This report only contains the various proofs and derivations as well as the final results. Details about the implementation can be found in the (well commented) notebook `Image Denoising.ipynb`.

Mean Field Approximation

Warmup 1

We assume a $\mathcal{N}(0, \omega^{-1})$ prior for μ and a $\text{Gamma}(\alpha, \beta)$ prior for τ and begin with a few lemmas:

$$\begin{aligned}
p(y, \mu, \tau) &= p(y|\mu, \tau)p(\mu)p(\tau) \\
&= \prod_{i=1}^{n} p(y_i|\mu, \tau)p(\mu)p(\tau) \\
&= \prod_{i=1}^{n} \mathcal{N}(y_i; \mu, \tau^{-1})\mathcal{N}(\mu; 0, \omega^{-1}) \,\text{Gamma}(\tau; \alpha, \beta). \qquad (1)
\end{aligned}$$

From this, we have

$$\begin{aligned}
p(\mu|y, \tau) &\propto p(y, \mu, \tau) \\
&\propto \prod_{i=1}^{n} \mathcal{N}(y_i; \mu, \tau^{-1})\mathcal{N}(\mu; 0, \omega^{-1}) \\
&\propto \prod_{i=1}^{n} \exp\Big(-\frac{\tau}{2}(y_i - \mu)^2\Big) \exp\Big(-\frac{\omega}{2}\mu^2\Big) \\
&= \exp\left(-\frac{1}{2}\left(\omega\mu^2 + \tau\sum_{i=1}^{n}(y_i - \mu)^2\right)\right) \\
&\propto \exp\left(-\frac{1}{2}\left(\omega\mu^2 + \tau\sum_{i=1}^{n}(-2y_i\mu + \mu^2)\right)\right) \\
&= \exp\left(-\frac{1}{2}\left(\omega\mu^2 + n\tau\mu^2 - 2\mu\tau\sum_{i=1}^{n} y_i\right)\right) \qquad (2)
\end{aligned}$$

Again using Equation (1), we have

$$\begin{aligned}
p(\tau|y,\mu) &\propto p(y,\mu,\tau)\\
&\propto \prod_{i=1}^{n} \mathcal{N}(y_i;\mu,\tau^{-1})\,\mathrm{Gamma}(\tau;\alpha,\beta)\\
&\propto \prod_{i=1}^{n} \sqrt{\tau}\exp\Big(-\frac{\tau}{2}(y_i-\mu)^2\Big)\tau^{\alpha-1}\exp(-\beta\tau)\\
&= \tau^{\frac{n}{2}+\alpha-1}\exp\Bigg(-\Big(\beta+\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2\Big)\tau\Bigg),
\end{aligned} \tag{3}$$

We approximate the posterior by $p(\mu,\tau|y)\approx q(\mu,\tau)=q(\mu)q(\tau)$. We can then use Equation (2) to obtain

$$\begin{aligned}
q^*(\mu) &\propto \exp\Big(\mathbb{E}_\tau[\ln p(\mu,\tau,y)]\Big)\\
&\propto \exp\Bigg(\mathbb{E}_\tau\Big[-\frac{1}{2}\Big(\omega\mu^2+n\tau\mu^2-2\mu\tau\sum_{i=1}^{n}y_i\Big)\Big]\Bigg)\\
&= \exp\Bigg(-\frac{1}{2}\Big(\omega\mu^2+n\mathbb{E}_\tau[\tau]\mu^2-2\mu\mathbb{E}_\tau[\tau]\sum_{i=1}^{n}y_i\Big)\Bigg)\\
&= \exp\Bigg(-\frac{n\mathbb{E}_\tau[\tau]+\omega}{2}\mu^2+\mu\mathbb{E}_\tau[\tau]\sum_{i=1}^{n}y_i\Bigg)\\
&\propto \exp\Bigg(-\frac{n\mathbb{E}_\tau[\tau]+\omega}{2}\Big(\mu-\frac{\mathbb{E}_\tau[\tau]}{n\mathbb{E}_\tau[\tau]+\omega}\sum_{i=1}^{n}y_i\Big)^2\Bigg),
\end{aligned}$$

which is a $\mathcal{N}(m_{new},s_{new}^{-1})$ distribution with $m_{new}=\frac{\mathbb{E}_\tau[\tau]}{n\mathbb{E}_\tau[\tau]+\omega}\sum_{i=1}^{n}y_i$ and $s_{new}=n\mathbb{E}_\tau[\tau]+\omega$.

