Homework 10

Due May 3 at 11 pm

Yves Greatti - yg390

1. (Hann window) In this problem we analyze the Hann window in the frequency domain.

   (a) Prove that for any vector $x \in \mathbb{C}^N$, if $y \in \mathbb{C}^N$ is defined as

   $$y[j] := x[j] \exp\left(\frac{i2\pi mj}{N}\right), \tag{1}$$

   for some integer m, then the DFT of y equals $\hat{y} = \hat{x}^{\downarrow m}$, where $\hat{x}$ is the DFT of x.
   We have

   $$\begin{aligned}\hat{y}[k] &= \sum_{j=1}^{N} x(j) \exp\left(\frac{-i2\pi kj}{N}\right) \exp\left(\frac{i2\pi mj}{N}\right) \\ &= \sum_{j=1}^{N} x(j) \exp\left(\frac{-i2\pi(k-m)j}{N}\right) \\ &= \hat{x}^{\downarrow m}[k]\end{aligned}$$

   (b) The Hann window $h \in \mathbb{C}^N$ of width $2w$ equals

   $$h[j] := \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{\pi j}{w}\right)\right) & \text{if } |j| \leq w, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

   Use the result from part (a) to show that the DFT of h can be expressed as

   $$\hat{h} = \frac{1}{2}\hat{\pi} + \frac{1}{4}\hat{\pi}^{\downarrow -N/2w} + \frac{1}{4}\hat{\pi}^{\downarrow N/2w}. \tag{3}$$

   The rectangular window $\hat{\pi} \in \mathbb{C}^N$ with width $2w$ is defined by

   $$\hat{y}[j] = \begin{cases} 1 & \text{if } |j| \leq w, \\ 0 & \text{otherwise} \end{cases}$$

   For $|j| \leq w$, $h[j] = \frac{1}{2}\left(1 + \cos\left(\frac{\pi j}{w}\right)\right) = \left(\frac{1}{2} \cdot 1 + \frac{\exp\left(\frac{i\pi j}{w}\right) + \exp\left(\frac{-i\pi j}{w}\right)}{2}\right) = \frac{1}{2} \cdot 1 + \frac{1}{4}\exp\left(\frac{i\pi j}{w}\right) \cdot 1 + \frac{1}{4}\exp\left(\frac{-i\pi j}{w}\right) \cdot 1$. Using part a, with $m = \frac{N}{2w}$, the DFT of h is then $\hat{h} = \frac{1}{2}\hat{\pi} + \frac{1}{4}\hat{\pi}^{\downarrow -N/2w} + \frac{1}{4}\hat{\pi}^{\downarrow N/2w}$.

   (c) Plot $\hat{h}$ as well as the different components in Eq. (3). Interpret what you see in terms of the desired properties of a windowing function.
   The sum of all the three dotted lines give the solid line. And we see that

- there are some cancellations of the side lobes coming from the two smaller sinc functions so less distortion for the Hann window
- the width of the main lobe is doubled

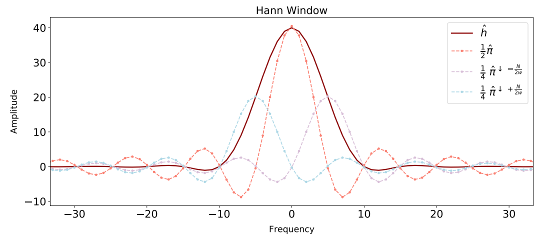


Figure 1: Frequency response of the Hann window

2. (STFT inverse) In this problem we show a simple way to invert the STFT.

   (a) In the definition of the STFT set $w_{[\ell]}$ to be a rectangular window where all entries are equal to one, and let $\alpha_{\text{ov}} = 0.5$. Show that the STFT can be inverted using just two operations: applying the inverse DFT and subsampling.
   In the same definition we have

$$
\begin{aligned}
\text{STFT}_{[l]}(x)[k,s] &= \left\langle x, \xi_k^{\downarrow \frac{sl}{2}} \right\rangle \\
&= \sum_{j=1}^{N} x[j] w_{[l]}[j - \frac{sl}{2}] \exp\left(\frac{-i2\pi kj}{l}\right) \\
&= \sum_{j=\frac{l(s-2)}{2}}^{\frac{l(s+2)}{2}} x[j] \exp\left(\frac{-i2\pi kj}{l}\right)
\end{aligned}
$$

   which leads to the definition of the STFT as the multiplication of DFT matrices of size l with identity matrices of size l.

