Introduction to Data Processing and Representation - HW1

Submit by: 315017277 , 315234351 (I.Ds)

1. Solving the L^p problem using the L^2 norm

(a) $p=1$: for each interval I_i take the median of f in it, means: for $x \in I_i$: $\hat{f}(x) = y_i$ where y_i holds that $\int_{\substack{a \le x \le b \\ f(x) \le y_i}} dx = \int_{\substack{a \le x \le b \\ y_i < f(x)}} dx$.

$p=2$: now we take the average in each interval. For $x \in I_i$: $\hat{f}(x) = \frac{1}{|I_i|} \int_{I_i} f(x)\,dx$.

(b) As we are looking for $\hat{f}$ which is constant on every interval I_i, we can reduce the problem to finding a value y_i that minimize $\int_{I_i} (f(x) - y_i)^2 w(x)\,dx$ for each $i = 1, \dots, N$ (see part (d) of this question).
To find that y_i we will derive the target function with respect to y_i, and find where it equals zero:

$$\frac{d}{dy_i} \int_{I_i} (f(x) - y_i)^2 w(x)\,dx = \int_{I_i} \frac{d}{dy_i} (f(x) - y_i)^2 w(x)\,dx$$

$$= -\int_{I_i} 2w(x) \cdot (f(x) - y_i)\,dx \underset{\text{wanted}}{=} 0$$

$$\Rightarrow \int_{I_i} w(x) f(x)\,dx = y_i \cdot \int_{I_i} w(x)\,dx$$

$$\Rightarrow \boxed{y_i = \frac{1}{\int_{I_i} w(x)dx} \cdot \int_{I_i} w(x) f(x)\, dx}$$

ⓒ We will use the same approach - we will look for y_i that minimiz $\int_{I_i} w(x)|f(x) - y_i| dx$.

$$\frac{d}{dy_i} \int_{I_i} w(x)|f(x) - y_i| dx = \frac{d}{dy_i}\left(\int_{\substack{x \in I_i \\ f(x) < y_i}} w(x)(y_i - f(x)) dx + \int_{\substack{x \in I_i \\ y_i \le f(x)}} w(x)(f(x) - y_i) dx \right)$$

$$= \left(\int_{\substack{x \in I_i \\ f(x) < y_i}} w(x)\, dx - \int_{\substack{x \in I_i \\ y_i \le f(x)}} w(x)\, dx \right) \underset{\text{wanted}}{=} 0$$

$$\Rightarrow \int_{\substack{x \in I_i \\ f(x) < y_i}} w(x) dx = \int_{\substack{x \in I_i \\ y_i \le f(x)}} w(x) dx \qquad (*)$$

We will have to define $\hat{f}(x) = y_i$ for a number y_i that satisfies $(*)$.

ⓓ In the next page we prove a claim that states that

$$\min_{\hat{f}} \int_0^1 |f(x) - \hat{f}(x)|^p w(x)\, dx = \sum_{i=1}^{N} \min_{\hat{f}_i} \int_{I_i} |f_i(x) - \hat{f}_i(x)|^p w(x)\, dx.$$

Using it, we can define

$$\mathcal{E}_i^p(f_i, \hat{f}_i) = \min_{\hat{f}_i} \int_{I_i} |f_i(x) - \hat{f}_i(x)|^p w(x)\, dx$$

and get that $\mathcal{E}^p(f, \hat{f}) = \sum_{i=1}^{N} \mathcal{E}_i^p(f_i, \hat{f}_i)$.

claim:

$$\min_{\hat f} \int_0^1 |f(x)-\hat f(x)|^p w(x)\,dx = \sum_{i=1}^{N} \min_{\hat f_i} \int_{I_i} |f_i(x)-\hat f_i(x)|^p w(x)\,dx$$

proof:

$\geq$:

$$\min_{\hat f} \int_0^1 |f(x)-\hat f(x)|^p w(x)\,dx = \min_{\hat f} \sum_{i=1}^{N} \int_{I_i} |f_i(x) - \underbrace{\hat f|_{I_i}(x)}_{\hat f \text{ reduced to } I_i}|^p w(x)\,dx \,.$$

$$\geq \min_{\hat f} \sum_{i=1}^{N} \min_{\hat f_i} \int_{I_i} |f_i(x) - \hat f_i(x)|^p w(x)\,dx \overset{(*)}{=} \sum_{i=1}^{N} |f_i(x) - \hat f_i(x)|^p w(x)\,dx$$

(*) note that $\hat f$ does not participate in the term that we minimize with respect to it.

$\leq$:

Assume towards a contradiction that

$$\min_{\hat f} \int_0^1 |f(x)-\hat f(x)|^p w(x)\,dx > \sum_{i=1}^{N} \min_{\hat f_i} \int_{I_i} |f_i(x)-\hat f_i(x)|^p w(x)\,dx \,.$$

This means that there is a set $\{\hat g_i : I_i \to \mathbb{R}\}_{i=1}^{N}$ such that

$$\min_{\hat f} \int_0^1 |f(x) - \hat f(x)|^p w(x)\,dx > \sum_{i=1}^{N} |f_i(x) - \hat g_i(x)|^p w(x)\,dx \,.$$

Define $\hat g : [0,1] \to \mathbb{R}$ with $\hat g(x) = \hat g_i(x)$ if $x \in I_i$ (well defined as I_i's are disjoint and covers $[0,1]$).
Then

$$\min_{\hat f} \int_0^1 |f(x) - \hat f(x)|^p w(x)\,dx > \sum_{i=1}^{N} |f_i(x) - \hat g_i(x)|^p w(x)\,dx = \int_0^1 |f(x) - \hat g(x)|^p w(x)\,dx$$

and we got a contradiction.

$\blacksquare$

(e) (i) Assume $f_i(x) \neq \hat{f}_i(x)$ for all $x \in I_i$. Thus, the term $1/(f_i(x)-\hat{f}_i(x))^2$ is well defind for all $x \in I_i$. Define $w_{f_i,\hat{f}_i} : I_i \to \mathbb{R}$ by $w_{f_i,\hat{f}_i}(x) = |f_i(x)-\hat{f}_i(x)|^p \cdot \frac{1}{(f_i(x)-\hat{f}_i(x))^2}$ (and $w_{f_i,\hat{f}_i}$ is well defined and positive), and we get the desired.

(ii)

$$\mathcal{E}_i^p(f_i,\hat{f}_i) = \min_{\hat{f}_i} \int_{I_i} w(x)|f_i(x)-\hat{f}_i(x)|^p dx = \min_{\hat{f}_i} \int_{I_i} \underbrace{w(x) \cdot \frac{|f_i(x)-\hat{f}_i(x)|^p}{(f_i(x)-\hat{f}_i(x))^2}}_{:= w_{f_i,\hat{f}_i}} \cdot (f_i(x)-\hat{f}_i(x))^2 dx$$

(iii) Recall that $\hat{f}_i$ is constant on I_i. Than solving the optimization problem for $\hat{f}_i$ is like trying to find $y_i \in \mathbb{R}$ that minimize $\int_{I_i} w_{f,y_i}(x) \cdot |f(x)-y_i|^p dx$.
If w_{f,y_i} was independent from y_i we could derive the target function with respect to y_i and find the wanted value as we did before.
If, on the other hand, w_{f_i,y_i} is not independent of y_i we then we had to use derivation of a multiplication and then extract y_i from there, and that will be harder, as w_{f_i,y_i} is not necessarily simple.

(iv) In the case where f_i and $\hat{f}_i$ might be equal, the term $|f_i(x)-\hat{f}_i(x)|^p/(f_i(x)-\hat{f}_i(x))^2$ might not be defined in some part of the points in I_i. So, $w_{f_i,\hat{f}_i}$ can't have the discussed above form. Even so, we would believe that there is a sequence of functions $\{w^{(n)}_{f_i,\hat{f}_i}\}_{n=1}^{\infty}$ that "converge" to $w(x) \cdot \frac{|f_i(x)-\hat{f}_i(x)|^p}{(f_i(x)-\hat{f}_i(x))^2}$, where the last is defined to be ∞ in places where $f_i(x) = \hat{f}_i(x)$.

But then, solving for $\underset{\hat{f}}{\operatorname{argmin}} \int_{I_i} w_{f_i,\hat{f}_i}(x)\,|f(x)-\hat{f}(x)|^p dx$ might be impossible, or involve very high numbers.
By limiting the values w can get, the problem becomes feasible.

(v) <u>Input:</u> (1) f - the function to approximate.
(2) w - the weight function.
(3) ε - small fixed number.
(4) stoping criteria.

<u>The algorithm:</u>

<u>Init:</u> - $\hat{f}_i \leftarrow 1$ (some random initialization)
<u>loop:</u> until reaching some stoping criteria, do:

$$-w' \leftarrow \min\left\{\frac{1}{\varepsilon}, w(x)\cdot\frac{|f_i(x)-\hat{f}_i(x)|^p}{(f_i(x)-\hat{f}_i(x))^2}\right\}$$

$$-\hat{f}_i^{next} = \frac{1}{\int_{I_i} w'(x)dx}\cdot\int_{I_i} f_i(x)\cdot w'(x)dx$$

The solution to the L^2 problem (see part (b) of the question.

$$-\hat{f}_i \leftarrow \hat{f}_i^{next}$$

<u>Output:</u> $\hat{f}_i$

(f) To solve the L^p problem using L^2 optimization we will use the algorithm from part (e)(v), denoted here as $Alg^{(i)}$ for interval I_i:

<u>Input:</u> (1) f - the function to approximate.
(2) w - the weight function.

(3) ε - small fixed number.

(4) stoping criteria.

<u>The algorithm:</u>

For $i = 1, \dots, N$:

$$\hat{f}_i = Alg^{(i)}(f, w, \varepsilon, \text{stoping criteria})$$

<u>Output:</u> $f = \sum_{i=1}^{N} \hat{f}_i \cdot \mathbb{1}_{I_i}$

(9) IRLS - Iteratively Reweighted Least Square

2- Haar matrix and Walsh-Hadamard matrix

ⓐ (i) We need to show that $\langle u_i, u_j \rangle = \delta_{ij}$ where $1 \le i,j \le 4$ and u_1, u_2, u_3, u_4 are the columns of H_4. So:

$$\langle u_1, u_2 \rangle = \left\langle \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} \right\rangle = 0 \qquad \langle u_2, u_3 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \begin{bmatrix}\sqrt{2}\\-\sqrt{2}\\0\\0\end{bmatrix} \right\rangle = 0$$

$$\langle u_1, u_3 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}\sqrt{2}\\-\sqrt{2}\\0\\0\end{bmatrix} \right\rangle = 0 \qquad \langle u_2, u_4 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \begin{bmatrix}0\\0\\\sqrt{2}\\-\sqrt{2}\end{bmatrix} \right\rangle = 0$$

$$\langle u_1, u_4 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\0\\\sqrt{2}\\-\sqrt{2}\end{bmatrix} \right\rangle = 0 \qquad \langle u_3, u_4 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}\sqrt{2}\\-\sqrt{2}\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\\sqrt{2}\\-\sqrt{2}\end{bmatrix} \right\rangle = 0$$

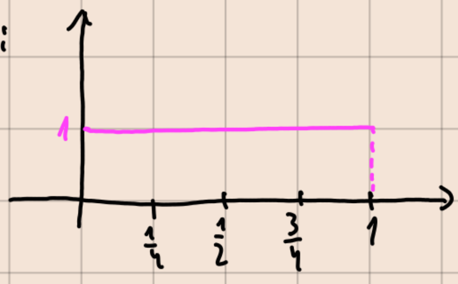
$$\langle u_1, u_1 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \right\rangle = 1 \qquad \langle u_3, u_3 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}\sqrt{2}\\-\sqrt{2}\\0\\0\end{bmatrix}, \begin{bmatrix}\sqrt{2}\\-\sqrt{2}\\0\\0\end{bmatrix} \right\rangle = 1$$

$$\langle u_2, u_2 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} \right\rangle = 1 \qquad \langle u_4, u_4 \rangle = \frac{1}{4}\left\langle \begin{bmatrix}0\\0\\\sqrt{2}\\-\sqrt{2}\end{bmatrix}, \begin{bmatrix}0\\0\\\sqrt{2}\\-\sqrt{2}\end{bmatrix} \right\rangle = 1$$

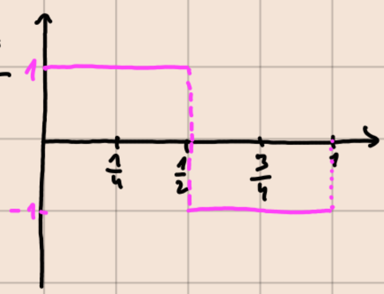
(ii)
$$\begin{bmatrix}\psi_1^H\\\psi_2^H\\\psi_3^H\\\psi_4^H\end{bmatrix} = H_4^T \cdot \begin{bmatrix}\psi_1^S\\\psi_2^S\\\psi_3^S\\\psi_4^S\end{bmatrix} = \cancel{\frac{1}{2}}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ \sqrt{2} & -\sqrt{2} & 0 & 0\\ 0 & 0 & \sqrt{2} & -\sqrt{2}\end{bmatrix}\begin{bmatrix}\cancel{\sqrt{4}} \cdot \mathbb{1}_{[0,\frac{1}{4})}\\ \cancel{\sqrt{4}} \cdot \mathbb{1}_{[\frac{1}{4},\frac{1}{2})}\\ \cancel{\sqrt{4}} \cdot \mathbb{1}_{[\frac{1}{2},\frac{3}{4})}\\ \cancel{\sqrt{4}} \cdot \mathbb{1} \cdot [\frac{3}{4},1]\end{bmatrix} = \begin{bmatrix}\mathbb{1}_{[0,1]}\\ \mathbb{1}_{[0,\frac{1}{2})} - \mathbb{1}_{[\frac{1}{2},1]}\\ \sqrt{2}\cdot\mathbb{1}_{[0,\frac{1}{4}]} - \sqrt{2}\cdot\mathbb{1}_{[\frac{1}{4},\frac{1}{2})}\\ \sqrt{2}\cdot\mathbb{1}_{[\frac{1}{2},\frac{3}{4})} - \sqrt{2}\cdot\mathbb{1}_{[\frac{3}{4},1]}\end{bmatrix}$$

graphs

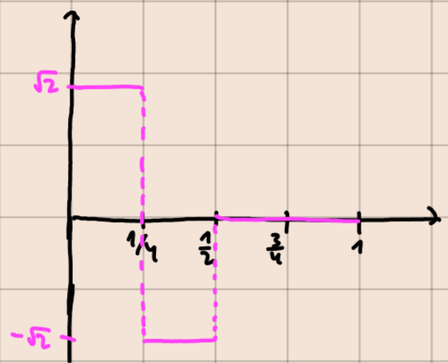
ψ_1^H:



ψ_2^H:



ψ_3^H:



ψ_4^H:

(iii) The approximation of ϕ using this Haar basis is given by $\hat{\phi}_H(t) = \sum_{i=1}^{4} \langle \phi, \psi_i^H \rangle \cdot \psi_i^H(t)$.