Similarly, we can use Equation (3) to obtain

$$\begin{aligned}
q^*(\tau) &\propto \exp\Big(\mathbb{E}_\mu[\ln p(\mu,\tau,y)]\Big)\\
&\propto \exp\Bigg(\mathbb{E}_\mu\Big[\Big(\alpha+\frac{n}{2}-1\Big)\ln\tau-\beta\tau-\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\mu)^2\Big]\Bigg)\\
&= \tau^{\frac{n}{2}+\alpha-1}\exp\Bigg(-\beta\tau-\frac{\tau}{2}\sum_{i=1}^{n}\mathbb{E}_\mu[(y_i-\mu)^2]\Bigg)\\
&= \tau^{\frac{n}{2}+\alpha-1}\exp\Bigg(-\beta\tau-\frac{\tau}{2}\sum_{i=1}^{n}(y_i^2-2y_i\mathbb{E}_\mu[\mu]+\mathbb{E}_\mu[\mu^2])\Bigg)\\
&= \tau^{\frac{n}{2}+\alpha-1}\exp\Bigg(-\Big(\beta+\frac{1}{2}\sum_{i=1}^{n}y_i^2-\mathbb{E}_\mu[\mu]\sum_{i=1}^{n}y_i+\frac{n}{2}\big(\mathrm{Var}_\mu[\mu]+\mathbb{E}_\mu[\mu]^2\big)\Big)\tau\Bigg),
\end{aligned}$$

which is a $\mathrm{Gamma}(\alpha_{new},\beta_{new})$ distribution with $\alpha_{new}=\alpha+\frac{n}{2}$ and $\beta_{new}=\beta+\frac{1}{2}\sum_{i=1}^{n}y_i^2-\mathbb{E}_\mu[\mu]\sum_{i=1}^{n}y_i+\frac{n}{2}(\mathrm{Var}_\mu[\mu]+\mathbb{E}_\mu[\mu]^2)$.

Finally, we utilise the known identities $\mathbb{E}_\tau[\tau] = \frac{\alpha_{new}}{\beta_{new}}$, $\mathbb{E}_\mu[\mu] = m_{new}$ and $\text{Var}_\mu[\mu] = s_{new}^{-1}$ to obtain the update equations

$$\begin{aligned}
m_{new} &= \frac{\alpha_{new}/\beta_{new}}{n \cdot \alpha_{new}/\beta_{new} + \omega} \sum_{i=1}^{n} y_i \\
s_{new} &= n\frac{\alpha_{new}}{\beta_{new}} + \omega \\
\alpha_{new} &= \alpha + \frac{n}{2} \\
\beta_{new} &= \beta + \frac{1}{2}\sum_{i=1}^{n} y_i^2 - m_{new}\sum_{i=1}^{n} y_i + \frac{n}{2}(s_{new}^{-1} + m_{new}^2).
\end{aligned}$$

