Finding Sparse Solutions via Orthogonal Matching Pursuit (OMP)

Mazeyu Ji, PID:A59023027

***Abstract*—This project aims to explore the efficacy of the Orthogonal Matching Pursuit (OMP) algorithm in recovering sparse signals and decoding compressed information by implementing and evaluating the OMP algorithm. The project is divided into four parts: the first part investigates the capability of recovering sparse signals in a noiseless environment; the second part proceeds with recovery in the presence of noise, exploring both known and unknown sparsity scenarios; the third part decodes hidden information from a compressed image using the OMP algorithm; the fourth part decodes hidden information from a compressed audio signal similarly. Through these experiments, the performance of the OMP algorithm under different scenarios is assessed, including its adaptability to changes in compression levels and noise conditions. The findings not only demonstrate the powerful recovery capabilities of the OMP algorithm but also provide deep insights into sparse signal processing.**

***Index Terms*—Orthogonal Matching Pursuit (OMP), Sparse Signal Recovery, Compressive Sensing, Audio Signal Processing, Image Decoding**

I. INTRODUCTION

THE Orthogonal Matching Pursuit (OMP) algorithm is widely used in the field of signal processing, particularly in sparse signal recovery and compressive sensing. The goal of the algorithm is to select a subset of atoms from an overcomplete dictionary that best represents the original signal. It iteratively selects the dictionary atom most correlated with the residual, updating the residual to approximate the original signal. OMP is favored for its simplicity and efficiency, especially in dealing with sparse signals and image recovery issues. This project implements and evaluates the OMP algorithm with the aim of recovering sparse signals and decoding compressed information, deeply exploring the performance and efficiency of the OMP algorithm under various conditions.

II. PROBLEM FORMULATION

Consider the measurement model$y = Ax + n$, where $y \in \mathbb{R}^M$ is the (compressed, $M < N$) measurement, $A \in \mathbb{R}^{MxN}$ is the measurement matrix, and $n \in \mathbb{R}^M$ is the additive noise. Here, $x \in \mathbb{R}^N$ is the unknown signal (to be estimated) with $s \leq N$ non-zero elements. The indices of the non-zero entries of x (also known as the support of x) is denoted by $S = \{i | x_i \neq 0\}$, with $|S| = s$.

Performance Metrics:

Let $\hat{x}$ be the estimate of x obtained from OMP. To measure the performance of OMP, we consider the Normalized Error defined as

$$\frac{\|x - \hat{x}\|_2}{\|x\|_2}$$

The average Normalized Error is obtained by averaging the Normalized Error over 2000 Monte Carlo runs.

Experimental setup:

(a) Generate A as a random matrix with independent and identically distributed entries drawn from the standard normal distribution. Normalize the columns of A.

(b) Generate the sparse vector x with random support of cardinality s (i.e. s indices are generated uniformly at random from integers 1 to N), and non-zero entries drawn as uniform random variables in the range $[-10, -1] \cup [1, 10]$.

(c) The entries of noise n are drawn independently from the normal distribution with standard deviation σ and zero mean.

(d) For each cardinality $s \in [1, s_{max}]$, the average Normalized Error should be computed by repeating steps (a) to (c) 2000 times and averaging the results over these 2000 Monte Carlo runs.

(e) For the memory limitation and faster processing speed, we set $M_{max} = N/2$ and $s_{max} = N/2$, which is enough to analyze the problem.

III. NOISELESS CASE

From the "Noiseless Phase Transition Plot," it is evident that as the sparsity s_{max} increases, the required number of measurements M also increases. This is because, within the OMP algorithm framework, the greater the sparsity of the signal, the more measurements are needed to acquire enough information to recover the original signal. In other words, the greater the sparsity of the signal, the more challenging it is for the OMP algorithm to determine the non-zero elements of the signal (also known as the support set). It can be observed from the plot that the probability of successfully recovering a sparse signal increases with the number of measurements. When the number of measurements is less than the sparsity, it is almost impossible to recover successfully, which is consistent with the intuition of information theory.

For the "Average Normalized Error Plot", when the number of measurements M is smaller and s_{max} is larger, the error is larger. This is due to the OMP algorithm attempting to recover a more sparse signal with limited measurements, and the available information is insufficient to accurately reconstruct the original signal, leading to a larger error. The complementary colors of the two plots reflect the opposing relationship between the number of measurements and the probability of successful recovery and error.

Additionally, as the value of N changes to 20, 50, and 100, the resolution of the generated phase transition plots becomes higher, and the boundary lines are more pronounced. This change in trend is not linear and appears to be logarithmic.

