SPP Assignment Report: Total Variational Regularization

SAPTARSHI MANDAL

18th February 2024

1 Exercises

1.1 Exercise 1

The total variation (TV) functional $\varphi : \mathbb{R}^n \to \mathbb{R}$ is defined as:

$$\forall s = (s_1, \ldots, s_n) \in \mathbb{R}^n : \quad \varphi(s) = \sum_{i=2}^{n} |s_i - s_{i-1}|. \tag{1}$$

If we represent a one-dimensional signal as a vector $s \in \mathbb{R}^n$, where n is the length of the signal, then $\varphi(s)$ is the ℓ_1 norm of a discrete form of the derivative of s. It measures the size of "jumps" in a signal. In particular, $\varphi(s)$ is small if s is a piecewise constant signal with small jumps.
It can be shown that $\varphi(s) = \|Hs\|_1$, where $H \in \mathbb{R}^{(n-1)\times n}$ and $\|\cdot\|_1$ is the ℓ_1 norm on $\mathbb{R}^{n-1}$.
The required matrix H will be

$$\begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & -1 & 1 \end{pmatrix}$$

The spectral norm of H is the $\sqrt{\text{highest eigenvalue of} HH^T}$
HH^T is a tri-diagonal matrix. T The eigenvalue of HH^T is given by

$$\lambda_{max} = 2 + 2cos\frac{\pi}{n+1}$$

So,

$$||H||_2 = \sqrt{2 + 2cos\frac{\pi}{n+1}}$$

The TV denoising operator $D_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ for some $\lambda > 0$ is defined as:

$$D_\lambda(s) = \underset{z\in\mathbb{R}^n}{argmin} \left\{ \frac{1}{2}\|s - z\|_2^2 + \lambda\varphi(z) \right\} \tag{2}$$

where:

- $\|\cdot\|_2$ denotes the ℓ_2 norm.
- s represents the noisy signal.
- z represents the denoised signal.
- $\mathbb{R}^n$ denotes the space of n-dimensional real vectors.
- λ is a regularization parameter controlling the trade-off between data fidelity and signal smoothness.
- $\varphi(z)$ is a function penalizing lack of smoothness in the signal, known as the Total Variation (TV) functional.

This minimization problem seeks to find the signal z closest to the noisy signal s (data fidelity term) while minimizing its total variation ($\varphi(z)$ encourages smoothness). This approach balances preserving the essential information in the noisy signal with removing unwanted high-frequency noise.

In the imaging community, this process is referred to as TV denoising due to its focus on promoting smoothness in the reconstructed image.

1.2 Exercise 4

Compressed sensing aims to recover the unknown signal ξ from the measurement x. This can be achieved using the reconstruction operator $R_\lambda : \mathbb{R}^n \to \mathbb{R}^n$, defined as:

$$R_\lambda(s) = \underset{z \in \mathbb{R}^n}{argmin} \left\{ \frac{1}{2}\|s - Az\|_2^2 + \lambda\varphi(z) \right\} \tag{3}$$

where:

- $\|\cdot\|_2$ denotes the ℓ_2 norm.
- A is an (m $\times$ n) matrix representing the measurement process.
- λ is a regularization parameter controlling the trade-off between data fidelity and signal prior.
- $\varphi(z)$ is a function penalizing lack of smoothness in the signal, known as the Total Variation (TV) functional.

This minimization problem, referred to as TV reconstruction, seeks a signal z with small total variation (smoothness) while simultaneously fitting the measurement model represented by $x = Az$.

To find $R_\lambda(s)$ we implement proximal gradient descent.

$$
\begin{aligned}
R_\lambda(s) &= \underset{z\in\mathbb{R}^n}{argmin}\left\{\frac{1}{2}\|s-Az\|_2^2+\lambda\varphi(z)\right\}\\
&= \underset{z\in\mathbb{R}^n}{argmin}\left\{f(z)+\lambda\varphi(z)\right\}\\
\implies f(z) &= \frac{1}{2}\|s-Az\|_2^2\\
\implies \nabla f(z) &= -A^T(s-Az)
\end{aligned}
$$

