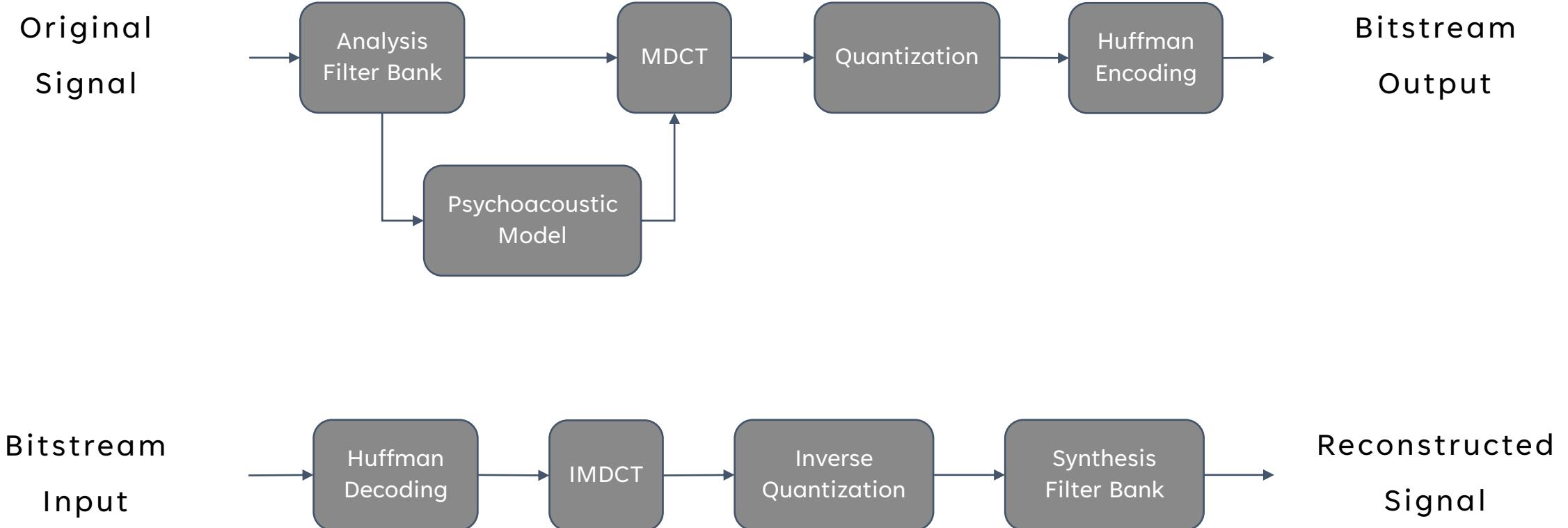


# Effects of Quantization Error in MP3

Tyler Hattori

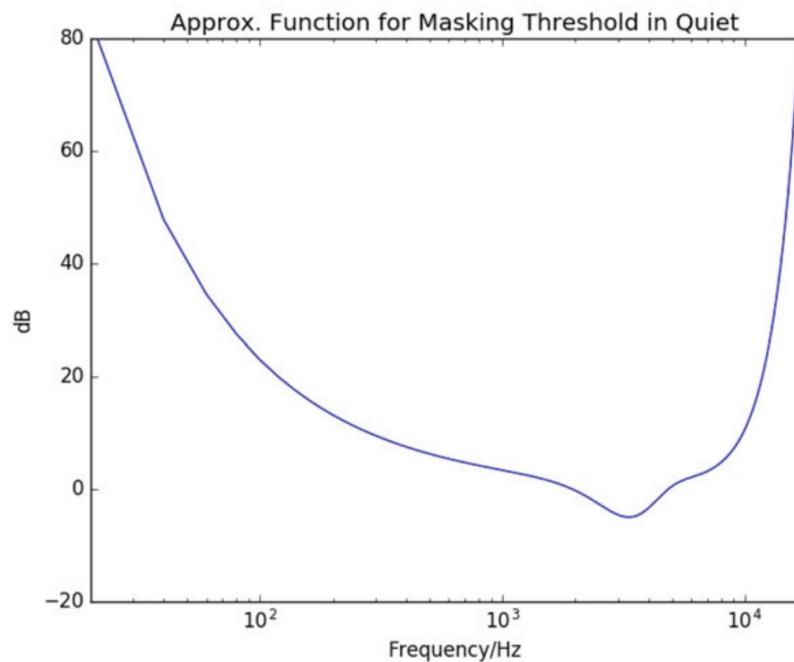
ECE 258 Fall 2023

# MP3 Audio Encoding and Decoding



# The Psychoacoustic Model

Catering to the Human Ear



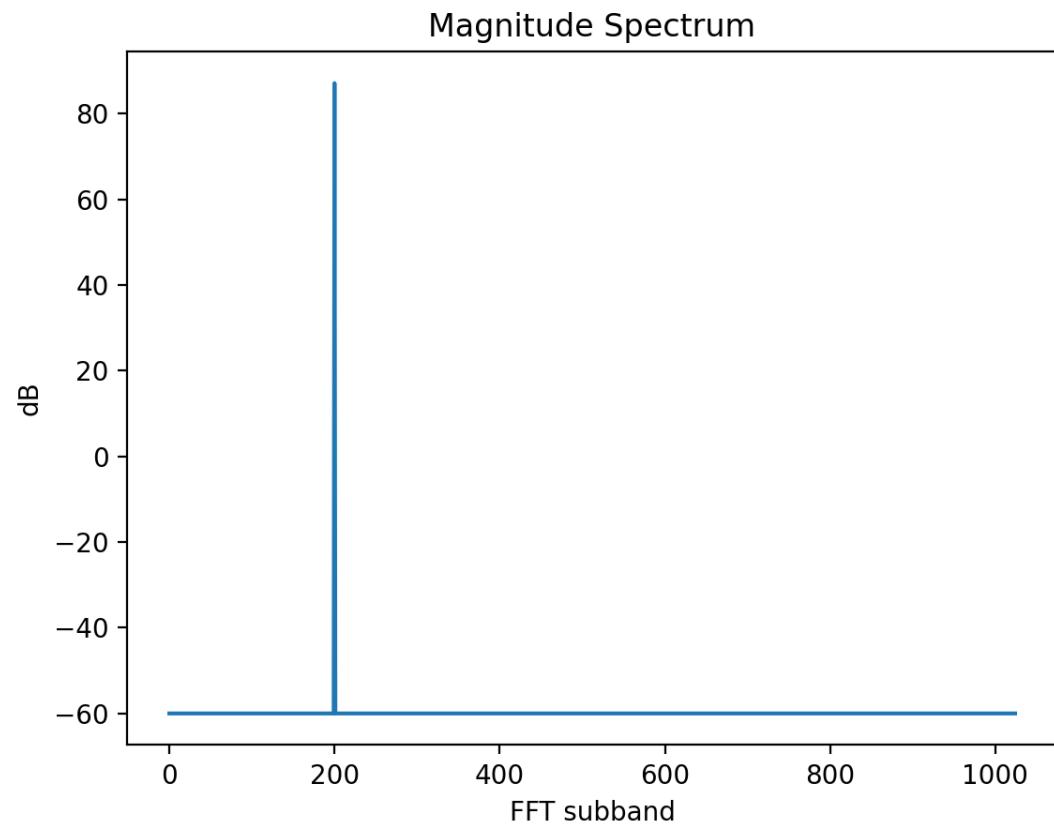
All of the plots and spectrograms in this presentation were obtained personally by modifying the code in Gerald Schuller's book *Hybrid Filter Banks and Audio Encoding*.

Schuller, Gerald. (2020). *Filter Banks and Audio Coding: Compressing Audio Signals Using Python*. Springer International Publishing.

# The Psychoacoustic Model

## Sine Wave Input Example

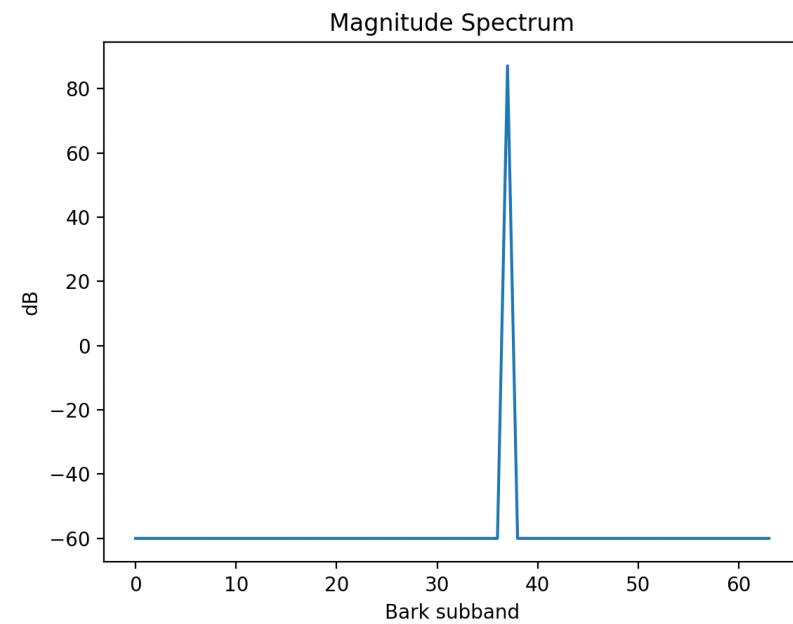
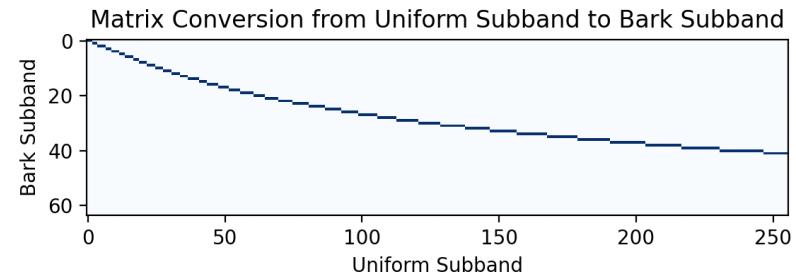
- Take FFT of input
- Convert frequency axis to bark axis
- Add spreading functions to the natural masking threshold of the ear
- Convert bark axis to frequency axis



# The Psychoacoustic Model

## Sine Wave Input Example

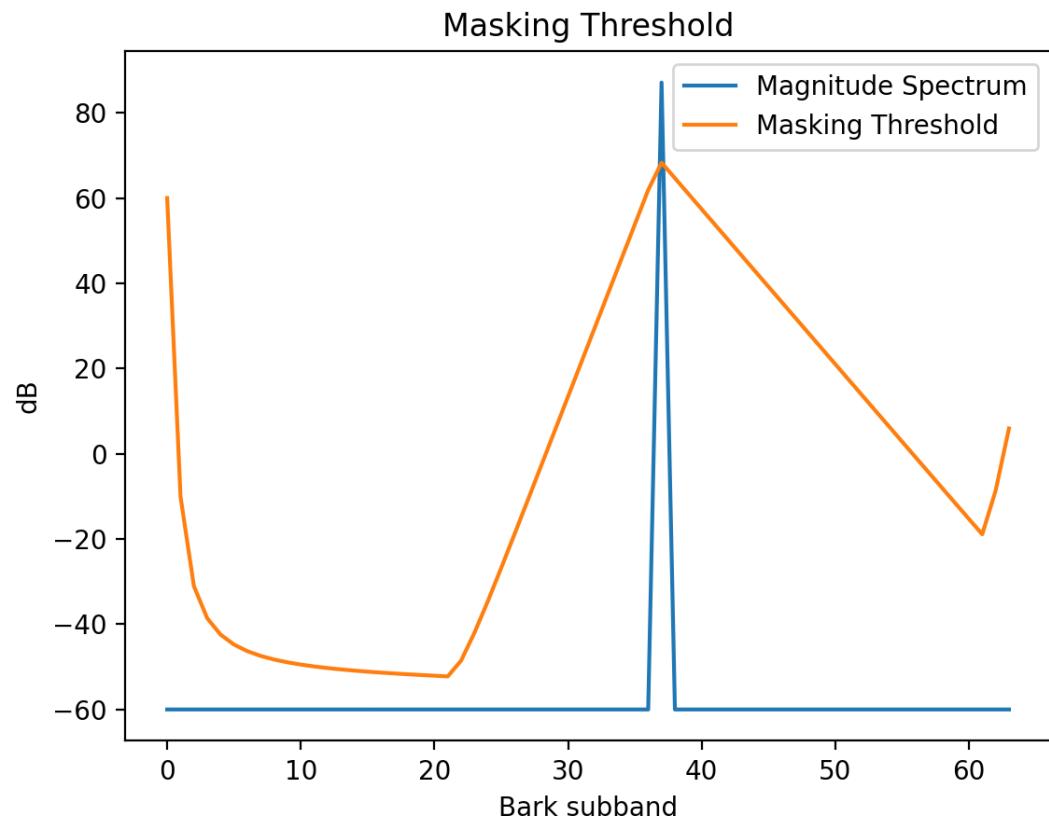
- Take FFT of input
- Convert frequency axis to bark axis
- Add spreading functions to the natural masking threshold of the ear
- Convert bark axis to frequency axis



# The Psychoacoustic Model

## Sine Wave Input Example

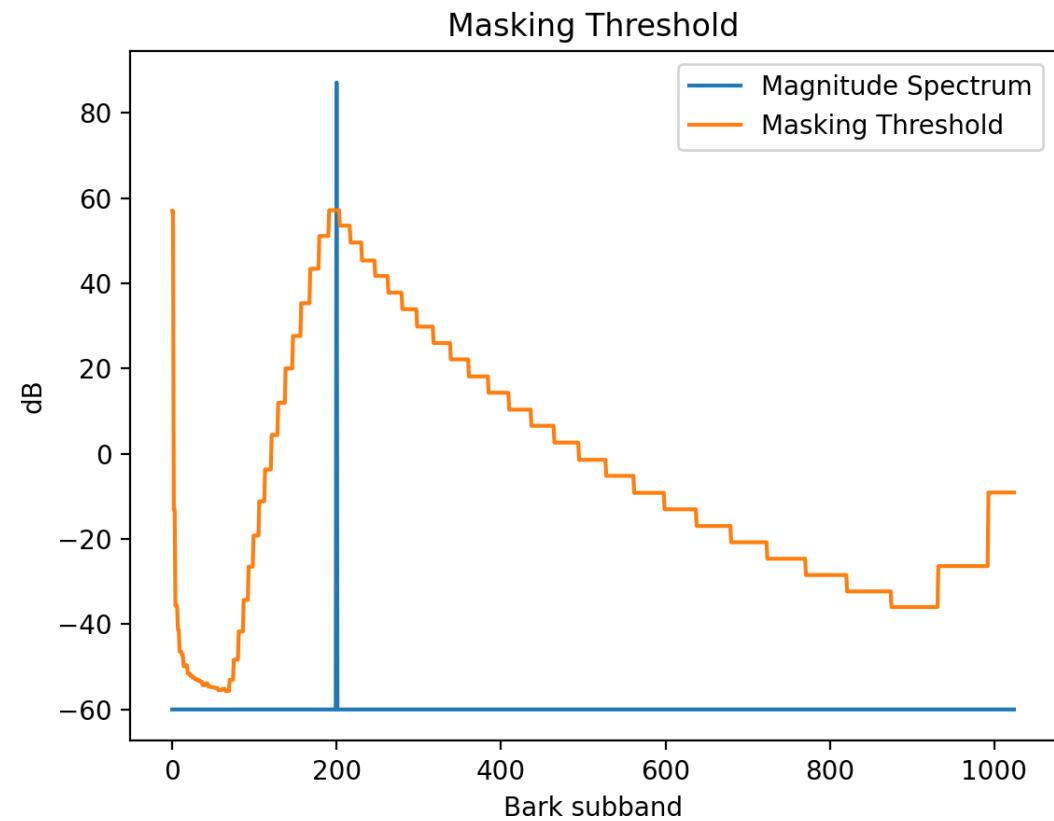
- Take FFT of input
- Convert frequency axis to bark axis
- Add spreading functions to the natural masking threshold of the ear
- Convert bark axis to frequency axis



# The Psychoacoustic Model

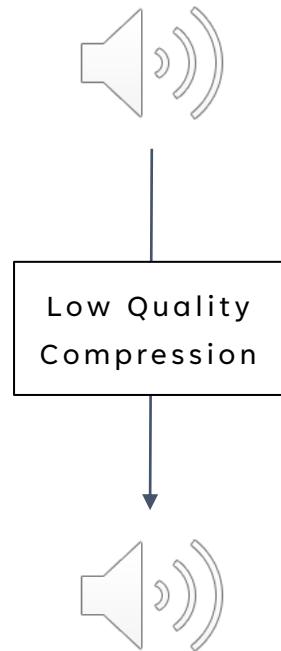
## Sine Wave Input Example

- Take FFT of input
- Convert frequency axis to bark axis
- Add spreading functions to the natural masking threshold of the ear
- Convert bark axis to frequency axis

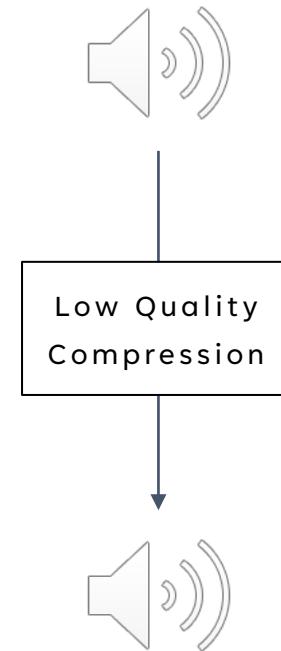


# Effects of Quantization Error

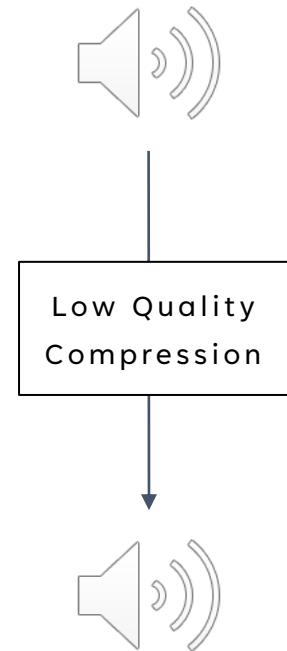
Castanets



Male Voice



Music (Mt Joy)



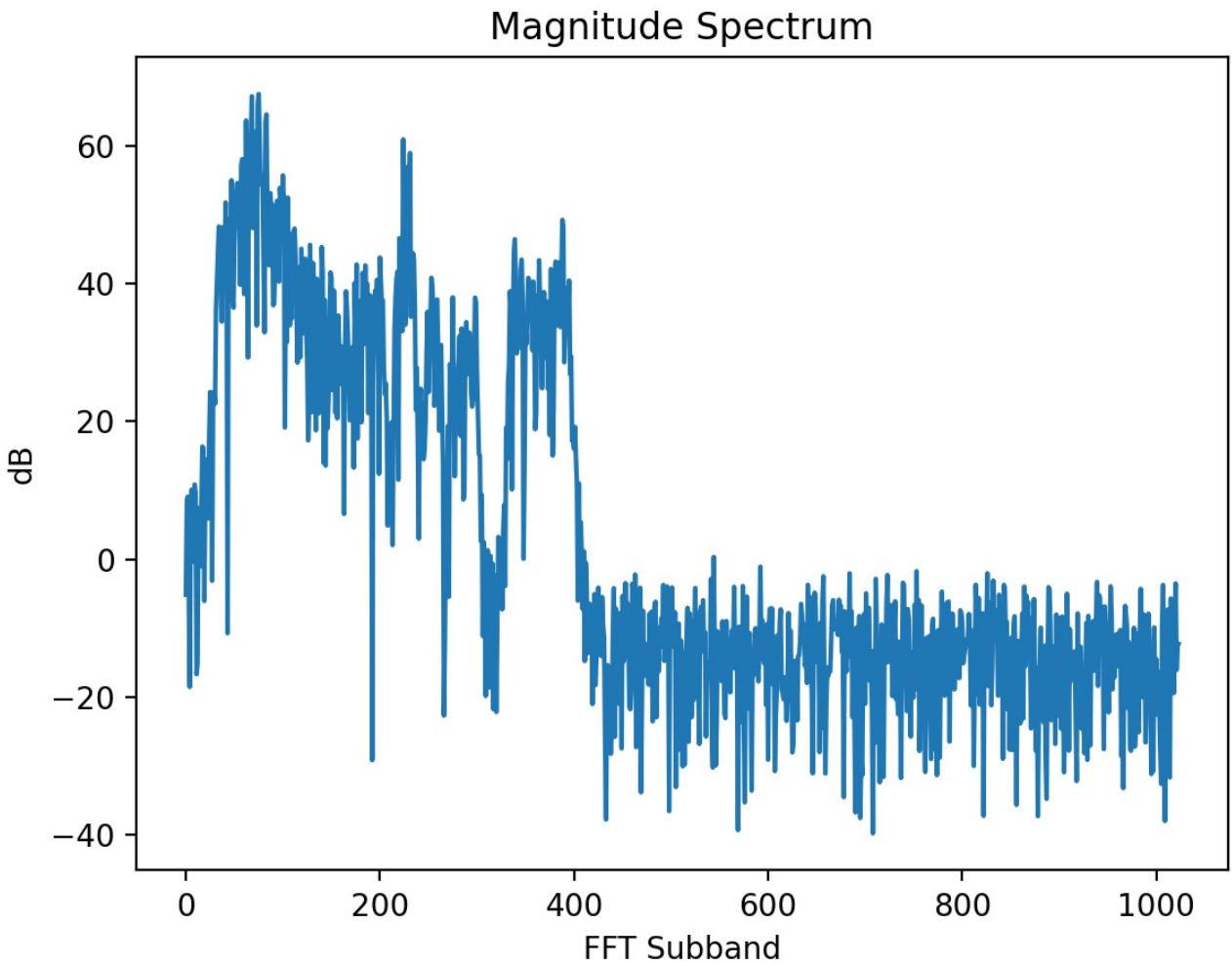
Pre-echo

Reverb and  
Pre-echo

Harmonic  
Distortion

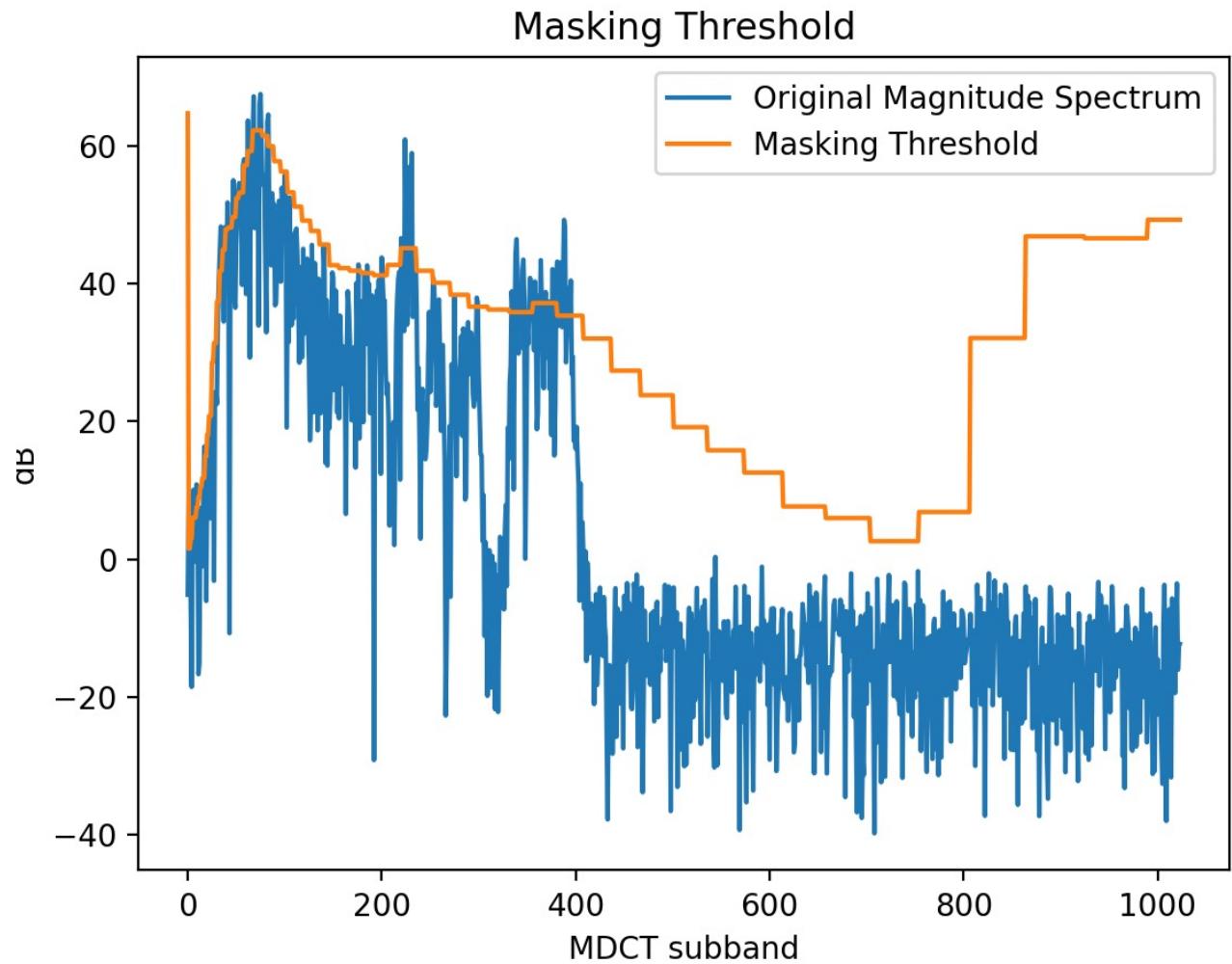
# Castanets Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