So, let us start by calculating $\langle \phi, \psi_i^H \rangle$ for $i=1,2,3,4$. To do so, we will start by calculating $\langle \phi, \mathbb{1}_{[\frac{j}{4}, \frac{j+1}{4})} \rangle$ for $j=0,1,2,3$, and then use additivity and linearity of the integral.

אין צורך ב-* כי הפונ' ממשי

$$\langle \phi, \mathbb{1}_{[\frac{j}{4}, \frac{j+1}{4})} \rangle = \int_0^1 \phi(t) \cdot \mathbb{1}_{[\frac{j}{4}, \frac{j+1}{4})}(t)\,dt = \int_{j/4}^{\frac{j+1}{4}} (a + b\cos(2\pi t) + c \cdot \cos^2(\pi t))\,dt$$

$$= \int_{\frac{j}{4}}^{\frac{j+1}{4}} \left(a + b\cos(2\pi t) + \frac{c}{2}\cos(2\pi t) + \frac{c}{2}\right) dt$$

$$= \frac{a + \frac{c}{2}}{4} + \frac{b + \frac{c}{2}}{2\pi} \cdot \sin(2\pi t) \Big|_{\frac{j}{4}}^{\frac{j+1}{4}}$$

$\underline{j=0}$: $\langle \phi, \mathbb{1}_{[0, \frac{1}{4})} \rangle = \frac{a}{4} + \frac{b}{2\pi} + \frac{c}{8} + \frac{c}{4\pi}$

$\underline{j=1}$: $\langle \phi, \mathbb{1}_{[\frac{1}{4}, \frac{1}{2})} \rangle = \frac{a}{4} - \frac{b}{2\pi} + \frac{c}{8} - \frac{c}{4\pi}$

$\underline{j=2}$: $\langle \phi, \mathbb{1}_{[\frac{1}{2}, \frac{3}{4})} \rangle = \frac{a}{4} - \frac{b}{2\pi} + \frac{c}{8} - \frac{c}{4\pi}$

$\underline{j=3}$: $\langle \phi, \mathbb{1}_{[\frac{3}{4}, 1)} \rangle = \frac{a}{4} + \frac{b}{2\pi} + \frac{c}{8} + \frac{c}{4\pi}$

$$\langle \phi, \psi_1^H \rangle = \sum_{j=0}^{3} \langle \phi, \mathbb{1}_{[\frac{j}{4}, \frac{j+1}{4})} \rangle = a + \frac{c}{2}$$

$$\langle \phi, \psi_2^H \rangle = \langle \phi, \mathbb{1}_{[0,\frac{1}{4})} \rangle + \langle \phi, \mathbb{1}_{[\frac{1}{4}, \frac{1}{2})} \rangle - \langle \phi, \mathbb{1}_{[\frac{1}{2}, \frac{3}{4})} \rangle - \langle \phi, \mathbb{1}_{[\frac{3}{4}, 1)} \rangle = 0$$

$$\langle \phi, \psi_3^H \rangle = \frac{\sqrt{2}\,b}{\pi} + \frac{\sqrt{2}\,c}{2\pi}$$

$$\langle \phi, \psi_4^H \rangle = -\frac{\sqrt{2}\,b}{\pi} - \frac{\sqrt{2}\,c}{2\pi}$$

$$\Rightarrow \boxed{\hat{\phi}_H^{(4)}(t) = \left(a + \frac{c}{2}\right) \cdot \mathbb{1}_{[0, \frac{1}{4})} + \left(\frac{\sqrt{2}\,b}{\pi} + \frac{\sqrt{2}\,c}{2\pi}\right)\left(\mathbb{1}_{[\frac{1}{2}, \frac{3}{4})} - \mathbb{1}_{[\frac{3}{4}, 1]}\right)}$$

The associated MSE is given by:

$$\Psi^H_{MSE}(\Phi(t) - \hat{\Phi}^{(4)}_H(t)) = \int_0^1 \phi^2(t)\,dt - \sum_{i=1}^{4} \langle \phi, \psi_i^H \rangle^2 .$$

So:

$$\int_0^1 \phi^2(t)\,dt = \int_0^1 (a + b\cos(2\pi t) + c\cdot\cos^2(\pi t))^2\,dt$$

$$= \int_0^1 (a^2 + 2ab\underbrace{\cancel{\cos(2\pi t)}}_{0} + 2ac\cdot\cos^2(\pi t) + b^2\cos^2(2\pi t) + c^2\cos^4(\pi t) + 2bc\cos(2\pi t)\cdot\cos^2(\pi t)\,dt$$

$$= a^2 + \int_0^1 ac\cdot(\underbrace{\cancel{\cos 2\pi t}}_{0} + 1) + \frac{b^2}{2}(\underbrace{\cancel{\cos(4\pi t)}}_{0} + 1) + \frac{c^2}{4}((\cos(2\pi t)+1)^2) + 2bc\cos(2\pi t)\cdot\cos^2(\pi t)\,dt$$

$$= a^2 + ac + \frac{b^2}{2} + \int_0^1 \frac{c^2}{4}\cdot(\cos^2(2\pi t) + 2\underbrace{\cancel{\cos(2\pi t)}}_{0} + 1) + bc\cos(2\pi t)\cdot(\cos(2\pi t)+1)\;dt$$

$$= a^2 + ac + \frac{b^2}{2} + \int_0^1 \frac{c^2}{4}\cdot(\underbrace{\cancel{\frac{\cos 4\pi t}{2}}}_{0} + \frac{1}{2} + 1) + bc(\cos^2(2\pi t) + \underbrace{\cancel{\cos 2\pi t}}_{0})\,dt$$

$$= a^2 + ac + \frac{b^2}{2} + \frac{3c^2}{8} + \int_0^1 bc(\underbrace{\cancel{\frac{\cos(4\pi t)}{2}}}_{0} + \frac{1}{2})\,dt$$

$$= a^2 + ac + \frac{b^2}{2} + \frac{3c^2}{8} + \frac{bc}{2}$$

$$\boxed{\begin{aligned}\Rightarrow \Psi^H_{MSE}(\Phi(t) - \Phi^{(4)}_H(t)) &= a^2 + ac + \frac{b^2}{2} + \frac{3c^2}{8} + \frac{bc}{2} - (a + \frac{c}{2})^2 - 2(\frac{\sqrt{2}b}{\pi} + \frac{\sqrt{2}c}{2\pi})^2 \\ &= \cancel{a^2} + \cancel{ac} + \frac{b^2}{2} + \frac{3c^2}{8} + \frac{bc}{2} - \cancel{a^2} - \cancel{ac} - \frac{c^2}{4} - \frac{4}{\pi^2}(b^2 + bc + \frac{c^2}{4}) \\ &= \frac{b^2}{2} + \frac{c^2}{8} + \frac{bc}{2} - \frac{4b^2}{\pi^2} + \frac{4bc}{\pi^2} - \frac{c^2}{\pi^2} \\ &= (\frac{1}{2} - \frac{4}{\pi^2})b^2 + (\frac{1}{2} + \frac{4}{\pi^2})bc + (\frac{1}{8} - \frac{1}{\pi^2})c^2\end{aligned}}$$

(iv) Assume $a \geq b \geq 0$ and $c \geq 0$.

From the expression of the MSE we can see that we need to sort $\{|\langle \phi, \psi_i^H \rangle|\}_{i=1}^{4}$ in a non-increasing order:

$1 > \frac{\sqrt{2}}{\pi}$ and $\begin{cases} a \geq b \geq 0 \\ c \geq 0 \end{cases}$ so $a + \frac{c}{2} > \frac{\sqrt{2}}{\pi} b + \frac{1}{2}\cdot\frac{\sqrt{2}}{\pi} c$.

Then:

<u>1-term approx.</u>: $\hat{\Phi}^{(1)}_H(t) = (a + \frac{c}{2})\cdot \mathbb{1}_{[0,1]}$

<u>2-term approx.</u>: $\hat{\Phi}^{(2)}_H(t) = (a + \frac{c}{2})\cdot \mathbb{1}_{[0,1]} + (\frac{2}{\pi} b + \frac{2}{2\pi} c)\cdot(\mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{1}{2})})$

<u>3-term approx.:</u>

$$\hat{\Phi}^{(3)}_H(t) = \left(a + \frac{c}{2}\right)\cdot \mathbb{1}_{[0,1]} + \left(\frac{2}{\pi}b + \frac{2}{2\pi}c\right)\cdot\left(\mathbb{1}_{[0,\frac14)} - \mathbb{1}_{[\frac14,\frac12)} - \mathbb{1}_{[\frac12,\frac34)} + \mathbb{1}_{[\frac34,1]}\right)$$

<u>4-term approx.:</u> the same as the 3-term approx., as $\langle \Phi, \Psi^2_H\rangle = 0$.

(v) Now assume that $a = \frac{1}{\pi}$, $b = 1$, $c = \frac{3}{2}$. We can compute $\langle \Phi, \Psi^H_i\rangle$ for $i = 1, \dots, 4$ exactly:

$$|\langle \Phi, \Psi^H_1\rangle| = \left|\frac{1}{\pi} + \frac{1}{2}\cdot\frac{3}{2}\right| = \frac{1}{\pi} + \frac{3}{4}$$

$$|\langle \Phi, \Psi^H_2\rangle| = 0$$

$$|\langle \Phi, \Psi^H_3\rangle| = |\langle \Phi, \Psi^H_4\rangle| = \frac{\sqrt2}{\pi}\cdot 1 + \frac{\sqrt2}{2\pi}\cdot\frac{3}{2} = \frac{\sqrt2}{\pi}\left(1 + \frac{3}{4}\right)$$

Comparing $|\langle \Phi, \Psi^H_1\rangle|$ and $|\langle \Phi, \Psi^H_3\rangle|$:

$$\left(\frac{1}{\pi} + \frac{3}{4}\right) - \left(\frac{\sqrt2}{\pi} + \frac{3\sqrt2}{4\pi}\right) = \frac{4 + 3\pi - 4\sqrt2 - 3\sqrt2}{4\pi} = \frac{4 + 3\pi - 7\sqrt2}{4\pi} = \circledast$$

and $7\sqrt2 < 7\cdot 1.5 = 11.5$, $4 + 3\pi > 4 + 3\cdot 3 = 13$

$\Rightarrow \circledast < 0$

$\Rightarrow |\langle \Phi, \Psi^H_3\rangle| > |\langle \Phi, \Psi^H_1\rangle|$.

<u>So:</u>

<u>1-term approx.:</u> $\hat{\Phi}^{(1)}_H(t) = \left(\frac{2b}{\pi} + \frac{2c}{2\pi}\right)\cdot\left(\mathbb{1}_{[0,\frac14)} - \mathbb{1}_{[\frac14,\frac12)}\right)$

<u>2-term approx.:</u> $\hat{\Phi}^{(1)}_H(t) = \left(\frac{2b}{\pi} + \frac{2c}{2\pi}\right)\cdot\left(\mathbb{1}_{[0,\frac14)} - \mathbb{1}_{[\frac14,\frac12)} - \mathbb{1}_{[\frac12,\frac34)} + \mathbb{1}_{[\frac34,1]}\right)$

<u>3-term approx.:</u>

$$\hat{\Phi}^{(3)}_H(t) = \left(a + \frac{c}{2}\right)\cdot \mathbb{1}_{[0,1]} + \left(\frac{2}{\pi}b + \frac{2}{2\pi}c\right)\cdot\left(\mathbb{1}_{[0,\frac14)} - \mathbb{1}_{[\frac14,\frac12)} - \mathbb{1}_{[\frac12,\frac34)} + \mathbb{1}_{[\frac34,1]}\right)$$

<u>4-term approx.:</u> the same as the 3-term approx., as $\langle \Phi, \Psi^2_H\rangle = 0$.

b (i) $<\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}> = \frac{1}{4}\cdot(1+1-1-1) = 0$, $<\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}> = 0$

$<\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}> = 0$, $<\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}> = 0$

$<\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}> = 0$, $<\frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}> = 0$

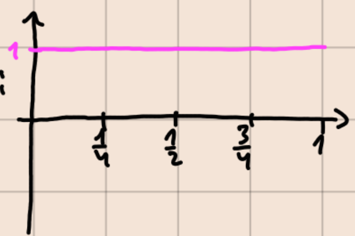
$<\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}> = \frac{1}{4}\cdot 4 = 1$, $<\frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}> = 1$

$<\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}> = \frac{1}{4}\cdot 4 = 1$, $<\frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}> = 1$

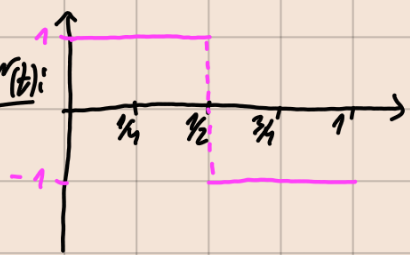
$\Rightarrow\ <\mathrm{col}_i(W_4), \mathrm{col}_j(W_4)> = \delta_{ij}$ for $(i,j) \in \{1,2,3,4\}^2$, and therfore W_4 is unitary.

