ECE 565 - Project 8 - Wavelet denoising

An Vuong, Tingwei Zhang

December 13, 2023

Abstract

This document contains the report for Wavelet denoising project. It contains derivations for the CRLB, MLE, MAP, MMSE and the Marginal ML estimators. It also contains some corresponding simulations and examples.

Code is available at https://github.com/anvuongb/ece565_final

1 Deriving pdf

We have the following setup:

$$x(n) = s(n) + v(n) \tag{1}$$

$$f_S(s) = \frac{1}{2\gamma} e^{-\frac{|s|}{\gamma}} \tag{2}$$

$$f_V(v) \sim \mathcal{N}(0, \sigma^2) \tag{3}$$

$$\begin{aligned} \mathbf{y} &= [x_1, x_2, \dots, x_N]^T \quad (\textit{observations}) \\ \boldsymbol{\theta} &= [s_1, s_2, \dots, s_N]^T \quad (\textit{parameters}) \end{aligned}$$

1.1 Consider $s(n)$ as deterministic unknown

In this case, $x(n)$ will be a Gaussian with mean s and variance σ^2

$$f_x(x_i|s_i) = \mathcal{N}(s_i, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}} \tag{4}$$

Hence, with iid assumption:

$$f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}} \tag{5}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}^N} e^{-\sum_{i=1}^{N} \frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}} \tag{6}$$

1.2 Both $s(n)$ and $v(n)$ are random variable

In this case, $f_X(x)$ is the convolution between $f_S(s)$ and $f_V(v)$

$$f_X(x) = f_S(s) \times f_V(v) \tag{7}$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\gamma} e^{-\frac{|\tau|}{\gamma}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\tau)^2}{\sigma^2}} d\tau \tag{8}$$

$$= \int_{-\infty}^{0} \frac{1}{2\gamma} e^{\frac{\tau}{\gamma}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\tau)^2}{\sigma^2}} d\tau + \int_{0}^{\infty} \frac{1}{2\gamma} e^{-\frac{\tau}{\gamma}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\tau)^2}{\sigma^2}} d\tau \tag{9}$$

The first integral is given as:

$$\int_{-\infty}^{0} \frac{1}{2\gamma} e^{\frac{\tau}{\gamma}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\tau)^2}{\sigma^2}} d\tau = \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} \int_{-\infty}^{0} e^{-\frac{1}{2}\frac{(x-\tau)^2}{\sigma^2}+\frac{\tau}{\gamma}} d\tau \tag{10}$$

$$= \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} \int_{-\infty}^{0} e^{-\frac{1}{2\sigma^2}(\tau^2+x^2-2\tau x-\frac{2\tau\sigma^2}{\gamma})} d\tau \tag{11}$$

$$= \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} \int_{-\infty}^{0} e^{-\frac{1}{2\sigma^2}(\tau^2+x^2-2\tau(x+\frac{\sigma^2}{\gamma}))} d\tau \tag{12}$$

$$= \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} \int_{-\infty}^{0} e^{-\frac{1}{2\sigma^2}(\tau^2-2\tau(x+\frac{\sigma^2}{\gamma})+(x+\frac{\sigma^2}{\gamma})^2-2\frac{x\sigma^2}{\gamma}-\frac{\sigma^4}{\gamma^2})} d\tau \tag{13}$$

$$= \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} \int_{-\infty}^{0} e^{-\frac{1}{2\sigma^2}\left(\tau-(x+\frac{\sigma^2}{\gamma})\right)^2+\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} d\tau \tag{14}$$

$$= \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} e^{\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \underbrace{\int_{-\infty}^{0} e^{-\frac{1}{2\sigma^2}\left(\tau-(x+\frac{\sigma^2}{\gamma})\right)^2} d\tau}_{\text{can be represented using } \mathbf{\Phi}(.)} \tag{15}$$

Similarly,

$$\int_{0}^{\infty} \frac{1}{2\gamma} e^{-\frac{\tau}{\gamma}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\tau)^2}{\sigma^2}} d\tau = \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} e^{-\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{0}^{\infty} e^{-\frac{1}{2\sigma^2}\left(\tau-(x-\frac{\sigma^2}{\gamma})\right)^2} d\tau \tag{16}$$

Thus,

$$f_X(x) = \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} e^{\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{-\infty}^{0} e^{-\frac{1}{2\sigma^2}\left(\tau-(x+\frac{\sigma^2}{\gamma})\right)^2} d\tau + \frac{1}{2\gamma\sqrt{2\pi\sigma^2}} e^{-\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{0}^{\infty} e^{-\frac{1}{2\sigma^2}\left(\tau-(x-\frac{\sigma^2}{\gamma})\right)^2} d\tau \tag{17}$$

$$= \frac{1}{2\gamma} e^{\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(\tau-(x+\frac{\sigma^2}{\gamma})\right)^2} d\tau + \frac{1}{2\gamma} e^{-\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left(\tau-(x-\frac{\sigma^2}{\gamma})\right)^2} d\tau \tag{18}$$

$$= \frac{1}{2\gamma} e^{\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \mathbf{\Phi}\big(\frac{-x-\sigma^2/\gamma}{\sigma}\big) + \frac{1}{2\gamma\sigma^2} e^{-\frac{x}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \mathbf{\Phi}\big(\frac{x-\sigma^2/\gamma}{\sigma}\big) \tag{19}$$

$$f_X(x_i) = \frac{1}{2\gamma} e^{\frac{\sigma^2}{2\gamma^2}} \left(e^{\frac{x_i}{\gamma}} \mathbf{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \mathbf{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big) \right) \tag{20}$$

where $\mathbf{\Phi}(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\tau^2} d\tau$