To check for convergence, we need to compute the ELBO which is given by

$$\begin{aligned}
\mathcal{L}(q^*) &= \mathbb{E}_{q^*}\big[\ln p(\mu,\tau,y)\big] - \mathbb{E}_{q^*}\big[\ln q^*(\mu,\tau)\big] \\
&= \mathbb{E}_{q^*}\big[\ln p(y|\mu,\tau)\big] + \mathbb{E}_{q^*(\mu)}\big[\ln p(\mu)\big] + \mathbb{E}_{q^*(\tau)}\big[\ln p(\tau)\big] - \mathbb{E}_{q^*(\mu)}\big[\ln q^*(\mu)\big] \\
&\quad -\mathbb{E}_{q^*(\tau)}\big[\ln q^*(\tau)\big] \\
&= \frac{n}{2}\mathbb{E}_{q^*(\tau)}[\ln\tau] - \mathbb{E}_{q^*}\left[\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right] - \frac{\omega}{2}\mathbb{E}_{q^*(\mu)}[\mu^2] + (\alpha-1)\mathbb{E}_{q^*(\tau)}[\ln\tau] \\
&\quad -\beta\mathbb{E}_{q^*(\tau)}[\tau] - \frac{s_{new}}{2}\mathbb{E}_{q^*(\mu)}[\mu^2] - \mathbb{E}_{q^*(\mu)}[\mu]\frac{\alpha_{new}}{\beta_{new}}\sum_{i=1}^{n} y_i - (\alpha_{new}-1)\mathbb{E}_{q^*(\tau)}[\ln\tau] \\
&\quad -\beta_{new}\mathbb{E}_{q^*(\tau)}[\tau] + \text{constant} \\
&= \left(\frac{n}{2} + \alpha - 1 - (\alpha_{new}-1)\right)\mathbb{E}_{q^*(\tau)}[\ln\tau] - \mathbb{E}_{q^*(\tau)}\left[\mathbb{E}_{q^*(\mu)}\left[\frac{\tau}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right]\right] \\
&\quad -\frac{\omega}{2}\left(\frac{1}{s_{new}} + m_{new}^2\right) - \beta\frac{\alpha_{new}}{\beta_{new}} - \frac{s_{new}}{2}\left(\frac{1}{s_{new}} + m_{new}^2\right) - m_{new}\frac{\alpha_{new}}{\beta_{new}}\sum_{i=1}^{n} y_i - \alpha_{new} \\
&\quad +\text{constant} \\
&= -\mathbb{E}_{q^*(\tau)}\big[(\beta_{new}-\beta)\tau\big] - \frac{1}{2}\left(\frac{\omega}{s_{new}} + (\omega + s_{new})m_{new}^2\right) - (\beta+\beta_{new})\frac{\alpha_{new}}{\beta_{new}} \\
&\quad -m_{new}\frac{\alpha_{new}}{\beta_{new}}\sum_{i=1}^{n} y_i + \text{constant} \\
&= (\beta-\beta_{new})\frac{\alpha_{new}}{\beta_{new}} - \frac{1}{2}\left(\frac{\omega}{s_{new}} + (\omega+s_{new})m_{new}^2\right) - (\beta+\beta_{new})\frac{\alpha_{new}}{\beta_{new}} \\
&\quad -m_{new}\frac{\alpha_{new}}{\beta_{new}}\sum_{i=1}^{n} y_i + \text{constant} \\
&= -2\alpha_{new} - \frac{1}{2}\left(\frac{\omega}{s_{new}} + (\omega+s_{new})m_{new}^2\right) - m_{new}\frac{\alpha_{new}}{\beta_{new}}\sum_{i=1}^{n} y_i + \text{constant}.
\end{aligned}$$

We ran the algorithm until the ELBO converged and plotted histograms for samples of μ and τ from their final distributions, shown in Figures 1 and 2.

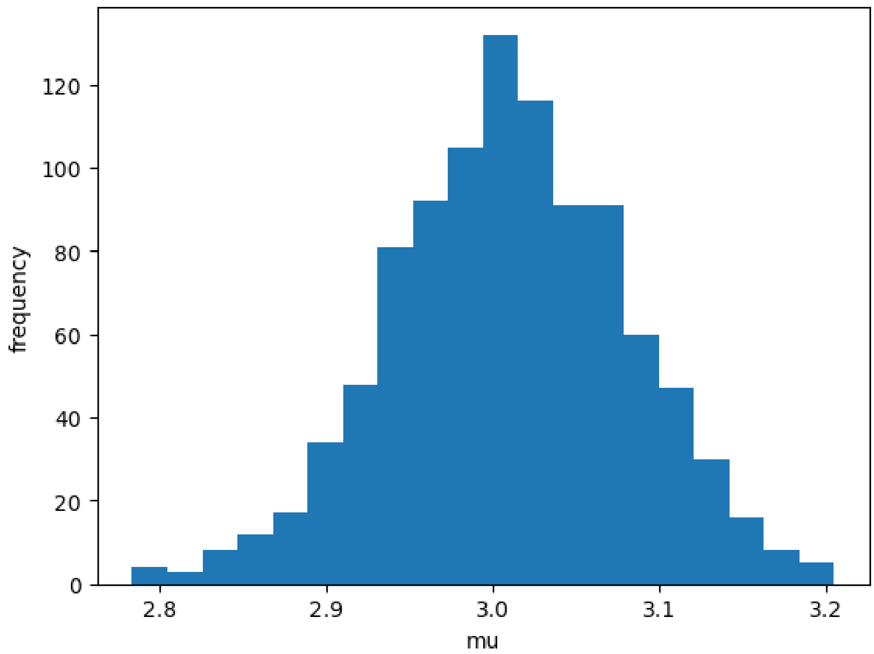

Figure 1: Histogram of samples of μ.

