Matching Pursuit Alghorithm for Sparse Signal Decomposition

Ghazaleh Hadian Ghahfarokhi,7216547  
*Dortmund University of Applied Sciences and Arts*   
ghazaleh.hadianghahfarokhi001@stud.fh-dortmund.de

*Abstract*— This paper presents the implementation of the Matching Pursuit algorithm, replicating the animation of signal decomposition as depicted on the Matching Pursuit Wikipedia page. The algorithm iteratively decomposes a signal into a sparse representation using a redundant dictionary of basis functions, selecting the best matching atom at each step. The implementation, developed in MATLAB, demonstrates this decomposition process through dynamic animation, illustrating how the signal is progressively approximated over successive iterations. This approach highlights the flexibility and adaptability of Matching Pursuit in sparse signal representation, particularly in non-stationary signal analysis, and showcases its potential applications in modern signal processing.

Keywords— *Matching Pursuit Algorithm, Signal Decomposition, Redundant Dictionary, Sparse Representation, Signal Processing*

# Introduction

In recent years, the demand for efficient signal processing techniques has increased due to the explosive growth of data in various fields such as communications, image processing, and machine learning. A key approach to processing high-dimensional signals is sparse signal representation, where a signal is expressed as a linear combination of a small number of basis functions or atoms from an overcomplete dictionary. Among the methods for sparse signal decomposition, the **Matching Pursuit (MP) algorithm** has garnered significant attention for its simplicity and adaptability.

Originally introduced by Mallat and Zhang in 1993, the Matching Pursuit algorithm iteratively decomposes a signal into a linear combination of dictionary elements that best approximate the signal's structure. Unlike traditional transforms, such as the Fourier or wavelet transforms that rely on fixed bases, MP adaptively selects atoms that match the signal's features, making it highly effective for sparse signal decomposition. MP has found extensive application in areas such as image compression, denoising, and compressed sensing.

A key feature of the MP algorithm is its ability to iteratively refine the signal approximation by subtracting the contribution of the selected atom from the signal at each step. This process continues until a stopping criterion is met, typically based on the residual error or the number of iterations. The result is a sparse representation of the signal, where only a few nonzero coefficients are required to accurately approximate the original signal.

This paper focuses on the implementation and visualization of the Matching Pursuit algorithm for sparse signal decomposition. In particular, we replicate the dynamic animation of the MP algorithm shown on the Wikipedia page for Matching Pursuit. The animation illustrates the iterative decomposition of a signal into its sparse components, highlighting the algorithm's step-by-step selection of atoms and the resulting sparse representation.

Using MATLAB, I developed a script to simulate the animation, visualizing the signal approximation process and the evolution of the sparse coefficients over time. The implementation demonstrates the effectiveness of MP in decomposing signals and provides insights into the behavior of the algorithm, particularly in terms of the impact of dictionary selection and convergence properties. Through this project, we aim to bridge the theoretical aspects of the MP algorithm with its practical application and visualization.

This paper is organized as follows: Section 2 provides a brief overview of the Matching Pursuit algorithm, including its mathematical formulation. Section 3 details our implementation in MATLAB, focusing on the construction of the dynamic animation. Section 4 presents the results of our experiments, comparing the animation with theoretical expectations and highlighting key observations. Finally, Section 5 concludes the paper with reflections on the MP algorithm's performance and suggestions for future improvements.

## Matching Persuit applications

In the Matching pursuit has been applied to signal, image and video coding, shape representation and recognition, 3D objects coding, and in interdisciplinary applications like structural health monitoring. It has been shown that it performs better than [DCT](https://en.wikipedia.org/wiki/Discrete_cosine_transform) based coding for low bit rates in both efficiency of coding and quality of image. The main problem with matching pursuit is the computational complexity of the encoder. In the basic version of an algorithm, the large dictionary needs to be searched at each iteration. Improvements include the use of approximate dictionary representations and suboptimal ways of choosing the best match at each iteration (atom extraction). The matching pursuit algorithm is used in MP/SOFT, a method of simulating quantum dynamics.

MP is also used in [dictionary learning](https://en.wikipedia.org/wiki/Sparse_dictionary_learning). In this algorithm, atoms are learned from a database (in general, natural scenes such as usual images) and not chosen from generic dictionaries.

A very recent application of MP is its use in linear computation coding to speed-up the computation of matrix-vector products‎[1].

