Supplementary Material

Cochlear Transform

Hessa Alfalahi, Ahsan H. Khandoker, Georgios K. Apostolidis, Leontios J. Hadjileontiadis

correspondence: hessa.ealfalahi@ku.ac.ae

June 20, 2025

(S1) Simulated $f_{sig,s}$

The synthetic test signal $f_{sig,s}(t)$ comprised of three components, i.e.:

$$f_{c1}(t) = 6t; f_{c2}(t) = \cos(40\pi t); f_{c3}(t) = 0.5\cos(300\pi t), \tag{1}$$

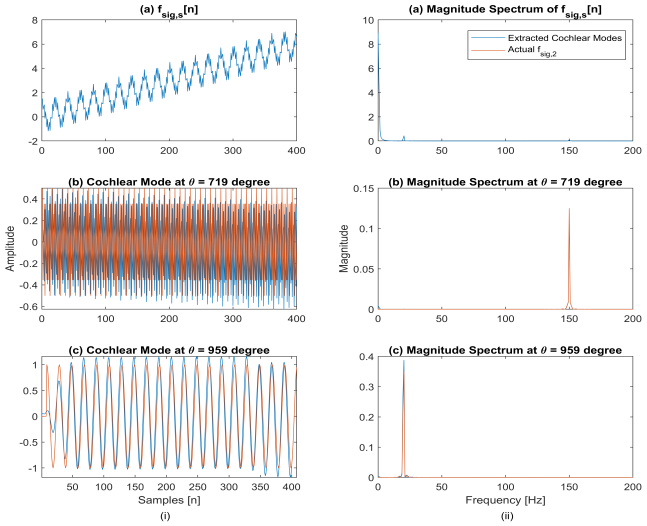
and

$$f_{sig,s}(t) = f_{c1}(t) + f_{c2}(t) + f_{c3}(t). \tag{2}$$

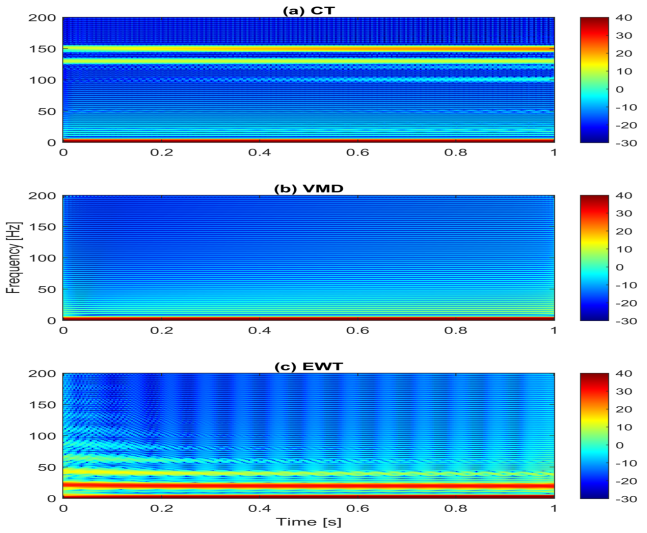
The sampling frequency was set to 400 Hz with $t \in [0, 1]$s. Using α of 0.03, two dominant frequencies were identified, namely 20.3125 Hz and 150 Hz, with errors 1.56% and 0%, respectively. We subsequently applied two Cochlear wavelet functions on the signal. Supplementary Figure 1 shows the signal f_{sig2}, its power spectral density, and the detected spectral peaks. We performed the Cochlear Filtering to evaluate the accuracy in extracting the synthetic signal mono-components. Although the frequencies of the mono-components fall far from the hearing range of high sensitivity (i.e., between 380Hz and 3400Hz [1]), we report the ability of the CD to extract the mono-components with nearly perfect reconstruction. Regarding TFR, we compared the CT-TFR to VMD-TFR and EWT-TFR, as illustrated in Supplementary Fig. 2. The CT-TFR results in higher TFR resolution, especially at the low amplitude mode at 150 Hz, amplified by the cochlea's active tuning property for low amplitude signals. Regarding the component at 130 Hz, we believe that it results from the acoustic distortion mechanism of the cochlea; a byproduct of the non-linearity of the basilar membrane that results in acoustic energy spread at $150 - 20 = 130$Hz.

