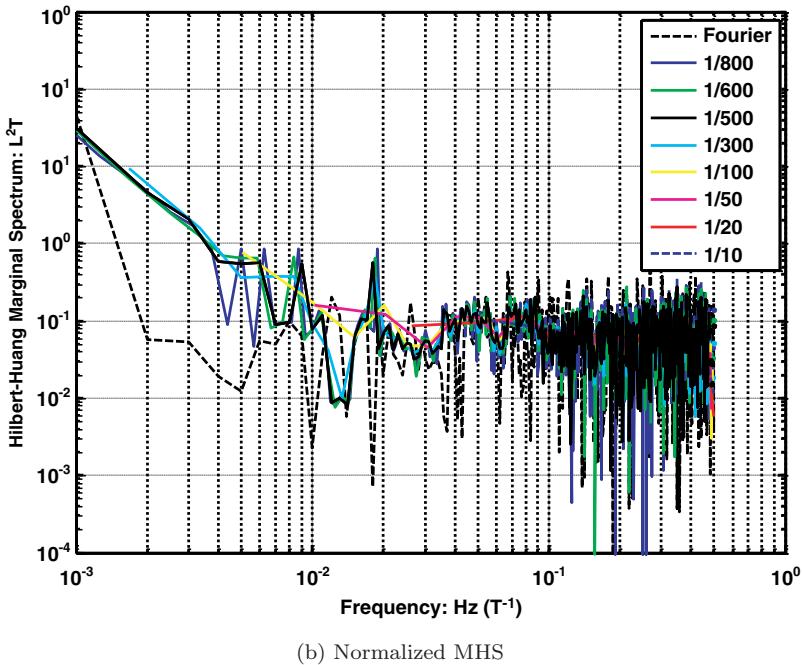


Fig. 6. (Continued)

and 800 as shown in Fig. 6. Although the general trends of these finer resolution spectra are similar to the 500 bin case, there are indeed additional details and subtle variations in the finer results. We return to this point later.

To examine the effects of the bin size and their relationship with the Fourier window size, we use another realization of a uniform white time series of 1,024 points. The results of Fourier spectra with window of different lengths and Hilbert marginal spectra with different bin numbers are given in Figs. 7(a) and 7(b). It is clear that the effect seems to be very similar. This is especially true for the low frequency cut-off limit. The appearance actually is misleading; the difference, in fact, is fundamental: in Fourier analysis, the short window would make the frequency resolution power almost totally disappear on the longer than the window length waves. For the HSA, the bin size serves as a smoothing function. By smoothing, the frequency resolution power is also enhanced.

Finally, we examine the case with the amplitude spectra as given in Fig. 8. Here, the amplitude spectra show a distinct slope uniformly over the whole frequency range. This abundance of energy in the lower frequency amplitude density is the one identified by Huang *et al.* (1998). It is also the persistent tilting that Gledhill (2003) had tried to get rid of through an empirical multiplier. The problem is not in the results, but in the fact that we were comparing totally different spectra: Fourier spectra are always in terms of energy (or power density), while the Hilbert

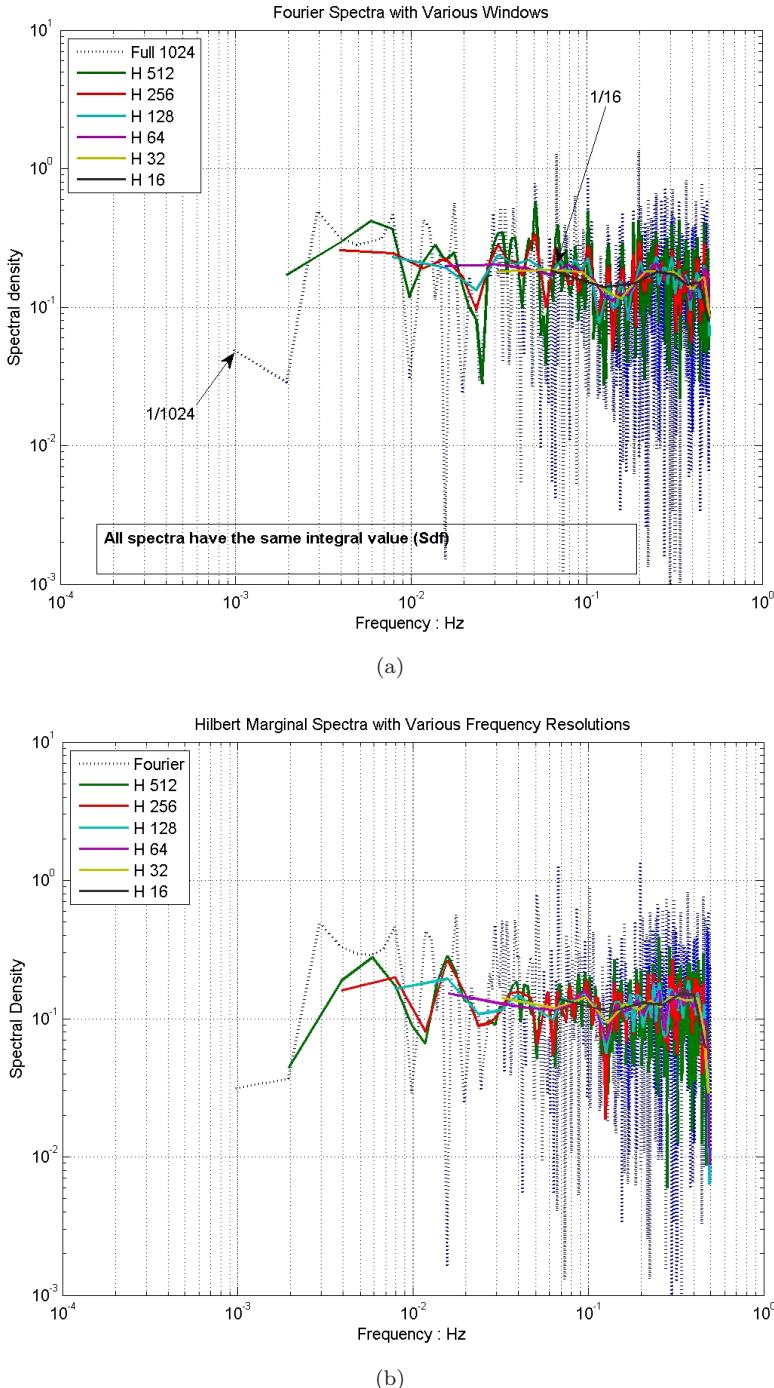


Fig. 7. The effects of window size in Fourier spectral analysis (a) and the numbers of bins in Hilbert marginal spectra (b). Although the causes are different, the end results of both smoothing and window sizes are the deterioration of resolution.

