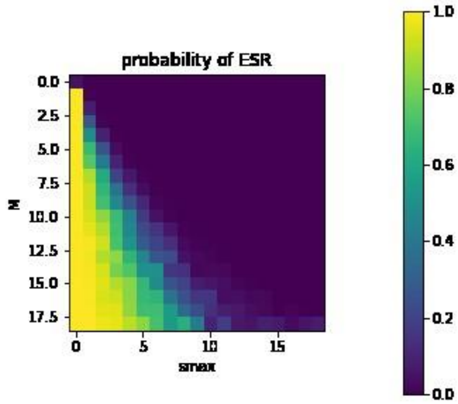
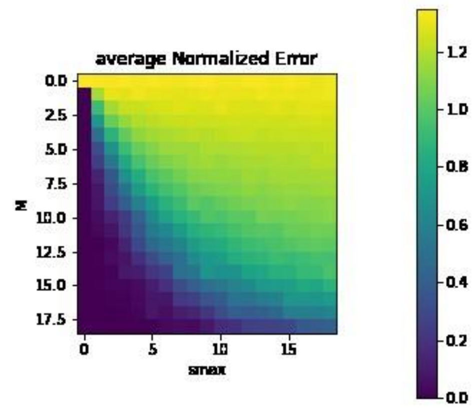
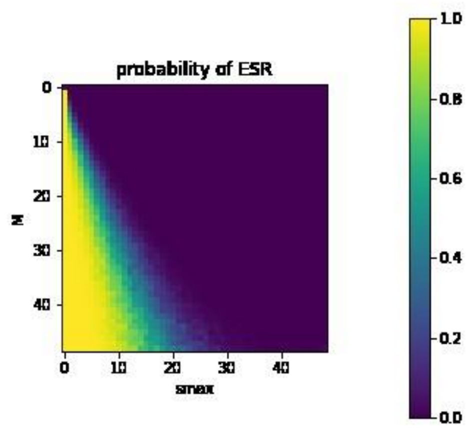
3) **Noiseless Case:**

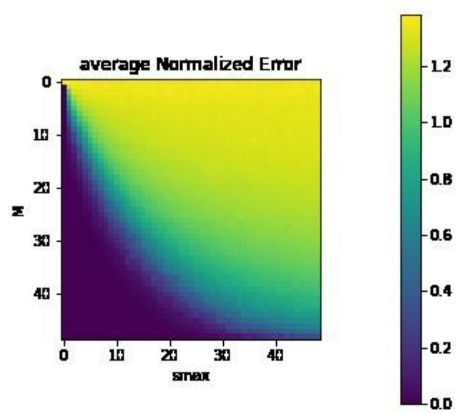
1) N=20



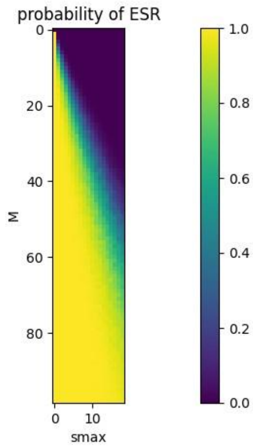


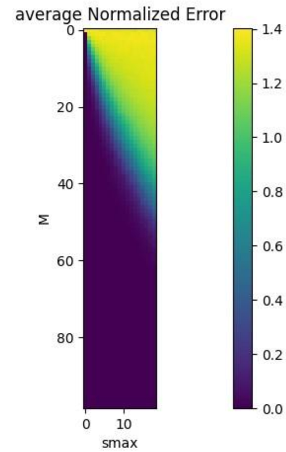
2) N=50





3) N=100





Comment: For the ESR: There's a clear threshold behaviour observable as a diagonal line from the bottom left to the upper middle right marking the transition between the high and low probability of ESR. To the left and below this line, you have a higher chance of successful recovery. The optimal measurement strategy seems to lie along the transition zone where you have just enough measurements to ensure high probability of recovery without making unnecessary measurements that do not contribute to improving the recovery probability. As the signal size increases it takes more measurements to recover a larger signal, even if the sparsity ratio (sparsity level relative to signal size) remains the same. As N increases, the number of measurements required for accurate reconstruction increases as well. The sharp transition can be typically observed when Smax>=M.

