

Spectrotemporal Modeling of Biomedical Signals: Theoretical Foundation and Applications

Raymundo Cassani and Tiago H Falk, Institut National de la Recherche Scientifique (INRS-EMT), University of Quebec, Montreal, QC, Canada

© 2018 Elsevier Inc. All rights reserved.

Introduction	2
Conventional Time and Frequency Domains	3
Spectrotemporal Representation	5
Short-Time Fourier Transform Approach	6
Continuous Wavelet Transform Approach	7
Hilbert Transform Approach	8
Spectrotemporal Computation Methods Comparison	10
Modulation Spectrogram	12
Biomedical Applications	15
Analysis Applications	15
Detection of respiration frequency from the ECG signal	15
Quality assessment for ECG	17
Neuronal amplitude modulations	17
Synthesis Applications	18
Separation of heart and lung sounds from digital auscultation	19
ECG enhancement	19
Conclusion	19
Further Reading	20

Glossary

Analytic signal The analytic signal is a complex temporal representation of a real-valued signal. The real and imaginary parts of the analytic signal are real-valued functions, with the imaginary part corresponding to the Hilbert transform of the real part. Negative frequency components in the spectrum of the analytic signal are equal to zero.

Amplitude modulation Process where the amplitude of a high frequency signal, called carrier, varies according to a lower frequency signal called the modulating signal.

Gabor uncertainty principle Also known as the duration-bandwidth product, it states that the product of temporal and spectral uncertainties must exceed a fixed constant, therefore implying that reducing temporal uncertainty increases frequency uncertainty and vice versa. The exact value of the constant depends on how the uncertainties are measured.

Modulation Process where a high frequency signal, called carrier, varies its properties such as amplitude, phase, and frequency according to a lower frequency signal called the modulating signal.

Modulation spectrogram Complex-valued function of a time signal that portrays the over-time amplitude changes for the spectral components of the time signal in the frequency modulation domain.

Short-time Fourier transform Method based on the Fourier transform to calculate the spectrotemporal representation of a signal. It consists in obtaining the Fourier transform for small segments of the signal.

Spectrum Complex-valued function that corresponds to the frequency-domain representation of a time signal.

Spectrogram Complex-valued function that corresponds to the representation of a time signal in the time-frequency domain.

Nomenclature

FT	Fourier transform
CWT	Continuous wavelet transform
CWT	Hilbert transform
STFT	Short-time Fourier transform
IFT	Inverse Fourier transform
AM	Amplitude modulation
EEG	Electroencephalogram
ECG	Electrocardiogram

PPG	Photoplethysmogram
$x(t)$	Real-valued time signal
$X(f)$	Complex-valued spectrum of a time signal
$X(t, f)$	Complex-valued spectrogram of a time signal
$X(f, f_{mod})$	Complex-valued modulation spectrogram of a time signal
$\hat{x}(t)$	Analytic signal for a time signal
$\hat{x}(t)$	Hilbert transform for a time signal
$\text{sgn}(x)$	Sign function
$H(x)$	Heaviside function

Introduction

Commonly, signal processing techniques and methods assume (and frequently require) that the signals be stationary, that is, that their statistical properties do not change over time. Stationarity, however, is a strict assumption rarely met by naturally occurring signals. As such, for practical purposes signals are assumed stationary if they fulfill a softer stationarity definition termed "wide-sense stationarity," which requires a constant mean and a time-invariant autocorrelation function. For a zero-mean signal, the equivalent of these conditions in the frequency domain is that of a time-invariant spectrum. Henceforth, the term stationary will be used for strict-sense and wide-sense stationary signals. In practice, in order to analyze nonstationary signals (e.g., biomedical signals) using conventional signal processing tools, shorter time segments are considered in which the signal can be treated as stationary. The duration of these stationary segments varies depending on the nature of the signal. For example, sound signals are assumed stationary in segments between 20 and 50 ms. Spontaneous electroencephalography signals, in turn, can present stationary behavior for segments as long as 20 s. Therefore, in order to analyze nonstationary signals, it is necessary to have a representation capable of registering the changes in frequency content over time, that is, a spectrotemporal representation.

Interestingly, for a great variety of nonstationary signals, although their spectral content is not constant over time, changes occur in single or multiple periodic ways, thus resulting in a property called cyclostationarity, which is a second-order periodicity. Thus, a nonstationary signal can be modeled as the result of the interaction of two independent signals, a low-frequency signal that changes (modulates) the properties (such as amplitude, phase, and/or frequency) of a higher frequency signal. This interaction is a well-known nonlinear process called modulation. When the signal is modeled as the product between a modulating signal (low-frequency), and a carrier signal (high-frequency), the modulating signal changes exclusively the amplitude of the carrier signal; this process is denominated amplitude modulation (AM).

Given the simplicity of its formulation, the amplitude modulation (AM) modeling approach has shown to be an important tool for the analysis and synthesis of nonstationary processes in diverse fields such as telecommunications, oceanography, speech, and music perception, as well in the study of biological signals, such as the electrocardiogram (ECG), electroencephalogram (EEG), photoplethysmogram (PPG), respiration patterns, and neuronal perception of stimuli, among others. Indeed, it has been argued that the presence of amplitude modulation in biological signals is a consequence of the processes of control, synchronization, regulation, and intersystem interaction found in biological systems; therefore the AM analysis plays a pivotal role on modeling, interpretation, and understanding of those processes. An example of AM in biomedical signals is the case of the ECG, depicted by Fig. 1. As can be seen, the modulating effect of the slow-changing respiration pattern signal (Fig. 1A) can be noticed in the ECG signal amplitude (Fig. 1B). This modulation process has its origin in the relative movement between the heart and the electrodes

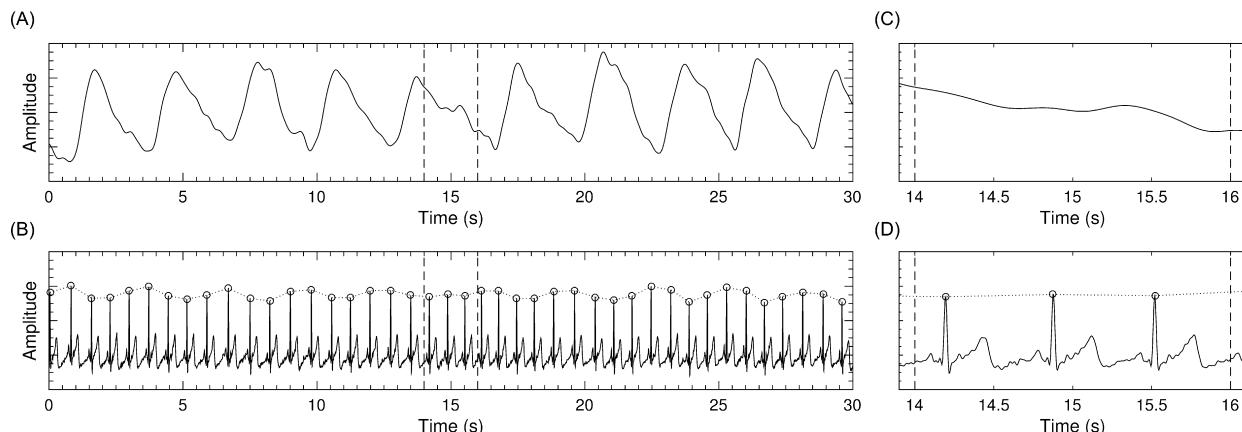


Fig. 1 Amplitude modulation of the ECG signal by the respiration process. Panel A shows the respiration signal. Panel B depicts ECG signal, indicating the location of the R-peaks (circles) and their amplitude changes over time (dashed line). Panels C and D present the respiratory and ECG signal during a shorter segment.

(causing a displacement of the projection of the cardiac electrical vector), and the changes in the thoracic impedance, both events caused by changes in the thorax volume at every breathing. By analyzing the amplitude changes in the R-peak in the ECG signal (dashed line in Fig. 1B) it is possible to retrieve information about the undergoing respiration process. Not surprisingly, the analysis in short segments where the ECG signal is assumed stationary fails to represent the long-term dynamics of the process, as shown in Fig. 1C and D. Unlike this traditional approach, the AM analysis requires longer signal segments.

The analysis of spectral changes over time is essential for the study and understanding of biological systems from their measured nonstationary signals. The ultimate goal of this article is to present the theoretical foundation and show existing applications of the field of spectrotemporal signal analysis applied to biomedical signals. The article is organized as follows: First, a brief overview of conventional time and frequency domains is provided in section “Conventional Time and Frequency Domains”. This serves as preamble to section “Spectrotemporal Representation”, which introduces the concept of the spectrotemporal signal representation and presents three of the most utilized approaches to calculate the representation. In section “Modulation Spectrogram”, the modulation spectrogram is presented as a tool to characterize second-order periodicities in nonstationary signals and an open-source toolbox for modulation spectral analysis is introduced. Next, applications of the modulation spectrogram for analysis and synthesis of biomedical signals are presented in section “Biomedical Applications”. Lastly, the Conclusions section summarizes the main points on the use of spectrotemporal modeling for biomedical signal analysis.

Conventional Time and Frequency Domains

One of the pillars of signal processing is the back-and-forth transformation of signals between time and frequency domains. While there are several methods to perform these transformations, the most utilized are the Fourier transform (FT) and the inverse Fourier transform (IFT) for time-to-frequency and frequency-to-time transformation, respectively. The FT and IFT are mathematically defined as:

$$X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (1)$$

$$x(t) = \mathcal{F}^{-1}\{X(f)\} = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} df \quad (2)$$

where $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ are the FT and IFT operators, respectively; $x(t)$ is a signal in the time domain; and $X(f)$ is a complex-valued function called spectrum or Fourier transform of $x(t)$. As complex function $X(f)$ can be expressed as:

$$X(f) = |X(f)|e^{j\theta(f)} \quad (3)$$

where $|X(f)|$ is the amplitude spectral density or amplitude spectrum with an associated phase $\theta(f)$. Intuitively, the FT can be regarded as a spectral decomposition process that allows a time signal to be represented in terms of its frequency components (amplitude and phase). On the other hand, the IFT can be considered as the synthesis process where these frequency components are summed into a time signal, as shown in Fig. 2.

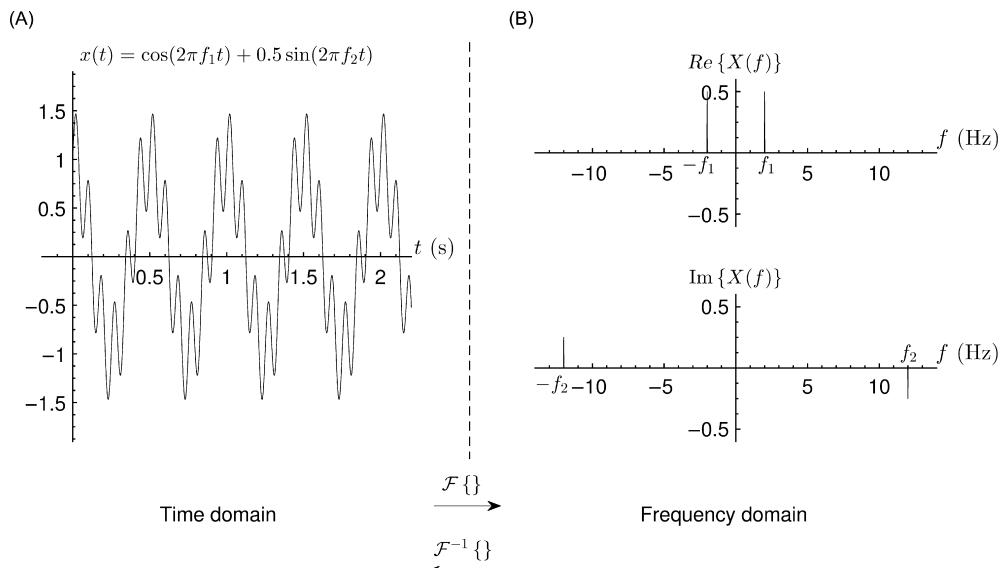


Fig. 2 Panel A depicts a time signal $x(t)$ comprised of cosine wave of frequency $f_1 = 2$ Hz and a sine wave with frequency $f_2 = 12$ Hz. Panel B shows the real and imaginary part of the spectrum $X(f)$ which is the frequency representation of the signal $x(t)$ (Panel A).

As $|X(f)|$ represents the amplitude across the different frequency components, a common representation of the spectrum of signal $x(t)$ is the energy spectral density or energy spectrum $|X(f)|^2$, which shows the distribution of the signal energy at different frequencies; thus the area under the curve $|X(f)|^2$ between $f_0 - df$ and $f_0 + df$ represents the energy for the frequency component at f_0 Hz.