(ii) $$\begin{bmatrix}\psi_1^W(t)\\ \psi_2^W(t)\\ \psi_3^W(t)\\ \psi_4^W(t)\end{bmatrix} = W_4 \cdot \begin{bmatrix}\psi_1^S(t)\\ \psi_2^S(t)\\ \psi_3^S(t)\\ \psi_4^S(t)\end{bmatrix} = \frac{1}{\cancel{2}}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & -1 & 1\\ 1 & -1 & 1 & -1\end{bmatrix} \cdot \begin{bmatrix}\cancel{\sqrt{4}} \cdot \mathbb{1}_{[0,\frac{1}{4})}\\ \cancel{\sqrt{4}} \cdot \mathbb{1}_{[\frac{1}{4},\frac{1}{2})}\\ \cancel{\sqrt{4}} \cdot \mathbb{1}_{[\frac{1}{2},\frac{3}{4})}\\ \cancel{\sqrt{4}} \cdot \mathbb{1}_{[\frac{3}{4},1]}\end{bmatrix} = \begin{bmatrix}\mathbb{1}_{[0,1]}\\ \mathbb{1}_{[0,\frac{1}{2})} - \mathbb{1}_{[\frac{1}{2},1]}\\ \mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{3}{4})} + \mathbb{1}_{[\frac{3}{4},1]}\\ \mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{1}{2})} + \mathbb{1}_{[\frac{1}{2},\frac{3}{4})} - \mathbb{1}_{[\frac{3}{4},1]}\end{bmatrix}$$

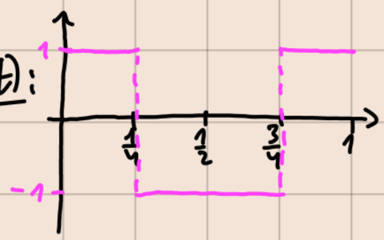
$\psi_1^W(t)$:



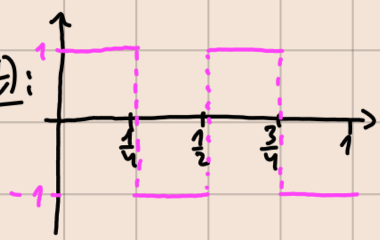
$\psi_2^W(t)$:



$\psi_1^W(t)$:



$\psi_1^W(t)$:



(iii) From ⓐ(iii) we already know $<\phi, \mathbb{1}_{[\frac{j}{4},\frac{j+1}{4})}>$ for $j = 0,1,2,3$.
We will now use this to compute $<\phi, \psi_i^W>$ for $i = 1,2,3,4$.
$<\phi, \psi_1^W> = <\phi, \psi_1^H> = a + \frac{c}{2}$

$\langle \phi, \psi_2^W \rangle = \langle \phi, \psi_2^H \rangle = 0$

$\langle \phi, \psi_3^W \rangle = \langle \phi, \mathbb{1}_{[0,\frac{1}{4})} \rangle - \langle \phi, \mathbb{1}_{[\frac{1}{4},\frac{1}{2})} \rangle - \langle \phi, \mathbb{1}_{[\frac{1}{2},\frac{3}{4})} \rangle + \langle \phi, \mathbb{1}_{[\frac{3}{4},1]} \rangle$

$= \frac{a}{4} + \frac{b}{2\pi} + \frac{c}{8} + \frac{c}{4\pi} - \frac{a}{4} + \frac{b}{2\pi} - \frac{c}{8} + \frac{c}{4\pi} - \frac{a}{4} + \frac{b}{2\pi} - \frac{c}{8} + \frac{c}{4\pi} + \frac{a}{4} + \frac{b}{2\pi} + \frac{c}{8} + \frac{c}{4\pi}$

$= \frac{2b+c}{\pi}$

$\langle \phi, \psi_4^W \rangle = \langle \phi, \mathbb{1}_{[0,\frac{1}{4})} \rangle - \langle \phi, \mathbb{1}_{[\frac{1}{4},\frac{1}{2})} \rangle + \langle \phi, \mathbb{1}_{[\frac{1}{2},\frac{3}{4})} \rangle - \langle \phi, \mathbb{1}_{[\frac{3}{4},1]} \rangle$

$\frac{a}{4} + \frac{b}{2\pi} + \frac{c}{8} + \frac{c}{4\pi} - \frac{a}{4} + \frac{b}{2\pi} - \frac{c}{8} + \frac{c}{4\pi} + \frac{a}{4} - \frac{b}{2\pi} + \frac{c}{8} - \frac{c}{4\pi} - \frac{a}{4} - \frac{b}{2\pi} - \frac{c}{8} - \frac{c}{4\pi}$

$= 0$

$$\boxed{\Rightarrow \hat{\phi}_W^{(4)} = \sum_{i=1}^{4} \langle \phi, \psi_i^W \rangle \cdot \psi_i^W(t) = \left(a + \frac{c}{2}\right) \cdot \mathbb{1}_{[0,1]} + \frac{2b+c}{\pi} \left(\mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{3}{4})} + \mathbb{1}_{[\frac{3}{4},1]} \right)}$$

$$\psi_{MSE}^W\left(\phi(t) - \phi_W^{(4)}(t)\right) = \int_0^1 \phi^2(t)\,dt - \sum_{i=1}^{4} \langle \phi, \psi_i^W \rangle^2$$

$$= a^2 + ac + \frac{b^2}{2} + \frac{3c^2}{8} + \frac{bc}{2} - \left(a + \frac{c}{2}\right)^2 - \left(\frac{2b+c}{\pi}\right)^2$$

$$= \cancel{a^2} + \cancel{a}c + \frac{b^2}{2} + \frac{3c^2}{8} + \frac{bc}{2} - \cancel{a^2} - \frac{c^2}{4} - \cancel{a}c - \frac{4b^2}{\pi^2} - \frac{c^2}{\pi^2} - \frac{4bc}{\pi^2}$$

$$= \left(\frac{1}{2} - \frac{4}{\pi^2}\right) b^2 + \left(\frac{1}{8} - \frac{1}{\pi^2}\right) c^2 + \left(\frac{1}{2} - \frac{4}{\pi^2}\right) bc$$

(iv) Assume $a \geq b \geq 0$, $c \geq 0$. As in (a)(iv), we will sort $\{\langle \phi, \psi_i^W \rangle\}_{i=1}^4$ in a non-increasing order:

$a \geq b \geq 0$ and $1 > \frac{2}{\pi}$ so $a > \frac{2b}{\pi}$. Also, $\frac{1}{2} > \frac{1}{\pi}$ and $c \geq 0$, so $\frac{c}{2} > \frac{c}{\pi}$, so $\langle \phi, \psi_1^W \rangle = a + \frac{c}{2} > \frac{2b+c}{\pi} = \langle \phi, \psi_3^W \rangle$.

Then:

1-term approx.: $\hat{\phi}_W^{(1)}(t) = \left(a + \frac{c}{2}\right) \cdot \mathbb{1}_{[0,1]}$

2-term approx.: $\hat{\phi}_W^{(2)}(t) = \left(a + \frac{c}{2}\right) \cdot \mathbb{1}_{[0,1]} + \frac{2b+c}{\pi} \cdot \left(\mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{3}{4})} + \mathbb{1}_{[\frac{3}{4},1]} \right)$

3-term and 4-term approx. are the same, as
$\langle \phi, \psi_2^w \rangle = \langle \phi, \psi_4^w \rangle = 0$.

(iv) Now $a = \frac{1}{\pi}$, $b = 1$, $c = \frac{3}{2}$.

$\Rightarrow \langle \phi, \psi_1^w \rangle = a + \frac{c}{2} = \frac{1}{\pi} + \frac{3}{4}$

$\langle \phi, \psi_4^w \rangle = \frac{2b+c}{\pi} = \frac{2 + \frac{3}{2}}{\pi} = \frac{2}{\pi} + \frac{3}{2\pi}$

$\langle \phi, \psi_4^w \rangle - \langle \phi, \psi_1^w \rangle = \frac{2}{\pi} + \frac{3}{2\pi} - \frac{1}{\pi} - \frac{3}{4} = \frac{5}{2\pi} - \frac{3}{4} = \frac{10 - 3\pi}{4\pi} > 0$

$\Rightarrow \langle \phi, \psi_4^w \rangle > \langle \phi, \psi_1^w \rangle$

Then:

1-term approx.: $\left(\frac{2}{\pi} + \frac{3}{2\pi}\right) \cdot \left(\mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{3}{4})} + \mathbb{1}_{[\frac{3}{4},1]}\right)$

2-term approx.: $\hat{\phi}_w^{(2)}(t) = \left(a + \frac{c}{2}\right) \cdot \mathbb{1}_{[0,1]} + \frac{2b+c}{\pi} \cdot \left(\mathbb{1}_{[0,\frac{1}{4})} - \mathbb{1}_{[\frac{1}{4},\frac{3}{4})} + \mathbb{1}_{[\frac{3}{4},1]}\right)$

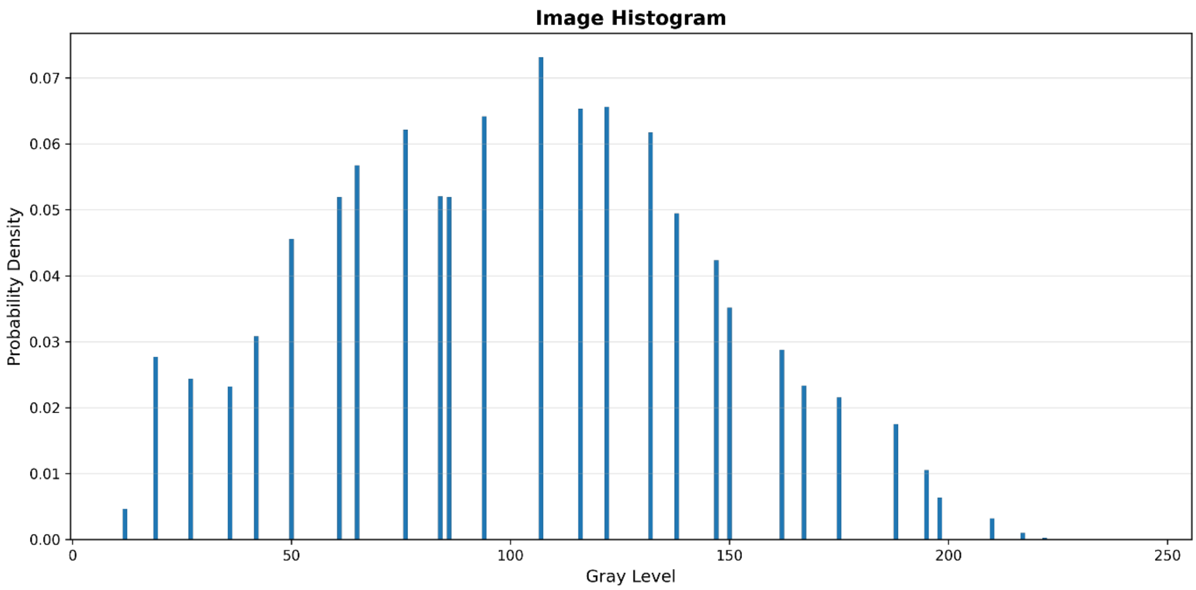
3-term and 4-term approx. are the same, as
$\langle \phi, \psi_2^w \rangle = \langle \phi, \psi_4^w \rangle = 0$.

Implementation part (2)

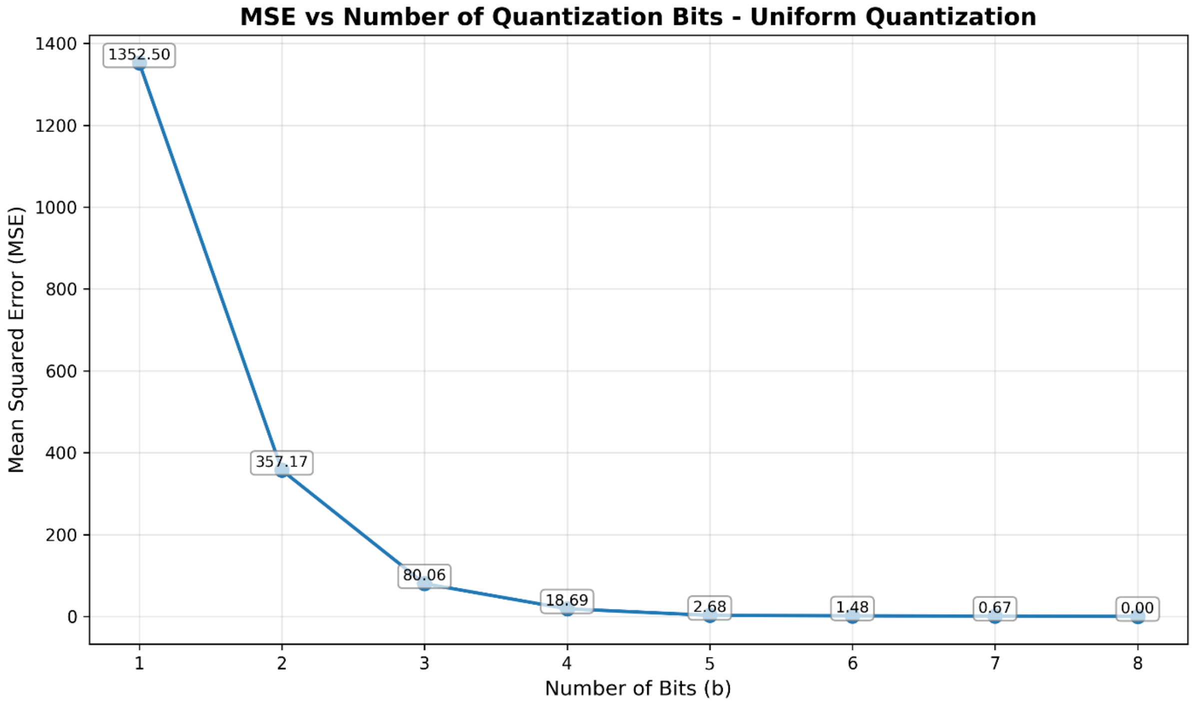
Question 1:

1. The 512x512 gray scale picture we choose is:

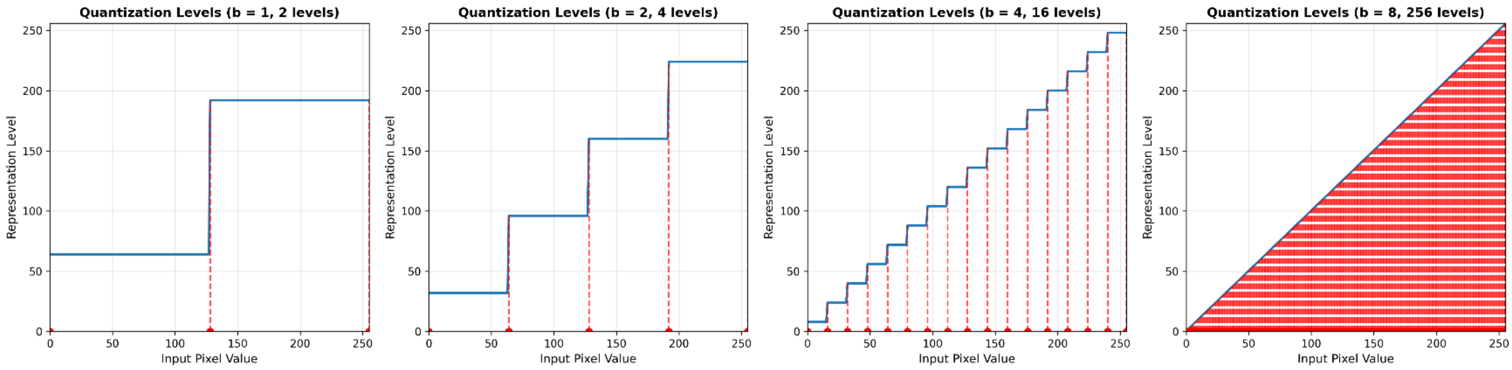
Its PDF of the gray levels is (not uniform):