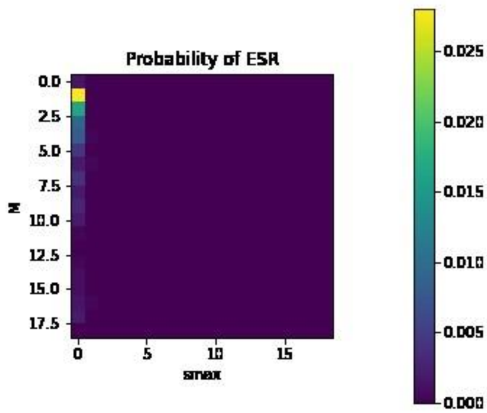
For the Error Curve: As the sparsity level decreases, it becomes harder to maintain a low error rate with the same number of measurements. This is visible from the overall shift from blue to yellow as you move right, indicating that less sparse signals inherently have a higher reconstruction error, given the same number of measurements.

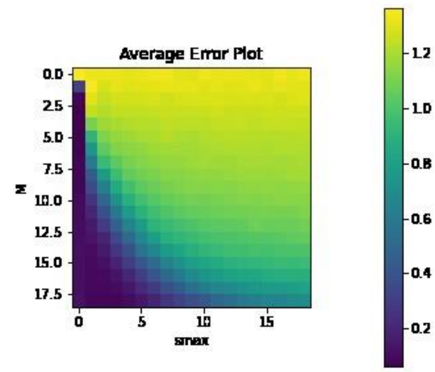
Both the curves are inverses of each other in this case.