# matching pursuit theory

The Matching Pursuit (MP) algorithm is an iterative greedy algorithm used for sparse signal decomposition. The goal of the algorithm is to approximate a given signal as a linear combination of a few elements (or "atoms") chosen from a larger dictionary. This approach is particularly useful when the signal can be represented sparsely, i.e., when only a small number of atoms are needed to describe the signal adequately.

## Mathematical Formulation

Given a signal and a dictionary consisting of normalized atoms , the Matching Pursuit algorithm seeks to express as a sum of a few dictionary elements. Mathematically, we approximate as:

where are the coefficients, are the selected atoms from the dictionary, and is the number of iterations or the desired level of sparsity.

The MP algorithm proceeds as follows:

1. **Initialization**: Start with the original signal and define the initial residual . Set the iteration counter.
2. **Atom Selection**: At each iteration , select the atom from the dictionary that maximizes the absolute inner product with the current residual:

This step identifies the atom that has the greatest correlation with the residual signal, which means it captures the largest remaining portion of the signal's structure.

1. **Coefficient Computation**: Compute the coefficient corresponding to the selected atom:

This coefficient quantifies how much of the selected atom contributes to the signal.

1. **Residual Update**: Update the residual by subtracting the contribution of the selected atom from the current residual:

This step removes the part of the signal already represented by the selected atom, leaving behind a new residual for the next iteration.

1. **Iteration**: The process is repeated by incrementing , selecting new atoms, and updating the residual until a stopping criterion is met. Common stopping criteria include reaching a predefined number of iterations or achieving a sufficiently small residual norm .

## Convergence and Properties

The Matching Pursuit algorithm guarantees that the residual norm decreases with each iteration. However, since it is a greedy algorithm, MP does not always yield the sparsest possible representation of the signal. Its performance depends on the structure of the dictionary and the signal being analyzed. In practice, MP is often used when computational efficiency and simplicity are more critical than obtaining the absolute sparsest solution.

A variation of the algorithm, Orthogonal Matching Pursuit (OMP), enforces orthogonality between the residual and the chosen atoms, leading to potentially faster convergence and sparser solutions. However, OMP is computationally more demanding.

Matching Pursuit is widely applied in signal processing tasks such as compression, denoising, and compressed sensing, were signals exhibit sparsity in specific domains.

## Algorithm Summary

To summarize, the key steps of the Matching Pursuit algorithm are:

* Initialize the residual as the original signal.
* Iteratively select atoms from the dictionary that best match the residual.
* Compute the corresponding coefficients and update the residual.
* Repeat until the signal is sufficiently decomposed into a sparse set of atoms.

The ability to adaptively select atoms that match the signal's structure makes MP a powerful tool for sparse signal decomposition and approximation.

# matching pursuit Extensions

Matching Pursuit (MP) has several important variants that address different needs in terms of accuracy, computational efficiency, and performance in signal decomposition. Below are the three primary forms of the algorithm:

## Standard Matching Pursuit (MP):

MP is a greedy algorithm that iteratively selects the atom (a basis function from a predefined dictionary) which has the highest correlation with the current residual of the signal. At each iteration, the residual signal is updated by subtracting the contribution of the selected atom. It is simple and computationally efficient, making it effective for quick approximations and real-time applications. On the other hand, since it doesn’t ensure orthogonality between selected atoms, some redundancy can occur, resulting in less optimal reconstructions in comparison to more refined methods.

## Orthogonal Matching Pursuit (OMP):

OMP extends MP by ensuring that each selected atom is orthogonal to the previously chosen atoms. This is achieved by using orthogonal projection methods during the selection process, ensuring that atoms do not interfere with each other. The orthogonality constraint improves accuracy, reducing redundancy and providing a more stable and efficient signal decomposition. It often converges faster in terms of reaching an accurate approximation. In contract, the cost of ensuring orthogonality increases computational complexity, making it less suitable for real-time or large-scale applications where speed is critical.

## Weak Matching Pursuit (WMP):

WMP relaxes the selection criteria of MP. Instead of choosing the atom with the highest correlation, it selects any atom that meets a weaker correlation condition. This allows for faster computation by reducing the strictness of atom selection. WMP offers a balance between speed and accuracy. It is faster than OMP and can be more efficient than standard MP in some cases where finding the exact maximum correlation is less important. The downside of WMP is the relaxation in selection criteria can lead to a less accurate representation of the signal. The weaker selection condition may result in a slower convergence to the optimal solution compared to OMP.