Using Proximal Gradient Descent:

$$z_k = D_\lambda(z_{k-1}\gamma\nabla_z f(z))$$

1.3 Exercise 5

We are required to show that $\forall\theta\in\mathbb{R}^n : ||\theta||_1 = \max_{y\in\mathbb{B}^n} y^T\theta$
where $\mathbb{B}^n = \{y\in\mathbb{R}^n : |y_i|\leq 1, i = 1,\ldots,n\}$
Proof:

$$
\begin{aligned}
\max_{y\in\mathbb{B}^n} y^T\theta &= \max_{||y||_\infty\leq 1}|\sum_i y_i\theta_i|\\
&= \max_{max\{|y_i|\}\leq 1}|\sum_i y_i\theta_i|\\
&\leq \max_{max\{|y_i|\}\leq 1}\sum_i |y_i||\theta_i|\\
&\leq \sum_i|\theta_i|\\
&= ||\theta||_1\\
\implies \max_{y\in\mathbb{B}^n} y^T\theta = ||\theta||_1
\end{aligned}
$$

1.4 Exercise 6

Let $\psi(y,s) = \min_{z\in\mathbb{R}^n}\left\{\underbrace{\frac{1}{2}\|s-z\|_2^2+\lambda y^T Hz}_{F}\right\}$

$$
\begin{aligned}
\nabla_z F = -(s-z')+\lambda(H^Ty) &= 0\\
\implies z' = s-\lambda(H^Ty)
\end{aligned}
$$

Putting the value of z' in $\psi(y,s)$

$$
\begin{aligned}
\psi(y,s) &= \frac{1}{2}\|\lambda H^T y\|_2^2 + \lambda y^T H(s - \lambda H^T y) \\
&= \frac{1}{2}\|\lambda H^T y\|_2^2 + \lambda y^T H s - \lambda^2 y^T H H^T y \\
&= \lambda y^T H s - \frac{1}{2}\|\lambda H^T y\|_2^2
\end{aligned}
$$

Now, Hessian of $\psi(y,s) = -\lambda^2 H H^T$ which is clearly NSD $\implies \psi(y,s)$ is a concave function.

$$
\begin{aligned}
||\nabla_y \psi(y_1, s) - \nabla_y \psi(y_2, s)||_2 & \\
&= ||\lambda^2 H H^T (y_2 - y_1)||_2 \\
&\leq \underbrace{\lambda^2 ||H H^T||_2}_{L} ||y_2 - y_1||_2 \\
\implies L = \lambda^2 ||H H^T||_2 &
\end{aligned}
$$

1.5 Exercise 7

To solve the optimization problem : $\max\limits_{y \in B^{n-1}} \psi(y,s)$

we use projected gradient ascent given by

$$
\begin{aligned}
y_k = \Pi_{\mathbb{B}}(y_{k-1} + \alpha \nabla_{y_k-1,s}\psi) & \\
&= \Pi_{\mathbb{B}}(y_{k-1} + \frac{1}{L}\nabla_{y_k-1,s}\psi)
\end{aligned}
$$

Now Previously we have seen in Eq.**??** that:

$$
z' = s - \lambda(H^T y)
$$

Now if the solution to the above projected gradient ascent is y^* then we need to plug that in the above equation to get z^*

$\implies z^* = s - \lambda(H^T y^*)$

1.6 Exercise 8

Consider the projection operator $\Pi_B^n : \mathbb{R}^n \to \mathbb{R}^n$ defined for all $x \in \mathbb{R}^n$ as:

$$
\Pi_B^n(x) = \underset{z \in B^n}{argmin} \|z - x\|_2^2. \tag{4}
$$

where:

* $\|\cdot\|_2$ denotes the ℓ_2 norm. * B^n is the unit hypercube in $\mathbb{R}^n$, meaning all vectors with ℓ_∞ norm less than or equal to 1.

This minimization problem finds the closest point to x within the unit hypercube B^n. There is a closed-form solution for this projection:

$$\Pi_B^n(x) = \frac{x}{||x||_\infty}, \tag{5}$$

In simpler terms, the projection onto the unit hypercube is achieved by first normalizing the input vector x by the max absolute value of the elements of x . So, the resulting vector has a max absolute value of 1 which ensures that the vector is on the hypercube. This also preserved the direction of the vector.

2 Signal Generation

Let n=256 represent the signal length. The following one-dimensional signals of length n are generated and plotted:

1. Piece wise Constant: A signal is constructed with the following structure:
   - 128 samples with a magnitude of 1.
   - 64 samples with a magnitude of 2.
   - 64 samples with a magnitude of 3.
2. Sawtooth: A sawtooth signal with 4 cycles and an amplitude of 1 is generated.
3. Sinusoid: A sinusoidal signal with 4 cycles and an amplitude of 1 is generated.