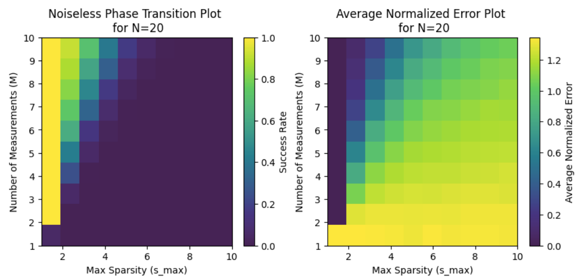

Fig. 1. Noiselesss phase transition plots for N=20

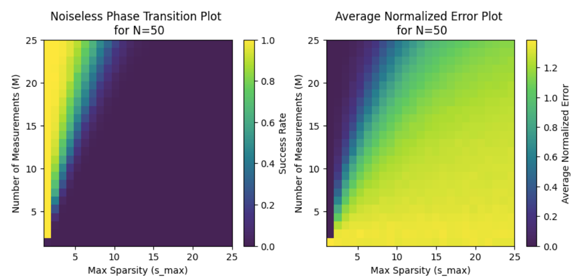

Fig. 2. Noiseless phase transition plots for N=50

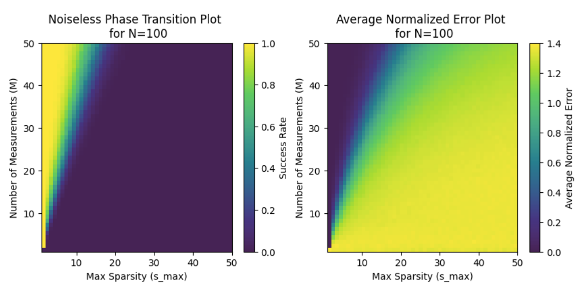

Fig. 3. Noiseless phase transition plots for N=100

IV. Noisy case

A. The sparsity is known

When the noise level is low (with $\sigma = 0.0001$), the recovery rate of the OMP algorithm is relatively high and closer to the noiseless case, but there is still a slight overall decrease in the success rate. The boundary line between success and failure tends to be more linear, suggesting that the relationship between the required number of measurements and the sparsity level is becoming more direct as the noise decreases.

When the noise level is increased to $\sigma = 0.001$, there is a significant drop in the recovery rate, and only at very low sparsity levels does the possibility of recovery remain. An interesting observation is that the maximum error occurs when the sparsity s and the number of measurements are the same. This might be because, at high sparsity levels, the OMP algorithm can still select a certain number of correct entries first, resulting in a relatively lower error.