A special case of the FT occurs if the time signal $x(t)$ is real-valued, then its corresponding spectrum $X(f)$ presents Hermitian symmetry. In this case, the spectral components for negative frequencies are the complex conjugate of the components for positive frequencies, that is,

$$X(f) = X^*(-f) \quad (4)$$

where $*$ denotes the complex conjugate, then $|X(f)| = |X(-f)|$. As a consequence, for real signals, it is common to present their amplitude and energy spectrum ($|X(f)|$ and $|X(f)|^2$, respectively) only for positive frequencies. Therefore, in order to keep the correct relation between spectral components and their respective amplitude and power, $|X(f)|$ and $|X(f)|^2$ values need to be doubled when only the positive frequencies are depicted.

Time and frequency domains provide complementary information about the same signal. For some applications handling the signal in the time domain is preferred over the frequency domain and vice versa. A clear example of these domain preferences is the processing of the ECG signal, as depicted by Fig. 3. For example, abnormal beat morphologies can be easily detected and characterized in the time domain (Fig. 3A) whereas the presence of 60 Hz powerline interference is more clearly noticeable in the frequency domain (Fig. 3B).

As the FT process can be reversed by the IFT, certain operations are more conveniently performed in one domain compared with the other. Among these operations, one that will be used throughout this article is the convolution, which is defined in time domain as:

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad (5)$$

where $f(t)$ and $g(t)$ are time signals and the $*$ symbol represents the convolution operator. While (5) appears overwhelming, it can be worded as the dot product of a signal $f(t)$ and the time-reversed version of $g(t)$ in function of a time shift τ between them. An important theorem derived from the FT is the convolution theorem, which indicates that convolution in one domain corresponds to the multiplication in the other domain. Being $F(f)$ and $G(f)$ the spectra for the above-mentioned time signals, the equations for the convolution theorem are:

$$f(t) * g(t) = \mathcal{F}^{-1}\{F(f) \cdot G(f)\} \quad (6)$$

$$F(f) * G(f) = \mathcal{F}\{F(f) \cdot G(f)\} \quad (7)$$

where the \cdot (dot) symbol denotes the point-wise product. Moreover convolution and multiplication operations share some relevant properties such as commutativity and associativity.

In the analysis of finite signals, the duration of the signal is related to the minimum differentiable frequency value such that at least one complete cycle of the frequency component is present to be able to be correctly differentiated, that is, resolved. More specifically, for a signal with duration of T s, the lowest resolvable frequency interval, that is, the frequency uncertainty Δf is given by $1/T$ Hz; moreover, for such a signal, its time uncertainty $\Delta t = T$ s. In this sense, having a longer signal, the more accurate the frequency measurement will be. For example, for a signal spread in time over a duration $\Delta t = 2 T$ s, its Δf is equal to $1/(2 T)$ Hz. In the opposite scenario, a signal with duration $T/2$ s, the time uncertainty Δt is reduced by half, but the uncertainty in frequency Δf is

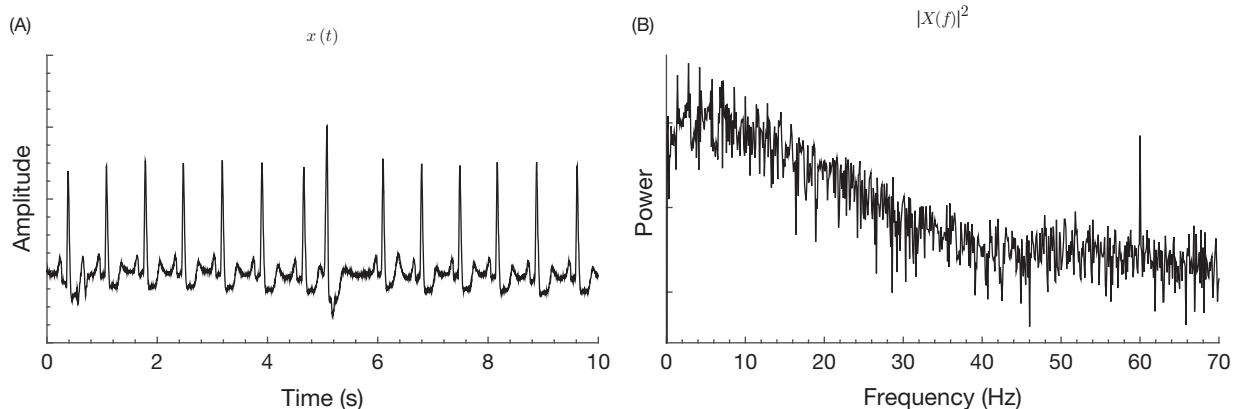


Fig. 3 Time and frequency representation of an ECG signal segment. Panel A presents the ECG signal in time domain where a premature ventricular contraction can be observed around 5 s. The frequency representation is depicted in Panel B, where the powerline interference is visible as the spike at 60 Hz.

doubled. This relationship is known as the Gabor uncertainty principle, where the product of uncertainties or “spreads” in time and frequency for a signal is greater or equal to a fixed constant c

$$\Delta t \Delta f \geq c \quad (8)$$

This principle is crucial in the spectrotemporal representation of signals described next.

Spectrotemporal Representation

Biomedical signals present nonstationary behavior, that is, their spectral content change over time; thus the FT does not sufficiently describe the dynamics of these signals. In this sense, a time-dependent spectral representation is needed. The fundamental idea to obtain the spectrotemporal representation is to split the signal in shorter segments and perform the time-to-frequency transformation for each segment. Fig. 4A and B depict a nonstationary time signal and its spectral representation, respectively. The signal is comprised of two major oscillations which are present in the spectrum; however, the information on when each oscillation occurred is lost. By dividing the signal in shorter segments (Fig. 4C, E, and G) and calculating their respective spectral representations (Fig. 4D, F, and H) it is possible to study the dynamics of the signal. As the spectral representation is calculated from shorter segments of the signal, the temporal uncertainty is reduced; thus it is possible to localize time changes in the spectral content. However, the value of the minimum resolvable frequency increases, that is, the frequency uncertainty augments, as per (8).

The spectrotemporal representation or spectrogram of a time signal is a complex-valued time-frequency function $X(t, f)$ that can be written as:

$$X(t, f) = |X(t, f)| e^{j\theta(t, f)} \quad (9)$$

with $|X(t, f)|$ being the amplitude spectrogram and $\theta(f, t)$ its associated phase. Analog to the power spectrum, the power spectrogram, is defined as $|X(t, f)|^2$. In the literature, it is common to find terms “power spectrogram” and the (complex-valued) “spectrogram” utilized as synonyms.

In the remainder of this section, three of the most utilized approaches to calculate the spectrotemporal representation are presented, namely based on the short-time Fourier transform, the continuous wavelet transform, and the Hilbert transform. The three methods are then compared.

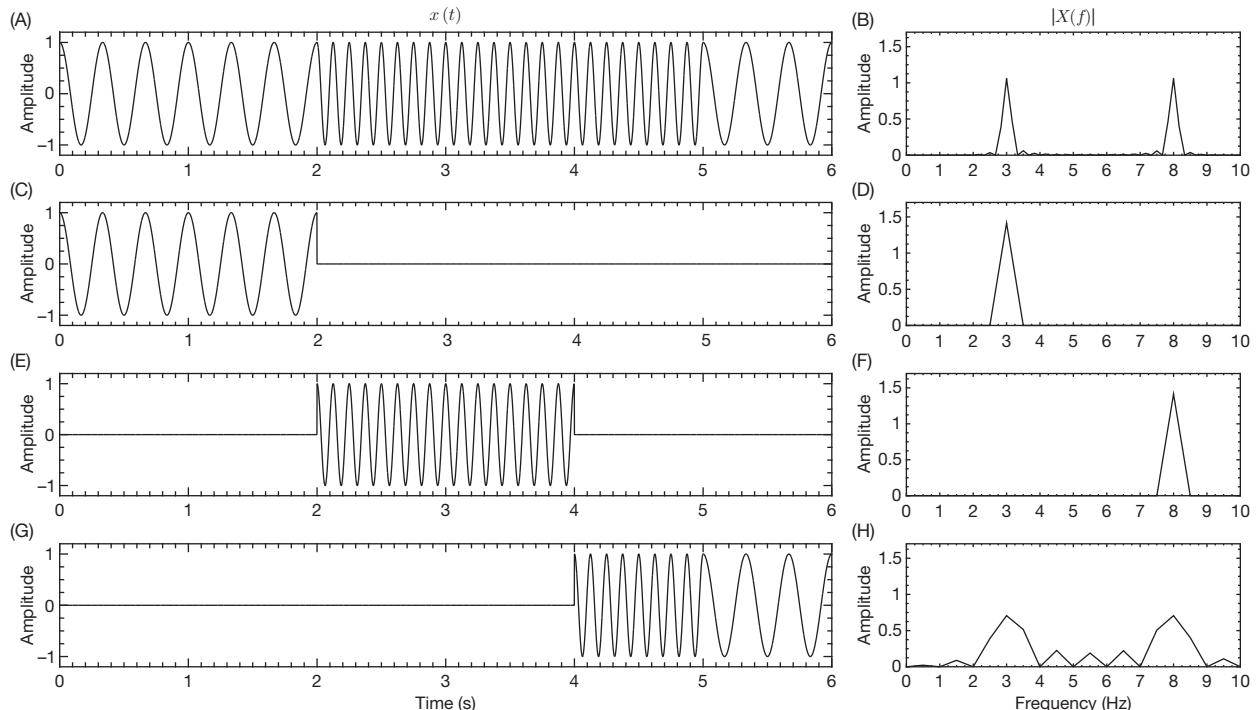


Fig. 4 Concept of spectrotemporal representation of a signal. Panel A and Panel B depict a time signal and its related amplitude spectrum respectively. Panels C, E, and G present shorter segments of the complete signal (Panel A) with their corresponding amplitude spectra Panels D, F, and H.

Short-Time Fourier Transform Approach

As presented in Fig. 4, the natural approach to obtain the spectrotemporal representation of a signal $x(t)$ is to obtain the FT for contiguous shorter segments of the signal; this is in fact the main principle in the short-time Fourier transform (STFT) approach. Then, the time-dependent spectral representation of $x(t)$ is defined as:

$$X(t, f) = \int_{-\infty}^{\infty} x(t_1 + t)w(t_1)e^{-j2\pi ft_1}dt_1 \quad (10)$$

where $w(t)$ is a window function that possesses two main purposes: segment the signal $x(t)$ into a finite duration T s, and reduce frequency leakage in the FT by reducing the sharp edges in the signal segment. While the first tentative window choice would be a rectangular window, as shown in Fig. 4, such window choice presents edge effects due to discontinuities that lead to frequency leakage. In order to reduce the frequency leakage, bell-shaped windows have been proposed, with the Hamming window being the most widely used. It is defined as:

$$w(t) = \begin{cases} 0 & \text{if } |t| > T/2 \\ 0.54 + 0.46 \cos\left(\frac{2\pi t}{T}\right) & \text{if } |t| < T/2 \end{cases} \quad (11)$$

In order to maintain the energy of the original segment of the signal tapered by the window function, the latter is commonly scaled such that its total energy is equal to unity, that is,

$$\int_{-\infty}^{\infty} |w(t)|^2 dt = 1 \quad (12)$$

The signal processing steps to obtain the spectrotemporal representation using the STFT approach are summarized in Fig. 5.

The duration of the window controls the temporal and, consequently, the frequency resolution, as per (8). For a window with duration T s, the minimum resolvable frequency is $1/T$ Hz. The effects of the duration of the window function are presented in the Fig. 6.

From Fig. 6 we can observe that the uncertainties in time and frequency vary with the window length. Temporal localization (low time uncertainty) is increased by using short windows, but this implies high uncertainty in the frequency domain (Fig. 6B). On the other extreme, high frequency localization occurs by using long windows, but the temporal uncertainty is augmented (Fig. 6D). A good trade-off in the time and frequency localizations is achieved by selecting the window length close to the duration of the stationary behavior of the signal under analysis (Fig. 6C).