$$
\text{STFT} = \begin{bmatrix} F_{[l]} & 0 & 0 \cdots \\ 0 & F_{[l]} & 0 \cdots \\ 0 & 0 & F_{[l]} \cdots \end{bmatrix} \begin{bmatrix} \text{diag}(w_{[l]}) & 0 & 0 \cdots \\ 0 & \text{diag}(w_{[l]}) & 0 \cdots \\ 0 & 0 & \text{diag}(w_{[l]}) \cdots \end{bmatrix} x
$$

   If there was no overlap between the windows, the initial signal x can be fully recovered by multiplying the STFT by

$$
\begin{aligned}
x &= \begin{bmatrix} \text{diag}(w_{[l]})^{-1} & 0 & 0 \cdots \\ 0 & \text{diag}(w_{[l]})^{-1} & 0 \cdots \\ 0 & 0 & \text{diag}(w_{[l]})^{-1} \cdots \end{bmatrix} \begin{bmatrix} F_{[l]}^* & 0 & 0 \cdots \\ 0 & F_{[l]}^* & 0 \cdots \\ 0 & 0 & F_{[l]}^* \cdots \end{bmatrix} \text{STFT} \\
&= \begin{bmatrix} F_{[l]}^* & 0 & 0 \cdots \\ 0 & F_{[l]}^* & 0 \cdots \\ 0 & 0 & F_{[l]}^* \cdots \end{bmatrix} \text{STFT}
\end{aligned}
$$

   Because of the window overlap we have redundant signal components which we need to eliminate by subsampling. With $\alpha_{\text{ov}} = 0.5$, every component of the initial signal x is doubled, so we need to keep half of the recovered signal, so we can subsample the recovered signal $\hat{x}$ by multiplying it with by an identity matrix which has only half of its entries equal to 1.

   (b) What is the disadvantage of using this rectangular window? Small perturbations in the STFT can become amplified in the recovered signal $\hat{x}$, as we saw in question 1, a Hann window is more stable and has less fluctuations.

3. (Haar wavelet) Define the discrete Haar wavelet $\mu_{2^s,p} \in \mathbb{R}^{2^n}$ at scale 2^s and position p by

$$\mu_{2^s,p}[j] := \begin{cases} -1/\sqrt{2^s} & \text{if } j \in \{p \cdot 2^s, p \cdot 2^s + 1, \ldots, p \cdot 2^s + 2^{s-1} - 1\}, \\ 1/\sqrt{2^s} & \text{if } j \in \{p \cdot 2^s + 2^{s-1}, 2^s + 2^{s-1} + 1, \ldots, (p+1) \cdot 2^s - 1\}, \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < s \leq n$ and $0 \leq p \leq 2^{n-s} - 1$. Define the discrete Haar scaling function $\varphi_{2^s,p} \in \mathbb{R}^{2^n}$ at scale 2^s and position p by

$$\varphi_{2^s,p}[j] = \begin{cases} 1/\sqrt{2^s} & \text{if } j \in \{p \cdot 2^s, p \cdots 2^s + 1, \ldots, (p+1) \cdot 2^s - 1\}, \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < s \leq n$ and $0 \leq p \leq 2^{n-s} - 1$. The code for this exercise is contained in the haar.py file. Include all generated plots in your submission.

(a) Define $V_0 := \mathbb{R}^{2^n}$. For $k > 0$, let $V_k \subset \mathbb{R}^{2^n}$ denote the subspace of all vectors that are constant on segments of size 2^k. That is

$$V_k := \{x \in \mathbb{R}^{2^n} : x[i] = x[j] \text{ if } \lfloor i/2^k \rfloor = \lfloor j/2^k \rfloor\}.$$

Give an orthonormal basis for V_k. What is the dimension of V_k?
V_k is the subspace of all vectors $x \in \mathbb{R}^{2^n}$ for which all the contiguous components modulo 2^k of the vector, are constant. Let define one of such vector of $\mathbb{R}^{2^n}$

$$\psi_{2^k,p}[j] = \begin{cases} 2^{-\frac{k}{2}} & \text{if } j \in \{p \cdot 2^k, p \cdot 2^k + 1, \ldots, (p+1) \cdot 2^k - 1\}, \\ 0 & \text{otherwise,} \end{cases}$$

where $0 \leq p \leq 2^{n-k} - 1$. These vectors are orthogonal since they are not overlapping for different position p, and they have unit norm. The number of these vectors is equal to the number of position p thus the dimension of V_k is 2^{n-k}.