Hence, with iid assumption:

$$f_{\mathbf{Y}}(\mathbf{y}|\gamma,\sigma) = \prod_{i=1}^{N} f_X(x_i|\gamma,\sigma) \tag{21}$$

$$= \prod_{i=1}^{N} \frac{1}{2\gamma} e^{\frac{\sigma^2}{2\gamma^2}} \left(e^{\frac{x_i}{\gamma}} \mathbf{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \mathbf{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big) \right) \tag{22}$$

2 CRLB for deterministic unknown $s(n)$

$$\log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta}) = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma^2) - \frac{\sum_{i=1}^{N}(x_i-s_i)^2}{2\sigma^2} \tag{23}$$

$$\frac{d\log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})}{d\boldsymbol{\theta}} = -\Big[\frac{(s_1-x_1)}{\sigma^2}, \frac{(s_2-x_2)}{\sigma^2}, \ldots, \frac{(s_N-x_N)}{\sigma^2}\Big]^T \tag{24}$$

$$\frac{d^2\log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})}{d\boldsymbol{\theta}d\boldsymbol{\theta}^T} = -\begin{bmatrix} \frac{1}{\sigma^2} & 0 & \ldots & 0 \\ 0 & \frac{1}{\sigma^2} & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \frac{1}{\sigma^2} \end{bmatrix} \tag{25}$$

$$\text{FIM} = \mathbb{E}\Big[-\frac{d^2\log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})}{d\boldsymbol{\theta}d\boldsymbol{\theta}^T}\Big] = \begin{bmatrix} \frac{1}{\sigma^2} & 0 & \ldots & 0 \\ 0 & \frac{1}{\sigma^2} & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \frac{1}{\sigma^2} \end{bmatrix} \tag{26}$$

Hence the CRLB for each estimate of s_i is the same:

$$\boxed{\text{CRLB}_i = (\text{FIM}_{ii})^{-1} = \sigma^2} \tag{27}$$

3 MLE for deterministic unknown $s(n)$

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta}) \tag{28}$$

$$\frac{d \log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})}{d\boldsymbol{\theta}} = \left[\frac{(s_1 - x_1)}{\sigma^2}, \frac{(s_2 - x_2)}{\sigma^2}, \dots, \frac{(s_N - x_N)}{\sigma^2}\right]^T = \mathbf{0} \tag{29}$$

$$\Rightarrow \boxed{\hat{s}_{i,ML} = x_i} \tag{30}$$

For completeness, we have $\frac{d^2 \log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})}{d\boldsymbol{\theta} d\boldsymbol{\theta}^T} = -\text{ddiag}(\frac{1}{\sigma^2}, \dots, \frac{1}{\sigma^2}) \preccurlyeq \mathbf{0}$, the maximum is indeed achieved.

4 MAP and MMSE estimators

4.1 MAP

When s is a stochastic unknown, $f(x_i|s_i) = \mathcal{N}(s_i, \sigma^2)$. Hence $f(x_i|s_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}}$. Given s_i follows $f_S(s_i) = \frac{1}{2\gamma} e^{-\frac{|s_i|}{\gamma}}$, the posterior is:

$$f(s_i|x_i) = \frac{f(x_i|s_i) f(s_i)}{f(x_i)} \tag{31}$$

$$= \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}} \frac{1}{2\gamma} e^{-\frac{|s_i|}{\gamma}}}{\int_{-\infty}^{\infty} f(x_i|\vartheta_i) f(\vartheta_i) d\vartheta_i} \tag{32}$$

The denominator of (32) does not depend on s_i, so:

$$\hat{s}_{i,MAP} = \arg\max_{s_i} f(s_i|x_i) = \arg\max_{s_i} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}} \frac{1}{2\gamma} e^{-\frac{|s_i|}{\gamma}} \tag{33}$$

$$= \arg\max_{s_i} \log \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2}} \frac{1}{2\gamma} e^{-\frac{|s_i|}{\gamma}} \quad (\textit{monotonicity of } \log) \tag{34}$$

$$= \arg\max_{s_i} -\frac{1}{2}\frac{(x_i - s_i)^2}{\sigma^2} - \frac{|s_i|}{\gamma} \tag{35}$$

$$= \arg\min_{s_i} \frac{1}{2}(x_i - s_i)^2 + \frac{\sigma^2}{\gamma}|s_i| \tag{36}$$

Since the observations are iid, (36) holds for all observations. Equation (36) is in the form of L1-regularized Least Squares, and optimization can be performed independently for each s_i. Let $l(s_i) = \frac{1}{2}(x_i - s_i)^2 + \frac{\sigma^2}{\gamma}|s_i|$.