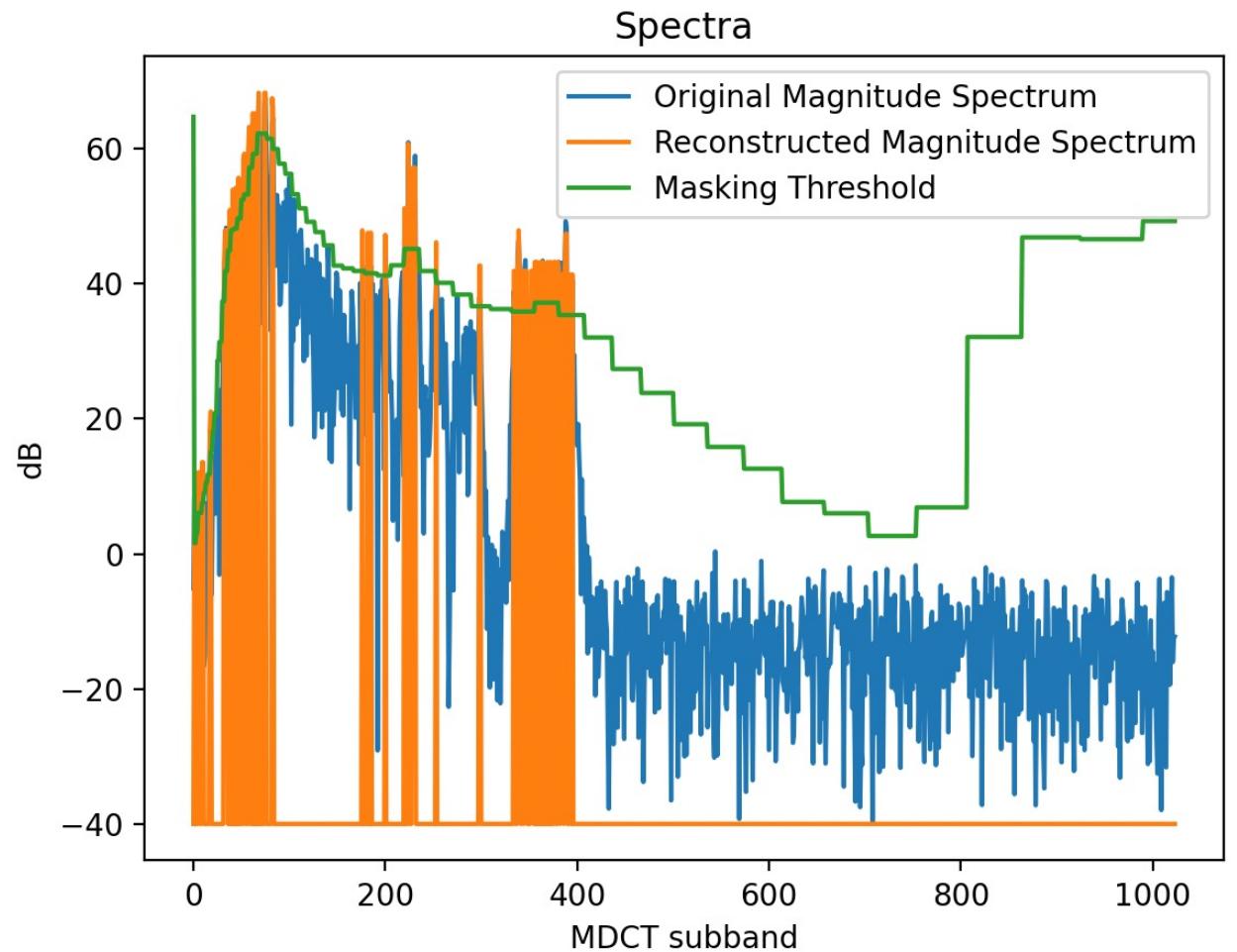
# Castanets Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



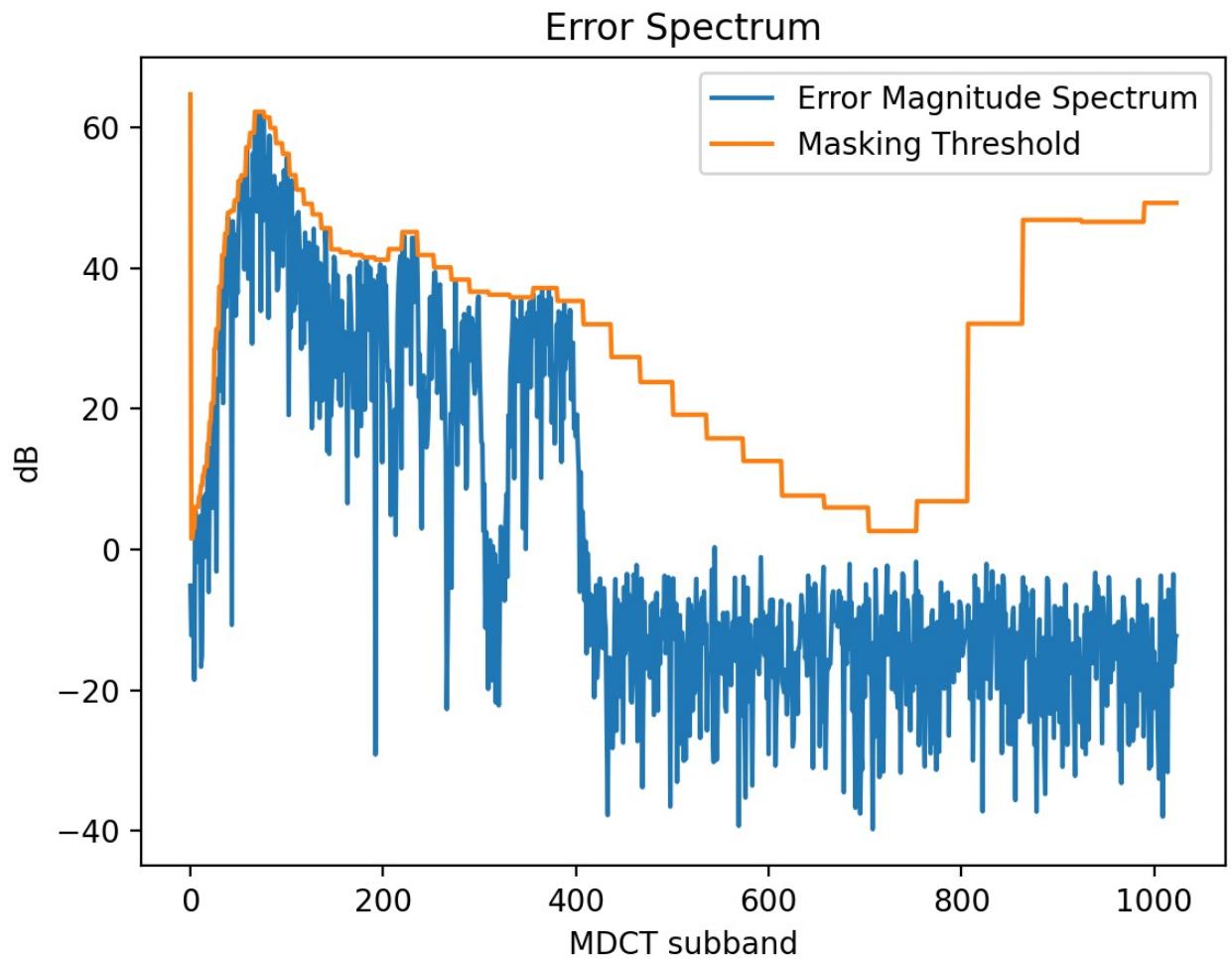
# Castanets Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



# Castanets Quantization Error

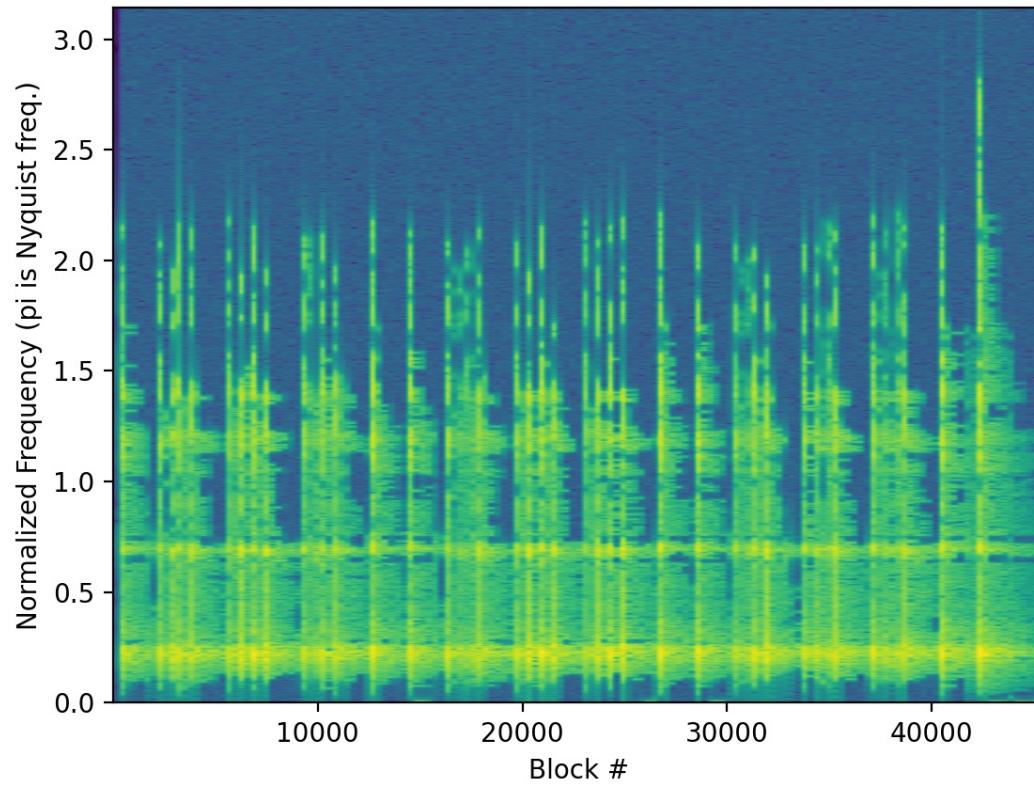
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



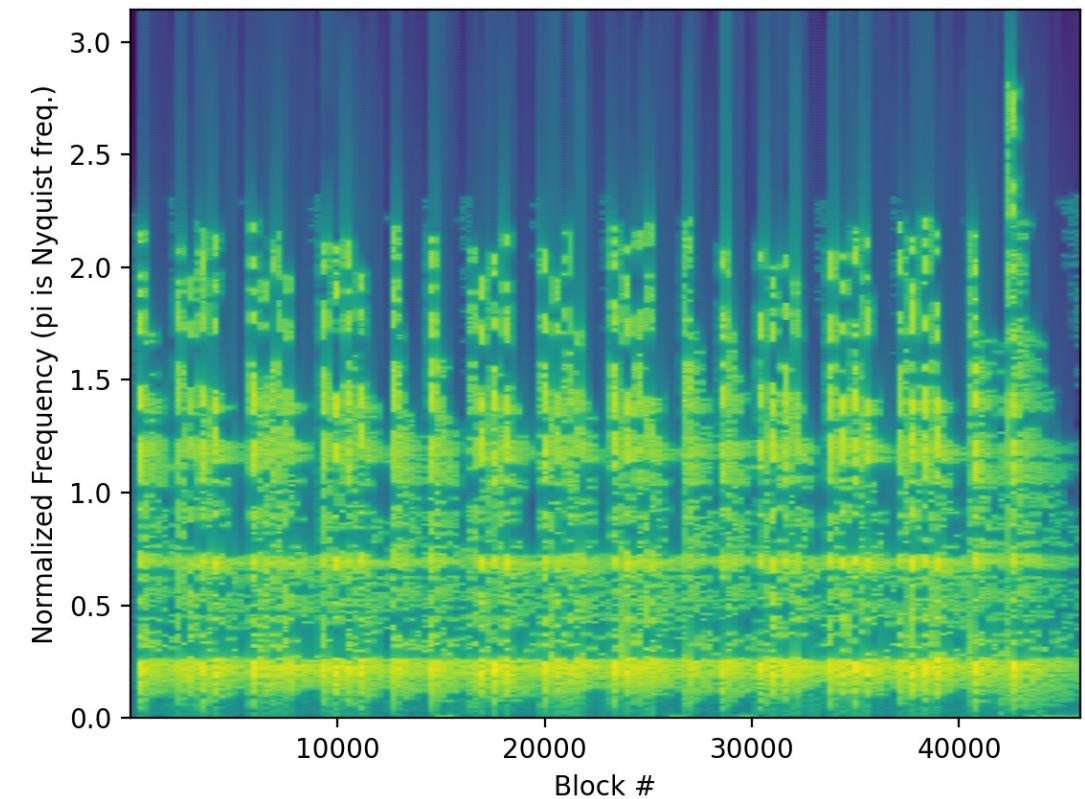
# Castanets Quantization Error

60 Quality, 1024 MDCT Subbands

Original



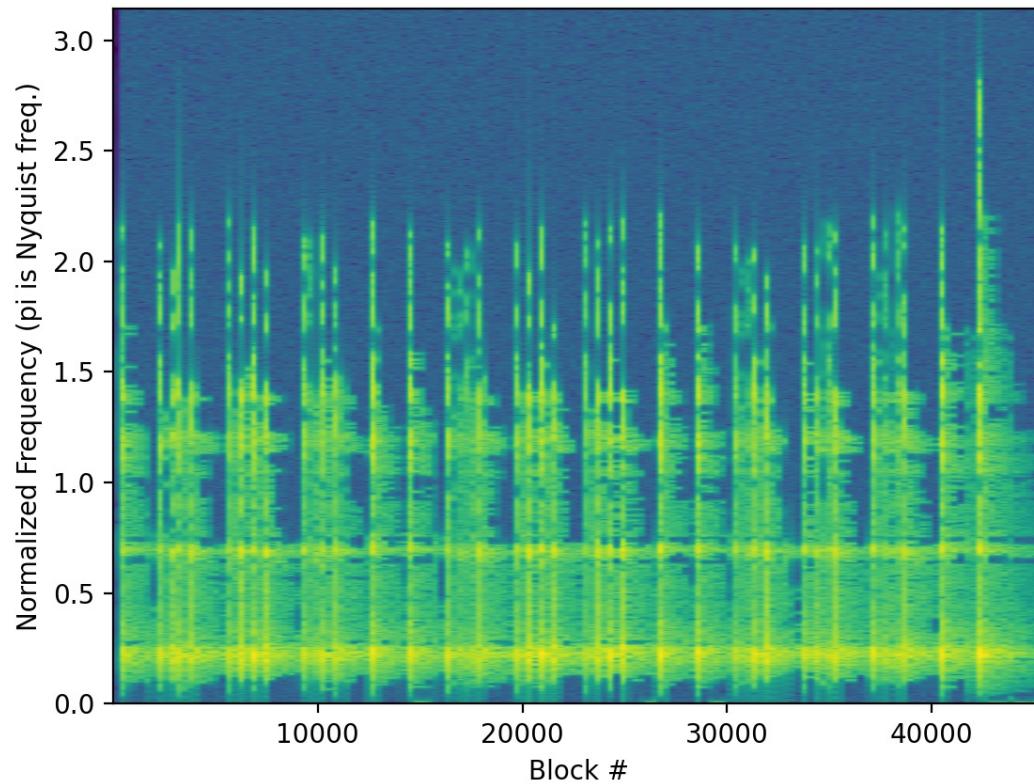
Reconstructed



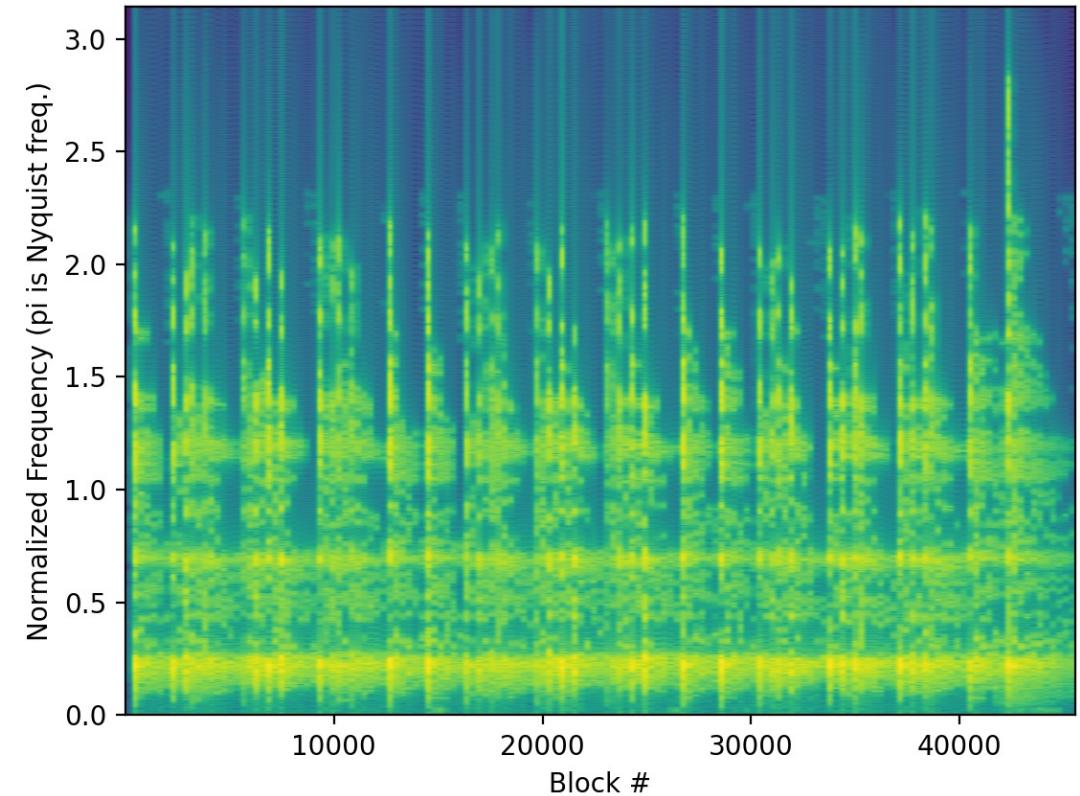
# Castanets Quantization Error

60 Quality, 256 MDCT Subbands

Original



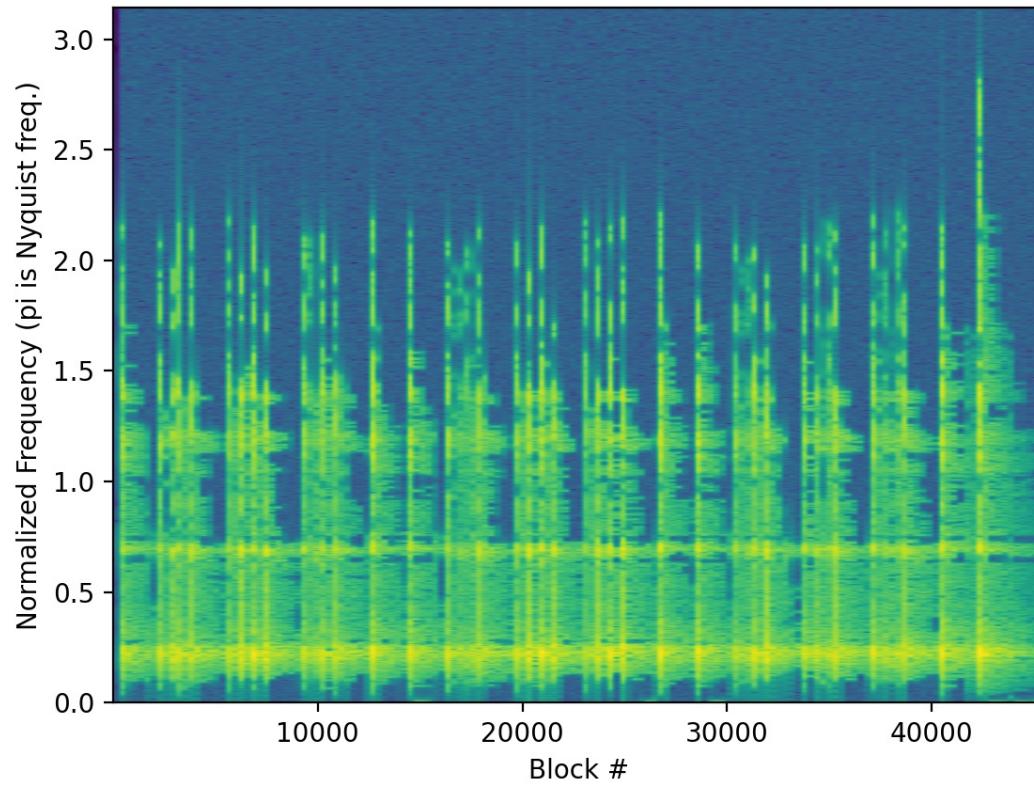
Reconstructed



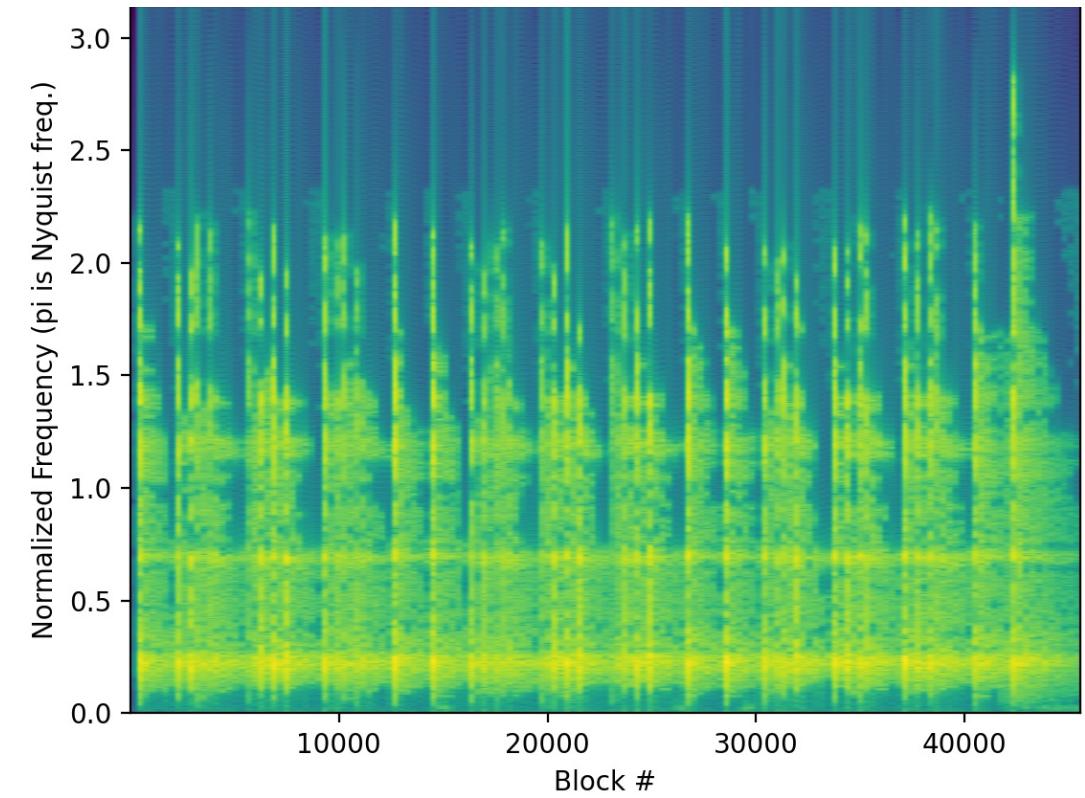
# Castanets Quantization Error

100 Quality, 256 MDCT Subbands

Original



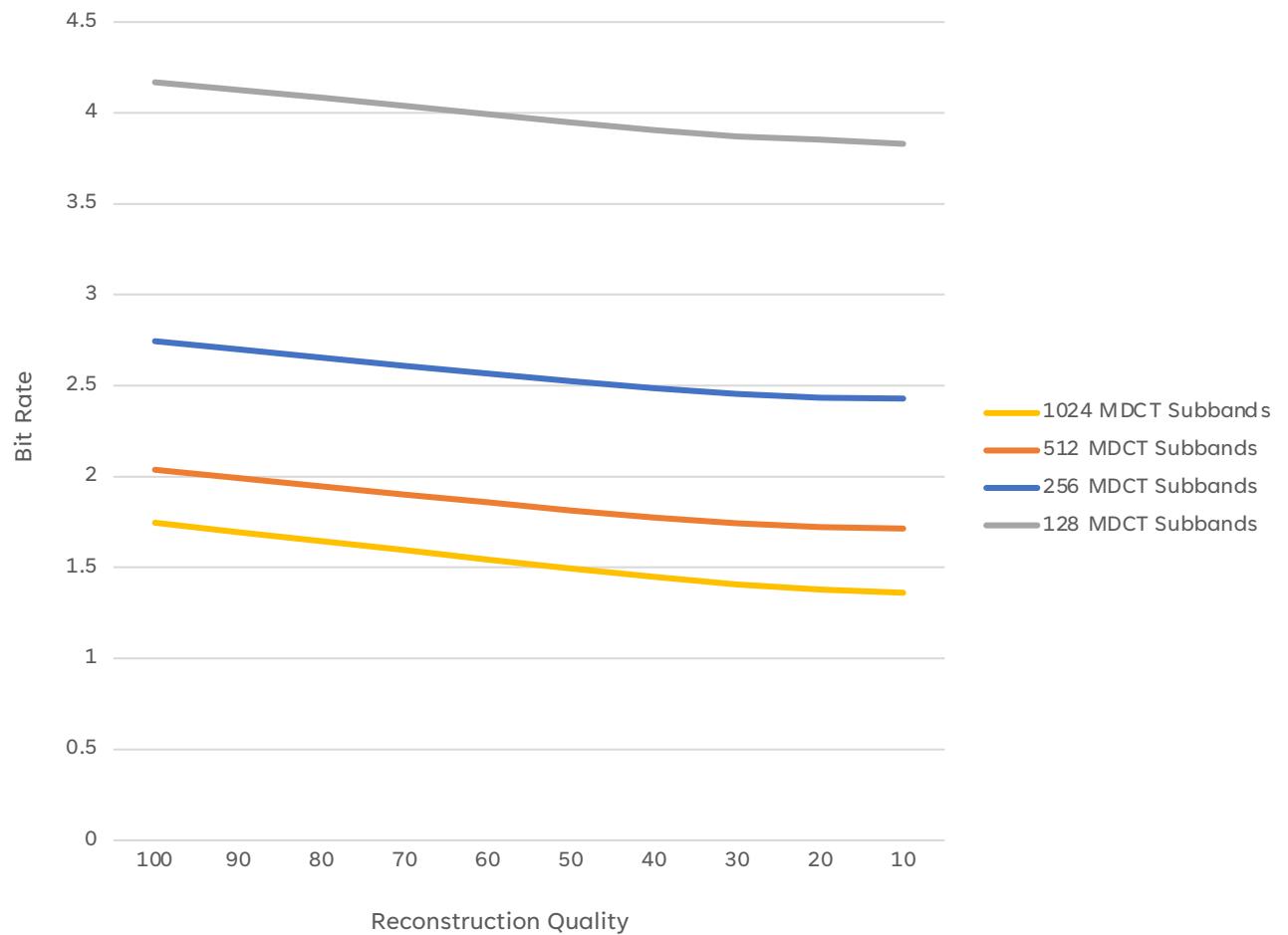
Reconstructed



# Bit Rate vs Reconstruction Quality

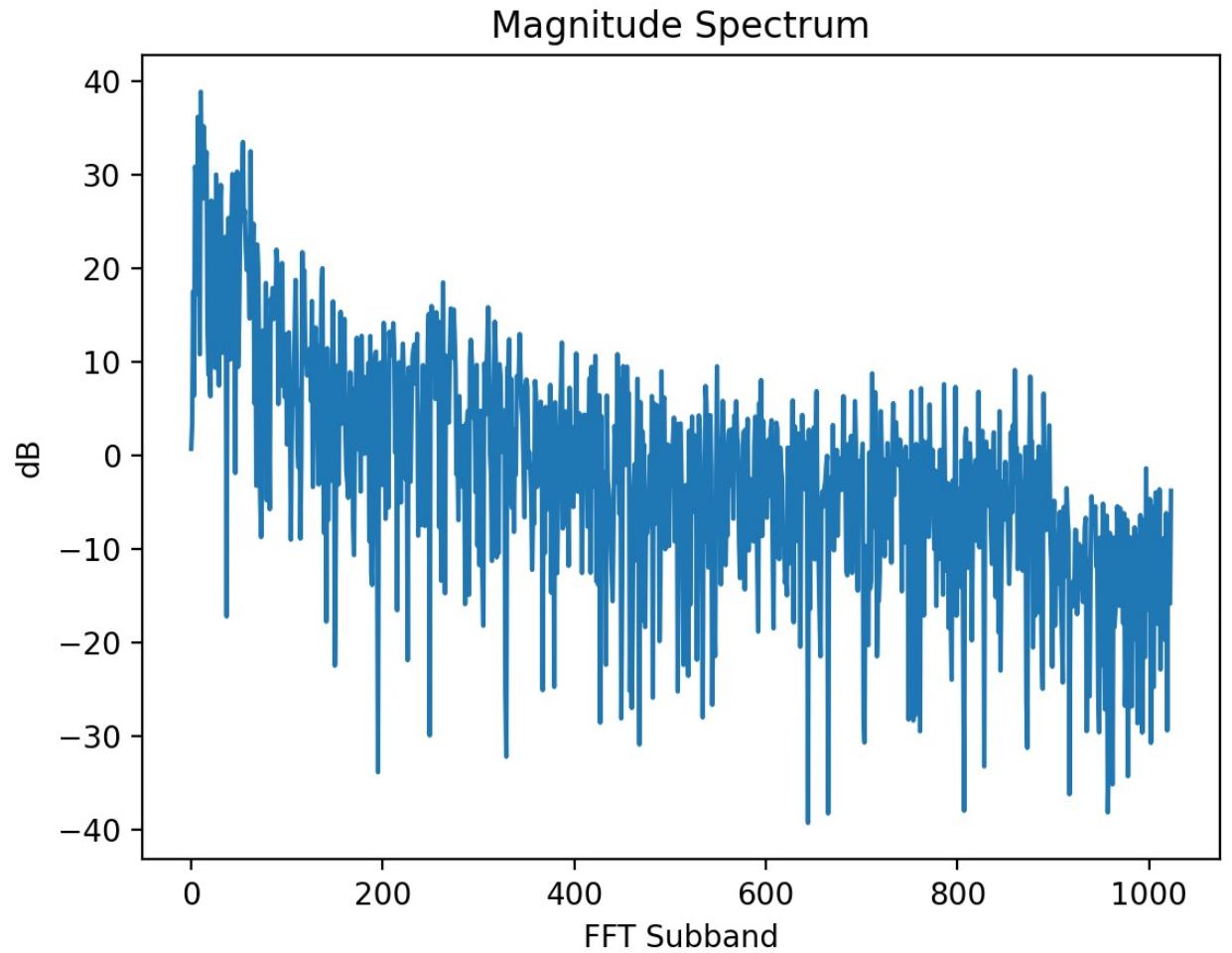
Castanets Input

- Decreasing MDCT Subbands removes pre-echo but increases bit rate
- Decreasing reconstruction quality (increasing the masking threshold) decreases bit rate but increases quantization distortion