(b) Fix $0 \leq k < n$, and note that $V_k \supset V_{k+1}$. Give an orthonormal basis for the set

$$W_{k+1} = \{x \in V_k : \langle x, y \rangle = 0 \text{ for all } y \in V_{k+1}\},$$

the orthogonal complement of V_{k+1} in V_k. Thus $V_k = V_{k+1} \oplus W_{k+1}$. What is the dimension of W_{k+1}?

Vector of the form

$$\xi_{2^{k+1},p}[j] := \begin{cases} -1/\sqrt{2^{k+1}} & \text{if } j \in \{p \cdot 2^{k+1}, p \cdot 2^{k+1} + 1, \ldots, p \cdot 2^{k+1} + 2^k - 1\}, \\ 1/\sqrt{2^{k+1}} & \text{if } j \in \{p \cdot 2^{k+1} + 2^k, 2^{k+1} + 2^k + 1, \ldots, (p+1) \cdot 2^{k+1} - 1\}, \\ 0 & \text{otherwise,} \end{cases}$$

belong to V_k (scaling by $\frac{1}{2}$), are orthonormal (unit norm and orthogonal to each other for different p) and they are also orthogonal to the vectors $\psi_{2^{k+1},p}$. Also any vector of V_k that is orthogonal to V_{k+1} is in W_{k+1} thus we can write $V_k = V_{k+1} \oplus W_{k+1}$. Dimension of W_{k+1} is $2^{n-(k+1)}$.

(c) For $1 \leq k \leq n$ give an orthonormal basis for the set

$$W_{\leq k} = \{x \in \mathbb{R}^{2^n} : \langle x, y \rangle = 0 \text{ for all } y \in V_k\},$$

the orthogonal complement of V_k in $\mathbb{R}^{2^n}$. Thus $\mathbb{R}^{2^n} = V_k \oplus W_{\leq k}$. What is the dimension of $W_{\leq k}$?
Using part b we can

(d) Complete the wavelet and scaling functions in *haar.py* that implement μ and φ above, respectively. See the comments for more details.

(e) Complete the projectV function that orthogonally projects a given vector onto V_k. [Hint: Consider averaging the values on each segment.]

(f) Complete the projectW function that orthogonally projects a given vector onto W_k. [Hint: You can use projectV.]

(g) Complete the function *wavelet_coeffs* which computes all of the (non-overlapping) wavelet coefficients of a given data vector at a given scale. See the comments for more details.

(h) Report the plots generated by the code, which apply your wavelet transform to some electrocardiogram data.