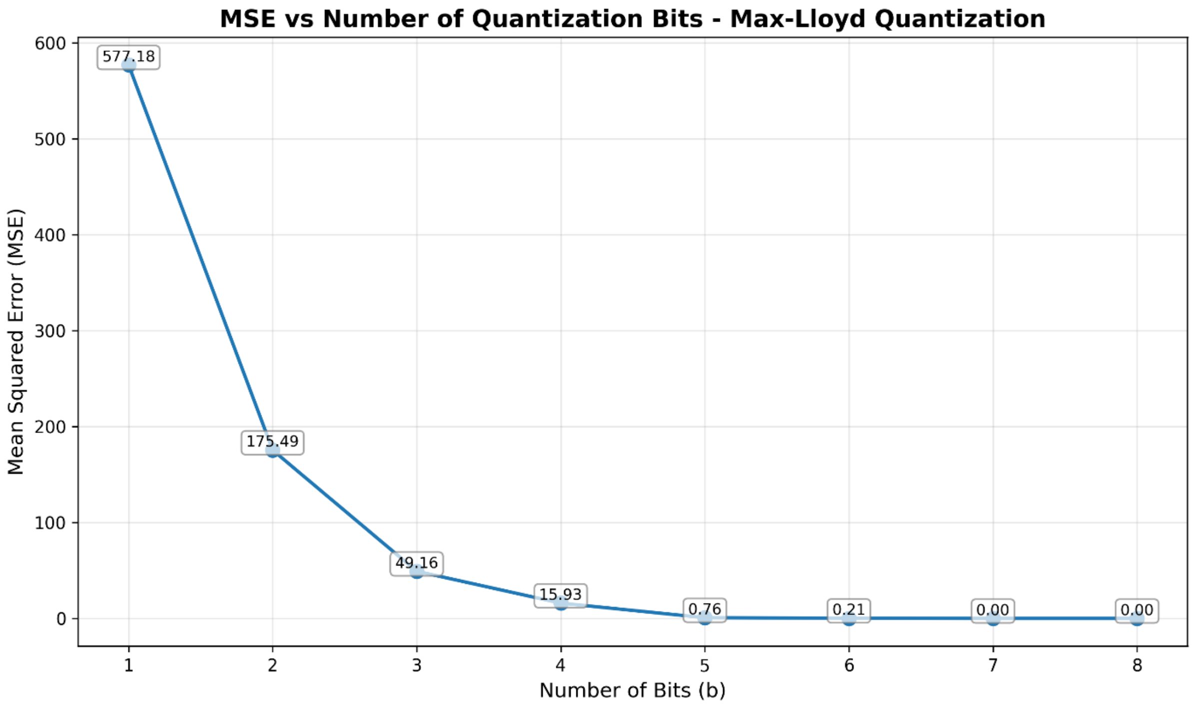
2. We applied uniform quantization on the image using b bits per pixel
    a. This is the MSE as function of bit budget b for $b = 1, \ldots, 8$ (code in quantization.py)



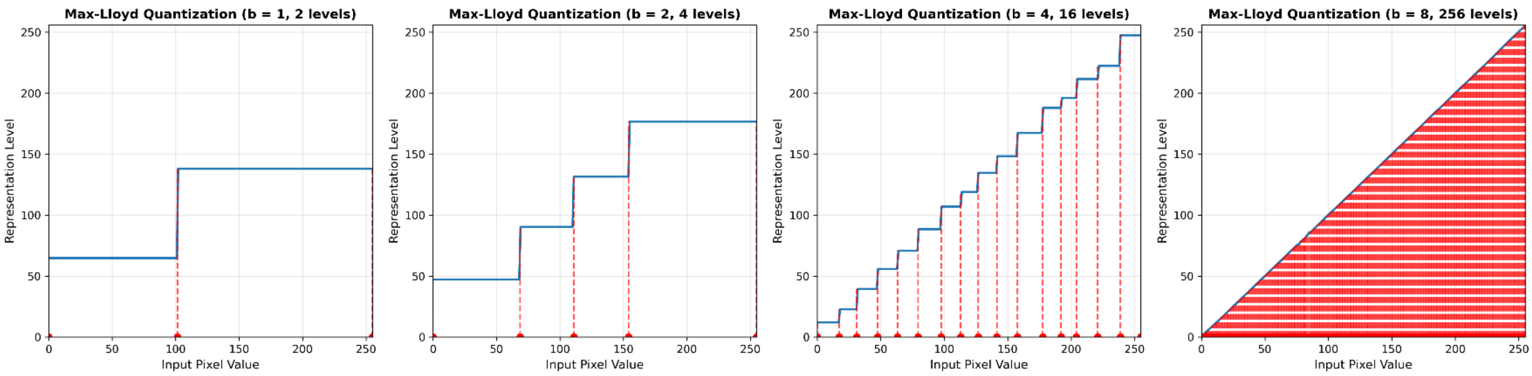
    b. Decision and representation levels for representative b values:



3. The implementation of Max-Loyd algorithm is in quantization.py file.
4. Implementation in quantization.py file
    a. This is the MSE as function of bit budget b for $b = 1, \ldots, 8$ (code in quantization.py)



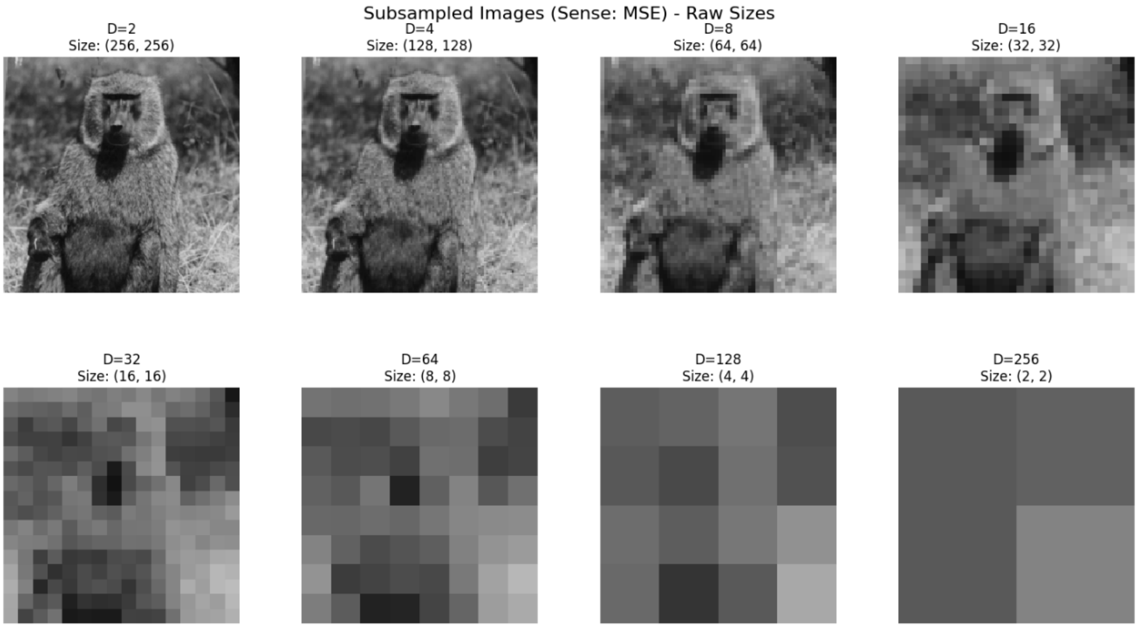
    b. Decision and representation levels for representative b values:



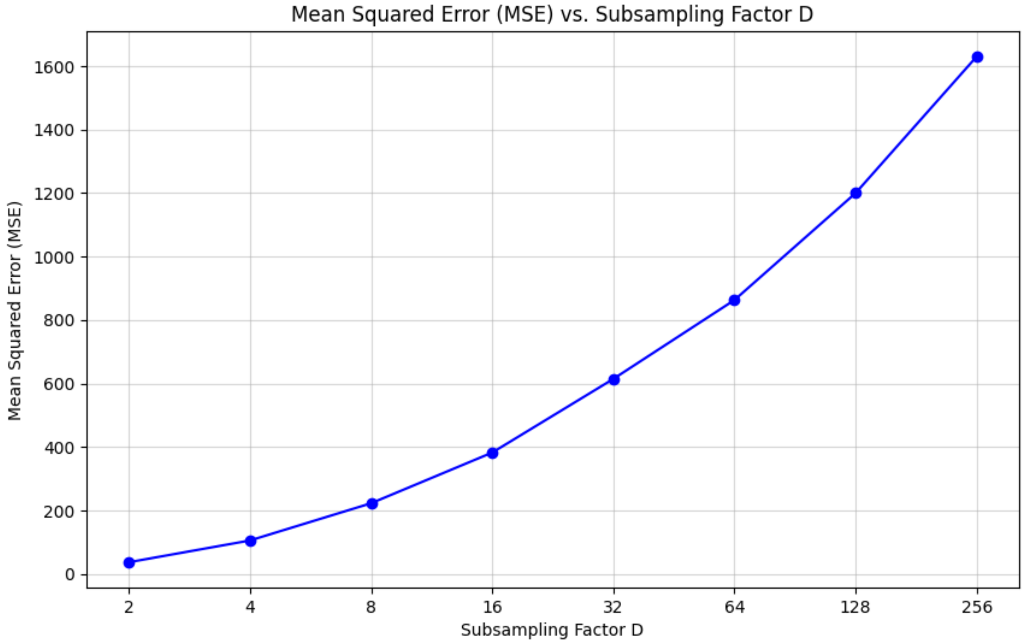
5. We observe that Max-Lloyd quantization consistently yields significantly lower MSE values than uniform quantization across all bit budgets (b) . This is expected, as uniform quantization divides the pixel range into equal intervals without regarding the specific image statistics. In contrast, the Max-Lloyd algorithm explicitly minimizes the MSE by iteratively adapting the decision and representation levels to fit the image's specific probability density function.

Question 2:

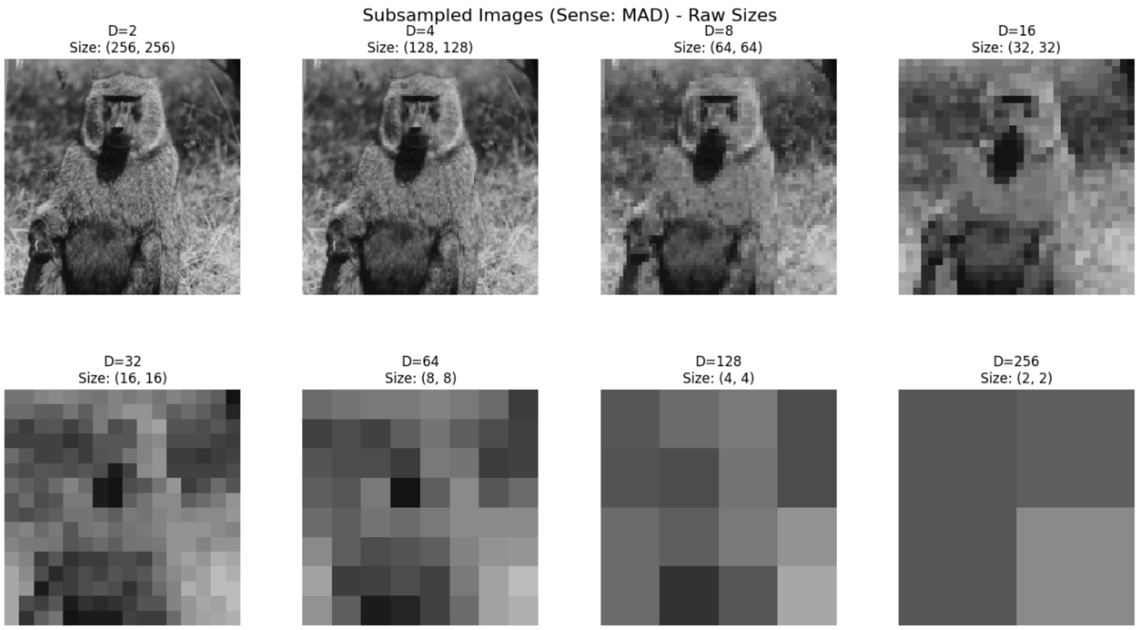
1. As proved in class, in the MSE sense, the optimal number for a sub-sample is the average of the function over the samples in the sub-sample. In the MAD sense, the optimal number for a sub-sample is the median of the function over the samples in the sub-sample.
    a. Sub-sampled images in MSE sense, for different sub-sampling factor, D:



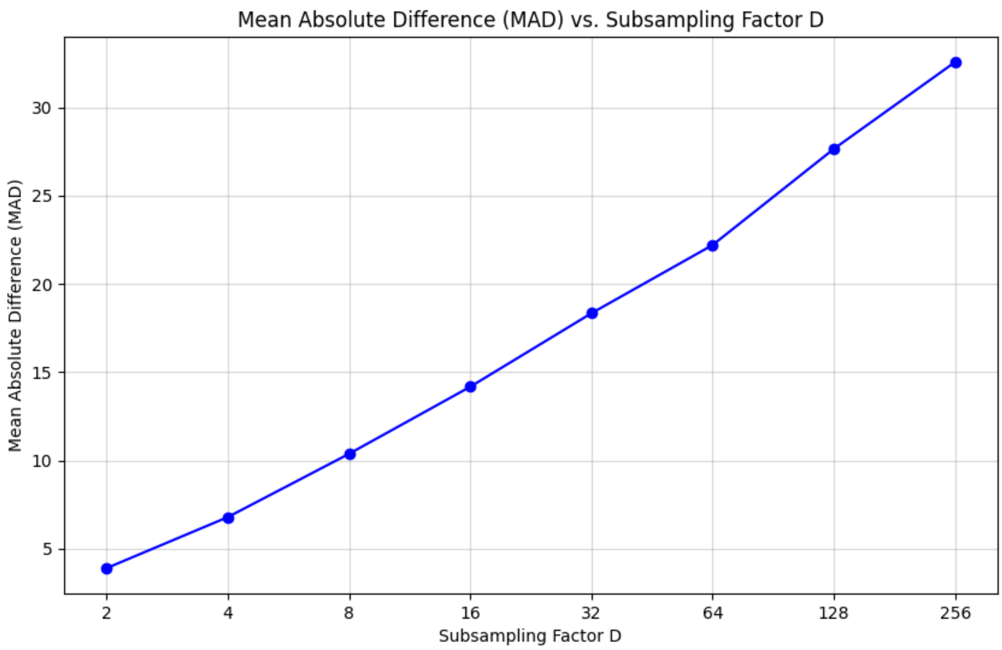
In the graph of MSE as function of Sub-sampling factor D, the horizontal axis (Sub-sampling factor D) is taken in log scale so the different values of D will be evenly spaced.