4 a) **Noisy Case: S is known, Noise is unknown**

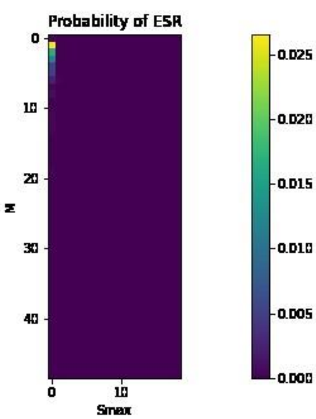
Sigma=0.1

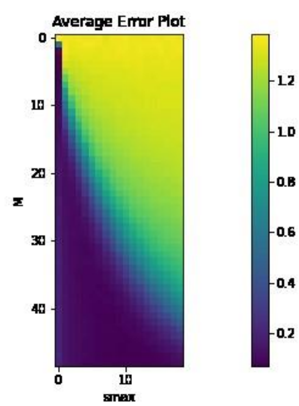
1) N=20



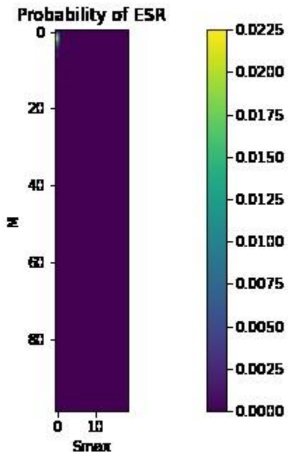


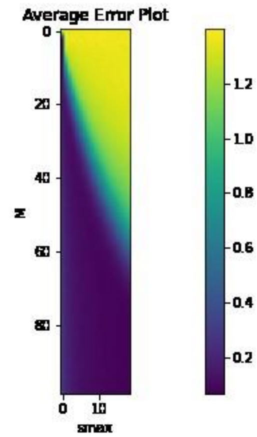
2) N=50





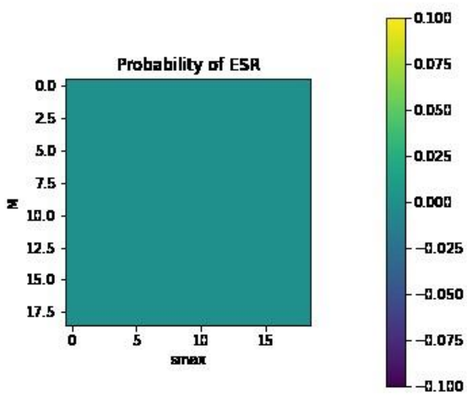
3) N=100

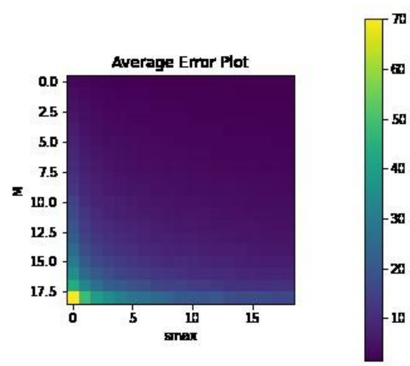




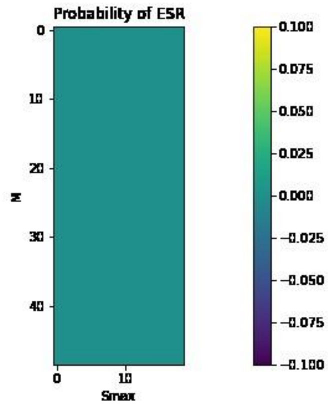
Sigma=10

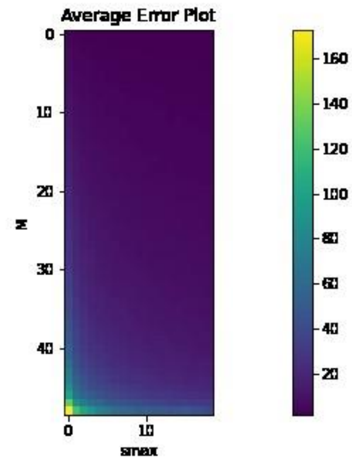
1) N=20



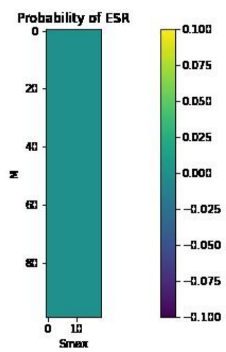


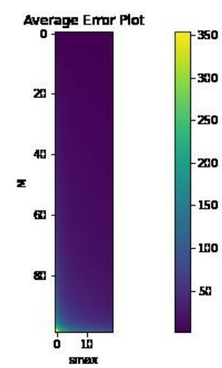
2) N=50





3) N=100



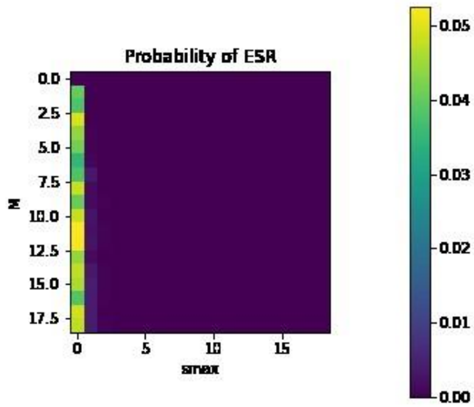


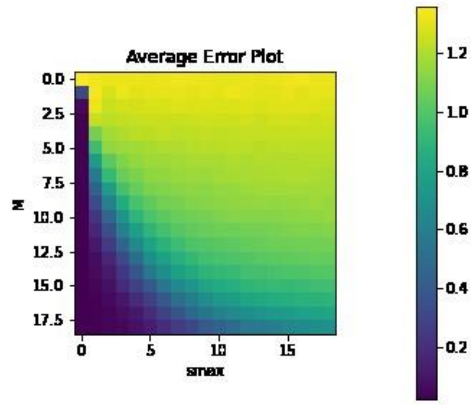
Comment: With added noise, and observing that even in Smax>>M, the probability of exact support recovery is low with the current definition of ESR. Also as the value of sigma increases, The probability of ESR is just a constant value which might indicate that it fails to perform OMP. In case of low sigma we can see that the error is large in the heatmap, but when sigma is significantly higher, and it fails to perform the OMP based on the ESR results, there's no significant output I can interpret here for the error.