When $s_i \geq 0$:

$$l(s_i) = \frac{1}{2}(x_i - s_i)^2 + \frac{\sigma^2}{\gamma} s_i \tag{37}$$

$$\frac{dl(s_i)}{ds_i} = s_i - x_i + \frac{\sigma^2}{\gamma} = 0 \tag{38}$$

$$\Rightarrow s_i = \begin{cases} x_i - \frac{\sigma^2}{\gamma}, & x_i - \frac{\sigma^2}{\gamma} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{39}$$

When $s_i < 0$:

$$l(s_i) = \frac{1}{2}(x_i - s_i)^2 - \frac{\sigma^2}{\gamma} s_i \tag{40}$$

$$\frac{dl(s_i)}{ds_i} = s_i - x_i - \frac{\sigma^2}{\gamma} = 0 \tag{41}$$

$$\Rightarrow s_i = \begin{cases} x_i + \frac{\sigma^2}{\gamma}, & x_i + \frac{\sigma^2}{\gamma} < 0 \\ 0, & \text{otherwise} \end{cases} \tag{42}$$

From (39)(42), we have the MAP estimator for s_i:

$$\boxed{\hat{s}_{i,MAP} = \begin{cases} x_i - \frac{\sigma^2}{\gamma}, & x_i > \frac{\sigma^2}{\gamma} \\ 0, & |x_i| < \frac{\sigma^2}{\gamma} \\ x_i + \frac{\sigma^2}{\gamma}, & x_i < -\frac{\sigma^2}{\gamma} \end{cases}} \tag{43}$$

For completeness, the second derivatives of $l(s_i)$ are 1 in both cases, hence the minimum is achieved.

4.2 MMSE

The MMSE of s is the expected value of s given the observations. Since observations are iid, each s_i can be estimated as:

$$\hat{s}_{iMMSE} = \mathbb{E}[s_i|x_i] \tag{44}$$

$$= \int_{-\infty}^{\infty} s_i f(s_i|x_i) ds_i \tag{45}$$

$$= \int_{-\infty}^{\infty} s_i \frac{f(x_i|s_i)f(s_i)}{f(x_i)} ds_i \tag{46}$$

$$= \int_{-\infty}^{\infty} \vartheta \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} \frac{1}{2\gamma} e^{-\frac{|\vartheta|}{\gamma}}}{\frac{1}{2\gamma} e^{\frac{\sigma^2}{2\gamma^2}} \left(e^{\frac{x_i}{\gamma}} \mathbf{\Phi}\left(\frac{-x_i-\sigma^2/\gamma}{\sigma}\right) + e^{-\frac{x_i}{\gamma}} \mathbf{\Phi}\left(\frac{x_i-\sigma^2/\gamma}{\sigma}\right)\right)} d\vartheta \tag{47}$$

$$= \frac{1}{\frac{1}{2\gamma} e^{\frac{\sigma^2}{2\gamma^2}} \left(e^{\frac{x_i}{\gamma}} \mathbf{\Phi}\left(\frac{-x_i-\sigma^2/\gamma}{\sigma}\right) + e^{-\frac{x_i}{\gamma}} \mathbf{\Phi}\left(\frac{x_i-\sigma^2/\gamma}{\sigma}\right)\right)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{2\gamma} \vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} e^{-\frac{|\vartheta|}{\gamma}} d\vartheta \tag{48}$$

Consider the integral:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{2\gamma} \int_{-\infty}^{\infty} \vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} e^{-\frac{|\vartheta|}{\gamma}} d\vartheta = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{2\gamma} \int_{-\infty}^{0} \vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} e^{\frac{\vartheta}{\gamma}} d\vartheta + \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{2\gamma} \int_{0}^{\infty} \vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} e^{-\frac{\vartheta}{\gamma}} d\vartheta \tag{49}$$

Similar to section 1, the integral in the right can be rewritten as:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{2\gamma} \int_{0}^{\infty} \vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} e^{-\frac{\vartheta}{\gamma}} d\vartheta = \frac{1}{2\gamma} e^{-\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \vartheta e^{-\frac{1}{2\sigma^2}\left(\vartheta - (x_i - \frac{\sigma^2}{\gamma})\right)^2} d\vartheta \tag{50}$$

$$= \frac{1}{2\gamma} e^{-\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \int_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \left(z\sigma + (x_i - \frac{\sigma^2}{\gamma})\right) e^{-\frac{1}{2}z^2} \sigma dz \;\; (\textit{let } z = \frac{\vartheta - (x_i - \frac{\sigma^2}{\gamma})}{\sigma}) \tag{51}$$

$$= \frac{1}{2\gamma} e^{-\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \left(\int_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} z\sigma e^{-\frac{1}{2}z^2} dz + (x_i - \sigma^2/\gamma) \int_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \right) \tag{52}$$

$$= \frac{1}{2\gamma} e^{-\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \left(\int_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \frac{\sigma}{\sqrt{2\pi}} z e^{-\frac{1}{2}z^2} dz + (x_i - \sigma^2/\gamma) \mathbf{\Phi}\left(\frac{x_i - \sigma^2/\gamma}{\sigma}\right) \right) \tag{53}$$