Plots for each generated signal are then created to visualize the results. The plots are given in Figure.1

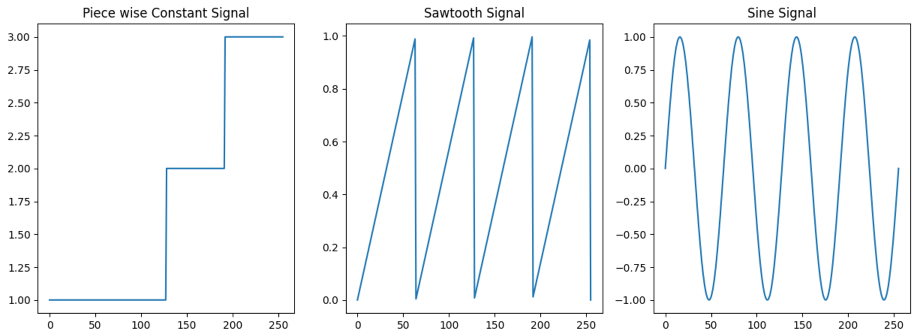

Figure 1: Generated Clean Signals

3 Signal Denoising

The ability to denoise signals using a prior information opens the door to applying the same approach to broader inverse problems. In this work, we focus on the compressed sensing problem, where the goal is to recover an unknown signal ξ from compressed measurements:

$$x = A\xi + \eta\sigma, \tag{6}$$

where:

- A is an (m×n) Gaussian random matrix with each element $A_{i,j}$is sampled independently from a zero-mean, unit-variance Gaussian distribution.
- η is white Gaussian noise with i.i.d. components $\eta_1, \ldots, \eta_m$ following a zero-mean, unit-variance Gaussian distribution.
- σ represents the noise level.

In essence, we leverage the prior knowledge that signals tend to have small total variation to guide the reconstruction process from compressed measurements corrupted by noise.

Let the measurement matrix A in model (6) be set to the identity matrix (A = I).The noisy signals are generated and plotted for each noise level: $\sigma \in \{0, 0.5, 1\}$.

The noisy signals are given in Figure.2, Figure.3, Figure.4.

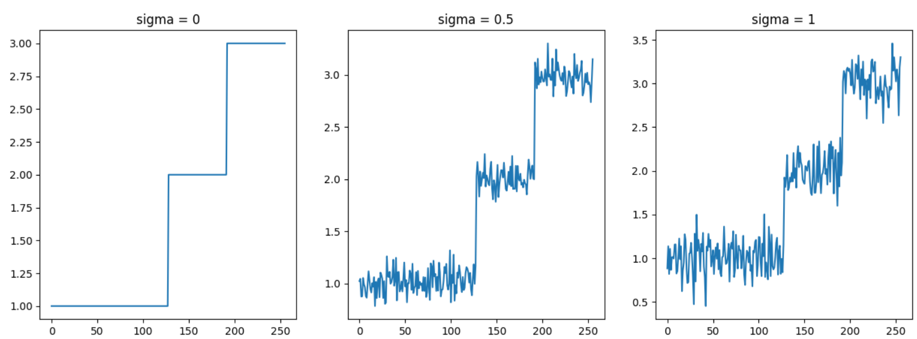

Figure 2: Noisy Piece-wise Constant Signals with varying noise levels

Exercise 2 is applied to each of the measured signals generated in part (a). Experiments with various values of the regularization parameter λ is done to identify suitable settings for reconstruction.The optimal value of λ for each case is reported and the corresponding reconstructed signal is plotted. λ values are varied and Mean Square Error with the original signal is calculated. The λ corresponding to minimum MSE is decalred as optimal The Optimal Value of λ found is given in Table.3