4b) **S is Unknown, Noise is Known**

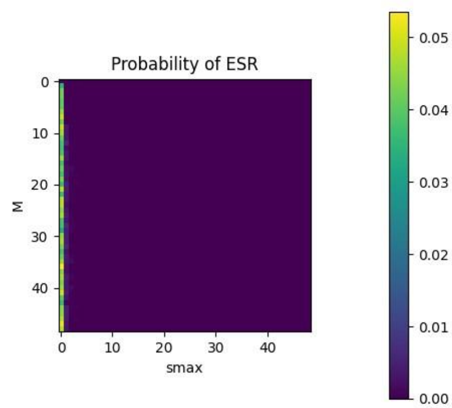
Sigma=0.1

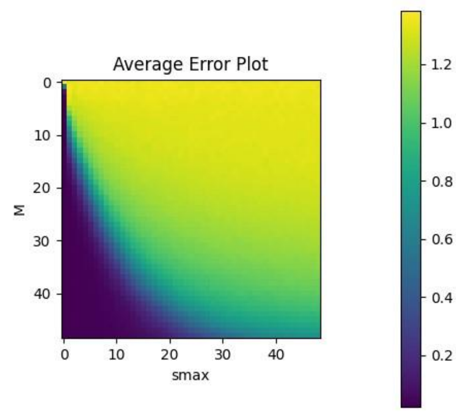
1) N=20



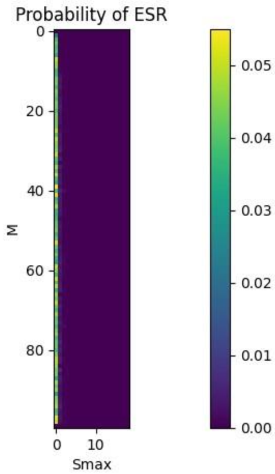


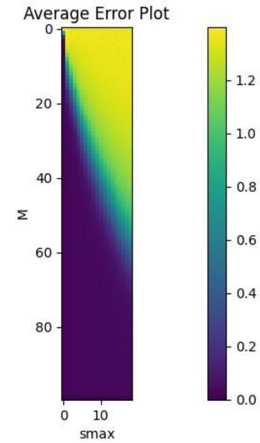
2) N=50





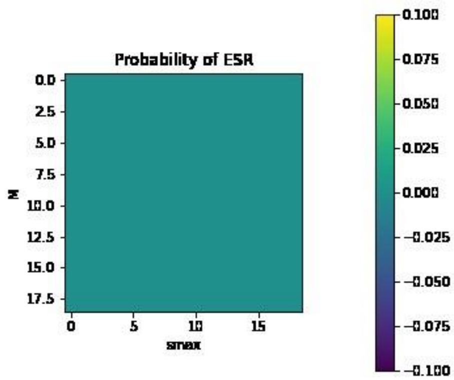
3) N=100

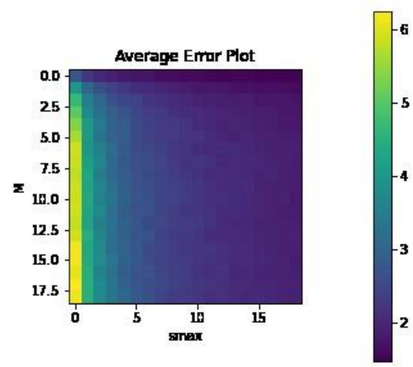




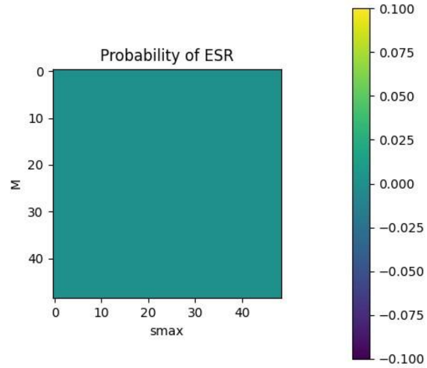
Sigma = 10

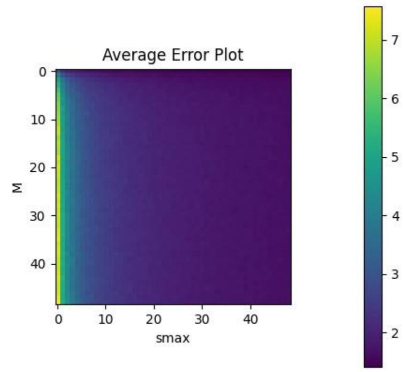
1) N=20



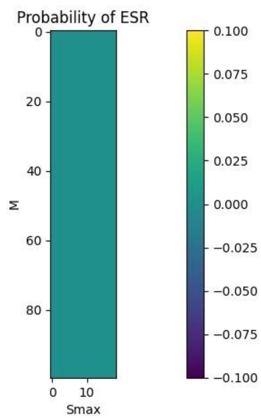


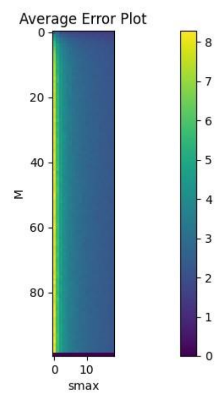
2) N=50





3) N=100



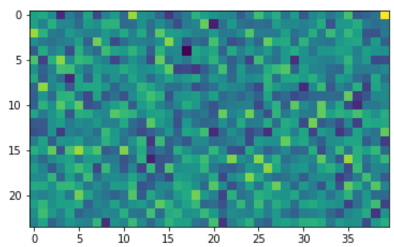


Comment: This one was faster to compute and I think for the lower value sigma the performance of the OMP is quite similar in case of 