Consider the integral:

$$\int_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \frac{\sigma}{\sqrt{2\pi}} z e^{-\frac{1}{2}z^2} dz = -\frac{\sigma}{\sqrt{2\pi}} \int_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \frac{d}{dz} e^{-\frac{1}{2}z^2} dz \tag{54}$$

$$= -\frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \Big|_{-(x_i-\frac{\sigma^2}{\gamma})/\sigma}^{\infty} \;\; (\textit{Fundamental Theorem of Calculus}) \tag{55}$$

$$= -\frac{\sigma}{\sqrt{2\pi}} (0 - e^{-\frac{1}{2\sigma^2}(x_i - \frac{\sigma^2}{\gamma})^2}) \tag{56}$$

$$= \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i - \frac{\sigma^2}{\gamma})^2} \tag{57}$$

Hence,

$$\int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{2\gamma} \vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}} e^{-\frac{\vartheta}{\gamma}} d\vartheta = \frac{1}{2\gamma} e^{-\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}} \left(\frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\frac{x_i}{\sigma} - \frac{\sigma}{\gamma})^2} + (x_i - \sigma^2/\gamma) \mathbf{\Phi}\left(\frac{x_i - \sigma^2/\gamma}{\sigma}\right) \right) \tag{58}$$

In a similar manner:

$$\int_{-\infty}^{0}\frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{2\gamma}\vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}}e^{\frac{\vartheta}{\gamma}}d\vartheta = \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\int_{-\infty}^{0}\vartheta\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}\left(\vartheta-(x_i+\frac{\sigma^2}{\gamma})\right)^2}d\vartheta \tag{59}$$

$$= \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\int_{-\infty}^{-(x_i+\frac{\sigma^2}{\gamma})/\sigma}\frac{1}{\sqrt{2\pi\sigma^2}}(z\sigma+(x_i+\frac{\sigma^2}{\gamma}))e^{-\frac{1}{2}z^2}\sigma dz \quad (let\ z=\frac{\vartheta-(x_i+\frac{\sigma^2}{\gamma})}{\sigma}) \tag{60}$$

$$= \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(\int_{-\infty}^{-(x_i+\frac{\sigma^2}{\gamma})/\sigma}\frac{\sigma}{\sqrt{2\pi}}ze^{-\frac{1}{2}z^2}dz+(x_i+\frac{\sigma^2}{\gamma})\int_{-\infty}^{-(x_i+\frac{\sigma^2}{\gamma})/\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz\right) \tag{61}$$

$$= \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(\int_{-\infty}^{-(x_i+\frac{\sigma^2}{\gamma})}\frac{\sigma}{\sqrt{2\pi}}ze^{-\frac{1}{2}z^2}dz+(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right) \tag{62}$$

$$= \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(-\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{-(x_i+\frac{\sigma^2}{\gamma})}\frac{d}{dz}e^{-\frac{1}{2}z^2}dz+(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right) \tag{63}$$

$$= \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(-\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\Big|_{-\infty}^{-(x_i+\frac{\sigma^2}{\gamma})}+(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right) \tag{64}$$

$$= \frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(\frac{\sigma}{\sqrt{2\pi}}(0-e^{-\frac{1}{2\sigma^2}(x_i+\frac{\sigma^2}{\gamma})^2})+(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right) \tag{65}$$

$$= -\frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i+\frac{\sigma^2}{\gamma})^2}-(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right) \tag{66}$$

Thus,

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{2\gamma}\vartheta e^{-\frac{1}{2}\frac{(x_i-\vartheta)^2}{\sigma^2}}e^{-\frac{|\vartheta|}{\gamma}}d\vartheta = \frac{1}{2\gamma}e^{-\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i-\frac{\sigma^2}{\gamma})^2}+(x_i-\sigma^2/\gamma)\mathbf{\Phi}(\frac{x_i-\sigma^2/\gamma}{\sigma})\right) \tag{67}$$

$$-\frac{1}{2\gamma}e^{\frac{x_i}{\gamma}+\frac{\sigma^2}{2\gamma^2}}\left(\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i+\frac{\sigma^2}{\gamma})^2}-(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right) \tag{68}$$

By plugging (68) into (48), we obtain the MMSE estimator for s_i.

$$\boxed{\hat{s}_{i,MMSE} = \frac{e^{-\frac{x_i}{\gamma}}\left(\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i-\frac{\sigma^2}{\gamma})^2}+(x_i-\sigma^2/\gamma)\mathbf{\Phi}(\frac{x_i-\sigma^2/\gamma}{\sigma})\right)-e^{\frac{x_i}{\gamma}}\left(\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i+\frac{\sigma^2}{\gamma})^2}-(x_i+\sigma^2/\gamma)\mathbf{\Phi}(-\frac{x_i+\sigma^2/\gamma}{\sigma})\right)}{e^{-\frac{x_i}{\gamma}}\mathbf{\Phi}\left(\frac{x_i-\sigma^2/\gamma}{\sigma}\right)+e^{\frac{x_i}{\gamma}}\mathbf{\Phi}\left(-\frac{x_i+\sigma^2/\gamma}{\sigma}\right)}} \tag{69}$$

From (69), we can observe that as γ approaches infinity, the equation approaches:

$$\frac{x_i\mathbf{\Phi}(\frac{x_i}{\sigma})-x_i\mathbf{\Phi}(-\frac{x_i}{\sigma})}{\mathbf{\Phi}(\frac{x_i}{\sigma})+\mathbf{\Phi}(-\frac{x_i}{\sigma})} \tag{70}$$

For large values of x_i, whose probability increases as $\gamma\to\infty$, this means (70) behaves as indicator functions $x_iI(x_i\geq 0)-x_iI(x_i\leq 0)$, which is also the MAP estimators for large values of γ. Thus, we should expect MMSE to perform similarly to MAP for large values of γ.