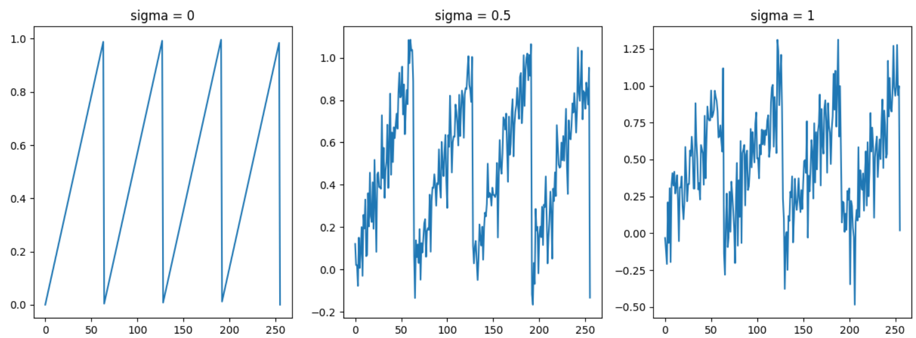

Figure 3: Noisy Sawtooth Signals with varying noise levels

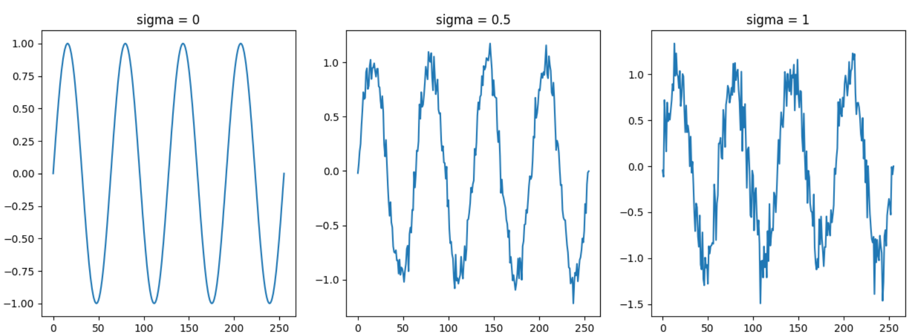

Figure 4: Noisy Sinusoidal Signals with varying noise levels

Signal	**Noise Level σ**	**Optimal λ**
Piece-wise Constant	0	0.1
	0.5	0.5
	1	1.5
Sawtooth	0	0.1
	0.5	0.4
	1	0.7
Sinusoidal	0	0.1
	0.5	0.4
	1	0.8

The corresponding denoised signal using the optimal λ is given in Figure.5, Figure.3 and Figure.7

