

Block Switching with Psychoacoustic Modifications

AJ Ferrick
Gabriele Carotti-Sha

Overview

- Codec extensions and discussion
- Objective results
- Subjective tests
- Future results

Features

Block Switching

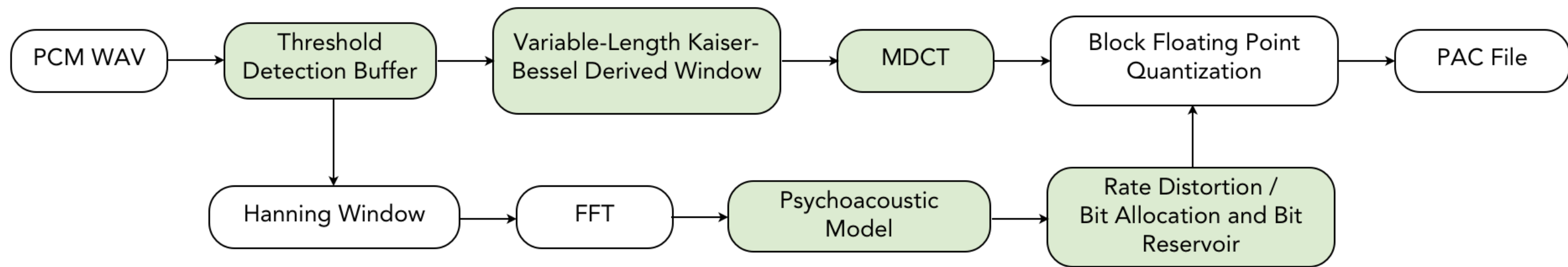
(because pre-echo is bad and it should make you feel bad)