5 Marginal ML

In this section, instead of s_i, we want to estimate γ and σ. From the first section, we have:

$$f_{\mathbf{Y}}(\mathbf{y}|\gamma,\sigma) = (\frac{1}{2\gamma}e^{\frac{\sigma^2}{2\gamma^2}})^N\prod_{i=1}^{N}\left(e^{\frac{x_i}{\gamma}}\mathbf{\Phi}(\frac{-x_i-\sigma^2/\gamma}{\sigma})+e^{-\frac{x_i}{\gamma}}\mathbf{\Phi}(\frac{x_i-\sigma^2/\gamma}{\sigma})\right) \tag{71}$$

$$\log f_{\mathbf{Y}}(\mathbf{y}|\gamma,\sigma) = N\log(\frac{1}{2\gamma})+N\frac{\sigma^2}{2\gamma^2}+\sum_{i=1}^{N}\log\left(e^{\frac{x_i}{\gamma}}\mathbf{\Phi}(\frac{-x_i-\sigma^2/\gamma}{\sigma})+e^{-\frac{x_i}{\gamma}}\mathbf{\Phi}(\frac{x_i-\sigma^2/\gamma}{\sigma})\right) \tag{72}$$

Differentiating with respect to γ, we have:

$$\frac{\partial \log f}{\partial \gamma} = -\frac{N}{\gamma} - \frac{N\sigma^2}{\gamma^3} + \sum_{i=1}^{N} \frac{\frac{\partial}{\partial \gamma}\big(e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big)\big) + \frac{\partial}{\partial \gamma}\big(e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)}{e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)\big)} \tag{73}$$

$$\frac{\partial}{\partial \gamma}\big(e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big)\big) = \frac{\partial e^{\frac{x_i}{\gamma}}}{\partial \gamma} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{\frac{x_i}{\gamma}} \frac{\partial \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big)}{\partial \gamma} \tag{74}$$

$$= -\frac{x_i}{\gamma^2} e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{\frac{x_i}{\gamma}} \frac{\partial}{\partial \gamma} \int_{-\infty}^{\frac{-x_i-\sigma^2/\gamma}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \tag{75}$$

$$= -\frac{x_i}{\gamma^2} e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{\frac{x_i}{\gamma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{-x_i-\sigma^2/\gamma}{\sigma})^2} \frac{\partial}{\partial \gamma} \frac{-x_i-\sigma^2/\gamma}{\sigma} \quad (\textit{Leibniz's integral rule}) \tag{76}$$

$$= -\frac{x_i}{\gamma^2} e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + \frac{\sigma}{\gamma^2} e^{\frac{x_i}{\gamma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{-x_i-\sigma^2/\gamma}{\sigma})^2} \tag{77}$$

Note: in (76), differentiation of the function at ∞ is undefined, but $e^{-\infty}$ tends to zero so we assume the differentiation with respect to the lower bound is zero.

Similarly,

$$\frac{\partial}{\partial \gamma}\big(e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)\big) = \frac{\partial e^{-\frac{x_i}{\gamma}}}{\partial \gamma} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \frac{\partial \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)}{\partial \gamma} \tag{78}$$

$$= \frac{x_i}{\gamma^2} e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big) + \frac{\sigma}{\gamma^2} e^{-\frac{x_i}{\gamma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_i-\sigma^2/\gamma}{\sigma})^2} \tag{79}$$

Thus,

$$\boxed{\frac{\partial \log f}{\partial \gamma} = -\frac{N}{\gamma} - \frac{N\sigma^2}{\gamma^3} + \sum_{i=1}^{N} \frac{\frac{x_i}{\gamma^2} e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big) + \frac{\sigma}{\gamma^2} e^{-\frac{x_i}{\gamma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_i-\sigma^2/\gamma}{\sigma})^2} - \frac{x_i}{\gamma^2} e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + \frac{\sigma}{\gamma^2} e^{\frac{x_i}{\gamma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{-x_i-\sigma^2/\gamma}{\sigma})^2}}{e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)\big)}} \tag{80}$$

Next, differentiating with respect to σ:

$$\frac{\partial \log f}{\partial \sigma} = \frac{N\sigma}{\gamma^2} + \sum_{i=1}^{N} \frac{\frac{\partial}{\partial \sigma}\big(e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big)\big) + \frac{\partial}{\partial \sigma}\big(e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)}{e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)\big)} \tag{81}$$

$$\frac{\partial}{\partial \sigma}\big(e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big)\big) = e^{\frac{x_i}{\gamma}} \frac{\partial}{\partial \sigma} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) \tag{82}$$

$$= e^{\frac{x_i}{\gamma}} \frac{\partial}{\partial \sigma} \int_{-\infty}^{\frac{-x_i-\sigma^2/\gamma}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \tag{83}$$

$$= e^{\frac{x_i}{\gamma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{-x_i-\sigma^2/\gamma}{\sigma})^2} \frac{\partial}{\partial \sigma} \frac{-x_i-\sigma^2/\gamma}{\sigma} \tag{84}$$