(S2) Electrocardiogram (ECG) signal

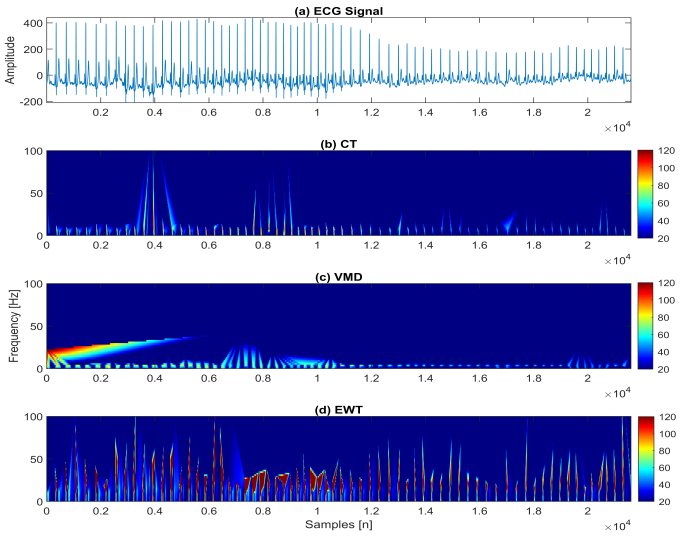
The ECG is a multi-resolution signal generated by the contractions of the heart. It is usually contaminated by noise, such as baseline wandering, 50/60 Hz powerline interference, and motion artifacts. Herein, we initially test the ECG de-noising performance of the CT [2]. Nevertheless, since the latter provides resolution adaptiveness by both frequency and IA, we expect it to reveal clinically plausible features of the ECG besides de-noising the signal. We, therefore, compared its performance to that of the VMD and EWT. We used an ECG signal from the MIT-BIH Physionet open access dataset [3], sampled at $f_s = 360$Hz. First, we up-sampled the ECG to 1440Hz before performing Cochlear modulation to ensure that the bandwidth of the modulated ECG signal falls in the audible range (the high-sensitivity cochlear region). To determine the number of modes and their center frequencies via initial spectral analysis, we set α to 0.02. The results are illustrated in Supplementary Figure. 3(a)-(c) for the CT-TFR, VMD-TFR, and EWT-TFR, respectively. From Supplementary Figure 3 it is clear that the CT-TFR results in smoother TFR (Supplementary Figure 3(a)) with greater localization potential of the clinically relevant information, such as the S-T segments of the ECG, when compared to VMD-TFR (Supplementary Figure 3(b)) and EWT-TFR (Supplementary Figure 3(c)). In particular, the CT-TFR best displays the transient information and the fast-varying IF between consecutive R and S peaks in the ECG signal. This is attributed to the fact that the estimation of fast-varying IFs of low amplitude modes by state-of-the-art methods is hard [2], while the adaptive characteristic of the spatial cochlear wavelets and the clustering of the resulting cochlear modes preclude this problem. The VMD-TFR and the EWT-TFR, on the other hand, both result in good localization of the R peaks of the ECG.



Supplementary Figure 1: CD of (i)-(a) $f_{sig,s}$ with the corresponding retrieved signal's mono-components at (i)-(b) $\theta = 719°$ and (i)-(c) $\theta = 959°$. (ii)-(a):(c) The corresponding magnitude spectra of (i)-(a):(c), respectively. In all cases, the blue and orange colors denote the original and the CD-based estimated output, respectively.

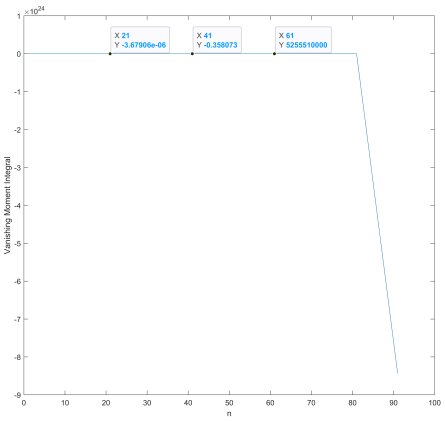


Supplementary Figure 2: TFR of $f_{sig,s}$ using (a) the CT, (b) the VMD, and (c) the EWT. The colormap indicates the intensity of the instantaneous amplitude in dB.



Supplementary Figure 3: TFR of (a) the ECG using (b) the CT, (c) the VMD, and (d) the EWT. The colormap indicates the intensity of the instantaneous amplitude in dB.