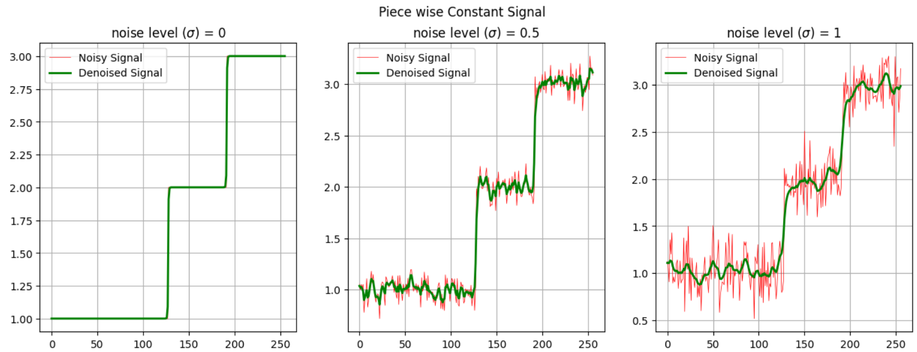

Figure 5: Denoised Piece-wise Constant signal

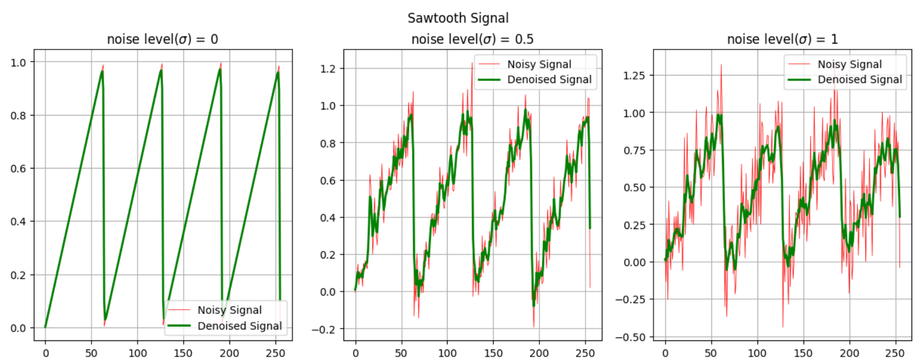

Figure 6: Denoised Sawtooth signal

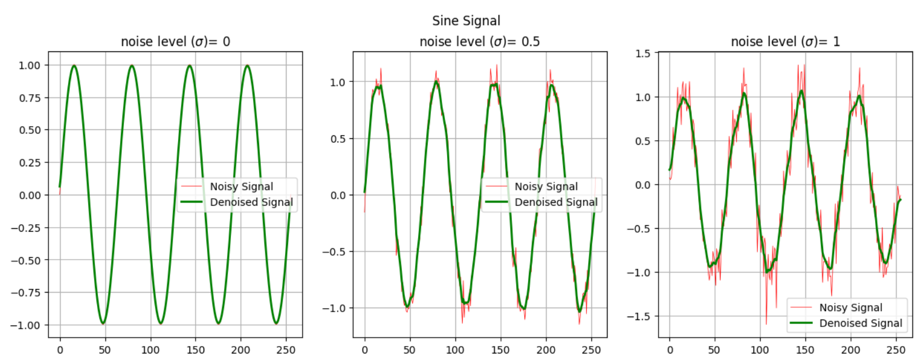

Figure 7: Denoised Sinusoidal signal

4 Signal Reconstruction

Equation.6 is applied as the measurement model on the generated signals. Different noise levels $\sigma \in \{0, 0.5, 1\}$ are considered and the resulting noisy,compressed signals for the measurement ratios $m \in \{0.5n, 0.75n, n\}$ are plotted.

The generated compressed, noisy signals are given in Figure.8, Figure.9, Figure.10.

Exercise 4 is solved for each of the measured signals. In each case, λ is tuned to find a good reconstruction.

The optimal λs are given in the following Table.4

Signal	Noise Level σ	m = 128	m = 192	m = 256
Piece-wise Constant	0	0.65	0.65	0.65
	0.5	0.65	0.65	0.65
	1	0.65	0.65	0.65
Sawtooth	0	0.6	0.6	0.6
	0.5	0.6	0.6	0.6
	1	0.6	0.6	0.65
Sinusoidal	0	0.6	0.6	0.6
	0.5	0.6	0.6	0.65
	1	0.6	0.6	0.65

The Reconstructed image using the optimal λ found is given in Figure.11, Figure.12 and Figure.13.