$$= \frac{1}{\sqrt{2\pi}}\big(\frac{x_i}{\sigma^2} - \frac{1}{\gamma}\big) e^{\frac{x_i}{\gamma}} e^{-\frac{1}{2}(\frac{-x_i-\sigma^2/\gamma}{\sigma})^2} \tag{85}$$

$$\frac{\partial}{\partial \sigma}\big(e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)\big) = -\frac{1}{\sqrt{2\pi}}\big(\frac{x_i}{\sigma^2} + \frac{1}{\gamma}\big) e^{-\frac{x_i}{\gamma}} e^{-\frac{1}{2}(\frac{x_i-\sigma^2/\gamma}{\sigma})^2} \tag{86}$$

Thus,

$$\boxed{\frac{\partial \log f}{\partial \sigma} = \frac{N\sigma}{\gamma^2} + \sum_{i=1}^{N} \frac{\frac{1}{\sqrt{2\pi}}\big(\frac{x_i}{\sigma^2} - \frac{1}{\gamma}\big) e^{\frac{x_i}{\gamma}} e^{-\frac{1}{2}(\frac{-x_i-\sigma^2/\gamma}{\sigma})^2} - \frac{1}{\sqrt{2\pi}}\big(\frac{x_i}{\sigma^2} + \frac{1}{\gamma}\big) e^{-\frac{x_i}{\gamma}} e^{-\frac{1}{2}(\frac{x_i-\sigma^2/\gamma}{\sigma})^2}}{e^{\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{-x_i-\sigma^2/\gamma}{\sigma}\big) + e^{-\frac{x_i}{\gamma}} \boldsymbol{\Phi}\big(\frac{x_i-\sigma^2/\gamma}{\sigma}\big)\big)}} \tag{87}$$

5.1 Gradient ascent to estimate γ and σ^2

From (80)(87), we can use Gradient Ascent to estimate γ and σ^2 that maximizes the Marginal ML, as shown in Algorithm 1. We also incorporated the backtracking line search into the algorithm for a more stable estimations of γ and σ. Without it, the estimations will not work for low SNR region.

Algorithm 1 Gradient Ascent for ML estimation of γ and σ^2

Require: Maximum iterations K, initial $\mathbf{v} = [\gamma_0, \sigma_0]^T$
  for $i = 1 \dots K$ **do**
    Use line search to get optimal step λ
    $\mathbf{v}_i = \mathbf{v}_{i-1} + \lambda * \frac{\text{grad}\mathbf{v}_{i-1}}{||\text{grad}\mathbf{v}_{i-1}||}$ ▷ from (80), (87)
  end for
Output: γ_K, σ_K

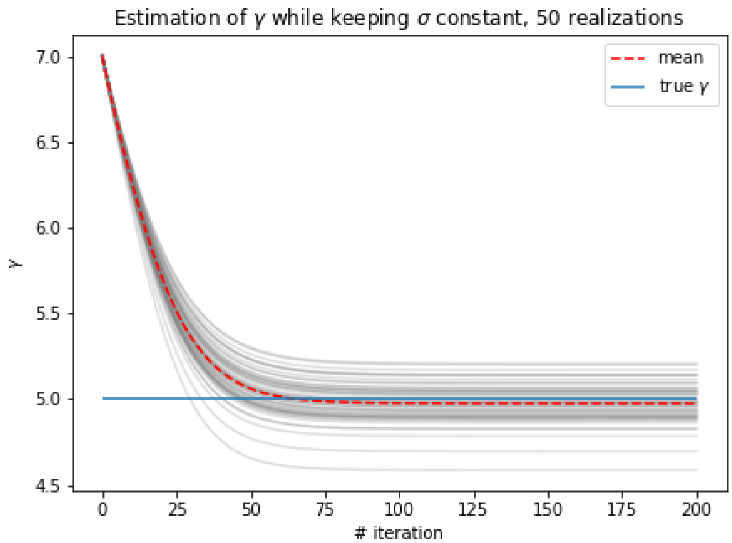

Figure 1: ML estimations of γ while keeping σ constant. Starting point is 7, the results were obtained after 50 realizations of data.

6 MSE of different estimators versus CRLB

Setup: 200 Monte Carlo runs with 2000 data points each.

6.1 Using true values of γ and σ^2

Figure 2 shows the MSE for ML, MAP, and MMSE estimators of $s(n)$. True values of γ and σ^2 were used in the simulations.

In general, all estimators approach the CRLB, which is σ^2, as γ gets larger. In the low SNR regime, MAP and MMSE both beat the CRLB, this is because MAP and MMSE are not unbiased estimators, so they do not have to adhere to the CRLB. Finally, MMSE, true to its name, is indeed the minimizer of MSE across all estimators.