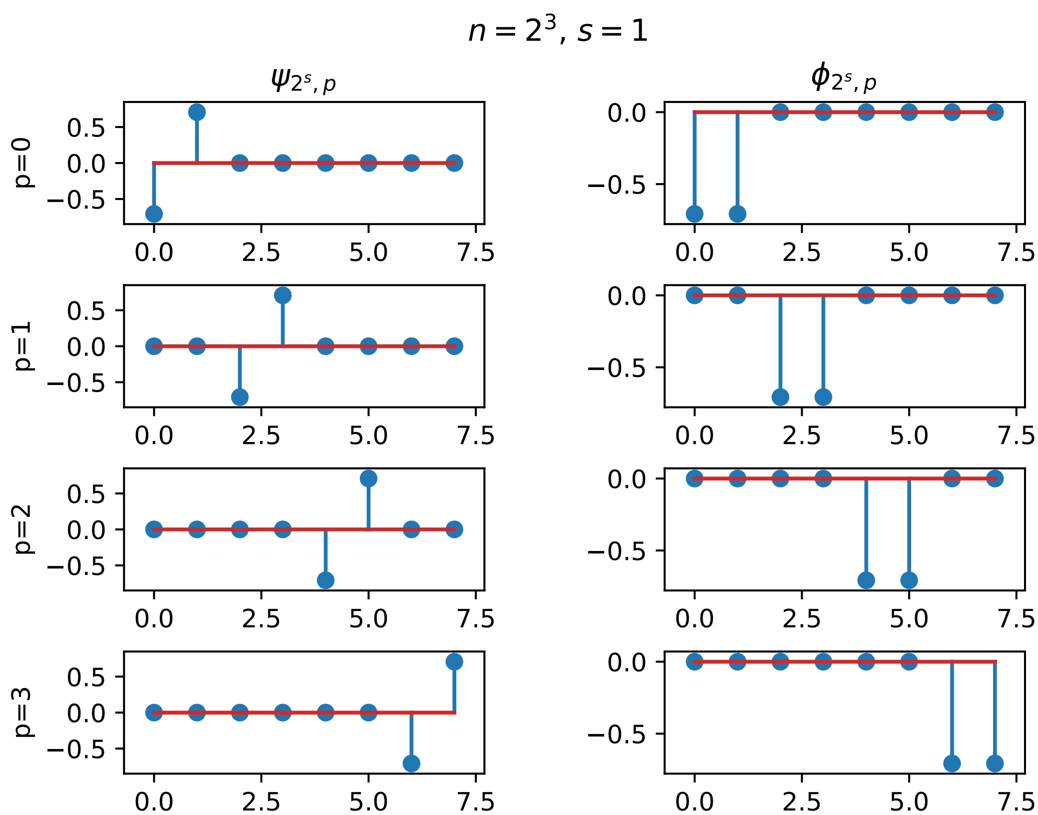


Figure 2: Wavelet $\psi_{2^s,p}$ and scaling $\phi_{2^s,p}$

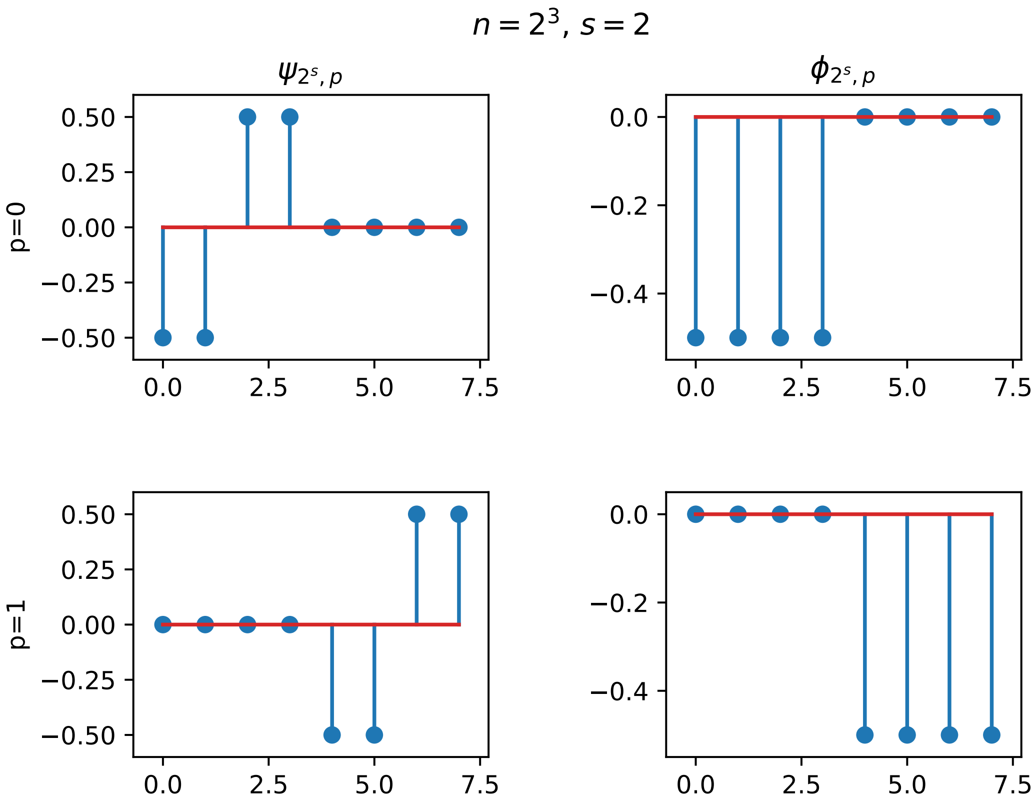
