# Theory

## Question 1:

**a)** $w =$ constant function.

We partition $[0,1]$ to $N$ uniform intervals

$$\Rightarrow \forall i \in [1,N]: \ I_i = \left[\frac{i-1}{N}, \frac{i}{N}\right]$$

**p=2:**

Problem: $$\min_{\hat{f}} \int_0^1 |f(x) - \hat{f}(x)|^p w(x)dx = \min_{\hat{f}} \int_0^1 |f(x) - \hat{f}(x)|^2 dx$$

$p=2$, $w=$const

Looking at each Interval seperately, this problem is equivilant to minimizing $\Psi_\Delta(\bar{\varphi}) \equiv \frac{1}{|\Delta|} \int_\Delta (\varphi(t) - \bar{\varphi})^2 dt$

Where: $t = x$, $\varphi(t) = f(x)$, $\bar{\varphi} = \bar{f}$ (=const in a specific interval).

We know the solution for this problem is:

$$\bar{\varphi}_{optimal} = \frac{1}{|\Delta|} \int_\Delta \varphi(t) dt$$

Back to our problem:

$$\hat{f}(x) = \frac{1}{|I_i|} \int_{I_i} f(x)dx \ \Big| \ x \in I_i \quad \text{or} \quad i = \left\lceil \frac{x}{1/N} \right\rceil$$

$$\Downarrow$$

$$\hat{f}(x) = N \int_{I_i} f(x)\,dx \ \Big| \ x \in I_i$$

**p = 1:**

$$\min_{\hat f} \int_0^1 |f(x)-\hat f|^p \underbrace{w(x)}_{\text{const}} dx = \min_{\hat f} \int_0^1 |f(x)-\hat f| w \, dx$$

Looking at each interval seperately, this problem is equivilant to minimizing $\Psi_\Delta(\bar\varphi) \equiv \frac{1}{|\Delta|} \int_\Delta (\varphi(t)-\bar\varphi)\, dt$

where: $t=x$, $\varphi(t)=f(x)$, $\bar\varphi = \hat f$ (=const in a specific interval).

we know the solution for this problem is:

$\bar\varphi \equiv$ median of $\varphi(t)$

Back to our problem:

$\hat f(x) =$ median of $\hat f$ over $I_i$ interval $\Big|\ x \in I_i$

b) $p=2$ & general $w$

$$\Psi_{wmse}(\hat f) = \int_0^1 |f(x)-\hat f(x)|^2 w(x)dx = \sum_{i=1}^{N} \int_{\frac{i-1}{N}}^{\frac{i}{N}} (f(x)-\hat f)^2 w(x) dx$$

for each interval $I_i$, $\hat f(x)$ is A const $\triangleq f_i$

$$\Rightarrow \Psi_{wmse}(\hat f) = \sum_{i=1}^{N} \int_{\frac{i-1}{N}}^{\frac{i}{N}} (f(x)-f_i)^2 w(x) dx$$

$$= \int_0^{\frac{1}{N}} (f(x)-f_1)^2 w(x)dx + \dots \int_{\frac{i-1}{N}}^{\frac{i}{N}} (f(x)-f_i)^2 w(x) dx + \dots \int_{\frac{N-1}{N}}^{1} (f(x)-f_1)^2 w(x)dx$$

to find $\hat f$ that minimizes $\Psi(\hat f)$ we'll demand:

$$\forall i \in [1,N]: \frac{\partial \Psi_{wMSE}(\hat f)}{\partial f_i} = 0$$

$$\frac{\partial \Psi_{WMSE}(\hat f)}{\partial f_i} = \int_{I_i} \cancel{(-1)}\cdot 2(f(x)-f_i)\,w(x)\,dx = 0$$

$$\int_{I_i} f(x)w(x)\,dx = \int_{I_i} f_i\, w(x)\,dx$$ / for each $I_i$, $f_i$ = const

$$\forall i\in[1,N]:\ f_i = \frac{\int_{I_i} f(x)w(x)\,dx}{\int_{I_i} w(x)\,dx} \qquad \hat f(x) = f_i \mid x\in I_i$$

C) $p=1$ & general $w$

$$\Psi_{wMAD}(\hat f) = \int_0^1 |f(x)-\hat f(x)|^1 w(x)\,dx = \sum_{i=1}^{N} \int_{\frac{i-1}{N}}^{\frac{i}{N}} |f(x)-\hat f|\, w(x)\,dx$$

for each interval $I_i$, $\hat f(x)$ is a const $\triangleq f_i$

$$\Rightarrow \Psi_{wMAD}(\hat f) = \sum_{i=1}^{N} \int_{\frac{i-1}{N}}^{\frac{i}{N}} |f(x)-f_i|\, w(x)\,dx$$

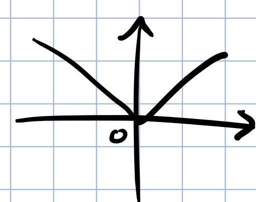
to find $\hat f$ that minimizes $\Psi(\hat f)$ we'll demand:

$$\forall i\in[1,N]:\ \frac{\partial \Psi_{wMAD}(\hat f)}{\partial f_i} = 0$$

The problem is - $|f(x)-\hat f(x)|$ is not differentiable at 0.

To use subdifferentials, we'd have to find $\hat f$

s.t there is an $s\in[-1,1]$ for which $\frac{\partial}{\partial f_i} MAD = 0$



$$\frac{\partial}{\partial f_i} MAD = \int_{\substack{I_i\\ f(x)>f_i}} (-1)\,w(x)\,dx \ \overset{+}{\cancel{-1}} \int_{\substack{I_i\\ f(x)<f_i}} \cancel{(-1)}\,w(x)\,dx - s\int_{\substack{I_i\\ f(x)=f_i}} w(x)\,dx,\ \forall s\in[-1,1] \quad \underline{=0}$$

$$\exists s,\ \frac{\partial}{\partial f_i} MAD = 0 \Leftrightarrow \int_{\substack{I_i\\ f(x)<f_i}} w(x)\,dx = \int_{\substack{I_i\\ f(x)>f_i}} w(x)\,dx + s\int_{\substack{I_i\\ f(x)=f_i}} w(x)\,dx$$

$\Rightarrow$ $f_i$ = A function over $I_i$ which satisfies the above equation (to put in words: $f_i$ is the weighted median of the interval $I_i$).

$\hat{f}(x) = f_i \mid x \in I_i$

d) $$\mathcal{E}^p(f,\hat{f}) = \int_0^1 |f(x)-\hat{f}(x)|^p w(x)dx = \sum_{i=1}^{N} \int_{\frac{i-1}{N}}^{\frac{i}{N}} |f(x)-\hat{f}(x)|^p w(x)dx$$

spliting the integral to $N$ equal intervals $I_i$ s.t $\sum_{i=1}^{N} I_i = [0,1]$.

for each interval $I_i = [\frac{i-1}{N}, \frac{i}{N}]$, $i \in [1,N]$: $\hat{f}(x)$ is a const function by it's definition as piecewise const.

We'll define: 1) $f_i$: $I_i \to$ Image of $f(x)$ as $f_i = f(x)$

2) $\hat{f}_i$: $I_i \to$ Image of $\hat{f}(x)$ as $\hat{f}_i = \hat{f}(x)$

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

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

we get: $\mathcal{E}^p(f,\hat{f}) = \sum_{i=1}^{N} \mathcal{E}^p_i(f_i,\hat{f}_i)$

Optimization of $\mathcal{E}^p$ is equivalent to optimizing each $\mathcal{E}_i^p$ seperatly, as we optimize with $\frac{\partial}{\partial f_i}\mathcal{E}^p$, and each $\mathcal{E}_i^p$ is independent of $f_j \mid \forall j \neq i$.

e) i Assuming $\forall x \in I_i : f_i(x) \neq \hat{f}_i(x)$

$$|f_i(x) - \hat{f}_i(x)|^p = \omega_{f_i,\hat{f}_i}(x) \cdot (f_i(x) - \hat{f}_i(x))^2 \quad / \ \forall t : t^2 = |t|^2$$

$$|f_i(x) - \hat{f}_i(x)|^p = \omega_{f_i,\hat{f}_i}(x) \cdot |f_i(x) - \hat{f}_i(x)|^2 \quad / \cdot \frac{1}{|f_i(x) - \hat{f}_i(x)|^2}$$