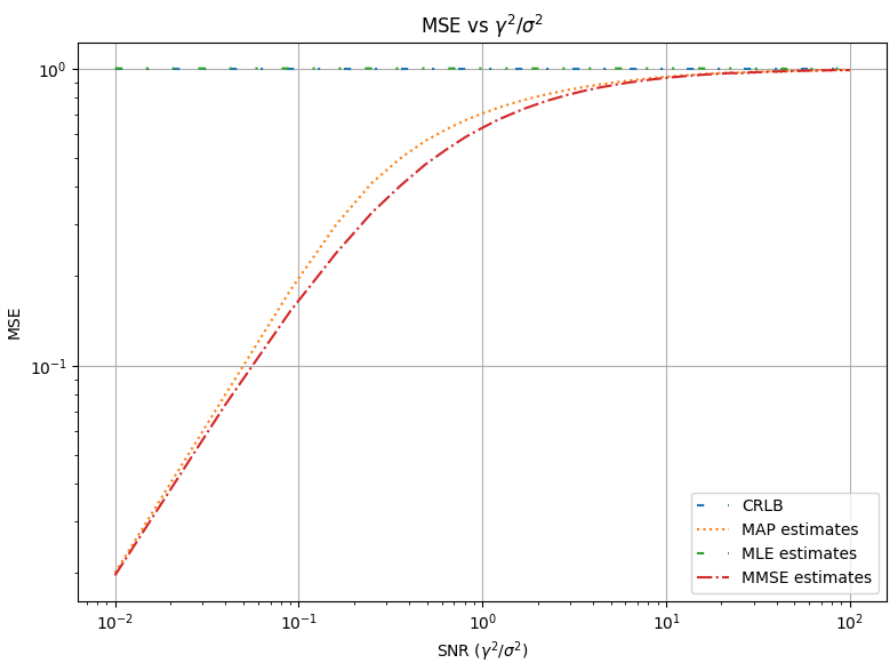

Figure 2: SNR vs MLE

6.2 γ and σ^2 are estimated by Algorithm 1

Figure 3 shows the MSE for ML, MAP, and MMSE estimators of $s(n)$. MML estimated values of γ and σ^2 were used in the simulations.

This figures tell a smilar story to that of Figure 2. The discrepancy at low and high SNR regimes is because of the noisy estimations of γ and σ. Since we are using Gradient Descent, when one parameter is much bigger than the other, the gradient in one direction will be much smaller than the other, this leads to the instability in low and high SNR regimes.

Another interesting observation is that, by doing nothing, we already achieves the CRLB. This illustrates how the frequentist approach can fail in estimating the parameters.

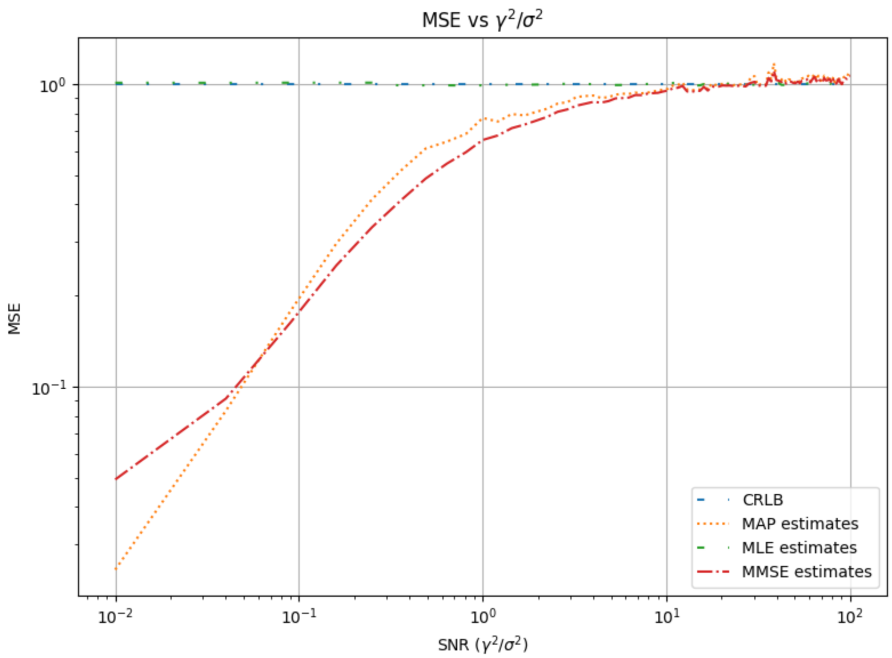

Figure 3: SNR vs MLE

7 Application to image denoising

7.1 Parameters estimation of Laplace distribution

To help with the experiment, we will need to know γ in the case of zero noise. Since we know empirically that the distribution of wavelet weights follow Laplace distribution. The ML estimator is:

$$f(\mathbf{x}|\gamma) = \prod_{i}^{N} \frac{1}{2\gamma} e^{-\frac{|x_i|}{\gamma}} \tag{88}$$

$$\log f(\mathbf{x}|\gamma) = -N \log 2\gamma - \sum_{i=1}^{N} \frac{|x_i|}{\gamma} \tag{89}$$

$$\frac{d \log f}{d\gamma} = -\frac{N}{\gamma} + \frac{1}{\gamma^2} \sum_{i=1}^{N} |x_i| = 0 \tag{90}$$

$$\Longrightarrow \frac{1}{\gamma}(1 - \frac{1}{\gamma}\frac{1}{N} \sum_{i=1}^{N} |x_i|) = 0 \tag{91}$$

$$\Longrightarrow \boxed{\hat{\gamma}_{ML} = \frac{1}{N} \sum_{i=1}^{N} |x_i|} \tag{92}$$