# Implementation

This section describes the MATLAB implementation of the Matching Pursuit algorithm and the process used to generate the dynamic animation that visually demonstrates the algorithm's iterative decomposition of a signal.

The goal of the implementation was to replicate the dynamic animation found on the Wikipedia page for Matching Pursuit, where the algorithm approximates a signal by selecting atoms from a predefined dictionary. The animation illustrates the evolution of the signal approximation over time as atoms are chosen and coefficients are computed‎[3].

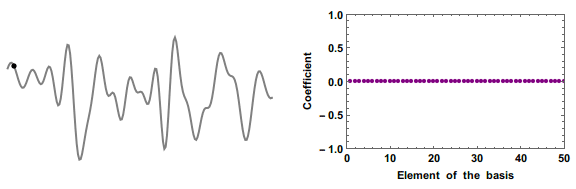


Figure 1: The retrieval of an unknown signal, the original intended animation‎[2]

Retrieving an unknown signal (black line) from a small number of measurements (black dots) is in general impossible. The signal can be represented with only a few nonzero elements of a basis, then “compressed sensing" can be used. By making the measurement in a basis that is incoherent with the basis the signal is sparse in, it is ensured that each measurement sample as many coefficients as possible. In this project, an "orthogonal-matching-pursuit" algorithm is implemented. The starting guess is that all coefficient is zero. At each iteration we add one coefficient that is nonzero, choosing the one that will affect our error metric the most. The values of those few nonzero coefficients are then estimated by a least-square fitting, and the process is iterated until the error is smaller than a given threshold. Let’s delve more into coding structure‎ [‎VI].

## Initial signal generation

The initialization of signals presented here replicates the known values and overall formulation implemented by the Mathematica 11.0 code found in the animation of Orthogonal Matching Pursuit on Wikimedia Commons ‎[2].

In this MATLAB implementation, we define several key elements for the matching pursuit algorithm. The signal's dimensionality and the number of non-zero coefficients are set, with random positions and values assigned to these coefficients within a sparse vector. The original signal is constructed using the sum of a weighted combination of Hermite polynomials functions, applied over a set of values ranging from -7 to 7. A measurement matrix is created to capture 30 sampled positions from the original signal, simulating the signal acquisition process. Additionally, a transformation matrix is constructed using Hermite polynomials, forming the basis for projecting the signal into a new space. A measurement matrix is then formed by multiplying the measurement matrix with the transposed Hermite basis, and the observed signal is calculated using the sampled positions of the original signal.

### Hermite polynomials

The Hermite polynomials are a sequence of orthogonal polynomials that can be defined using the following general recursion formula:

with the initial conditions:

I compute using a function that sums the contributions of Hermite polynomials.

## Error Calculation and Update Algorithm

**Initialization:** The algorithm begins with an initial residual, which represents the difference between the actual measurements () and the predicted values.

**Error Calculation**: While the norm of the residual vector is greater than a small threshold (), the algorithm performs the following steps:

**Compute Error** Vector: For each candidate index in the complement of the support set, the error vector is calculated. For each candidate index , the error is computed using the formula:

where:

* is the -th column of the measurement projection matrix .
* residual is the current residual vector.
* denotes the norm of the column vector .

**Select Best Fit**: Identify the index with the minimum error value. This index corresponds to the column that best reduces the residual.

**Update Error**: Set the current error () to the minimum error value found.

**Update Sets**: Remove the selected index from the complement set and update the support set to include this index. This index is added to the support set and removed from the complement set.

**Update Solution**: Using the new support set, waiting factors are computed by solving the least-squares problem that minimizes the error between the observed measurements and the projections of the signal, given the current set of basis vectors

Where, ​ is the matrix restricted to the current support set columns.

**Update Residual**: Calculate the new residual vector:

**Store Results**: Append the current solution to a list of results.

**Termination**: The loop continues until the norm of the residual is less than or equal to , indicating that the approximation is sufficiently close to the original data.

## **Visualization and Plot Customization**

To ensure clarity in the visualization, the original signal is plotted in **black**. The evolving signal approximation is plotted in **orange** to highlight its progression. The selected measurements are marked as **filled black dots** to show which parts of the signal are being used at each iteration. The coefficients of the selected atoms are plotted in **magenta** using stem plots to depict their evolution. The visualizations are adjusted to have consistent scaling, ensuring that the full range of the signal and coefficients are visible throughout the animation.