When the STFT is continuous, or in the discrete case, the shift t equals to one sample, then (10) can be further simplified by applying a change to the reference of time and noting that $d(t_1 - t) = dt_1$, that is:

$$X(t, f) = \int_{-\infty}^{\infty} x(t_1)w(t_1 - t)e^{-j2\pi f(t_1 - t)}dt_1 \quad (13)$$

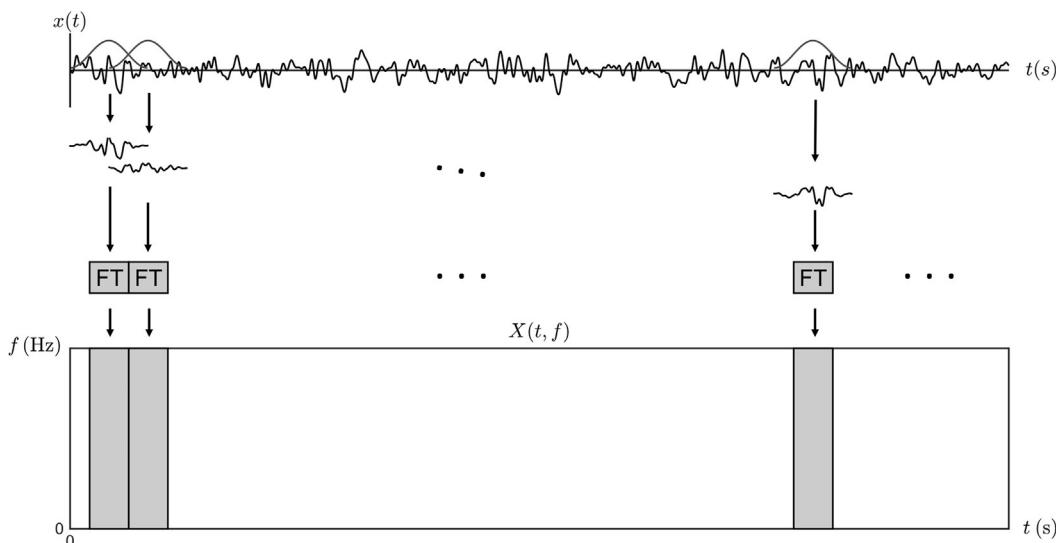


Fig. 5 Signal processing steps involved in the calculation of the spectrogram using the STFT approach; a Hamming window is depicted.

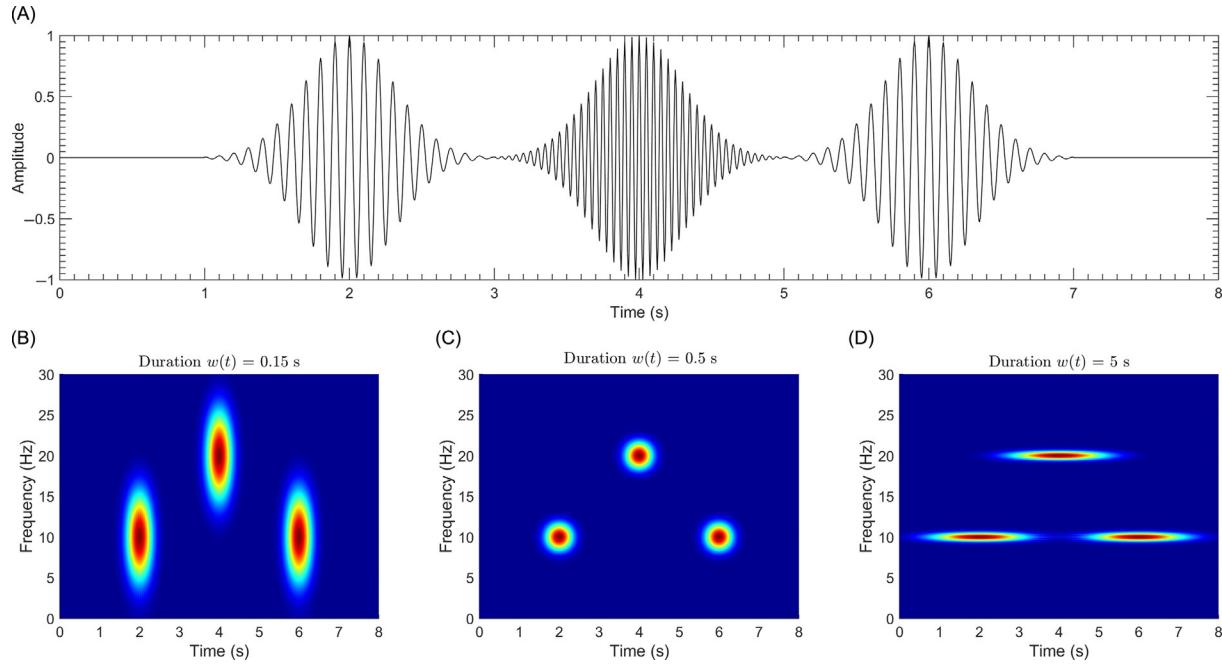


Figure 6 Effects of window duration in the calculated spectrogram for the time signal depicted in Panel A. A short window increases the time localization while the frequency localization is diminished (Panel B). Using a long window increases the frequency localization at cost of time localization (Panel D). A good time and frequency localizations can be obtained by utilizing a proper window duration (Panel C).

Given that the function $w(t)$ is time symmetric, $w(t_1 - t) = w(t - t_1)$, and changing the sign in the exponential, then (13) can be rewritten as a convolution; thus the formula for the STFT becomes:

$$X(t, f) = \int_{-\infty}^{\infty} x(t_1)w(t - t_1)e^{j2\pi f(t-t_1)}dt_1 = x(t) * w(t)e^{j2\pi ft} \quad (14)$$

In practice, the shift between consecutive windows t can be larger than one sample. In such cases, the result is a downsampled version of the spectrogram. By utilizing a window length of T s, low frequencies, close to $1/T$ Hz, will have only a few cycles to be accurately resolved, while higher frequencies will have a good number of cycles, and thus will be accurately frequency-resolved. An alternative to “equalize” this number of cycles as a function of the frequency is addressed by the use of the continuous wavelet transform, as detailed next.

Continuous Wavelet Transform Approach

The continuous wavelet transform (CWT) provides an alternative to extract time-localized information from a signal by calculating the time convolution of the signal with the analyzing wavelet. To calculate the spectrotemporal representation of a signal by using the CWT, it is necessary to perform the convolution between the signal $x(t)$ and a family (set) of wavelets $\tilde{\psi}_f(t)$ defined for a specific range of frequencies, that is:

$$X(t, f) = x(t) * \tilde{\psi}_f(t) \quad (15)$$

To extract amplitude and phase, the wavelet functions need to be complex-valued and well localized in time and frequency.

Among the wavelets that are suitable for complex spectrotemporal signal characterization, the most commonly used is the complex Morlet wavelet, which consists of a complex oscillation with a fixed frequency (frequency localization) tapered by a Gaussian window (time localization). The template or mother wavelet for the complex Morlet wavelet family is defined as:

$$\tilde{\psi}_{f_0}(t) = A(\sigma_t)e^{-t^2/2\sigma_t^2}e^{j2\pi f_0 t} \quad (16)$$

with

$$A(\sigma_t) = \frac{1}{\sqrt{\sigma_t \sqrt{\pi}}} \text{ and } \sigma_t(f_0) = \frac{n_c}{2\pi f_0} \quad (17)$$

where $e^{j2\pi f_0 t}$ is the complex oscillatory component with frequency of f_0 Hz, $e^{-t^2/2\sigma_t^2}$ is the Gaussian window with a temporal standard deviation of σ_t , $A(\sigma_t)$ is utilized to assure that the wavelet energy is equal to one (as per (12)), and the parameter n_c approximately determines the number of cycles at the frequency f_0 inside the Gaussian bell.

Substituting (16) in (15) we obtain:

$$X(tf) = x(t) * A(\sigma_t) e^{-t^2/2\sigma_t^2} e^{j2\pi ft} \quad (18)$$

Thus, the CWT approach to calculate the spectrotemporal representation of signal $x(t)$ can be summarized by Fig. 7.

As shown in (16), the parameter n_c is related to the time and frequency resolutions in the obtained spectrotemporal representation; in practice values between 4 and 6 provide a good trade-off for the temporal and spectral localizations. As such, the spectra for wavelets defined in (16) are given as:

$$\psi_{f_0}(f) = \mathcal{F}\{\psi_{f_0}(t)\} = A(\sigma_t) \sigma_t \sqrt{2\pi} e^{-f^2/2\sigma_f^2} \star \delta(f - f_0) \quad (19)$$

where the bandwidth of the spectrum $\psi_{f_0}(f)$ is given by the Gaussian component $e^{-f^2/2\sigma_f^2}$ with

$$\sigma_f = \frac{1}{2\pi\sigma_t} = f_0/n_c \quad (20)$$

In Fig. 8, the spectrotemporal representation obtained with CWT for different values of n_c is shown to illustrate the effects of this parameter.

Hilbert Transform Approach

Before expanding in the Hilbert transform approach, it is necessary to introduce a concept important for this approach, that is, that of the analytic signal. Let us start from the generalization that any real bandpass signal $u(t)$ can be rewritten as:

$$u(t) = a(t) \cos(\phi(t)) \quad (21)$$

where $a(t)$ and $\phi(t)$ are the instantaneous parameters of the signal; $a(t)$ is the instantaneous amplitude or time envelope; and $\phi(t)$ is the instantaneous phase. For a real signal $u(t)$ there is an infinite number of function pairs $[a(t), \phi(t)]$ that satisfy (21). However, with the use of the analytic signal it is possible to associate any real signal with a unique pair of functions $[a(t), \phi(t)]$. The analytic signal is a complex temporal representation of a real signal that has been useful in signal processing due its advantageous characteristics. A real signal $u(t)$ is uniquely associated to an analytic signal $\tilde{u}(t)$, if the latter fulfills the following conditions:

1. Is related to the real signal $u(t)$, as $u(t) = \operatorname{Re}\{\tilde{u}(t)\}$
2. Its spectrum is $\tilde{U}(f) = 2H(f)U(f)$, with $U(f) = \mathcal{F}\{u(t)\}$ and $H(f)$ is the Heaviside function defined as:

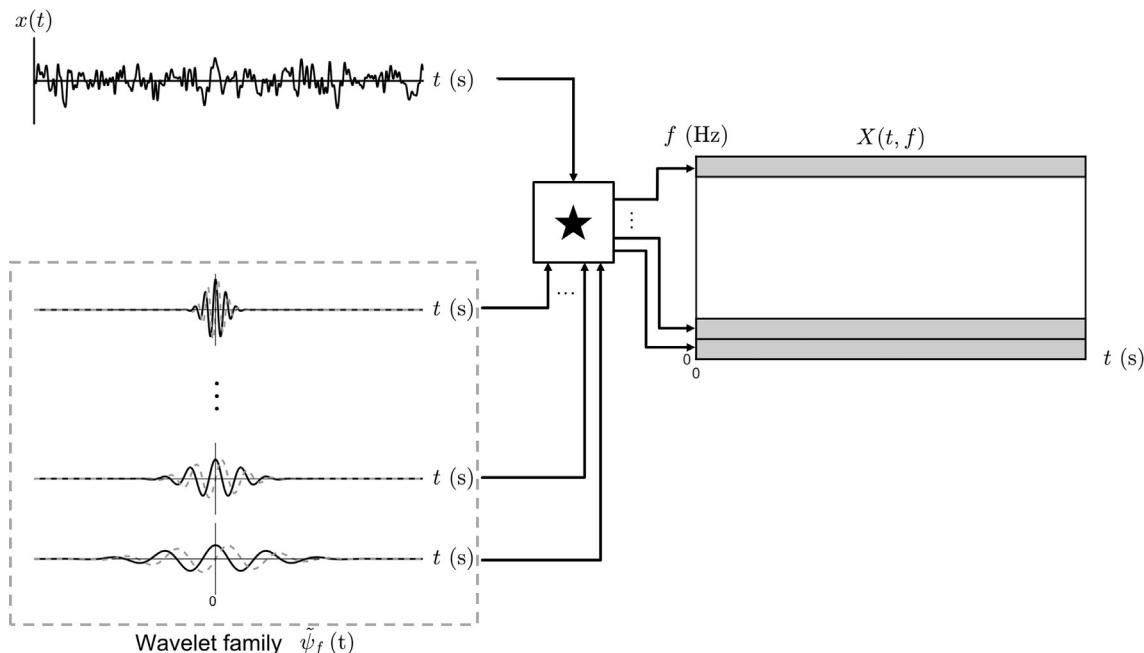


Fig. 7 Signal processing steps involved in the calculation of the spectrogram using the CWT approach; the \star symbol indicates the convolution of the time signal with each one of the complex wavelets (the black line is the real part and gray line is the imaginary part) in the wavelet family.