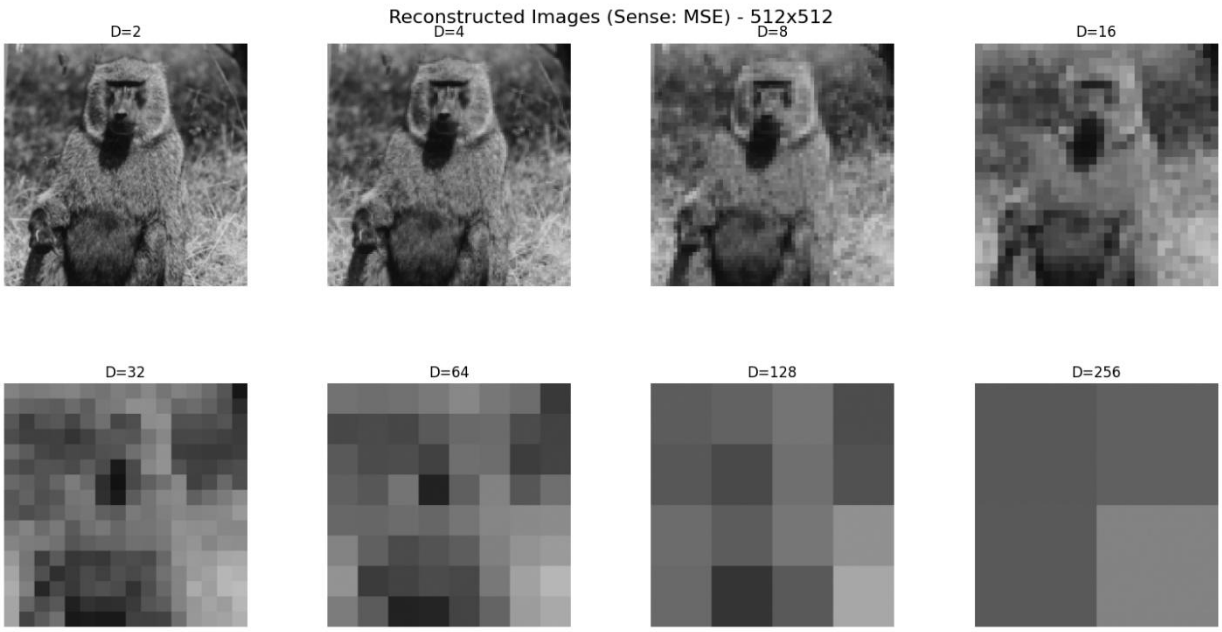
b. Sub-sampled images in MAD sense, for different sub-sampling factor, D:

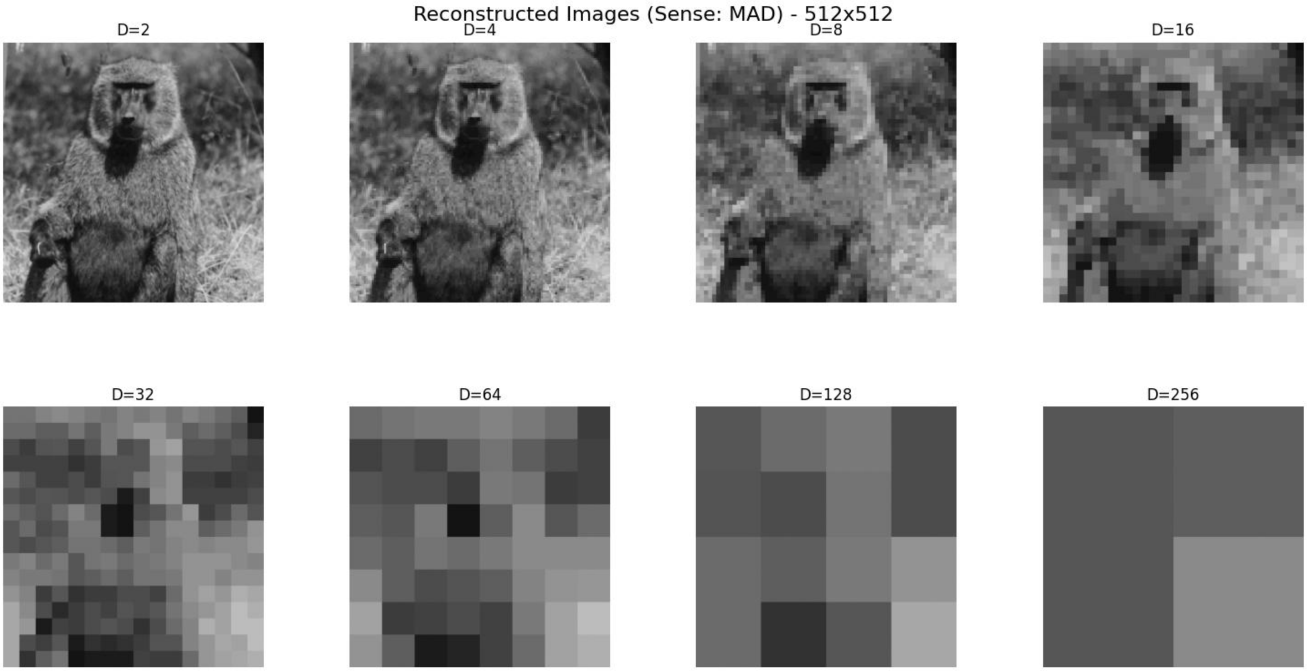


In the graph of MAD as function of Sub-sampling factor D, the horizontal axis (Sub-sampling factor D) is taken in log scale so the different values of D will be evenly spaced.



2. We reconstruct the image by taking each pixel and duplicate it number of times such that we get a 512x512 image (original size). For example, in case D=2, we get a sampled image in size 256x256. In the reconstructed image, we will take each pixel and duplicate it 4 times – meaning each 1x1 pixel becomes 2x2 four pixels, and the result is 512x512 image.



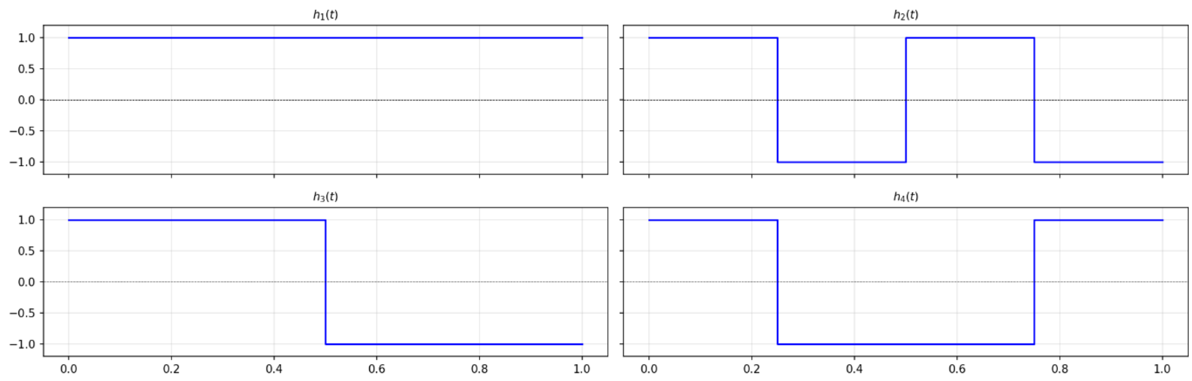


3. As the sub-sampling factor D increases, fewer pixels are retained from the original image, since each $D \times D$ block is represented by a single sample. This leads to a loss of spatial detail and high-frequency information. After reconstruction, the image becomes increasingly blurred, and the reconstruction error increases accordingly, as reflected by higher MSE and MAD values, as we can see in the graph presented in previous sections of this answer.

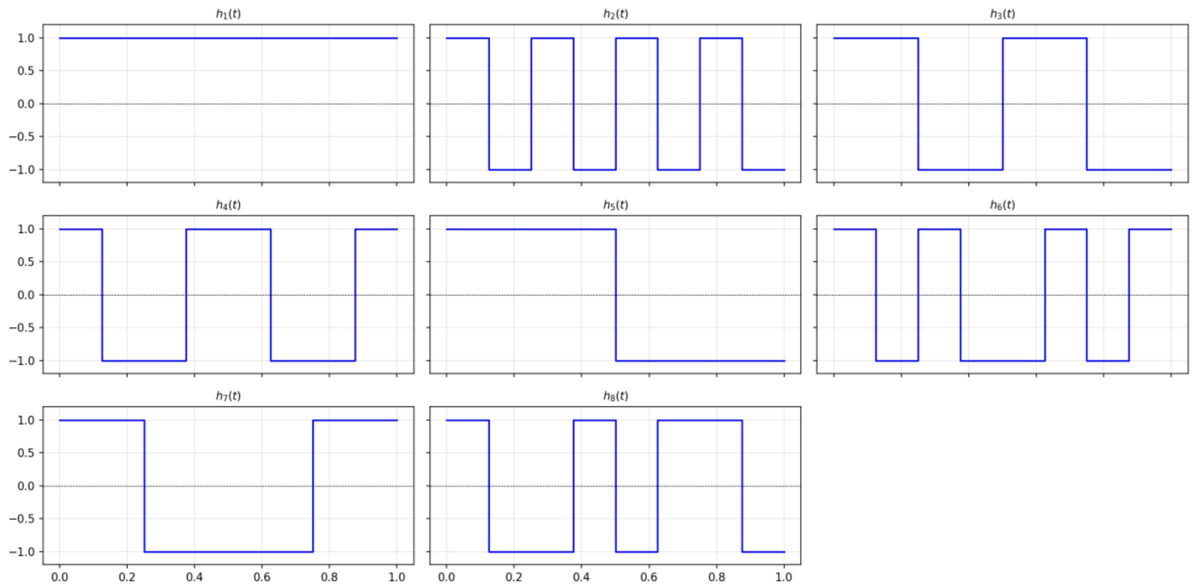
Question 3:

a. Implementation is in python file **matrices.py** by a function named **Hadamard_matrix** which get n as an input and returns the Hadamard matrix of size $2^n \times 2^n$.
b. We did it by taking the columns of the appropriate Hadamard matrix multiplied by normalization factor – square of N.

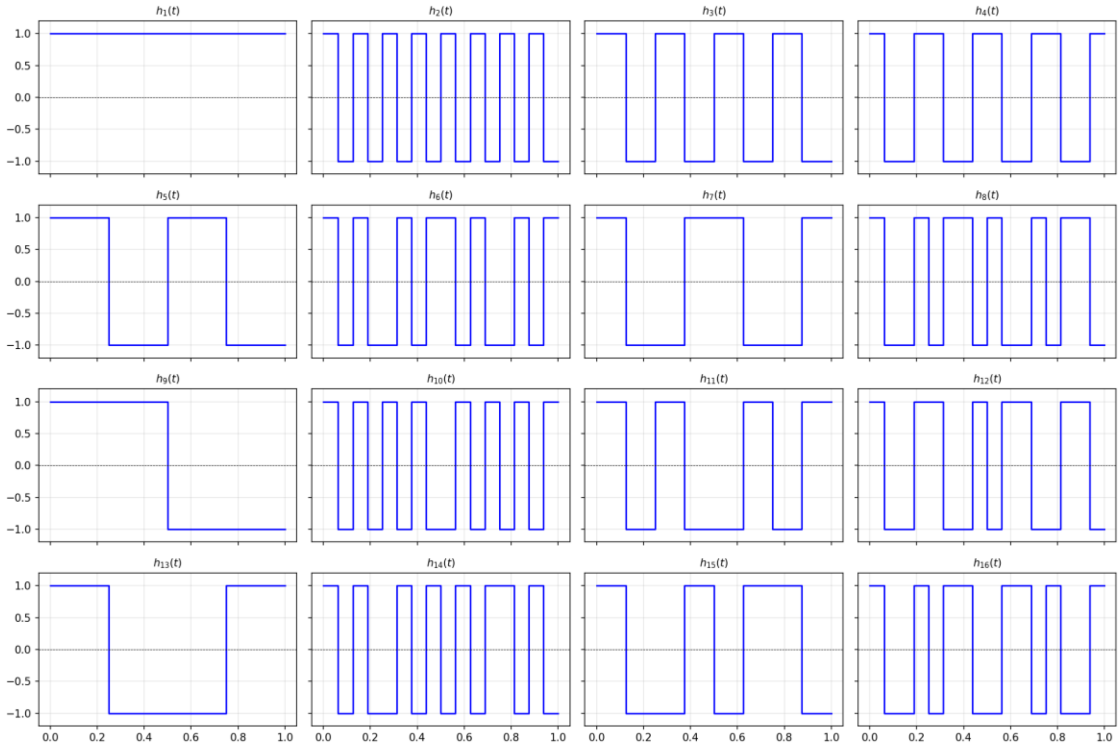
Basis Functions for n=2: Hadamard (Natural Order)



Basis Functions for n=3: Hadamard (Natural Order)

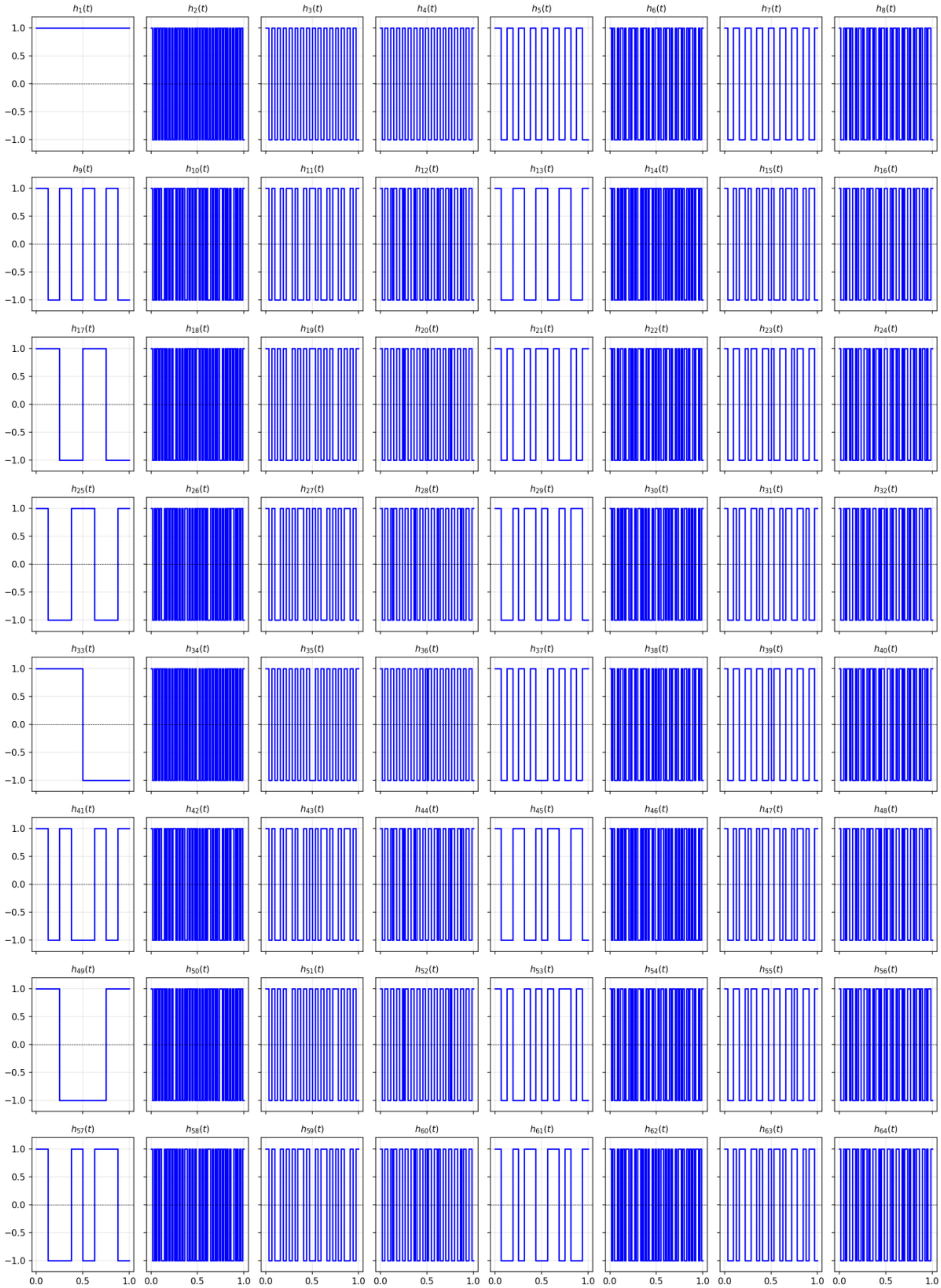


Basis Functions for n=4: Hadamard (Natural Order)