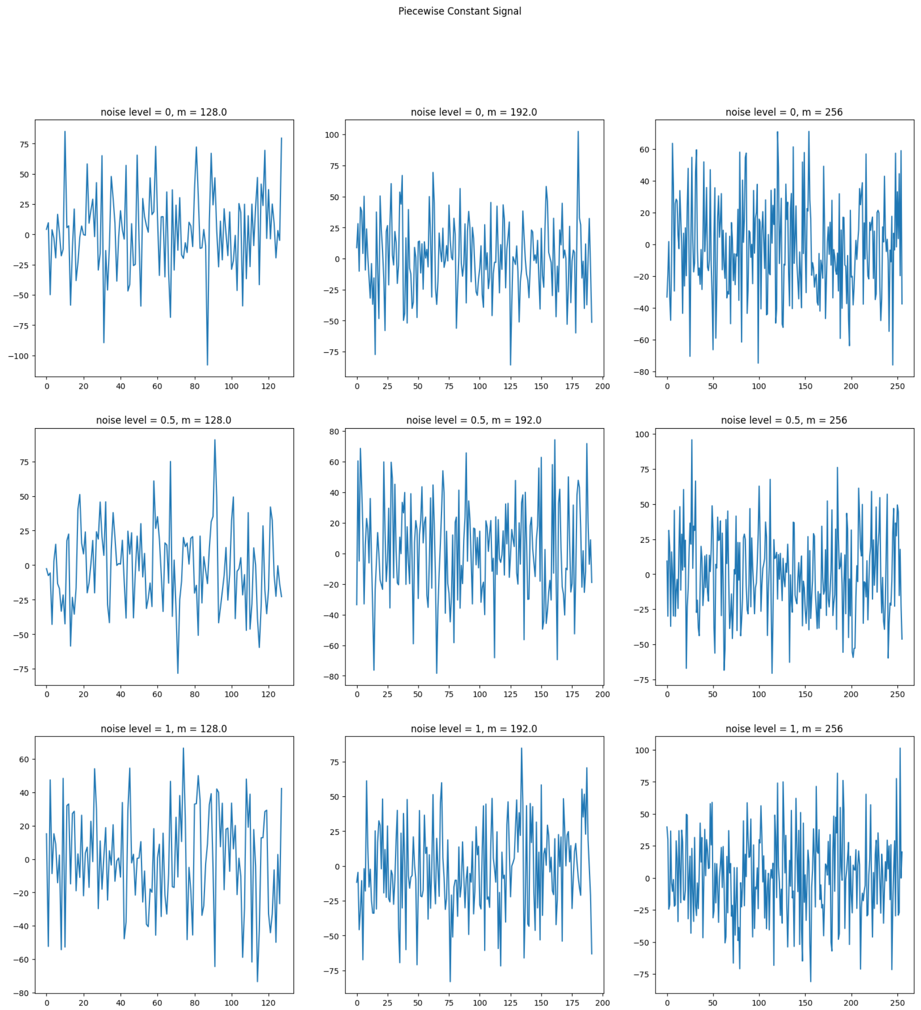

Figure 8: Compressed, Noisy Piece-wise Constant Signal

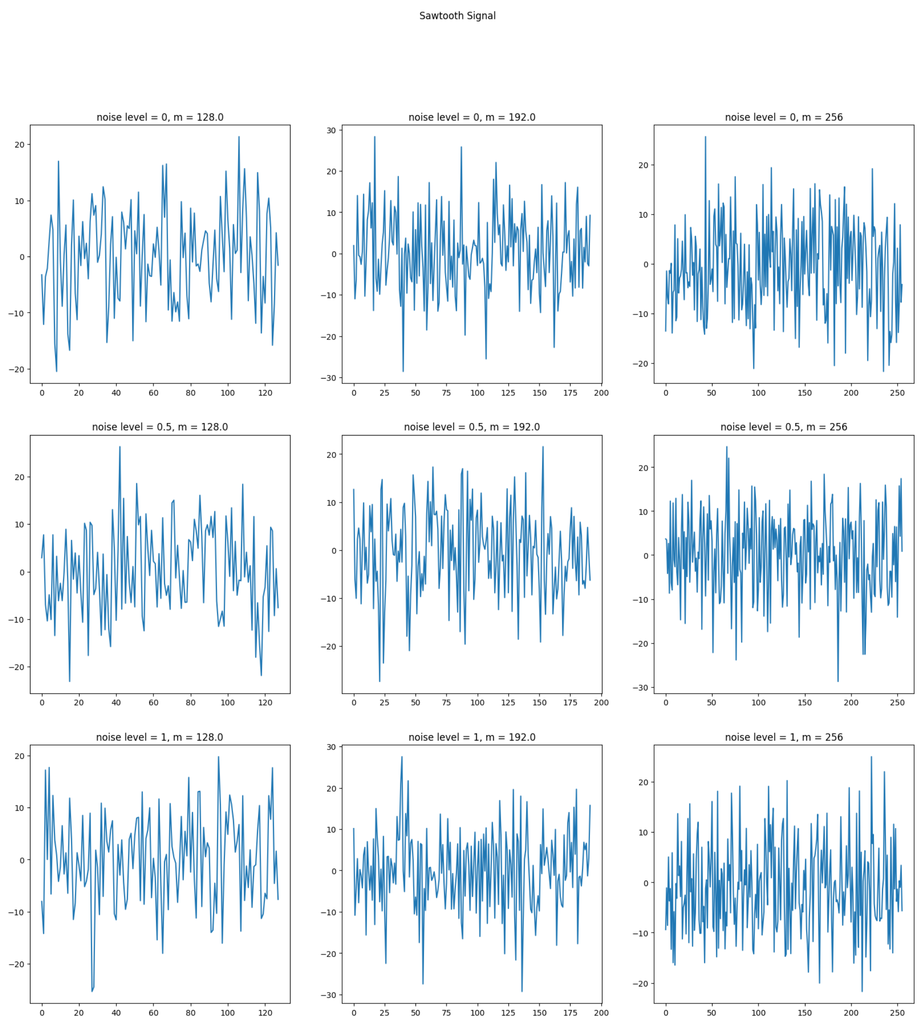

Figure 9: Compressed, Noisy Sawtooth Signal