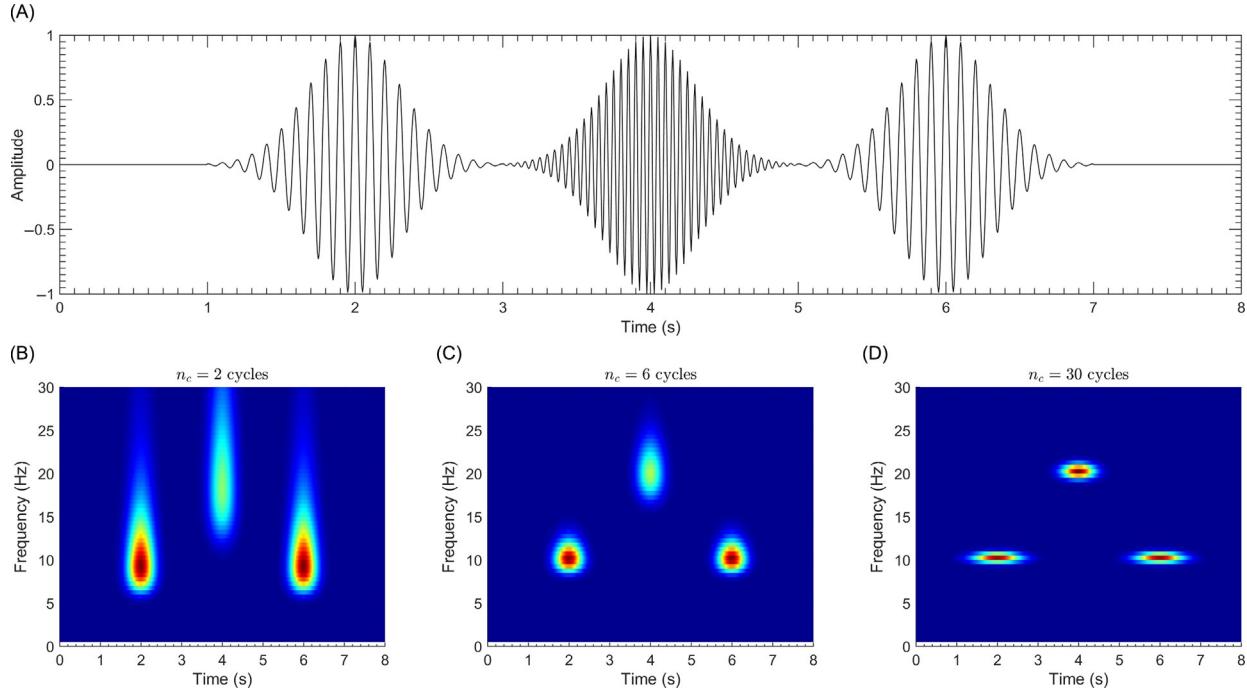


Fig. 8 Effects of the parameter n_c in the spectrotemporal representation obtained with the CWT approach. Panel A depicts a time signal. Small values of n_c increase the temporal resolution diminishing the spectral one (Panel B), while the opposite effect is observed for large values of n_c (Panel D). In practice a value of $n_c = 6$ provides good results (Panel C).

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2} & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \quad (22)$$

As the analytic signal $\tilde{u}(t)$ is a complex signal, it can be written as:

$$\tilde{u}(t) = u(t) + j\nu(t) = |\tilde{u}(t)|e^{j\arg(\tilde{u}(t))} \quad (23)$$

The characteristic that has been useful in signal processing is that the module and phase of the analytic signal correspond to the instantaneous amplitude and instantaneous phase of its associated real signal $u(t)$ presented in (21) by:

$$a(t) = |\tilde{u}(t)| = \sqrt{u(t)^2 + \nu(t)^2} \quad (24)$$

$$\phi(t) = \arg(\tilde{u}(t)) = \arctan\left(\frac{\nu(t)}{u(t)}\right) \quad (25)$$

While the analytic signal can be found for broadband signals, its usage for narrow-band limited signals provides meaningful results for the instantaneous parameters.

In order to find the analytic signal associated to $u(t)$, the imaginary part of the analytic signal is needed. Expressing the analytic signal in (23) in the frequency domain results in:

$$\tilde{U}(f) = U(f) + jV(f) \quad (26)$$

From the second condition of the analytic signal, we can determine the spectrum of $\nu(t)$ by using:

$$\tilde{U}(f) = \begin{cases} 0 & \text{for } f < 0 \text{ then : } V(f) = -jU(f) \\ U(f) & \text{for } f = 0 \text{ then : } V(0) = 0 \\ 2U(f) & \text{for } f > 0 \text{ then : } V(f) = jU(f) \end{cases} \quad (27)$$



thus we can rewrite (27) as:

$$V(f) = -j \operatorname{sgn}(f)U(f) \quad (28)$$

with $\operatorname{sgn}(f)$ being the sign function. Indeed, (28) is the frequency representation of the Hilbert transform (HT) of a signal, $\mathcal{H}\{\cdot\}$. Multiplying the spectrum $U(f)$ by $-j\operatorname{sgn}(f)$ produces a rotation by -90° in positive frequencies and a rotation of 90° in negative

frequencies. For this reason, the Hilbert transform is also frequently known as the quadrature filter. Obtaining the IFT of (28) results in

$$v(t) = \mathcal{H}\{u(t)\} = \hat{u}(t) \quad (29)$$

Then, the analytic signal for any real signal $u(t)$ can be obtained as:

$$\tilde{u}(t) = u(t) + j\mathcal{H}\{u(t)\} \quad (30)$$

or its equivalent in the frequency domain:

$$\tilde{u}(t) = \mathcal{F}^{-1}\{2H(f)U(f)\} \quad (31)$$

In summary, given a real signal $u(t)$, the Hilbert transform allows its corresponding analytic signal to be found, from which, in turn, instantaneous amplitudes and instantaneous phase can be computed in a straightforward manner.

In the Hilbert transform approach to calculate the spectrotemporal representation, a broadband signal $x(t)$ is first decomposed in band-neighboring signals by customized filter banks. The Hilbert transform is then utilized to find the analytic signal for each of the band-neighboring signals, and from their instantaneous amplitude and instantaneous phase. This process is depicted by Fig. 9.

Let us consider a narrow bandpass filter around f_0 Hz, with a real-valued impulse response $c_{f_0}(t)$, transfer function $C_{f_0}(f)$, and an associated bandwidth B_{f_0} , then, the result of filtering the signal $x(t)$ with such a filter is $x_{f_0}(t)$, given as:

$$x_{f_0}(t) = \mathcal{F}^{-1}\{X_{f_0}(f)\} = \mathcal{F}^{-1}\{X(f)C_{f_0}(f)\} \quad (32)$$

The analytic signal for $x_{f_0}(t)$ can then be obtained by substituting (32) in (31), thus resulting in:

$$\tilde{x}_{f_0}(t) = \mathcal{F}^{-1}\{2H(f)X_{f_0}(f)\} = \mathcal{F}^{-1}\{2H(f)X(f)C_{f_0}(f)\} \quad (33)$$

which can be written as a convolution operation

$$\tilde{x}_{f_0}(t) = x(t) * \mathcal{F}^{-1}\{2H(f)C_{f_0}(f)\} \quad (34)$$

Note that the point-wise product $2H(f)C_{f_0}(f)$ scales the transfer function $C_{f_0}(f)$ and only its positive frequency part is kept. This new one-side filter can be expressed as a low-pass filter $C(f)$ with bandwidth equal to B_{f_0} shifted by f_0 ,

$$\tilde{x}_{f_0}(t) = x(t) * \mathcal{F}^{-1}\left\{C_{B_{f_0}}(f - f_0)\right\} = x(t) * c_{B_{f_0}}(t)e^{2\pi f_0 t} \quad (35)$$

Note that the subindex B_{f_0} in the low-pass filter $c_{B_{f_0}}(t)$ does not indicate the central frequency of the filter, but the bandwidth $B_{f_0}(f)$ associated with it. Finally, when this process is repeated for all the filters in the filter bank, the spectrotemporal representation of $x(t)$ obtained by using the Hilbert transform approach can be expressed as:

$$X(tf) = x(t) * c_f(t)e^{j2\pi ft} \quad (36)$$

where the subindex f in the low-pass filter impulse response $c_f(t)$ indicates that the filter bandwidth is associated with a given value of f .

Spectrotemporal Computation Methods Comparison

In the previous section, three different approaches to calculate the spectrotemporal representation of a time signal were presented, namely STFT, CWT, and HT, and defined as (14), (18), and (36), respectively. From these equations, it is possible to observe that the spectrotemporal representation $X(t, f)$ is the result of the convolution between the time signal $x(t)$ and a tapered complex oscillatory function $e^{j2\pi ft}$. Such an algorithm can be written in a more general fashion as:

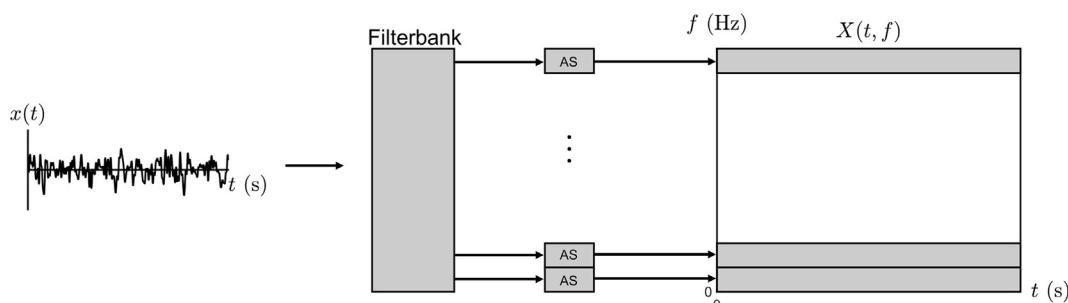


Fig. 9 Signal processing steps involved in the calculation of the spectrogram using the HT approach. The block AS uses the Hilbert transform to calculate the corresponding analytic signal.

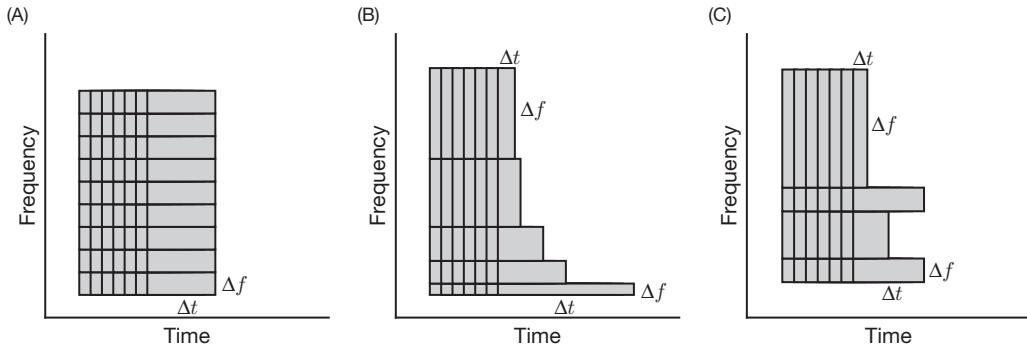


Fig. 10 Time-Frequency uncertainty for STFT, CWT, and HT approaches, Panels A, B, and C, respectively.

$$X(tf) = x(t) \star \lambda_f(t) e^{j2\pi ft} \quad (37)$$

where $\lambda_f(t)$ is a real-valued time window function (low-pass filter), which can be different for different values of f . As a consequence, the product $\lambda_f(t)e^{j2\pi ft}$ can be seen as a filter bank. The main difference among the STFT, CWT, and HT approaches lies in how the window $\lambda_f(t)$ is defined, and as consequence how the time and frequency uncertainties are handled.

The STFT approach can be obtained by making $\lambda_f(t) = w(t)$ in (37), where $w(t)$ is a window function with constant duration T s for all the values of f . This is reflected as a time resolution Δt and consequently a constant frequency resolution Δf in the time-frequency space (see Fig. 10A). The uniformity of the time-frequency bins presents a drawback for low frequency components, that is, those close to 1/T Hz, as they are not accurately represented since the window duration comprehends just a few cycles at those frequencies, thus resulting in frequency leakage. The division of the time-frequency space is optimized when the window duration is equal to the duration of the stationary segments of the signal. The CWT approach, in turn, uses an infinite time Gaussian window with a standard deviation inversely proportional to f , as per (17). Then, according to (18), $\lambda_f(t) = A(\sigma_t) e^{-t^2/2\sigma_t^2}$. Despite the fact that the duration of the Gaussian bell is infinite, it is possible to determine a nominal or effective duration Δt , which corresponds to the duration of a rectangular pulse with amplitude where the area of such a rectangle is equal to the area of the original signal, that is:

$$\Delta t = \frac{\int_{-\infty}^{\infty} A(\sigma_t) e^{-t^2/2\sigma_t^2} dt}{A(\sigma_t)} = \sigma_t \sqrt{2\pi} \quad (38)$$

As such, for the CWT approach, the minimum resolvable frequency Δf is equal to $1/\Delta t$ (Fig. 10B). Lastly, in the HT approach, $\lambda_f(t) = c_f(t)$, where $c_f(t)$ is a low-pass filter which is defined according to the application. If the window $c_f(t)$ is constant for different values of f , the sectioning of the time-space space is similar to that obtained with the STFT. On the other hand, if the duration of $c_f(t)$ is inversely proportional to f , then the time-frequency space sectioning will be similar to the one obtained with the CWT approach. However with the use of the HT approach, the duration of the window for each value of f is defined in an application-wise manner (Fig. 10C). Given the correct parameters, it is possible to obtain similar results with the three approaches. Overall, despite STFT, CWT, and HT approaches handle the uncertainty principle differently, neither is capable of avoiding the trade-off between time and frequency uncertainties.