Masker Tuning: Decimation, Frequency Band

Naive Bit Reservoir

Fairly Deep Refactor

Codec Structure

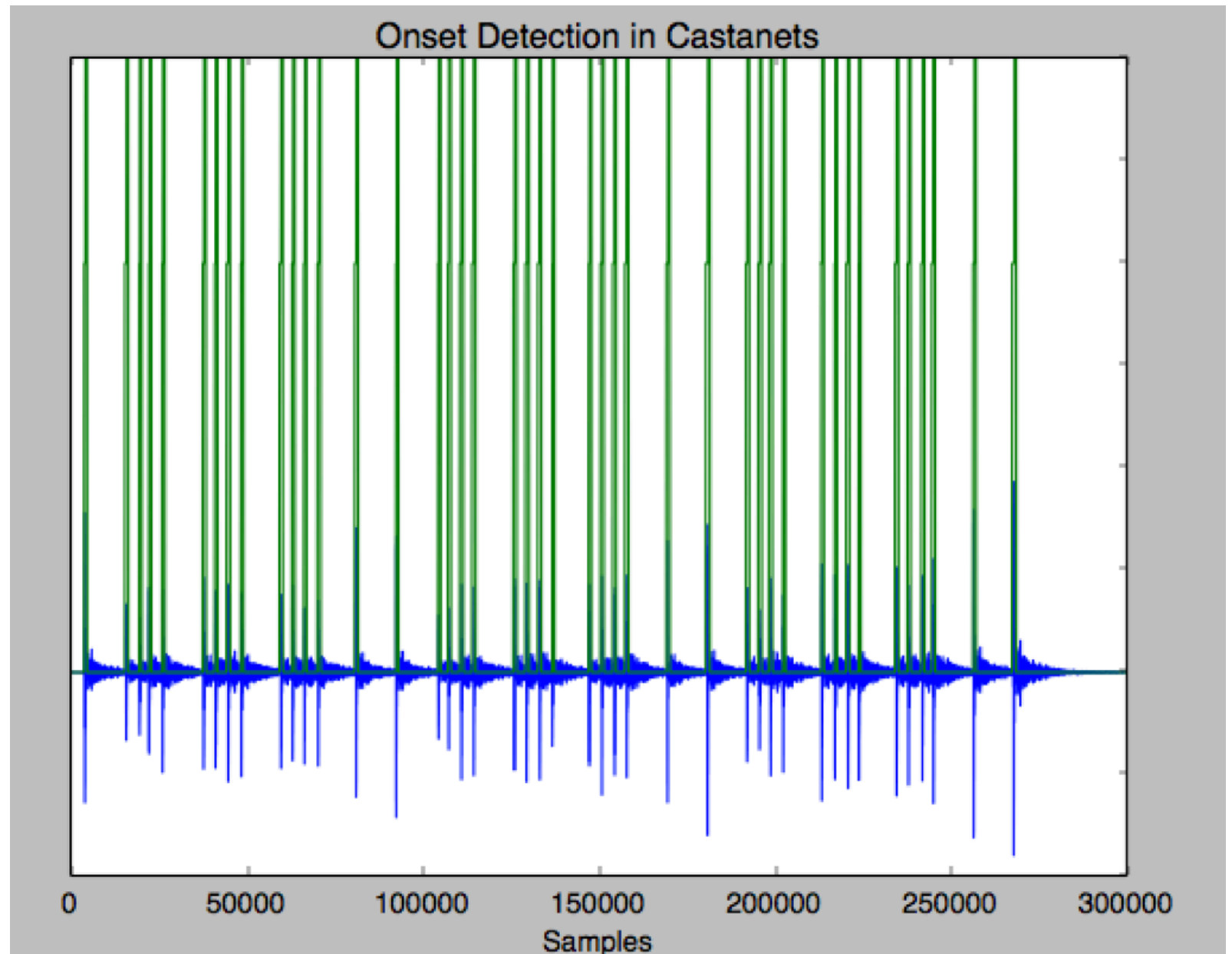


Block Switching

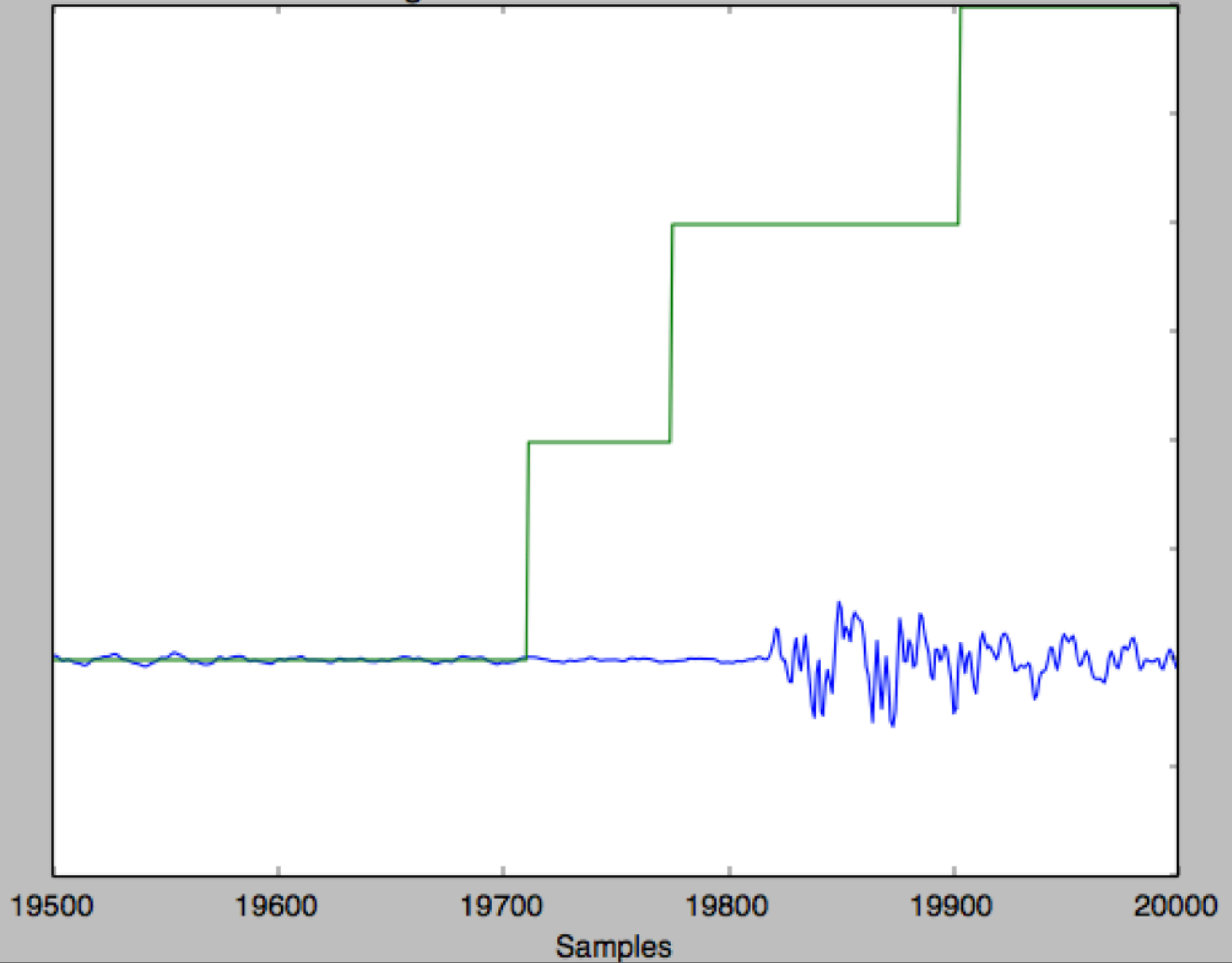
Onset Detection

Onset detection on high frequencies using a weighted energy measure.

Designed to operate on a buffer of samples and return a boolean result.