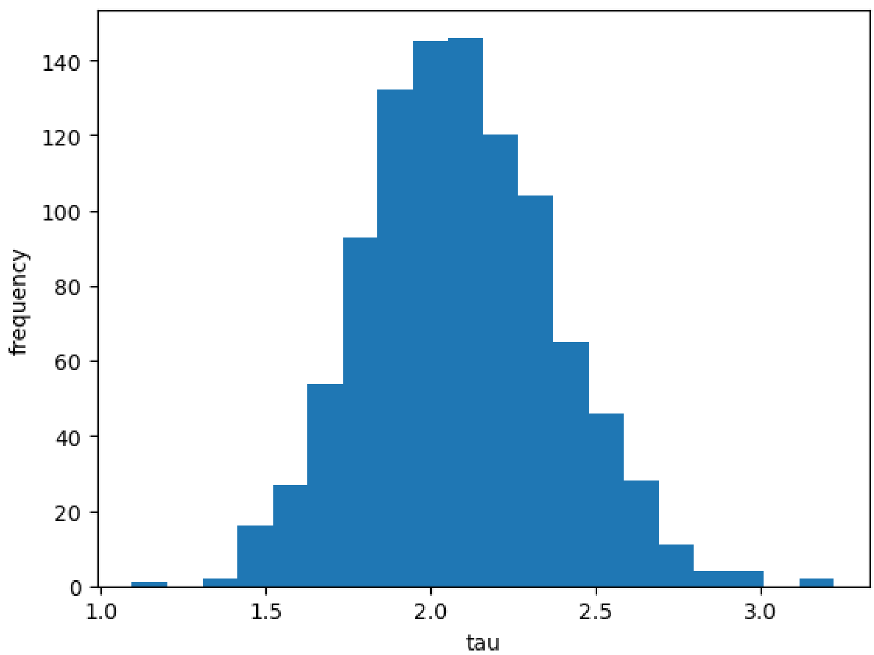

Figure 2: Histogram of samples of τ.

Warmup 2

Recall that assuming the priors $q(\lambda) = \text{Gamma}(a_\lambda, b_\lambda)$ and $q(\beta) = \text{Gamma}(a_\beta, b_\beta)$, we have the update equations

$$\begin{aligned}
a_\lambda &= 2 + \sum_i y_i \\
b_\lambda &= b_\lambda^0 \\
a_\beta &= 3 \\
b_\beta &= b_\beta^0
\end{aligned}$$

Again, we need to compute the ELBO to check for convergence. Using the known identities $\mathbb{E}_{q^*(\lambda)}[\lambda] = a_\lambda / b_\lambda$ and $\mathbb{E}_{q^*(\beta)}[\beta] = a_\beta / b_\beta$, we have

$$\begin{aligned}
\mathcal{L}(q^*) &= \mathbb{E}_{q^*}\big[\ln p(y, \lambda)\big] - \mathbb{E}_{q^*}\big[\ln q^*(\lambda, \beta)\big] \\
&= \mathbb{E}_{q^*(\lambda)}\big[\ln p(y|\lambda)\big] + \mathbb{E}_{q^*}\big[\ln p(\lambda|\beta)\big] + \mathbb{E}_{q^*(\beta)}[\ln p(\beta)] - \mathbb{E}_{q^*(\lambda)}\big[\ln q^*(\lambda)\big] \\
&\quad - \mathbb{E}_{q^*(\beta)}\big[\ln q^*(\beta)\big] \\
&= \sum_{i=1}^n \mathbb{E}_{q^*(\lambda)}[y_i \ln \lambda - \lambda] + \mathbb{E}_{q^*}[\ln \lambda - \beta\lambda + 2 \ln \beta] + \mathbb{E}_{q^*(\beta)}[-\beta] \\
&\quad - \mathbb{E}_{q^*(\lambda)}\big[(a_\lambda - 1)\ln \lambda - b_\lambda \lambda\big] - \mathbb{E}_{q^*(\beta)}\big[(a_\beta - 1)\ln \beta - b_\beta \beta\big] + \text{constant} \\
&= \mathbb{E}_{q^*(\lambda)}[\ln \lambda] \sum_{i=1}^n y_i - \mathbb{E}_{q^*(\lambda)}[\lambda] + \mathbb{E}_{q^*(\lambda)}[\ln \lambda] - \mathbb{E}_{q^*}[\beta\lambda] + 2\mathbb{E}_{q^*(\beta)}[\ln \beta] - \mathbb{E}_{q^*(\beta)}[\beta] \\
&\quad - (a_\lambda - 1)\mathbb{E}_{q^*(\lambda)}[\ln \lambda] - b_\lambda \mathbb{E}_{q^*(\lambda)}[\lambda] - (a_\beta - 1)\mathbb{E}_{q^*(\beta)}[\ln \beta] - b_\beta \mathbb{E}_{q^*(\beta)}[\beta] \\
&\quad + \text{constant} \\
&= \Big(\sum_{i=1}^n y_i + 1 - (a_\lambda - 1)\Big)\mathbb{E}_{q^*(\lambda)}[\ln \lambda] + \big(2 - (a_\beta - 1)\big)\mathbb{E}_{q^*(\beta)}[\ln \beta] \\
&\quad - \frac{a_\lambda}{b_\lambda} - \frac{a_\lambda a_\beta}{b_\lambda b_\beta} - \frac{a_\beta}{b_\beta} - a_\lambda - a_\beta + \text{constant} \\
&= -\frac{a_\lambda}{b_\lambda} - \frac{a_\lambda a_\beta}{b_\lambda b_\beta} - \frac{a_\beta}{b_\beta} - a_\lambda - a_\beta + \text{constant},
\end{aligned}$$