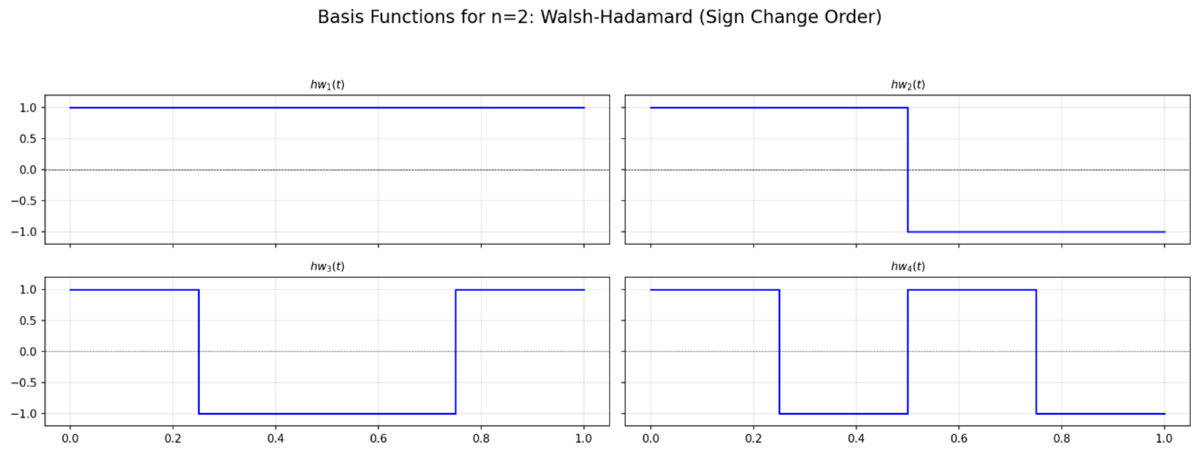
Basis Functions for n=5: Hadamard (Natural Order)

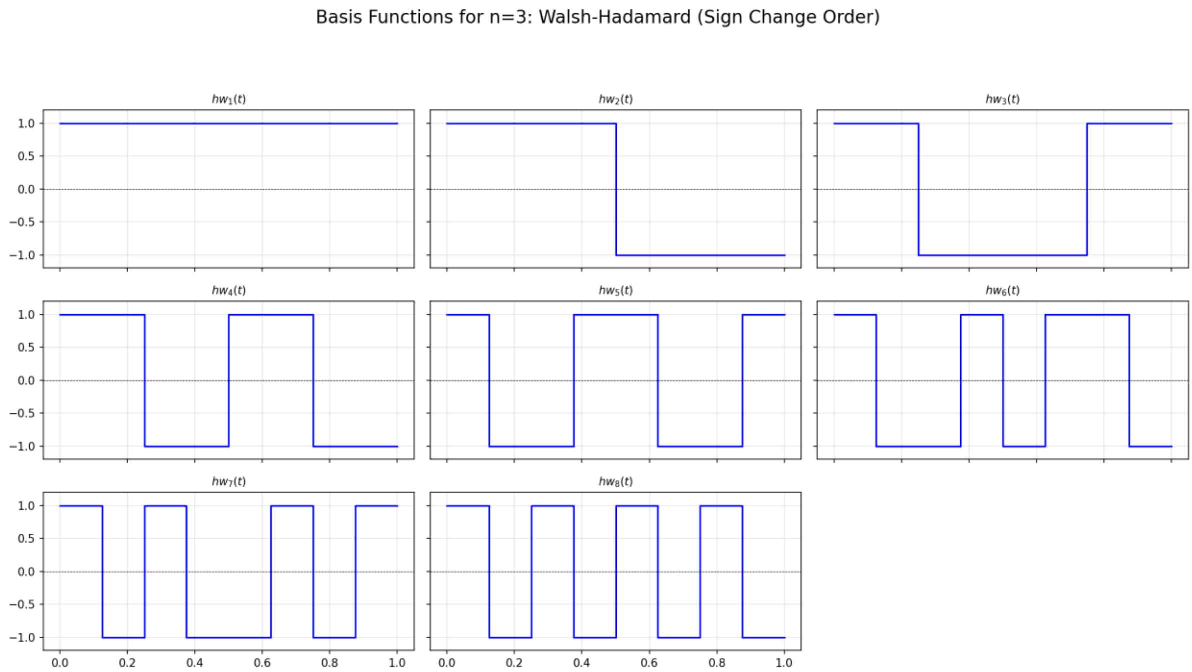
Basis Functions for n=6: Hadamard (Natural Order)



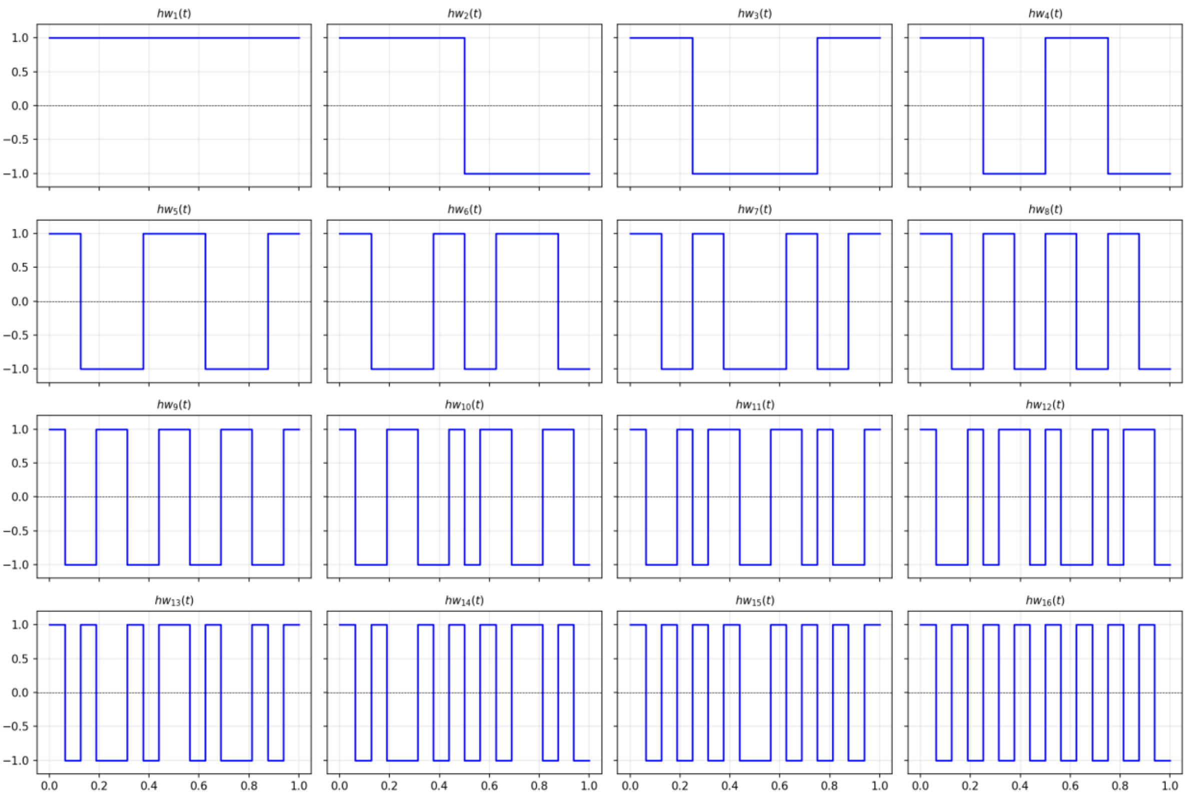
c. For Walsh-Hadamard, we took the Hadamard matrix in the desired size, and sorted its rows based on sign changes – in increasing order.

d. We did it by taking the columns of the appropriate Walsh-Hadamard matrix multiplied by normalization factor – square of N:



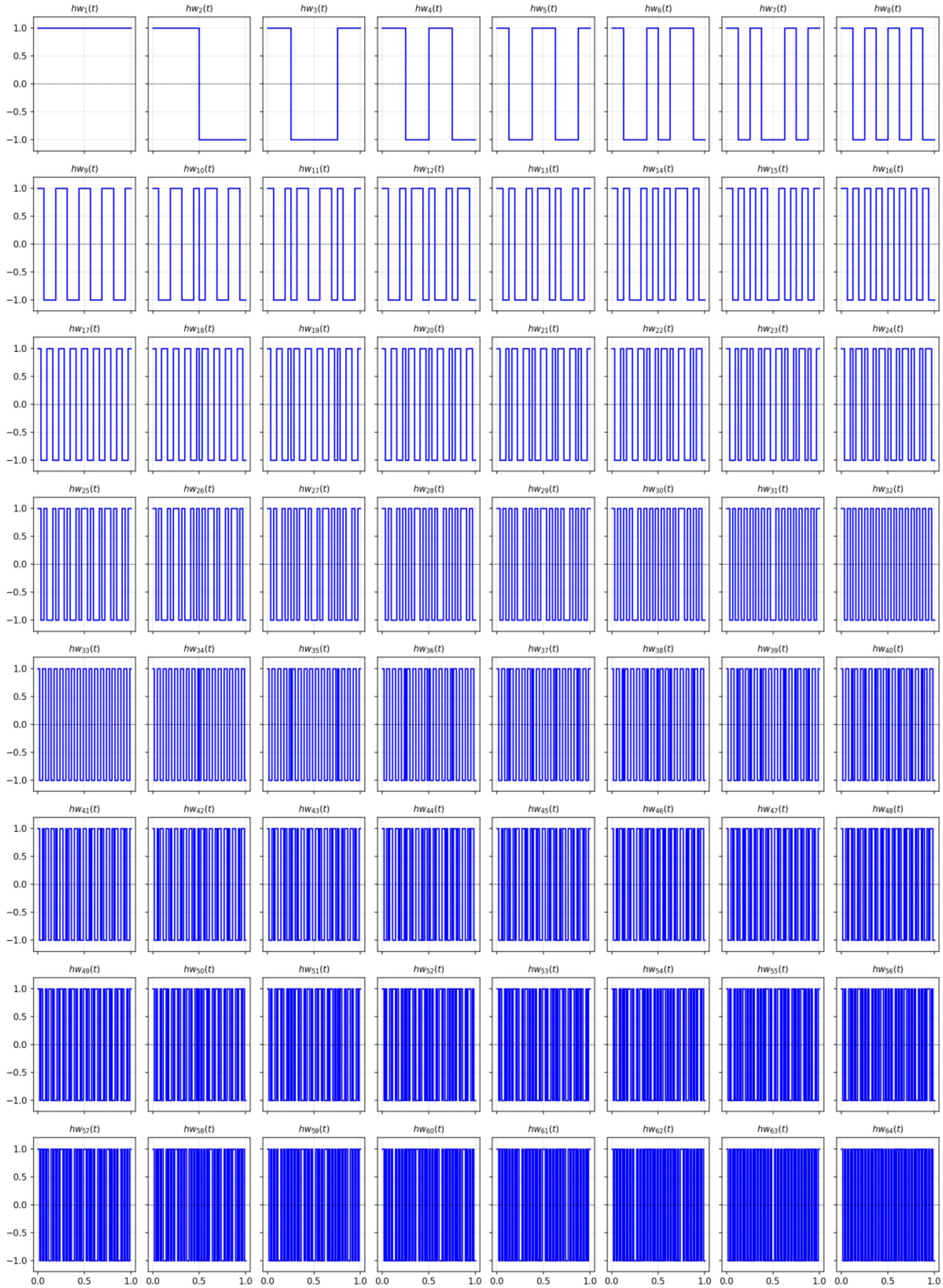


Basis Functions for n=4: Walsh-Hadamard (Sign Change Order)



Basis Functions for n=5: Walsh-Hadamard (Sign Change Order)

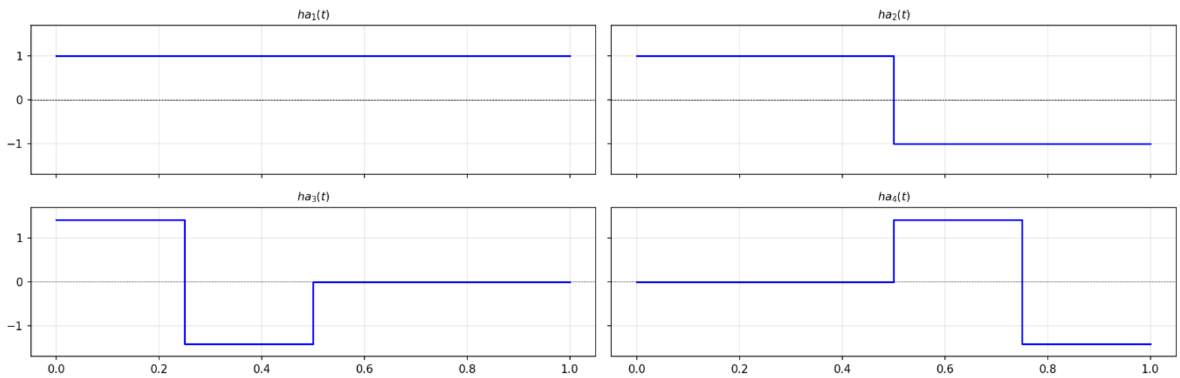
Basis Functions for n=6: Walsh-Hadamard (Sign Change Order)