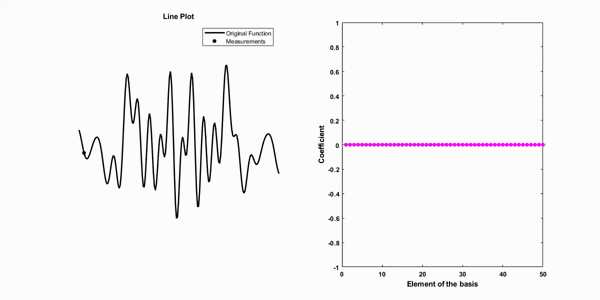


Figure 2: The replicated animation results of OMP algorithm

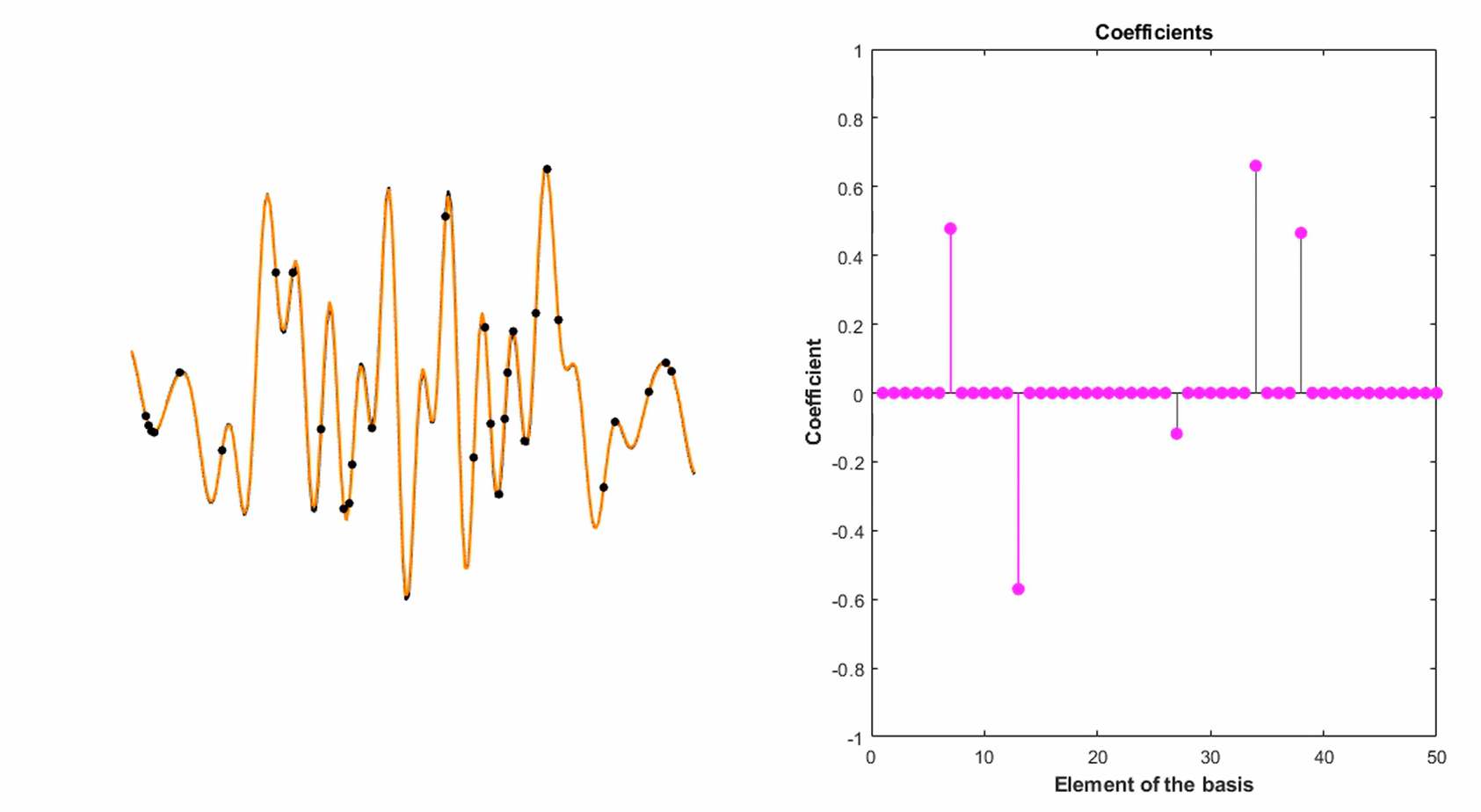


Figure 3: Final result of approximated signal

The implementation of the Matching Pursuit (MP) algorithm successfully decomposed the original signal into a sparse representation using a dictionary of Hermite basis functions. The iterative process illustrated in the animation shows how the algorithm selects the best-fitting atoms at each step to reconstruct the signal. The accuracy of the reconstructed signal, ​, closely matched the original signal ​, validating the effectiveness of MP in sparse signal decomposition. However, some limitations emerged, such as the algorithm's sensitivity to noise and the inherent dependency on the choice of dictionary. These factors may influence the convergence speed and reconstruction accuracy in more complex or noisy signal environments.

Despite these limitations, the results demonstrate the potential of MP in various applications, particularly in signal processing, where sparse representations are critical for compression and denoising tasks. The decomposition of signals into meaningful components highlights MP’s utility in fields like machine learning, where sparse coding can lead to better interpretability, and in communications, where efficient signal representation is essential. Future work could focus on enhancing the computational efficiency of the algorithm or exploring alternative dictionaries to improve performance in diverse signal conditions.

# Summary

This paper implemented the Matching Pursuit (MP) algorithm to replicate the signal decomposition process shown in an animation. Through iterative selection of the best-matching basis functions from a redundant dictionary, the algorithm successfully decomposed the original signal into a sparse representation. This demonstrated the flexibility and effectiveness of MP in processing non-stationary signals and its utility in applications like compression and denoising. While the implementation revealed challenges such as sensitivity to dictionary selection, noise, and computational cost, MP remains a valuable tool with potential for future applications in machine learning, communications, and real-time signal processing.