$|f_i(x) - \hat{f}_i(x)|^2 \neq 0$

As $\forall x \in I_i : f_i(x) \neq \hat{f}_i(x)$

$$\omega_{f_i,\hat{f}_i}(x) = \frac{|f_i(x) - \hat{f}_i(x)|^p}{|f_i(x) - \hat{f}_i(x)|^2}$$

$$\omega_{f_i,\hat{f}_i}(x) = |f_i(x) - \hat{f}_i(x)|^{p-2}$$

ii) $$\mathcal{E}_i^p = \int_{I_i} |f_i(x) - \hat{f}_i(x)|^p w(x)\, dx \underset{e(i)}{=} \int_{I_i} \omega_{f_i,\hat{f}_i}(x) \cdot (f_i(x) - \hat{f}_i(x))^2 \, w(x)\, dx$$

we'll define $w'_{f_i,\hat{f}_i} = w(x) \cdot \omega_{f_i,\hat{f}_i}(x)$

$$\Rightarrow \mathcal{E}_i^p = \int_{I_i} \omega'_{f_i,\hat{f}_i} \cdot (f_i(x) - \hat{f}_i(x))^2 \, dx$$

iii) Optimizing the $L^p$ problem requires partial derivatives of $\hat{f}_i$'s. Having $w$ dependent on $\hat{f}_i$ makes the derivative calculation way harder compared to an independent $w$!

iv) Without the assumption of $f_i(x) \neq \hat{f}_i(x) \; \forall x \in I_i$, we prefer to use $\tilde{\omega}_{f_i,\hat{f}_i}(x) = \min\{\frac{1}{\varepsilon}, \omega_{f_i,\hat{f}_i}(x)\}$ instead of $\omega_{f_i,\hat{f}_i}(x)$ where $\varepsilon > 0$ a small fixed number to avoid dividing by 0 in the case where $p=1$.

explanation:

$$\underbrace{\omega_{f_i,\hat{f}_i}(x)}_{e(i)} = |f_i(x) - \hat{f}_i(x)|^{p-2} \underset{\text{for } p=1}{=} \frac{1}{|f_i(x) - \hat{f}_i(x)|} \underset{\text{for } f_i(x) = \hat{f}_i(x)}{=} \frac{1}{0} = \text{very bad math}$$

replacing solves this problem as $\min\{\frac{1}{\varepsilon}, \omega_{f_i,\hat{f}_i}(x)\}$ would return $\frac{1}{\varepsilon}$ in this case where "$\omega \to \infty$".

In addition to avoiding dividing by zero, this replacement is somewhat propotional as we replace <u>zero</u> with a very <u>small number</u> $\varepsilon$.

v) assumptions:

1) Initial $\hat{f}_i$ and $\omega_i$ are given.
2) We have a machine that calculates integrals.
3) At each iteration, $\omega_i$ is independent of $\hat{f}_i$.

Based on these assumptions, we can solve the $\varepsilon^p$ problem as $\varepsilon^2$.

Pseudo code ↓

init:

- $(\hat{f}_i, w'_i) =$ provided $(\hat{f}_i, w'_i)$
- error $= \int_{I_i} w'_i |f_i(x) - \hat{f}_i(x)|^p dx$

repeat:

- $f_i^{next} = \arg\min_{\hat{f}_i} \int_{I_i} w_i |f_i(x) - \hat{f}_i|^2 dx = \dfrac{\int_{\Delta_i} w'(x) f_i dx}{\int_{\Delta_i} w'(x) dx}$
- $\hat{f}_i = f_i^{next}$
- $w'_i = \min\{w(x)|f_i(x) - \hat{f}_i|^{p-2}, \varepsilon\}$
- prev_error = error
- error $= \int_{I_i} w'_i |f_i(x) - \hat{f}_i(x)|^p dx$

stop condition: error - prev_error $< \delta$ of our choice.

Note: The above pseudo-code is for the specific interval $I_i$. If you meant solving for all intervals, minor changes are requiered (i.e: iterating on all intervals...).

f) In this section, we are not given init vals. We'll init to an arbitrary values, and then use prev pseudo code:

for each interval $I_i$:

init
- $w'_i = 1$
- $\hat{f}_i = \dfrac{\int_{\Delta_i} w'(x) f_i dx}{\int_{\Delta_i} w'(x) dx}$

- run pseudo code from 1.e(v)

## Question 2

$\Phi(t): [0,1) \to [\phi_L, \phi_H]$ , $\Delta_i = [\frac{i-1}{N}, \frac{i}{N}] \mid i \in [1, n]$

$\hat{\Phi}(t) = a_i(t-t_i) + c_i \mid t \in \Delta_i$ , $a_i, c_i = \text{const} \in \mathbb{R}$, $t_i$ = center of $\Delta_i$

evaluating the approximatin with MSE.

**a)** $$\int_{t\in\Delta_i} (t-t_i)^k dt = \frac{(t-t_i)^{k+1}}{k+1}\Big|_{\frac{i-1}{N}}^{\frac{i}{N}} = \frac{1}{k+1}\left[\left(\frac{i}{N}-t_i\right)^{k+1} - \left(\frac{i-1}{N}-t_i\right)^{k+1}\right]$$

* $t_i$ is the center of $\Delta_i$, hence $\frac{i}{N} - t_i = \frac{|\Delta_i|}{2} = -\left(\frac{i-1}{N} - t_i\right)$

where $|\Delta_i|$ is the size of the $\Delta_i$ interval.

we'll denote $X_i = \frac{|\Delta_i|}{2}$

$$\Rightarrow \int_{t\in\Delta_i} (t-t_i)^k dt = \frac{1}{k+1}\left(X_i^{k+1} - (-X)^{k+1}\right)$$

* if $k$ is odd $\to$ $k+1$ is even $\to$ $(-X_i)^{k+1} = X_i^{k+1}$

$$\int_{t\in\Delta_i} (t-t_i)^k dt = \frac{1}{k+1}\left(X_i^{k+1} - X_i^{k+1}\right) = \frac{1}{k+1}\cdot 0 = \underline{0}$$

* if $k$ is even $\to$ $k+1$ is odd $\to$ $(-X_i)^{k+1} = -(X_i)^{k+1}$

$$= \frac{1}{k+1}\left(X_i^{k+1} - (-)(X_i)^{k+1}\right) = \frac{2X_i^{k+1}}{k+1} \underset{X_i = \frac{|\Delta_i|}{2}}{=} \frac{2\left(\frac{|\Delta_i|}{2}\right)^{k+1}}{k+1} = \frac{\cancel{2}|\Delta_i|^{k+1}}{2^{k\cancel{+1}}(k+1)}$$

$$= \frac{|\Delta_i|^{k+1}}{2^k(k+1)}$$

to conclude:
$$\int_{t\in\Delta_i}(t-t_i)^k dt = \begin{cases} 0 & k \text{ is odd} \\ \frac{|\Delta_i|^{k+1}}{2^k(k+1)} & k \text{ is even} \end{cases}$$

as requested.

■

**b)** $$\Psi_{MSE}(\hat{\phi}) = \int_0^1 \left(\phi(t)-\hat{\phi}(t)\right)^2 dt = \sum_{i=1}^{N}\int_{\Delta_i}\left(\phi(t)-[a_i(t-t_i)+c_i]\right)^2 dt$$

$$= \sum_{i=1}^{N}\int_{\Delta_i}\left(\phi(t)-a_i(t-t_i)-c_i\right)^2 dt$$

to minimize $\Psi_{MSE}(\hat{\phi})$, we'll demand $\begin{cases} ① \ \frac{\partial}{\partial c_i}\Psi_{MSE}(\hat{\phi})=0 \\ ② \ \frac{\partial}{\partial a_i}\Psi_{MSE}(\hat{\phi})=0 \end{cases}$

① $$\frac{\partial}{\partial c_i}\Psi_{MSE}(\hat{\phi}) = \int_{\Delta_i}\cancel{(-1)}\cdot 2\left(\phi(t)-a_i(t-t_i)-c_i\right)dt = 0$$

$$\int_{\Delta_i}(\phi(t)dt - \int_{\Delta_i}a_i(t-t_i)dt - \int_{\Delta_i}c_i dt = 0$$

/ for each interval $\Delta_i$, $a_i, c_i$ = const

$$\int_{\Delta_i}(\phi(t)dt - a_i\underbrace{\int_{\Delta_i}(t-t_i)^1 dt}_{=0} - c_i|\Delta_i| = 0$$