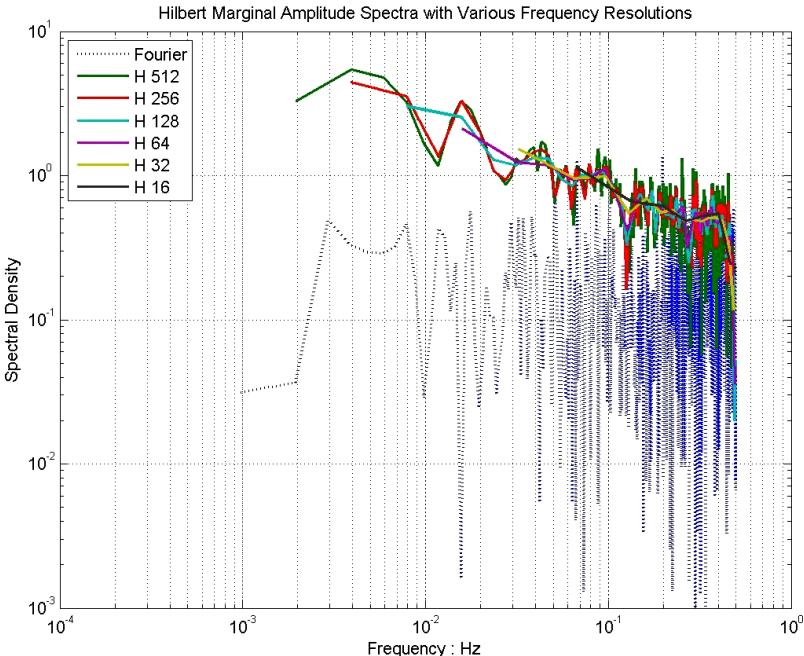


Fig. 8. Hilbert amplitude spectra for various bin numbers. Notice the tilting of the spectra making the form no longer flat. This should not be confused with the flat spectra of Fourier or Hilbert energy spectra for white noise.

amplitude spectrum was defined then in terms of the density of the amplitude but not the squared power. Therefore, the disagreement was natural and should be expected. It should be pointed out that the amplitude spectra are useful if our interest is in the lower-frequency range as discussed by Huang *et al.* (2005). In that case, the abundance of density over this range would emphasize the low-frequency components. However, we should always be clear on what we are computing to avoid confusion.

3.2. Delta function

Delta function is the most extreme, temporally local signal. It offers a useful tool to measure the time spread of the time-frequency analysis method. If we define a function consisting of 1,000 data points with zero everywhere except a unit value at the time point 500, we would have an empirical delta function. To decompose the delta function, we have to add noise as in Flandrin *et al.* (2005) or use the ensemble EMD as in Wu and Huang (2009). The resulting 5 IMFs and a residue are shown in Fig. 9. Then, with various number of bins at 10, 20, 30, 50, 100, 300, 500, 600, and 800 respectively, we have the different Hilbert spectra as given in Fig. 10. Even though the time spread is quite limited for any of these spectra, the frequency spread of all the spectra is considerable. The marginal spectra derived from each Hilbert

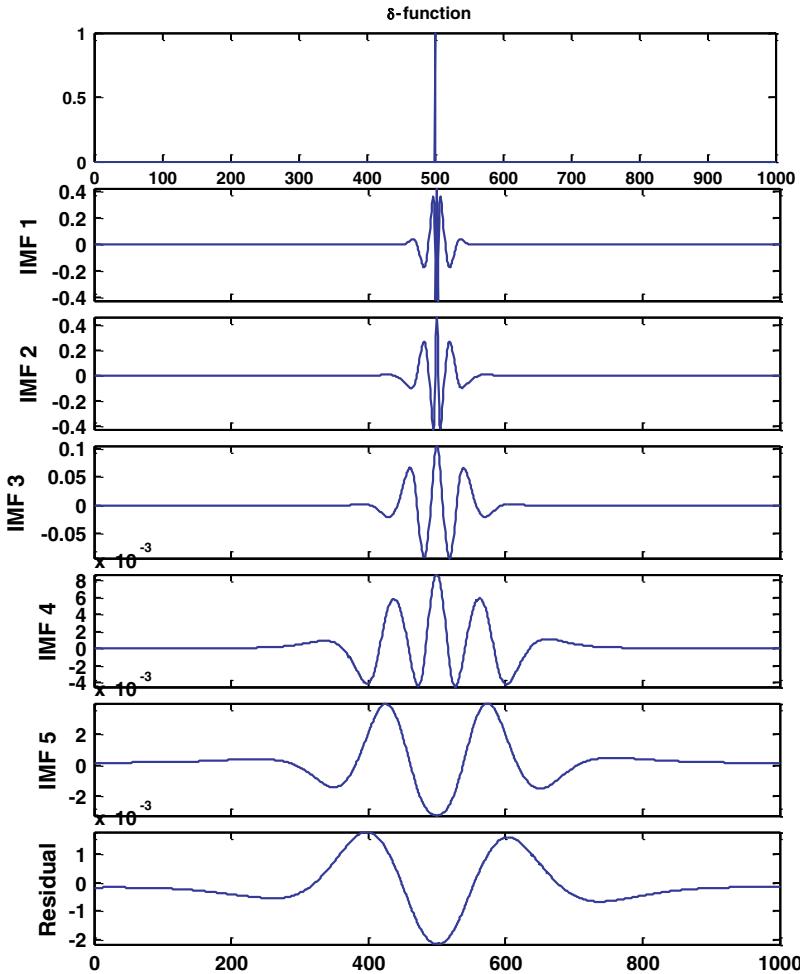


Fig. 9. Delta function data and the IMFs obtained through ensemble EMD.

spectra are given in Fig. 11. In Fig. 11, all the spectra have a flat spectral form in agreement with the Fourier spectrum. These results again validate the conversion formula given above.

It is interesting when we compare the spectra of white noise and delta function. The marginal spectra indeed look alike, and the time-frequency spectra are very different. The reason is obvious, for Fourier spectral representation is technically a marginal spectrum: because Fourier transform totally eliminates temporal variable, in the time-frequency presentation it is a cylindrical form. The only information resides in the projection on the frequency-energy space. Based on this comparison, we can conclude that, for nonstationary processes, the Fourier spectrum give limited, if not erroneous, information. It should never be used to represent such data.

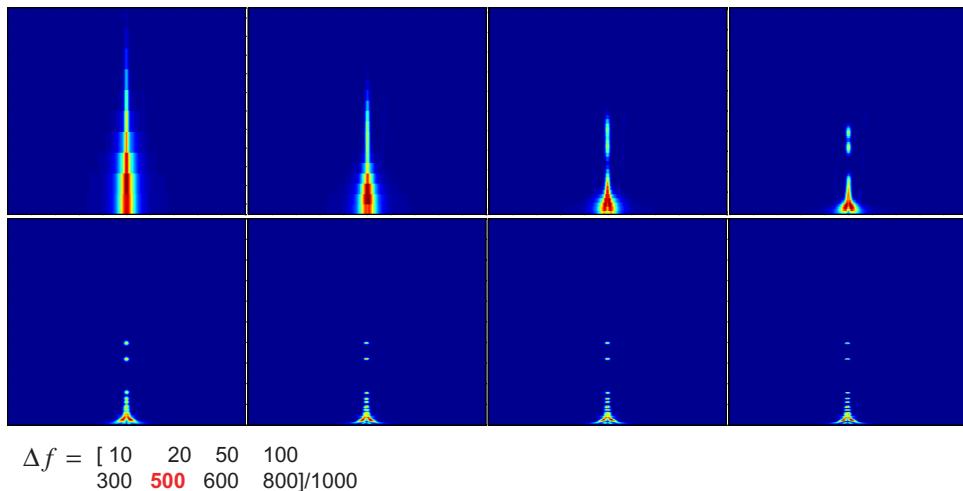


Fig. 10. Hilbert spectra for the delta function with different bin numbers equal to 10, 20, 50, 100, 300, 500, 600, and 800, respectively.

