

IIITH

TEMPORAL SPEECH ENHANCEMENT FOR LANGUAGE IDENTIFICATION IN NOISY ENVIRONMENTS

SSP

P10

Arshini Govindu 2020102009

Sankeerthana Venugopal 2020102008

Shruti Kolachana 2020102053

TABLE OF CONTENTS

01

INTRODUCTION

02

**GROSS WEIGHT
FUNCTION**

03

**FINE WEIGHT
FUNCTION**

04

**FINAL WEIGHT
FUNCTION**

05

RESULT

06

**LANGUAGE
IDENTIFICATION
RESULTS**

ABOUT INTRODUCTION

The basis for this approach is that human beings perceive speech by capturing features present from the high signal-to-noise ratio (SNR) regions and then extrapolating the features in the low SNR regions

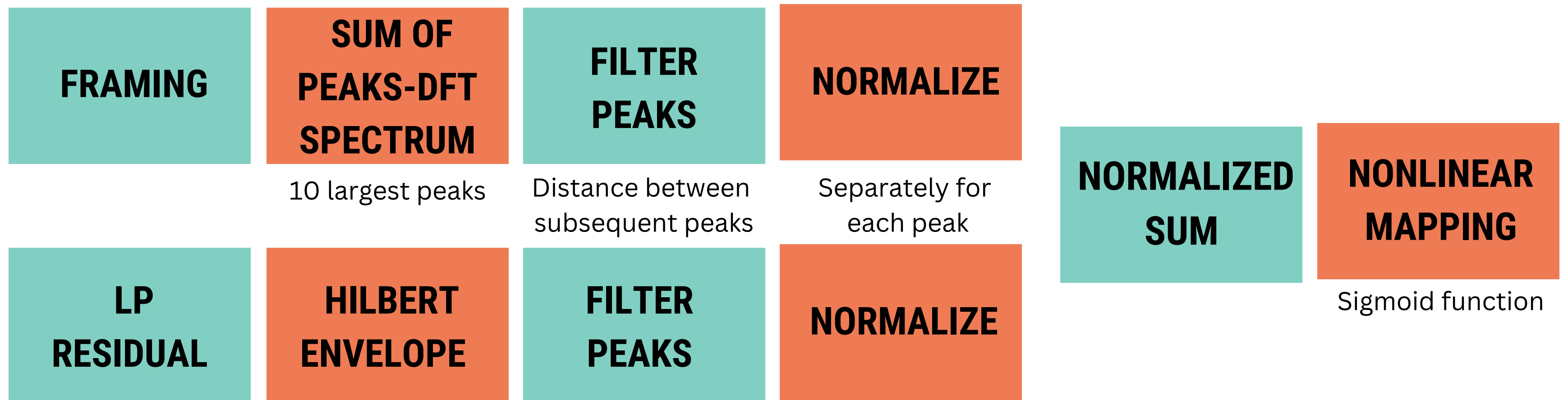
Accordingly, the approach for speech enhancement is to identify the high SNR regions in the noisy speech and enhance them relative to the low SNR regions without causing significant distortion in the enhanced speech.

