Probability Theory and Stochastic Processes

Voice Recognition using Cross-Corelation in MATLAB

Assignment – 3

Name: Shiva kalyan Gupta

Roll no: 1602-21-735-115

As we all know that cross correlation gives us the similarity between two signals, I have used the same concept of finding similarities between the two voice signals and recognising the voice.

This is basically done in 5 Steps and they are:

* Making Source files

In this step we take the voice of 10 persons and store them in .wav files by naming each file with the person name.

* Taking Test input

Once the Source files are included into the code hen we take the test input and store it with a variable.

* Finding Correlation

In this step we find the correlation between the Test input and all the available source files and store the maximum possible correlation with each source file in a variable.

* Finding Maximum

From all the available maximum corelation values we choose the maximum value amongst them and store it as an answer.

* Matching and displaying result

Now the answer is matched with all the maximum corelation values of all the source files, now the source file which is matched with the answer i.e.. which has got the highest correlation value with the test input is treated to be the most probable signal and the given signal is identified to be the person to whom the source file belongs.

MATLAB CODE

close all;

clear;

clc;

% audio recording

Fs =16000;

Channels = 1;

bits = 16;

X1=audioread("anirudh.wav");

X2=audioread("sharanya.wav");

X3=audioread("arun.wav");

X4=audioread("nithin.wav");

X5=audioread("saicharan.wav");

X6=audioread("salma.wav");

X7=audioread("rajnesh.wav");

X8=audioread("mahesree.wav");

X9=audioread("shivakalyan.wav");

pause(5);

rt = audiorecorder(Fs,bits, Channels);

duration = 10; disp('recording started test');

recordblocking(rt,duration);

disp('audio recording stopped test');

Xt = getaudiodata(rt);

%graph the audio xt

t = 0:1/Fs:(length(Xt)-1)/Fs;

figure(1); plot(t,Xt,'LineWidth',1.5);

xlabel('time(sec)'); ylabel('Amplitude');

title('audio time domain - xt');

%Find whether they both are of same person or not

[c,lags]=xcorr(Xt,X1);

figure(2);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("anirudh");

anirudh=max(c);

[c,lags]=xcorr(Xt,X2);

figure(3);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("Sharanya");

sharanya=max(c);

[c,lags]=xcorr(Xt,X3);

figure(4);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("nithin")

nithin=max(c);

[c,lags]=xcorr(Xt,X4);

figure(5);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("arun");

arun=max(c);

[c,lags]=xcorr(Xt,X5);

figure(6);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("saicharan");

saicharan=max(c);

[c,lags]=xcorr(Xt,X6);

figure(7);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("salma");

salma=max(c);

[c,lags]=xcorr(Xt,X7);

figure(8);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("rajnesh");

rajnesh=max(c);

[c,lags]=xcorr(Xt,X8);

figure(9);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("mahesree");

mahesree=max(c);

[c,lags]=xcorr(Xt,X9);

figure(10);

stem(lags,c);

xlabel("LAGS");

ylabel("corelation");

title("shivakalyan");

shivakalyan=max(c);

maxvalues=[anirudh sharanya nithin arun saicharan salma mahesree rajnesh shivakalyan];

answer=max(maxvalues);

switch answer

case anirudh

disp("The voice is anirudh's");

case sharanya

disp("The voice is sharanya's");

case nithin

disp("The voice is nithin's");

case arun

disp("The voice is arun's");

case saicharan

disp("The voice is saicharan's");

case salma

disp("The voice is salma's");

case mahesree

disp("The voice is mahesree's");

case rajnesh

disp("The voice is rajnesh's");

case shivakalyan

disp("The voice is shivakalyan's");

otherwise

disp("The Voice is not yet registered");

end

The Output of the above code is:

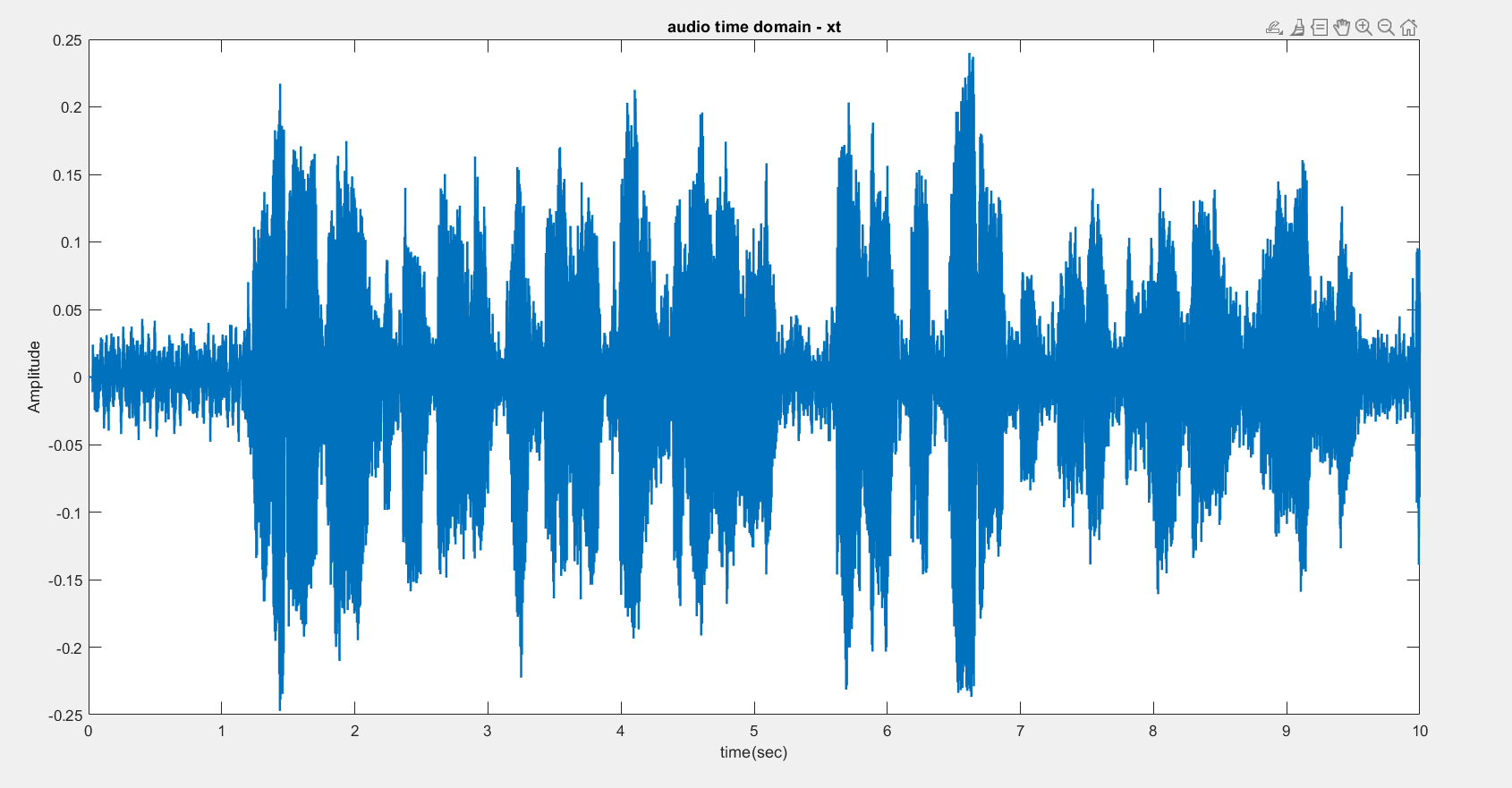