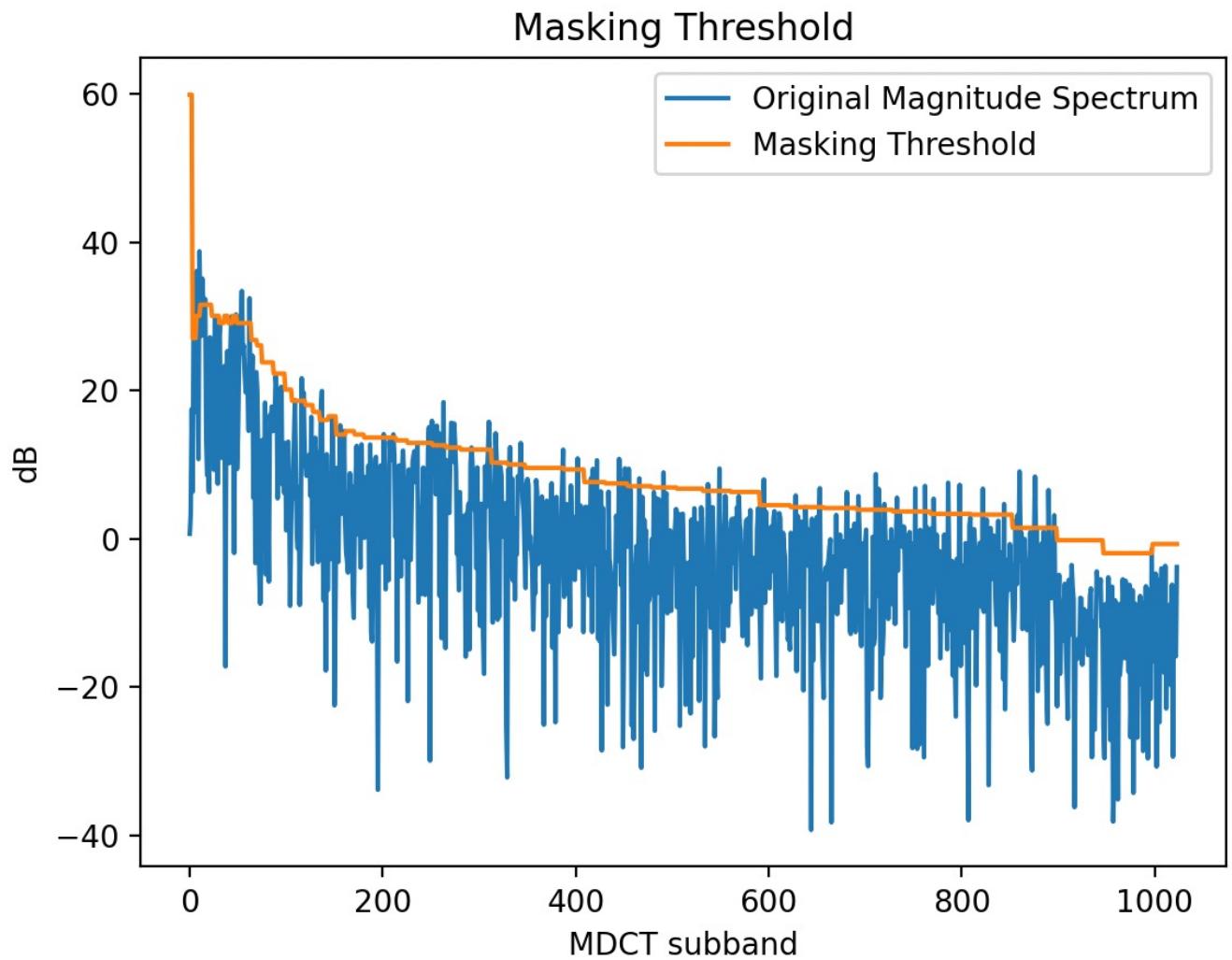
# Male Voice Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



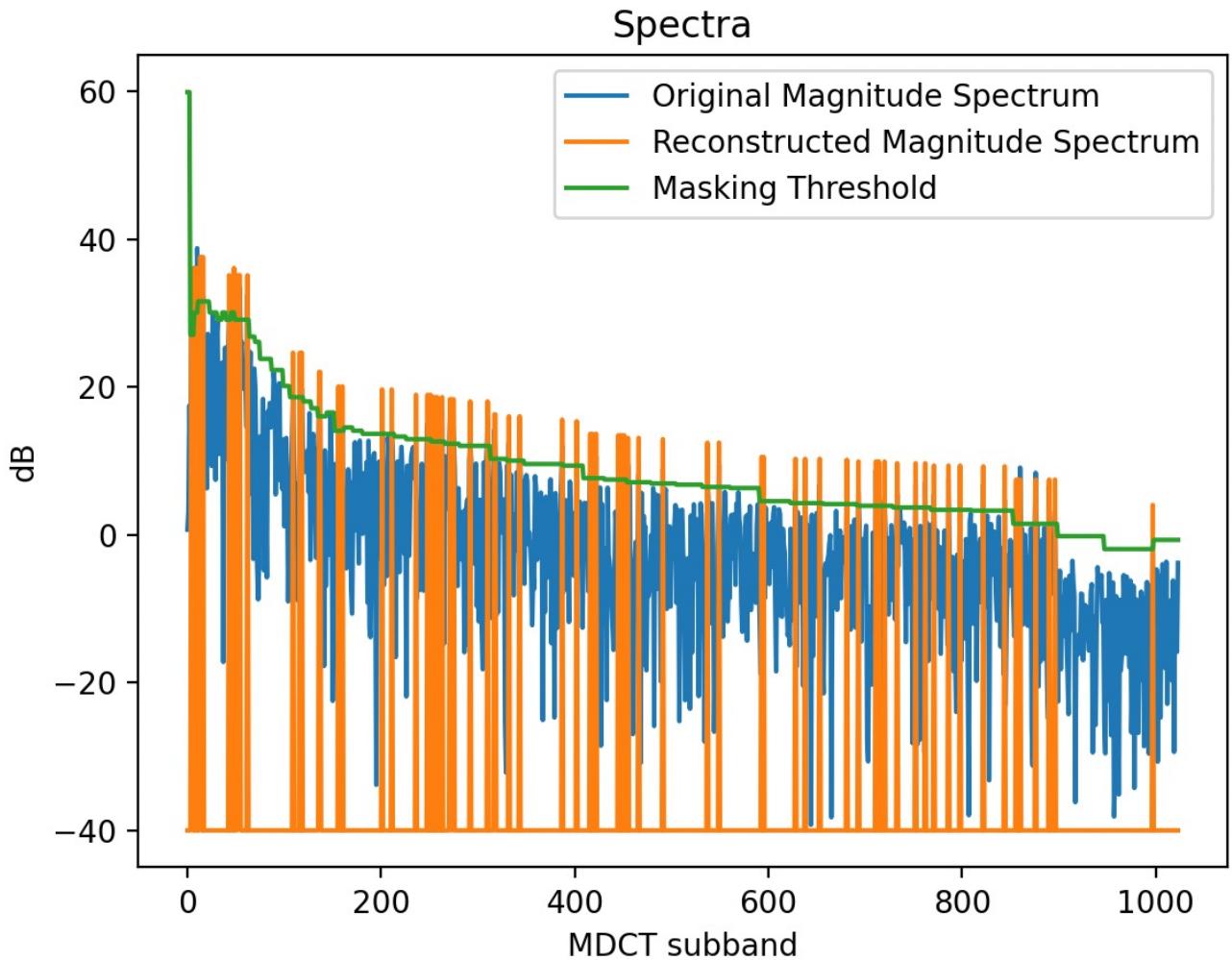
# Male Voice Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



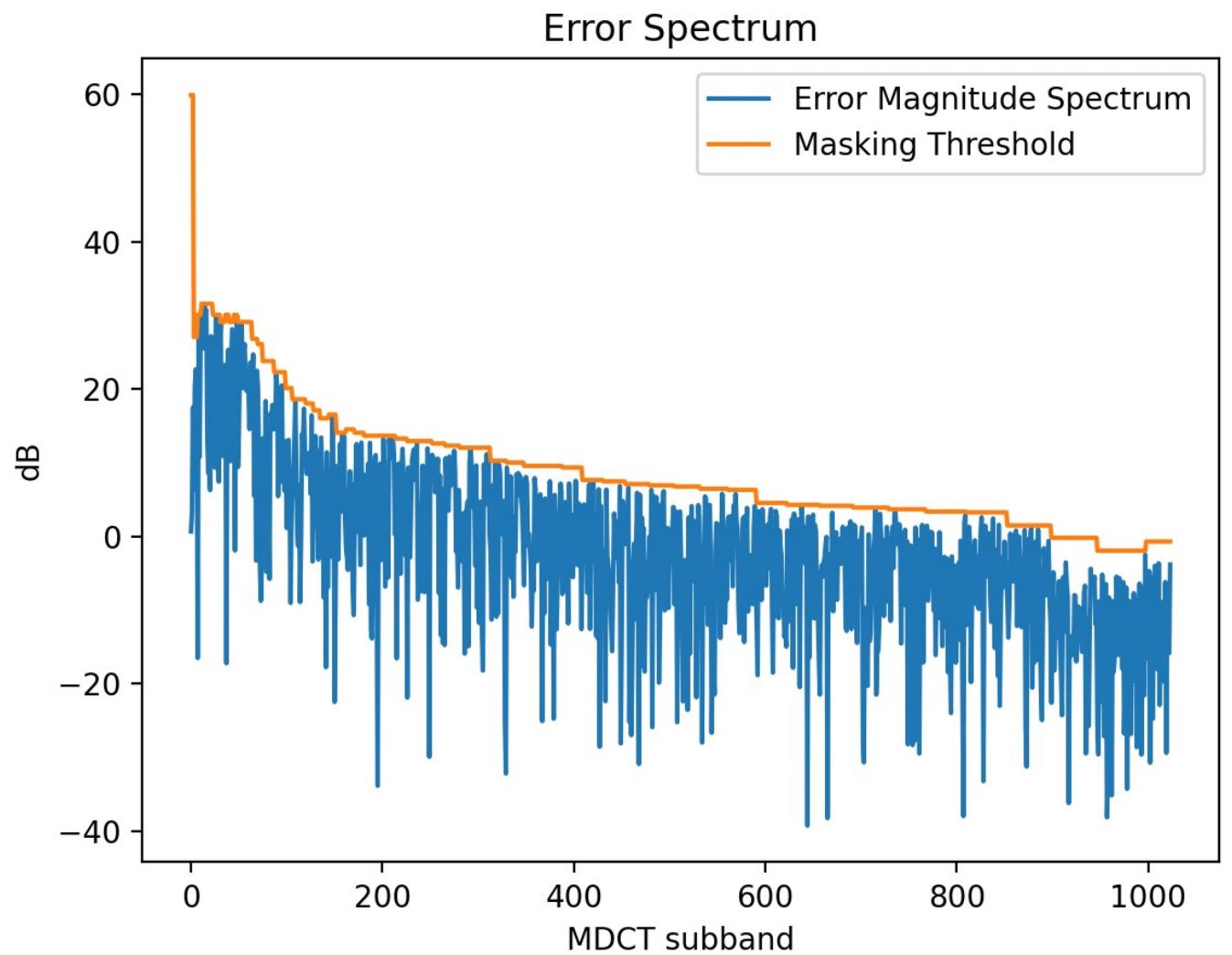
# Male Voice Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



# Male Voice Quantization Error

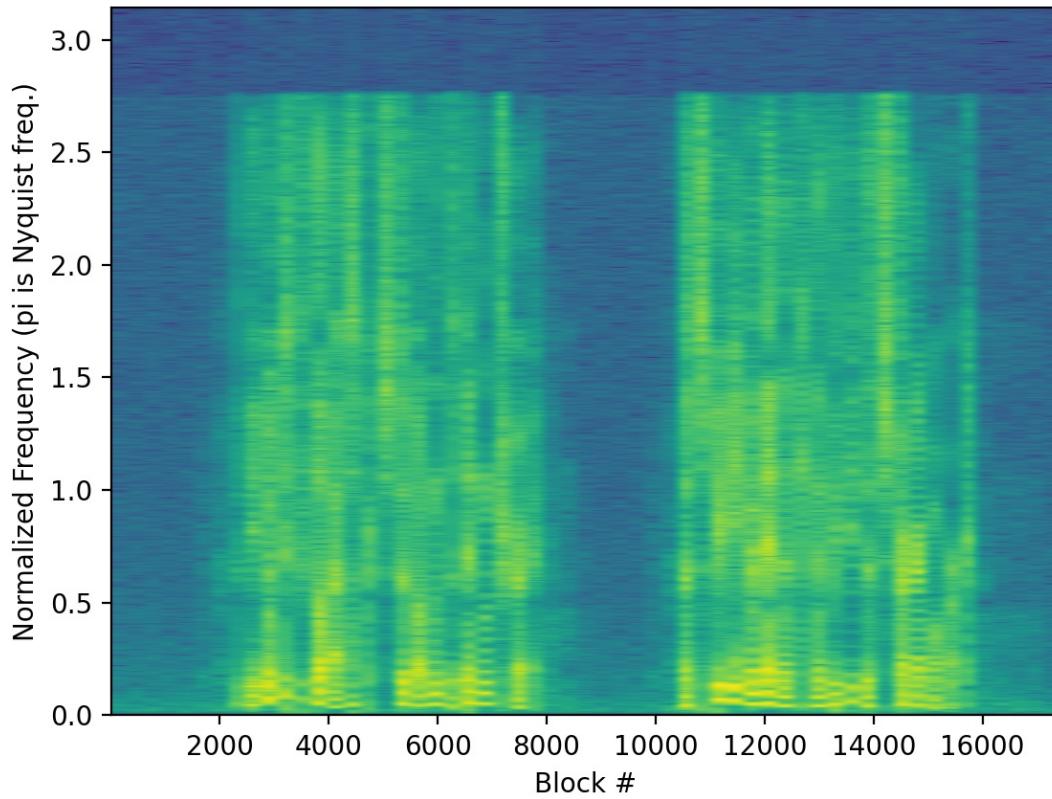
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