**GROSS
LEVEL
PROCESSING**

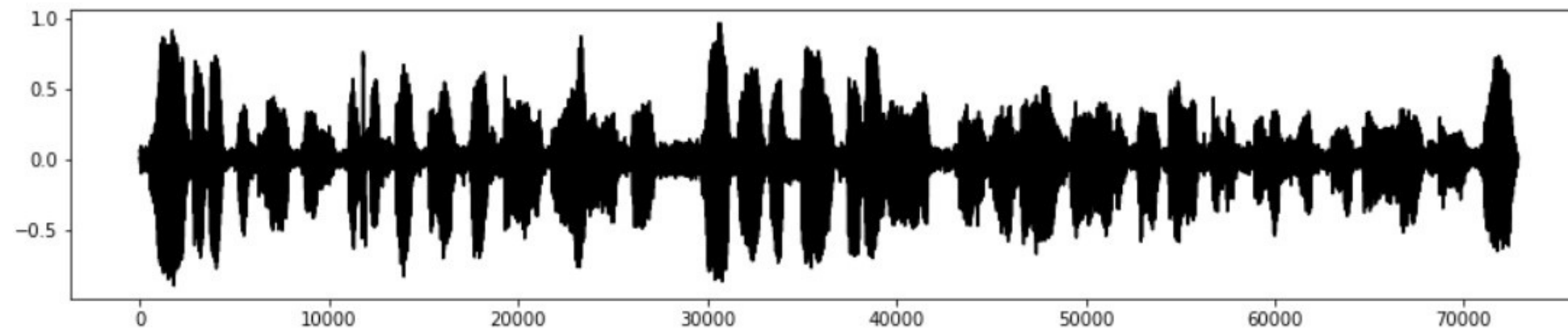
**FINE
LEVEL
PROCESSING**



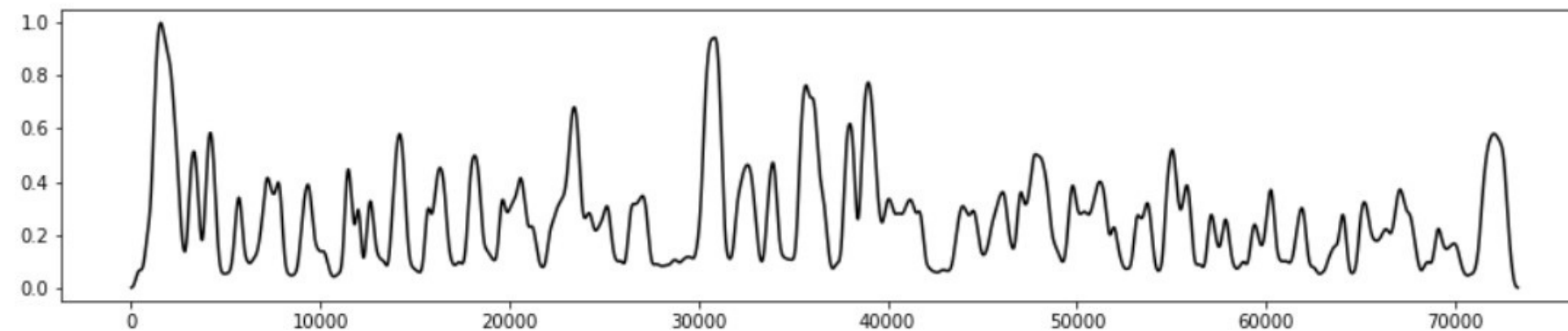
GROSS WEIGHT FUNCTION



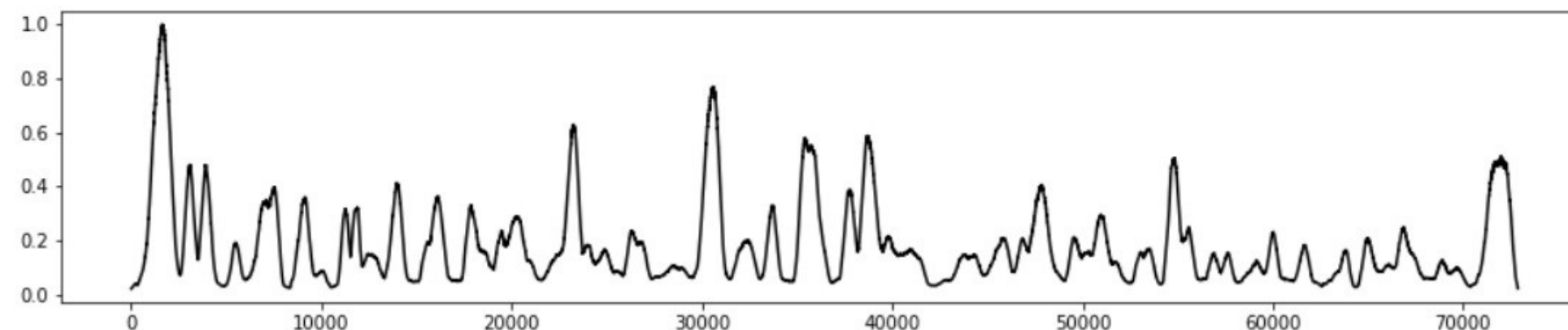
GROSS LEVEL FEATURES