where the last equality comes from the given update equations.

We ran the algorithm until either the parameters or the ELBO converged and plotted histograms for samples of λ and β from their final distributions, shown in Figures 3 and 4.

Image Denoising with Gibbs Sampling

Effect of different noise levels vs normalized mean squared error (NMSE)

The image denoised using Gibbs sampling is shown in Figure 5. The plot of the NMSEs

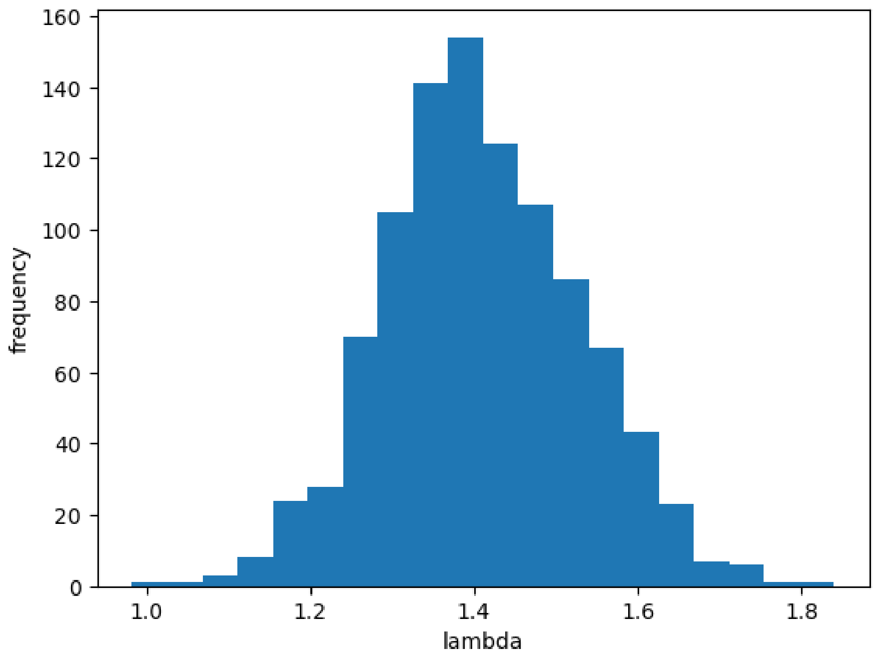

Figure 3: Histogram of samples of λ.