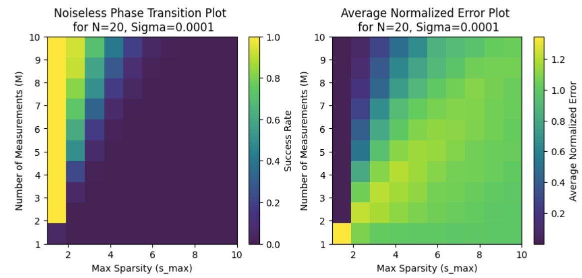

Fig. 4. Noisy phase transition plot for N=20, Sigma=0.0001 with known sparsity

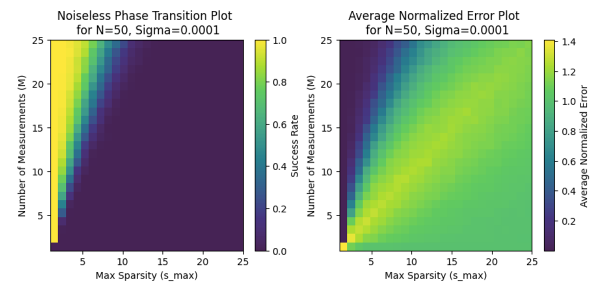

Fig. 5. Noisy phase transition plot for N=50, Sigma=0.0001 with known sparsity

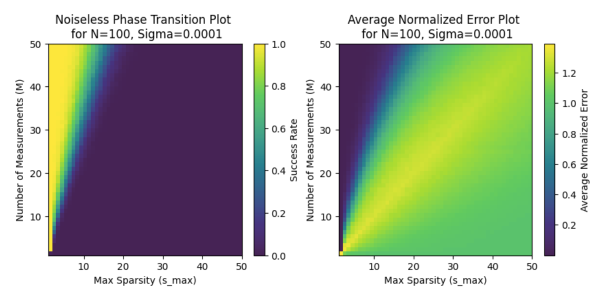

Fig. 6. Noisy phase transition plot for N=100, Sigma=0.0001 with known sparsity

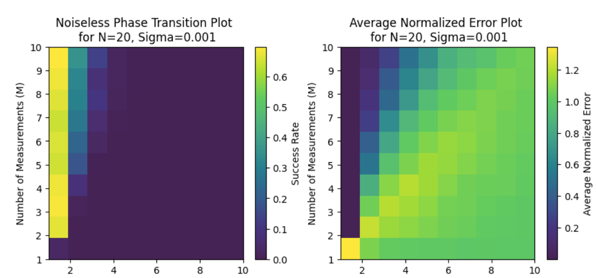

Fig. 7. Noisy phase transition plot for N=20, Sigma=0.001 with known sparsity

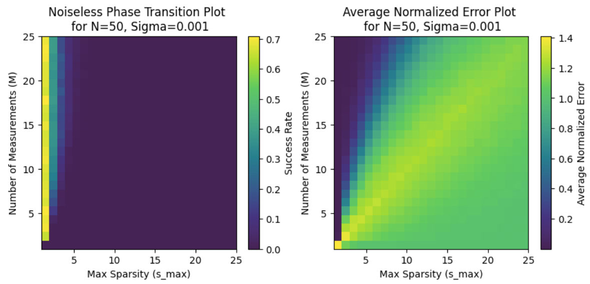

Fig. 8. Noisy phase transition plot for N=50, Sigma=0.001 with known sparsity

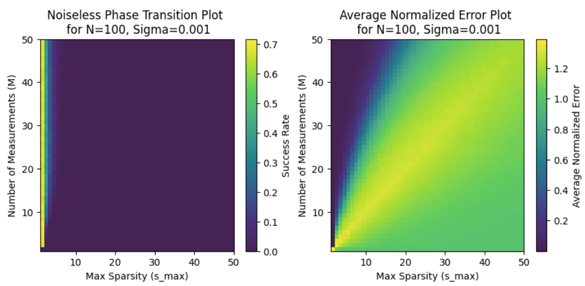

Fig. 9. Noisy phase transition plot for N=100, Sigma=0.001 with known sparsity

B. The norm of the noise is known

When the noise level is low (e.g., $\sigma = 0.0001$), the recovery rate is relatively high, even higher than the case where the sparsity s is known, almost akin to the noiseless situation. This indicates that even without knowledge of the sparsity, the OMP algorithm can achieve a high recovery rate at low noise levels with a good stopping criterion. However, when the noise level increases to $\sigma = 0.001$, the recovery rate significantly drops, and recovery is only probable at extremely low sparsities, similar to the results when sparsity is known. Unlike the known sparsity scenario, the error plot becomes closer to the noiseless condition under this unknown sparsity condition, reflecting the improved recovery effect when the algorithm has an appropriate stopping criterion.