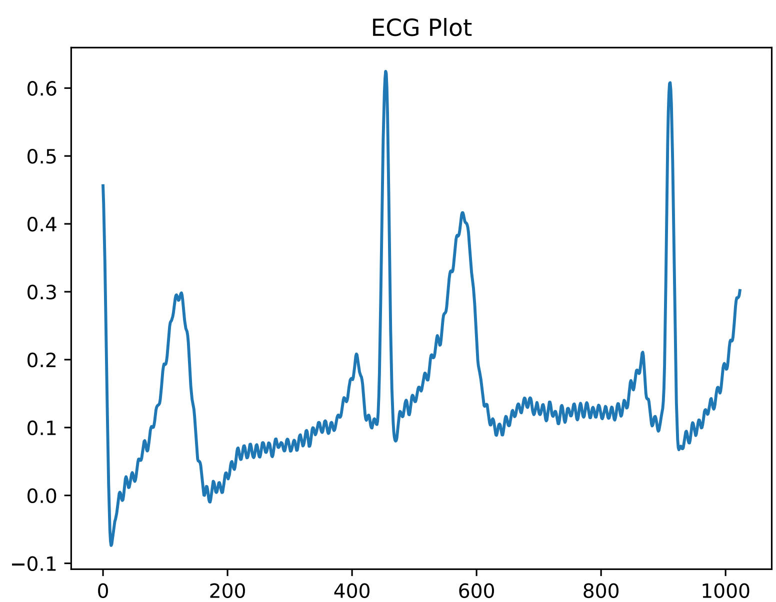
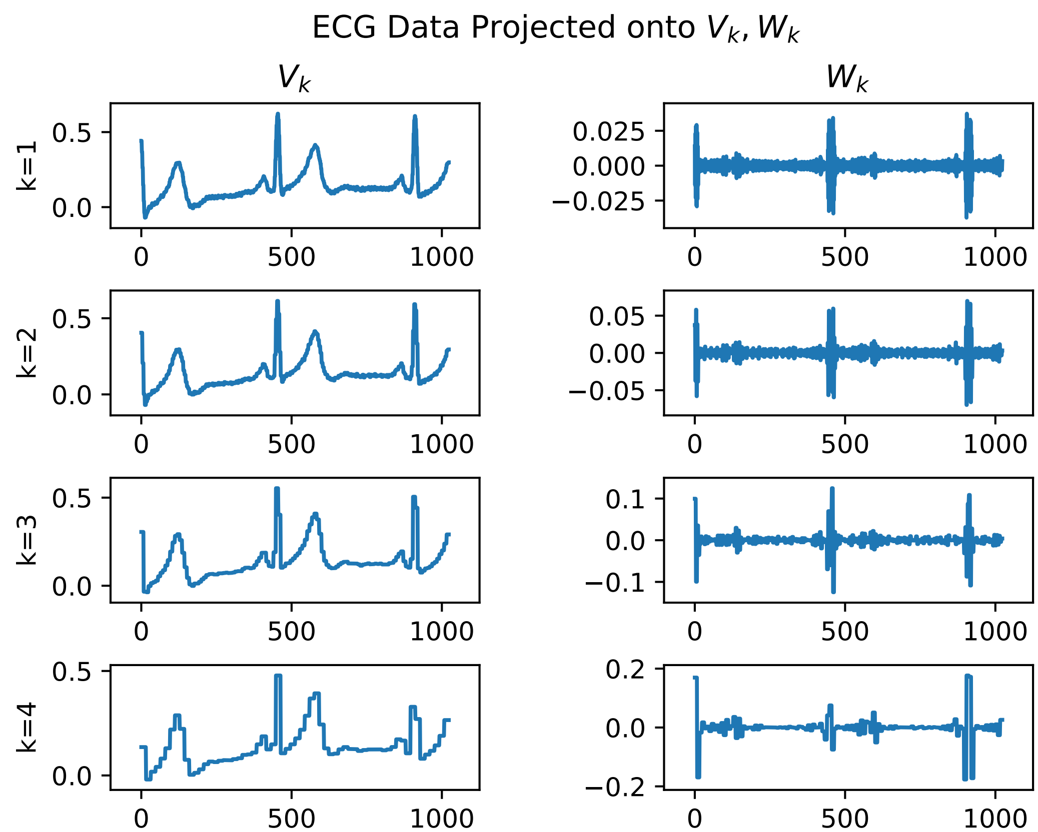
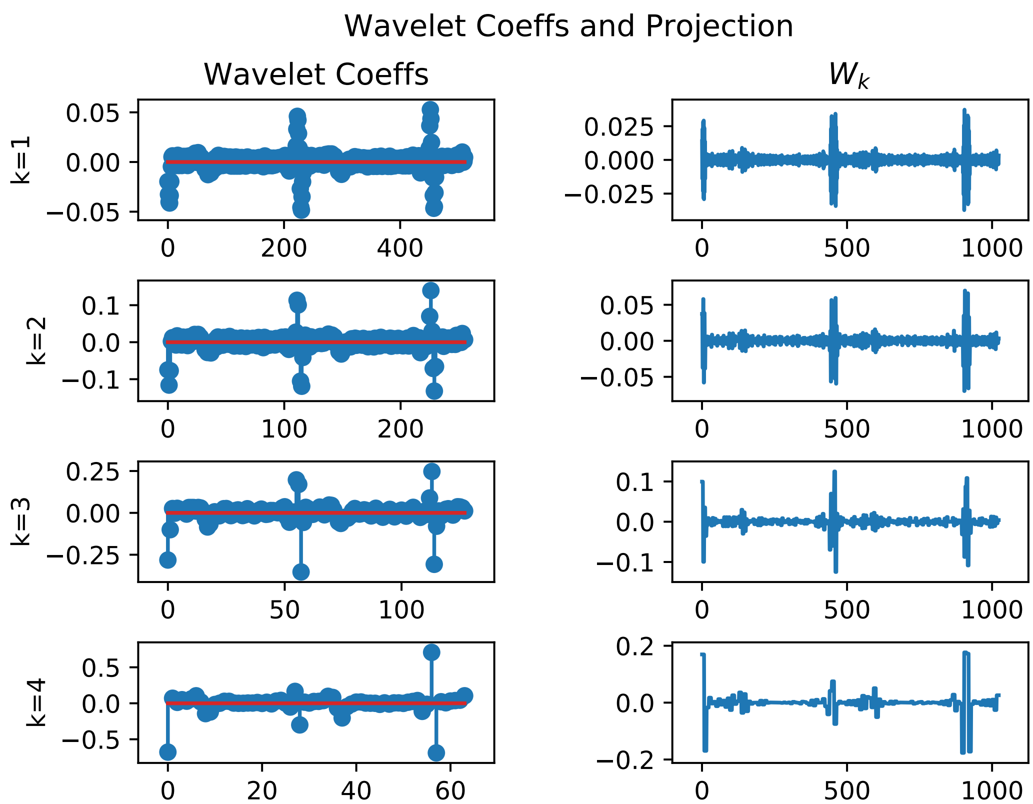