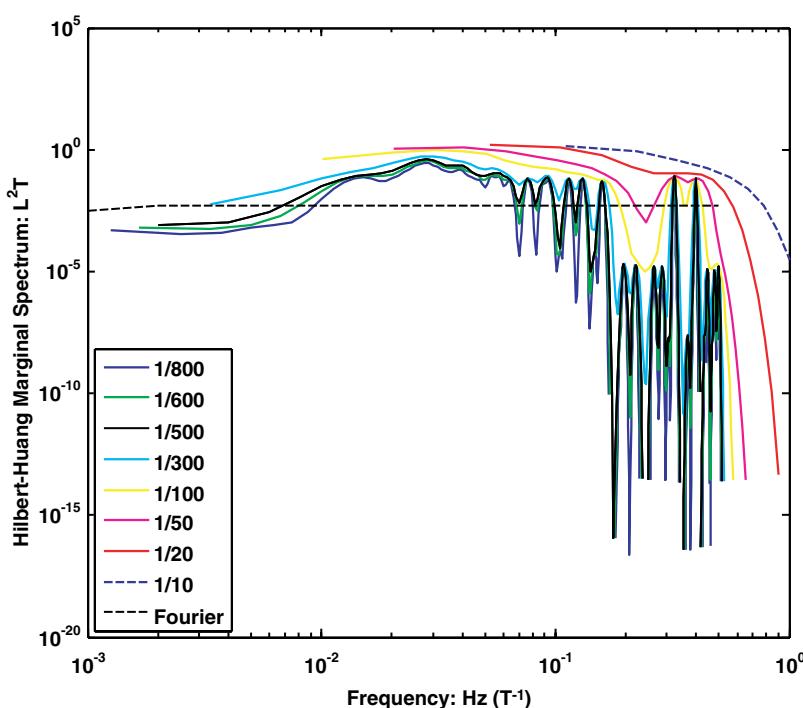


Fig. 11. Converted Hilbert marginal spectra for the Hilbert spectra given in Fig. 10 and the Fourier spectrum, which is a straight flat line.

The time spread presented here for delta function indicates that EMD as implemented here has considerable amount of leakage. This leakage could be eliminated if more precise decomposition method is used. An alternative definition for EMD has been developed by Hou and Shi (2011) that would treat the delta function as a single IMF component, which actually would limit the time to the local data point and the frequency to a single value of the Nyquist. In other words, the delta function represents a fastest change in phase function; therefore, it should have the highest frequency allowed, or the Nyquist value.

3.3. Application to a earthquake data set

Earthquake motions are all transient, therefore, nonstationary. The ground motions from strong near field earthquakes are far from linear. Consequently, the logical choice to represent the earthquake data should be a time-frequency representation that could accommodate both nonstationary and nonlinear characteristics in the ground motions. HHT seems to be a unique choice to fulfill this need. In fact, the earthquake signals have been studied by Loh *et al.* (2001) and Huang *et al.* (2001) using HHT. Having established the validity of the conversion formula, we apply the method to an earthquake data set and compare them with the traditional methods. The data used here had actually been used once already by Huang *et al.* (2001). Unfortunately, Huang *et al.* (2001) had used the then default definition of the Hilbert amplitude spectrum and made comparisons between it with the response spectrum. As the response spectrum is defined in terms of energy and it indeed is equivalent to the Fourier spectrum, the comparison is incompatible and not very meaningful. We now make the proper comparisons between the Hilbert energy spectrum and the corresponding Fourier spectrum.

The data used here is the set for the east-west direction acceleration at station TCU129 during the 21 September 1999 Chi-Chi earthquake. The data shown in Fig. 12 are the raw data sampled at 200 Hz for the duration of 70 s. With regular EMD, we obtained the IMF components as given in Fig. 13, where a total of 13 components were obtained. The Hilbert spectrum is given in Fig. 14 and the marginal Hilbert and Fourier spectra are given in Figs. 15(a) and 15(b). The marginal spectrum is obtained with the simple conversion factor of $2N/T$ as given in Eq. (13). One can multiply this factor with Fourier spectrum to get the Hilbert energy distribution unit, or divide this factor through the Hilbert spectrum to get the Fourier energy density unit. At any rate, from Figs. 14 and 15 it is easily seen that the marginal Hilbert spectrum is almost identical to the Fourier spectrum. Indeed, in principle this should be the case, for the marginal spectrum, in a sense, is the time integral of the time-frequency full spectrum; consequently, the temporal variations are averaged out. Mathematically, the integration actually gives the mean value over the full duration, which is essentially what the Fourier spectrum is designed to represent. On detailed inspection, however, one still could find subtle but important differences between the Fourier and the marginal Hilbert spectra. The differences are discussed in the following paragraphs.

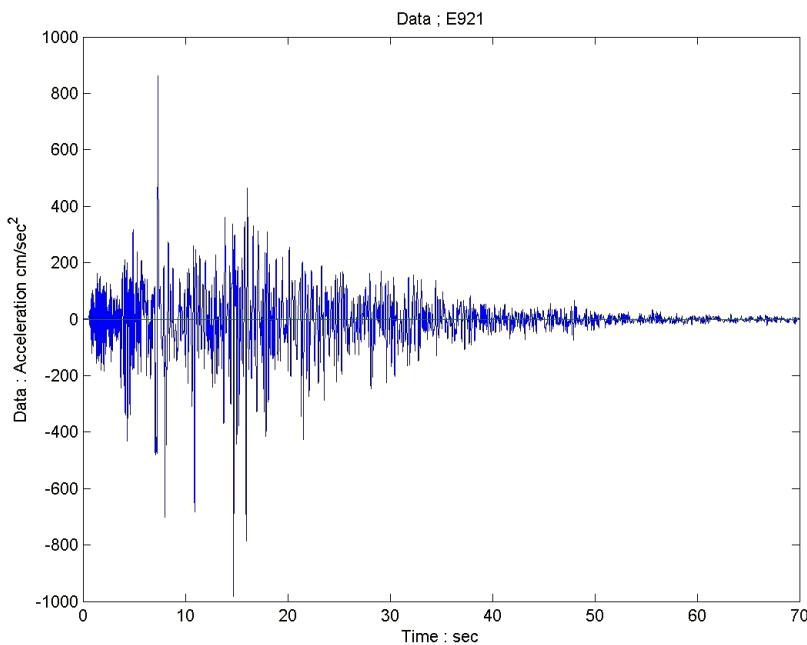


Fig. 12. The east-west acceleration of the earthquake data at station TCU129 during the 21 September 1999 Chi-Chi earthquake.

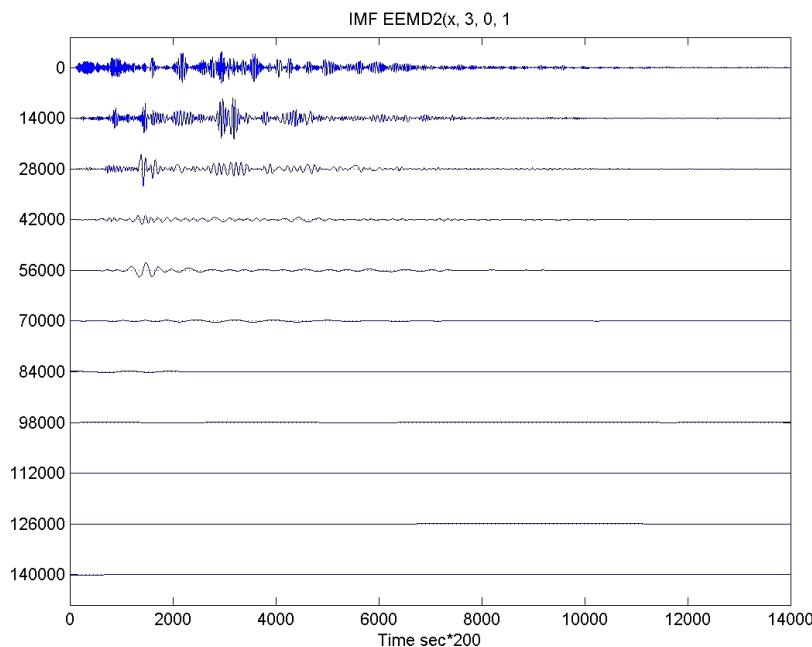


Fig. 13. IMF for the earthquake data given in Fig. 12.