The power spectrogram, which is defined as $|X(t, f)|^2$, can be utilized to compute the power of the signal for a specific time interval (τ) in the range of frequencies (v) as:

$$\text{power}_{\tau v} = \int_{\tau} \int_v |X(t, f)|^2 dt df \quad (39)$$

However, this only holds if filters in filter bank $\lambda_f(t)e^{j2\pi ft}$ are orthogonal, as in the STFT approach (Fig. 11A) and the HT approach, if the filters are designed with such properties. For the CWT, when the central frequency for the kernels is linearly spaced, Fig. 11B, the filter bank $\lambda_f(t)e^{j2\pi ft}$ is not orthogonal; this is appreciated as spectral overlap or correlation between the different filters in the filter bank, which has as consequence that the time series related to each wavelet kernel are highly correlated; therefore redundant information is present. The filter bank associated to the CWT can be orthogonal if the central frequencies for the kernels are logarithmically spaced, Fig. 11C. In the analysis of biomedical signals, this orthogonality requirement can be loosened and relevant information can be derived directly from the spectrotemporal representation. While the spectrotemporal representation $|X(f, t)|$ is a helpful tool to identify time-occurring spectral changes in nonstationary signals, it does not provide information on the nature of these changes. As such, in the next section, the modulation spectrogram will be introduced as tool to further analyze the spectrotemporal representation of a signal.

In order to illustrate the similarities and difference in the approaches to calculate the spectrotemporal representation of a signal, the power spectrogram is calculated for a 10-s segment of EEG signal acquired in the electrode Cz during closed-eyes resting-state condition (Fig. 12A) by using the STFT (Fig. 12B), CWT (Fig. 12C), and HT (Fig. 12D) approaches.

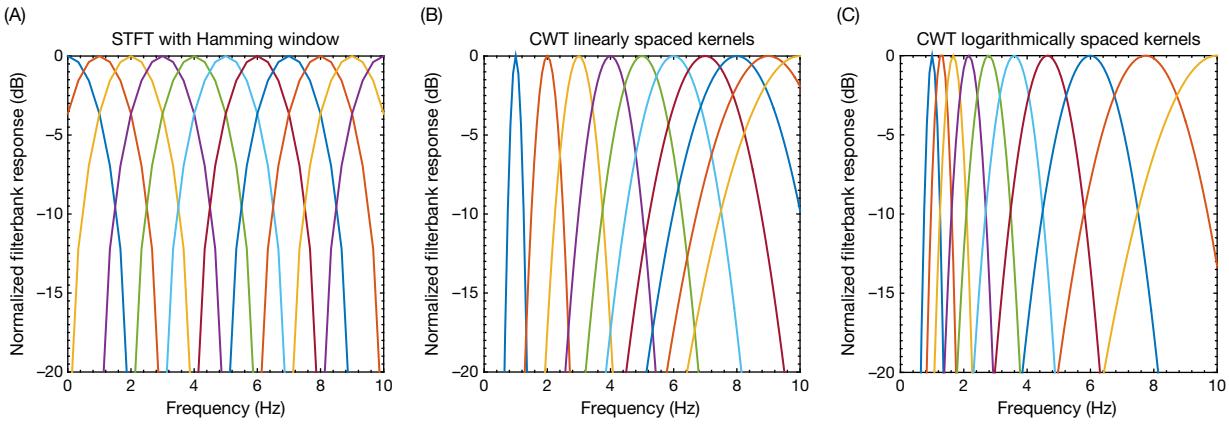


Fig. 11 Normalized filter response for the filter bank $\lambda_f(t)e^{j2\pi f_0 t}$ for STFT with Hamming window (Panel A), for CWT kernels with linearly spaced central frequencies (Panel B), and for CWT kernels with logarithmically spaced central frequencies (Panel C).

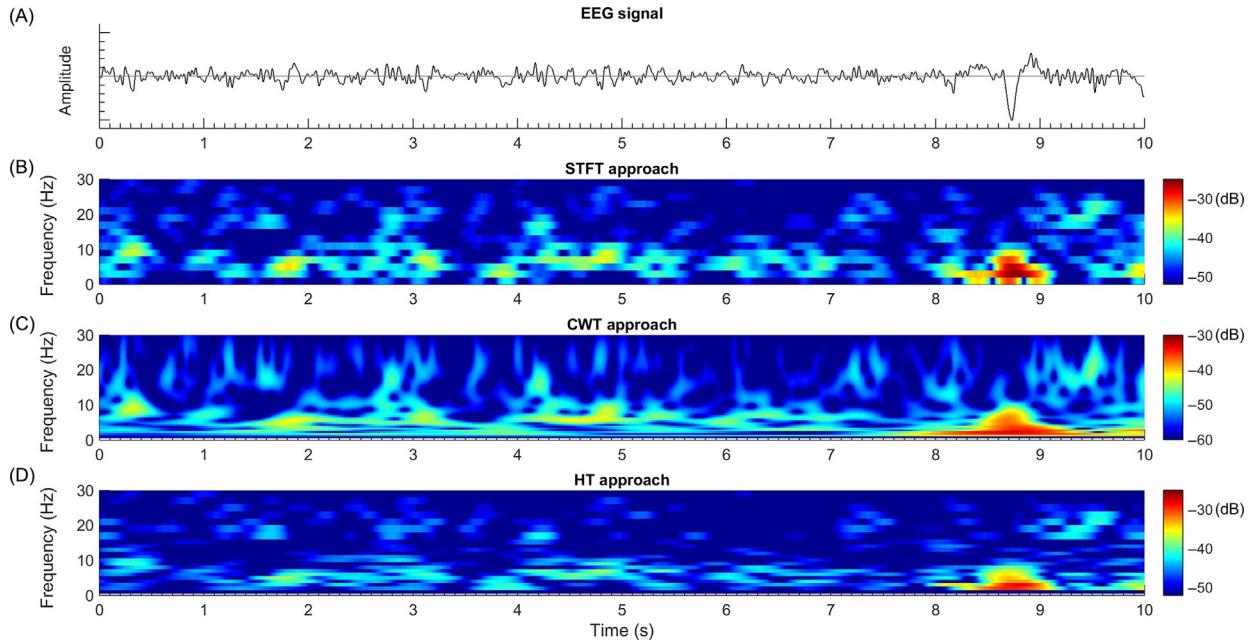


Fig. 12 EEG signal at electrode Cz during eyes-closed resting-state condition (Panel A). Panel B shows the power spectrogram obtained with the short-time Fourier transform approach using a window duration of 0.5 s. Panel C presents the power spectrogram calculated using the continuous wavelet transform, with a parameter $n_c = 6$. Panel D shows the power spectrogram obtained with the Hilbert transform approach with a filter bank comprised of 23 nonoverlapping filters, of which 15 are 1-Hz-bandwidth filters in the range 0 to 15 Hz, and 8 are 2-Hz-bandwidth filters in the range 15–31 Hz.

Modulation Spectrogram

The spectrotemporal representation of a signal $x(t)$ presents information on how amplitude and phase change over time for different frequency components. A natural extension of that representation is the modulation spectrogram, which characterizes those amplitude changes in the framework of amplitude modulation analysis, providing a representation of second-order periodicities, that is, modulation frequencies. These modulation frequency components are sometimes referred to as “hidden periodicities” as they are not present as spectral components in the “conventional” frequency representation of the $x(t)$ signal.

In the framework of amplitude modulation analysis, a signal $x(t)$ can be expressed as the product of a low-frequency modulating signal $m(t)$, and a high-frequency or carrier signal $c(t)$:

$$x(t) = m(t)c(t) \quad (40)$$

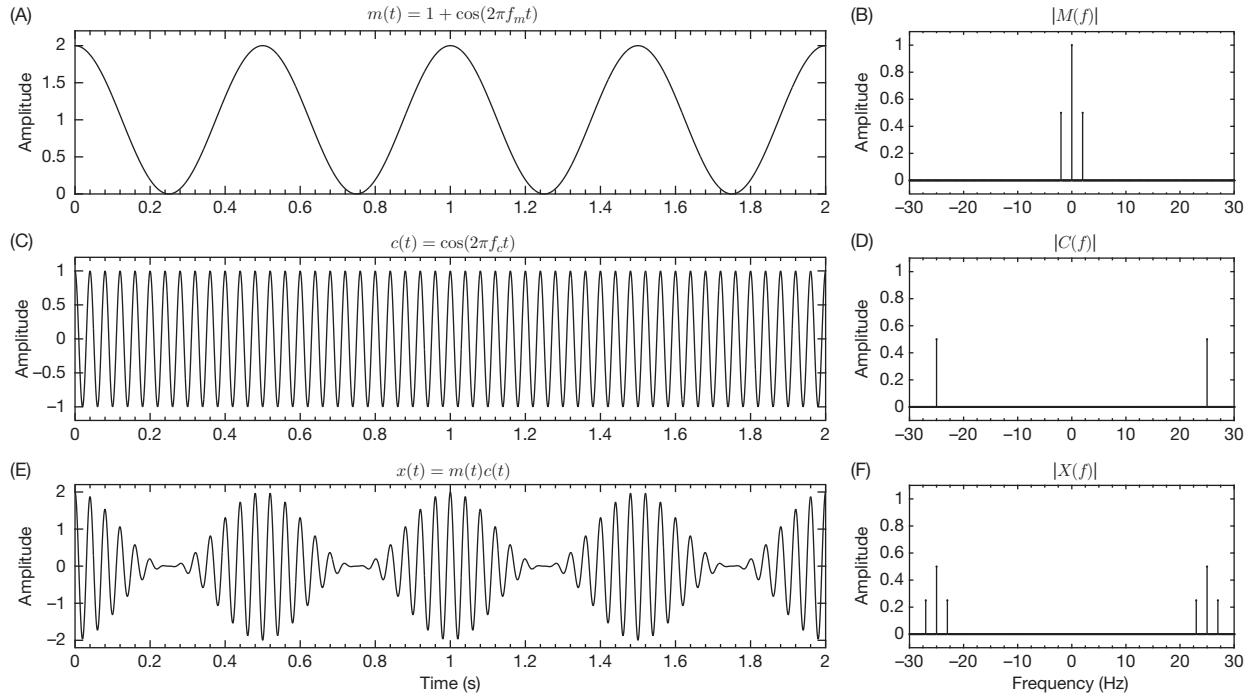


Figure 13 Temporal representation for the modulating signal $m(t)$ (Panel A), the carrier signal $c(t)$ (Panel C), and the resulting signal $x(t)$ (Panel E). Panels D, B, and F show the spectral representations for the signals $m(t)$, $c(t)$, and $x(t)$, respectively.

with the following two assumptions: (1) there is no spectral overlap between $m(t)$ and $c(t)$, and (2) $m(t)$ is a real nonnegative signal. When $m(t)$ takes negative values, the amplitude modulation process becomes overmodulated, that is, possess a modulation index larger than 1. In this scenario, the modulating signal $m(t)$ cannot be recovered by analyzing only the envelope of the carrier signal $c(t)$.

Let us use the following numerical example to introduce the concept of the modulation frequency domain. Consider a time signal $x(t)$ defined as:

$$x(t) = (1 + \cos(2\pi f_m t))(\cos(2\pi f_c t)) \quad (41)$$

with $f_m = 2$ Hz and $f_c = 25$ Hz. Comparing (41) with the AM model presented in (40), it can be seen that the modulating signal is $\cos(2\pi f_m t)$ (an offset of amplitude 1 is added to ensure that the modulating signal is nonnegative), and the carrier signal is $\cos(2\pi f_c t)$. The time and frequency representations for $x(t)$, $m(t)$, and $c(t)$ are depicted by Fig. 13. The spectrum for $x(t)$ can be found by using the product-of-cosines trigonometric identity where:

$$\cos(\alpha)\cos(\beta) = \frac{1}{2}[\cos(\alpha + \beta) + \cos(\alpha - \beta)] \quad (42)$$

then,

$$\begin{aligned} x(t) &= \cos(2\pi f_c t) + \cos(2\pi f_m t) \cos(2\pi f_c t), \\ x(t) &= \cos(2\pi f_c t) + \frac{1}{2}[\cos(2\pi(f_c + f_m)t) + (\cos(2\pi(f_c - f_m)t))] \text{ and} \\ X(f) &= \frac{1}{2}[\delta(f - f_c) + \delta(f + f_c)] + \frac{1}{4}[\delta(f - f_c - f_m) \\ &\quad + \delta(f - f_c + f_m) + \delta(f + f_c - f_m) + \delta(f + f_c + f_m)] \\ X(f) &= \frac{1}{2}[\delta(f - 25) + \delta(f + 25)] + \frac{1}{4}[\delta(f - 27) + \delta(f - 23) + \delta(f + 23) + \delta(f + 27)] \end{aligned}$$

From Fig. 13A, it is possible to infer that the modulation spectrum is the frequency representation of the modulating signal $m(t)$, and it can be obtained with the Fourier transform. As such, the modulation spectrum $M(f)$ is an indicator of the second-order periodicities present in the signal $x(t)$. Unlike the carrier signal spectral components of $C(f)$, the spectral components of modulating signal $M(f)$ are not represented in the spectrum $X(f)$, but can be processed by analyzing the changes in amplitude for the first periodicity (see Fig. 14).