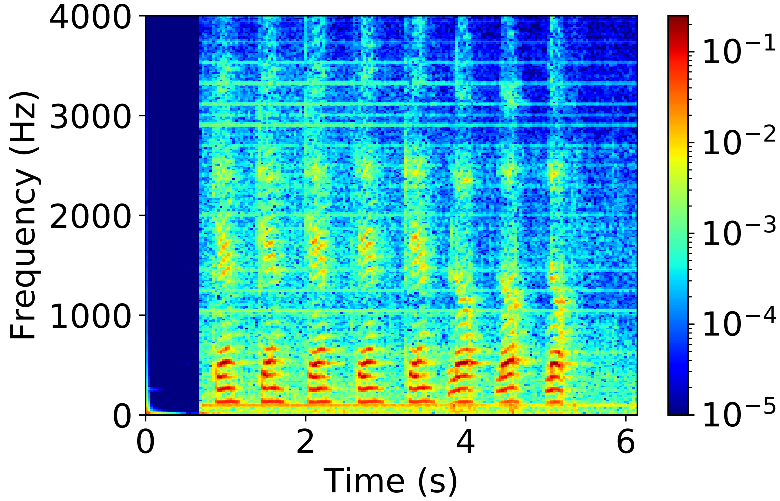
Figure 3: Wavelet $\psi_{2^s,p}$ and scaling $\phi_{2^s,p}$