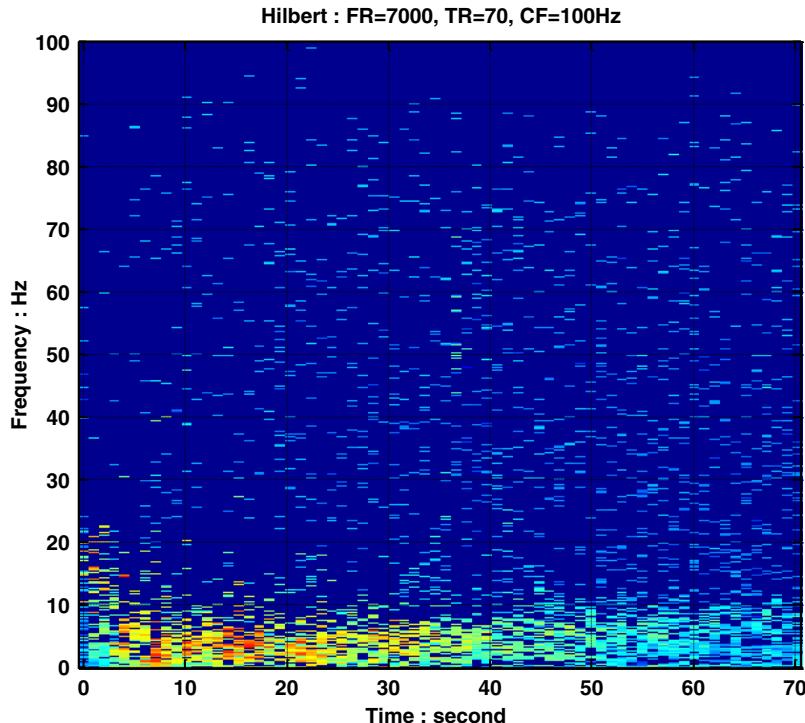


Fig. 14. Full resolution Hilbert spectrum with the frequency range covering 0–100 Hz as depicted by the Fourier analysis.

As the earthquake is nonstationary and nonlinear, this spectral representation in the mean does not give the full representation of the complicated underlying processes as shown in Fig. 14, where one could easily see the details of how the different waves arrive from the epicenter to the station: the first to arrive were the compression waves of small amplitude but higher frequency (from 10 to 20 Hz with a mean at 15 Hz) for the first 3–4 s or so. These compression waves, in general, are not capable of causing any damages. Immediately following these compression waves, the shear and surface waves would arrive. The magnitude is much stronger and the frequency is lower (with strongest energy below 5 Hz). The shear and surface wave will eventually separate through the dispersive property of the surface waves. The lower-frequency wave arrived sooner and the slower high frequency waves arrived later showing an energy band with a gradual increasing frequency from around 5 Hz to 8 or 9 Hz over the full duration of the data span. The other body shear wave is not dispersive and maintained a near constant frequency around 4 Hz through the time span. These details are all obliterated through the time integrated marginal spectrum. Therefore, for data from a genuine nonstationary and nonlinear process the detailed time-frequency spectral representation should definitely be the choice.

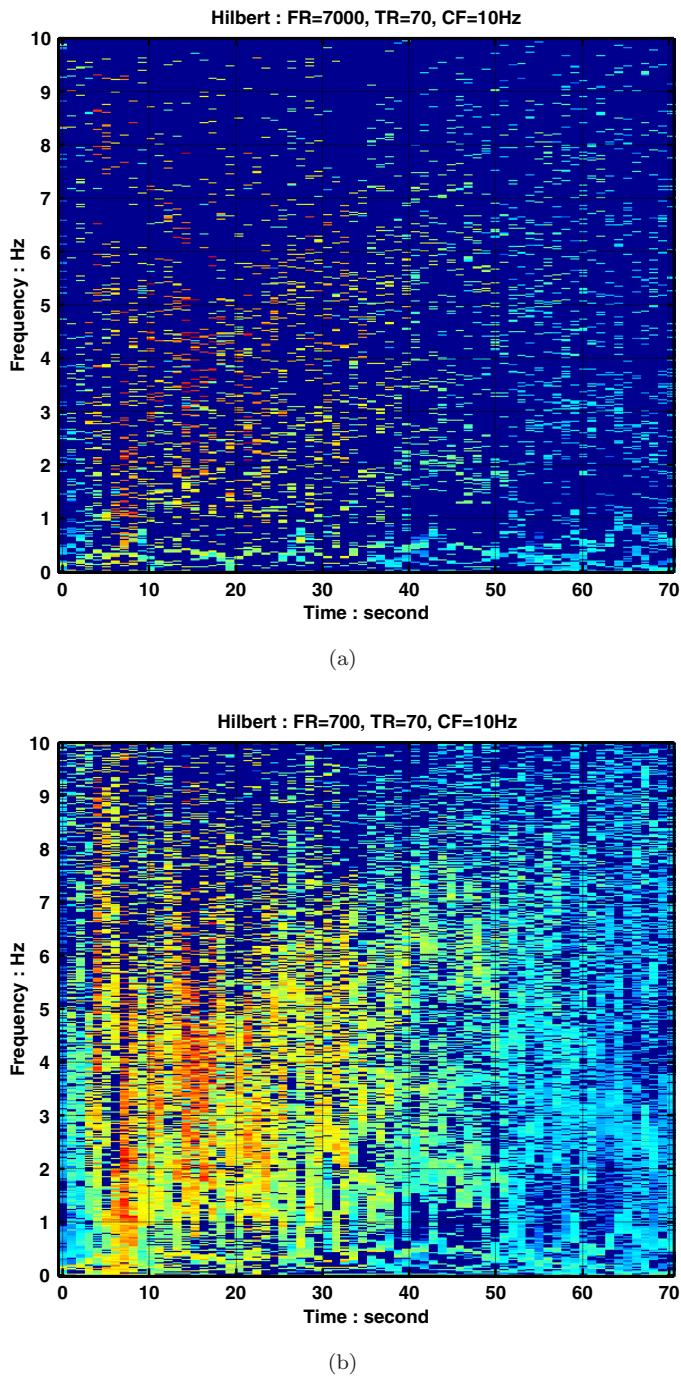
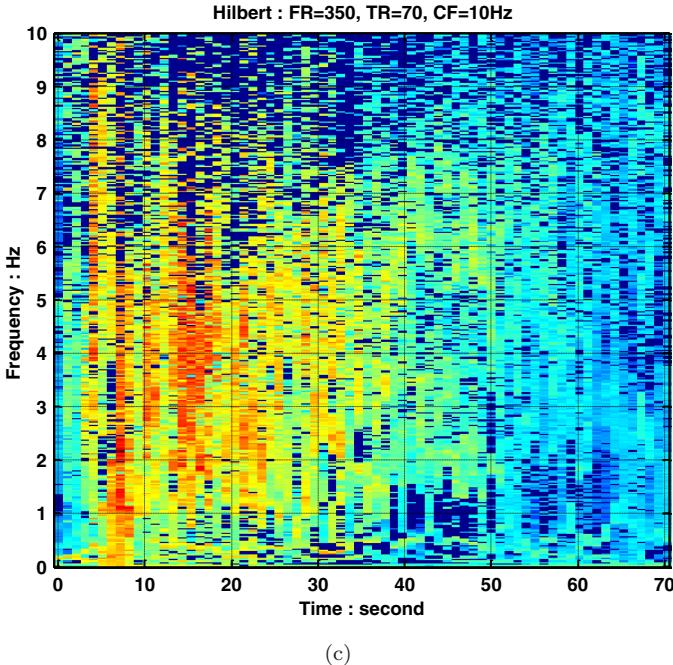


Fig. 15. Fourier and Hilbert marginal spectra with full Fourier resolution (a) and with the 7×7 smoothed Hilbert spectrum (b).