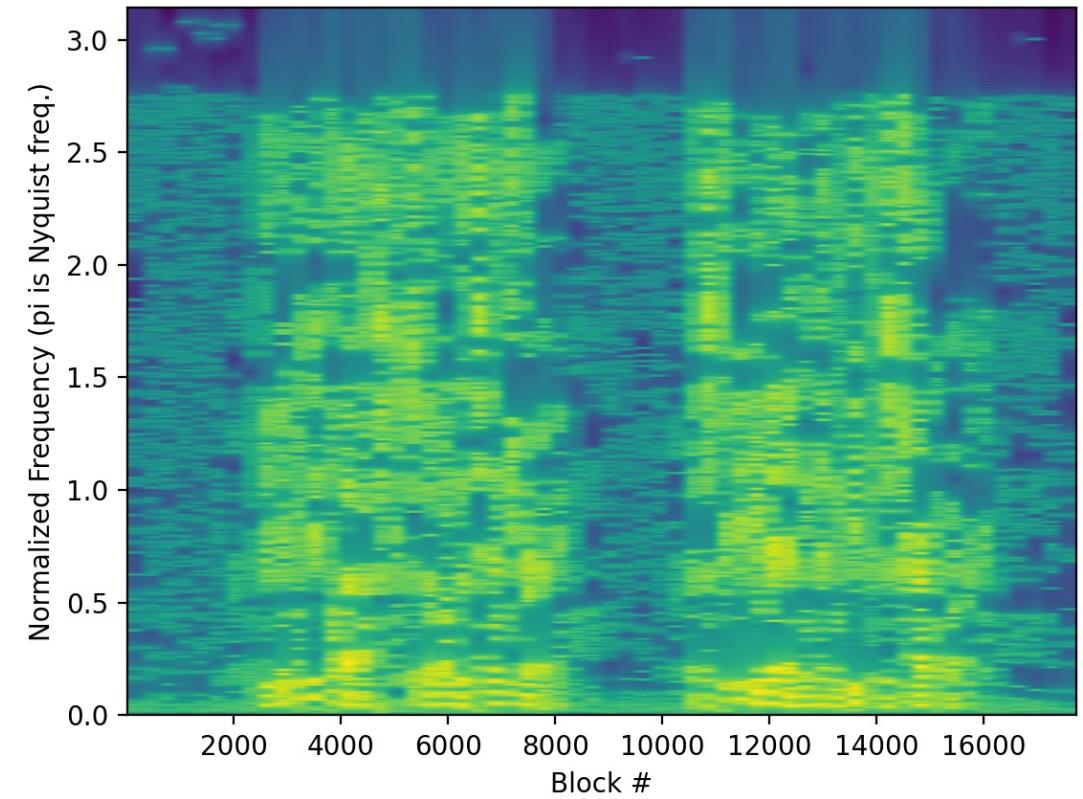
# Male Voice Quantization Error

60 Quality, 1024 MDCT Subbands

Original



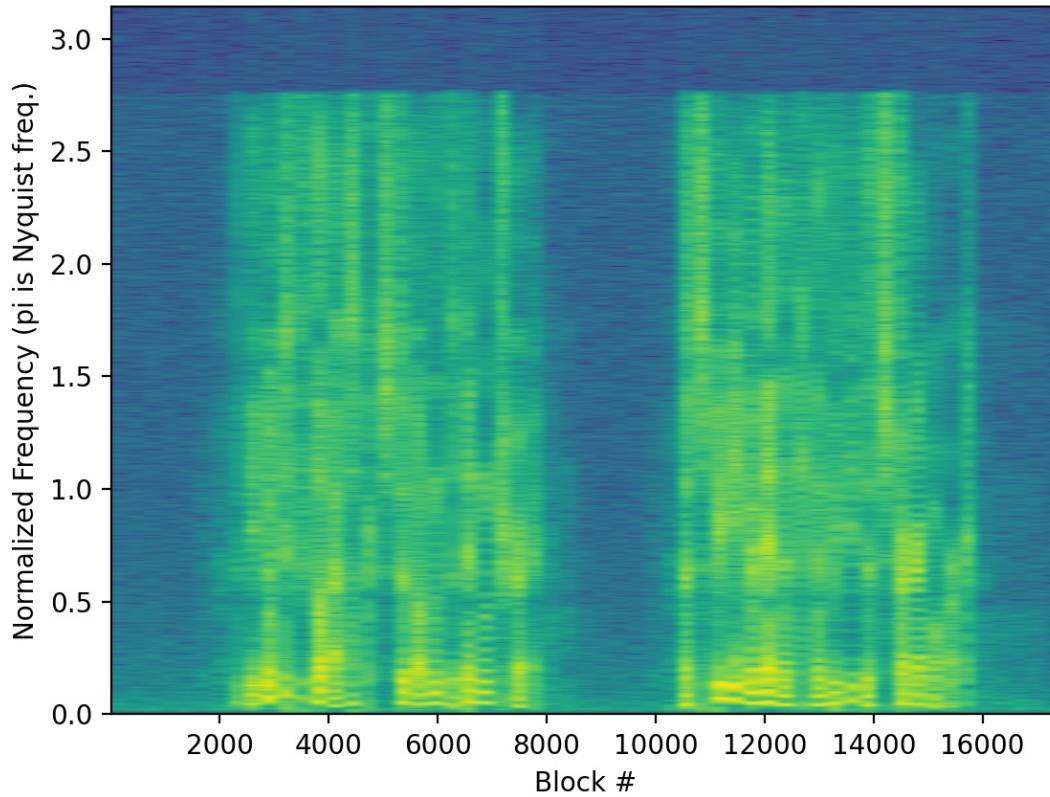
Reconstructed



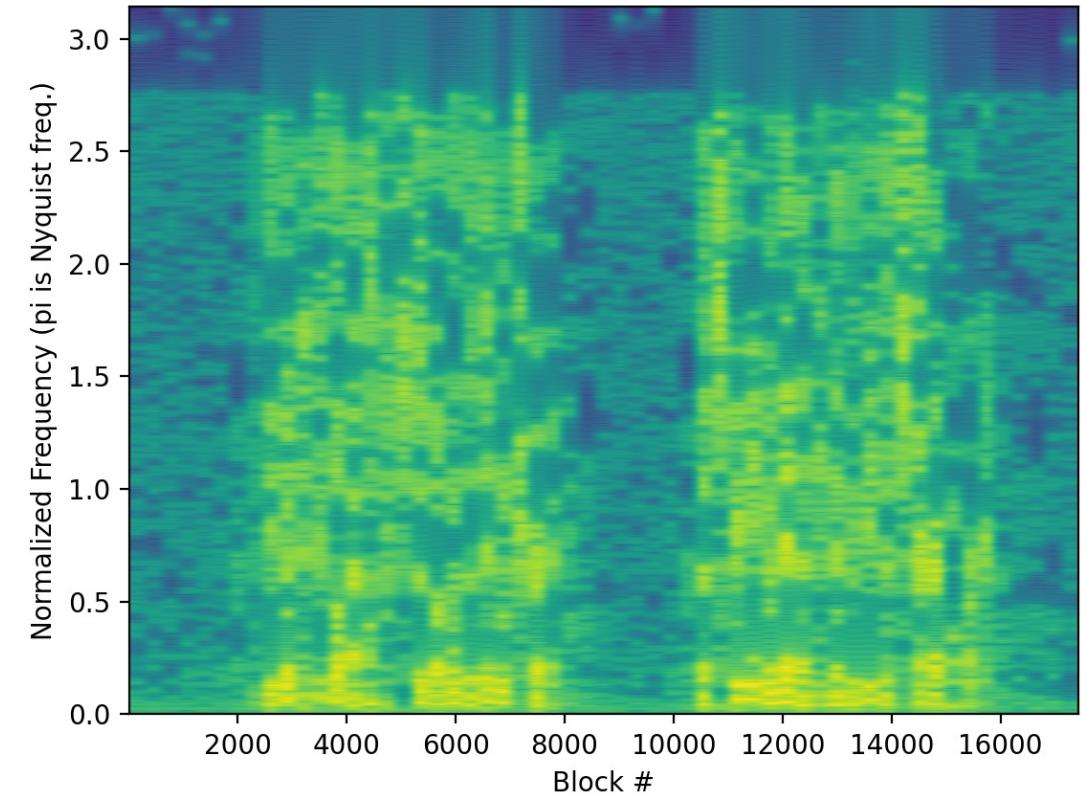
# Male Voice Quantization Error

60 Quality, 256 MDCT Subbands

Original



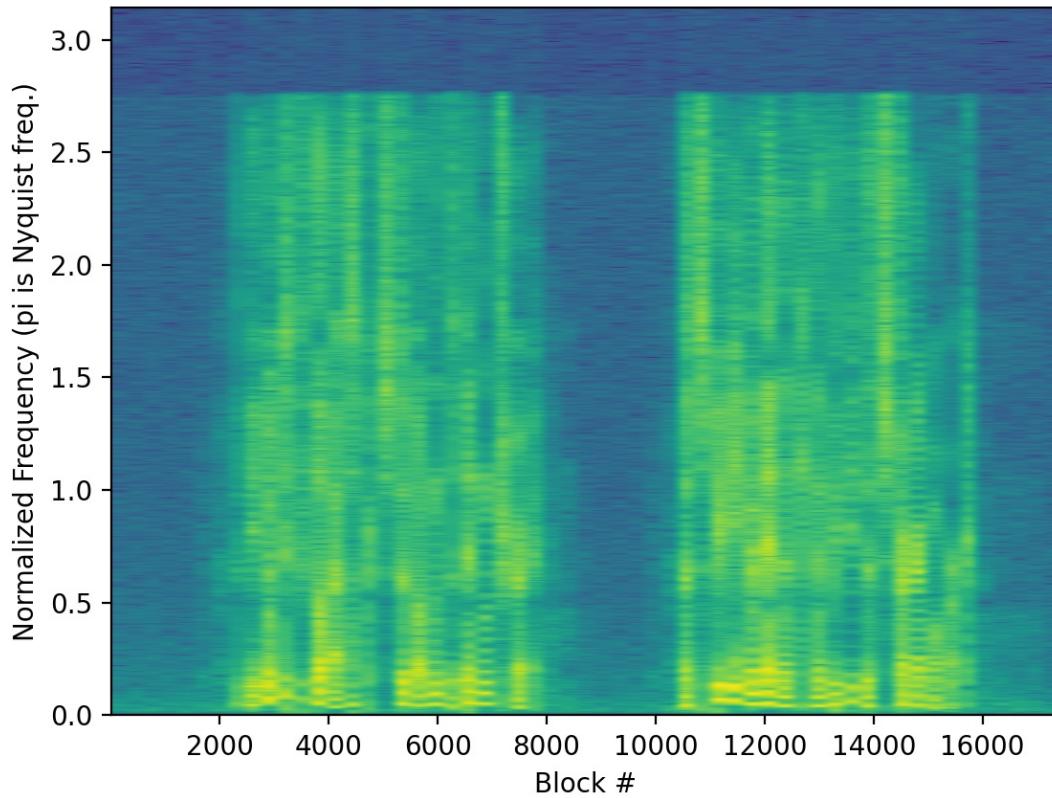
Reconstructed



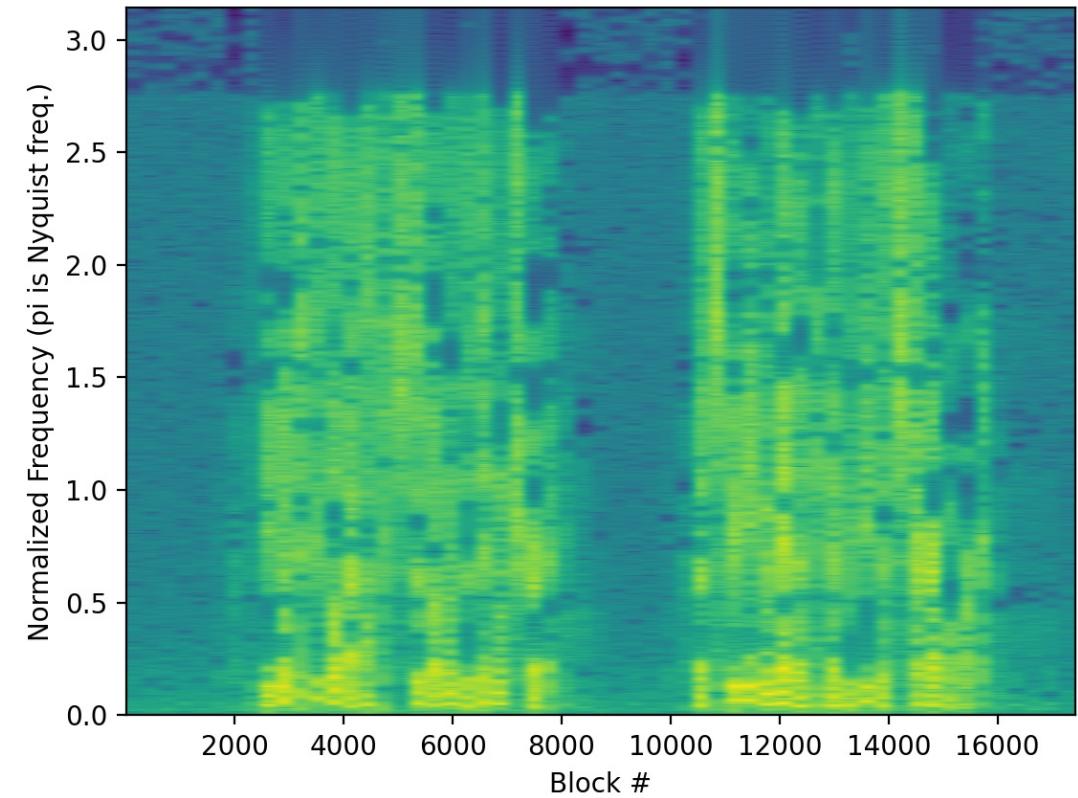
# Male Voice Quantization Error

100 Quality, 256 MDCT Subbands

Original



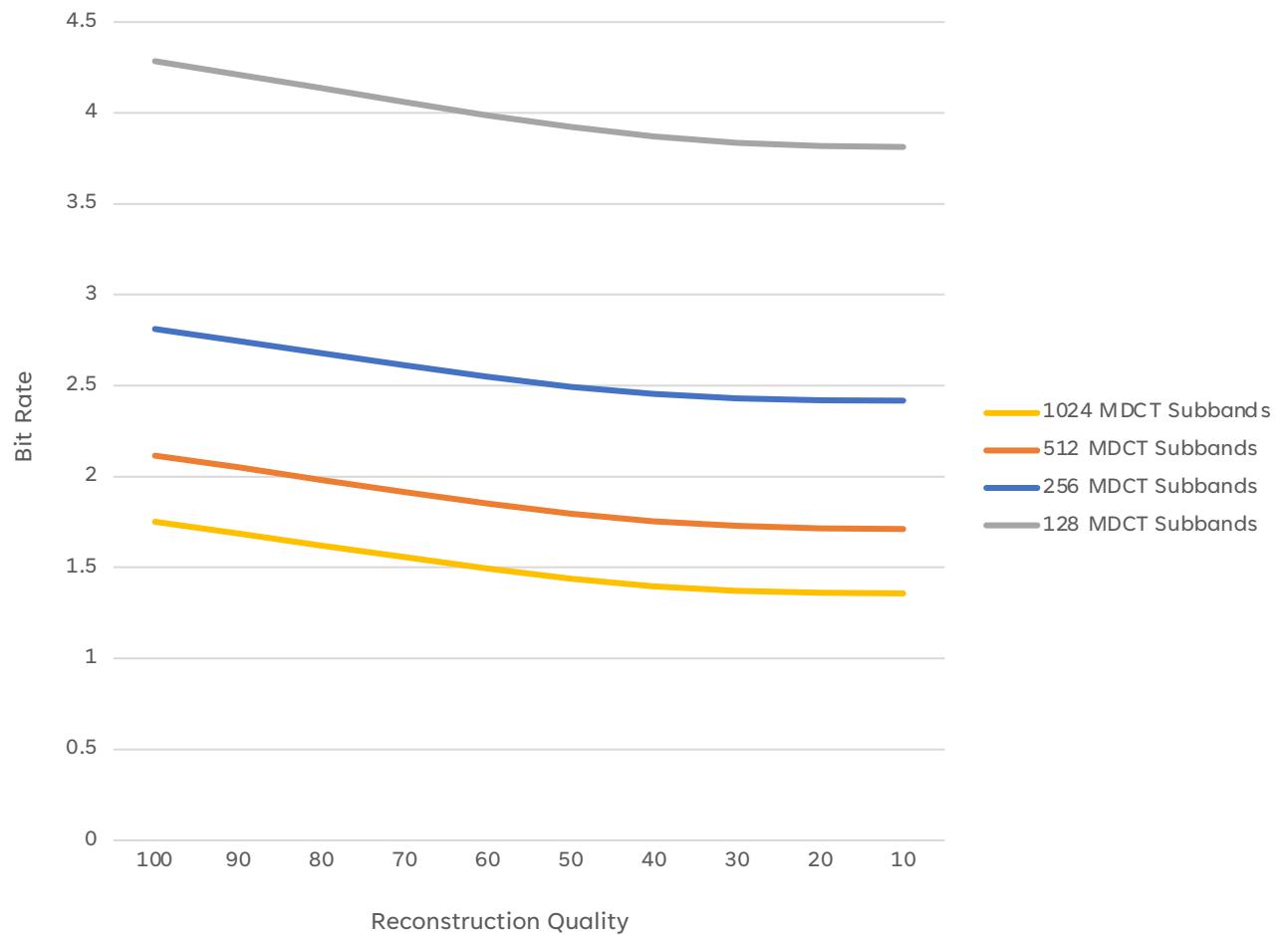
Reconstructed



# Bit Rate vs Reconstruction Quality

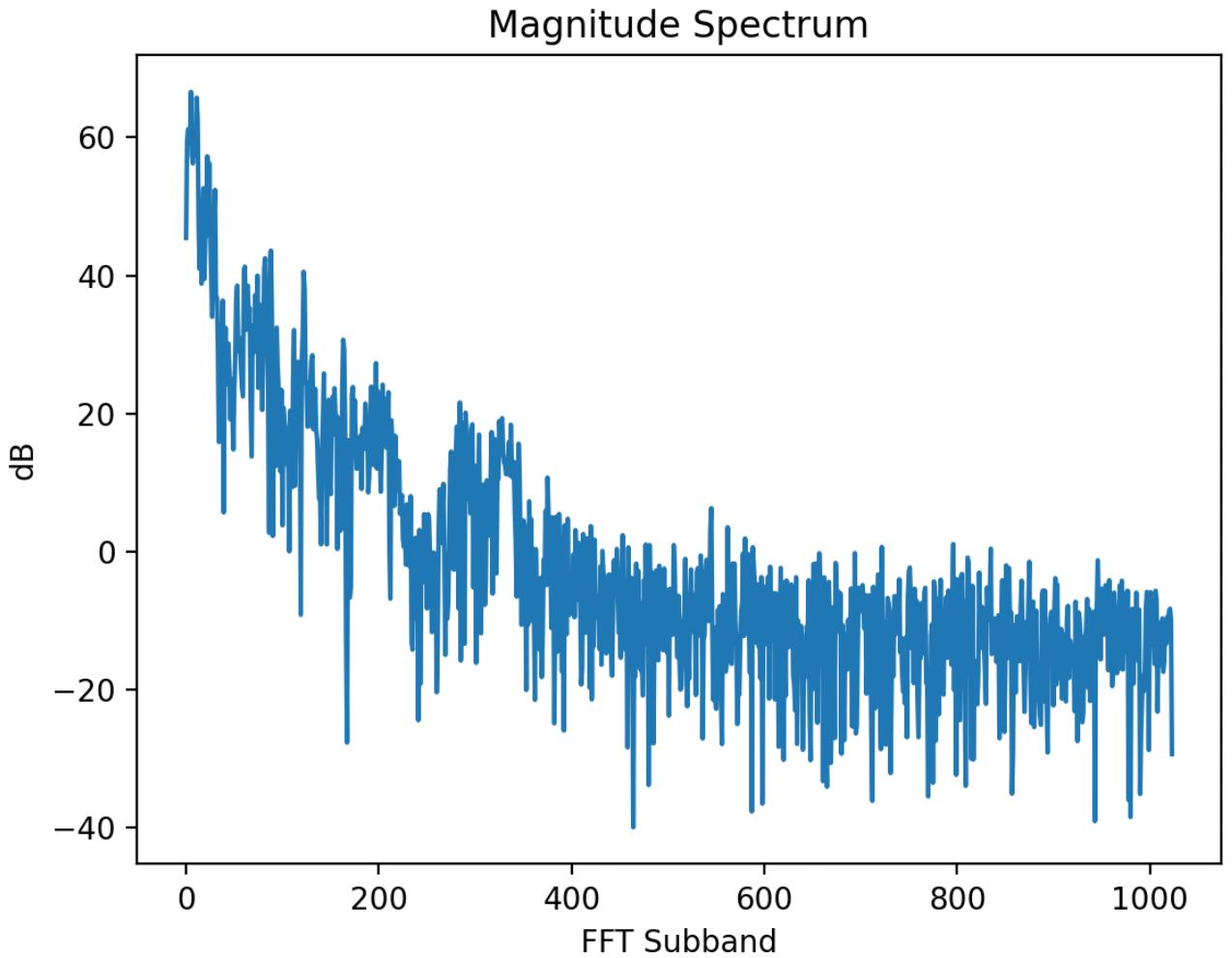
Male Voice Input

- Decreasing MDCT Subbands removes reverb but increases bit rate
- Decreasing reconstruction quality (increasing the masking threshold) decreases bit rate but increases quantization noise



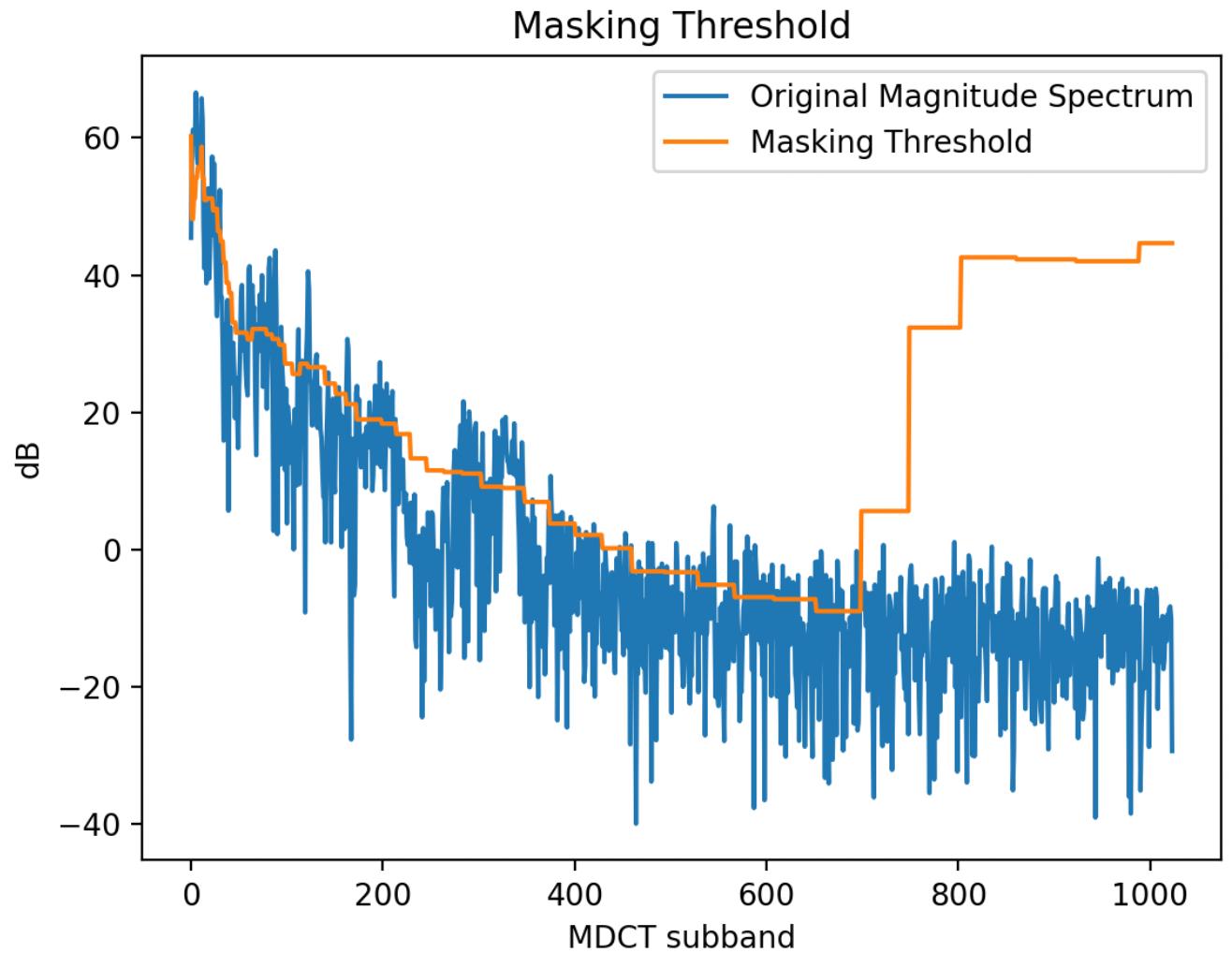
# Music Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



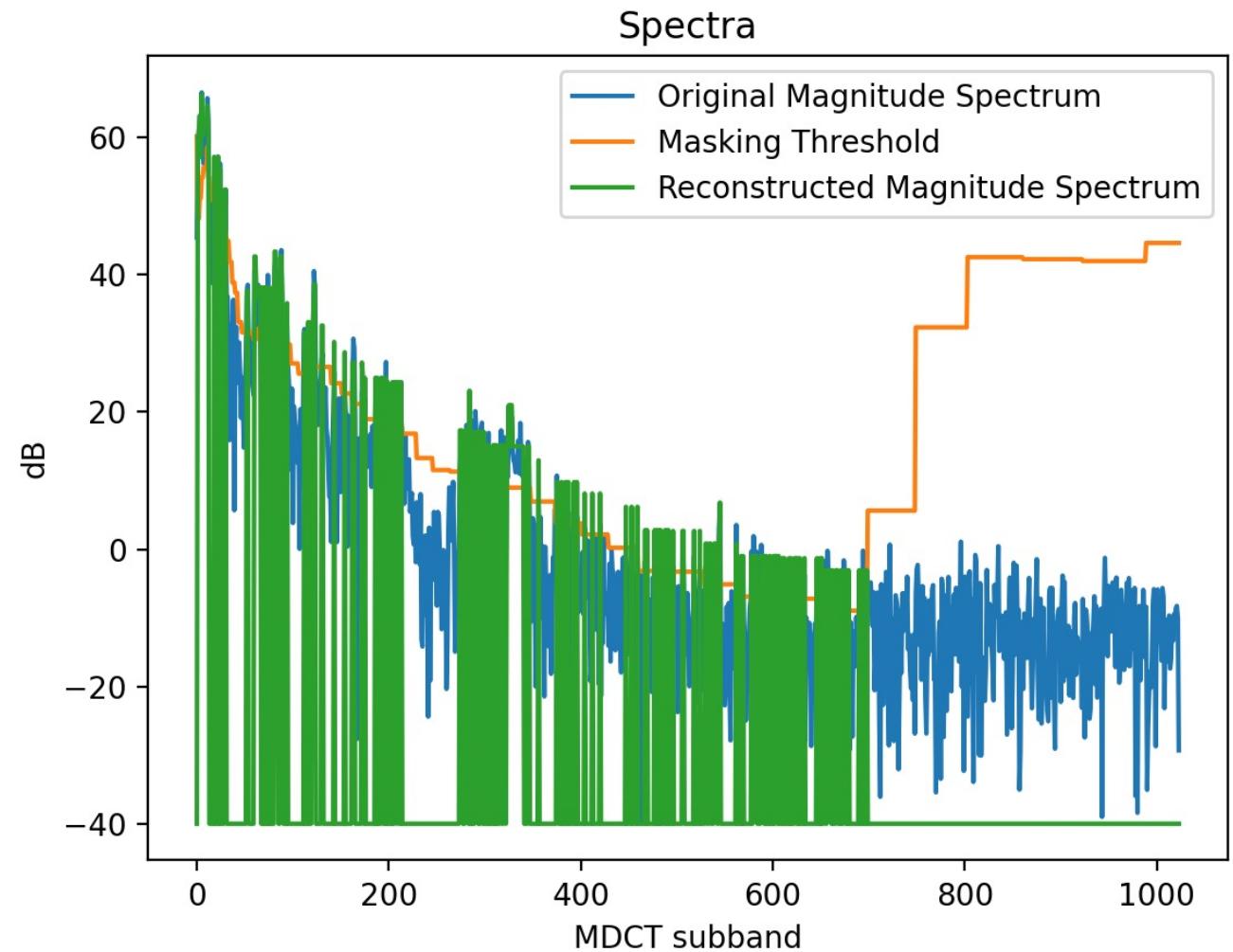
# Music Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



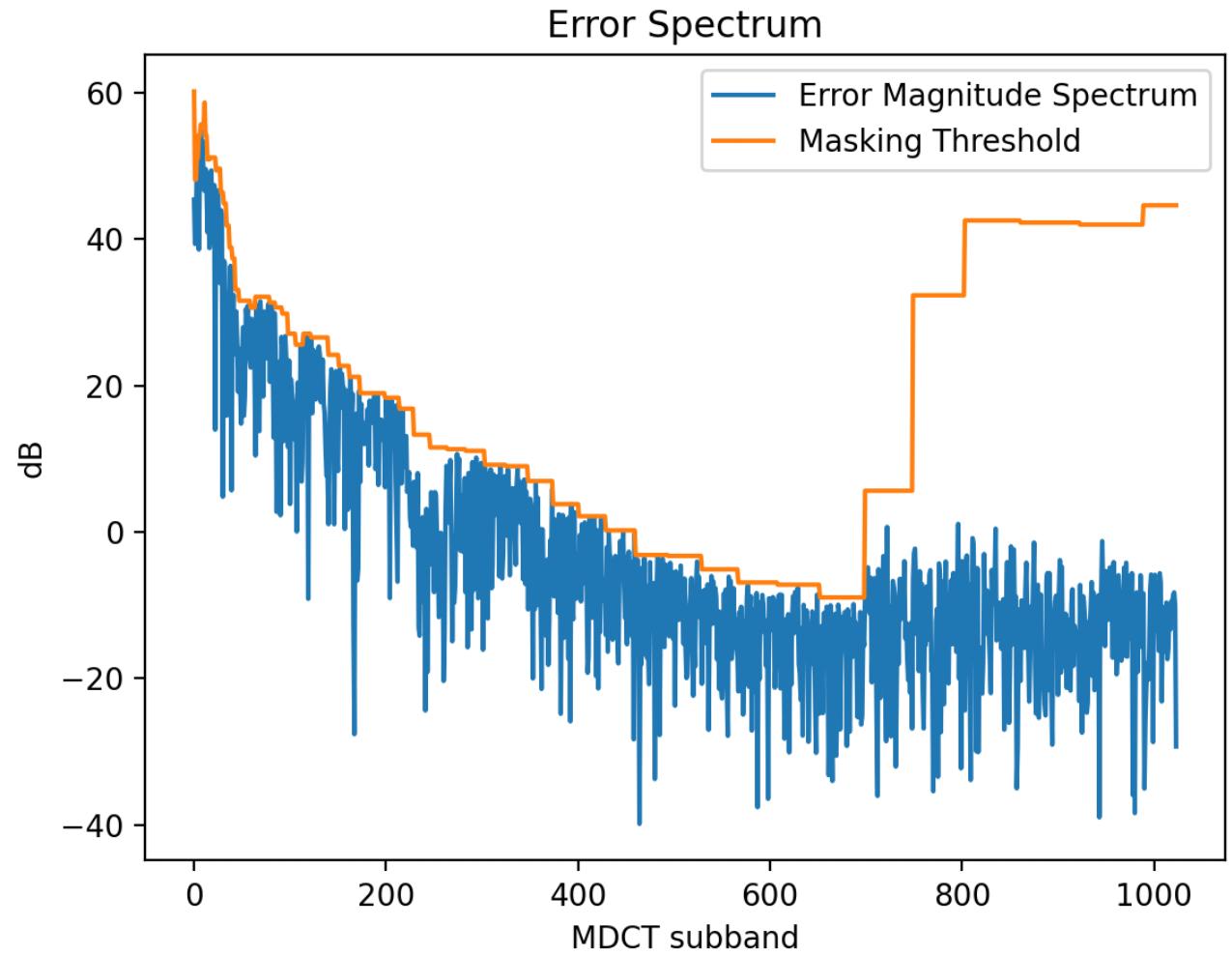
# Music Quantization Error

- FFT
- Masking Threshold
- **Reconstructed Signal**
- Quantization Error



# Music Quantization Error

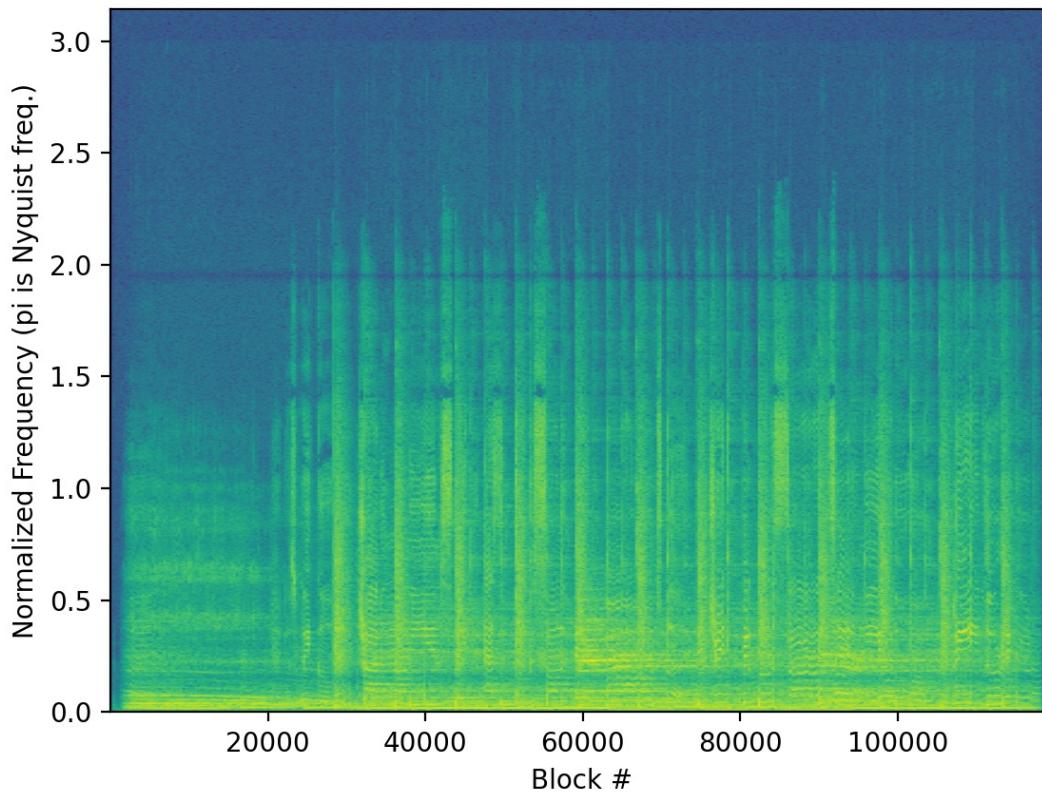
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