(S3) Vanishing Moments



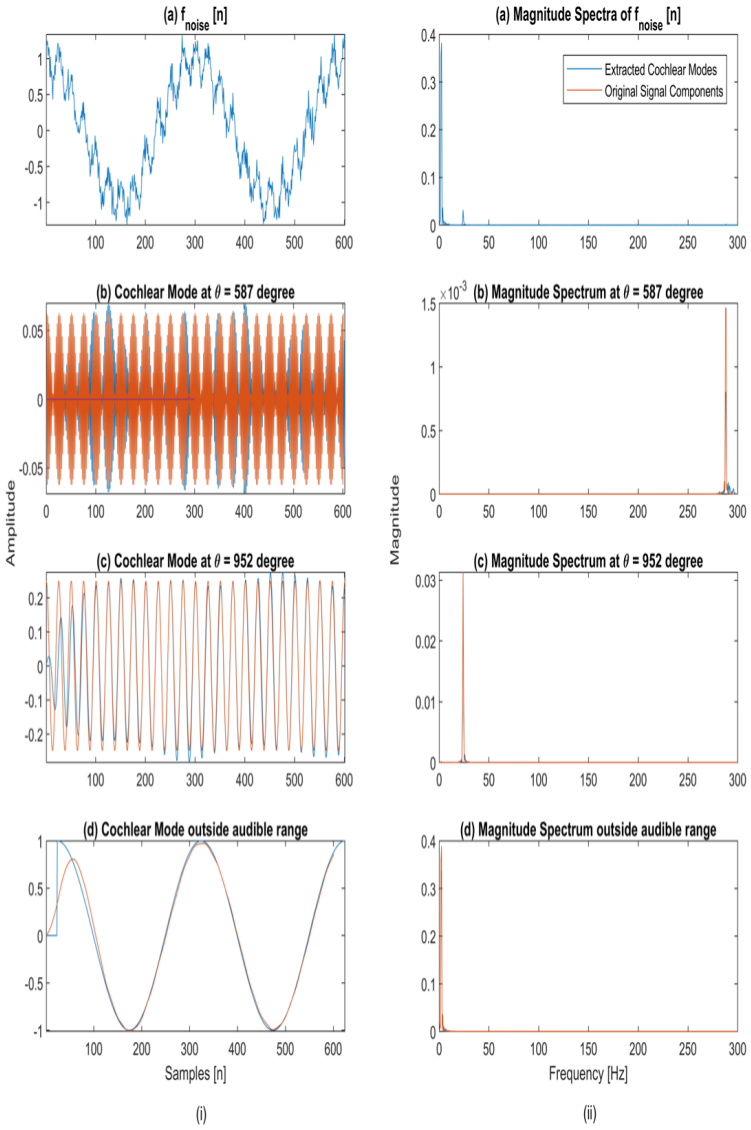
Supplementary Figure 4: Vanishing Moment Integral for different n.

(S4) Noise Robustness

To illustrate the noise robustness of the CD, we used as input to the CD the following tri-harmonic signal with additive WGN, adopted from [4] for comparative analysis:

$$f_{noise}(t) = \cos(4\pi t) + \frac{1}{4}\cos(48\pi t) + \frac{1}{16}\cos(576\pi t) + \eta, \tag{3}$$

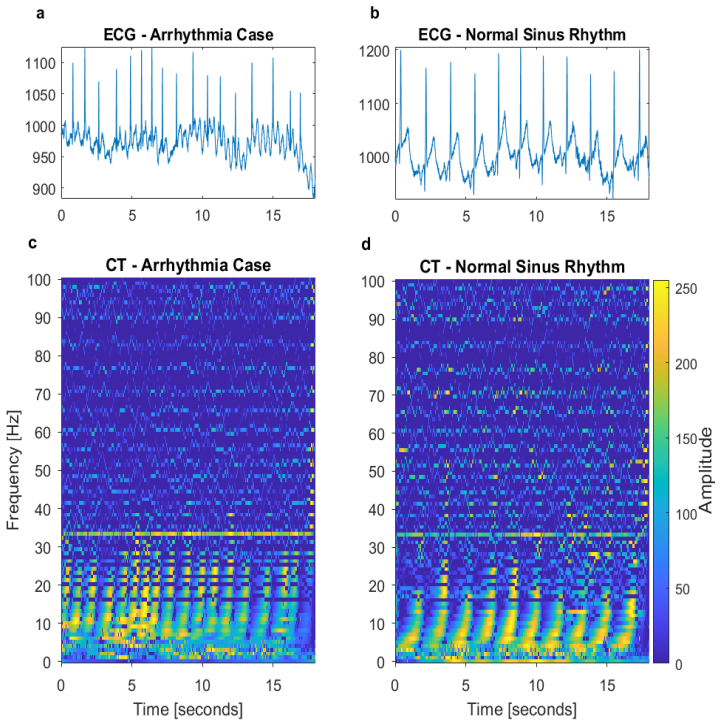
where $\eta \sim N(0, \sigma)$ is the additive WGN with σ set to 0.1 and sampling frequency of 600 Hz [4]. We first started by spectral peaks selection to determine *a priori* the number of modes and their corresponding dominant frequency. We set the peak selection threshold α to 0.003. The noise analysis results are shown in Fig. 5; three peaks were detected at $2.34Hz$, $23.44Hz$, and $288.28Hz$. The estimation of the dominant frequencies errors were 1.7%, 0.23%, and 0.09% for the three components, respectively, compared to 0.27%, 1.11%, and 0.18% with VMD for the same input signal and noise settings [4]. Despite the low amplitude, we show that the proposed peak selection and the CD approach localized and retrieved the weak, high-frequency component with higher accuracy compared to VMD, in addition to the low-frequency mode, which is outside the audible range, even under noise existence, as illustrated in Supplementary Figure5.