SIGNAL

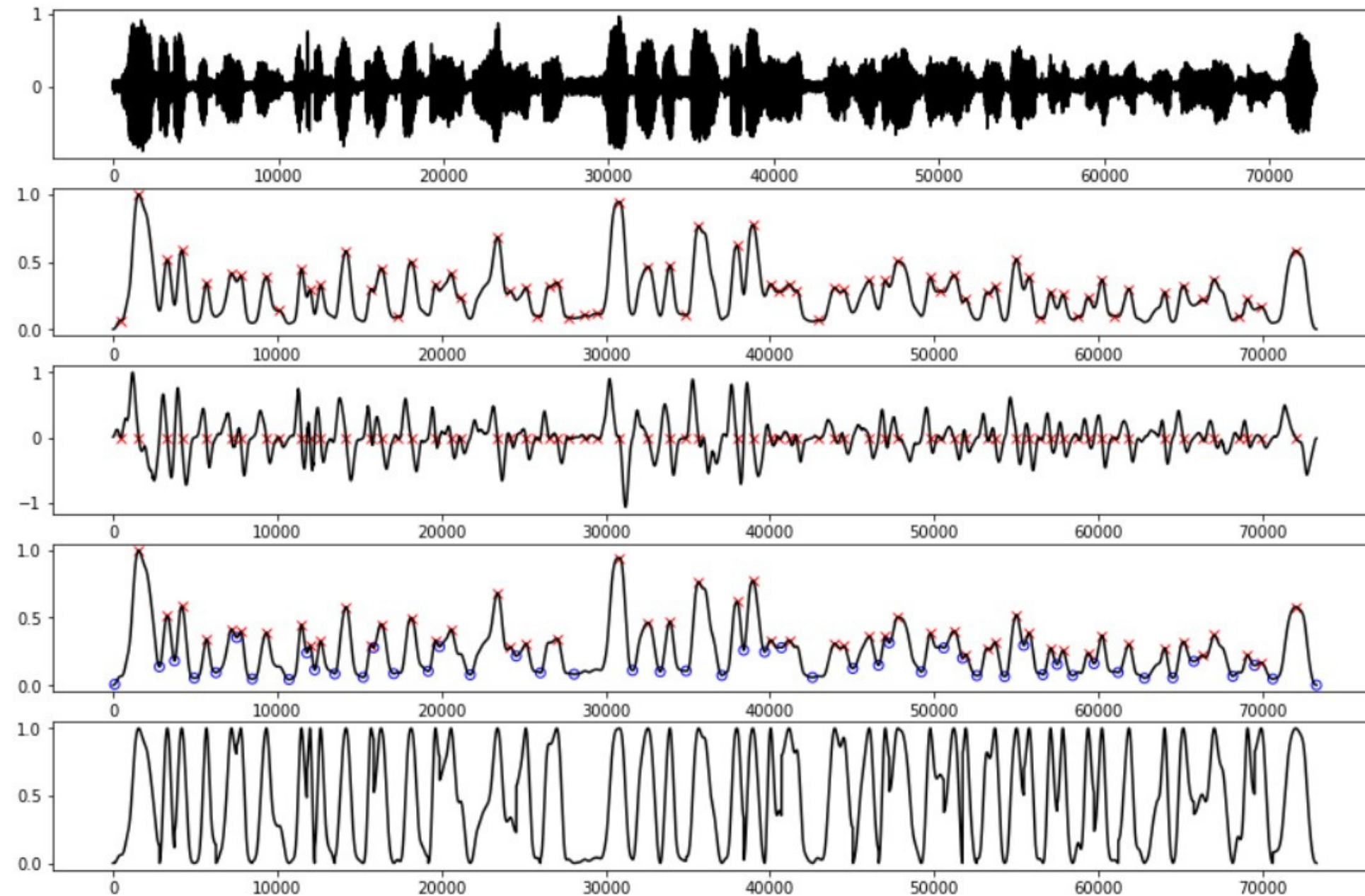


SUM OF DFT PEAKS



**HILBERT ENVELOPE OF
LP RESIDUAL**

IDENTIFYING HIGH SNR REGIONS



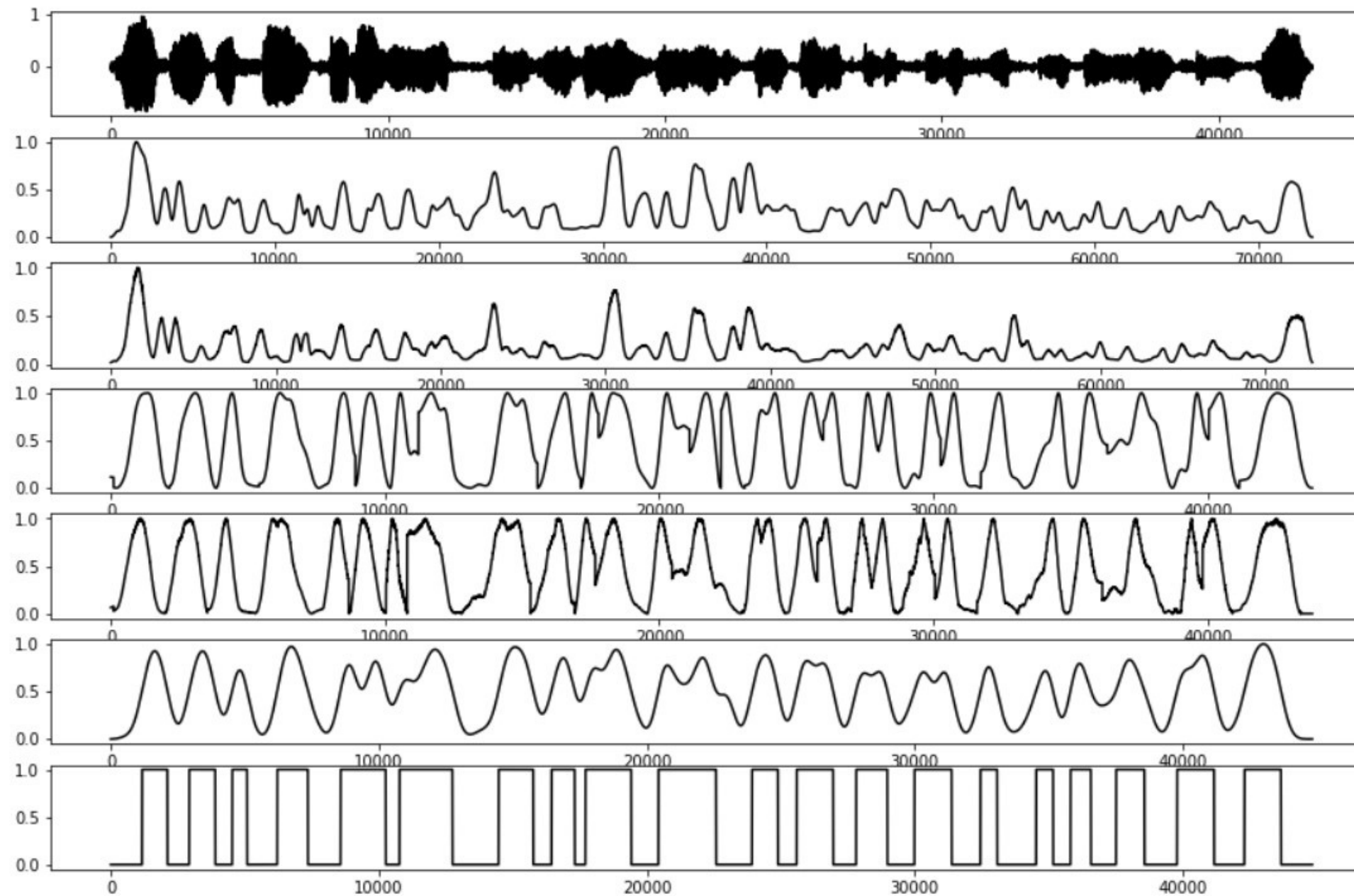
PEAKS

FIRST ORDER DIFFERENCE