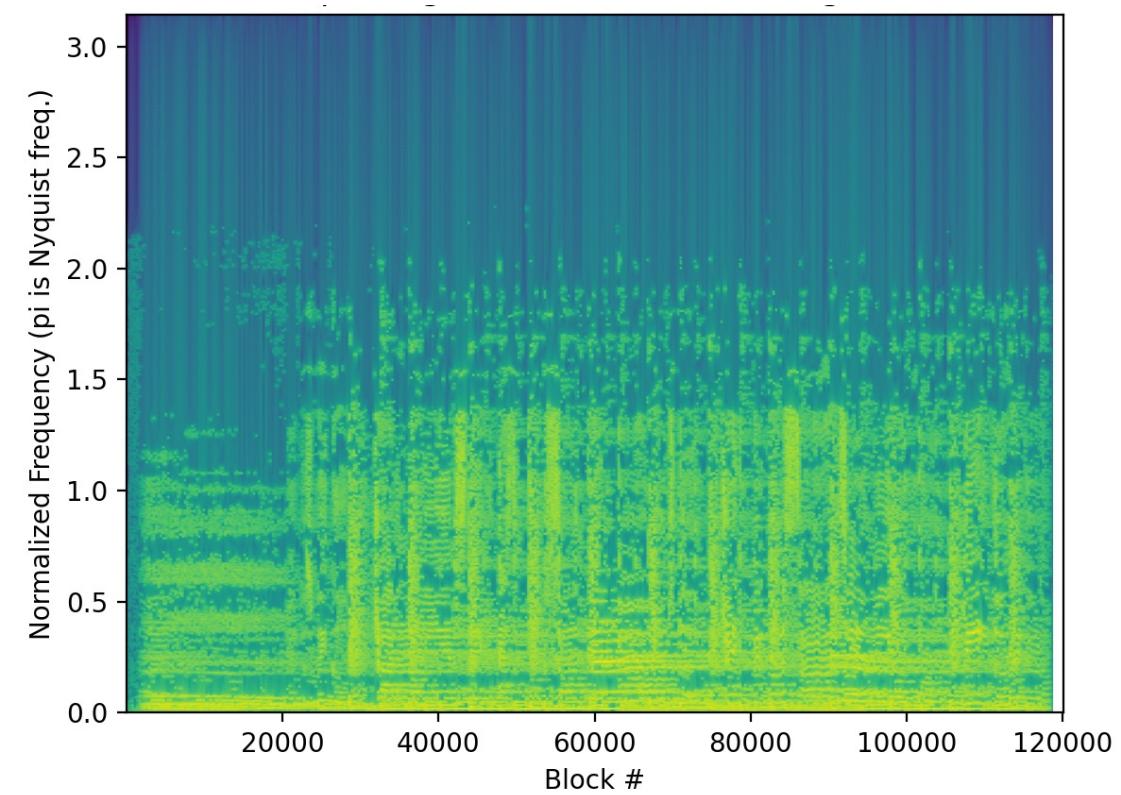
# Music Quantization Error

60 Quality, 1024 MDCT Subbands

Original



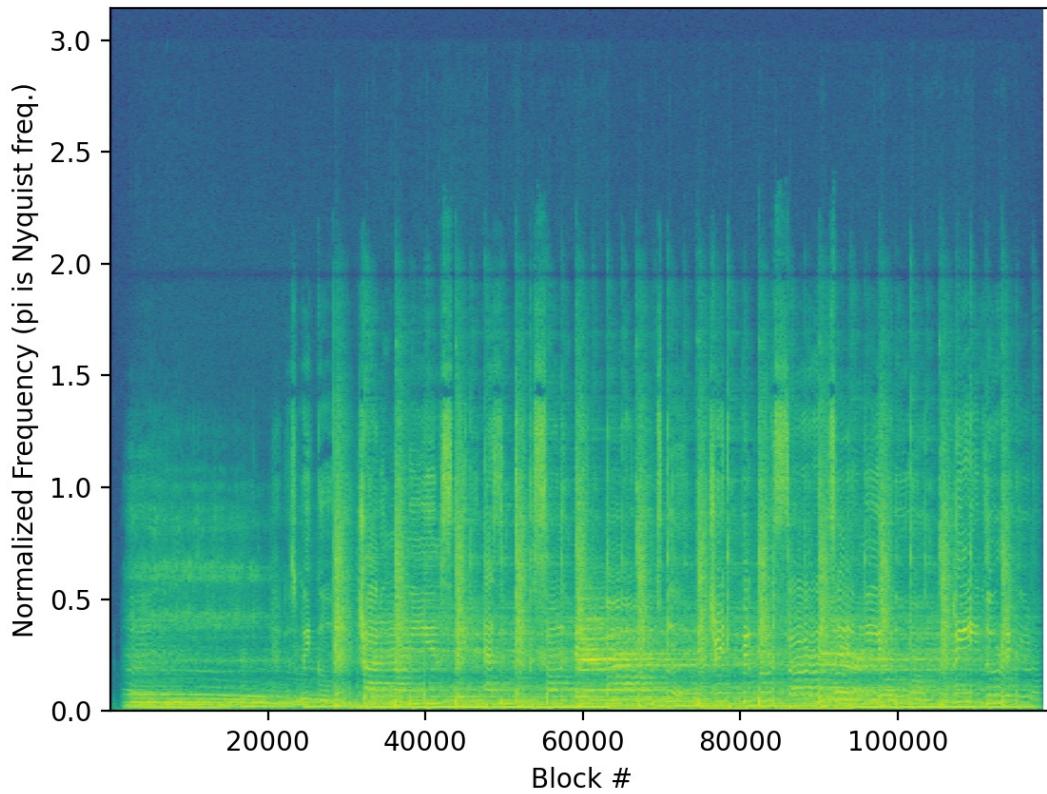
Reconstructed



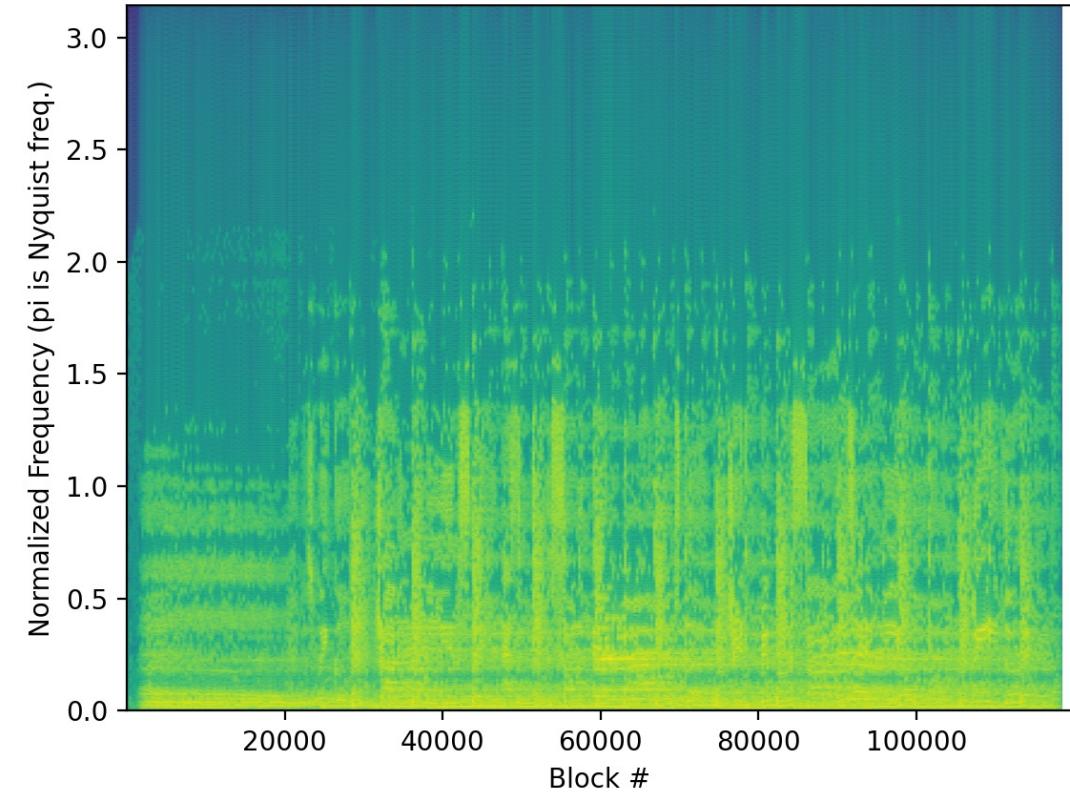
# Music Quantization Error

60 Quality, 256 MDCT Subbands

Original



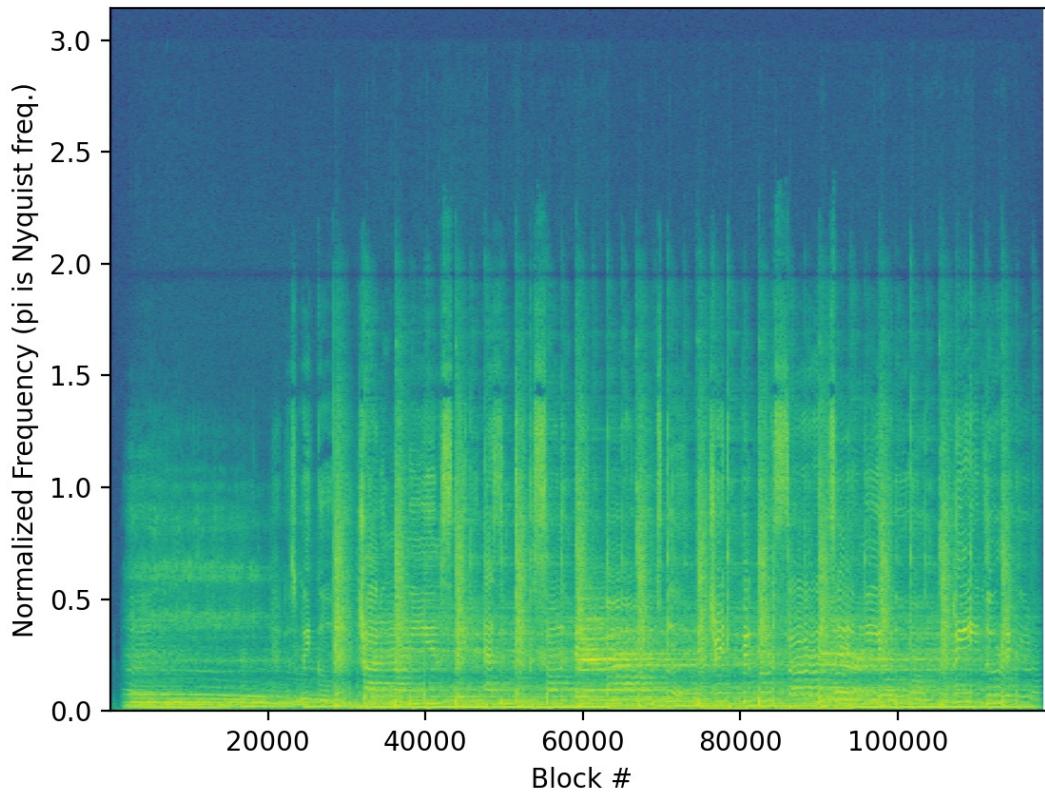
Reconstructed



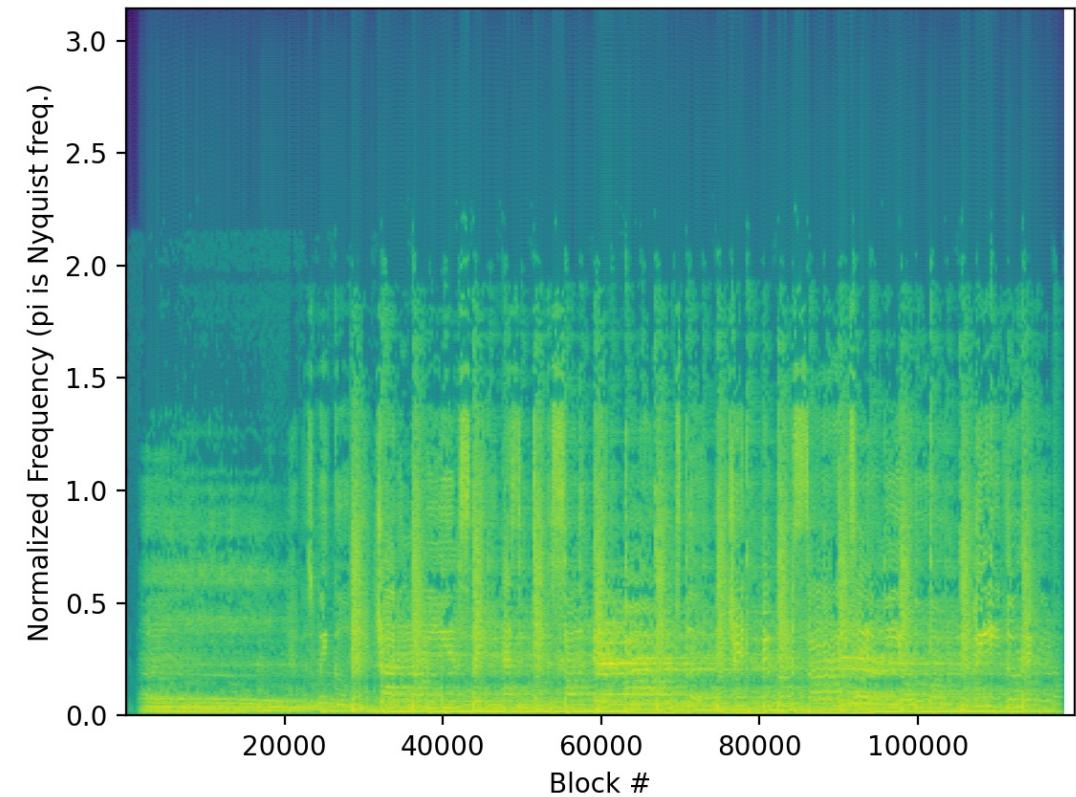
# Music Quantization Error

100 Quality, 256 MDCT Subbands

Original



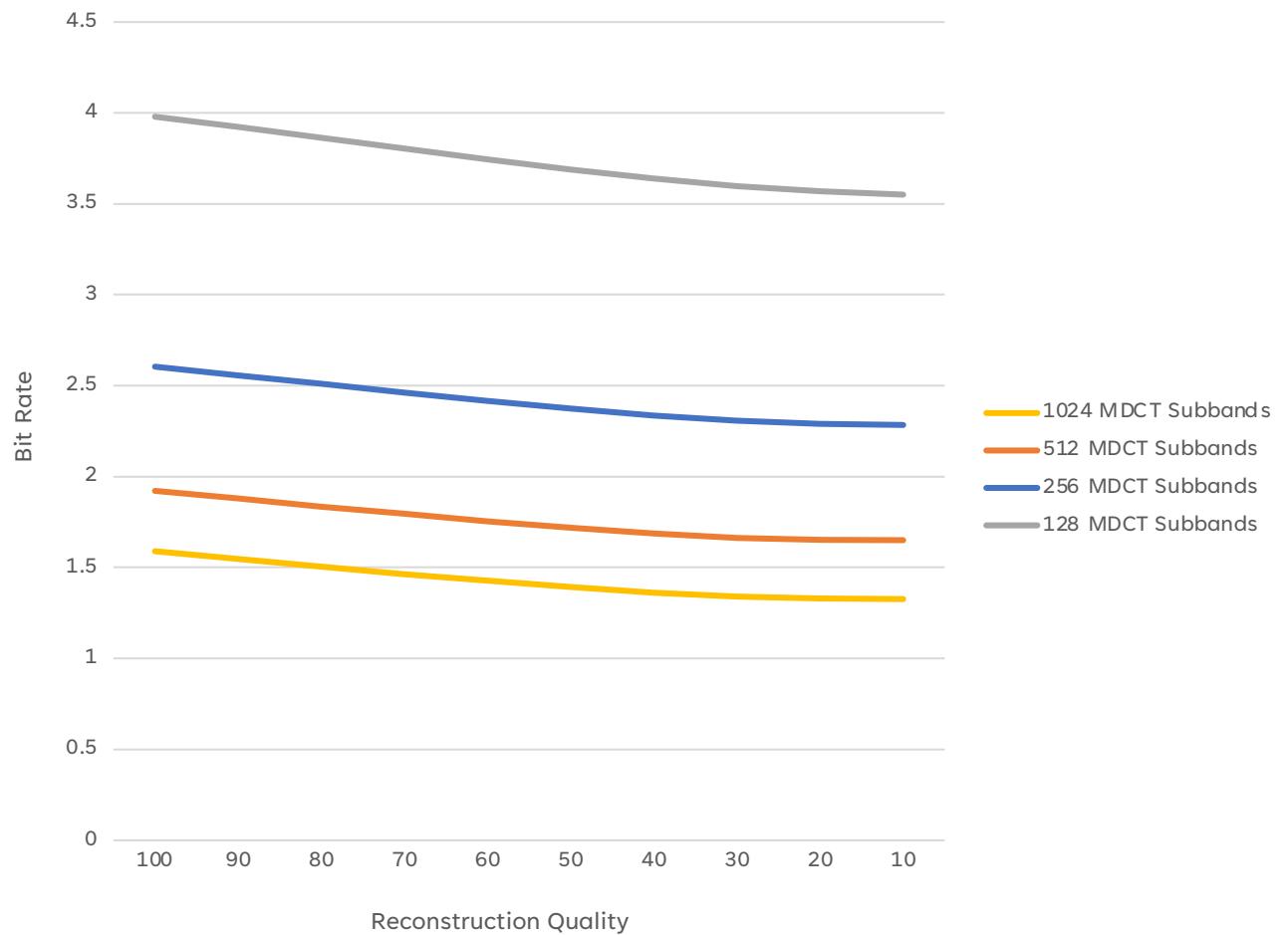
Reconstructed



# Bit Rate vs Reconstruction Quality

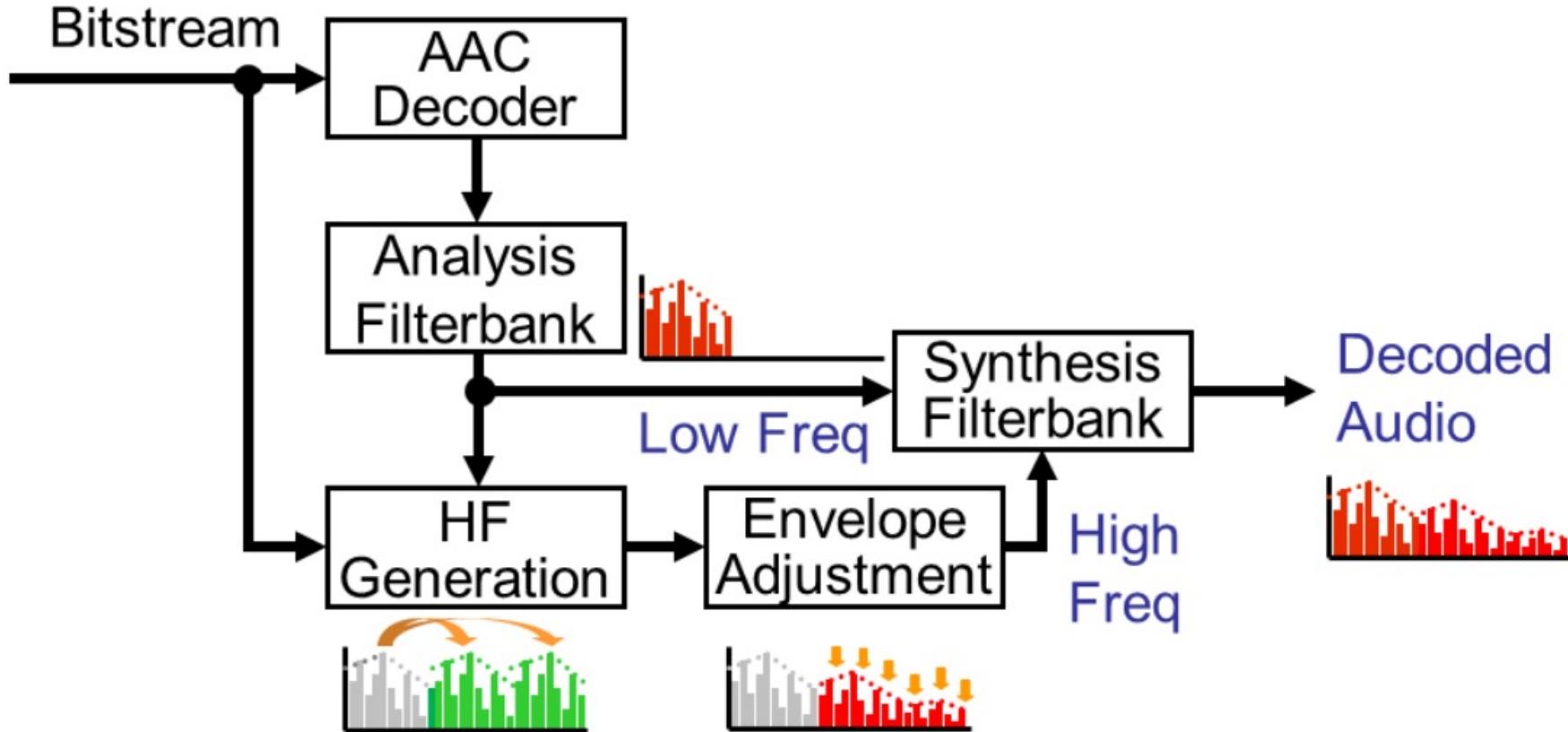
Music Input

- Decreasing MDCT Subbands increases distortion and bit rate
- Decreasing reconstruction quality (increasing the masking threshold) decreases bit rate but increases flanging



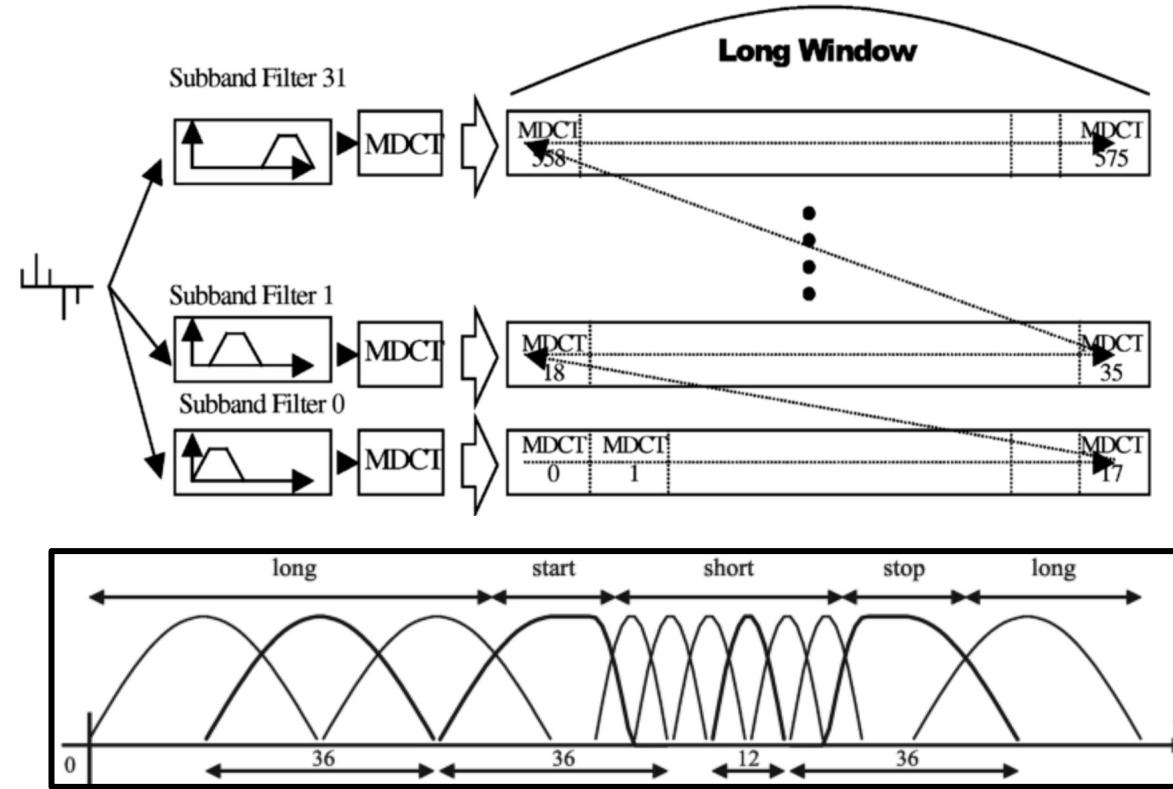
# Current Solutions

Bandwidth Extension (Subband Replication)



# Current Solutions

## Adaptive Windowing



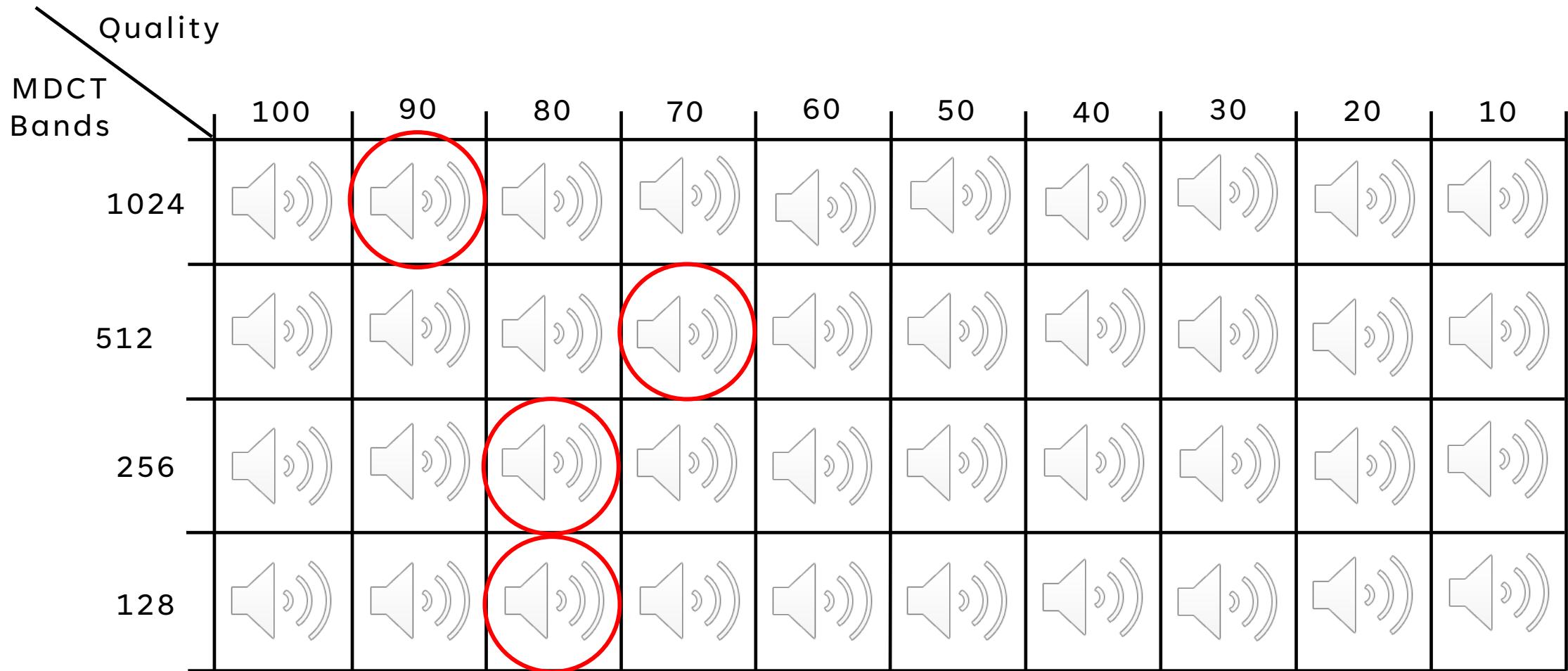
(top) Kiranyaz, Serkan & Qureshi, Ahmad & Gabbouj, Moncef. (2004). A Generic Audio Classification and Segmentation Approach for Multimedia Indexing and Retrieval. *Audio, Speech, and Language Processing, IEEE Transactions on*. 14. 1062 - 1081. 10.1109/TSA.2005.857573.