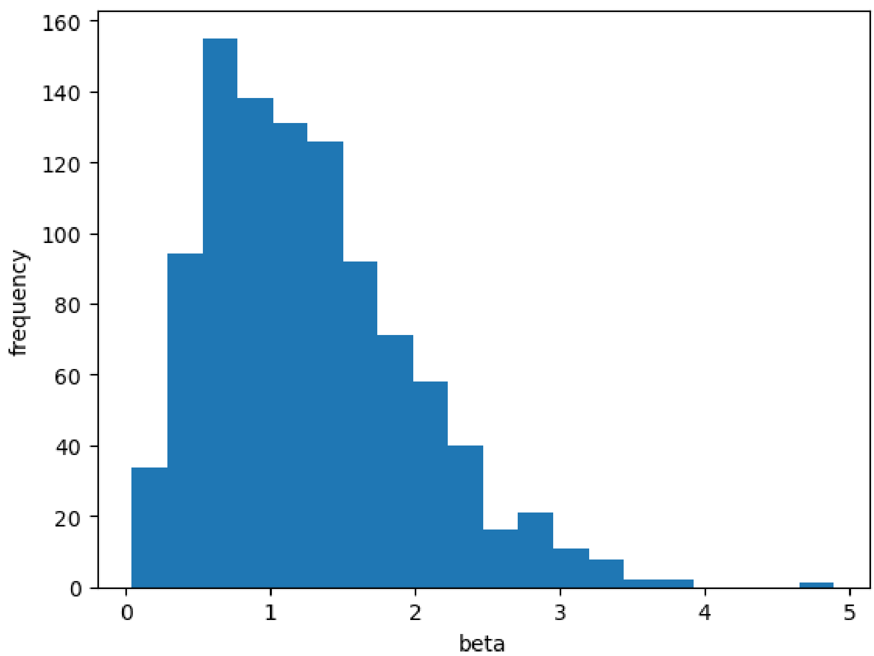

Figure 4: Histogram of samples of β.

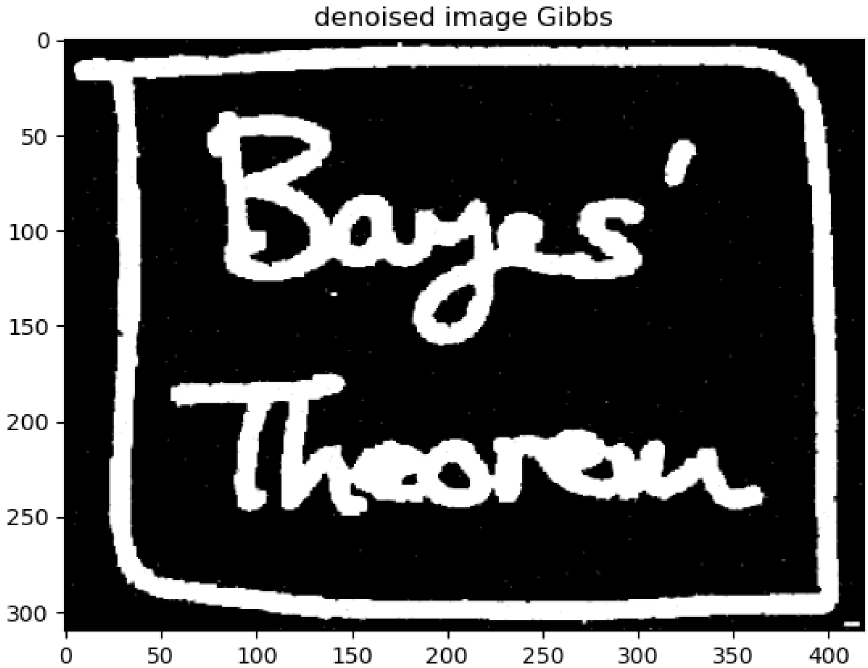

Figure 5: Denoised image using Gibbs sampling.

against the various parameters is shown in Figure 6.

Image Denoising with Mean Field Approximation

Noise Level Analysis (Again)

The ELBO is given by

$$\begin{aligned}
\mathcal{L}(q^*) &= \mathbb{E}_{q^*}\big[\ln p(\mathbf{x},\mathbf{y})\big] - \mathbb{E}_{q^*}\big[\ln q^*(\mathbf{x})\big] \\
&= \mathbb{E}_{q^*}\Big[\sum_i \sum_{j\in N(i)} \beta x_i x_j + \sum_i \eta x_i y_i\Big] - \sum_i \mathbb{E}_{q^*(x_i)}\big[\ln q^*(x_i)\big] + \text{constant} \\
&= \mathbb{E}_{q^*}\Big[\sum_i \sum_{j\in N(i)} \beta x_i x_j\Big] + \sum_i \mathbb{E}_{q^*(x_i)}[\eta x_i y_i] - \sum_i \mathbb{E}_{q^*(x_i)}\Big[x_i \sum_{j\in\mathcal{N}(i)} \beta\mu_j + \eta x_i y_i\Big] \\
&\quad + \text{constant} \\
&= \sum_i \sum_{j\in N(i)} \beta\mu_i\mu_j + \eta\sum_i \mu_i y_i - \sum_i \sum_{j\in\mathcal{N}(i)} \beta\mu_i\mu_j - \eta\sum_i \mu_i y_i + \text{constant} \\
&= \text{constant},
\end{aligned}$$

which means that the ELBO only converges once all μ_i converge.