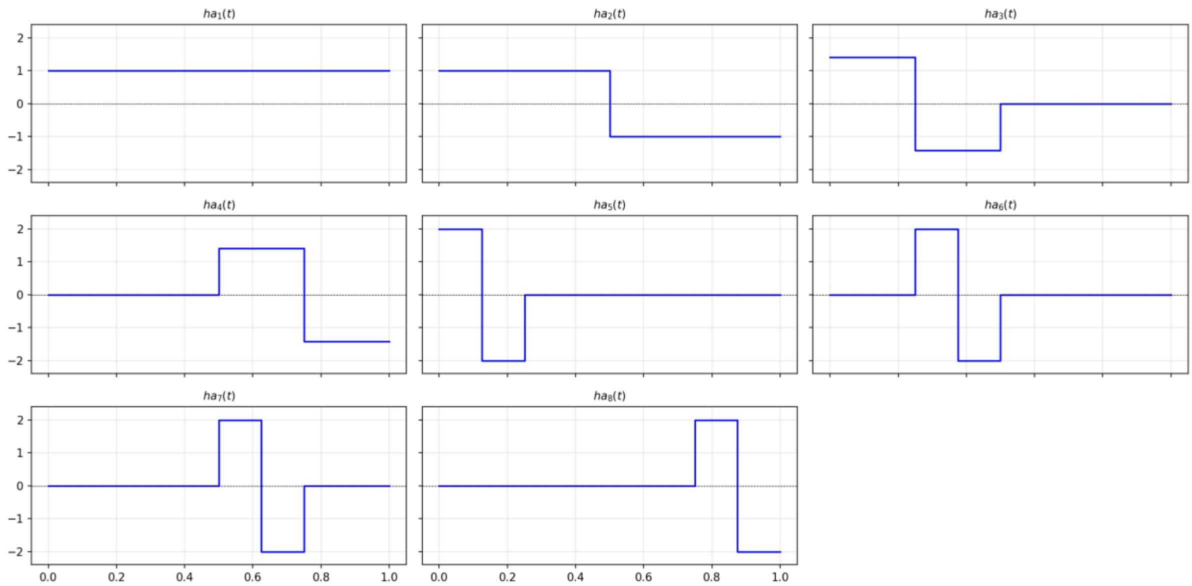
e. We implemented Haar matrix generation as function of n as requested, while preserving the matrix orthonormal. See **haar_matrix** function in **matrices.py** file.

f. We did it by taking the columns of the appropriate Haar matrix multiplied by normalization factor – square of N:

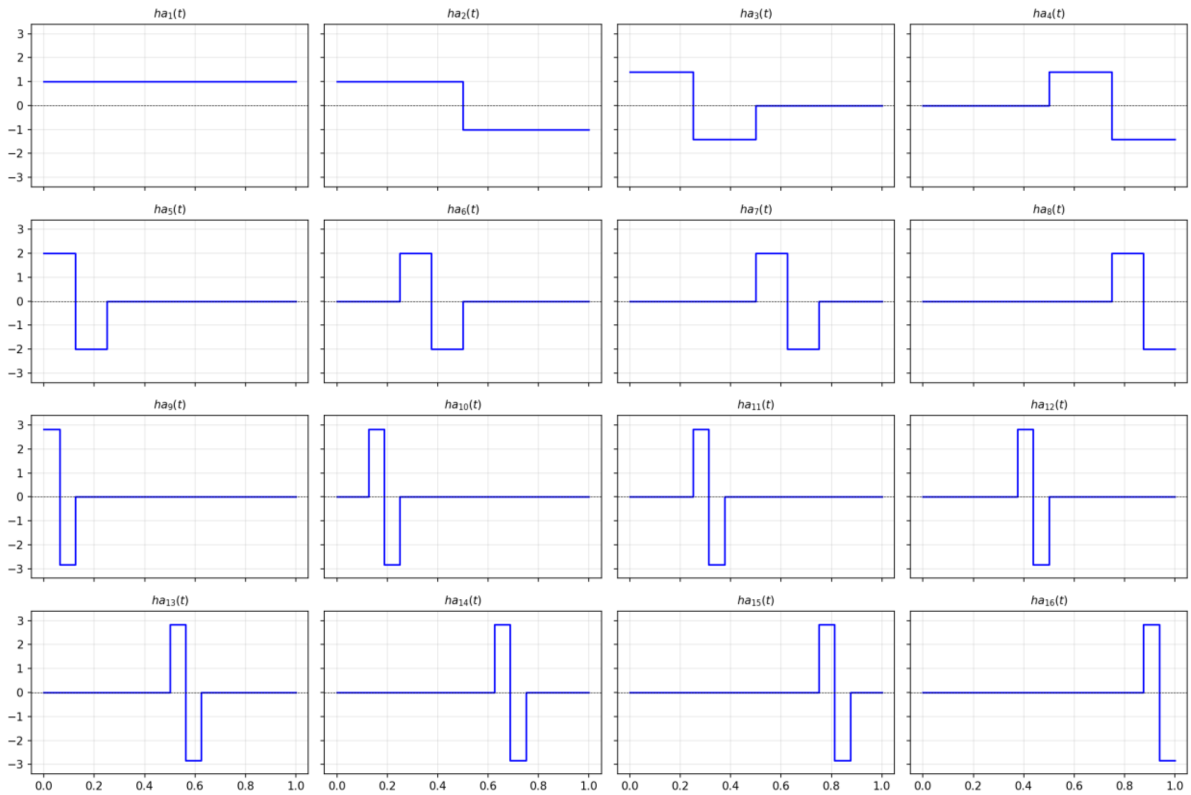
Basis Functions for n=2: Haar Basis Functions



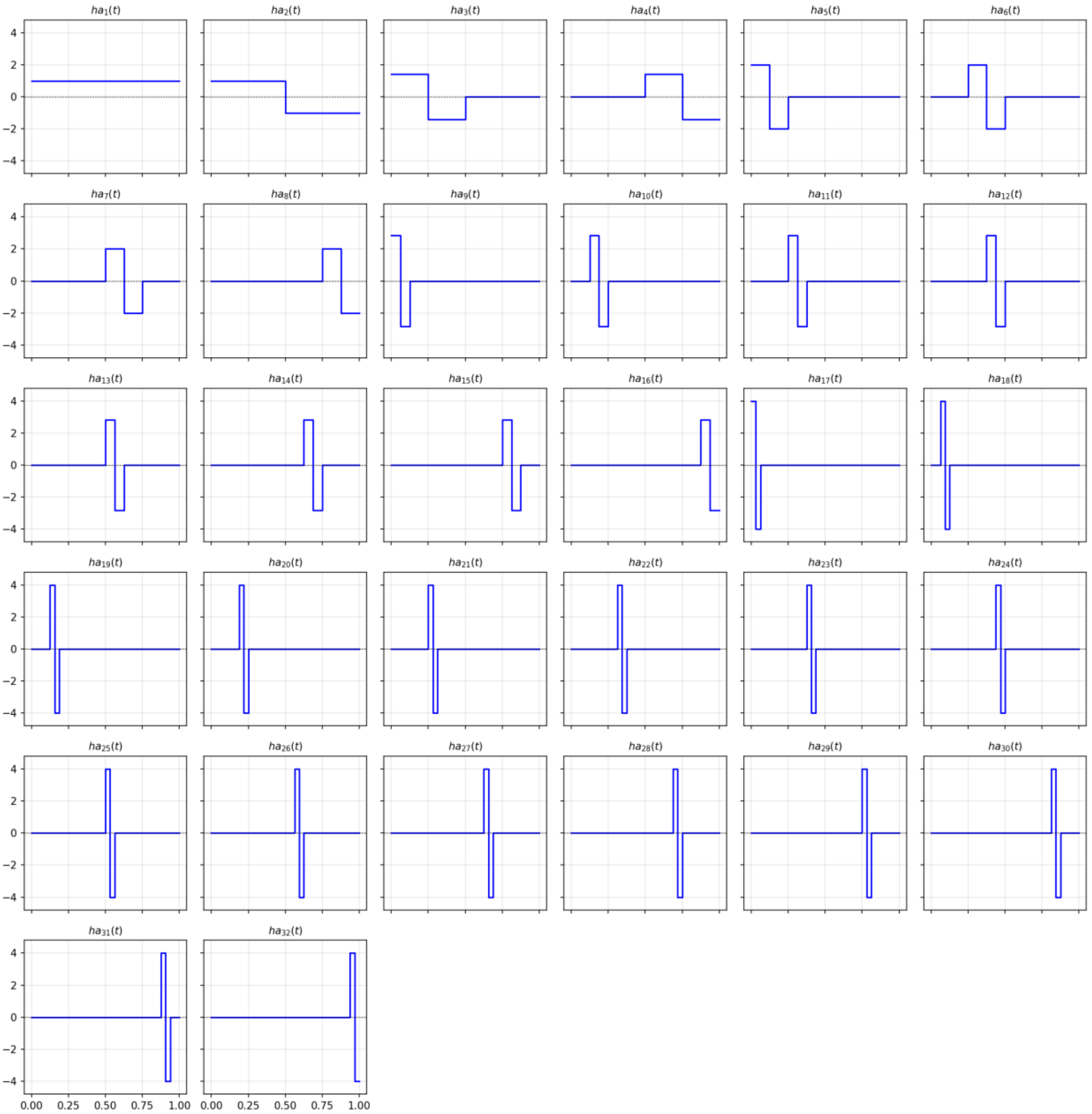
Basis Functions for n=3: Haar Basis Functions



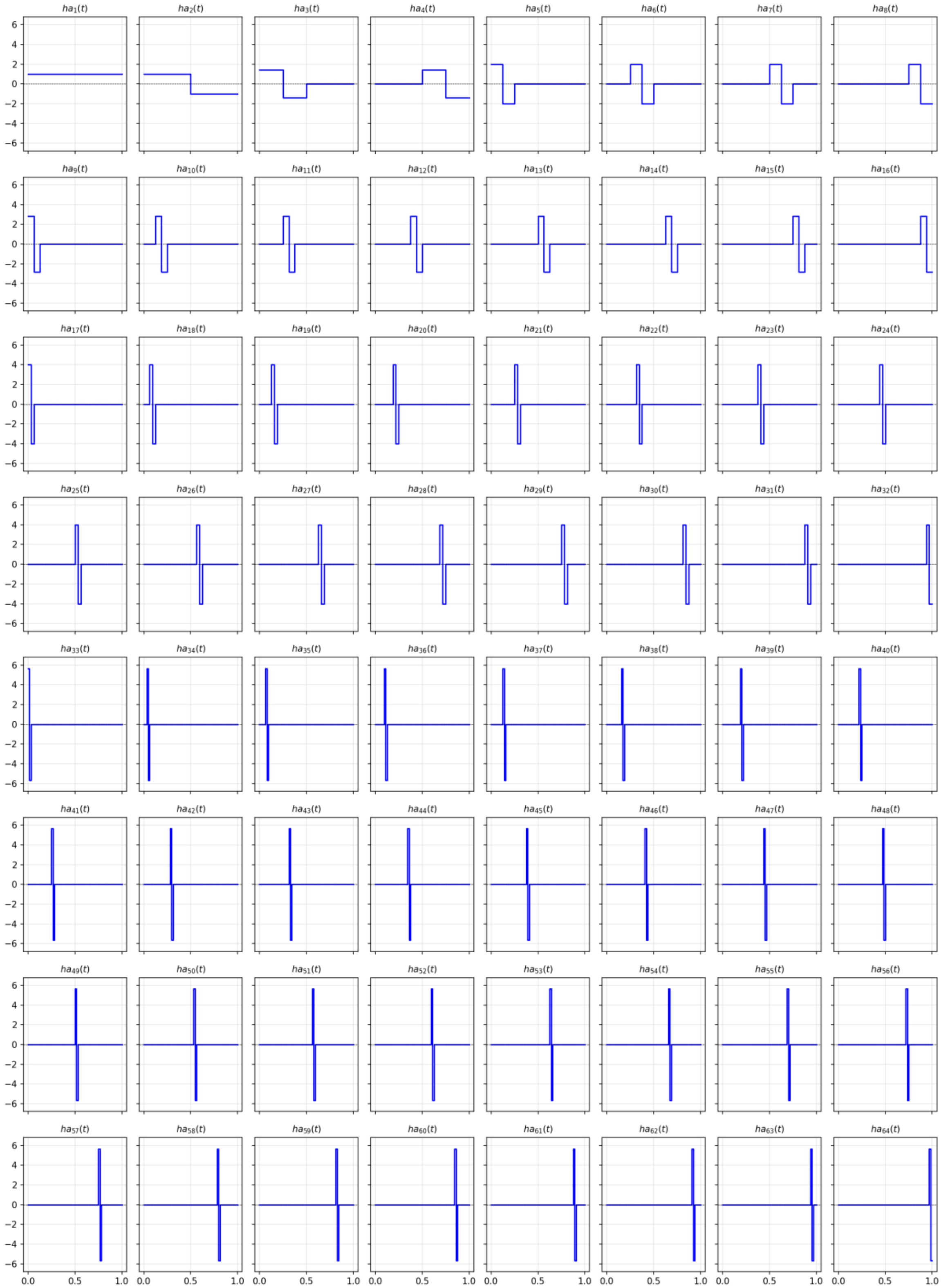
Basis Functions for n=4: Haar Basis Functions