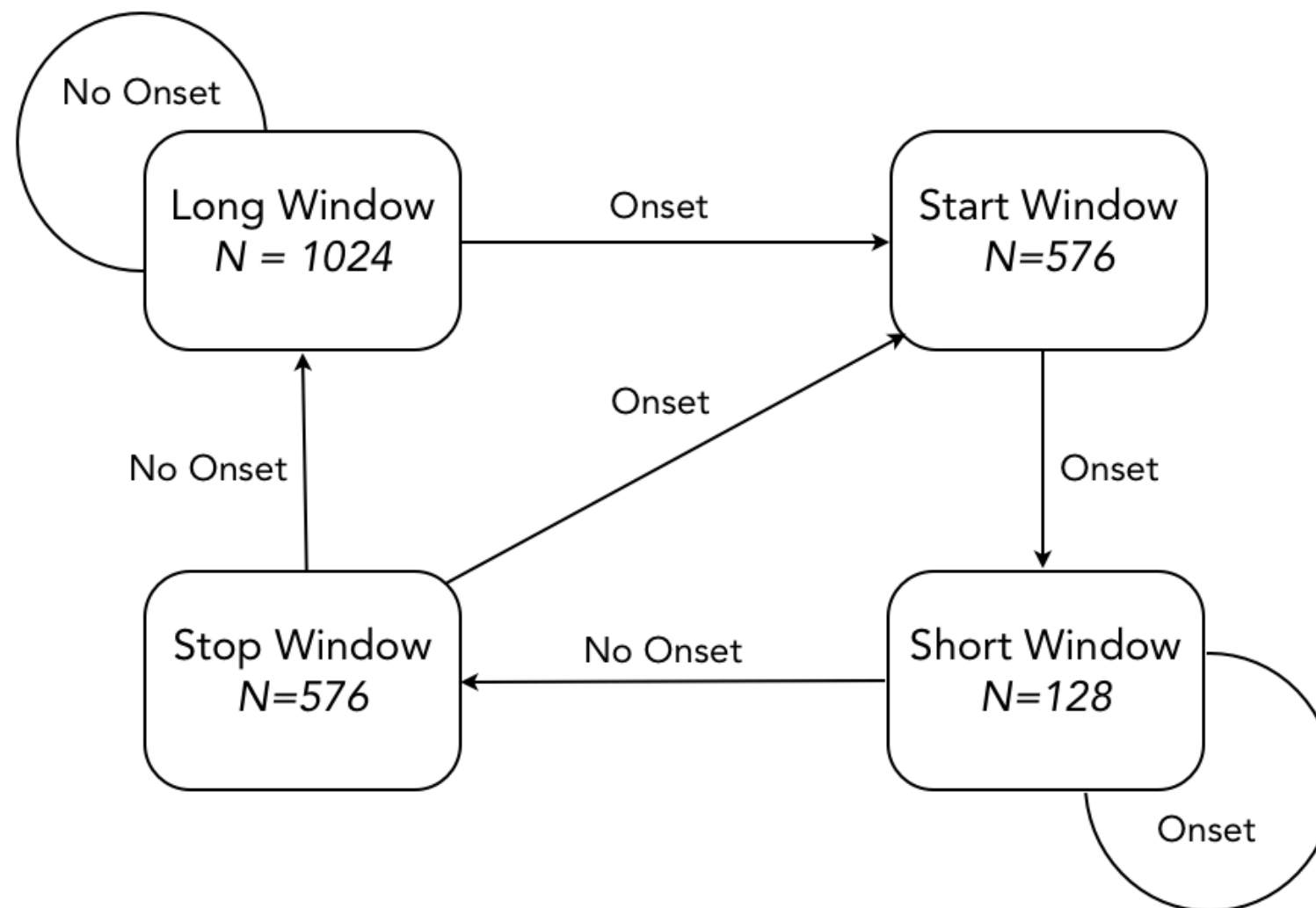


Single Castanet Onset Detection



Block Control

State machine controls the block size the codec receives when the compressor pops the next data block.



Usage?

All this built into a queue data structure that can perform onset detection on the contents of its buffer.