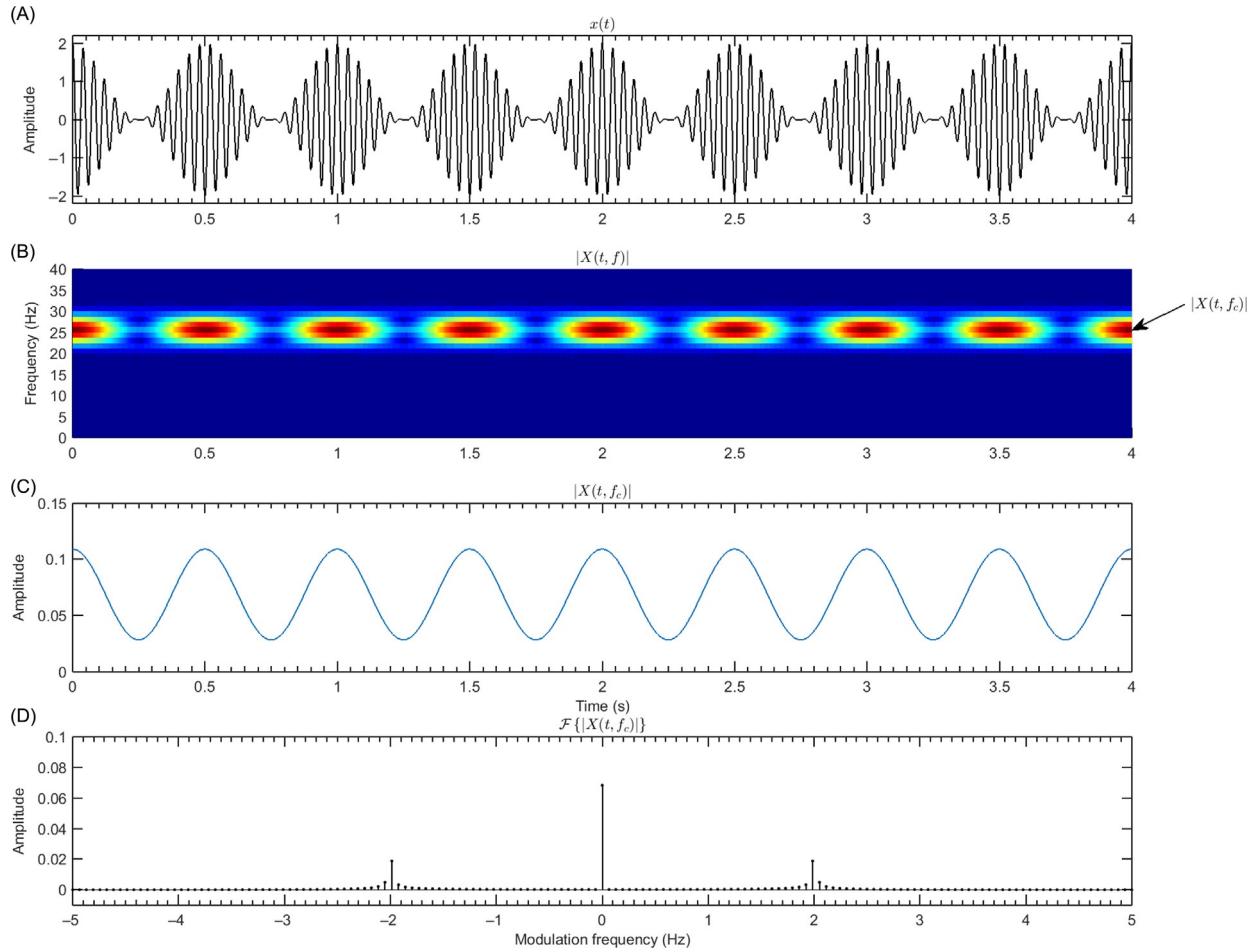


Fig. 14 Temporal representation for signal $x(t)$ (Panel A); this signal is the result of the amplitude modulation process presented in 13; corresponding amplitude spectrogram $|X(t, f)|$ is depicted in Panel B, where the row corresponding to the amplitude temporal changes in the frequency of the carrier signal f_c is indicated. The time series for the amplitude changes for $f = f_c$ is presented in Panel C. Panel D shows the spectral representation of the time signal in Panel C.

Note that according to (40) the modulating and carrier signals are not necessarily monochromatic waves, as in the example signal described in (41). Nevertheless, the instantaneous amplitude and phase need to be calculated in narrow-band signals to possess a straightforward physical meaning. In the spectrogram $X(t, f)$, for a given frequency component f_0 , its amplitude changes over time are given as $|X(t, f_0)|$; these changes can be studied in the framework of amplitude modulation. This is to say, the spectrum for the amplitude time series of each frequency component in $|X(t, f)|$ is calculated; thus, the modulation spectrogram is a 2D representation of the conventional frequencies components versus their modulation spectra, that is:

$$X(f, f_{mod}) = \mathcal{F}_t\{|X(t, f)|\} \quad (43)$$

The signal processing steps required to obtain the modulation spectrogram are depicted in Fig. 15. In essence, the modulation spectrogram consists of a double transformation: given a time signal $x(t)$, its corresponding spectrotemporal representation $X(t, f)$ is computed, then a time-frequency transformation is performed over the amplitude time series for all the frequency components of $|X(t, f)|$, hence obtaining the modulation spectrogram $X(f, f_{mod})$.

The amplitude modulation analysis differs from the one typically proposed in so called “cross-frequency coupling” analyses, as these latter ones typically measure the correlation between power-power, power-phase, or phase-phase between two frequency components of a signal. Therefore, cross-frequency coupling methods cannot be utilized to study modulation as the AM modulating frequencies do not appear as elements in the frequency axis in $|X(t, f)|$.

The modulation spectrogram is a tool that provides the possibility of studying nonlinear interactions in biological systems, such as amplitude modulation which may have its origin in the interaction among these systems. To facilitate this analysis, a Python-Octave-MATLAB toolbox has been developed and made available open-source at: <https://github.com/MuSAELab/amplitude-modulation-analysis-toolbox>.

The amplitude modulation analysis toolbox is a collection of functions for the calculation and visualization of the spectrum, spectrogram, and modulation spectrogram representations for time signals. These functions are compatible with common scientific

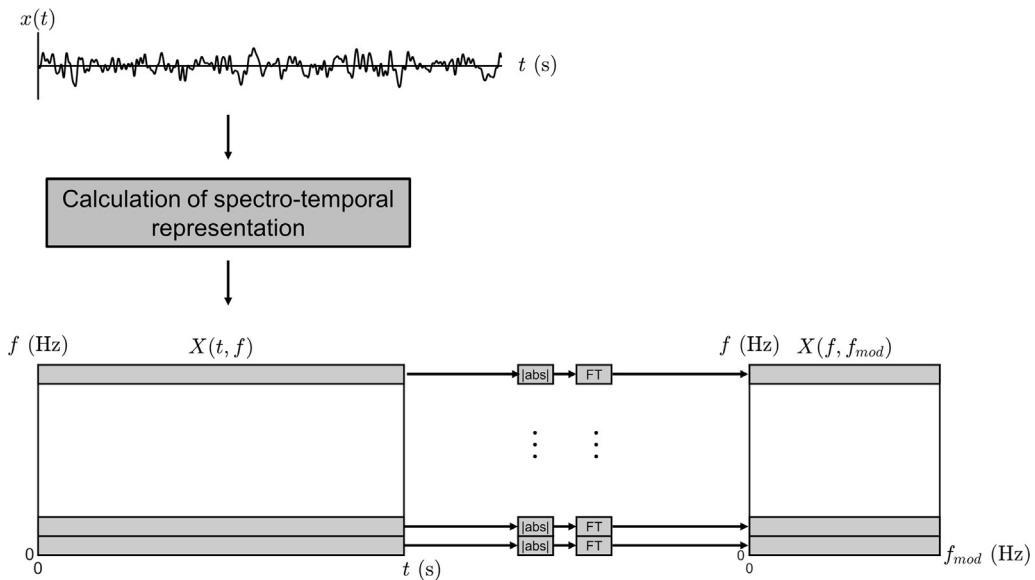


Fig. 15 Signal processing steps involved in the calculation of the modulation spectrogram from the amplitude spectrogram of signal. The block `abs` indicates the absolute value, and the `FT` indicates the use of the Fourier transform.

programming languages such as MATLAB, Octave and, Python3 and Python2. Besides the implementation of the spectrotemporal methods, a graphical user interface is provided to facilitate exploratory amplitude modulation analysis by allowing online changes in the parameters for its calculation (see Fig. 16). Moreover, example data and scripts are provided to the interested reader. Next, we review biomedical applications that have relied on modulation spectral analysis and showcase some of the functionalities of the amplitude modulation analysis toolbox.

Biomedical Applications

As mentioned previously, the modulation spectrogram is a 2D representation of a time signal in the conventional frequency versus modulation frequency domain, that is, it portrays the over-time changes in amplitude for the spectral components in the time signal in the modulation frequency domain. This representation has proven very useful in processing nonstationary signals such as speech for diverse tasks: quality assessment in reverberant speech, speaker identification, and emotion recognition, among others. In the case of biomedical signals, the modulation spectrogram is of great importance as it is a helpful tool to assess the interaction processes between biomedical signals and the mechanisms that give origin to them.

Applications relying on the use of the modulation spectrogram can be divided in two main groups, namely analysis and synthesis. In analysis applications, the modulation spectrogram is used to extract novel features to characterize and/or classify signals, thus providing a deeper understanding of the nature of the signals. In synthesis applications, on the other hand, signal processing, such as filtering, is performed in the frequency-modulation-frequency domain, and then the processed signal is transformed back to the time domain. In the following sections, example applications using the modulation spectrogram for analysis and synthesis of biomedical signals are presented.

Analysis Applications

If we are interested in solely extracting innovative features and/or insights from biomedical signals, analysis-only processing of the modulation spectrogram can be performed. Such features/insights can be used for example, assessment, classification, and/or diagnosis. Below are four representative examples of analysis-only applications.

Detection of respiration frequency from the ECG signal

Respiration influences the amplitude of the ECG signal by diverse mechanisms. By analyzing the periodic changes in the amplitude of the R-peak in the ECG signal, it is possible to obtain information about the respiration rate. Fig. 17A depicts a raw ECG signal acquired when the subject was breathing at a fixed rate of 15 breathings per minute. This respiration rate is equal to 0.25 Hz, and can be seen in the modulation spectrogram of the ECG signal (Fig. 17B). The respiration rates obtained with this method present a percentage of error of 17% for noncontrolled, and 10% for a fixed respiration rate of 15 breathings per minute, compared with the ground-of-truth respiration rate derived from measurements with a chest-located strain gauge.

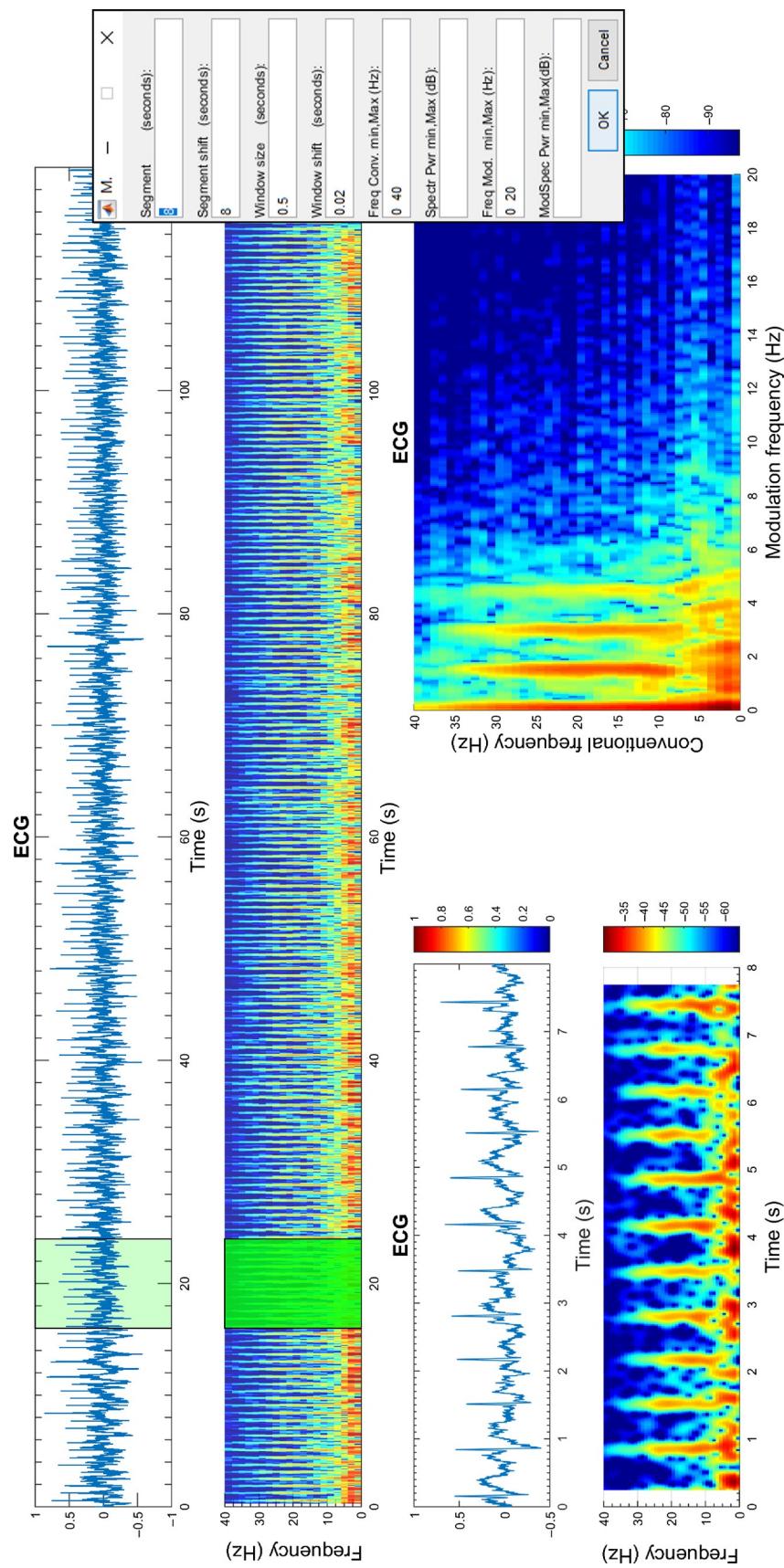


Fig. 16 Graphical user interface for the amplitude modulation analysis toolbox.