FILTERED PEAKS

NORMALIZED SIGNAL

GROSS WEIGHT FUNCTION



SUM OF PEAKS OF DFT

HILBERT ENVELOPE OF LP RESIDUAL

NORMALIZED DFT

NORMALIZED HILBERT

NORMALIZED SUM

NONLINEARLY MAPPED VALUES

FINE WEIGHT FUNCTION

**LP
RESIDUAL**

**HILBERT
ENVELOPE**

**CONVOLVE
WITH NEG
FOGD**

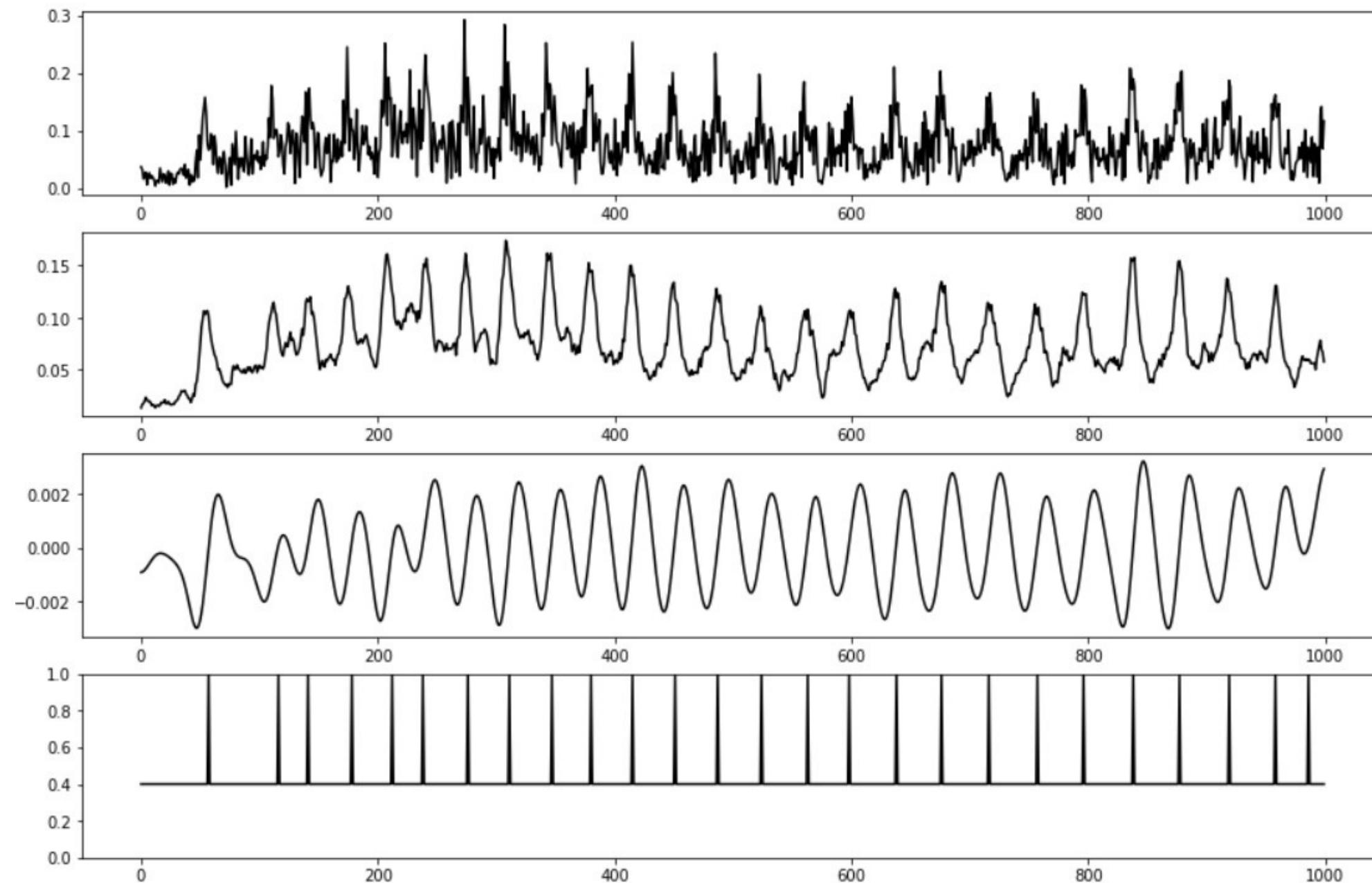
First order Gaussian
differentiator

**ZERO
CROSSINGS**

Keep ZCs with negative
to positive transition

**FILTER ZERO
CROSSINGS**