The FIM can be obtained by differentiating (91) one more time:

$$\mathrm{FIM} = \mathbb{E}\Big[-\frac{N}{\gamma^2} + \frac{2}{\gamma^3}\sum_{i=1}^{N}|x_i|\Big] \tag{93}$$

$$= -\frac{N}{\gamma^2} + \frac{2}{\gamma^3}\sum_{i=1}^{N}\mathbb{E}[|x_i|] = -\frac{N}{\gamma^2} + \frac{2}{\gamma^3}N\gamma = \frac{N}{\gamma^2} \tag{94}$$

$$\mathrm{CRLB} = \frac{\gamma^2}{N} \tag{95}$$

Hence $\hat{\gamma}_{ML}$ achieves CRLB because:

$$\mathrm{FIM} \times (\hat{\gamma}_{ML} - \gamma) = \frac{N}{\gamma^2}\Big(\frac{1}{N}\sum_{i=1}^{N}|x_i| - \gamma\Big) \tag{96}$$

$$= -\frac{N}{\gamma} + \frac{1}{\gamma^2}\sum_{i=1}^{N}|x_i| = \frac{d\log f}{d\gamma} \tag{97}$$

7.2 A denoising application

Figure 4: Original image

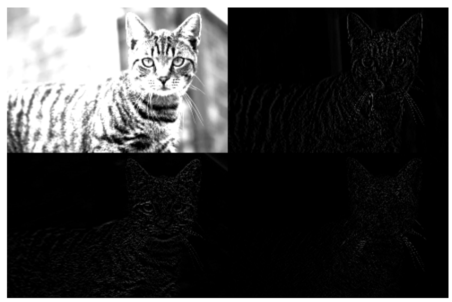

Figure 5: Haar DWT of original image

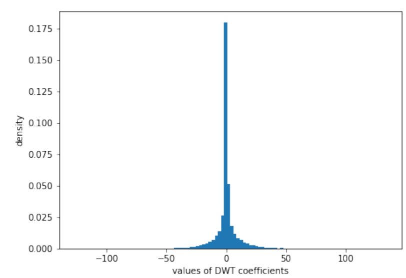

Figure 6: Distribution of the DWT coefficients

Our demonstrative example follows these steps:

- Obtain a clean image (Figure 4). Figure 5 shows the corresponding Haar Discrete Wavelet Transform (DWT) coefficients. By applying MLE from Section 7.1 on these coefficients, we obtain value of $\gamma \sim 7$, as shown in Figure 6.
- Add iid Gaussian noise $\sigma^2 = 400$ to the image, pixel wise (Figure 7). σ^2 is selected such that $\frac{\gamma^2}{\sigma^2}$ lies around 0.1. From Figure 2, this helps us to predict the MSE performance of MAP to be about $0.5 * \sigma^2 = 200$, and that of MMSE is approximately $0.9 * \mathrm{MSE}_{MAP} = 180$.
- Apply MLE (do nothing), MAP, and MMSE for DWT coefficients. Figure 8 shows the DWT coefficients of the noisy image.
- Apply Inverse DWT to recover the image, hopefully with reduced noise.

Figure 9 and 10 shows the reconstructed images from MAP and MMSE, respectively. The MLE estimate is simply the same noisy image from Figure 7. Visually, they all look equally bad, at least to us. But if we calculate the respective MSE, they do agree with the theory as shown in Figure 2.

- 385 for MLE
- 194 for MAP
- 173 for MMSE

We believe the denoising process can be largely improved by using a more visual-oriented metric instead of MSE, and also by estimating the joint $f(\mathbf{x}|\boldsymbol{\theta})$ without the assumption of iid, because this assumption is simply not true for an image. This will lead to a much more complicated formulation.

Figure 7: Noisy image, $\sigma^2 = 400$, MSE=385

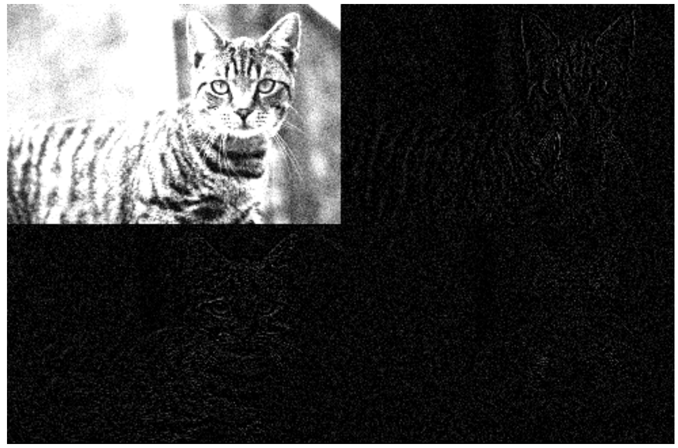

Figure 8: Haar DWT of noisy image

Figure 9: MAP reconstruction, MSE=194

Figure 10: MMSE reconstruction, MSE=173

8 Codes

Full source codes are available on Github: `https://github.com/anvuongb/ece565_final`, where:

- Section 5.1 is from **estimate_gamma.py**
- Section 6.1 is from **parallel_compute.py**
- Section 6.2 is from **parallel_compute_mml.py**
- Section 7 is from **application.py**