Leads to modular, testable code.

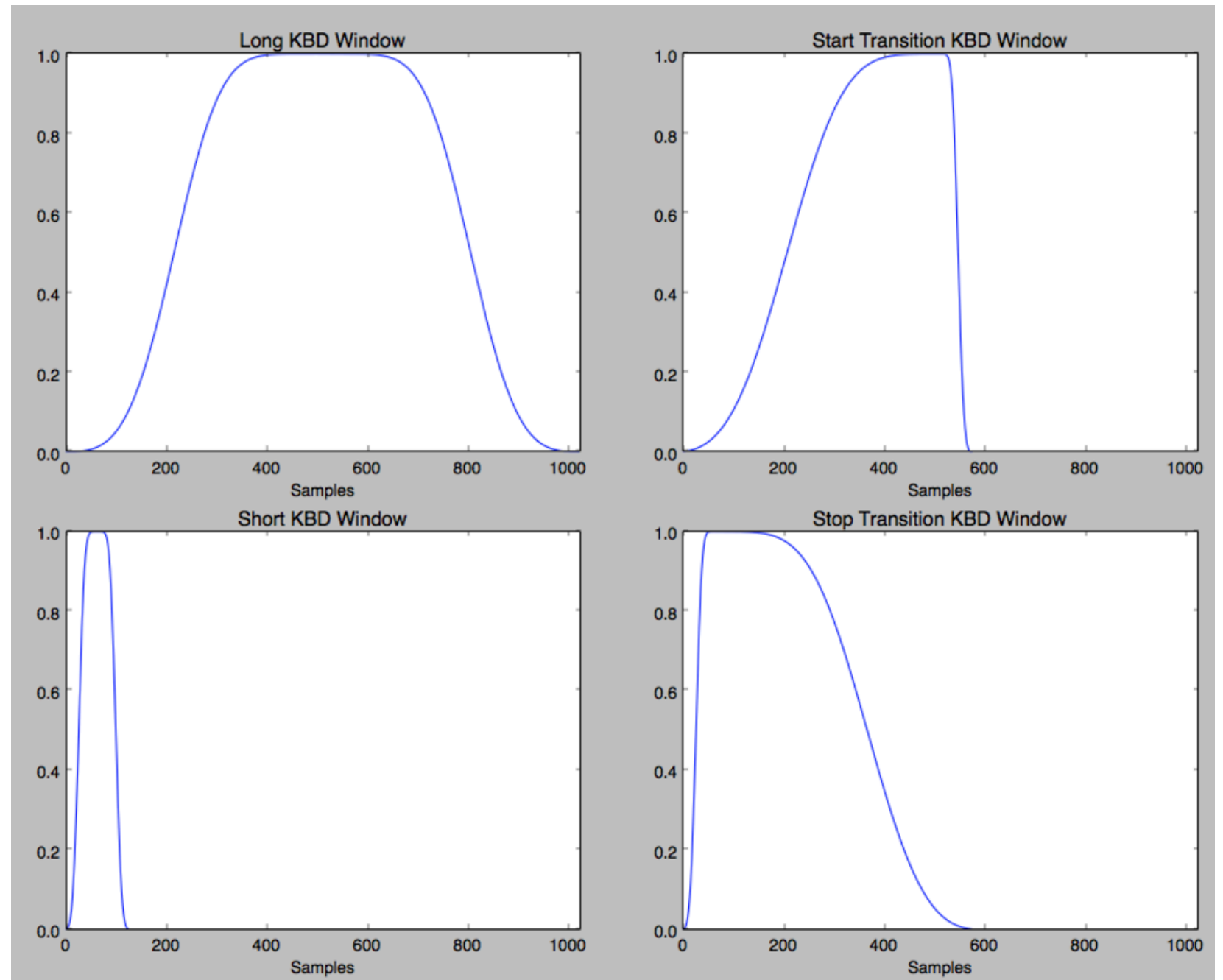
```
7  from codec.transbuffer import TransientBuffer
8
9  buf = TransientBuffer('samples/trmp/trumpet.wav')
10
```

```
25  done = False
26  while not done:
27      x, win = buf.next()
28      if len(x) == 0:
29          done = True
30      else:
```