Supplementary Figure 5: Noise robustness of the CD. (i)-(a) The noisy signal given by 3, and the resulting cochlear modes at (i)-(b):(d). (ii)-(a):(d) The corresponding magnitude spectra of (i)-(a):(d), respectively. In all cases, the blue line represents the original signal mono-components while the cochlear modes are given in orange.

(S5) Normal Sinus Rhythm vs. Cardiac Arrhythmia

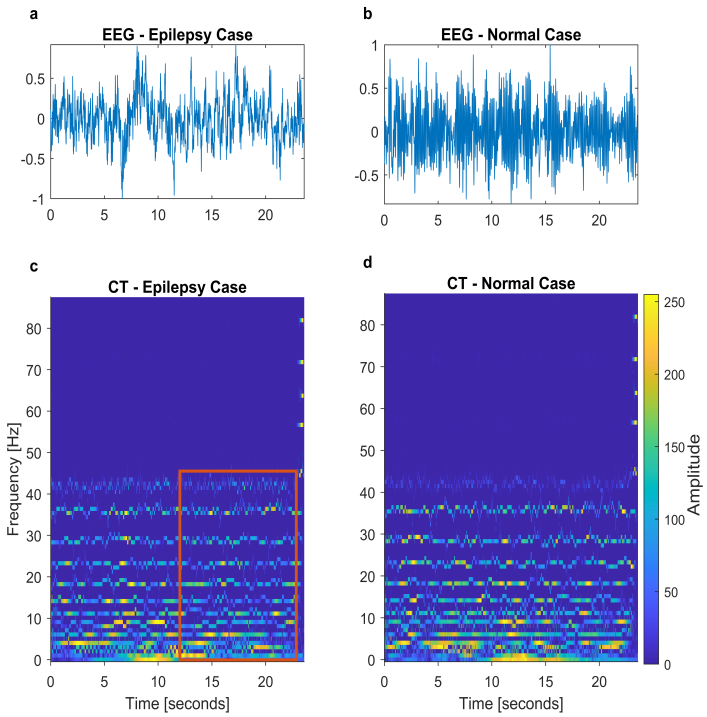
In order to examine the ability of the CT to reveal abnormalities in ECG signals, we compared the CT representation of and ECG signal with arrhythmia to normal sinus rhythm. As shown in Supplementary Figure 6, the CT reveals the unstable R peaks in arrhythmia, reflecting the fast palpitation of the heart.



Supplementary Figure 6: (a) ECG signal with arrhythmia episode. (b) ECG of normal sinus rhythm. (C) CT of (a). (d) CT of (b).

(S6) Normal EEG vs. Epileptic EEG

Here, we applied the CT to normal and epileptic EEG to investigate the ability of the CT in distinguishing the distinct features of Epilepsy. As shown in Supplementary Figure 7, the CT reveals the rhythmic generation of high frequency activity in epilepsy EEG, as well as the diminished alpha band (8-14Hz) as opposed to the normal EEG case.



Supplementary Figure 7: (a) EEG signal with epileptic seizure. (b) Normal ECG signal. (c) CT of epileptic EEG. (d) CT of normal EEG.

References

[1] T. M. Talavage, M. I. Sereno, J. R. Melcher, P. J. Ledden, B. R. Rosen, and A. M. Dale, "Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity," *Journal of neurophysiology*, vol. 91, no. 3, pp. 1282–1296, 2004.

[2] M. A. Colominas, S. Meignen, and D.-H. Pham, "Time-frequency filtering based on model fitting in the time-frequency plane," *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 660–664, 2019.

[3] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE engineering in medicine and biology magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[4] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE transactions on signal processing*, vol. 62, no. 3, pp. 531–544, 2013.