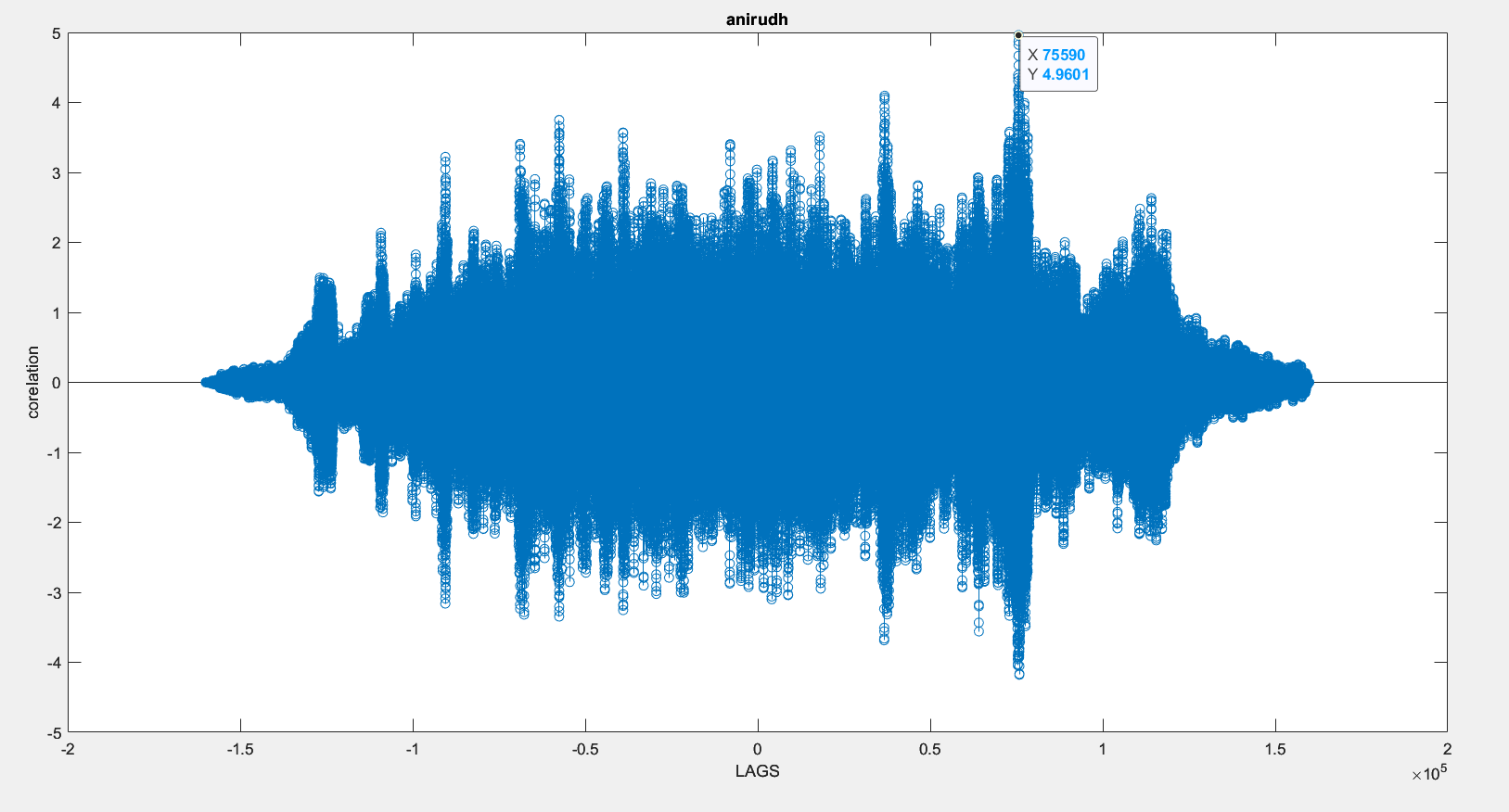
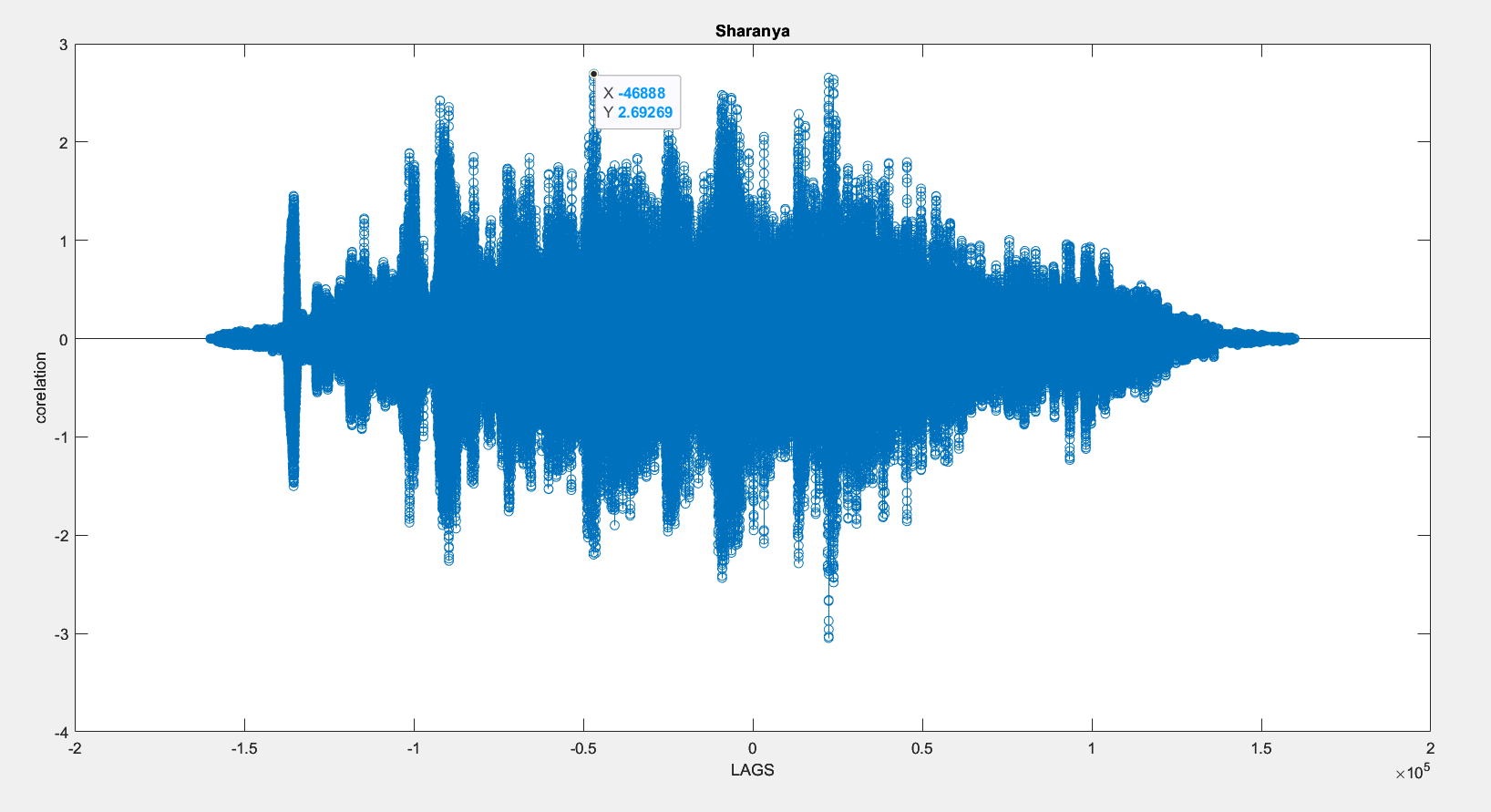