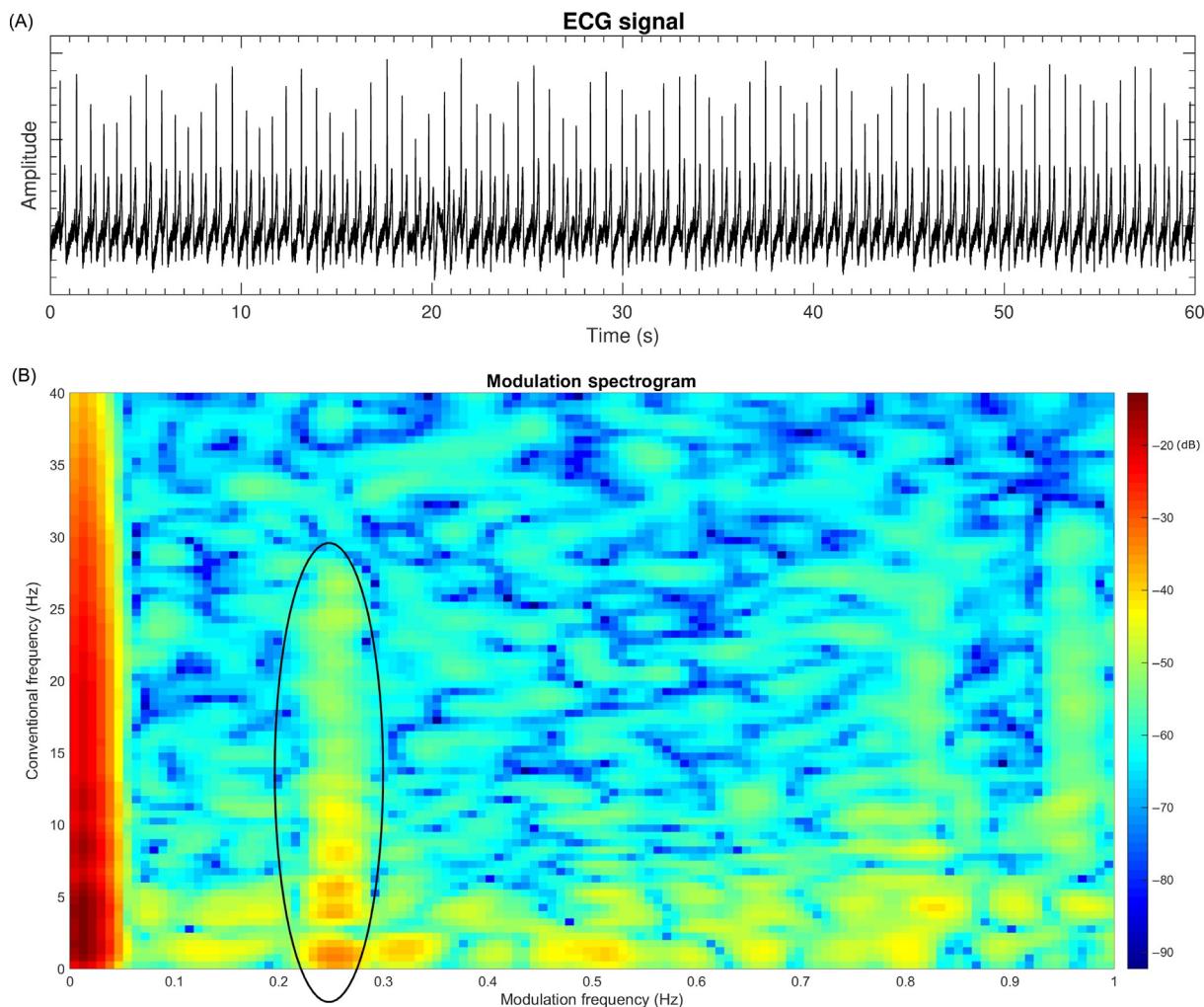


Fig. 17 Segment of raw ECG signal (Panel A) with its corresponding modulation spectrogram (Panel B), where the components related to the respiration rate, 0.25 Hz, are indicated.

Quality assessment for ECG

As the ECG is a (quasi-)periodic signal, its spectral components present (quasi-)periodic changes (related to the heart rate) that are different from the ones generated by conventional ECG artifacts. As a result, by analyzing the ECG signal with the modulation spectrogram representation, the quality of the ECG can be assessed. As an example, Fig. 18A presents the time series for a clean ECG signal and Fig. 18C a segment of ECG artifactual signal. Fig. 18B and D depict the modulation spectrograms for clean and artifactual ECG, respectively. As can be seen, measuring the energy ratio between the modulation-spectrogram components related to ECG and components related to artifacts, it is possible to obtain the so-called modulation spectrum-based quality index (MS-QI). The MS-QI metric has been shown to present higher correlation with the ECG SNR (signal-to-noise-ratio) relative to traditional ECG-quality metrics such as kurtosis and the QRS complex “in-band to out-of-band” spectral power ratio on synthesized, healthy real-world and pathological real-world ECG recordings. For example, correlation values of 0.94 were obtained with MS-QI, thus comparing favorably against the benchmarks, which achieved correlations of 0.81 and 0.71, respectively.

Neuronal amplitude modulations

The modulation analysis of spontaneous neuronal oscillations has been suggested as a promising approach for the characterization and assessment of different neurological disorders. For the study of Alzheimer’s disease (AD), for example, a neuromodulatory abnormality has been reported with potential origin in the impaired cerebral blood flow caused by plaques and tangles. Since blood flow plays an important role in the regulation and modulation of neural activity, it has been shown that modulation spectrogram analysis can be useful for AD diagnosis. To illustrate this, the average power modulation spectrogram for healthy controls and mild Alzheimer’s disease patients is presented in Fig. 19A and B, respectively. The performance of modulation-based features for the

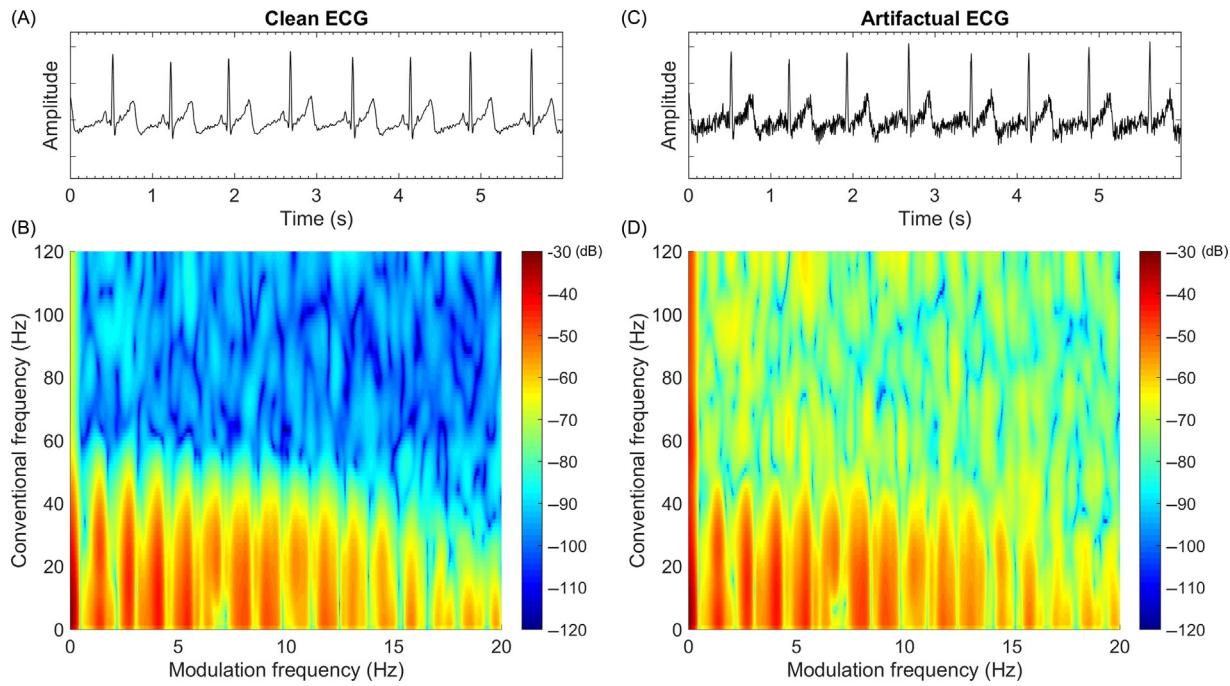


Fig. 18 Clean (Panel A) and artifactual ECG (Panel C) signals. The modulation spectrogram for the clean and artifactual ECG is depicted in panels B and D, respectively.

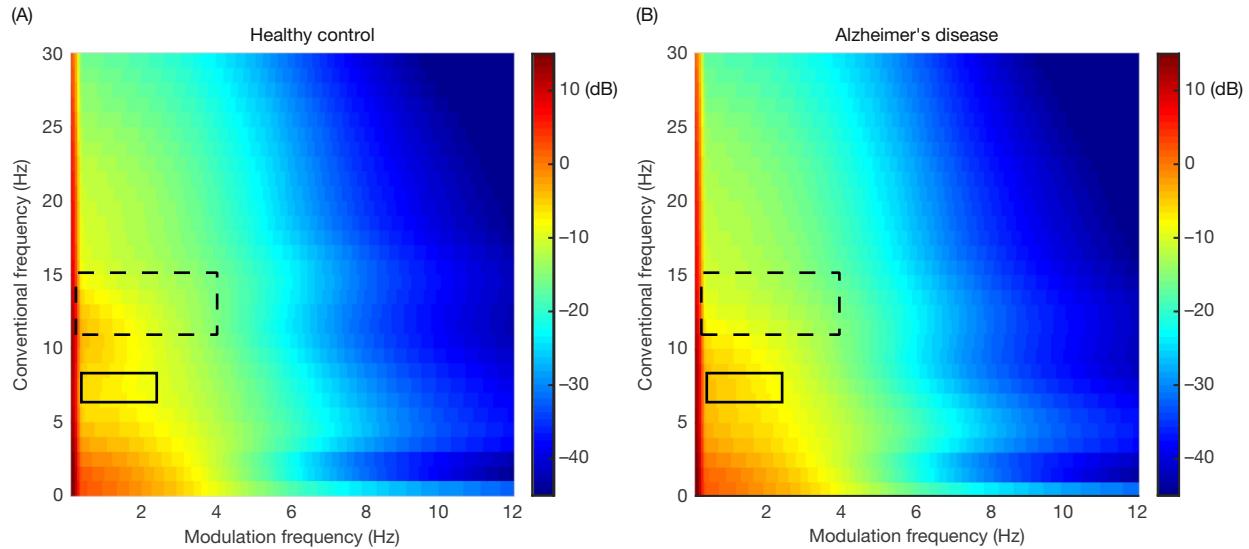


Fig. 19 Average power modulation spectrogram for healthy controls (Panel A) and mild Alzheimer's disease patients (Panel B). Main differences are indicated in the black boxes in both Panels.

classification of healthy controls versus mild/moderate AD (90.6% accuracy) surpasses the performance obtained with traditional features such as spectral-peak per EEG band (81.3% accuracy). The modulation features, in fact, were shown to provide complementary information to traditional features; thus they could be fused to obtain further gains.