/ $k=1$ is odd. Question 2.a

$$c_i = \frac{1}{|\Delta_i|}\int_{\Delta_i}\phi(t)\,dt$$

② $$\frac{\partial}{\partial a_i}\Psi_{MSE}(\hat{\phi}) = \int_{\Delta_i} \cancel{-}(t-t_i)\cdot \cancel{2}\left(\phi(t) - a_i(t-t_i) - c_i\right) dt = 0$$

$$\int_{\Delta_i}(t-t_i)\left(\phi(t) - a_i(t-t_i) - c_i\right) dt = 0$$ / $a_i, c_i, t_i$ are const for each interval $\Delta_i$

$$\int_{\Delta_i}(t-t_i)\phi(t)\,dt - a_i\underbrace{\int_{\Delta_i}(t-t_i)^2 dt}_{k=2 \text{ is even } = \frac{|\Delta_i|^{2+1}}{2^2(2+1)}} - c_i\underbrace{\int_{\Delta_i}(t-t_i)\,dt}_{k=1 \text{ is odd. } =0} = 0$$

$$\int_{\Delta_i}(t-t_i)\phi(t)\,dt - \frac{|\Delta_i|^3}{12}a_i = 0$$

$$a_i = \frac{12}{|\Delta_i|^3}\int_{\Delta_i}(t-t_i)\phi(t)\,dt$$

c) $$\Psi_{MSE}(\hat{\phi}) = \int_0^1\left(\phi(t)-\hat{\phi}(t)\right)^2 dt = \sum_{i=1}^{N}\int_{\Delta_i}\left(\phi(t) - [a_i(t-t_i)+c_i]\right)^2 dt$$

$$= \sum_{i=1}^{N}\int_{\Delta_i}\left(\phi(t) - a_i(t-t_i) - c_i\right)^2 dt$$

$$= \sum_{i=1}^{N}\Bigg[\int_{\Delta_i}\phi(t)^2 dt + a_i^2\int_{\Delta_i}(t-t_i)^2 dt + c_i^2\int_{\Delta_i}dt - 2a_i\int_{\Delta_i}(t-t_i)\phi(t)\,dt$$
$$-2c_i\int_{\Delta_i}\phi(t)\,dt + 2a_ic_i\underbrace{\int_{\Delta_i}(t-t_i)\,dt}_{=0 \text{ k is odd}}\Bigg]$$

$$= \sum_{i=1}^{N}\Bigg[\int_{\Delta_i}\phi(t)^2 dt + \overbrace{a_i^2\frac{|\Delta_i|^3}{12}}^{1} + \overbrace{c_i^2|\Delta_i|}^{2} \overbrace{-2a_i\int_{\Delta_i}(t-t_i)\phi(t)\,dt}^{3}$$
$$\overbrace{-2c_i\int_{\Delta_i}\phi(t)\,dt}^{4} + 0\Bigg]$$

① $$a_i^2 \frac{|\Delta_i|^3}{12} = \left(\frac{12}{|\Delta_i|^3}\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2 \cdot \frac{|\Delta_i|^3}{12} = \boxed{\frac{12}{|\Delta_i|^3}\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2}$$

② $$c_i^2|\Delta_i| = \left(\frac{1}{|\Delta_i|}\int_{\Delta_i}\phi(t)\,dt\right)^2 |\Delta_i| = \boxed{\frac{1}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2}$$

③ $$-2a_i\int_{\Delta_i}(t-t_i)\phi(t)\,dt = -2\cdot\left(\frac{12}{|\Delta_i|^3}\cdot\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)\cdot\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)$$

$$= \boxed{\frac{-24}{|\Delta_i|^3}\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2}$$

④ $$-2c_i\int_{\Delta_i}\phi(t)\,dt = \left(\frac{-2}{|\Delta_i|}\cdot\int_{\Delta_i}\phi(t)\,dt\right)\left(\int_{\Delta_i}\phi(t)\,dt\right) = \frac{-2}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2$$

$$\sum_{i=1}^{N}\left[\int_{\Delta_i}\phi(t)^2dt + \frac{12}{|\Delta_i|^3}\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2 + \frac{1}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2 \right.$$
$$\left. \frac{-24}{|\Delta_i|^3}\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2 - \frac{2}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2\right]$$

$$= \sum_{i=1}^{N}\left[\int_{\Delta_i}\phi(t)^2dt - \frac{12}{|\Delta_i|^3}\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2 - \frac{1}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2\right]$$

$$= \int_0^1\phi(t)^2dt - 12N^3\sum_{i=1}^{N}\left(\int_{\Delta_i}(t-t_i)\phi(t)\,dt\right)^2 - N\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2$$

d) $$MSE_{linear} = \int_0^1 \phi(t)^2 dt - 12N^3 \sum_{i=1}^{N}\left(\int_{\Delta_i}(t-t_i)\phi(t)dt\right)^2 - N\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)dt\right)^2$$

$$MSE_{const} = \underbrace{\int_0^1 \phi(t)^2 dt - \frac{1}{N}\sum_{i=1}^{N}(\hat{\phi}_i)^2}_{\text{from the tutorial}} \overset{*}{=} \int_0^1 \phi(t)^2 dt - N\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)dt\right)^2$$

$$\left( * \quad -\frac{1}{N}\sum_{i=1}^{N}(\hat{\phi}_i)^2 = -\frac{1}{N}\sum_{i=1}^{N}\left[N\int_{\Delta_i}\phi(t)dt\right]^2 = \frac{-1}{N}\sum_{i=1}^{N}N^2\left(\int_{\Delta_i}\phi(t)dt\right)^2 = \frac{-N^2}{N}\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)dt\right)^2 = -N\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)dt\right)^2 \right)$$

$$\left(MSE_{Linear}\right) - \left(MSE_{Const}\right) = \cancel{\int_0^1 \phi(t)^2 dt} - 12N^3 \sum_{i=1}^{N}\left(\int_{\Delta_i}(t-t_i)\phi(t)dt\right)^2 \cancel{- N\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)dt\right)^2} - \left[\cancel{\int_0^1 \phi(t)^2 dt} \cancel{- N\sum_{i=1}^{N}\left(\int_{\Delta_i}\phi(t)dt\right)^2}\right]$$

$$= -12N^3 \sum_{i=1}^{N}\left(\int_{\Delta_i}(t-t_i)\phi(t)dt\right)^2 \leq 0$$

① $\forall x$: $x^2 \geq 0$
② sum of positives is positive
③ minus this sum is $\leq 0$

so we get: $\left(MSE_{Linear}\right) - \left(MSE_{Const}\right) \leq 0$

$$\Longrightarrow \boxed{MSE_{Linear} \leq MSE_{Const}}$$

* This result makes sense as the linear aproximation is more "sophisticated" as it "includes" the const aprox. Thus, it's error should be smaller.