(bottom) Zieliński, T.P. (2021). Audio Compression. In: Starting Digital Signal Processing in Telecommunication Engineering. Textbooks in Telecommunication Engineering. Springer, Cham. [https://doi.org/10.1007/978-3-030-49256-4\\_15](https://doi.org/10.1007/978-3-030-49256-4_15)

## Appendix



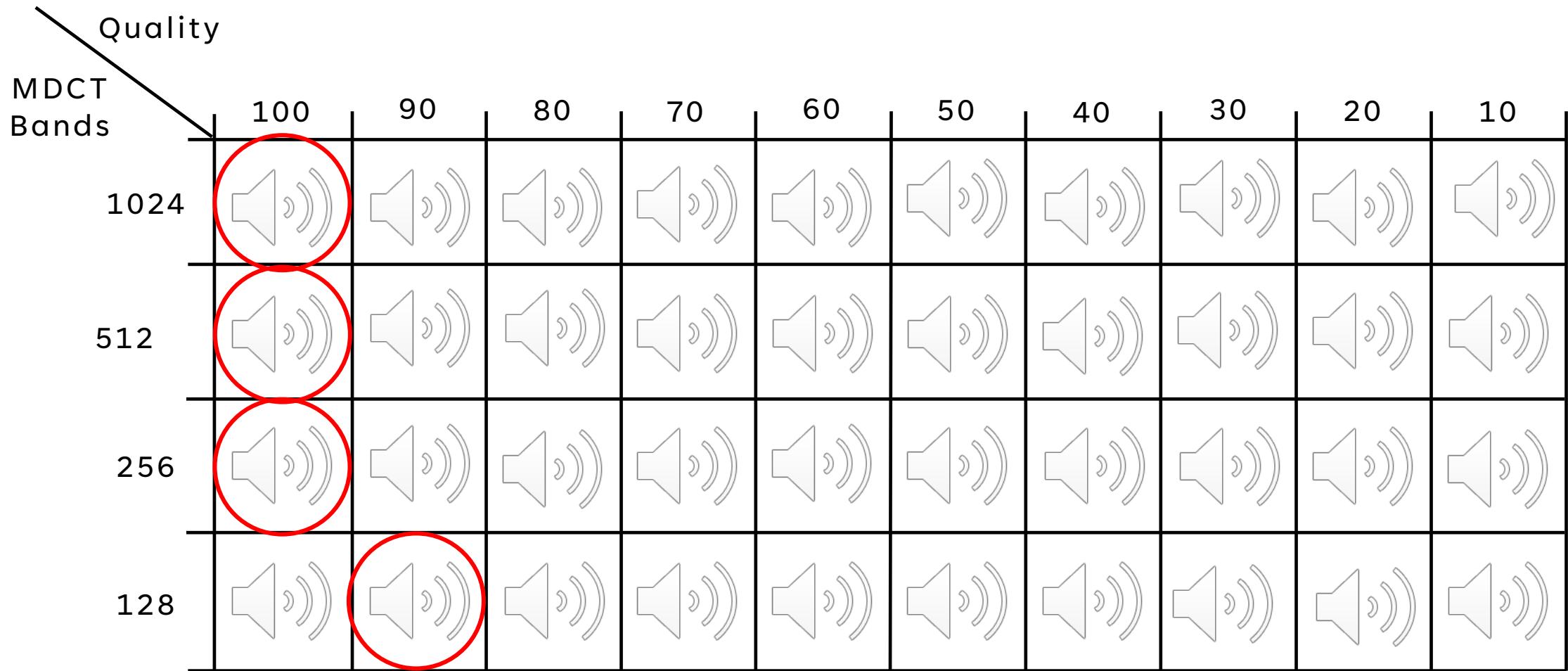
### i. Castanets Reconstructed Audio



## Appendix



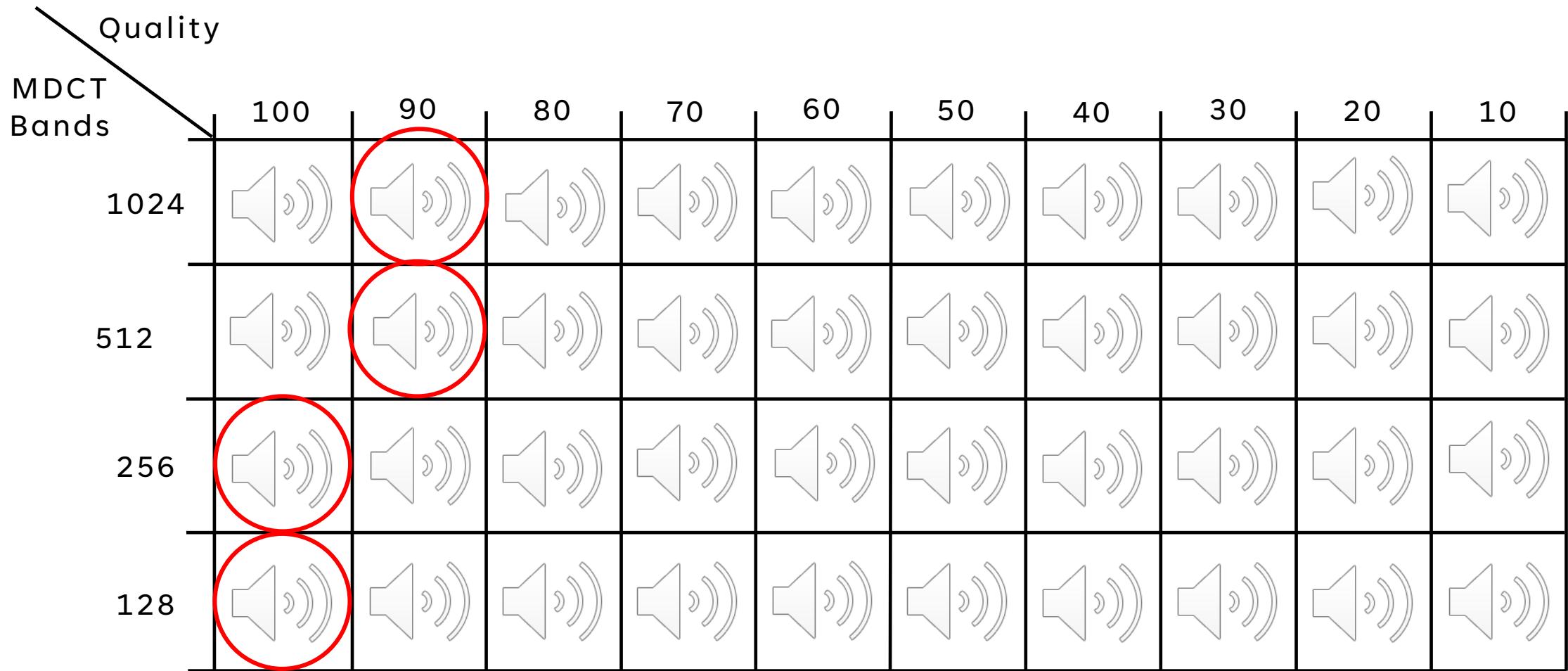
## ii. Male Voice Reconstructed Audio



## Appendix



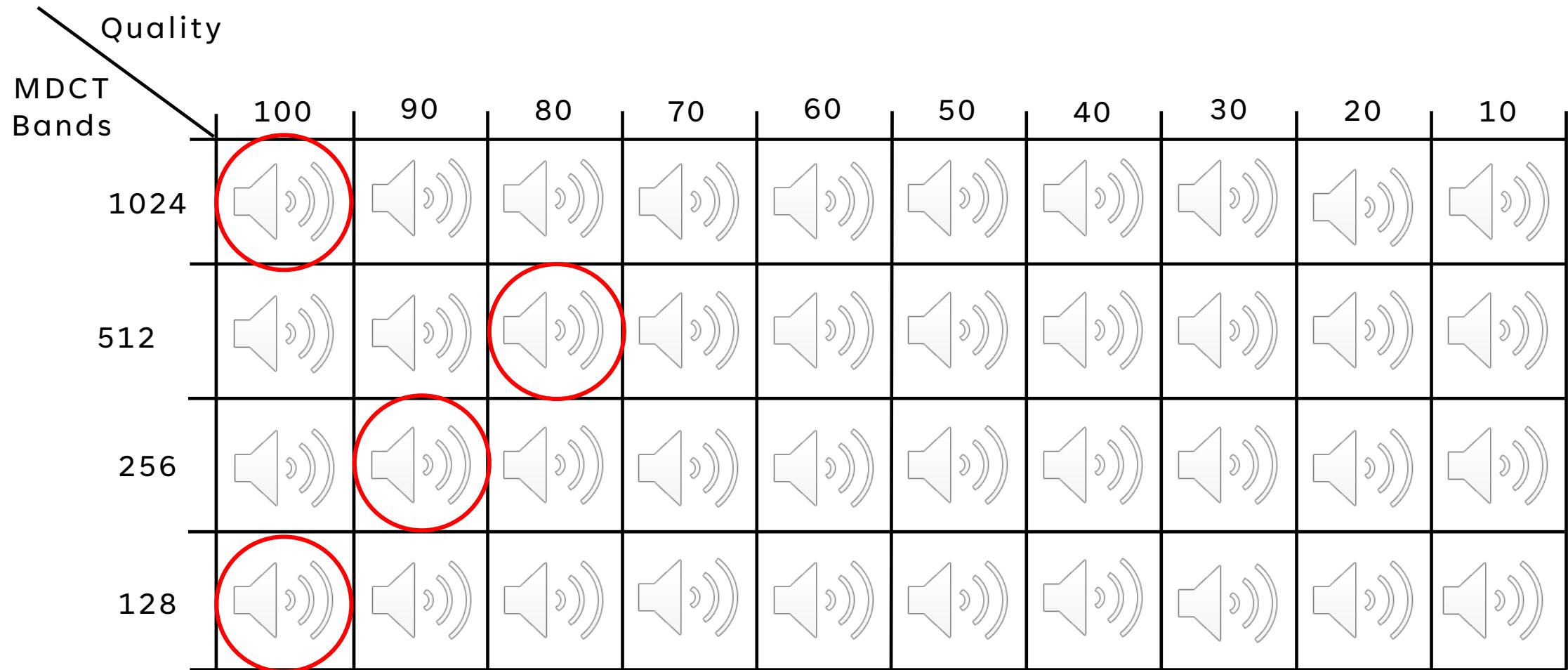
#### iv. Music Reconstructed Audio



## Appendix



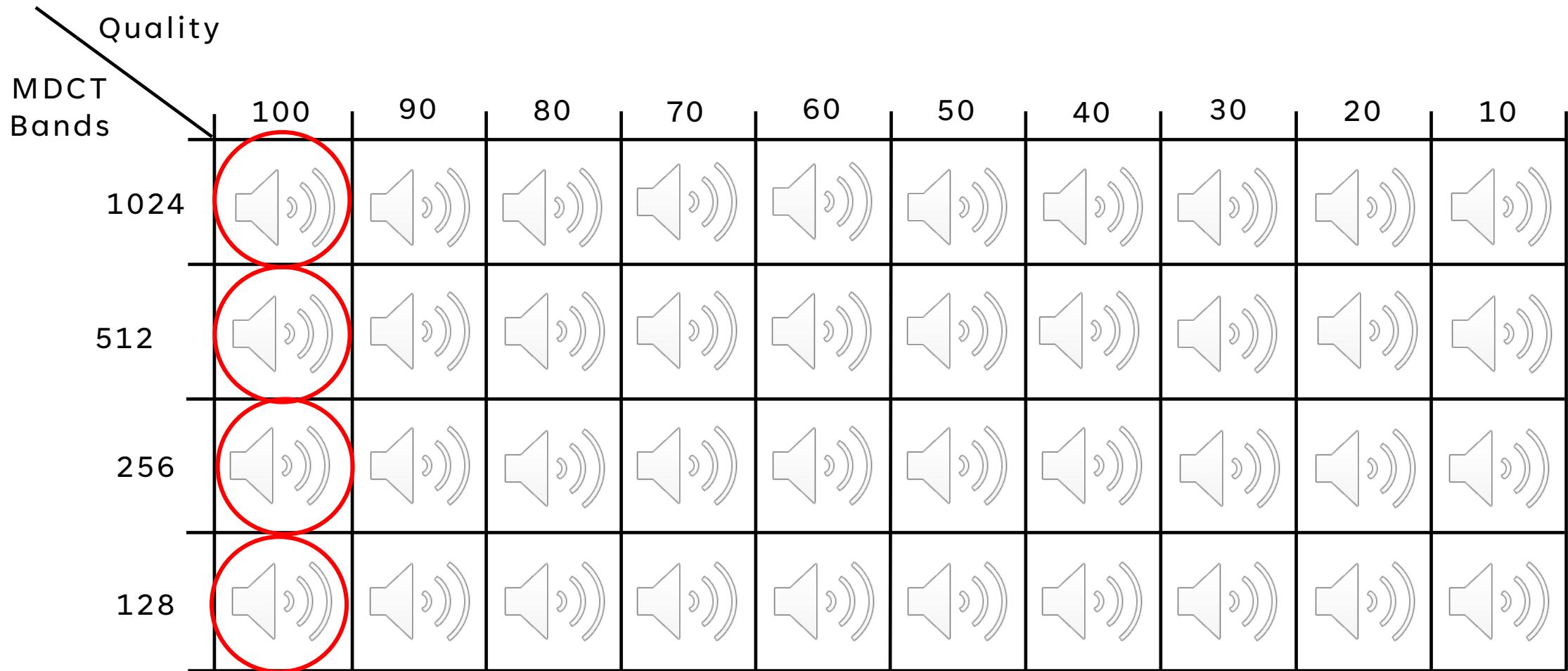
### iii. Female Voice Reconstructed Audio



## Appendix

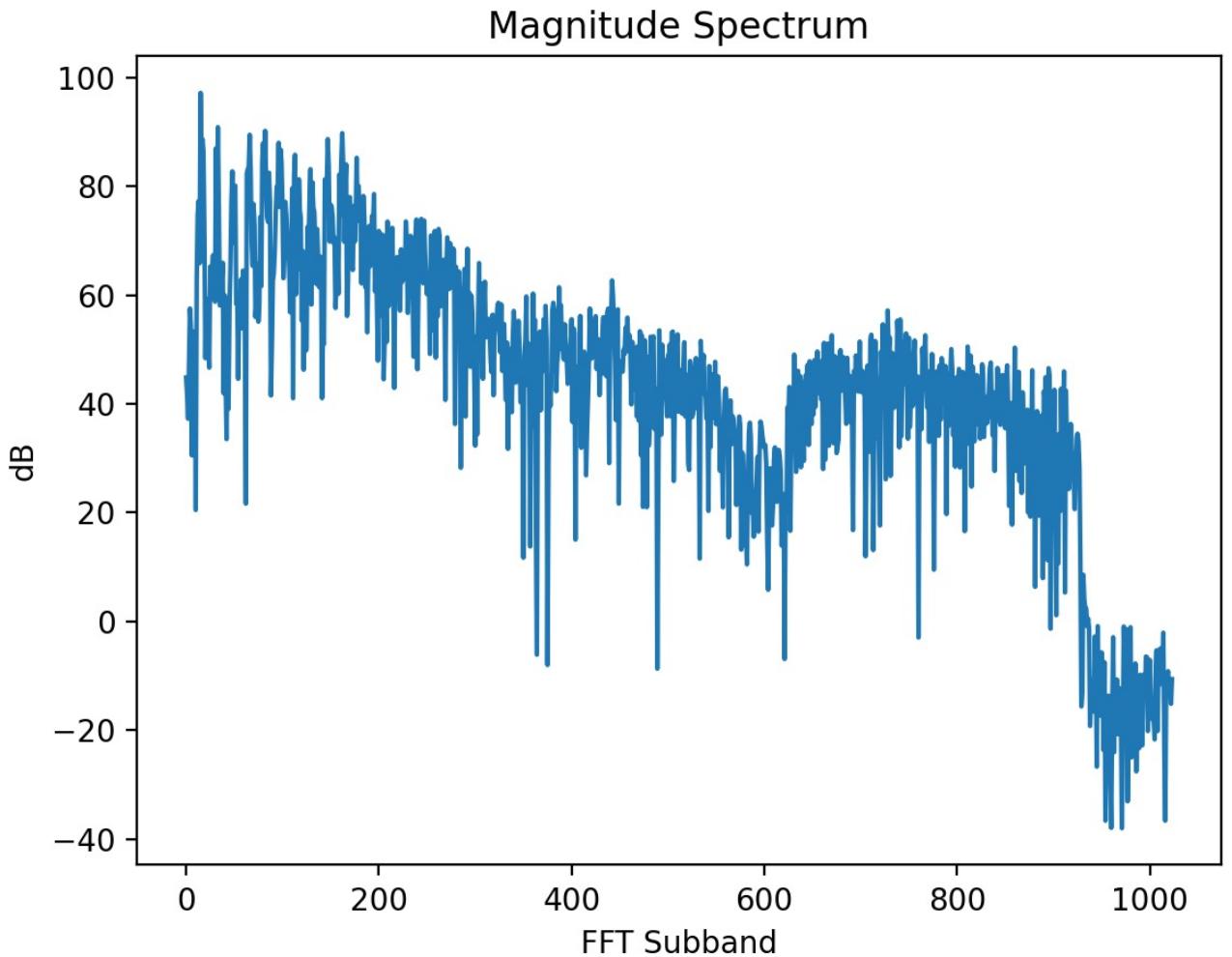


### v. Constant Tone Reconstructed Audio



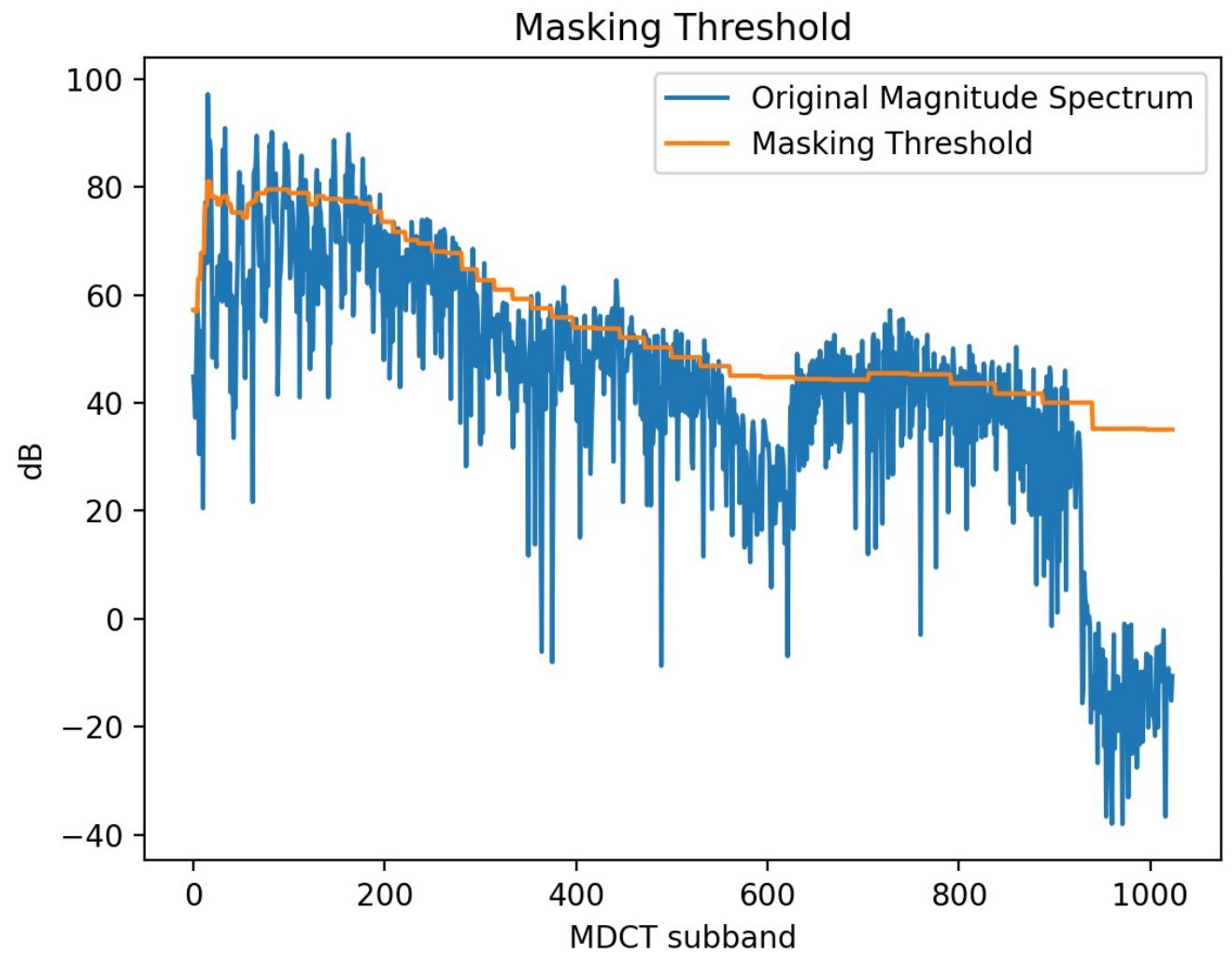
# Female Voice Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



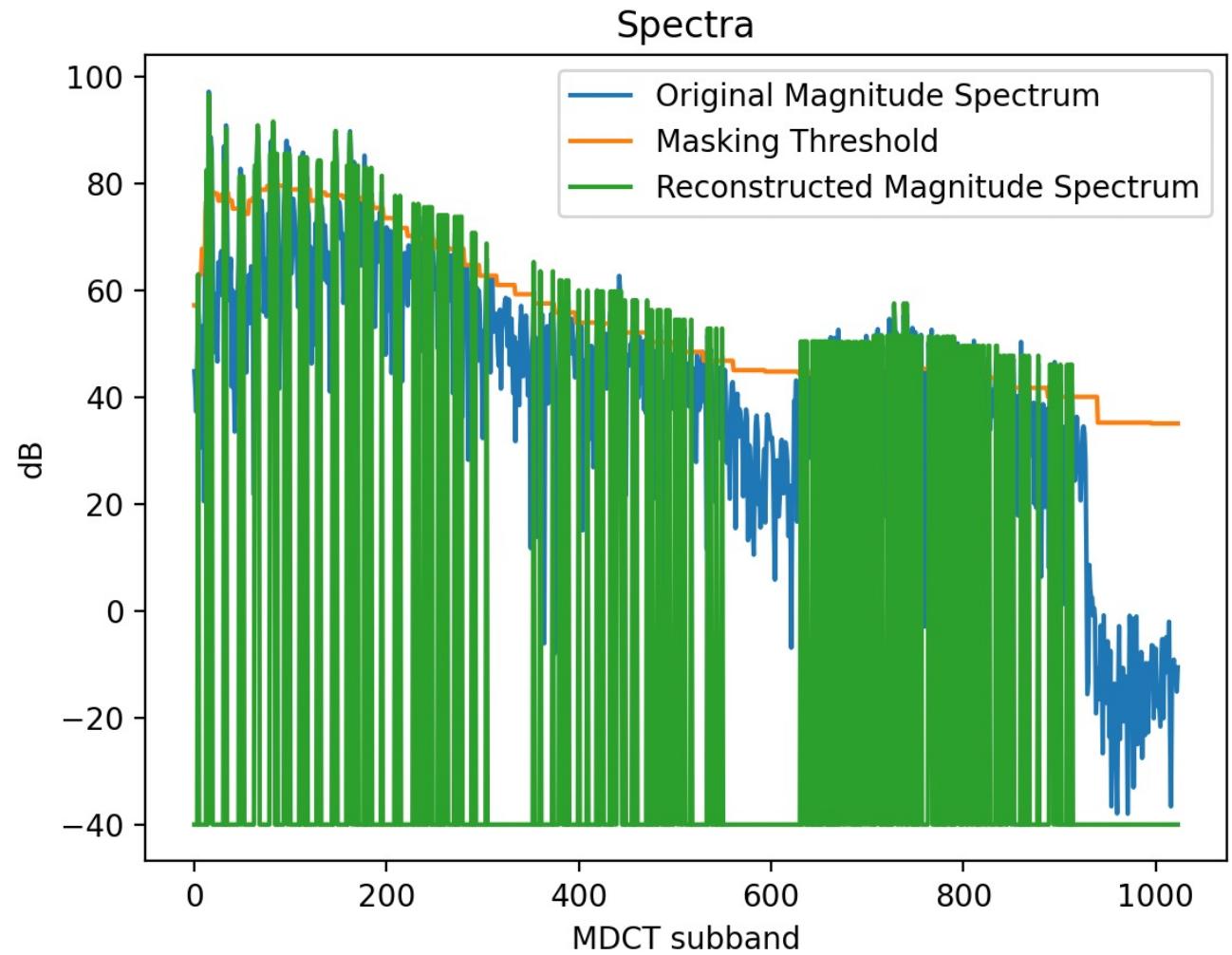
# Female Voice Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