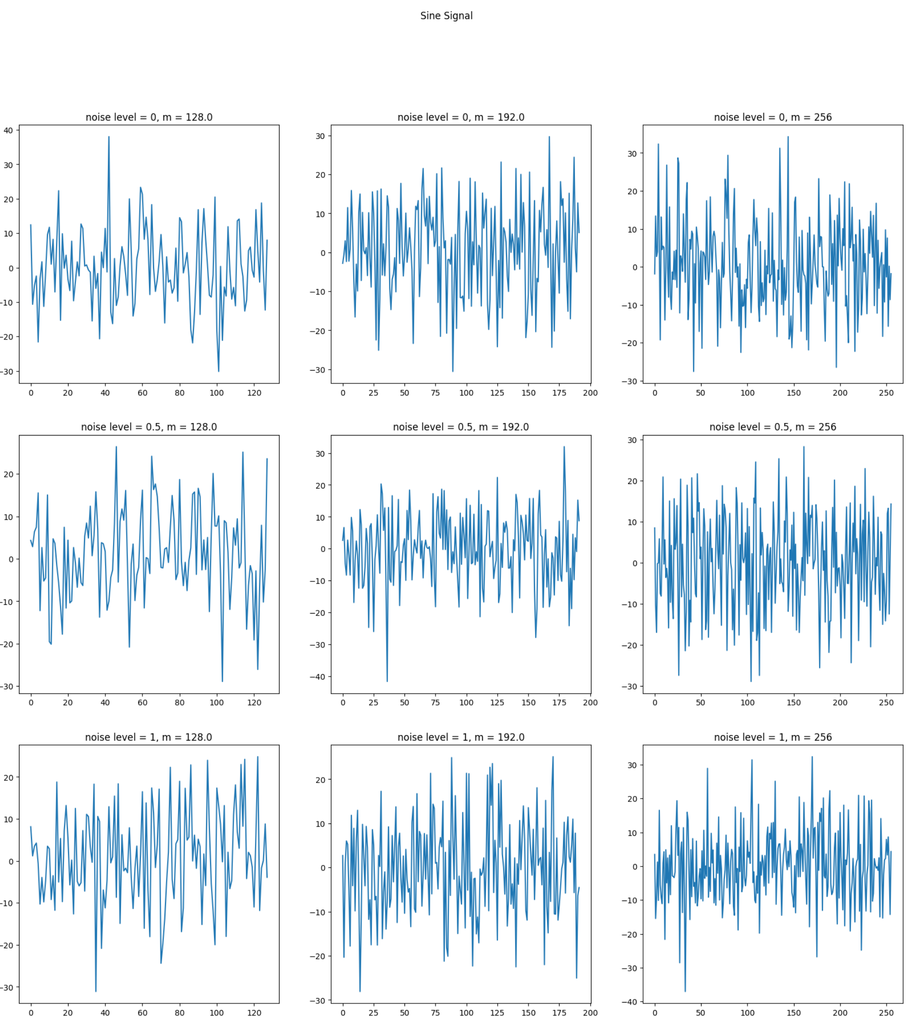

Figure 10: Compressed, Noisy Sinusoidal Signal

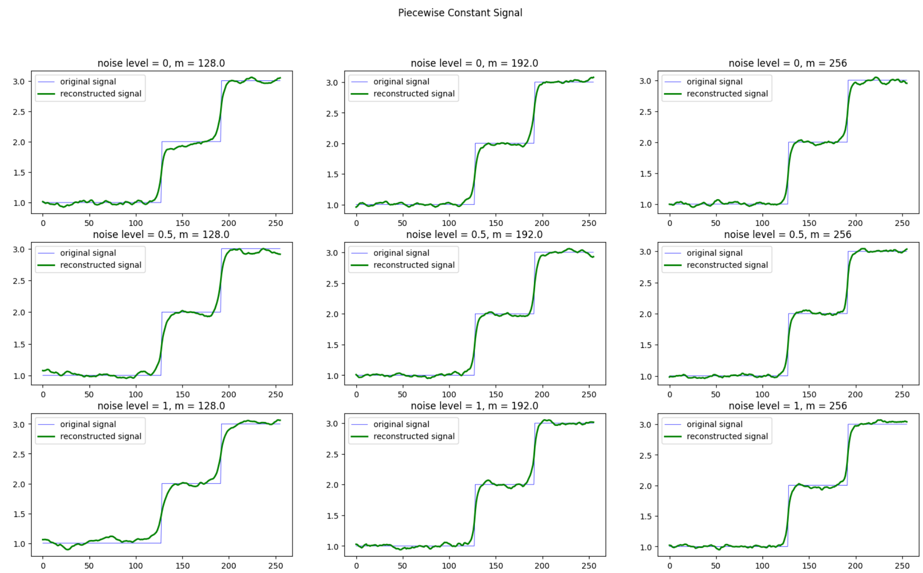

Figure 11: Reconstructed Piece-wise Constant Signal