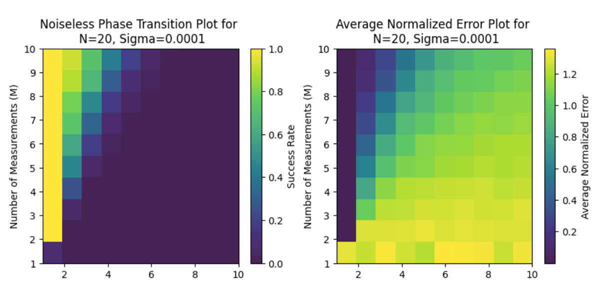

Fig. 10. Noisy phase transition plot for N=20, Sigma=0.0001 with known norm

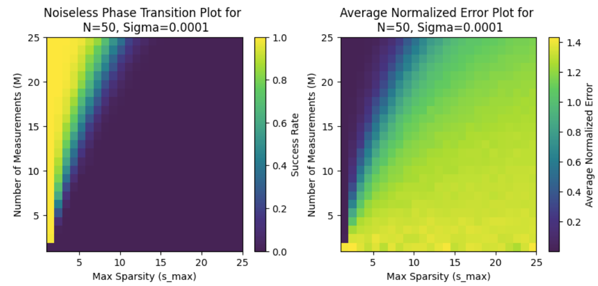

Fig. 11. Noisy phase transition plot for N=50, Sigma=0.0001 with known norm

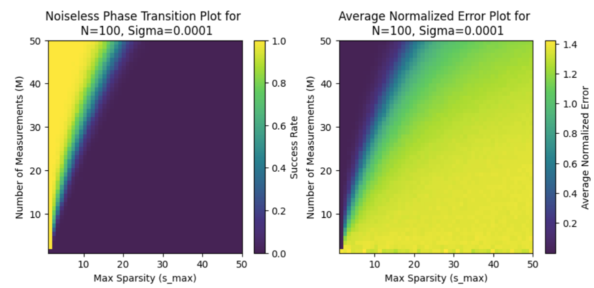

Fig. 12. Noisy phase transition plot for N=100, Sigma=0.0001 with known norm

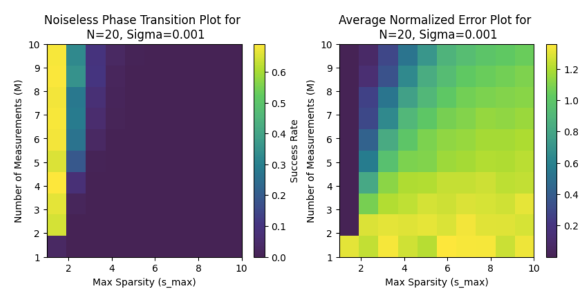

Fig. 13. Noisy phase transition plot for N=20, Sigma=0.001 with known norm

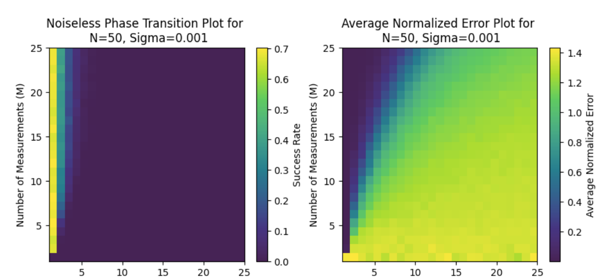

Fig. 14. Noisy phase transition plot for N=50, Sigma=0.001 with known norm

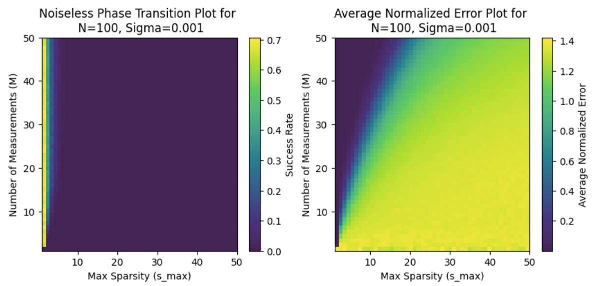

Fig. 15. Noisy phase transition plot for N=100, Sigma=0.001 with known norm

V. Decode a Compressed Message

A. Guess from the compressed images

As shown in the figure, by adjusting the original signal to an aspect ratio of approximately 9:16, no information can be extracted from the image, only mosaic patterns are visible.

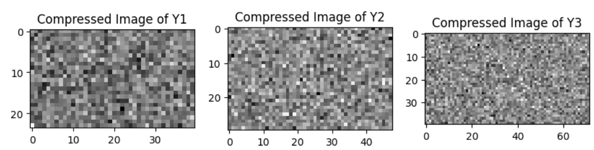

Fig. 16. The compressed images.

B. Results and comparison with the Least Squares Solution

In the image reconstruction using the Orthogonal Matching Pursuit (OMP) algorithm, I have established stopping criteria when the residual is less than 1e-3, and I have also considered limiting the maximum sparsity, which is the number of non-zero pixels. As we progress from A1 to A3, the number of iterations decreases successively. The outcomes demonstrate that, with the exception of the image compressed with A1, which could not be restored, the images compressed with A2 and A3 were almost perfectly reconstructed by the OMP algorithm. On the other hand, the images restored using the Least Squares Solution exhibited significant noise. Notably, the image compressed with A1 seemed to be better restored using the Least Squares Solution than with OMP, while the restorations of images compressed with A2 and A3 were much worse. This discrepancy might be due to an insufficient number of iterations in the OMP algorithm.

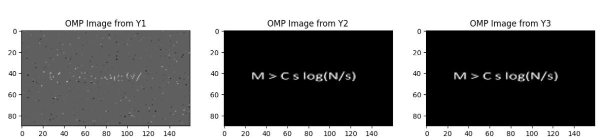

Fig. 17. The recovered images with OMP.

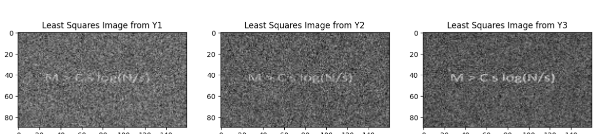

Fig. 18. The recovered images with Least Squares.