(c)

Fig. 15. (Continued)

This full resolution spectrum representation has wasted valuable time-frequency space, for most of the information critical to the engineering aspect of earthquake reside in the range below 10 Hz. If one use Fourier-based analysis, there is no way one could zoom in this range and study the details of the ground motion. Here, with HSA, we could zoom in all the resolution power to any specific frequency range. To show the effect of different bin numbers in the HSA, the marginal Hilbert spectra with different resolution are shown in Figs. 16(a) and 16(b). In Fig. 16(a), the different marginal Hilbert spectra are identified by the number of bins used in the analysis of all with the upper frequency limit set at the Nyquist value. The case with 70 bins is a 100 times coarser than the case with 7,000 bins. Consequently, the low frequency cut-off values differ by a factor of 100. As the same amount of energy is present in both representations, the energy distribution in the 70-bin case would be 100 times more than the 7,000-bin case. With this notion, the additional normalization factors are easily implemented as shown in Fig. 16(b): Factor = 2, 5, 10, 50, and 100 for the bin numbers at 3, 500, 1,400, 700, 140, and 70, respectively.

Now, let us examine the spectral resolution in more details to emphasize the difference between the Fourier and the Hilbert marginal spectra. From either Figs. 16(a) or 16(b), one can immediately see that all the Hilbert marginal spectra have a much sharp decay of energy density after the 20 Hz range; while, the Fourier spectrum has a near power law decay. The physical meaning is clear here: the Hilbert marginal spectrum reveals that the instantaneous frequency has eliminated

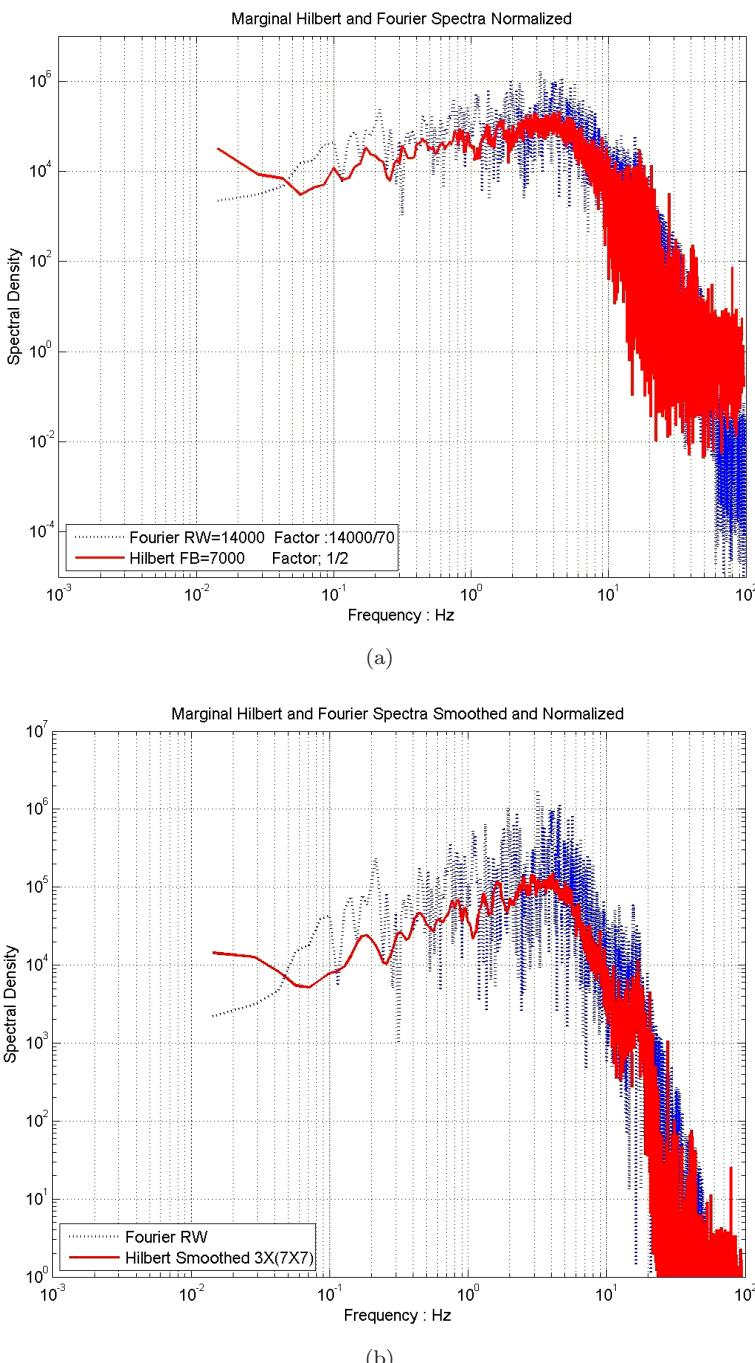


Fig. 16. Comparison of the effects of bin sizes: (a) Unnormalized Hilbert marginal spectra and Fourier spectrum for the earthquake data with various bin numbers; and (b) the corresponding normalized spectra.

the need of harmonics; therefore, beyond the initial compressive waves, there is no more earthquake signal. Consequently, the energy density drops precipitously immediate beyond 20 Hz to a flat level indicating the characteristics of noise. The Fourier spectrum, on the other hand, needs the harmonics to represent any nonlinear oscillations. As the near field earthquake motions are highly nonlinear, there will be a richness of harmonics. Consequently, the high frequency energy density drops off following a power law form. The spectral density in this range is not real, but it is the consequence of mathematical artifacts particular to the basis used in the time-frequency transform. It should be ignored in physical interpretation of the phenomenon.

Next, we should examine the effects of zooming. In many applications, instead of the full range of frequency representation up to the Nyquist value, we would like to zoom in some special frequency subrange. For example, in the earthquake data, in order to get detailed enough information, the sampling rate is set at 200 Hz and the Nyquist value is at 100 Hz. However, for engineering applications most of the information in this 100 Hz frequency range would be useless as shown in Fig. 14. The richest information and also the critical frequency range that would matter in the structure design are lower than 10 Hz. With Fourier spectral representation, we have no choice but to go for the full frequency range from 1/70 Hz all the way to the Nyquist value of 100 Hz. The result is that, with 90% of the data in the useless range, the useful part ends up with a very poor resolution. In the HSA, we have this additional choice: to put all the frequency resolution power in a frequency subrange of interest to us. We demonstrate both the zooming and bin size effect in Fig. 17. Figure 17(a) actually shows the Hilbert spectrum with full 7,000 bins covering the frequency range only up to 10 Hz. Figures 17(b) and 17(c) are the Hilbert spectra for bin sizes 700 and 350, respectively. As the bin size decreases, there will be more total energy in each bin and the pattern could show up clearly. This clear pattern, however, is gained at the price of resolution. The two marginal spectra with 7,000 and 700 bin sizes are given in Fig. 18. Here, the low frequency cut-off is extended to 10/7,000 Hz for the 7,000 bin size case. The details of the spectral energy variations in this range are in full view here. For tall-building design, the rule-of-thumb critical resonant periods are set at $S/10$ s, with S as the number of stories of the building of interest. For a 50-story building, the resonant wave would have a period of 5 s (or at 0.2 Hz). The sampling rate of 200 Hz for 70 s could barely cover it. Any frequency lower than this would be a problem. With the zooming power, we could use the 70 s data length to cover the range almost down to 0.01 Hz and still offer some frequency variations. Inherent with the zooming power is the complication of normalization factor. As presented in Eq. (16), all we have to add is another factor to cover the zooming effect. The normalized marginal spectra are also shown in Fig. 18. The details offered by the zooming are clearly seen: the low-frequency coverage even with 700 bins is comparable with the Fourier spectrum. The case with 7,000-bin size had extended the frequency down to 1/100 Hz. The richness of energy in the high-frequency range is overwhelming. Comparing with the Fourier spectrum, we