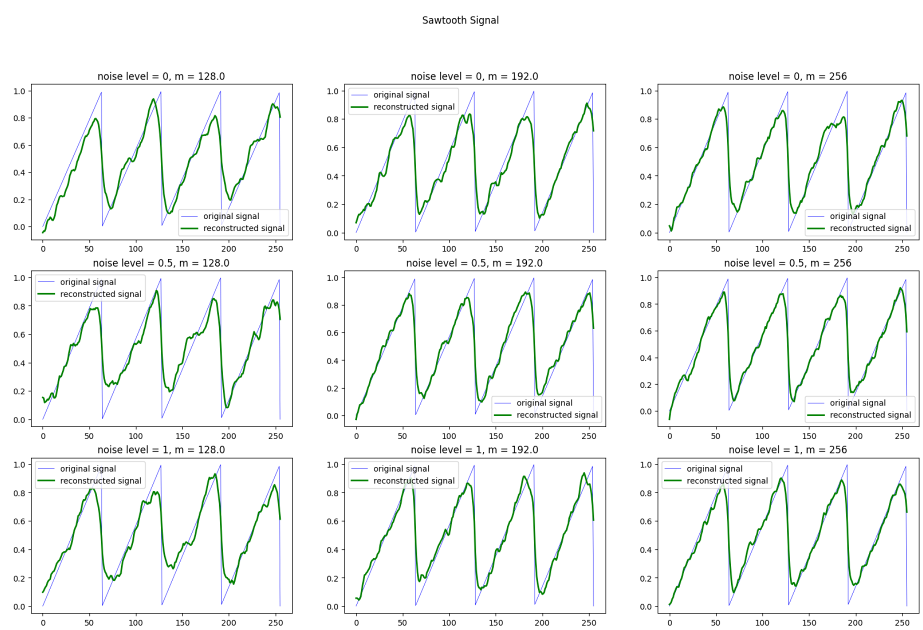

Figure 12: Reconstructed Sawtooth Signal

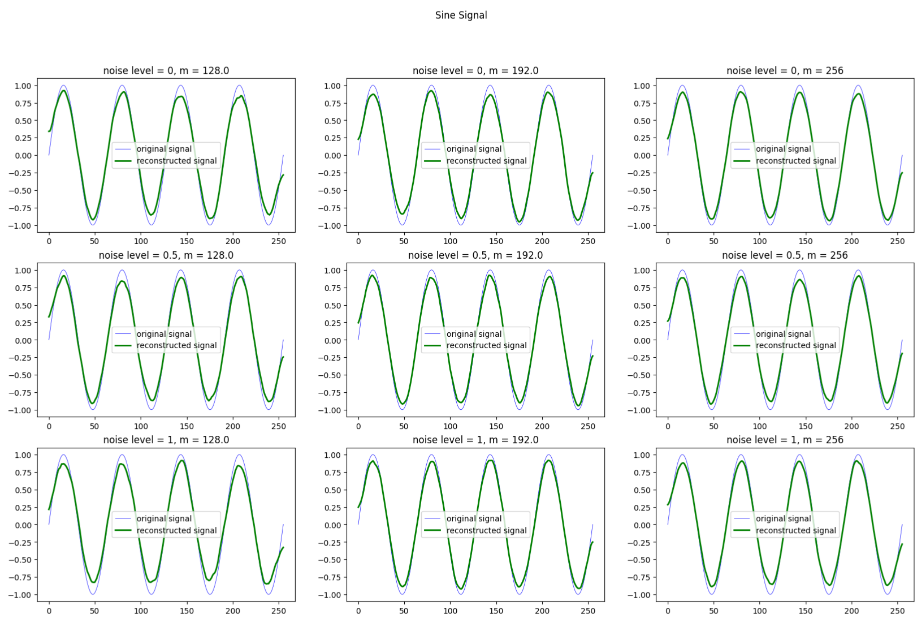

Figure 13: Reconstructed Sinusoidal Signal