##### References

1. Angoletta, Maria Elena. "Digital signal processor fundamentals and system design." (2008).
2. Jacopo Bertolotti, "Orthogonal Matching Pursuit," \*Wikimedia Commons\*. [Online]. Available: https://commons.wikimedia.org/wiki/File:Orthogonal\_Matching\_Pursuit.gif. [Accessed: Aug. 30, 2024].
3. "Matching Pursuit Algorithm," Wikipedia: The Free Encyclopedia. [Online]. Available: <https://en.wikipedia.org/wiki/Matching_pursuit_algorithm>. [Accessed: August 30, 2024]

# Matlab Script

clc;

close all;

clear all;

%% Step 1: Initialization

dim = 200;

numOfMeasurmentOfOrigSignal = 30;

numOfCoefficient = 6;

posisionOfCoffs= sort(randsample(40, numOfCoefficient));

valueOfCoeffs = -0.8 + (1.6) .\* rand(1, numOfCoefficient);

% Create a sparse vector

idealX\_CoeffSparseArray = sparse(posisionOfCoffs, 1, valueOfCoeffs, dim, 1);

z\_values = linspace(-7, 7, dim); %(1x200)

f\_original = zeros(1, dim);

for i = 1:dim

% Compute g(z) for each point in z\_values

f\_original(i) = compute\_g(z\_values(i),numOfCoefficient,valueOfCoeffs,posisionOfCoffs);

end

measuringPositionsOfOrigSignal = sort(randperm(dim, numOfMeasurmentOfOrigSignal));

% 30x200

AtomSparseMatrix = sparse(1:numOfMeasurmentOfOrigSignal, measuringPositionsOfOrigSignal, 1, numOfMeasurmentOfOrigSignal, dim);

% Matrix of change of basis from Hermite to canonical

Phi\_HermitePolyMatrix = zeros(dim, dim);

for i = 1:dim

% Basis matrix (Hermite polynomials)

Phi\_HermitePolyMatrix(i, :) = arrayfun(@(x) hermitFunc(i-1, x), z\_values);

end

% Measurement matrix (30x200)

A = AtomSparseMatrix \* transpose(Phi\_HermitePolyMatrix);

f\_hat = f\_original(measuringPositionsOfOrigSignal)';

%% Step 2: Orthogonal Matching Pursuit Algorithm

% Initialization

% Error threshold

epsilon0 = 1e-10;