C. Results analysis for recovery performance

Through experiments compressing images using three different matrices (A1, A2, and A3) and recovering them using the Orthogonal Matching Pursuit (OMP) algorithm, different outcomes can be observed. The results indicate that after compression with the A1 matrix, the OMP algorithm fails to effectively recover the image, while the A2 and A3 matrices successfully restore the image with clear quality. Further analysis reveals that the A1 matrix compresses the data to 960, while the A2 and A3 matrices compress the data to 1440 and 2880 respectively. From the extracted images, it can be inferred that the image signal is sparse, with only a few pixels having values. Because the measurements of the A2 and A3 matrices are sufficient, the compressed signals contain complete data, enabling successful image restoration. This demonstrates that even with a reduced compression ratio, when compression contains complete information, it does not greatly affect the recovery results. This experimental result underscores the importance of matrix selection in image recovery when using the OMP algorithm.

D. Meaning of the compressed message

The inequality $M > C \cdot s \cdot \log(\frac{N}{s})$ encapsulates a fundamental principle of compressive sensing (CS), a technique aimed at efficiently acquiring and reconstructing sparse signals. Here, M denotes the minimum number of measurements required to accurately recover a signal, N represents the original signal's dimensionality, s is the signal's sparsity indicating the number of non-zero coefficients in a suitable basis, and C is a constant that depends on the specifics of the recovery algorithm and the desired probability of successful recovery. This inequality describes the lower bound condition on the minimum number of measurements M required to successfully recover a length-N signal that is approximately s-sparse, i.e., has approximately s non-zero elements. This result suggests that as long as the number of measurements M is sufficiently large, exceeding a logarithmic threshold related to the signal length N and sparsity s, it is possible to accurately reconstruct the original sparse signal through some recovery algorithm.

VI. Decode a Compressed Audio Signal

A. Guess from the compressed signal

Playing the compressed signal directly results in hearing only noise, without obtaining any information.

B. Results and comparison with the Least Squares Solution

The message embedded within the audio signal is "I love linear algebra." Upon utilizing the Orthogonal Matching Pursuit (OMP) algorithm for signal recovery, the clarity of the message varied with the number of elements K selected from the measurement matrix A and the compressed signal y.

For $K = 10$, the OMP reconstruction produced only noise, and the message was indiscernible. However, as K increased to the range of 50 to 300, the spoken phrase "I love linear algebra" became increasingly audible, accompanied by diminishing noise. Notably, upon reaching $K = 1000$ and beyond, the noise was virtually eliminated, indicating a highly accurate recovery of the signal.

In contrast, the Least Squares Solution did not perform as effectively. Within the same range of $K = 10$ to $K = 300$, the output consisted solely of noise, with no discernible message. A slight improvement was observed at $K = 1000$, where a faint voice emerged amidst a significant level of noise. Further enhancement was noted for $K = 2000$ and $K = 3000$, where the message was clearer, yet the noise remained pronounced.

In comparison between the two methods, OMP demonstrated superior performance, effectively recovering the audio signal and clarifying the message at a lower value of K.

C. Results analysis for K

The minimum number of measurements required to understand the message from the reconstructed audio signal, referred to as K_{min}, is 50 based on the provided observations. This number is influenced by the sparsity level of the signal, the total size of the sparse vector, and the characteristics of the measurement matrix. In this case, since s is a 100-sparse vector of size 15980, K_{min} must be sufficient to capture enough non-zero elements of s to reconstruct the signal with adequate clarity.

Exceeding K_{min} typically enhances the signal quality, reduces the impact of noise, and provides a more stable and accurate reconstruction, especially in real-world scenarios where signals may be subject to various disturbances or deviations from ideal conditions. However, there is a trade-off in terms of computational cost and efficiency, as adding more measurements increases the complexity of the reconstruction process.

D. How to construct D

To construct a basis D that allows for a sparse representation of an audio signal x, one typically uses methods that identify a representation that captures the underlying structure or patterns in the data with as few non-zero coefficients as possible. Since D is of size $(15980, 15980)$, we are looking for a square matrix that can sparsely represent the signal when multiplied with the sparse vector s. Here are some methods and considerations for constructing such a basis:

Fourier Transform Basis: If the signal is composed of a few frequency components, a Fourier basis could efficiently sparsify such signals by representing them in the frequency domain.

Wavelet Transform Basis: Wavelets are particularly well-suited for signals with sharp discontinuities or non-periodic features. A wavelet basis could be used to sparsely represent signals that vary at multiple scales.

Learning-Based Approaches: Methods like dictionary learning can be used to learn a basis D from a large set of example signals that are representative of the type of signals we wish to sparsify. The learned dictionary D would be specifically tuned to provide a sparse representation of signals similar to x.

REFERENCES

[1] Tropp, Joel A. ”Greed is good: Algorithmic results for sparse approximation.” *IEEE Transactions on Information theory* 50.10 (2004): 2231-2242.