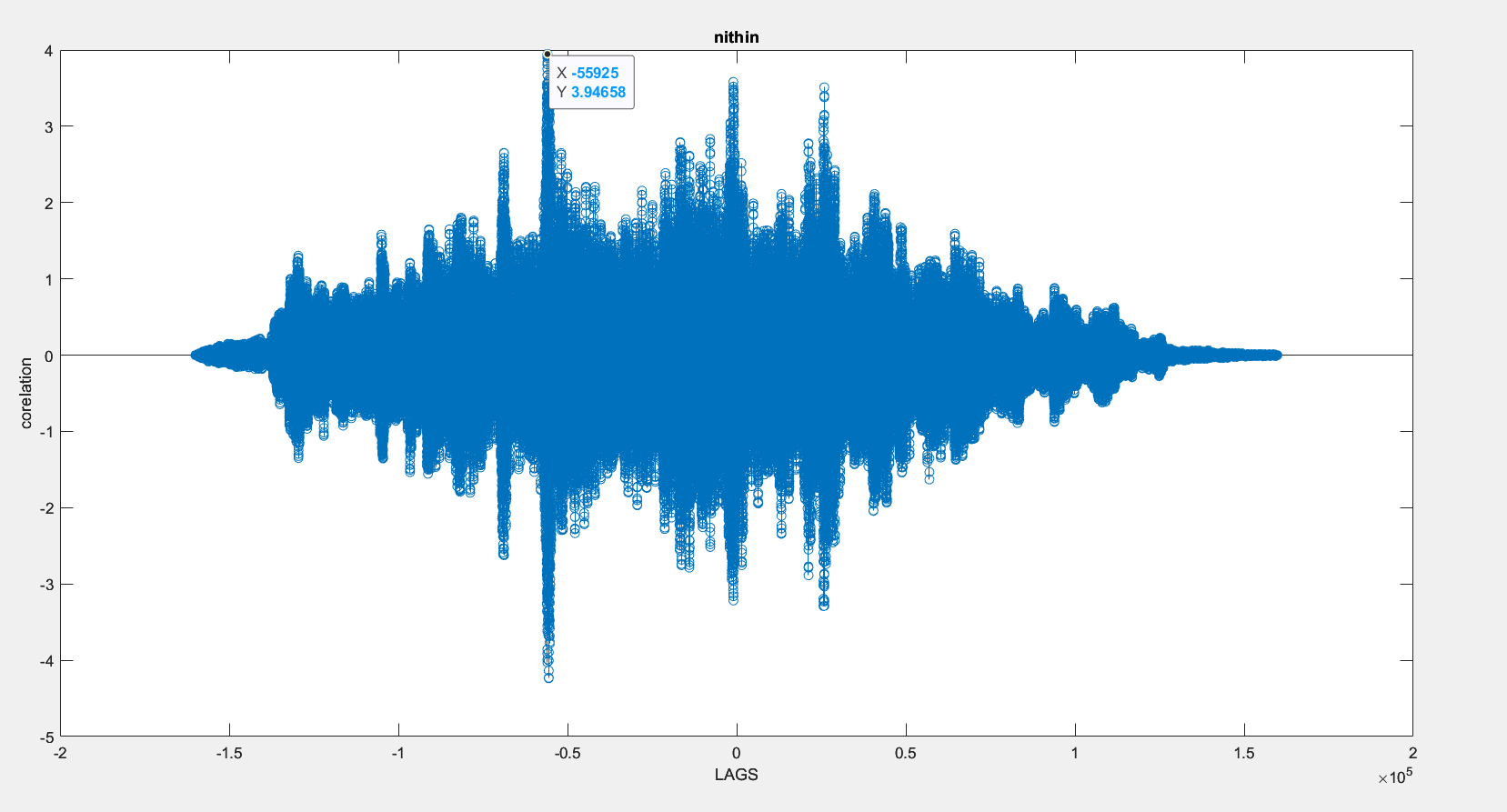