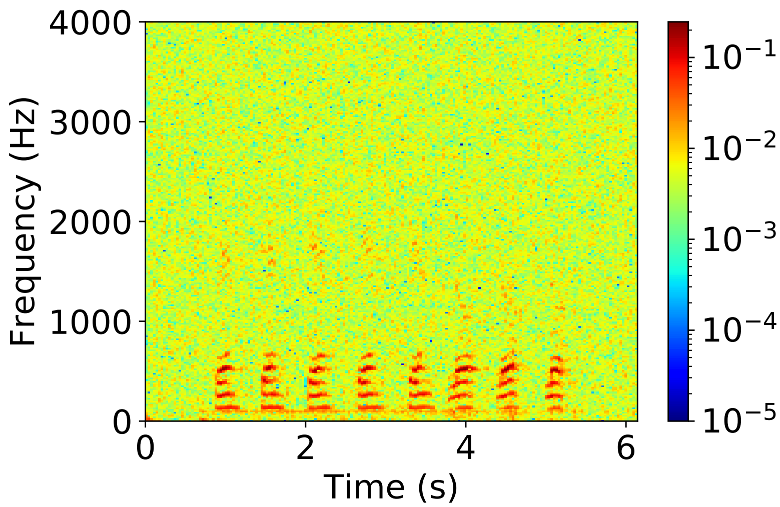


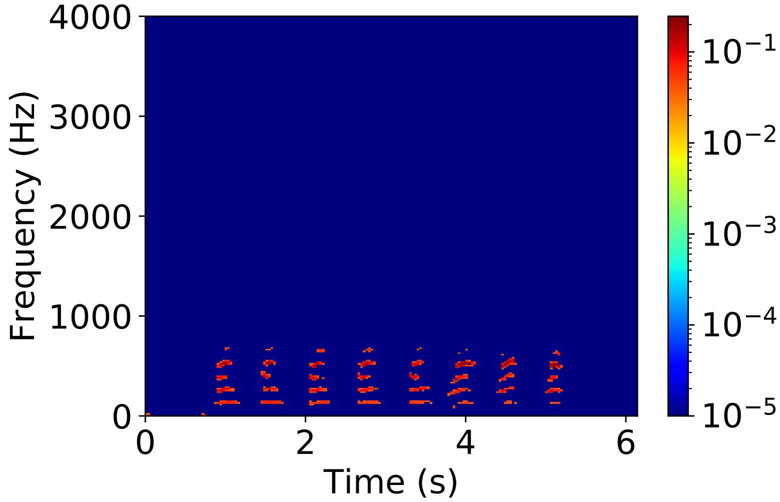

ECG Data Projected onto V_k, W_k
V_k
W_k
k=1
k=2
k=3
k=4
0.5
0.0
0.025
0.000
−0.025
0.05
0.00
−0.05
0.1
0.0
−0.1
0.2
−0.2
0
500
1000