4 a) . Both the trends of Error and ESR are similar. In the larger sigma case I see there is some change in the Error pattern initially as it stops when it reaches a certain level of noise threshold, but there is no improvement in ESR, it’s the same as in case 4a). So, with noise it is hard to recover the signal.

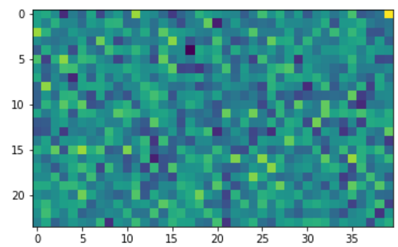
5)

a)

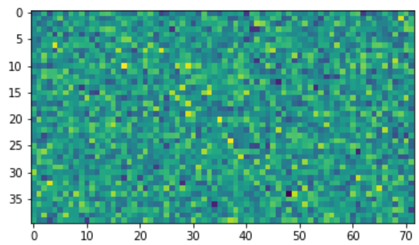
1) Y1



2) Y2



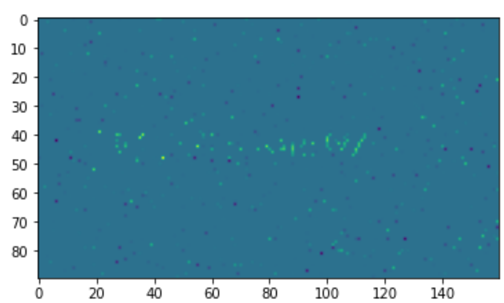
3) Y3

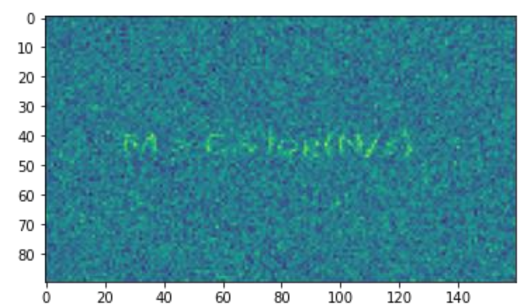


This is what I got when I plotted the Y’s. The image is not decodable.

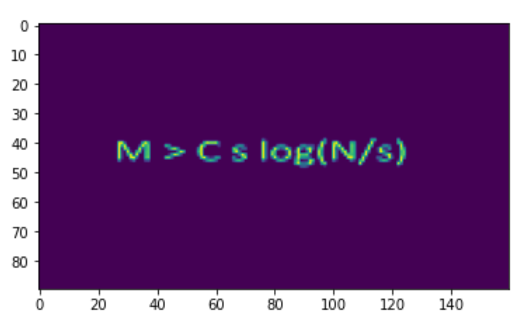
b)

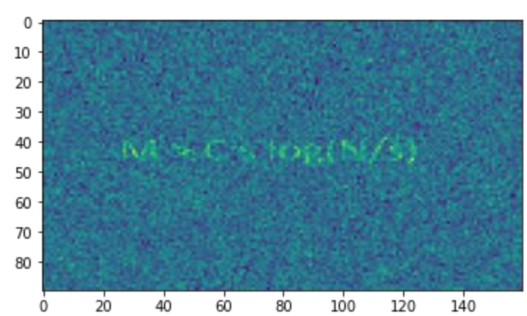
1) Y1 (RHS – OMP, LHS – Least squares solution)