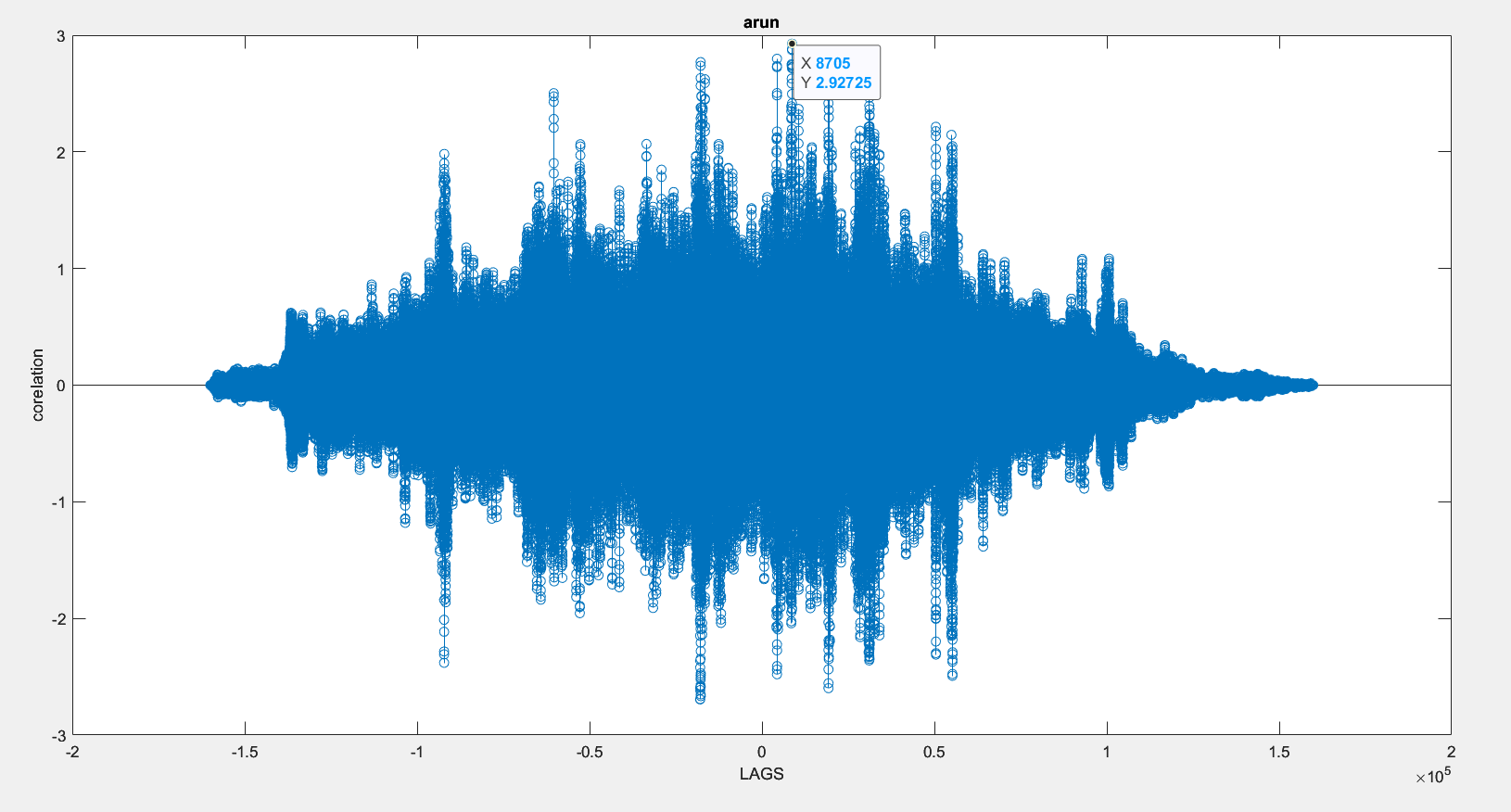
Fig: Test audio signal time domain plot

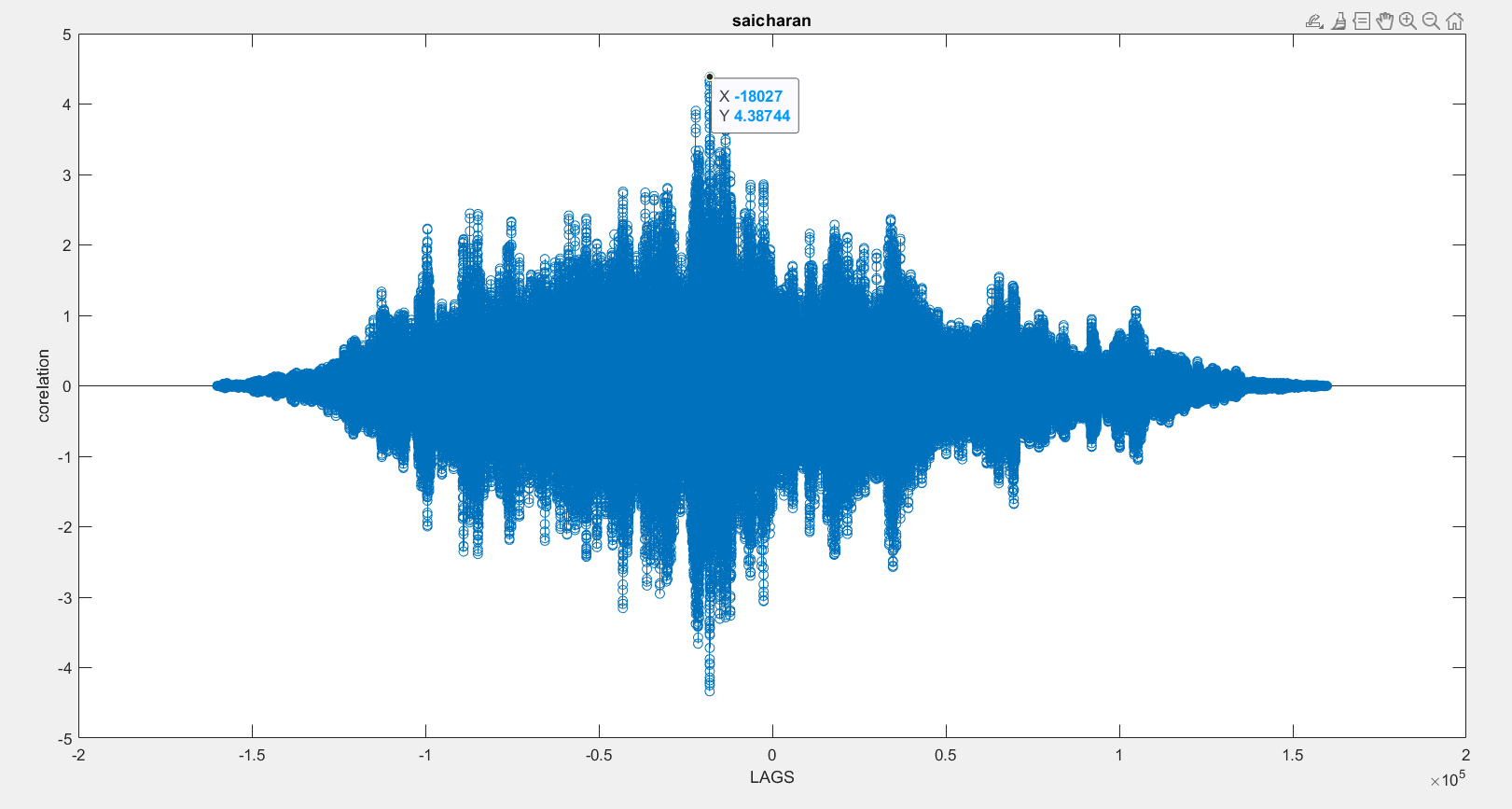
**Corelation plots with all the source files:**