FINE WEIGHT FUNCTION



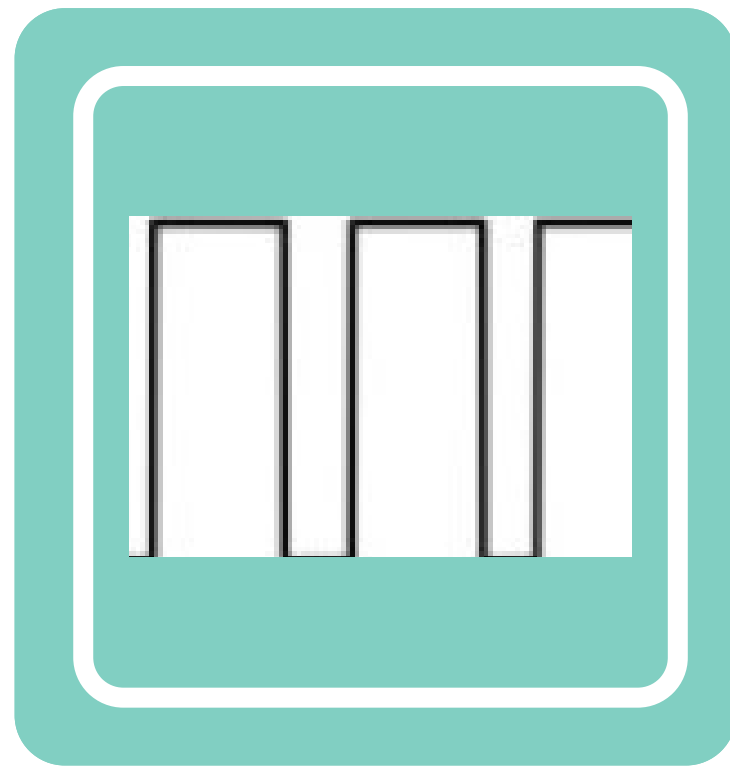
HILBERT ENVELOPE OF LP RESIDUAL

SMOOTHED HILBERT

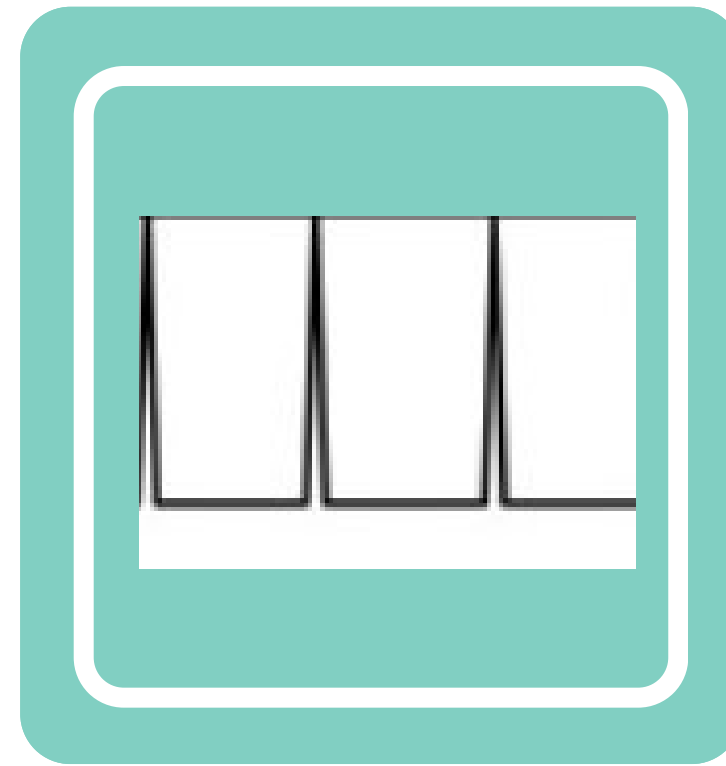
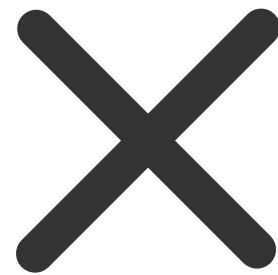
CONVOLVED WITH NEG FOGD