Basis Functions for n=5: Haar Basis Functions

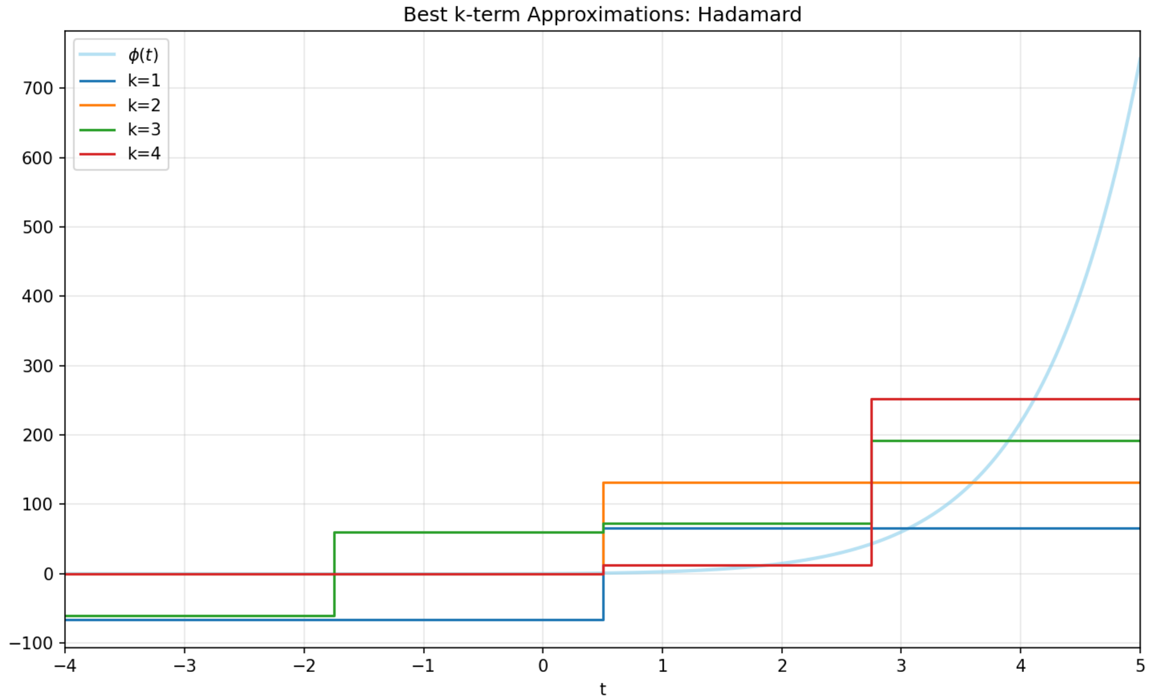


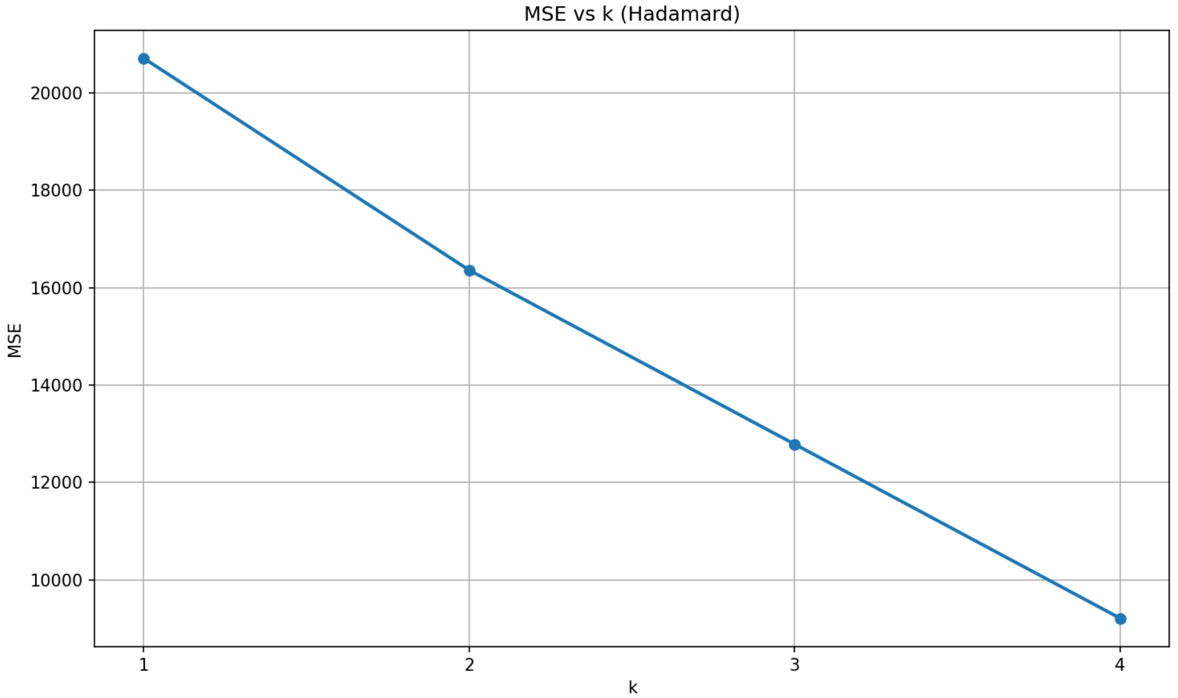
Basis Functions for n=6: Haar Basis Functions



g. We will introduce the results for each one of the bases:

i. In Hadamard basis:

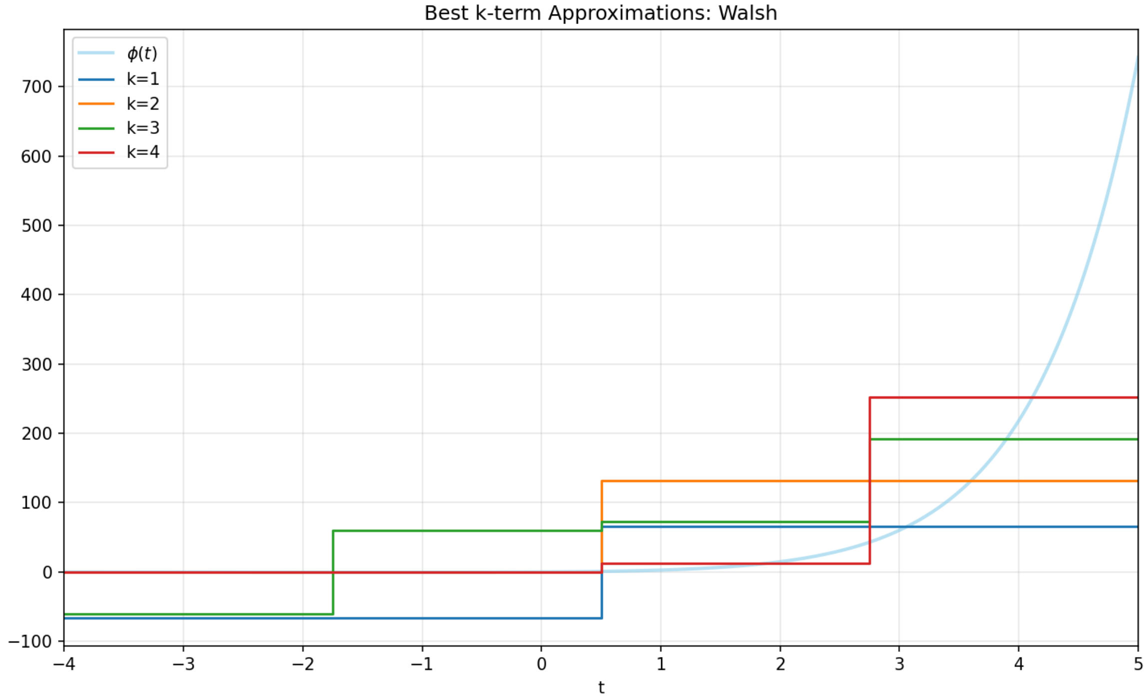


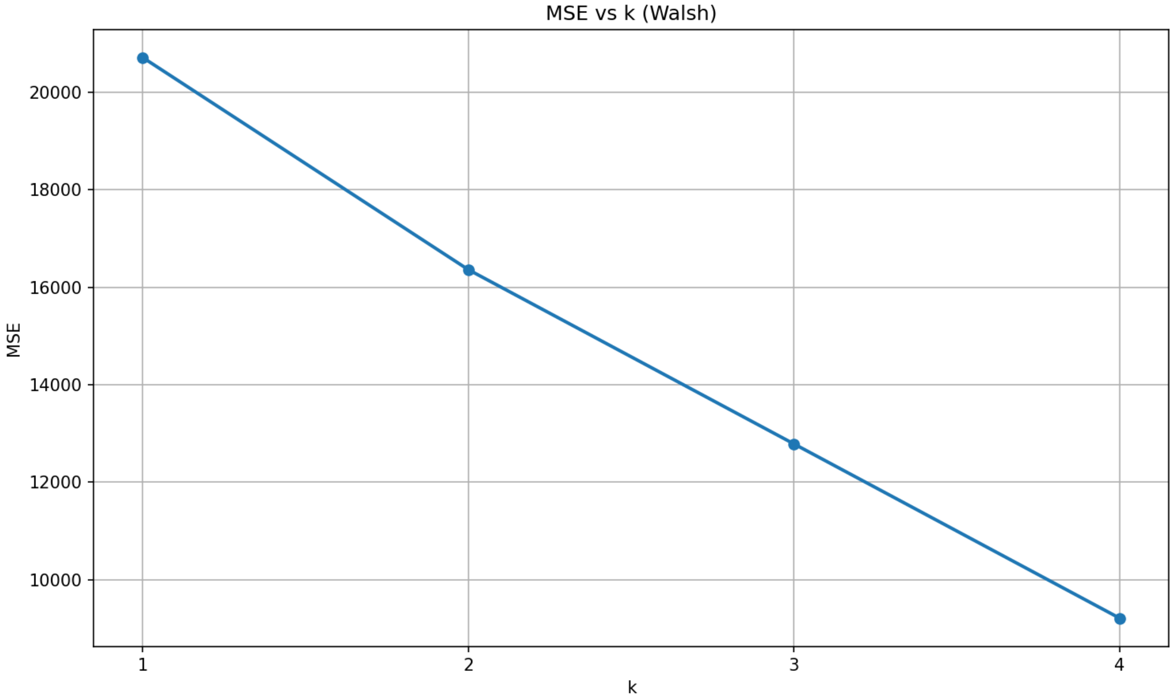


MSE results:

1. 1-term approx.: 20711.93
2. 2-term approx.: 16359.68
3. 3-term approx.: 12784.72
4. 4-term approx.: 9210.81

ii. For Walsh Hadamard basis:

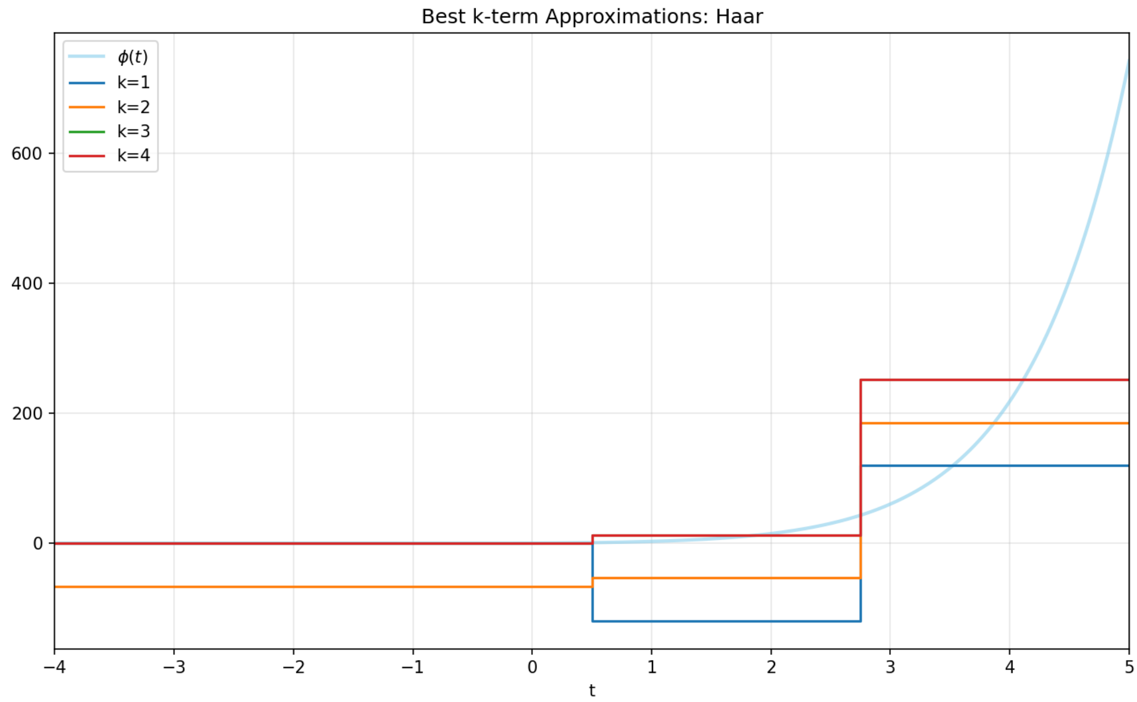


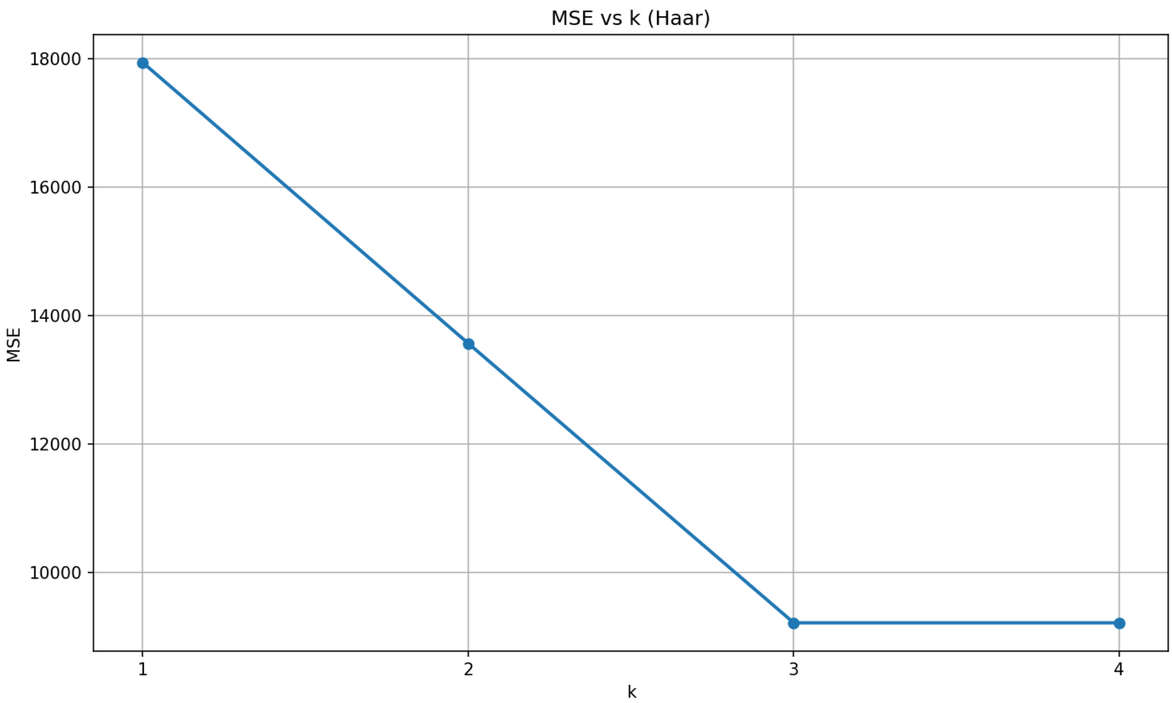


MSE results:

1. 1-term approx.: 20711.93
2. 2-term approx.: 16359.68
3. 3-term approx.: 12784.72
4. 4-term approx.: 9210.81

iii. For Haar basis





MSE results:

1. 1-term approx.: 17936.83
2. 2-term approx.: 13563.06
3. 3-term approx.: 9210.81
4. 4-term approx.: 9210.81

We see that the MSE results for the Hadamard and Walsh-Hadamard bases are identical for all k-term approximations. This is expected, as the Walsh-Hadamard basis is simply a permutation of the Hadamard basis. They consist of the exact same set of vectors, arranged in a different order. Consequently, sorting the coefficients by magnitude selects the same set of basis vectors, resulting in identical reconstruction errors.

In contrast, the Haar basis yields different MSE values for partial approximations (e.g., k=1, 2, 3) because its individual basis vectors differ structurally from the Hadamard family. However, for the full approximation (k=4), all three bases achieve the exact same MSE. This is because all three sets are complete bases that span the exact same vector space. Once all terms are included, the reconstructed signal is equivalent regardless of the basis used.