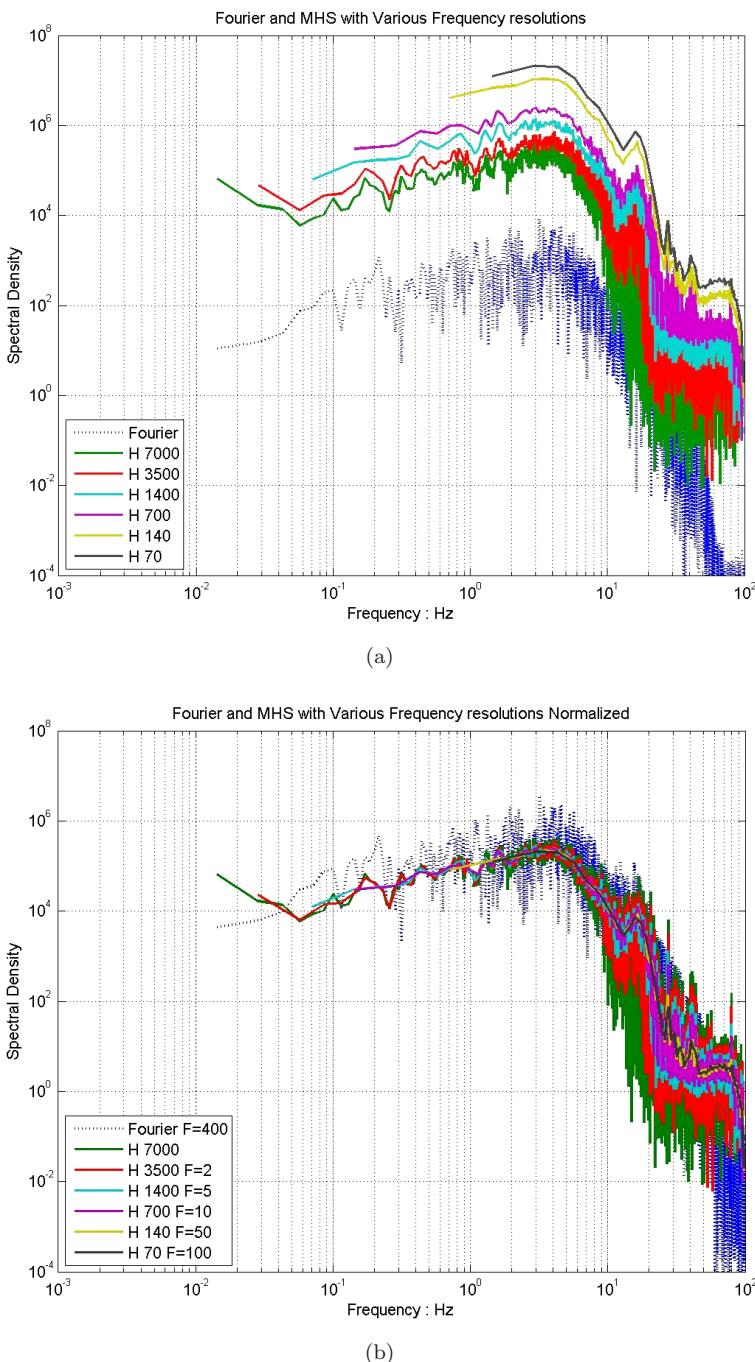


Fig. 17. Zoomed Hilbert spectra for frequency range 0–10 Hz with different bin sizes: (a) 7000 bin; (b) 700 bin; and (c) 350 bin.

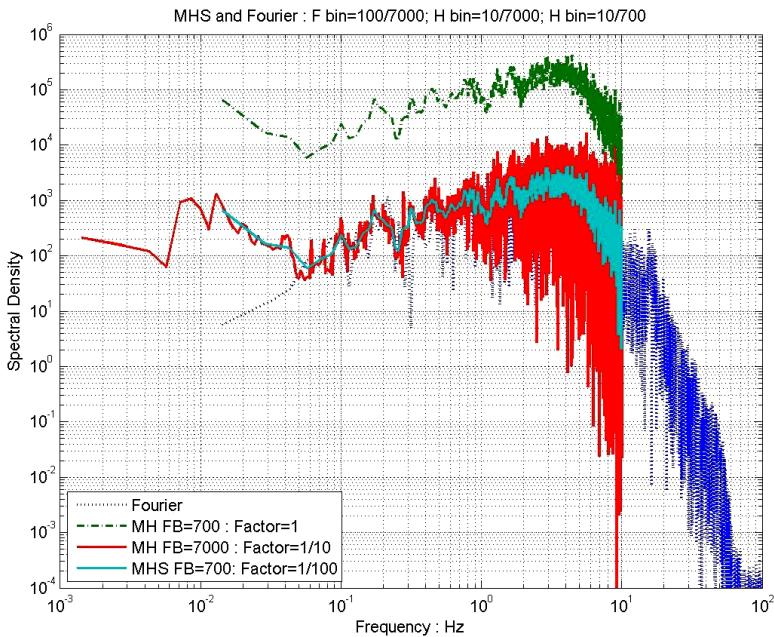


Fig. 18. Hilbert spectrum at 10 Hz cut-off frequency with various bin sizes.

found that the relative poor resolution case with 700 bins could offer comparable information already. The 7,000-bin size would be an over kill. With the zooming and bin number selection, we could save a lot of computation cost. Thus, with the above demonstrations, we could use the combination of zoom and selected bin number with full range of freedom.

Once we have the normalization rules settled, we should revisit the discrepancy presented in Huang *et al.* (1998) and (2001). In both cases, the Hilbert amplitude spectra were used to compare with Fourier power spectrum. If we plot the Hilbert power spectrum together with the Fourier power spectrum, we would have the results shown in Fig. 19. Here, an abundance of energy indeed shows up. The information offered though is interesting and could be important as shown by Huang *et al.* (2005), the comparison is meaningless.

4. Discussions and Conclusions

Having established the versatility of the Hilbert spectral representation, we provide a comparison with some current available method for time-frequency analysis methods. The example here is to compare among the spectrogram and wavelet. The signal is the recording of the word, “hello,” digitalized at a rate of 20,400 Hz as shown in Fig. 20. The top row shows the spectrogram: the first spectrogram on the left is for the narrow band representation (window length at 1,024 data points), designed based on the compromise between maximum frequency resolution and

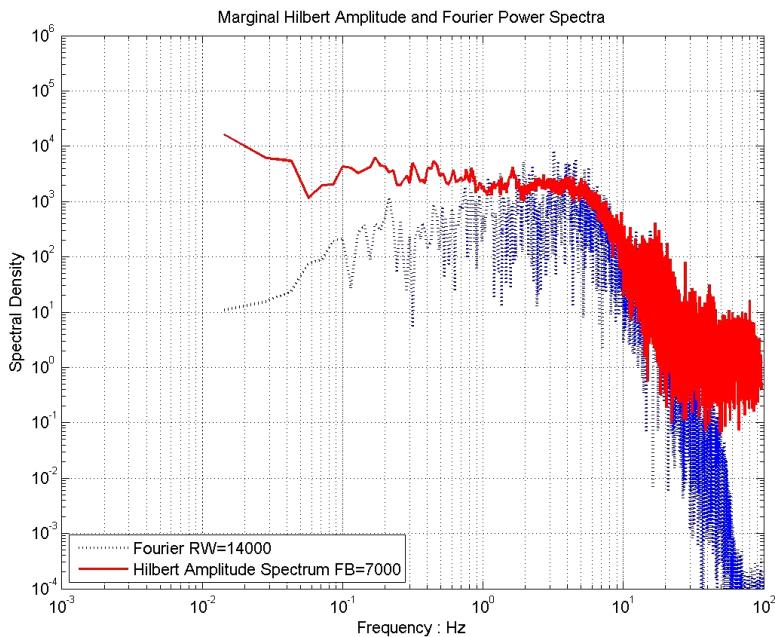


Fig. 19. Comparison of the Hilbert amplitude and Fourier spectra with the Fourier power spectrum. The Hilbert amplitude spectrum shows excessive energy in the low-frequency range.