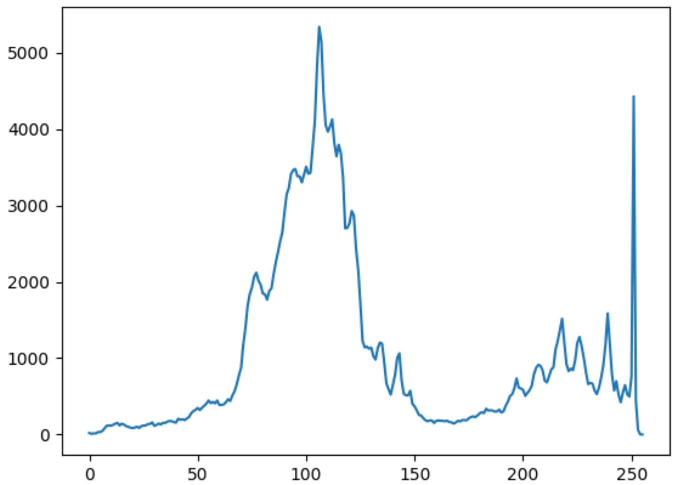
# Implementation

Noam Wolf and Murad Ektilat

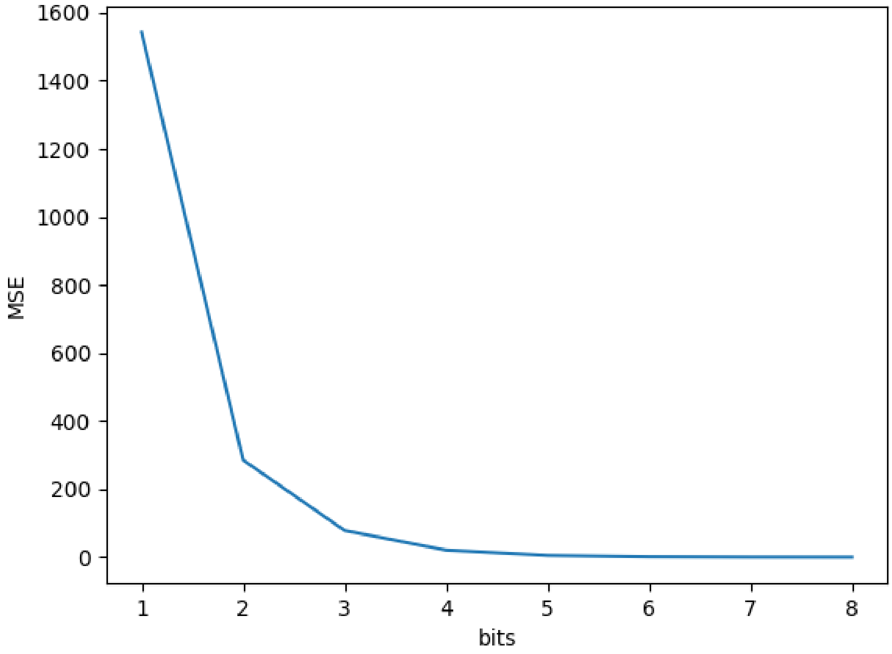
1) Quantization:

The image we chose to quantize:

1.1) The histogram, the pdf is obviously the same, with the y values normalized to be the probability values of the pixel values.

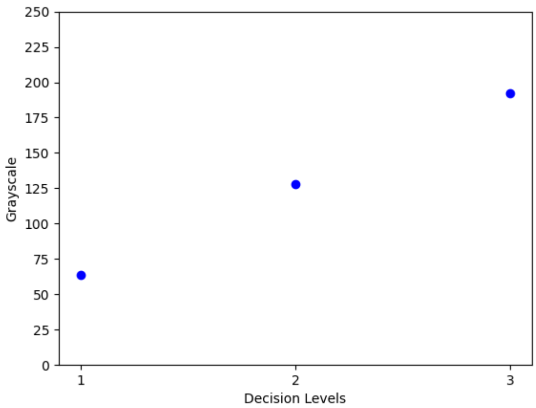


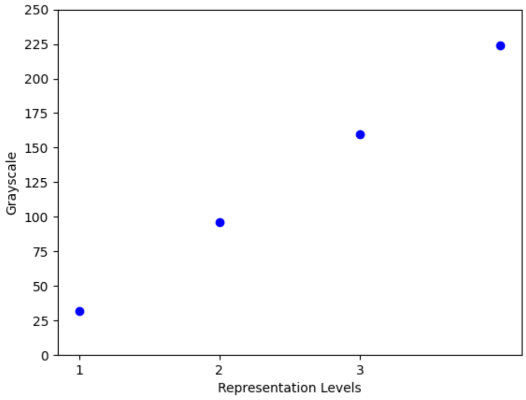
1.2) MSE as a function of the bit-budget, uniform quantizer:



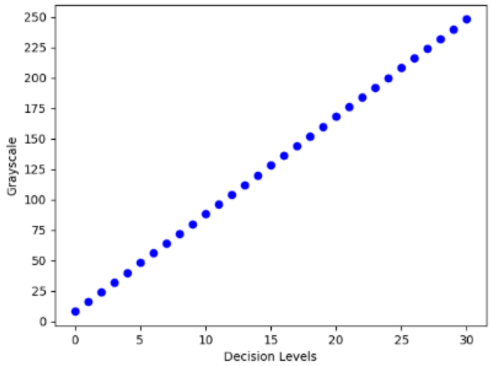

The bits chosen where 2,5,8 for both quantizers

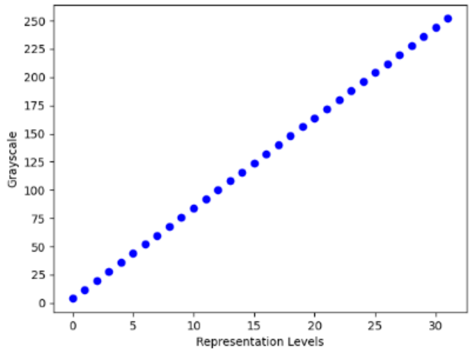
b= 2



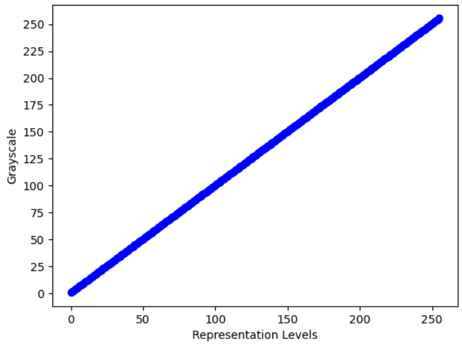



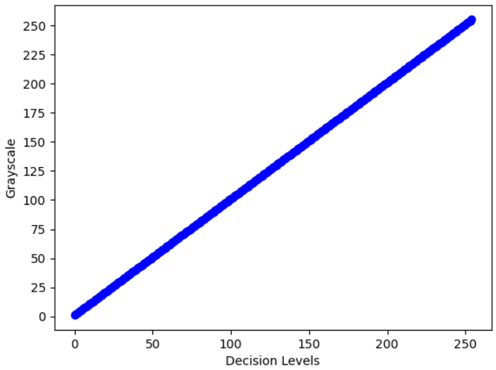
b = 5


Grayscale
Decision Levels

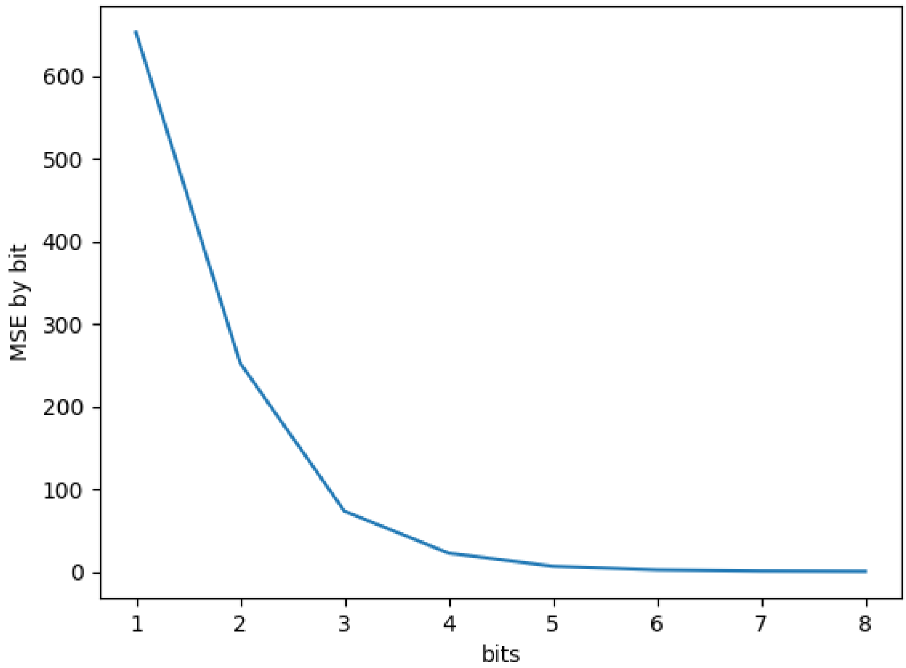

Grayscale
Representation Levels

b = 8

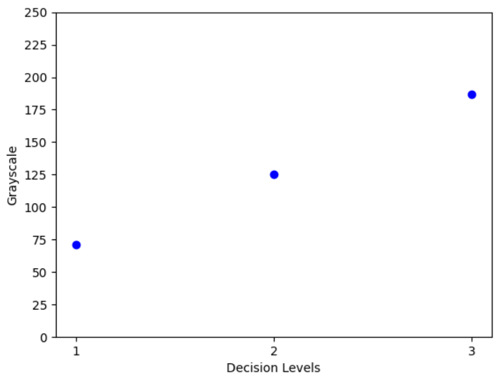

Grayscale
Representation Levels

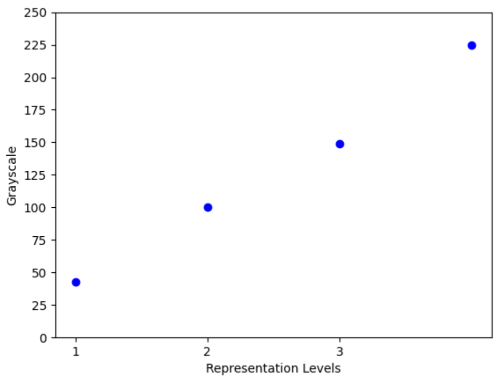

Grayscale
Decision Levels

4) MSE as a function of the bit-budget for Max-Lloyd algorithm:

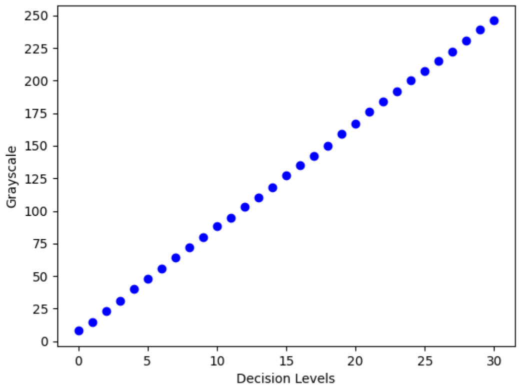


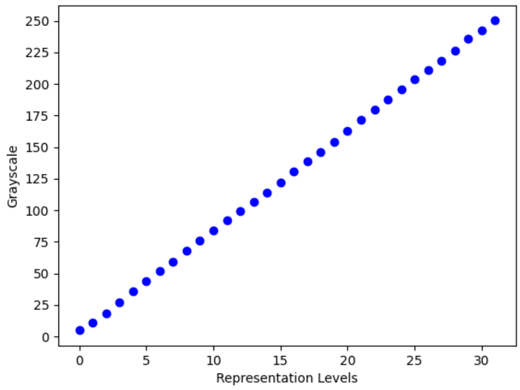

b = 2



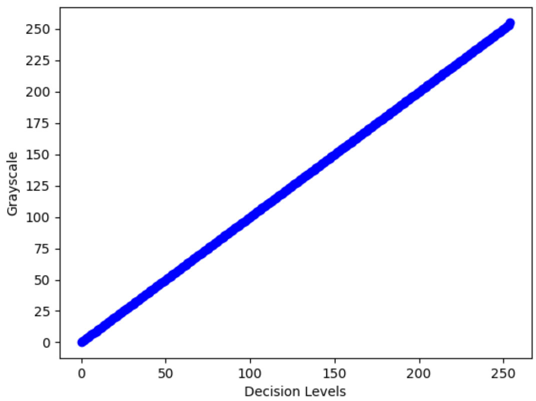



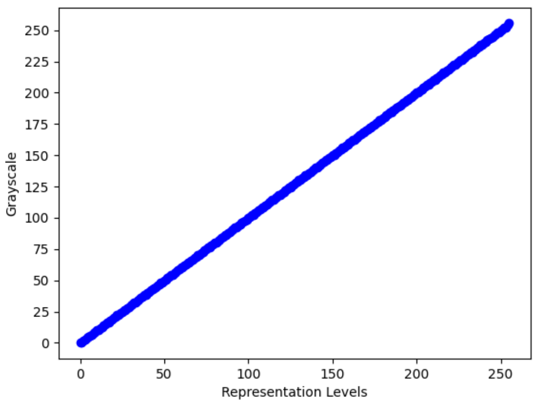
b = 5




b = 8

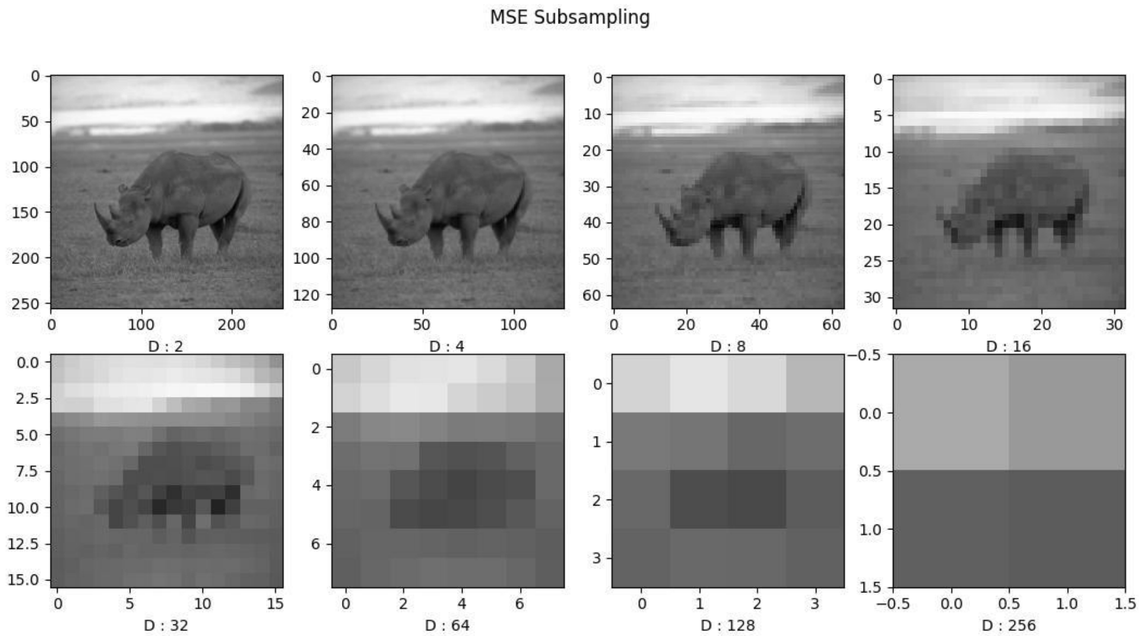


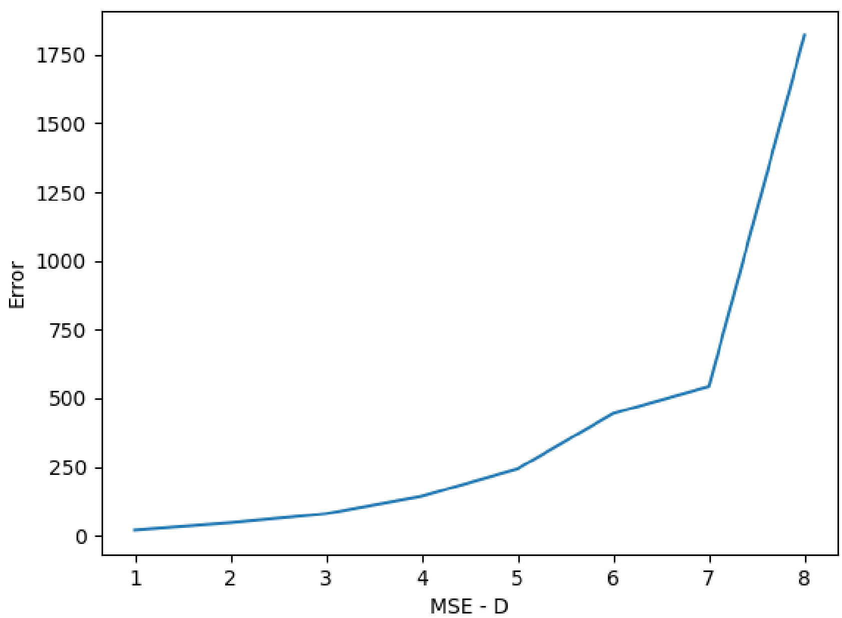

c. The Max-Lloyd works better than the uniform quantizer with smaller bit budgets, as it brings the MSE lower. However both quantizers work well with more bits. The decision and representation levels are also quite different, as they are not uniform, rather local minimums.