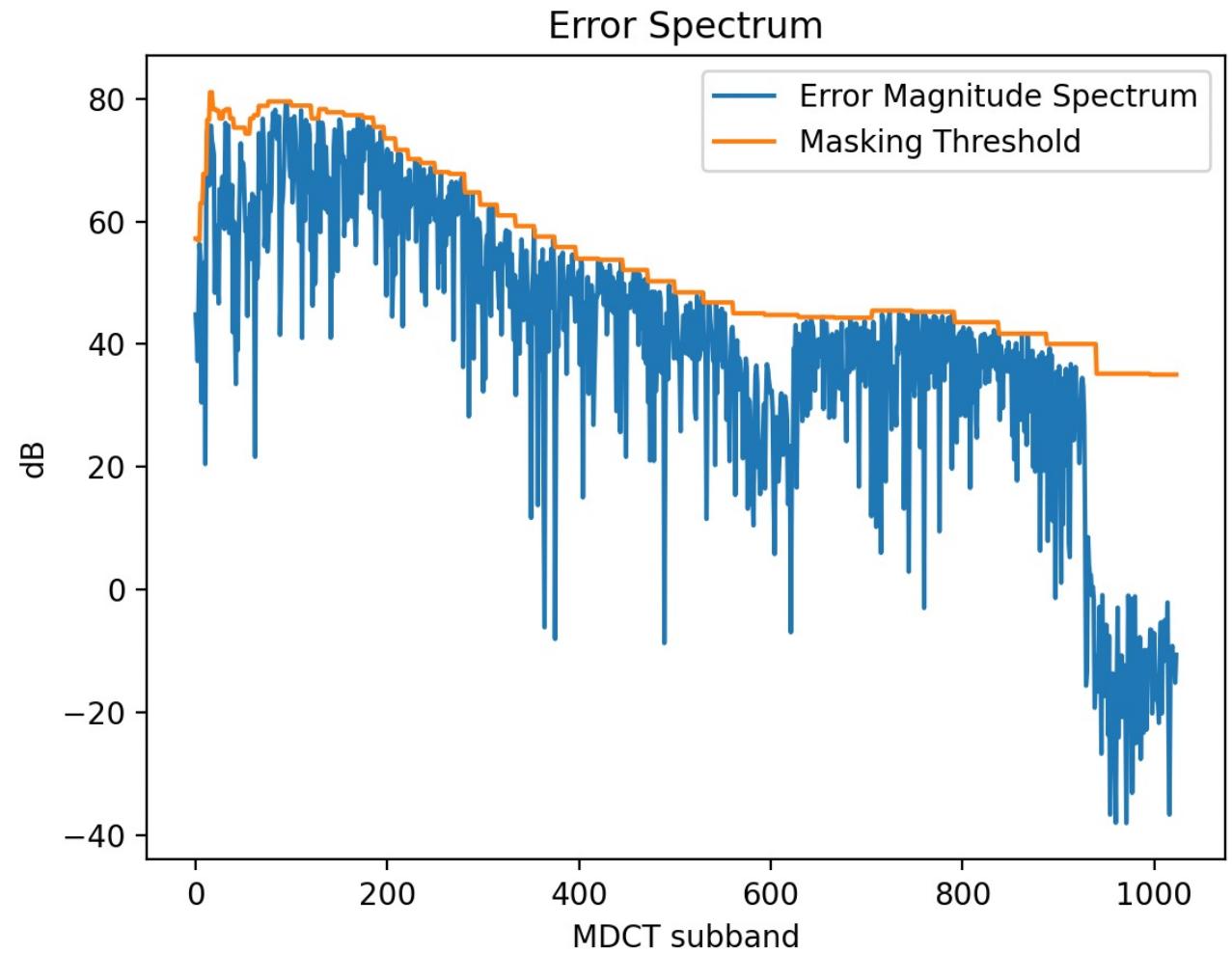
# Female Voice Quantization Error

- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



# Female Voice Quantization Error

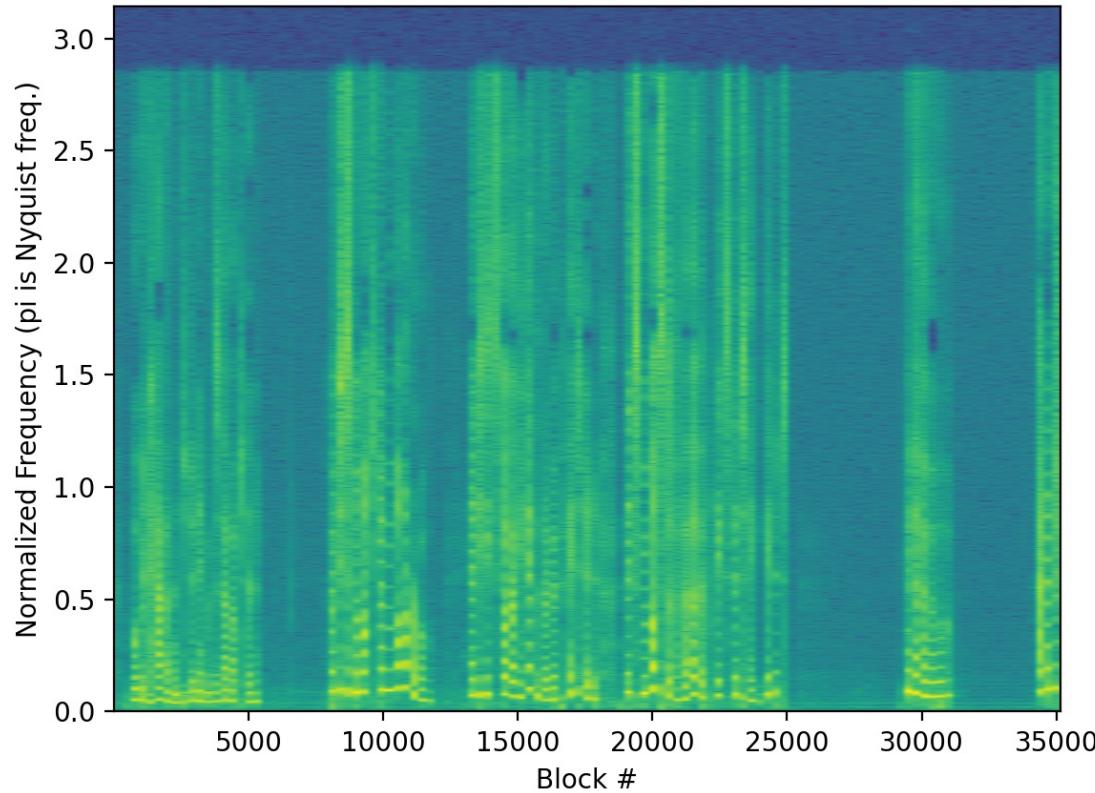
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



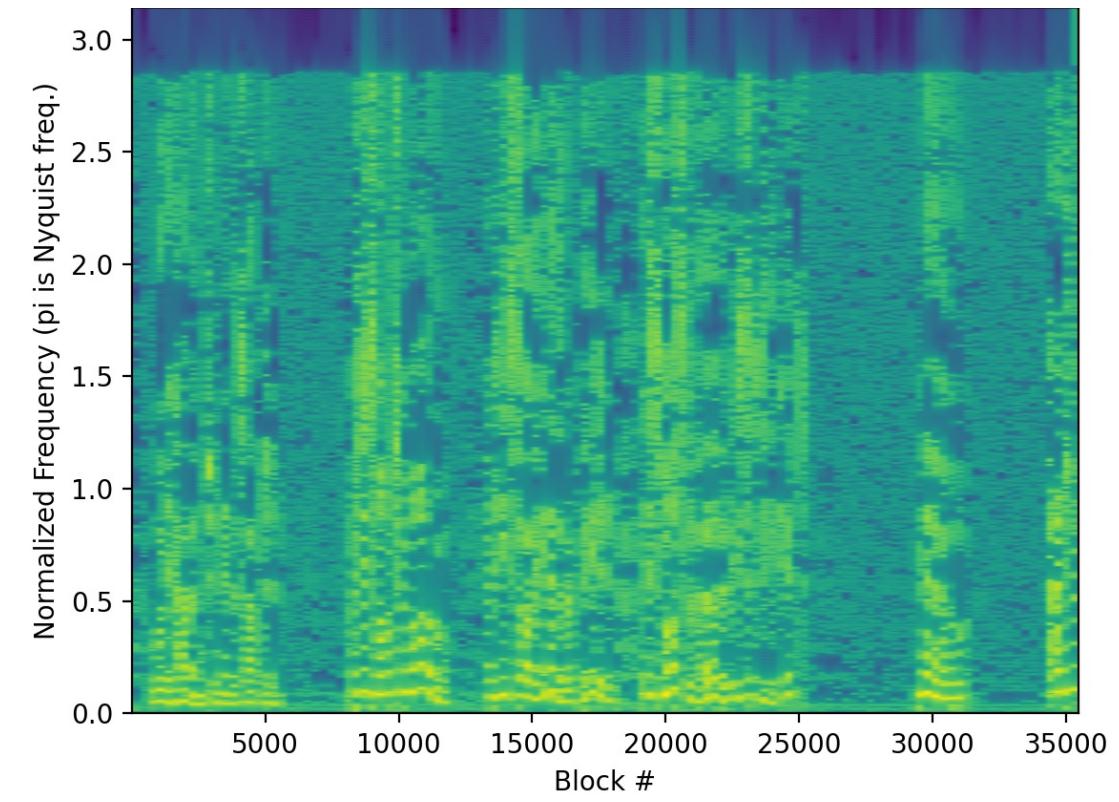
# Female Voice Quantization Error

60 Quality, 1024 MDCT Subbands

Original



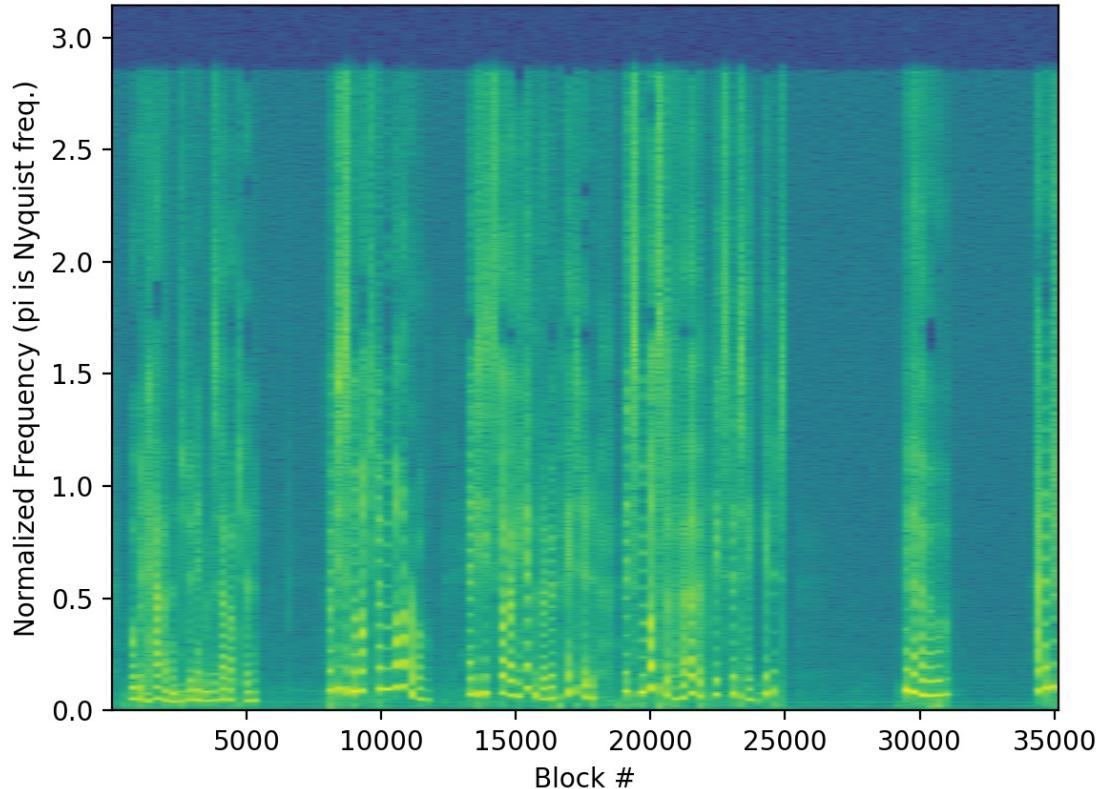
Reconstructed



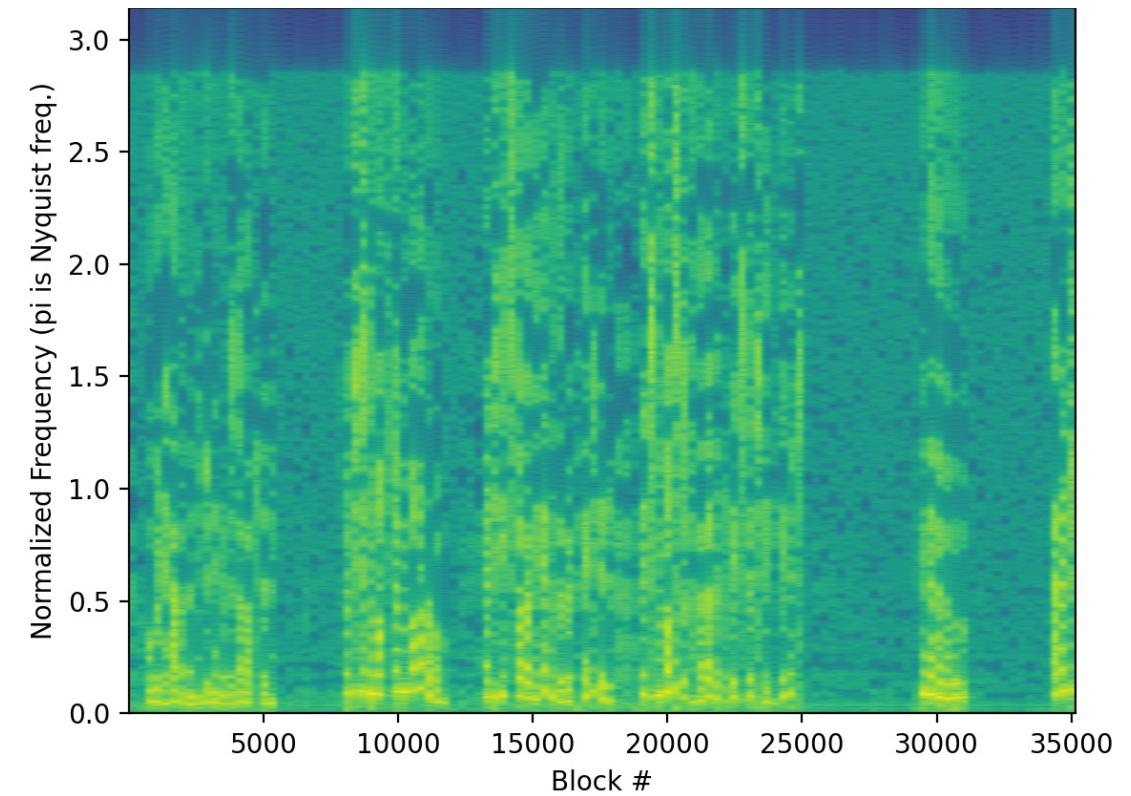
# Female Voice Quantization Error

60 Quality, 256 MDCT Subbands

Original



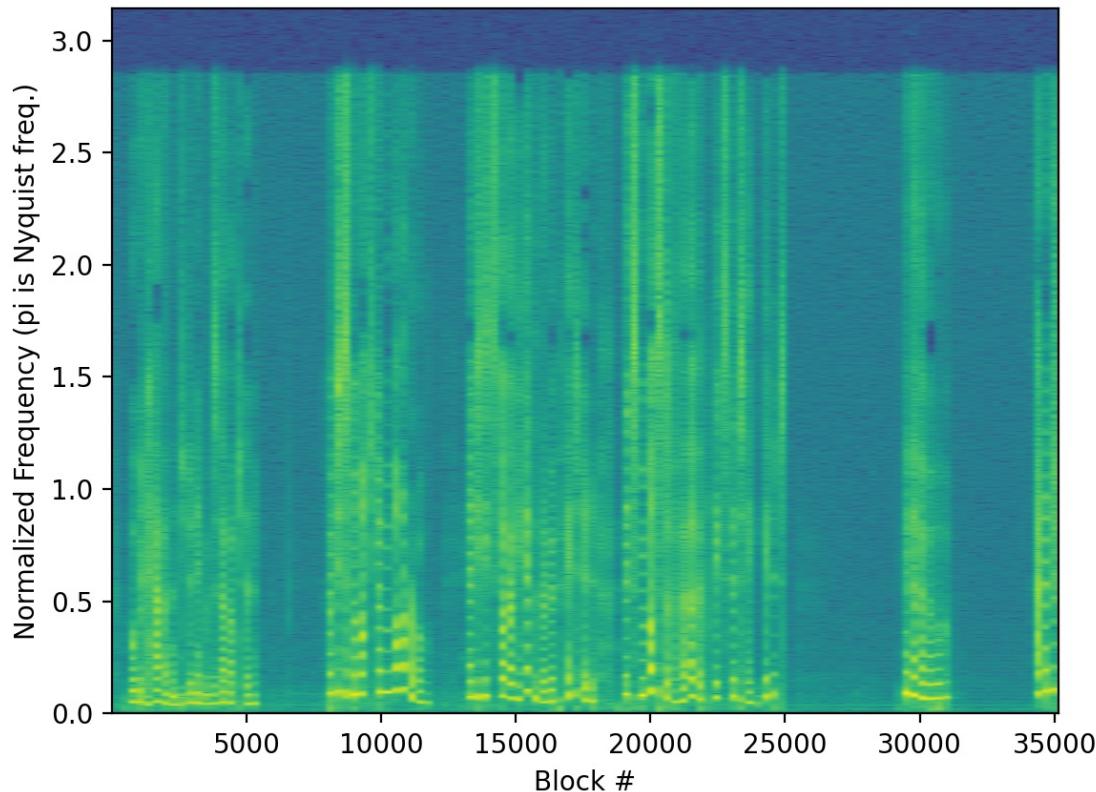
Reconstructed



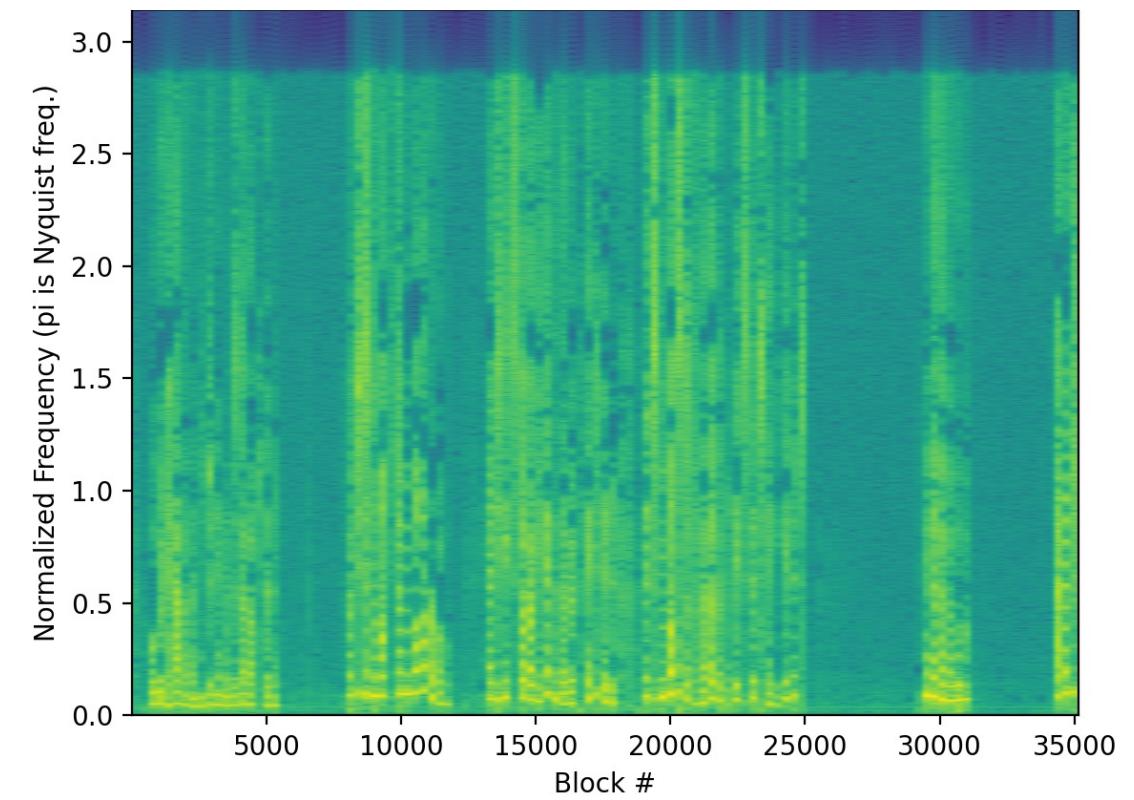
# Female Voice Quantization Error

100 Quality, 256 MDCT Subbands

Original



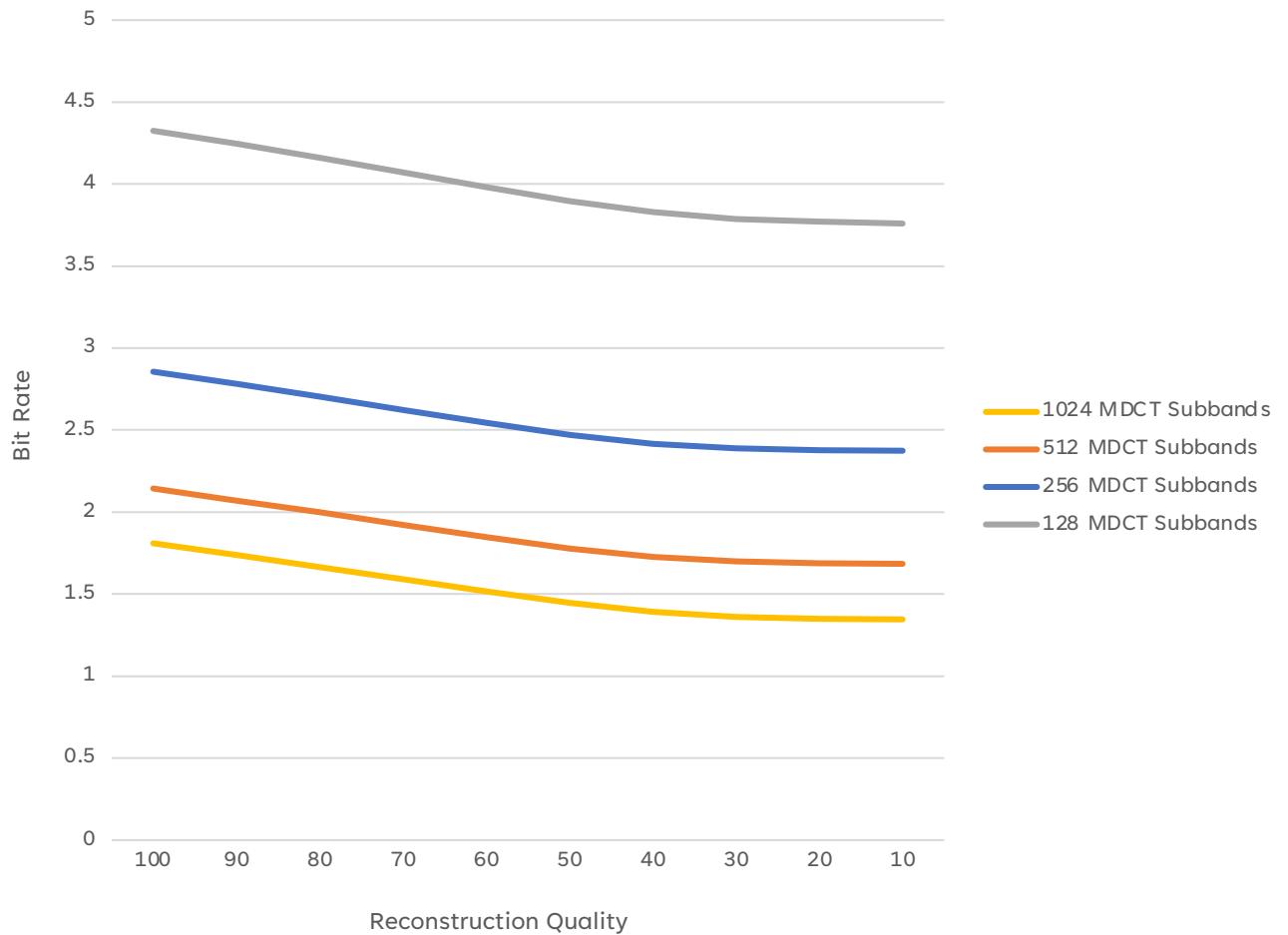
Reconstructed



# Bit Rate vs Reconstruction Quality

Female Voice Input

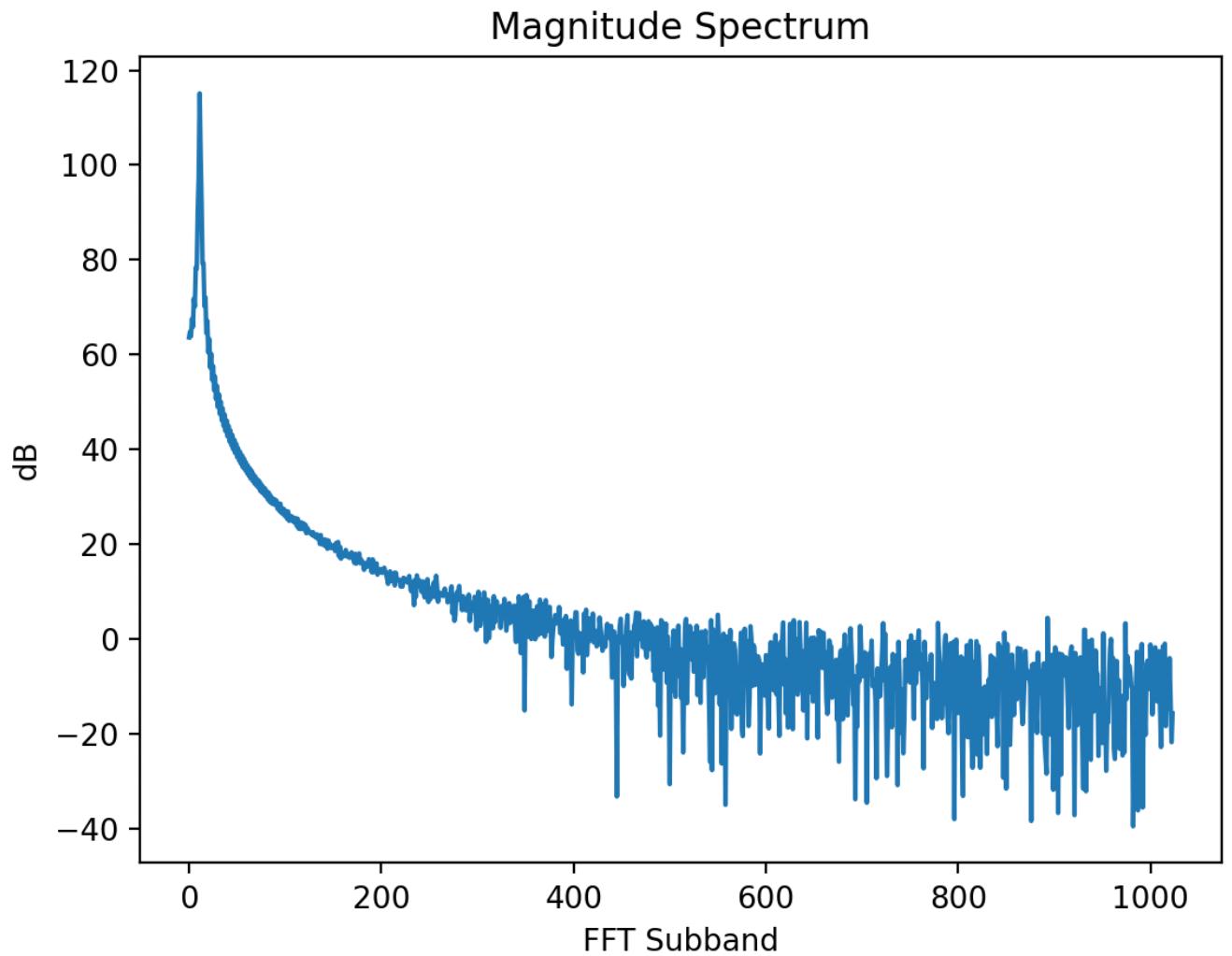
- Decreasing MDCT Subbands increases distortion and bit rate
- Decreasing reconstruction quality (increasing the masking threshold) decreases bit rate but increases distortion