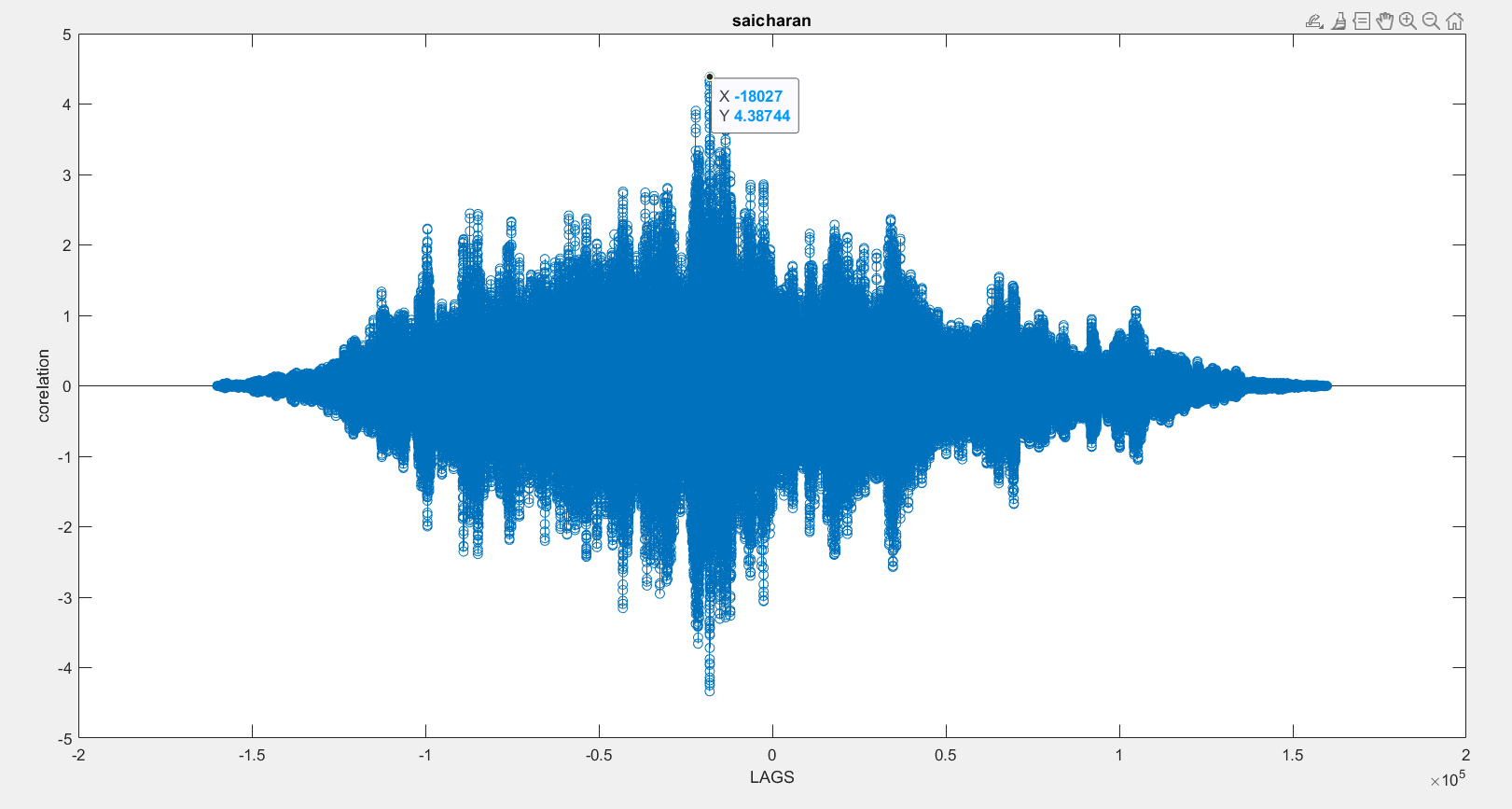


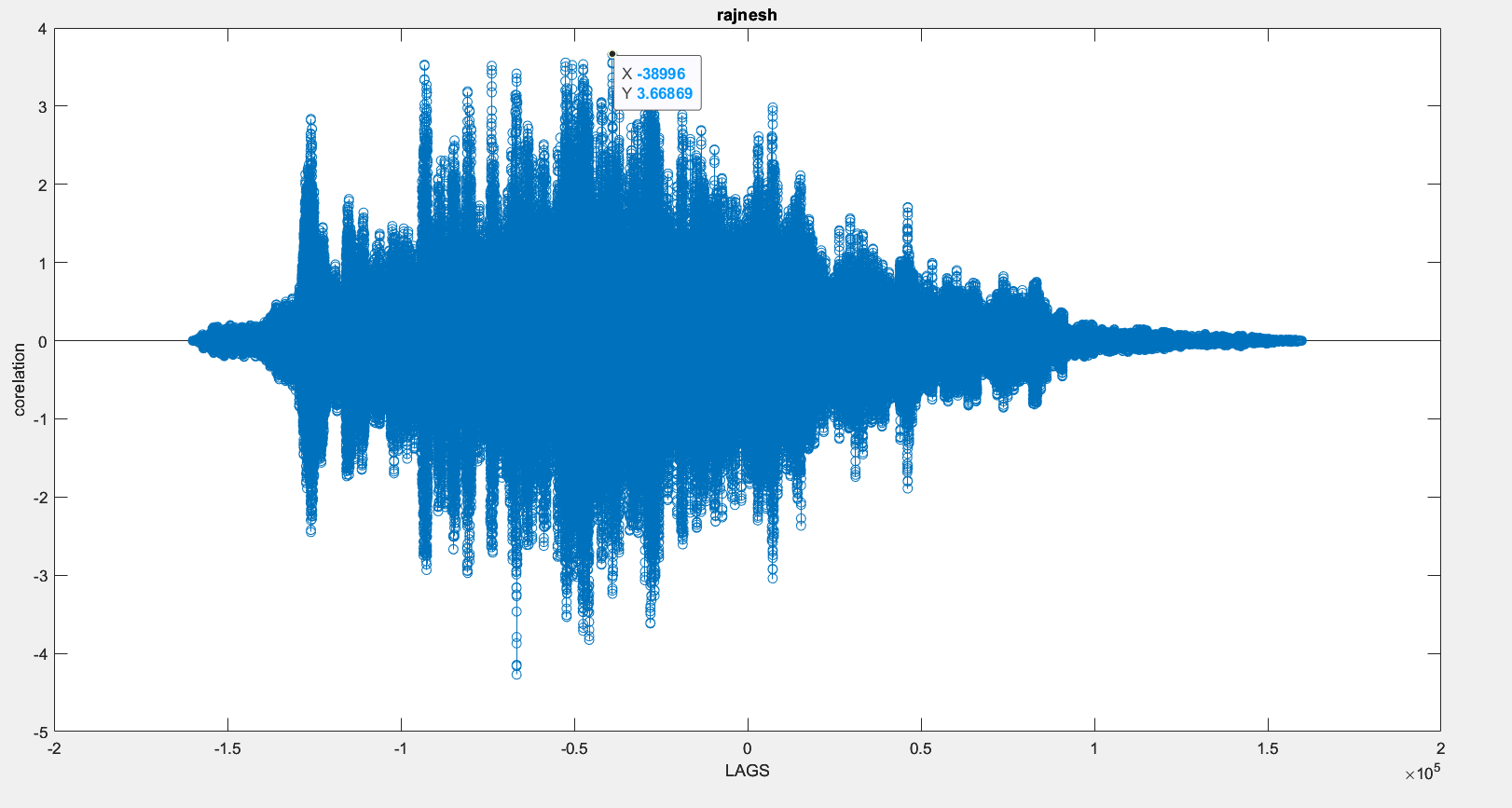


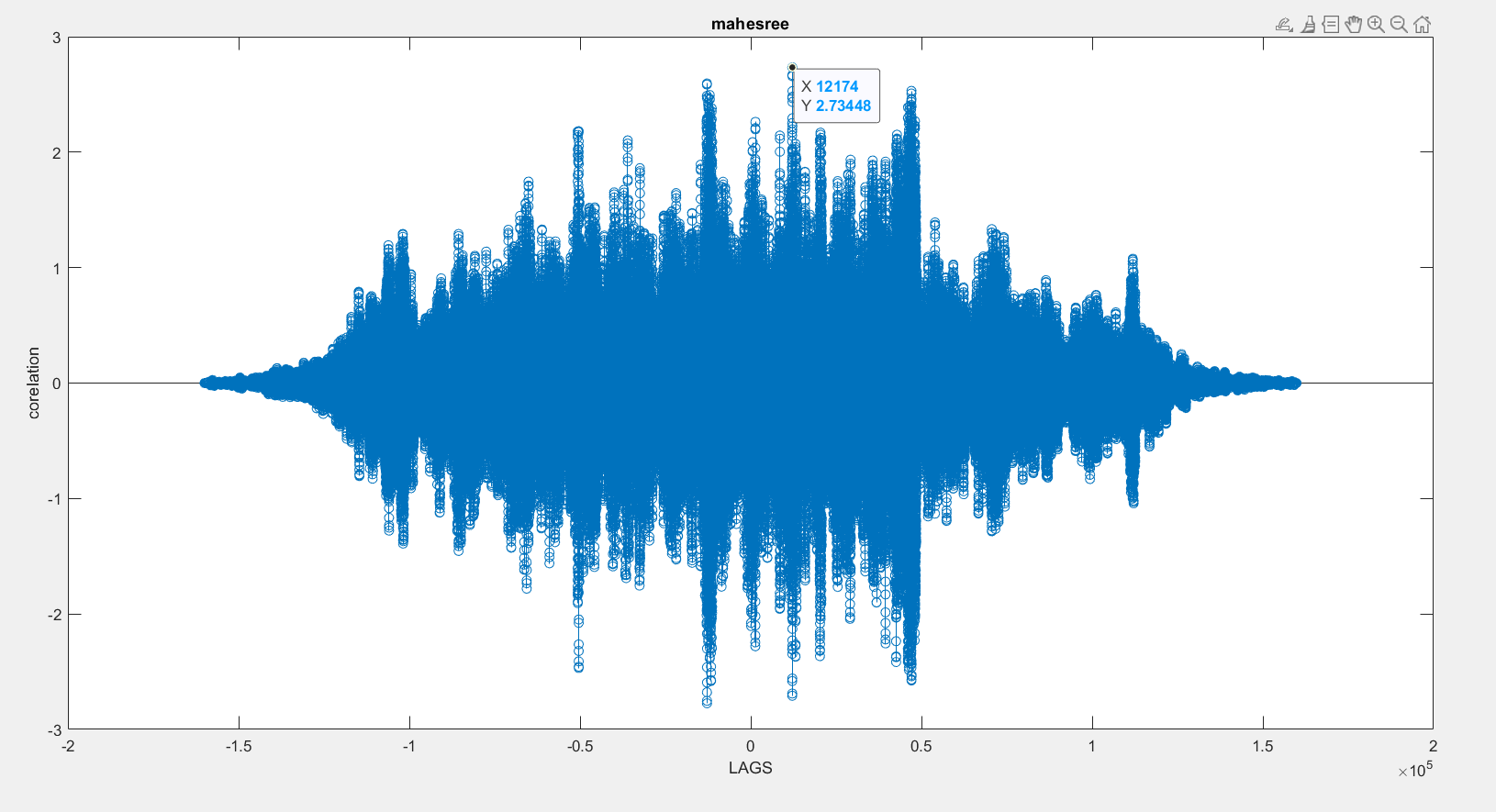


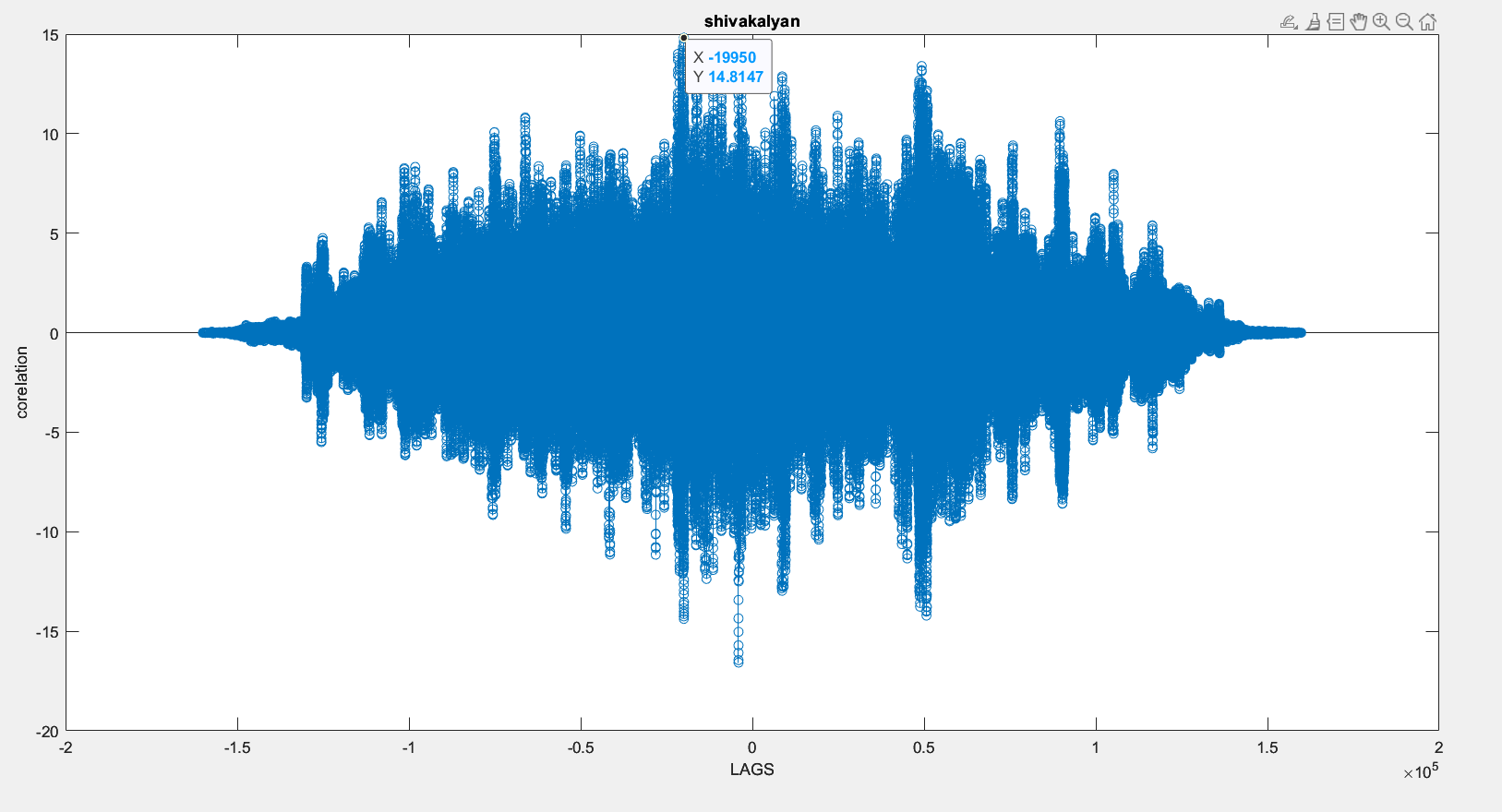




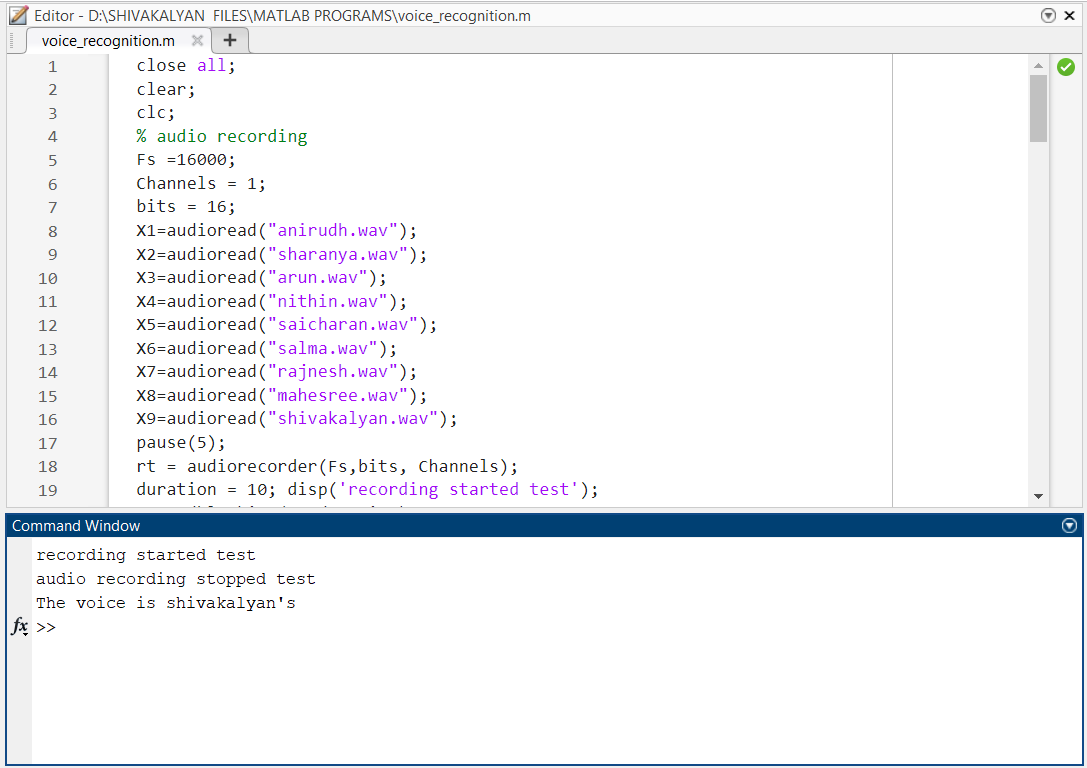




****

****

**Output:**

****

Limitations:

The major limitation of this approach is that if the signal is not there among the source files, then the output wont be “No voice found” Instead it shows the person who’s voice closely matches the given test input.

What can be done?

In order to avoid this we can set certain thresholds for every person just by observing the corelation values that are obtained when the person’s voice is corelated with his own voice and if the output is not up to the threshold then it must display that the “Test input voice cannot be matched wit any file in the source library!”.

Conclusion:

Using MATLAB one can perform correlation of two signals and find the similarity among the two and one of the applications of this is voice recognition which is based on the similarity levels of the test voice and the source voice.

Thank You