% Initial guess (all coefficients zero) 200x1

gaussed\_waiting\_factors = zeros(dim, 1);

% Initial residual is the measurement vector

residual = f\_hat;

% Complement of the support (all indices)

indexOfCompOfSupport = 1:dim;

% Calculate the error for all columns of A that are not already in the support

epsilon = arrayfun(@(j) norm((A(:, j)' \* residual / norm(A(:, j))^2)\* A(:, j) - residual)^2, indexOfCompOfSupport);

% Select the one with the smallest error

[~, min\_idx] = min(epsilon);

j\_MinValueErrorIndex = indexOfCompOfSupport(min\_idx);

% Update the error and complement of the support

epsilon = epsilon(j\_MinValueErrorIndex);

% Update the complement of the support

indexOfCompOfSupport(indexOfCompOfSupport == j\_MinValueErrorIndex) = [];

% Update the support

support = setdiff(1:dim, indexOfCompOfSupport);

% Extract the columns of A corresponding to the support set

A\_OfSupport = A(:, support);

% Find the best fit for the new estimate of x (least squares solution)

gaussed\_waiting\_factors(support) = (A\_OfSupport' \* A\_OfSupport) \ (A\_OfSupport' \* f\_hat);

% Update the residuals

residual = f\_hat - A \* gaussed\_waiting\_factors;

tmp = gaussed\_waiting\_factors;

% To store results

result\_waiting\_factors = {};

% Orthogonal Matching Pursuit

while norm(residual) > epsilon0

% Calculate error vector epsilon

epsilon = arrayfun(@(j) norm((A(:, j)' \* residual / norm(A(:, j))^2) \* A(:, j) - residual)^2, indexOfCompOfSupport);

% Select the index with the minimum error (best fit)

[~, min\_idx] = min(epsilon);

% Update epsilon

epsilon = epsilon(min\_idx);

% Remove the selected index from the complement set

indexOfCompOfSupport(min\_idx) = [];

% Update the support set

support = setdiff(1:dim, indexOfCompOfSupport);

% Update the solution for the new support set

A\_OfSupport = A(:, support);

% Least squares solution

gaussed\_waiting\_factors(support) = (A\_OfSupport' \* A\_OfSupport) \ (A\_OfSupport' \* f\_hat);

% Update the residuals

residual = f\_hat - A \* gaussed\_waiting\_factors;

result\_waiting\_factors{end+1} = gaussed\_waiting\_factors;

end

result\_waiting\_factors=[tmp,result\_waiting\_factors];

%% Step 3: Generate the animation

% % Define video dimensions

% Initialize video writer

videoFileName = 'matching\_pursuit\_animation.mp4';

frameRate = 6; % Frame rate

% Create a video writer object

v = VideoWriter(videoFileName, 'MPEG-4');

v.FrameRate = frameRate; % Set frame rate

open(v); % Open the video file for writing

p0 = cell(1, numOfMeasurmentOfOrigSignal);

for k = 1:numOfMeasurmentOfOrigSignal

fig = figure('Visible', 'off');

% Adjust figure size: [left, bottom, width, height]

set(fig, 'Position', [100, 100, 1200, 600]);

% Create the first subplot for the line plot

subplot(1, 2, 1);

plot(f\_original, 'k', 'LineWidth', 2);

hold on;

plot(measuringPositionsOfOrigSignal(1:k), f\_hat(1:k), 'ko', 'MarkerSize', 5, 'MarkerFaceColor', 'k');

ylim([-1, 1]);

legend('Original Function', 'Measurements');

title('Line Plot');

axis off;

% Second plot (horizontal line with shaded area)

subplot(1, 2, 2);

stem(1:50, zeros(1, 50), 'm', 'MarkerSize', 5, 'MarkerFaceColor', 'm');

ylim([-1, 1]);

xlim([0,50]);

xlabel('Element of the basis', 'FontWeight', 'bold');

ylabel('Coefficient', 'FontWeight', 'bold');

set(gca, 'XColor', 'k', 'YColor', 'k', 'FontWeight', 'bold');

p0{k} = gcf; % Save the current figure handle into the cell array

% Capture the current frame and write it to the video

frame = getframe(gcf);

writeVideo(v, frame);