# Appendix

## vi. Constant Tone Quantization Error

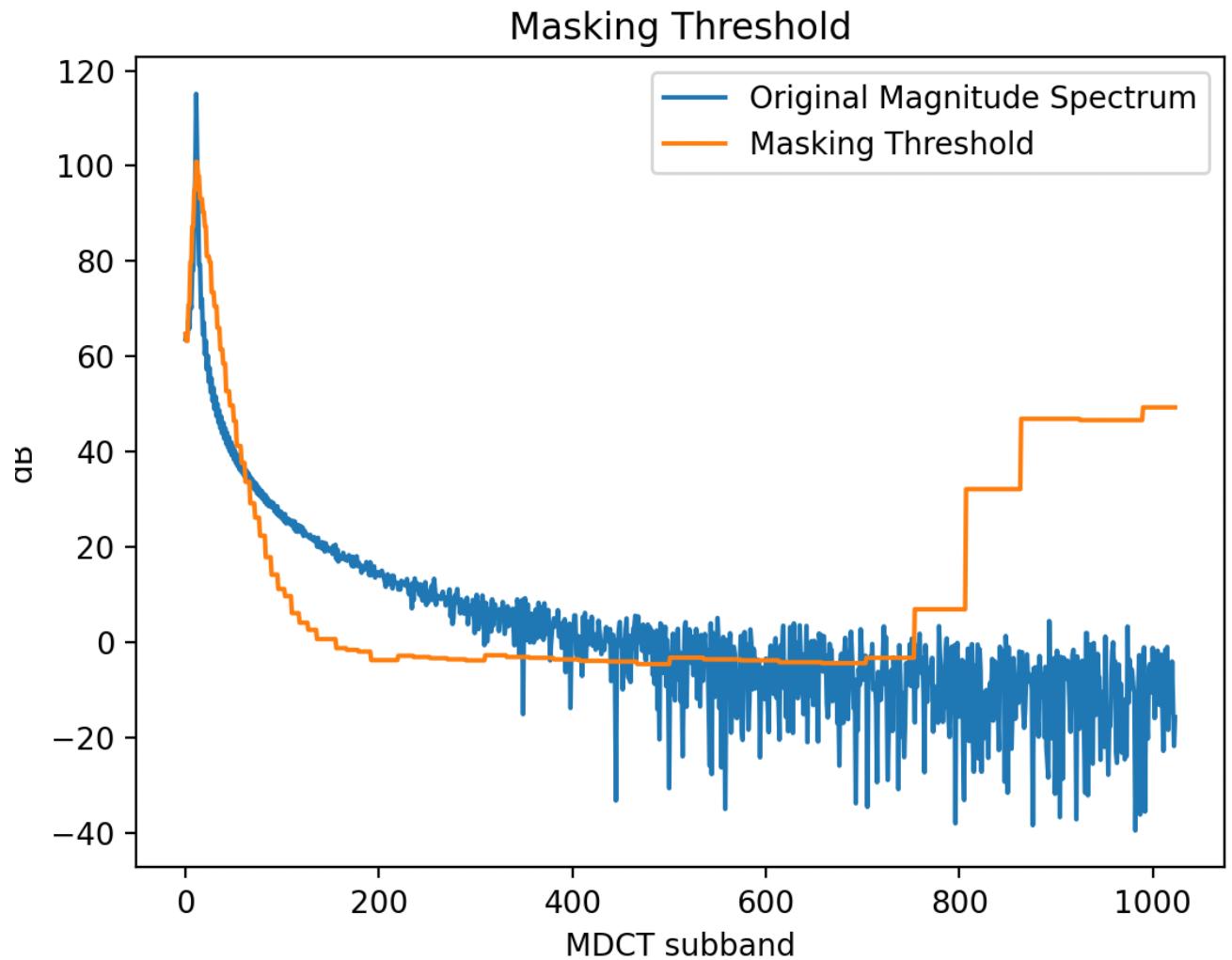
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



# Appendix

## vi. Constant Tone Quantization Error

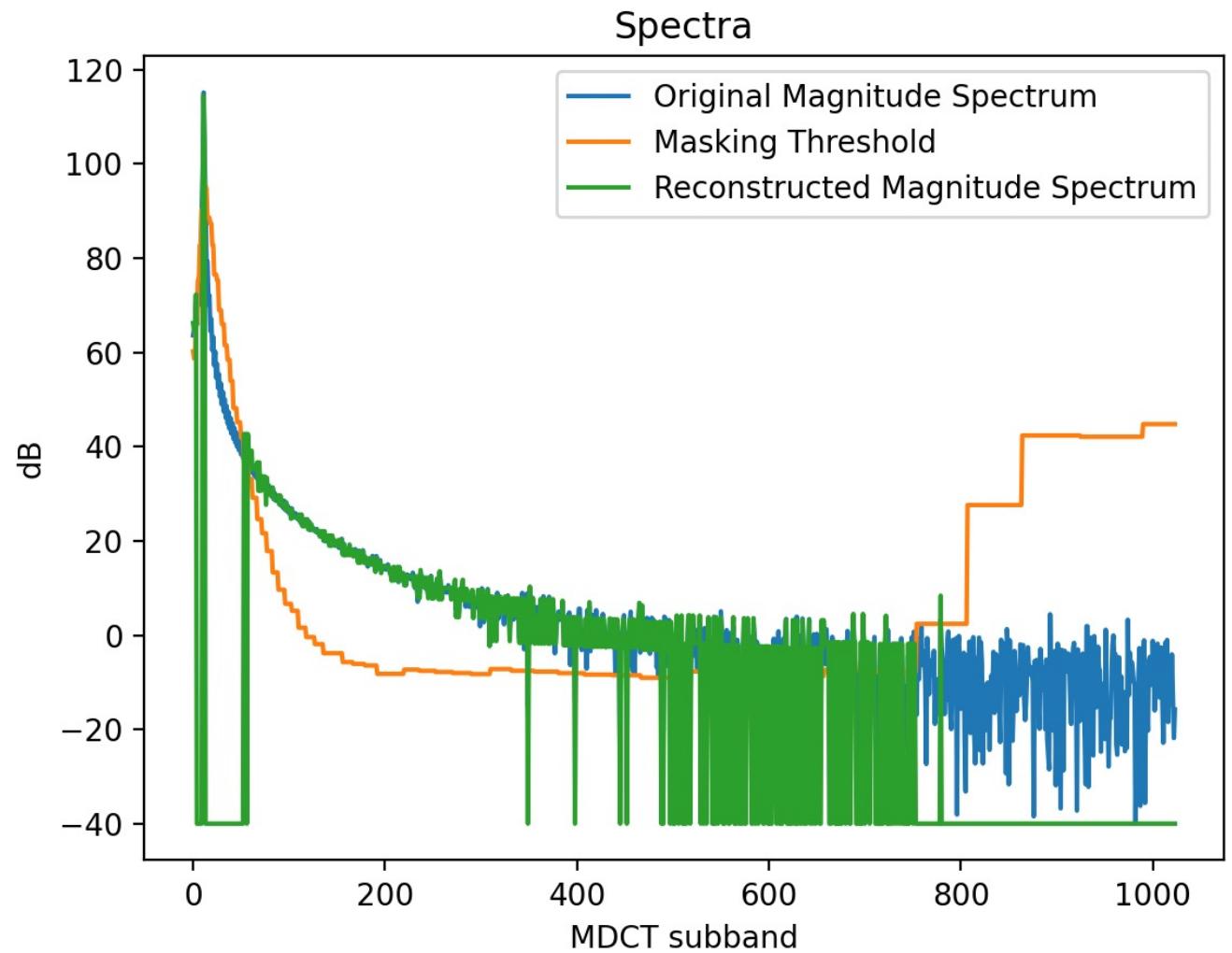
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



# Appendix

## vi. Constant Tone Quantization Error

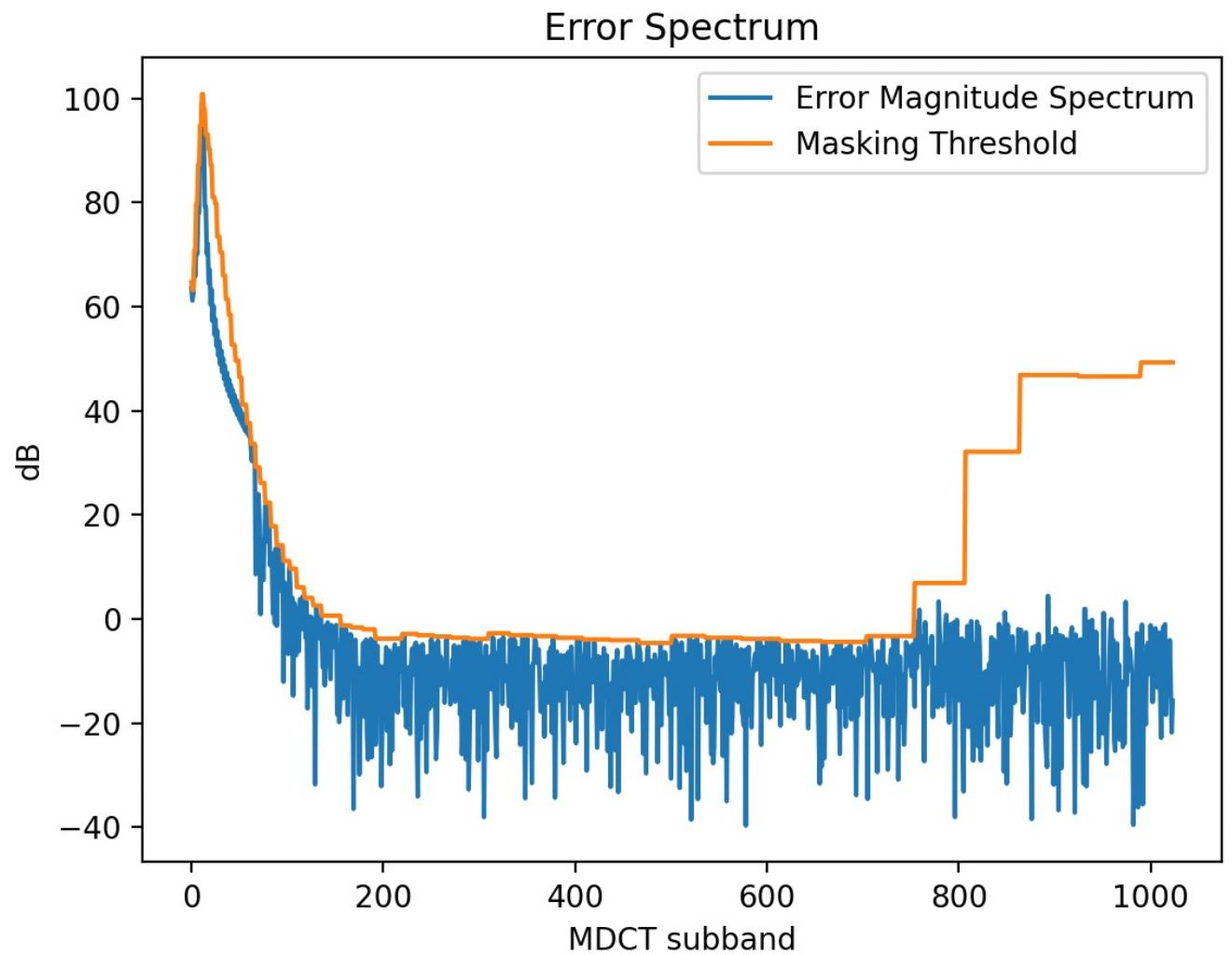
- FFT
- Masking Threshold
- **Reconstructed Signal**
- Quantization Error



# Appendix

## vi. Constant Tone Quantization Error

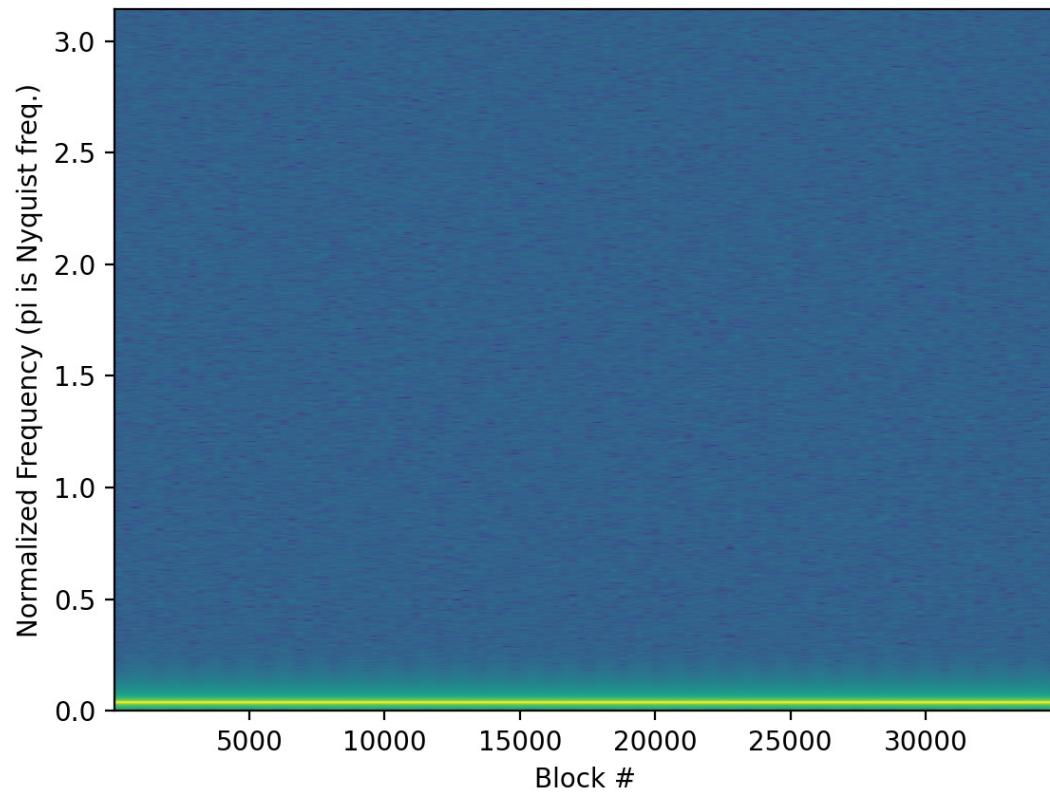
- FFT
- Masking Threshold
- Reconstructed Signal
- Quantization Error



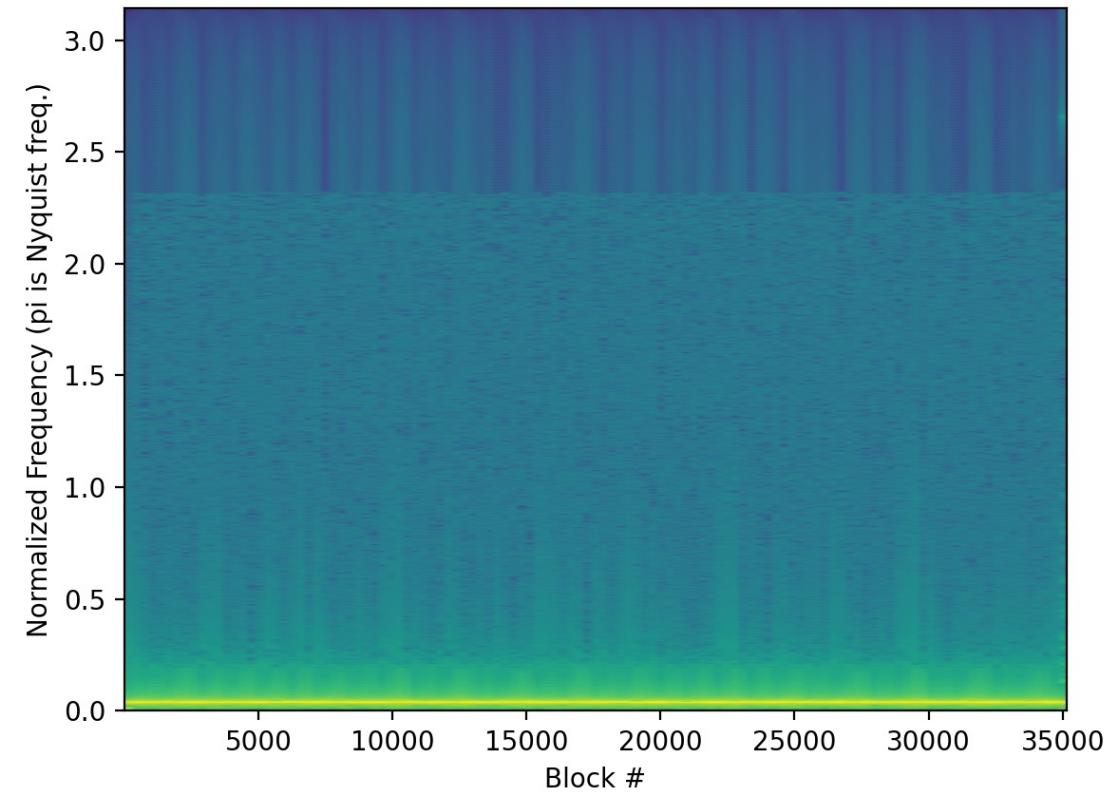
# Appendix

vii. Constant Tone, 60 Quality, 1024 MDCT Bands

Original



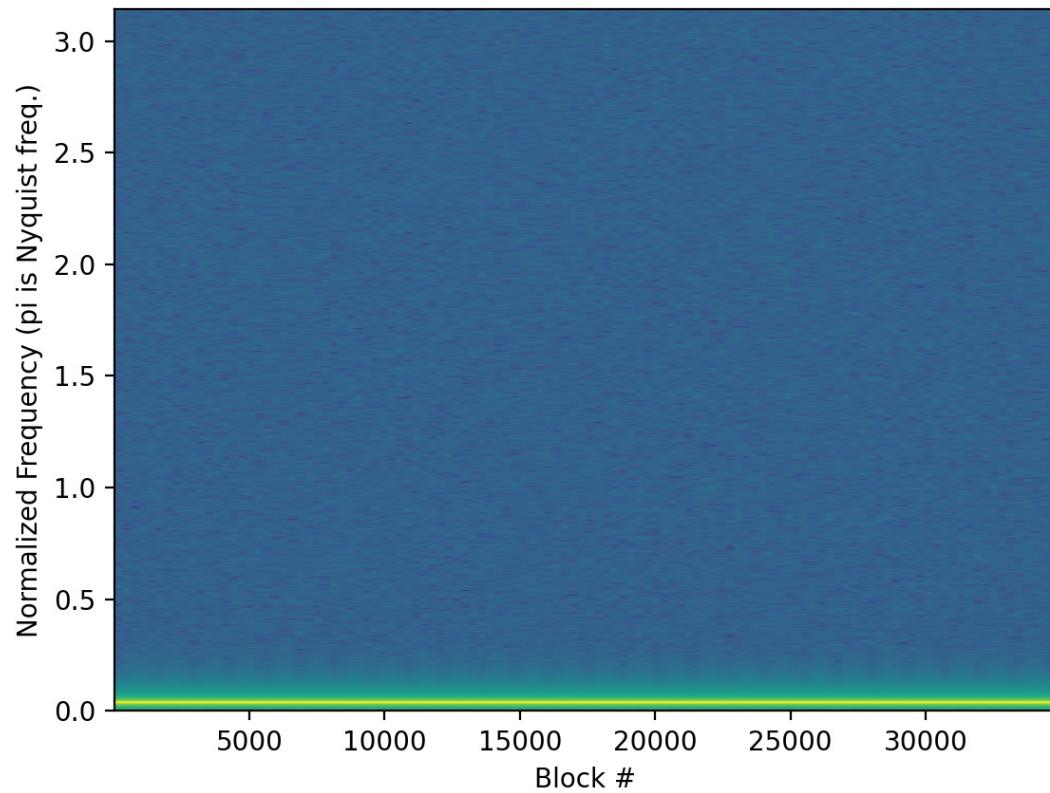
Reconstructed



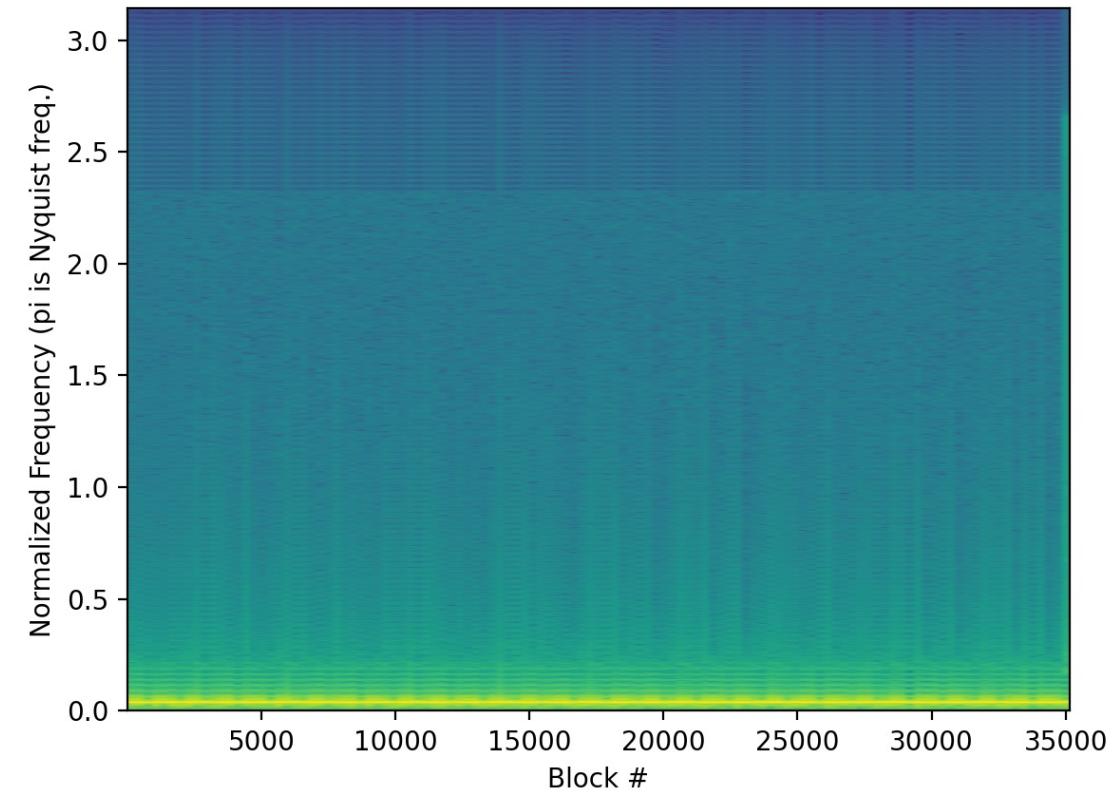
# Appendix

vii. Constant Tone, 60 Quality, 256 MDCT Bands

Original



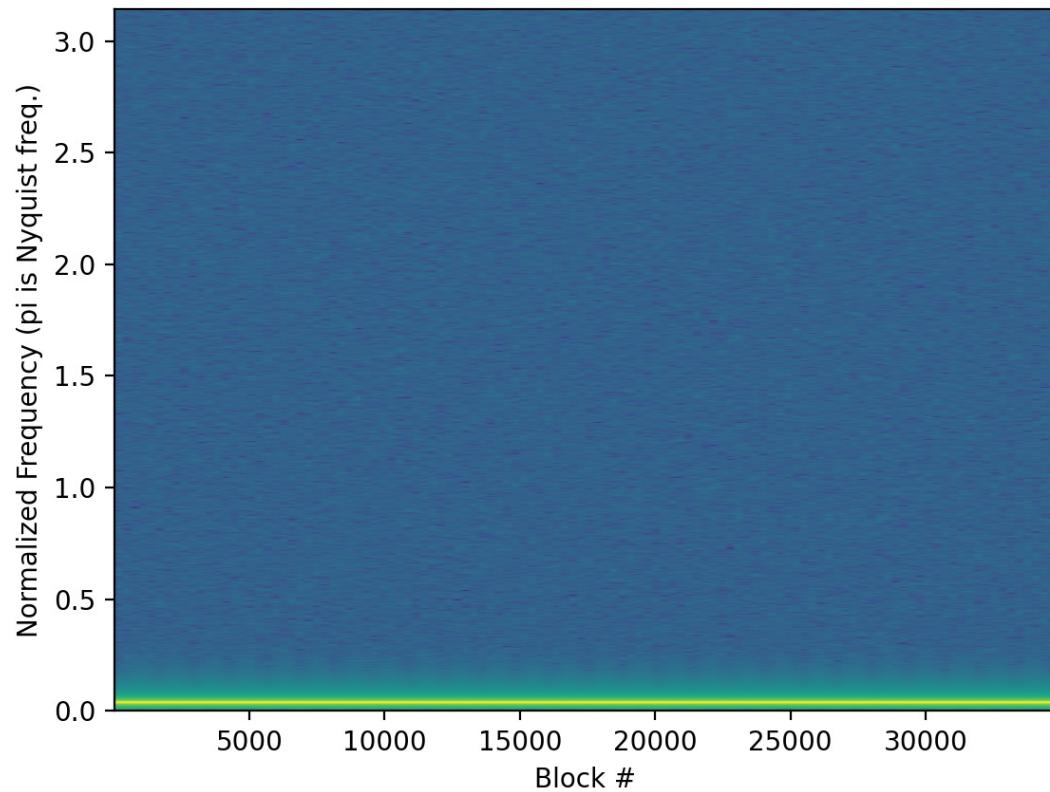
Reconstructed



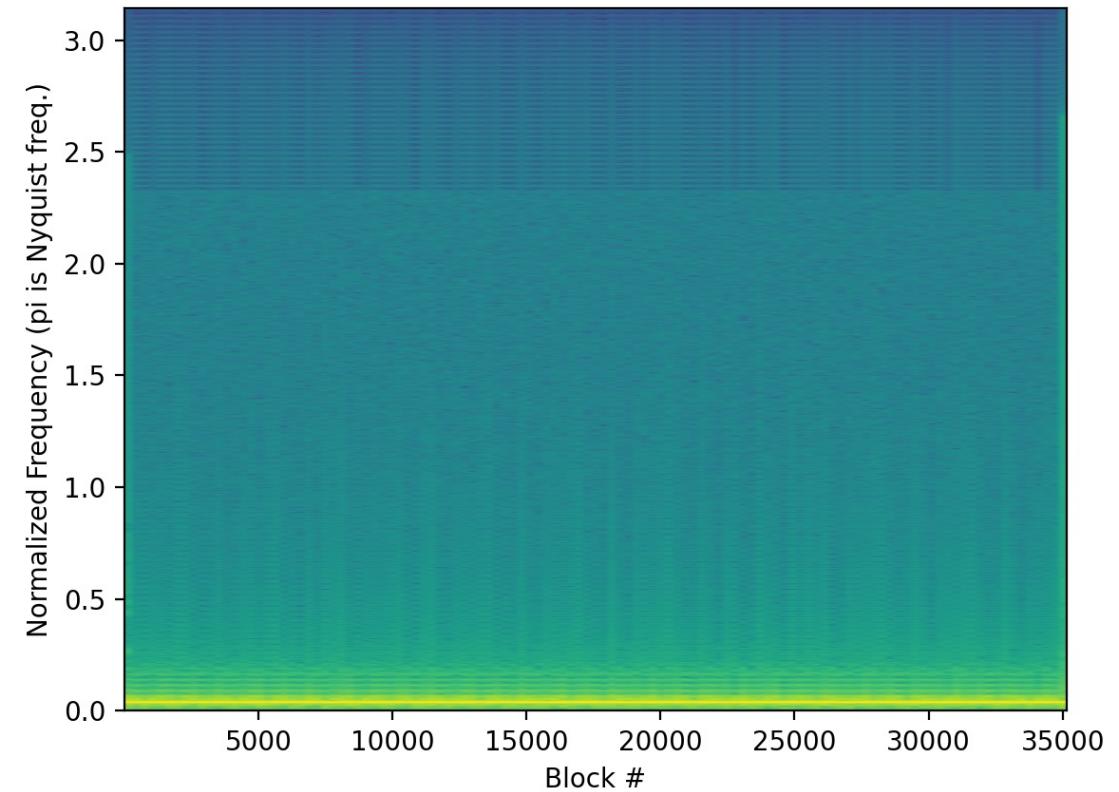
# Appendix

vii. Constant Tone, 100 Quality, 256 MDCT Bands

Original



Reconstructed



# Appendix

## viii. Constant Tone Bit Rate vs Reconstruction Quality

- Decreasing MDCT Subbands increases distortion and bit rate
- Decreasing reconstruction quality (increasing the masking threshold) decreases bit rate but increases distortion