ZCS WITH NEG-POS TRANSITION

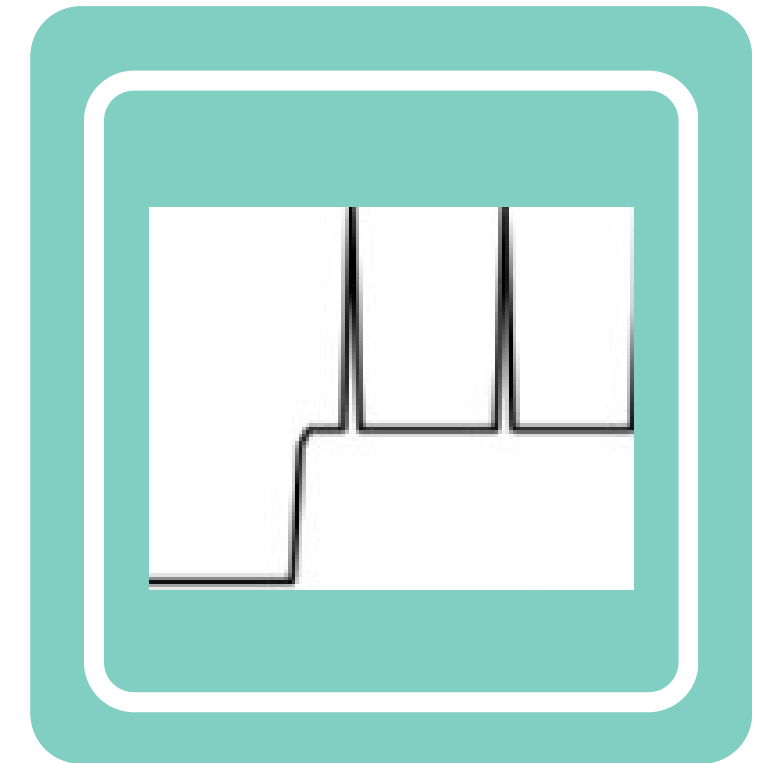
FINAL WEIGHT FUNCTION



**GROSS WEIGHT
FUNCTION**



**FINE WEIGHT
FUNCTION**

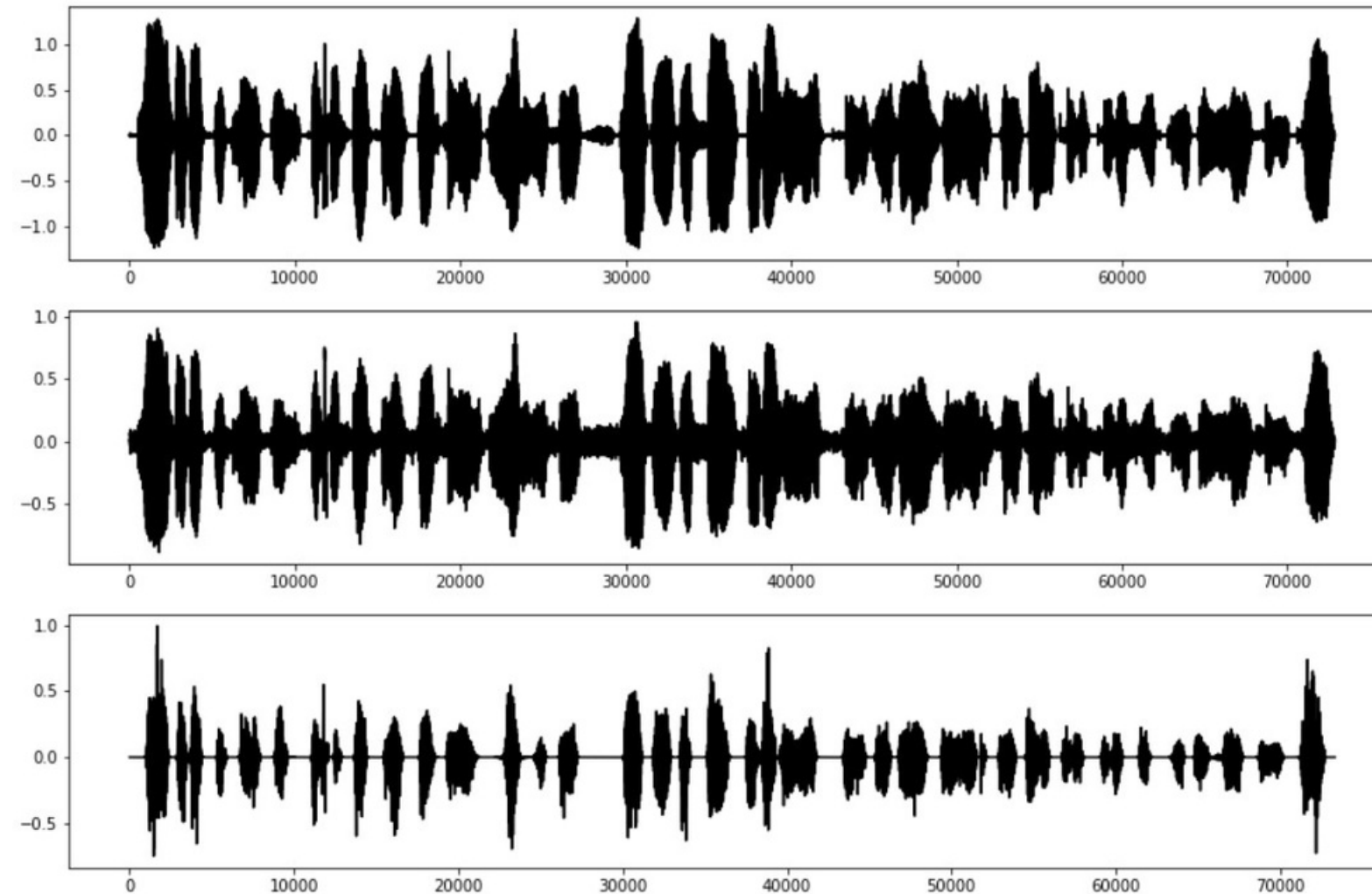


**FINAL WEIGHT
FUNCTION**

RESULT

The final weight function is multiplied with the LP residual of the noisy signal.

The LP coefficients of the initial signal are used to perform LP synthesis, to get the final output.



CLEAN SIGNAL

NOISY SIGNAL

OUTPUT SIGNAL

LID RESULTS

Accuracy: 85%

Accuracy for Noisy Speech: 41.67%

Accuracy for Temporally Processed Speech: 29.16%

Increase in language score:

- Babble noise: in 66.7% cases
- Factory noise: in 66.7% cases
- Pink noise: in 33.3% cases
- White noise: in 50% cases

*Thank
you!*