% Close the figure to free up memory

close(gcf);

end

tau\_values = 0:0.1:1; % Tau values for the second loop

% Second Loop: Create the second set of figures and add them to the video

for tau = tau\_values

% Create a new figure for each frame

fig = figure('Visible', 'off');

set(fig, 'Position', [100, 100, 1200, 600]);

subplot(1, 2, 1);

plot(f\_original, 'k', 'LineWidth', 1.5);

hold on;

evolving\_signal = (1 - tau)\*0 + tau\*transpose(Phi\_HermitePolyMatrix)\*result\_waiting\_factors{1};

plot(evolving\_signal, 'Color', [1, 0.5, 0] ,'LineWidth', 1.5); % Plot evolving signal

plot(measuringPositionsOfOrigSignal, f\_hat, 'ko', 'MarkerSize', 4, 'MarkerFaceColor', 'k');

ylim([-1, 1]); % Set y-axis limits

axis off; % Hide axes

% Second subplot (Right): Plot the evolution of coefficients

subplot(1, 2, 2); % Create second subplot

coeffs = tau \* result\_waiting\_factors{1}; % Compute the evolving coefficients

h = stem(1:200, coeffs, 'm', 'MarkerSize', 6, 'MarkerFaceColor', 'm'); % Plot coefficients

% Ensure the color is consistent

set(h, 'Color', 'm'); % Set color for the stem lines

set(h, 'MarkerEdgeColor', 'm'); % Set color for marker edges

set(h, 'MarkerFaceColor', 'm'); % Set color for marker faces

ylim([-1, 1]); % Set y-axis limits

xlim([0,50]);

xlabel('Element of the basis', 'FontWeight', 'bold');

ylabel('Coefficient', 'FontWeight', 'bold');

title('Coefficients'); % Add title

% Capture the current frame and write it to the video

frame = getframe(gcf);

writeVideo(v, frame);

% Close the figure to free up memory

close(gcf);

end

for k = 2:length(support) - 1

for tau = tau\_values

% Create a new figure for each frame

fig = figure('Visible', 'off');

set(fig, 'Position', [100, 100, 1200, 600]); % Adjust figure size: [left, bottom, width, height]

% First subplot: Plot df and evolving signal

subplot(1, 2, 1); % Create the first subplot

plot(f\_original, 'k' ,'LineWidth', 1.5); % Plot df as gray line

hold on;

evolving\_signal\_p2 = (1 - tau) \* Phi\_HermitePolyMatrix' \* result\_waiting\_factors{k - 1} ...

+ tau \* Phi\_HermitePolyMatrix' \* result\_waiting\_factors{k};

plot(evolving\_signal\_p2 ,'Color', [1, 0.5, 0], 'LineWidth', 1.5); % Plot evolving signal

plot(measuringPositionsOfOrigSignal, f\_hat, 'ko', 'MarkerSize', 4, 'MarkerFaceColor', 'k');

ylim([-1, 1]); % Set y-axis limits

axis off; % Hide axes

% Second subplot: Plot coefficients

subplot(1, 2, 2); % Create the second subplot

coeffs\_p2 = (1 - tau) \* result\_waiting\_factors{k - 1} + tau \* result\_waiting\_factors{k};

h2 = stem(1:200, coeffs\_p2, 'm', 'MarkerSize', 6, 'MarkerFaceColor', 'm');

% Ensure the color is consistent

set(h2, 'Color', 'm'); % Set color for the stem lines

set(h2, 'MarkerEdgeColor', 'm'); % Set color for marker edges

set(h2, 'MarkerFaceColor', 'm'); % Set color for marker faces

ylim([-1, 1]); % Set y-axis limits

xlim([0,50]);

xlabel('Element of the basis', 'FontWeight', 'bold');

ylabel('Coefficient', 'FontWeight', 'bold');

title('Coefficients'); % Add title

% Capture the current frame and write it to the video

frame = getframe(gcf);

writeVideo(v, frame);

% Close the figure to free up memory

close(gcf);

end

end

% Finalize and close the video file

close(v);

function g\_val = compute\_g(z,nb,c,nzc)

g\_val = 0;

for j = 1:nb

g\_val = g\_val + c(j) \* hermitFunc(nzc(j), z); % Sum of terms

end

end

function y = hermitFunc(n, x)

% Compute the normalization factor

norm\_factor = sqrt(factorial(n) \* 2^n \* sqrt(pi));

% Define the normalized Hermite function

y = exp(-x.^2 / 2) .\* hermiteH(n, x) / norm\_factor;

end