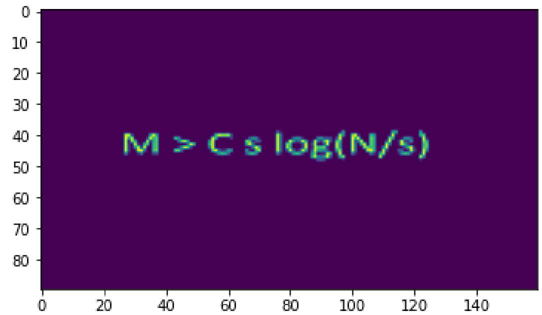


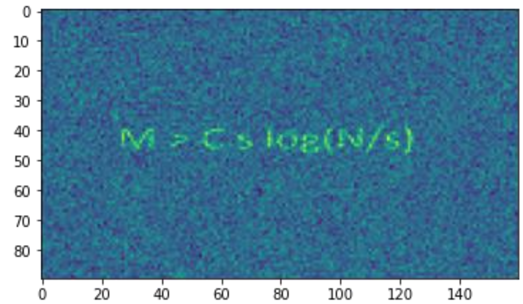
2) Y2 (RHS – OMP, LHS – Least squares solution)





3)Y3 (RHS – OMP, LHS – Least squares solution)





Comment: The Least squares solution is always noisy compared to the OMP for sparse signals. The exception can be made in case of Y1, where the message was decodable in case of least squares. One of the reasons that OMP outperforms Least squares here is that OMP is designed specifically for problems where the solution is sparse while Least squares does not inherently account for sparsity.

c) I got almost the same results for Y2 and Y3. If Y2 and Y3 were more sparsely represented than Y1 at the given measurement levels, this would've led to their better reconstruction quality as OMP excels in sparse signal recovery. Also, if Y1 is less sparse or it's sparsity is more spread out, 960 measurements may be insufficient for reconstruction. For Y2 and Y3, 1440 and 2880 measurements might have been closer to or exceeded the required number of measurements to capture their essential sparse features.

d) The inequality M>Cslog(N/2) might suggest that the number of measurements should be proportional to the sparsity level, multiplied by a logarithmic factor that depends on the size of the original signal N. It sets a lower bound on the number of measurements needed to ensure that a sparse signal can be recovered from its compressed form. C might be a constant that encapsulates factors like the incoherence of the measurement matrix. Maybe this implies that even if the original signal is very large you don't need to sample every component to recover it, instead it could be dictated by how sparse the signal is. The fewer the non-zero elements, the fewer measurements needed.

6)

a) No, I could not decode the message.

b) The message is **"I love Linear Algebra!".** The least squares solution output was very noisy and was audible around k=1000 while in case of OMP I could hear the message when k was equal to 50 but was noisy, gradually increasing k helped make the sound clearer.

c) I could hear the message at Kmin=50. Kmin being the number of measurements, depends on the following:

1) Sparsity of s: The fewer non-zero elements of s has, the fewer measurements are needed to recover it.

2) Size of s: Bigger the dimension of s, more measurements are needed to recover the signal, even if the sparsity is maintained.

3) Characteristics of the Matrix A and Basis D: How well the rows are spread out in terms of linear independence also affects the number of measurements needed in case of Matrix A. Also as A is being multiplied by D, D will also have an effect of the number of Measurements. If D is well-chosen such that it efficiently represents the signal x in a sparse manner, fewer measurements may be required. Also, minimum correlation among the columns of the basis D is preferred for sparse recovery as that makes each element of D more distinct.

d) As it's an audio signal, I think about Fourier Transform. If the audio signal is periodic or has a strong frequency component, a Fourier Basis might have been used.

Steps to make a Fourier Basis Matrix:

1) For a discrete signal of length N, create a basis matrix where each column is a basis vector corresponding to a specific frequency.

2) Each basis vector DFTnk can be represented as:

$$DFT_{nk} = e^{-i2\pi nk/N}$$

   Where n=0,1,.. N-1 and k is the frequency index.

3) Normalize the DFT matrix by 1/√N .

   Other methods that I think can be used involve Wavelet transform and Statistical approaches such as PCA.