2) Subsampling and Reconstruction

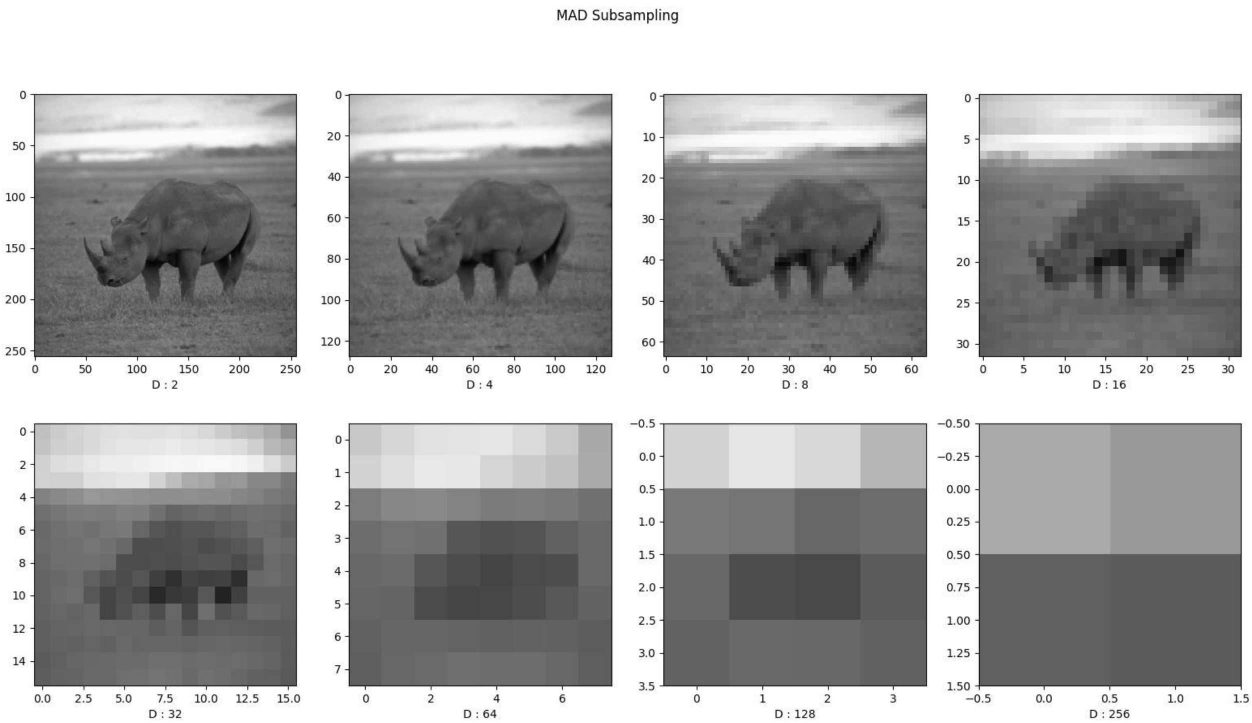
2.1) MSE subsampled images:



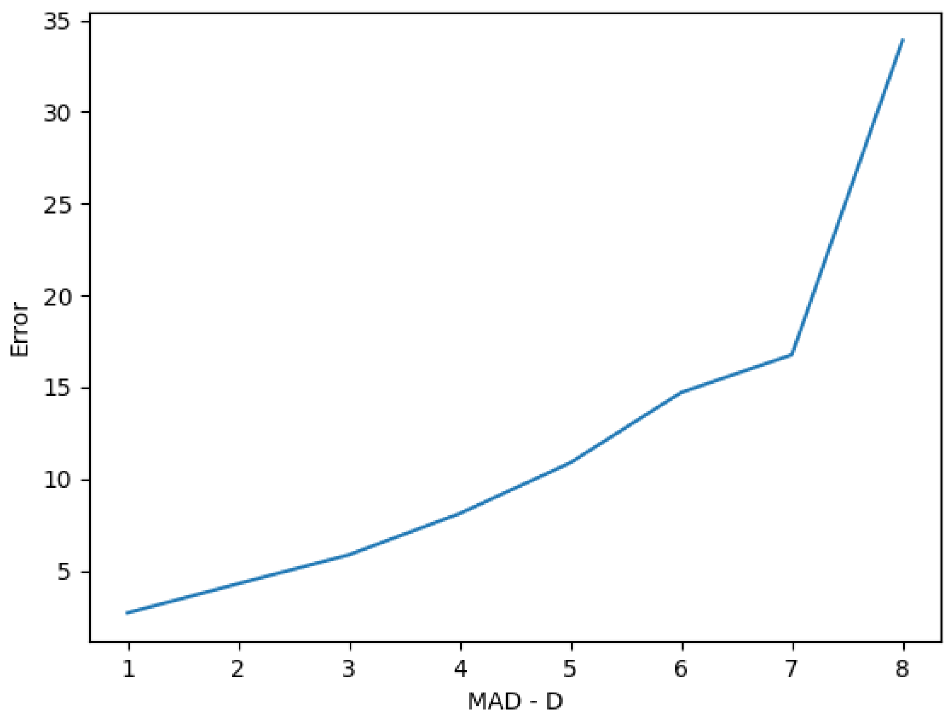

MSE by D:



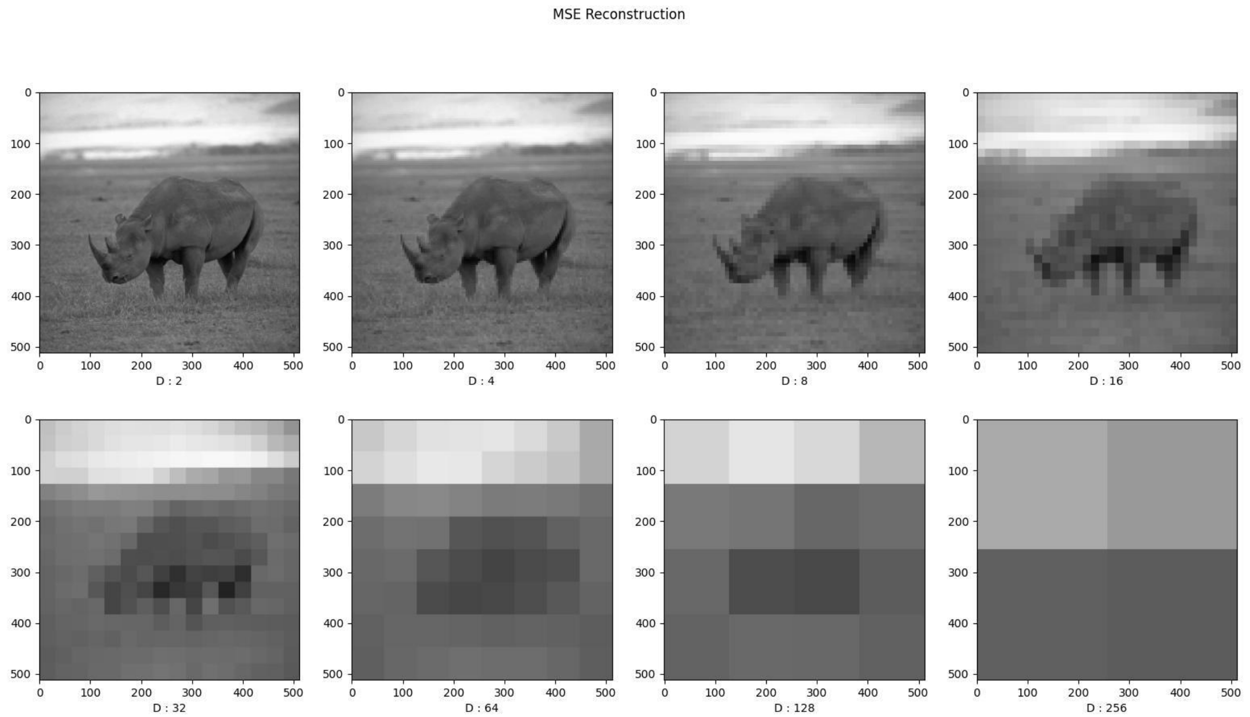
2.2) MAD subsampled images:



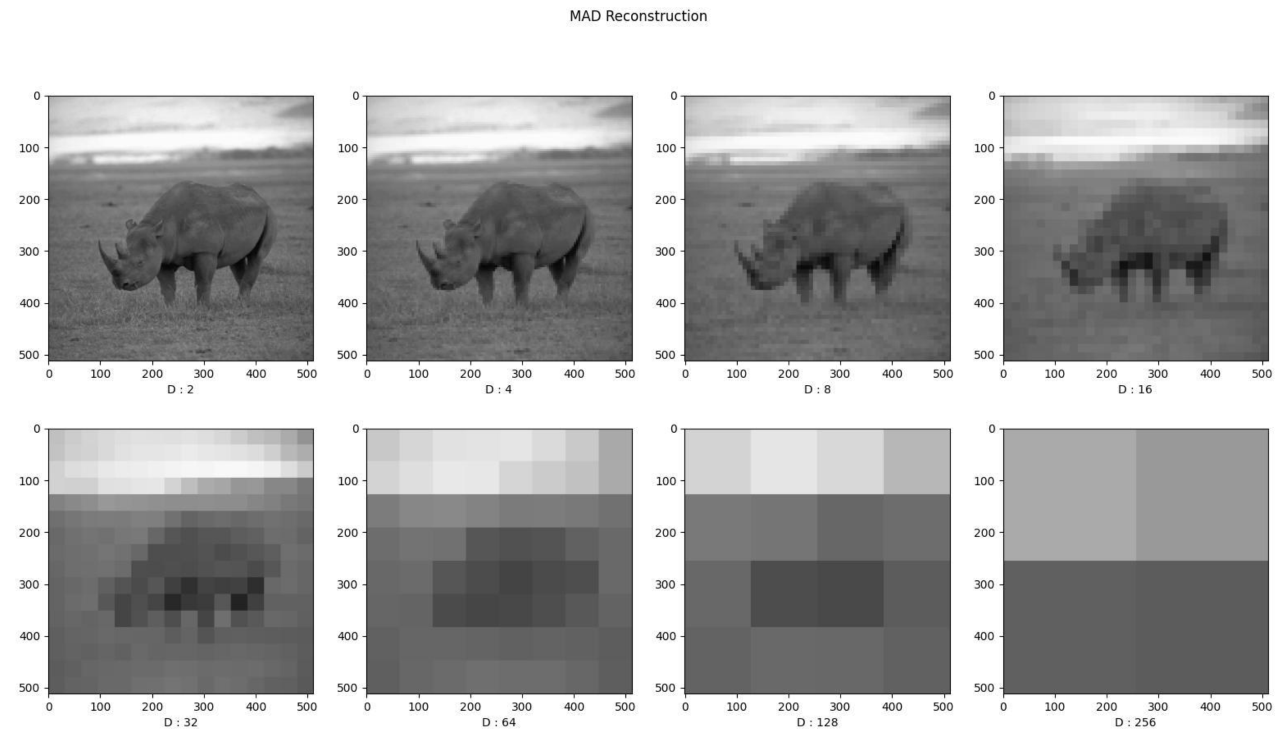

MAD by D:



## 2.2) Reconstruction from MSE subsampled images:




## Reconstruction from MAD subsampled images:



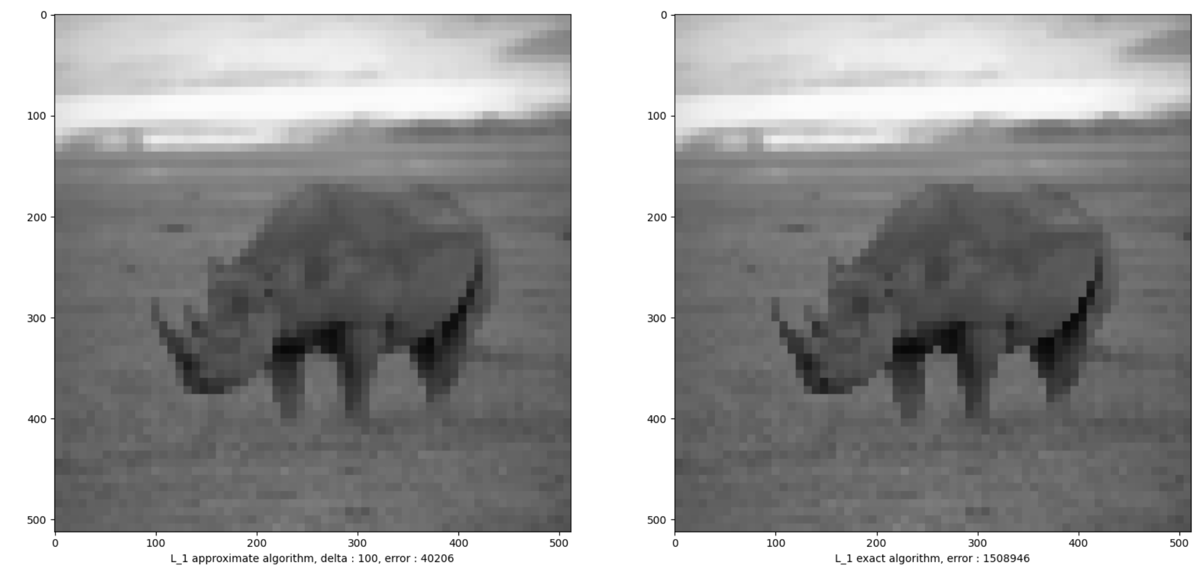
2.3) As the D grows, the error grows, which can be qualitatively seen by the blurriness of the pictures. The errors are significantly different but this makes sense as MSE creates an error raised by an exponent while MAD is the error between the median and the pixel value, which would understandably be lower. There seems to be no noticeable difference to the human eye between the two approaches. Both do quite a good job of subsampling and reconstructing the image, with reconstructed images that are recognizable up till around D = 16.

3) Solving the $L^p$ problem using the $L^2$ solution

3.1) The algorithm is inserted below

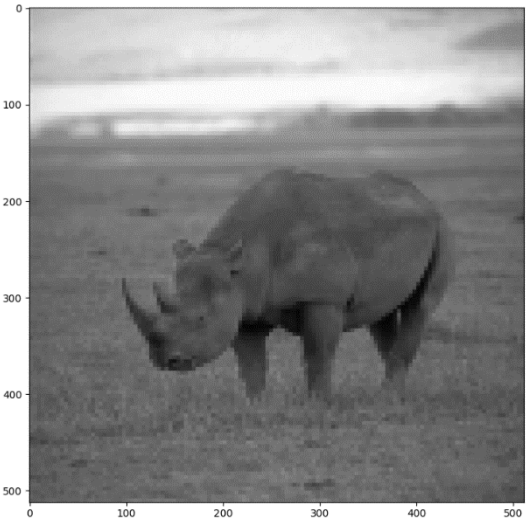
3.4) Comparison between $L^1$ exact algorithm and $L^1$ approximate algorithm, with multiple Ns and deltas.