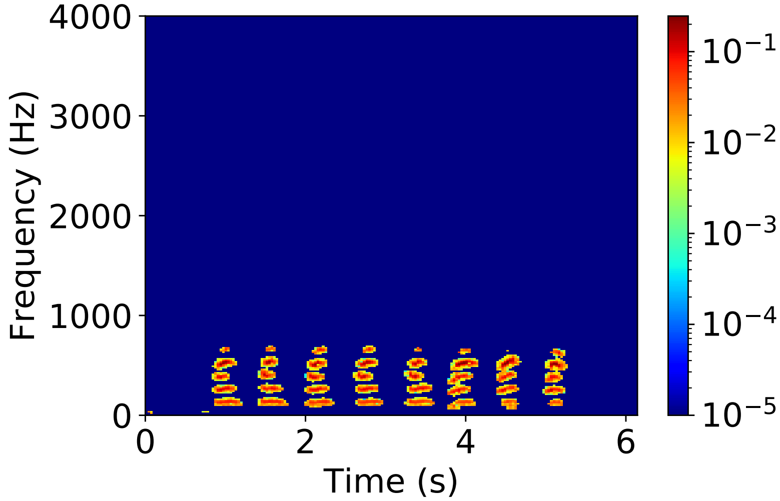

Wavelet Coeffs and Projection
Wavelet Coeffs
W_k
k=1
k=2
k=3
k=4

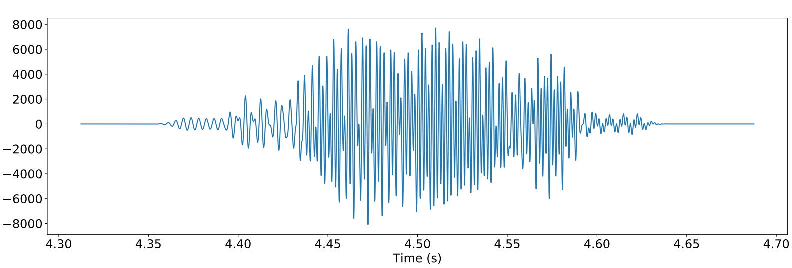
4. (Denoising with the STFT) In the lecture, we saw that STFT often yields sparse representation for a signal but dense representation for noise. Building on this, we derived hard thresholding (Algorithm 4.1 in notes) and block thresholding (Algorithm 4.2 in notes) to denoise signals. In this question, we will denoise audio signals. `audio_denoising.ipynb` contains skeleton code for the task. The notebook will download required dataset and contains other utility functions for loading data, plotting and playing the audio signals. You have to fill in the functions `get_block_L2_norm()` and `stft_denoising()`. Report all the plots generated by the script.

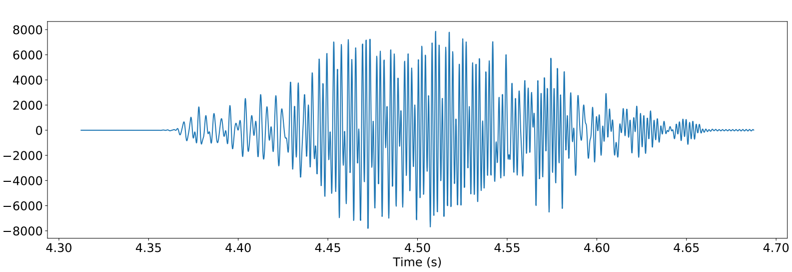


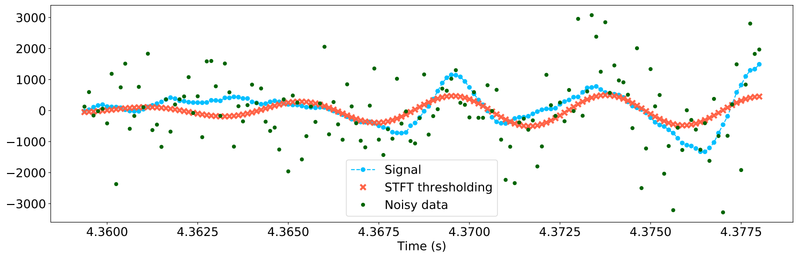



4000
3000
2000
1000
0
Frequency (Hz)
0
2
4
6
Time (s)
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}

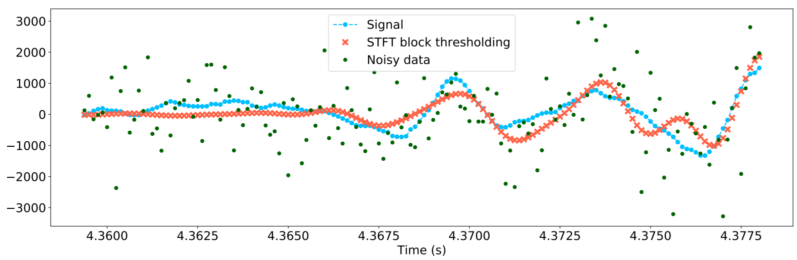

4000
3000
2000
1000
0
Frequency (Hz)
0
2
4
6
Time (s)
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}


8000
6000
4000
2000
0
−2000
−4000
−6000
−8000
4.30
4.35
4.40
4.45
4.50
4.55
4.60
4.65
4.70
Time (s)


8000
6000
4000
2000
0
−2000
−4000
−6000
−8000
4.30
4.35
4.40
4.45
4.50
4.55
4.60
4.65
4.70
Time (s)



3000
2000
1000
0
−1000
−2000
−3000
4.3600
4.3625
4.3650
4.3675
4.3700
4.3725
4.3750
4.3775
Time (s)
Signal
STFT thresholding
Noisy data



3000
2000
1000
0
−1000
−2000
−3000
4.3600
4.3625
4.3650
4.3675
4.3700
4.3725
4.3750
4.3775
Time (s)
Signal
STFT block thresholding
Noisy data