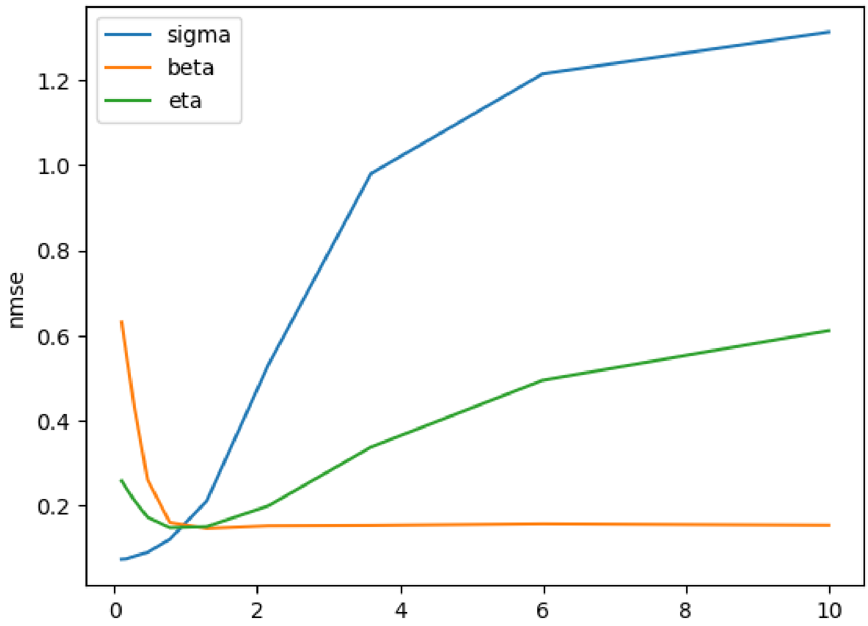

Figure 6: Plot of NMSEs against various parameters.

Running the algorithm until convergence, we obtain the denoised image shown in Figure 7. The plot of the NMSEs against the various parameters is shown in Figure 8.

0.1 NMSE Comparison Between Models

The image denoised using ICM is shown in Figure 9. The final NMSEs are the following:

$$\begin{aligned} \text{Gibbs NMSE:} \quad & 0.14667091262223317 \\ \text{Mean Field NMSE:} \quad & 0.12169590856594853 \\ \text{ICM NMSE:} \quad & 0.16758670277993076 \end{aligned}$$

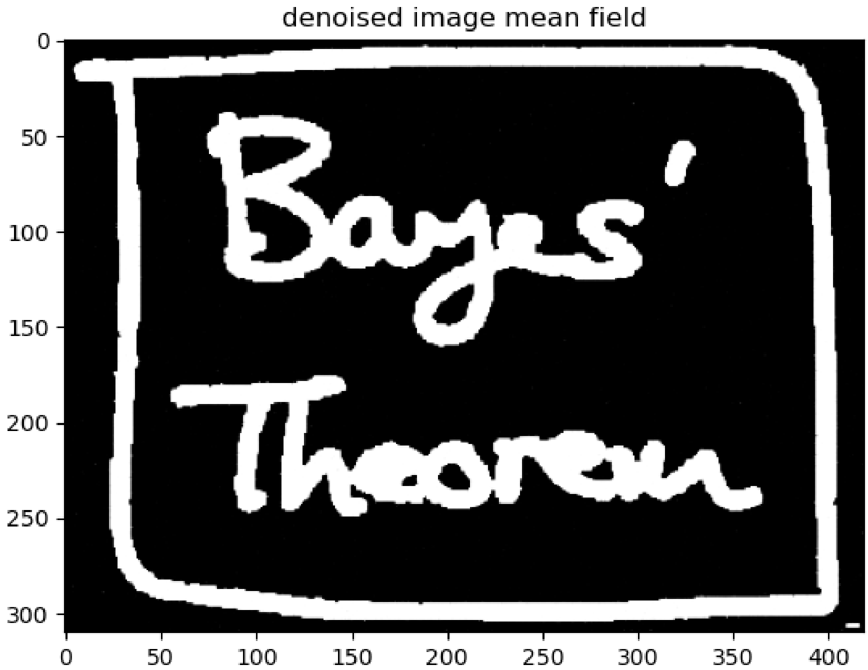

Figure 7: Denoised image using mean field approximation.