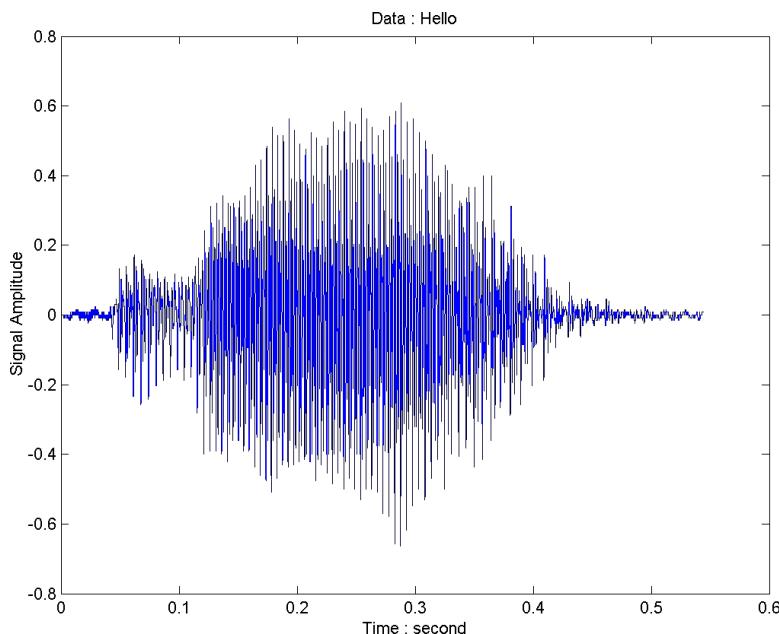


Fig. 20. Speech signal for the sound, "Hello."

the degree of temporal variations. With the emphasis on frequency resolution, one could see the regular multi-band harmonics, which are all spurious mathematical artifacts produced by using the linear Fourier analysis to represent the nonlinear speech producing sound signal as discussed by Huang *et al.* (1998, 1999). The second spectrogram on the left is the wide band representation (window length at 64 data points), designed based on the compromise between revealing the maximum degree of temporal variation and some degree of frequency resolution (Fig. 21). Here, the time step is much more precise, but the frequency resolution is smeared to the degree of nearly useless. The fact that we could not obtain high degree of precision in both frequency and time is known as the “uncertainty principle” in data analysis, which states

$$\Delta t \times \Delta \omega \geq \frac{1}{2}, \quad (20)$$

in which the Δt and $\Delta \omega$ are the time and frequency step selected, respectively, or the resolution or the degree of precision for the particular spectrogram. In physics, the uncertainty principle is based on sound arguments and supported by experimental

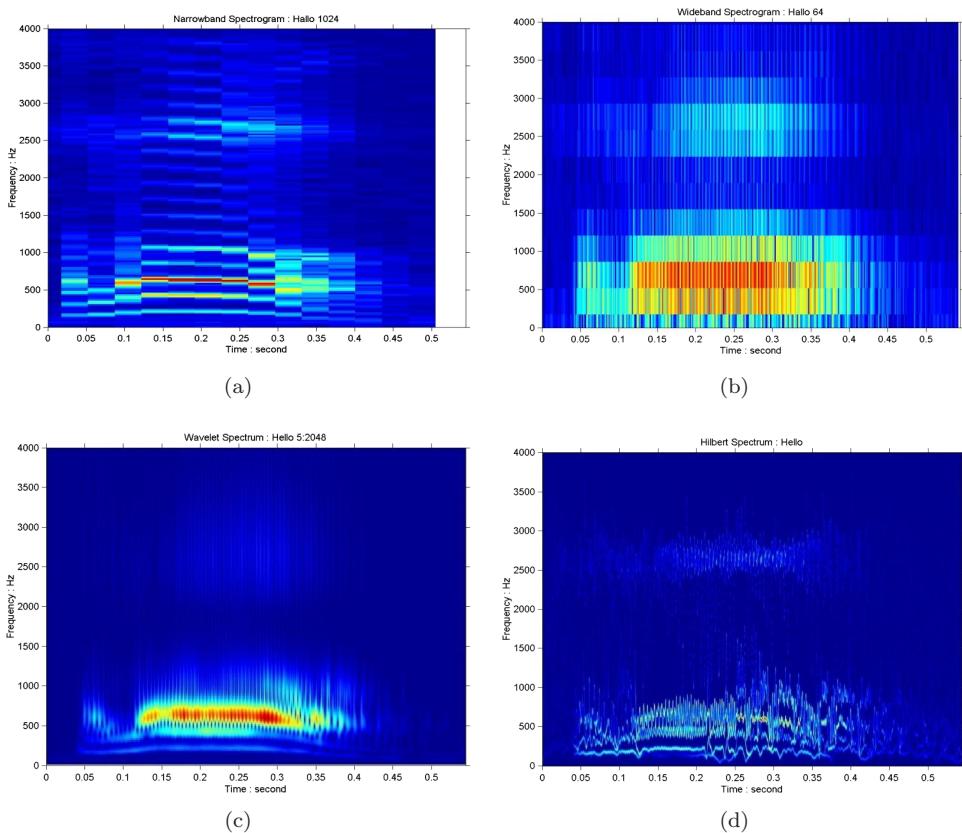


Fig. 21. Comparison among the spectrograms (a) narrow band and (b) wide band; (c) continuous Morlet wavelet and (d) Hilbert spectral representations for the speech signal given in Fig. 20.

observations. In data analysis, the “uncertainty principle” is purely the consequence of using integral transformation in time-frequency conversion: the integration operation would smear the temporal and frequency resolutions within the integral ranges. Therefore, the uncertainty principle is a tool-dependent phenomenon, not related to any fundamental physical law. If we eschew integral operations in time-frequency conversion, as in the HSA, where frequency is computed through differentiation rather than integration, we would not have such spurious restriction. As in the HSA, the frequency is determined by differentiation rather than the integral transform, the time could be determined within the resolution of digitalizing step, and the frequency could be precise to any number below the Nyquist value. Consequently, we could obtain arbitrary precision on time and frequency resolution subject only to the sampling rate. This approach reserves the maximum degree of locality; therefore, the frequency resolution here does not depend on the data length, except naturally the low frequency cut-off values and the high frequency Nyquist value.

The third representation is the continuous wavelet analysis on the lower left. Here, we can see the adjustable window used in the wavelet analysis seems to have offered some improvements for the visual impression of the time-frequency representation. Unfortunately, the actual quantitative values do not support this impression, for in the wavelet basis design, we have to reach a compromise between temporal and frequency resolution and thus reduce the power of frequency resolution. As the continuous wavelet is overly redundant, it is not amenable to quantitative analysis. The frequency values are always overly smoothed. The discrete wavelet, although improved on orthogonality, still suffers poor time-frequency resolution. As a result, wavelet analysis is not regarded as a standard time-frequency analysis tool. Furthermore, the wavelet analysis is based on integral transform, the same uncertainty principle restriction applied here too.