Synthesis Applications

Often times, biomedical applications rely on signals that are measured in a noisy manner or are corrupted by cooccurring phenomena. These competing signals overlap in time and frequency but, given their varying nature, become separable in the modulation spectral domain. As such, processing of the modulation spectrogram (e.g., filtering) can allow for more accurate

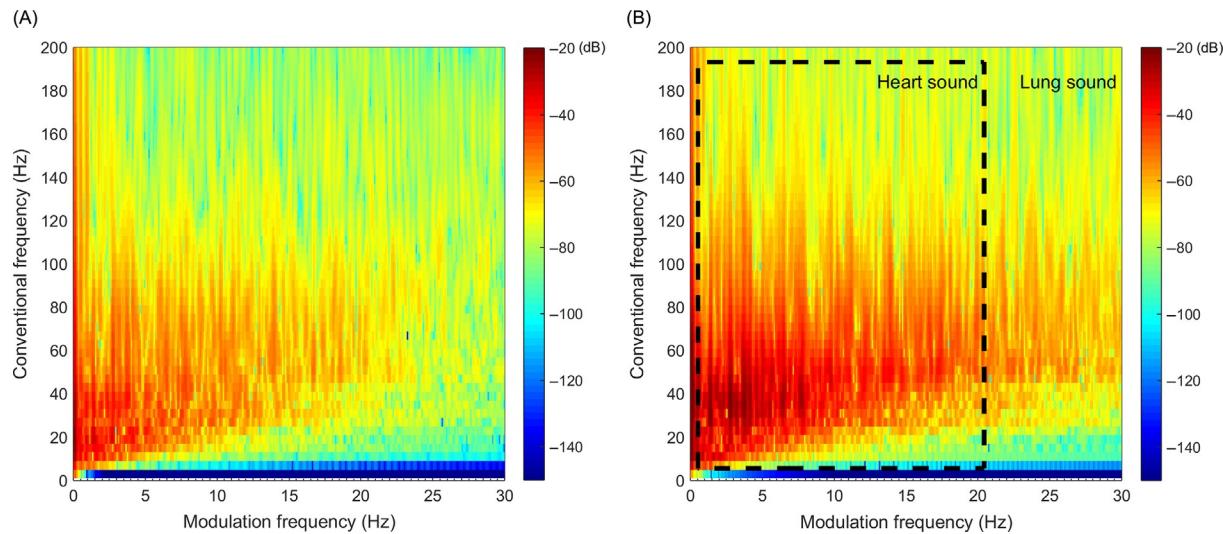


Fig. 20 Modulation spectrograms for lung sounds (Panel A) and lung sounds corrupted by heart sounds. (Panel B) The modulation frequency range where the heart sounds are more noticeable is indicated in the dashed box.

enhancement and/or source separation of competing sources. In this case, the two transformations involved in the computation of the modulation spectrogram need to be invertible, thus allowing for a time-domain digital to be reconstructed, or “synthesized.” With synthesis-only applications, processing is typically done within the amplitude spectrogram and the phase of the original mixed signal is used for reconstruction. Below are two representative examples of synthesis applications.

Separation of heart and lung sounds from digital auscultation

Auscultation of lung sounds is a useful procedure for detection of pulmonary diseases. Unfortunately, lung sounds are frequently corrupted by heart sounds, as both overlap in the time and frequency domains. However, due to the different origin of the two signals, their spectral content changes with different modulation frequencies; thus separability can be attained by filtering in the modulation spectrogram domain. Fig. 20A and B depict the modulation spectrogram for lung sounds, and lung sounds corrupted by heart sounds. The modulation-based analysis and synthesis of the auscultation sound signal to remove the heart sound provides an improvement in the separation of the lung and heart sounds compared with other utilized methods such as adaptive filtering, wavelet filtering, and independent component analysis, among others. A metric to compare the performance of the aforementioned separation methods is to measure the log-spectral distance (LSD) between the recovered lung sound and heart-sounds-free breathing sound. Comparing the modulation approach with wavelet filtering, the former achieves almost half the LSD value of wavelet filtering with almost 30 times lower signal processing time.

ECG enhancement

As shown previously, the ECG signal and its potential artifacts become separable in the modulation spectral domain. As such, filtering in such domain should allow for enhancement of the ECG signal. To illustrate this, Fig. 21A presents a segment of artifactual ECG with its modulation spectrogram (Fig. 21B). The ECG signal reconstructed after filtering in the modulation domain is depicted in Fig. 21C with its corresponding modulation spectrogram in Fig. 21D. For very noisy ECGs ($\text{SNR} = -10 \text{ dB}$), estimating heart rate becomes very challenging and errors of around 57% are seen. With the modulation-based enhancement, this error is reduced to 2% and with wavelet-based enhancement to 6%. The ECG signal is modulation-enhanced to a post-SNR of almost 1.5 dB (from -10 dB), whereas the wavelet filtering method achieves a post-SNR of -2.5 dB .

Conclusion

Biomedical signals are known to be nonstationary; thus conventional frequency-domain analysis techniques have their limitations. In this article, we have presented the theory behind spectrotemporal amplitude modulation analysis and compared three popular methods of computing the so-called modulation spectral representation, namely via short-time Fourier, continuous wavelet, and Hilbert transforms. Computation of the modulation spectrogram is discussed and the open-source Amplitude Modulation Toolbox is presented. Lastly, as can be seen, the modulation spectrogram can be a useful tool for biomedical signal analysis allowing us to extract new insights for diagnostics, as well as enhance noisy signals in conditions where conventional time- and frequency-domain analyses are not suitable.

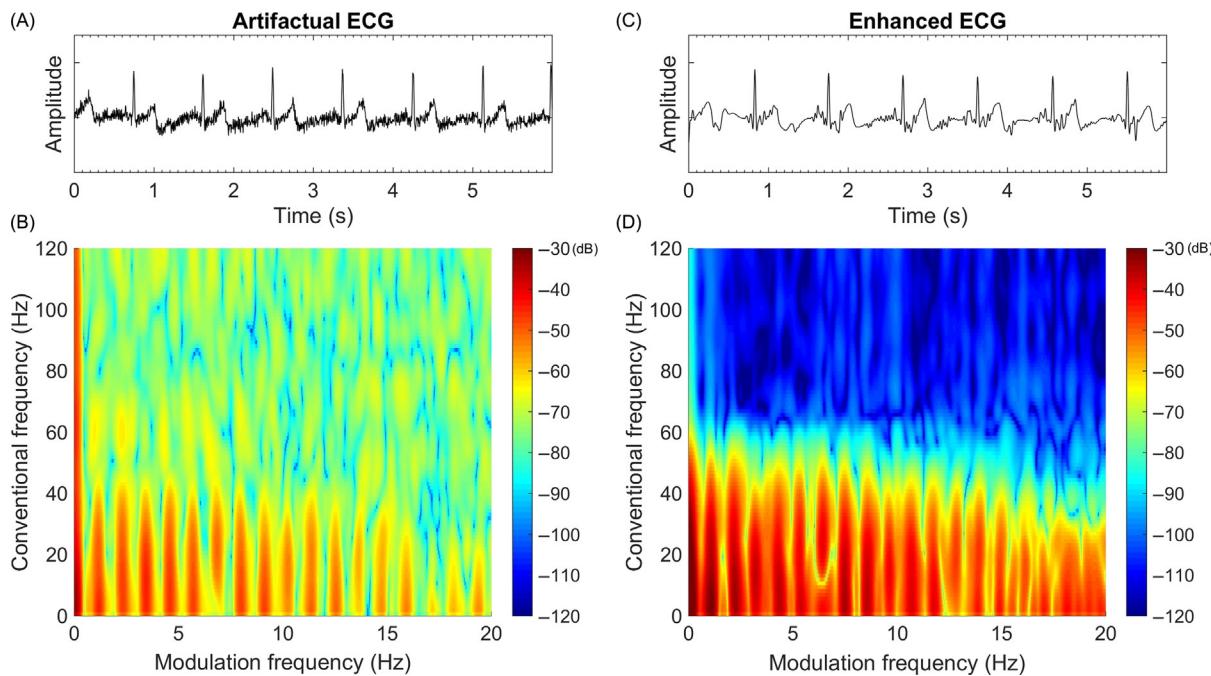


Fig. 21 Panels A and B show a segment of artifactual ECG signal and its modulation spectrogram. The reconstructed ECG signal obtained after being filtered in the modulation domain is depicted in Panel C with its corresponding modulation spectrogram in Panel D.

Further Reading

- Akay M (2012) *Biomedical signal processing*. Academic Press.
- Atlas L and Shamma SA (2003) Joint acoustic and modulation frequency. *EURASIP Journal on Advances in Signal Processing* 2003: 310290.
- Avila AR, Fraga FJ, Sarria-Paja M, and Falk TH (2014) In: *Investigating the use of modulation spectral features within an i-vector framework for far-field automatic speaker verification*. Telecommunications Symposium (ITS), 2014 International, pp. 1–5. IEEE.
- Bondar' AT and Fedotchev AI (2000) Concerning the amplitude modulation of the human EEG. *Human Physiology* 26: 393–399. <https://doi.org/10.1007/BF02760265>.
- Bruno A (2004) Fourier-, Hilbert- and wavelet-based signal analysis: Are they really different approaches? *Journal of Neuroscience Methods* 137: 321–332. <https://doi.org/10.1016/j.jneumeth.2004.03.002>.
- Cohen L (1995) *Time-frequency analysis: Theory and applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
- Cohen MX (2014) *Analyzing neural time series data: Theory and practice*.
- Falk TH and Chan W-Y (2010) Modulation spectral features for robust far-field speaker identification. *IEEE Transactions on Audio, Speech, and Language Processing* 18(1): 90–100. <https://doi.org/10.1109/TASL.2009.2023679>.
- Falk TH, Zheng C, and Chan W-Y (2010) A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech. *IEEE Transactions on Audio, Speech, and Language Processing* 18(7): 1766–1774. <https://doi.org/10.1109/TASL.2010.2052247>.
- Falk TH, Fraga FJ, Trambaioli L, and Anghinah R (2012) EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease. *EURASIP Journal on Advances in Signal Processing* 2012: 1–9.
- Fraga FJ, Falk TH, Kanda PAM, and Anghinah R (2013) Characterizing Alzheimer's disease severity via resting-awake EEG amplitude modulation analysis. *PLoS ONE* 8: e72240. <https://doi.org/10.1371/journal.pone.0072240>.
- Gardner WA (1991) Exploitation of spectral redundancy in cyclostationary signals. *IEEE Signal Processing Magazine* 8: 14–36. <https://doi.org/10.1109/79.81007>.
- Le Van Quyen M, Foucher J, Lachaux J, et al. (2001) Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *Journal of Neuroscience Methods* 111: 83–98. [https://doi.org/10.1016/S0165-0270\(01\)00372-7](https://doi.org/10.1016/S0165-0270(01)00372-7).
- Moody GB, Mark RG, Zuccola A, and Mantero S (1985) Derivation of respiratory signals from multi-lead ECGs. *Computers in Cardiology* 12: 113–116.
- Picinbono B (1997) On instantaneous amplitude and phase of signals. *IEEE Transactions on Signal Processing* 45: 552–560.
- Santos JF and Falk TH (2014) Updating the SRMR-Cl metric for improved intelligibility prediction for cochlear implant users. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)* 22: 2197–2206.
- Sarria-Paja M and Falk TH (2013) Whispered speech detection in noise using auditory-inspired modulation spectrum features. *IEEE Signal Processing Letters* 20: 783–786.
- Sörnmo L and Laguna P (2005) *Bioelectrical signal processing in cardiac and neurological applications*. Academic Press.
- Tobon Vallejo D and Falk T (2016) Adaptive spectro-temporal filtering for electrocardiogram signal enhancement. *IEEE Journal of Biomedical and Health Informatics* 1. <https://doi.org/10.1109/JBHI.2016.2638120>.
- Tobon VDP, Falk TH, and Maier M (2016) MS-QI: A modulation spectrum-based ECG quality index for telehealth applications. *IEEE Transactions on Biomedical Engineering* 63: 1613–1622. <https://doi.org/10.1109/TBME.2014.2355135>.
- van Vugt MK, Sederberg PB, and Kahana MJ (2007) Comparison of spectral analysis methods for characterizing brain oscillations. *Journal of Neuroscience Methods* 162: 49–63. <https://doi.org/10.1016/j.jneumeth.2006.12.004>.
- Wu S, Falk TH, and Chan W-Y (2011) Automatic speech emotion recognition using modulation spectral features. *Speech Communication* 53: 768–785. <https://doi.org/10.1016/j.specom.2010.08.013>.

Relevant Websites

- <https://github.com/MuSAELab/amplitude-modulation-analysis-toolbox>—Amplitude Modulation Toolbox for Python, Octave and MATLAB.
- <https://sites.google.com/a/uv.edu/isdl/projects/modulation-toolbox>—Modulation Toolbox.