Uniform sampling on a square grid of size: 64X64

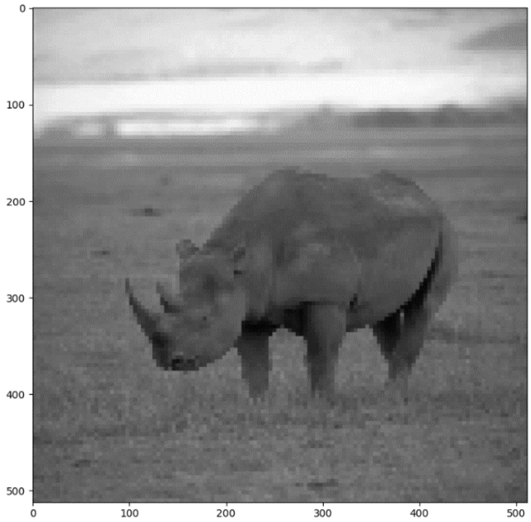



Uniform sampling on a square grid of size: 128X128

Uniform sampling on a square grid of size: 128X128

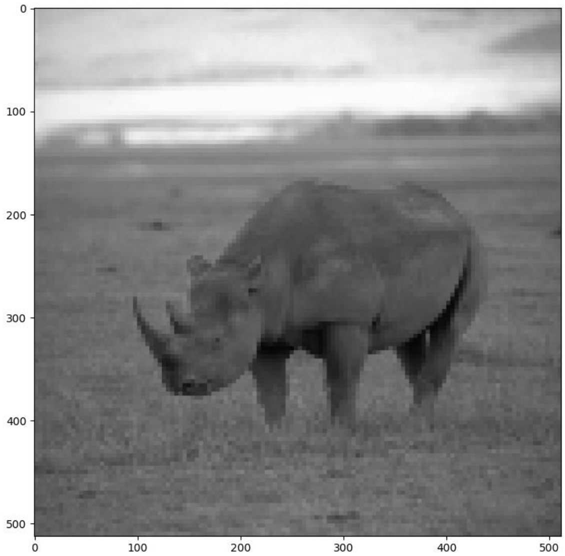


L_1 approximate algorithm, delta : 10, error : 54943

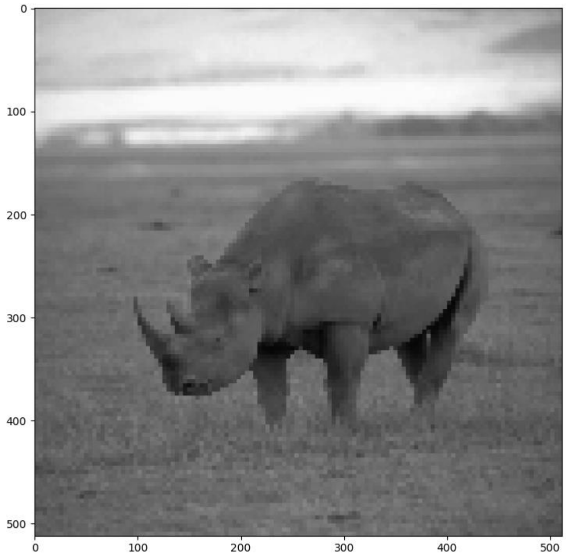


L_1 exact algorithm, error : 1097916

Uniform sampling on a square grid of size: 128X128

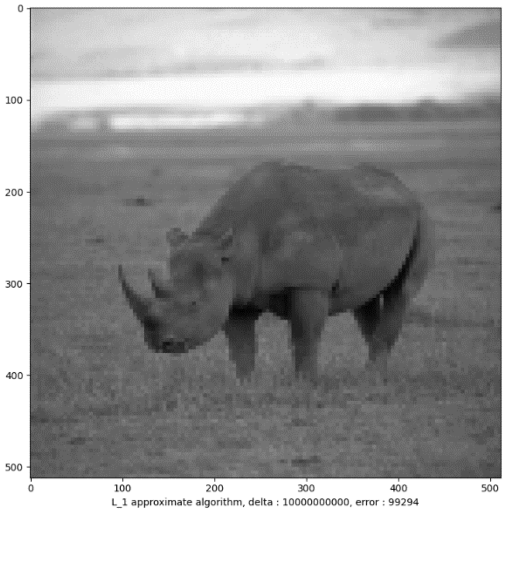


L_1 approximate algorithm, delta : 100, error : 55011

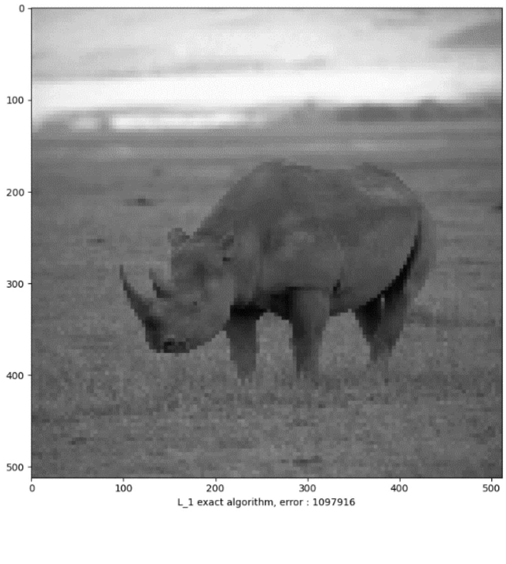


L_1 exact algorithm, error : 1097916

Uniform sampling on a square grid of size: 128X128



L_1 approximate algorithm, delta : 10000000000, error : 99294



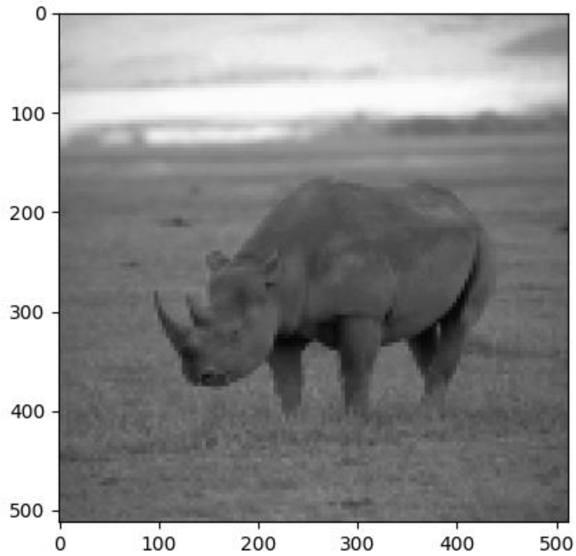
L_1 exact algorithm, error : 1097916

Uniform sampling on a square grid of size: 256X256


Uniform sampling on a square grid of size: 256X256

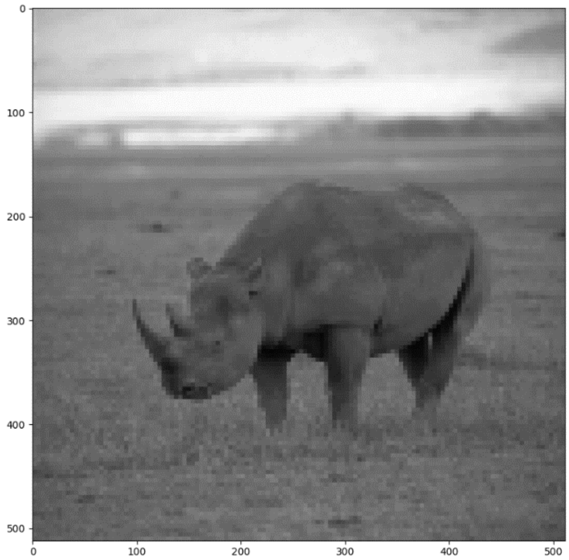

From a qualitive perspective there is very little difference between the two approaches. Both do quite an effective job at quantizing the signal. From a quantitative point of view, the L_1 exact algorithm is quite a bit faster, as it goes over the image exactly once. In the approximation the process is iterative and can take quite a bit longer because of this. The error for the approximate approach is lower, but as that does not translate into any qualitive result we don't believe this produce much benefit.

3.5)



L_1.5 approximate algorithm, delta : 100, error : 55158, N : 128

L_1.5 approximate algorithm, delta : 100, error : 76121, N : 256



L_4 approximate algorithm, delta : 100, error : 44367984557117936, N : 128

L_4 approximate algorithm, delta : 100, error : 2160497091635806, N : 256

We see that although the errors between the two are significantly different, again the qualitive difference is not very noticeable if at all. This makes sense as the error difference is a result of the large exponential factor difference. Both options result in quite good quantizations.

1) in order to generlize the pseudo code We will use the assumption in 1 that each $f_i^p$ is independant, and thus can be optimized individualy.

Init: ~~ass~~ initialize a weight-matrix for the signal f to 1. divide the image into NXN square grid. for each grid sample the pixels and generate a constant $\hat{f}_i$ function based on general $L^2$ solution

$$\hat{f}_i = \frac{\int_{\Delta_i} f(x)\, W(x)\, dx}{\int_{\Delta} W(x)\, dx} = \frac{\sum\sum f(x) W(x)}{\sum\sum W(x)}$$ avg weighted mean of pixel values in grid.

$$\text{error} = \iint_{[0,1]^2} W \left| f(x_1,x_2) - \hat{f}(x_1,x_2) \right|^p dx_1 dx_2$$

$$= \sum_1^N \sum_1^N \left| f(x_1,x_2) - \hat{f}(x_1,x_2) \right|^p w(x_1,x_2)$$ - sum of all weighted p errors

(or assume they are given)

$$W_i' = \min\left\{ W(x) \left| f_i(x) - \hat{f}_i(x) \right|^{p-2}, 1/\varepsilon \right\}$$

repeats:

$$\hat{f}_i^{next} = \underset{\hat{f}_i}{\arg\min} \int_{[0,1]^2} \left| f_i(x) - \hat{f}_i(x) \right|^2 W_i'(x)\, dx$$

$$= \underset{\hat{f}_i}{\arg\min} \sum^N \sum^N \left| f_i(x_1 x_2) - \hat{f}_i(x_1,x_2) \right|^2 W_i(x_1,x_2)$$

= piecewise const function over the grid in each grid $\hat{f}_i^{next}$ is as defined above

$$f_i = f_i^{next}$$

$$W'(x_1,x_2) = \min\left\{ W_i(x) \left| f(x) - \hat{f}_i \right|^{p-2}, 1/\varepsilon \right\}$$

prev_error = error

error = as above with updated vals

stop condition error - prev error $< \delta$