The HSA presented here in the lower left corner shows the sharp time-frequency resolution, which could be achieved at an arbitrary degree of precision subject only to the data sampling rate. The instantaneous frequency definition has enabled us to overcome the restriction of the uncertainty principle. The instantaneous frequency approach using the intrawave frequency has eliminated the spurious harmonics totally. Thus, the HSA has indeed offered us a practical, versatile, and physically meaningful time-frequency analysis tool.

There is still another popular time-frequency representation method, the Wigner–Ville distribution (Flandrin, 1995), we need to comment on. The Wigner–Ville representation, $W(\omega, t)$, is defined as

$$W(\omega, t) = \int_0^T \overline{x\left(t + \frac{\tau}{2}\right)} x\left(t - \frac{\tau}{2}\right) \exp(-i\omega\tau) d\tau \quad (21)$$

There is no substantive difference between the Wigner–Ville representation and a collection of the traditionally defined Fourier spectra

$$W(\omega, t) = \int_0^T \overline{x(t)x(t + \tau)} \exp(-i\omega\tau) d\tau \quad (22)$$

except the slight time shift. As a result, the time integral of Wigner–Ville distribution is identical to the Fourier power spectrum. This result is obvious, and it was erroneously and superficially used as a criterion in the so-called “the marginal requirement,” for judging time-frequency representation, implying the superiority of any representation if it could satisfy this superfluous integral requirement. This requirement is not logical: if Fourier spectral analysis should not be used for nonstationary processes, the marginal requirement should also be meaningless. As Fourier analysis could not resolve time variation, the instantaneous frequency in Wigner–Ville distribution defined by the moment method

$$\omega(t) = \frac{\int \omega W(\omega, t) d\omega}{\int W(\omega, t) d\omega} \quad (23)$$

would also have limited power to resolve time variation. With this definition, any signal should only have a single instantaneous frequency at any given specific time, which is the mean of all the available frequency at that time with no detailed differentiation of separate sources or processes. Based on this definition, the individual instruments that made the symphonic music beautiful as they are would be totally obliterated and reduced to a simplistic averaged sound. This conclusion is patently senseless. Other problems notwithstanding, because of this shortcoming, we decided not even to make a comparison with the results from this method.

The HSA could be made more statistically significant through the ensemble approach (Huang *et al.*, 2003), coupled with downsampling (Huang *et al.*, 2006). All these alternatives and the normalization to the energy density definition have made the Hilbert spectral representation a truly versatile time-frequency analysis method.

Acknowledgments

N. E. Huang has been supported by a grant from the Federal Highway Administration, DTFH61-08-00028, and the grants NSC 98-2627-B-008-004 (Biology) and NSC 98-2611-M-008-004 (Geophysical) from the National Science Council, Taiwan, and finally a grant from NCU 965941 that have made the conclusion of this study possible. He is also supported by a KT Lee endowed Chair at NCU. Z.W. was also sponsored by the NSF of USA grant ATM-0917743.

References

- Daubechies, I., Lu, J. and Wu, H.-T. (2009). Synchrosqueezed wavelet transforms: A tool for empirical mode decomposition. arXiv:0912.2437v1 [math.NA], <http://arxiv.org/abs/0912.2437>.
- Flandrin, P. (1995). *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, 386 pp.
- Flandrin, P., Rilling, G. and Gonçalves, P. (2004). Empirical mode decomposition as a filterbank, *IEEE Signal Proc. Lett.*, **11**: 112–114.
- Flandrin, P. and Gonçalves, P. (2004). Empirical mode decompositions as data-driven wavelet-like expansions. *Int. J. Wavelets, Multires. Inf. Process.*, **2**: 477–496.

- Flandrin, P., Gonçalves, P. and Rilling, G. (2005). EMD equivalent filter bank, from interpretation to applications. In *Hilbert-Huang Transform and Its Applications*, eds. N. E. Huang and S. Shen, World Scientific, Singapore, pp. 57–74.
- Gledhill, R. J. (2003). *Methods for Investigating Conformational Change in Biomolecular Simulations*, Ph.D. Dissertation, Department of Chemistry, University of Southampton, England.
- Hou, T. Y., Yan, M. P. and Wu, Z. (2009). A variant of the EMD method for multi-scale data. *Adv. Adap. Data Anal.*, **1**: 483–516.
- Hou, T. Y. and Shi, Z. Q. (2011). A total-variation-diminishing-based optimization method for analyzing nonlinear nonstationary data, *Adv. Adapt. Data Anal.*, **3**.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, E. H., Zheng, Q., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis. *Proc. Roy. Soc. Lond.*, **454A**: 903–993.
- Huang, N. E., Shen, Z. and Long, S. R. (1999). A new view of non-linear water waves — the Hilbert spectrum, *Ann. Rev. Fluid Mech.*, **31**: 417–457.
- Huang, N. E., Chern, C. C., Huang, K., Salvino, L., Long, S. R. and Fan, K. L. (2001). Spectral analysis of the Chi–Chi earthquake data: Station TUC129, Taiwan, September 21, *Bull. Seism. Soc. Am.*, **91**: 1310–1338.
- Huang, N. E., Wu, M. L., Long, S. R., Shen, S. S., Qu, W. D., Gloersen, P. and Fan, K. L. (2003). A confidence limit for the empirical mode decomposition and Hilbert spectral analysis, *Proc. Roy. Soc. Lond.*, **459A**: 2317–2345.
- Huang, N. E., Huang, K. and Chiang, W.-L. (2005). HHT-based bridge structural health-monitoring method. In *Hilbert-Huang Transform and Its Applications*, eds. N. E. Huang and S. Shen, World Scientific, Singapore, pp. 263–287.
- Huang, N. E., Brenner, M., Pak, C.-G. and Salvino, L. (2006). An application of Hilbert–Huang transform to the stability study of airfoil flutter, *AIAA J.*, **44**: 772–786.
- Huang, N. E. and Wu, Z. (2008). A review on Hilbert–Huang transform: The method and its applications on geophysical studies, *Rev. Geophys.*, **46**: RG2006, doi:10.1029/2007RG000228.
- Huang, N. E., Wu, Z., Long, S. R., Arnold, K. C., Chen, X. and Blank, K. (2009). On instantaneous frequency. *Adv. Adap. Data Analy.*, **1**: 177–229.
- Loh, C.-H., Wu, T.-C. and Huang, N. E. (2001). Application of EMD-HHT method to identify near-fault ground motion characteristics and structural responses, *Bull. Seism. Soc. Am.*, **91**: 1339–1357.
- Olhede, S. and Walden, A. T. (2004). The Hilbert spectrum via wavelet projections. *Proc. Roy. Soc. Lond. A*, **460**: 955–975.
- Pao, S. H., Young, C.-N., Tseng, C.-L. and Huang, N. E. (2010). Smoothing empirical mode decomposition: A patch to improve the decomposed accuracy. *Adv. Adap. Data Analy.*, **2** (in press).
- Wen, Y. K. and Gu, P. (2009). HHT-based simulation of uniform hazard ground motions. *Adv. Adap. Data Analy.*, **1**: 71–87.
- Wu, Z. and Huang, N. E. (2004). A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. Roy. Soc. Lond.*, **460A**: 1597–1611.
- Wu, Z. and Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adap. Data Analy.*, **1**: 1–41.
- Wu, Z., Huang, N. E. and Chen, X. (2009). Multi-dimensional empirical mode decomposition based on ensemble empirical mode decomposition. *Adv. Adap. Data Analy.*, **1**: 339–372.