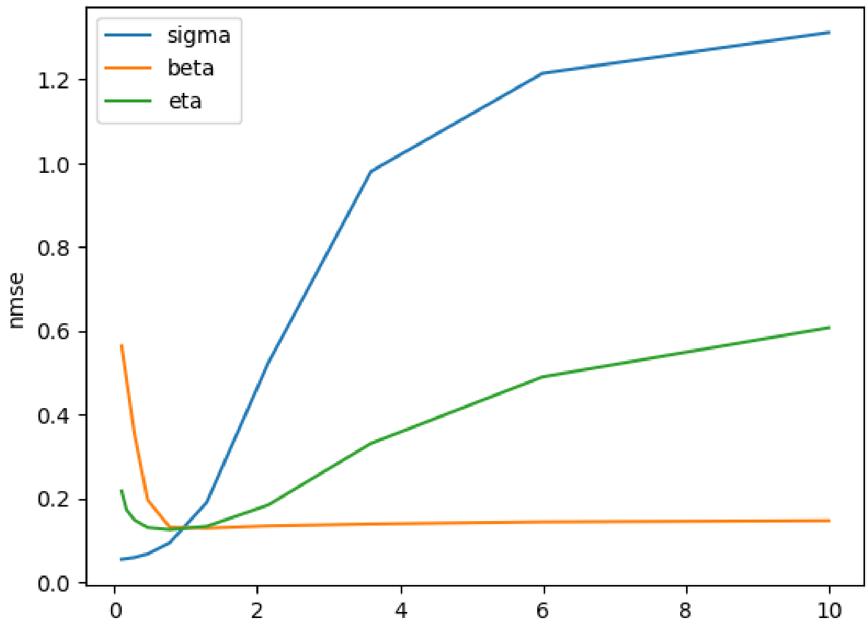

Figure 8: Plot of NMSEs against various parameters.

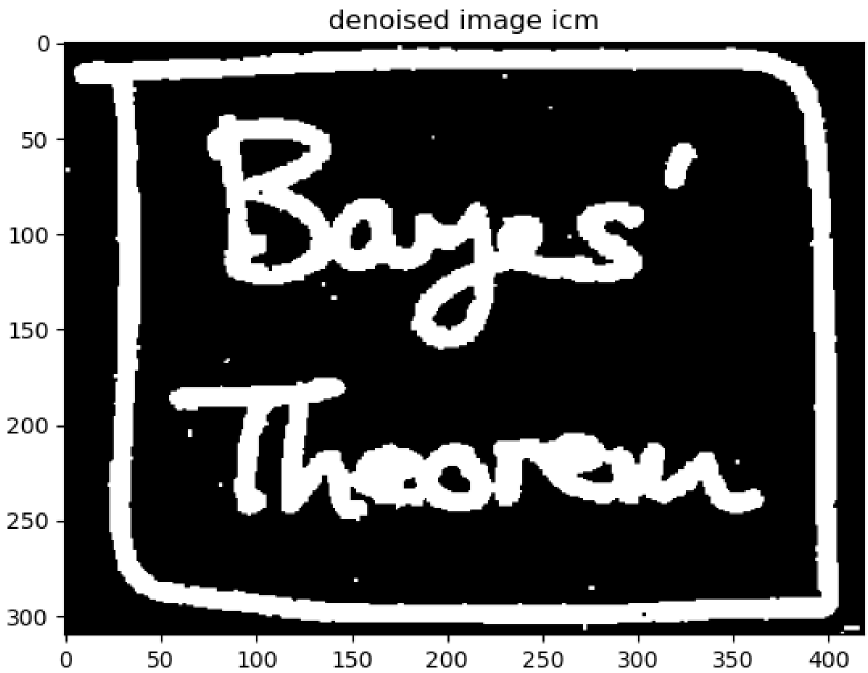

Figure 9: Denoised image using ICM.