KBD Window & MDCT

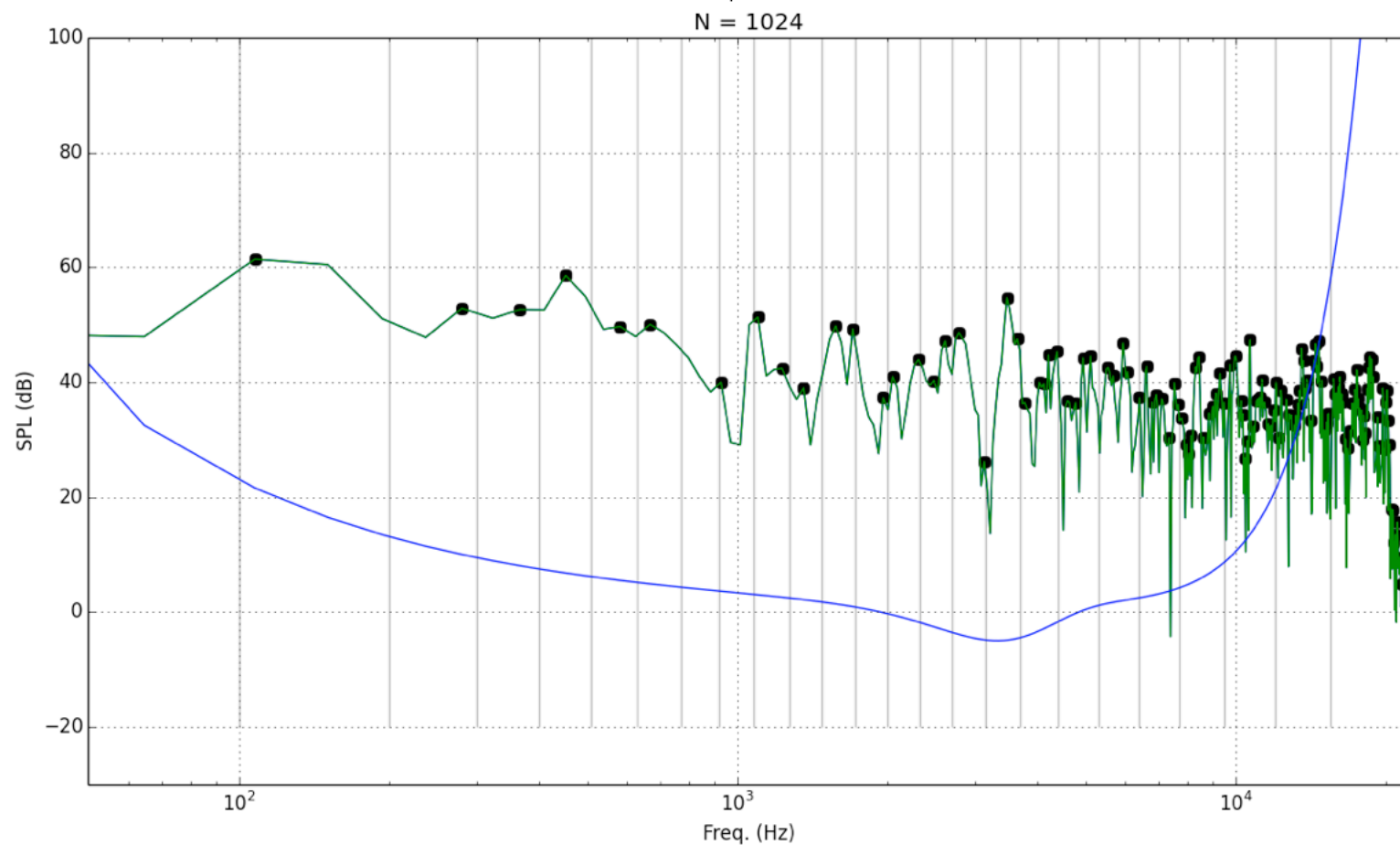
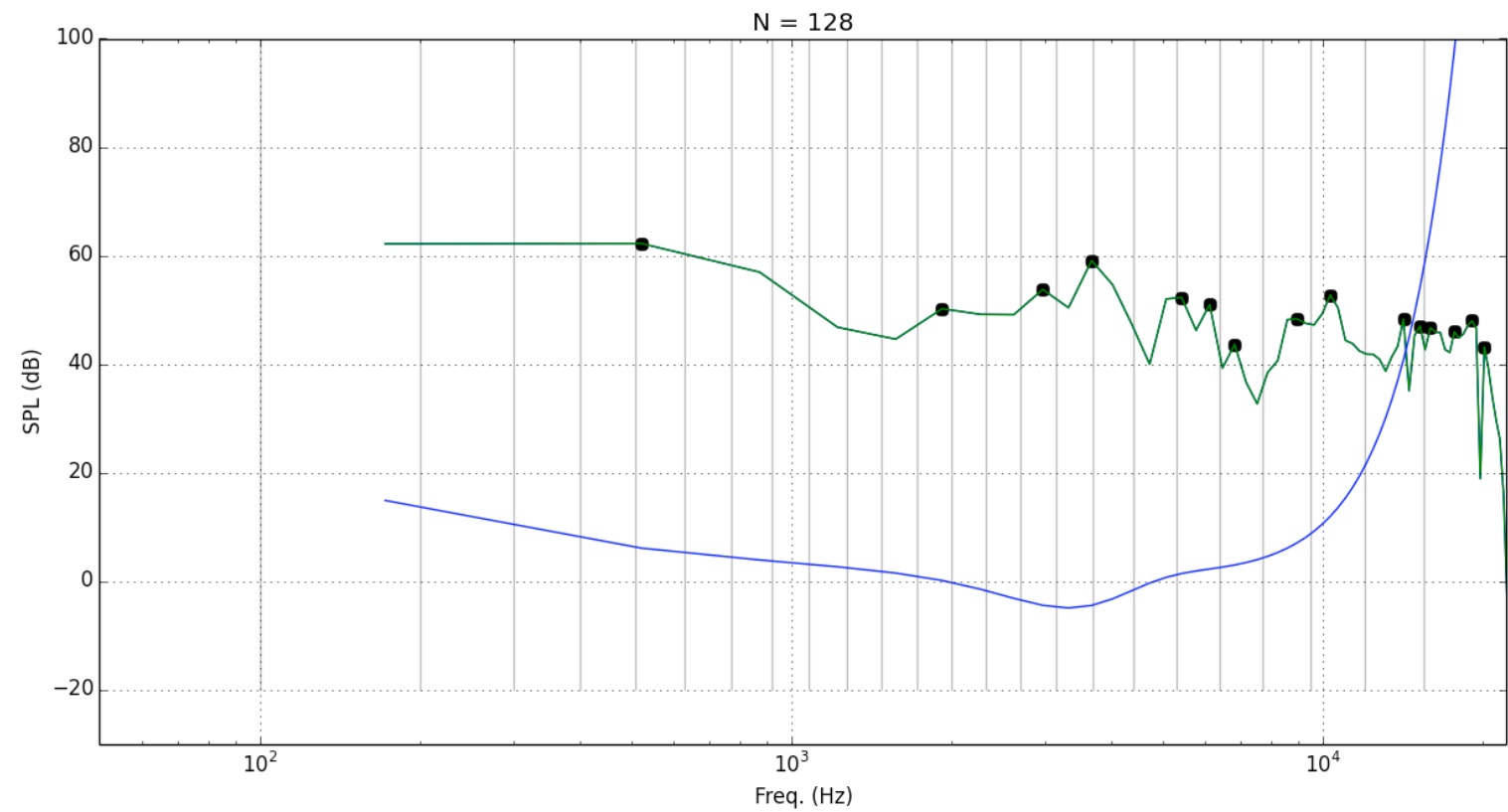
Windows must preserve the perfect reconstruction conditions of MDCT *and* operate on different block sizes.

This also means the MDCT kernel must change so it is offset by half the *right-hand window*.

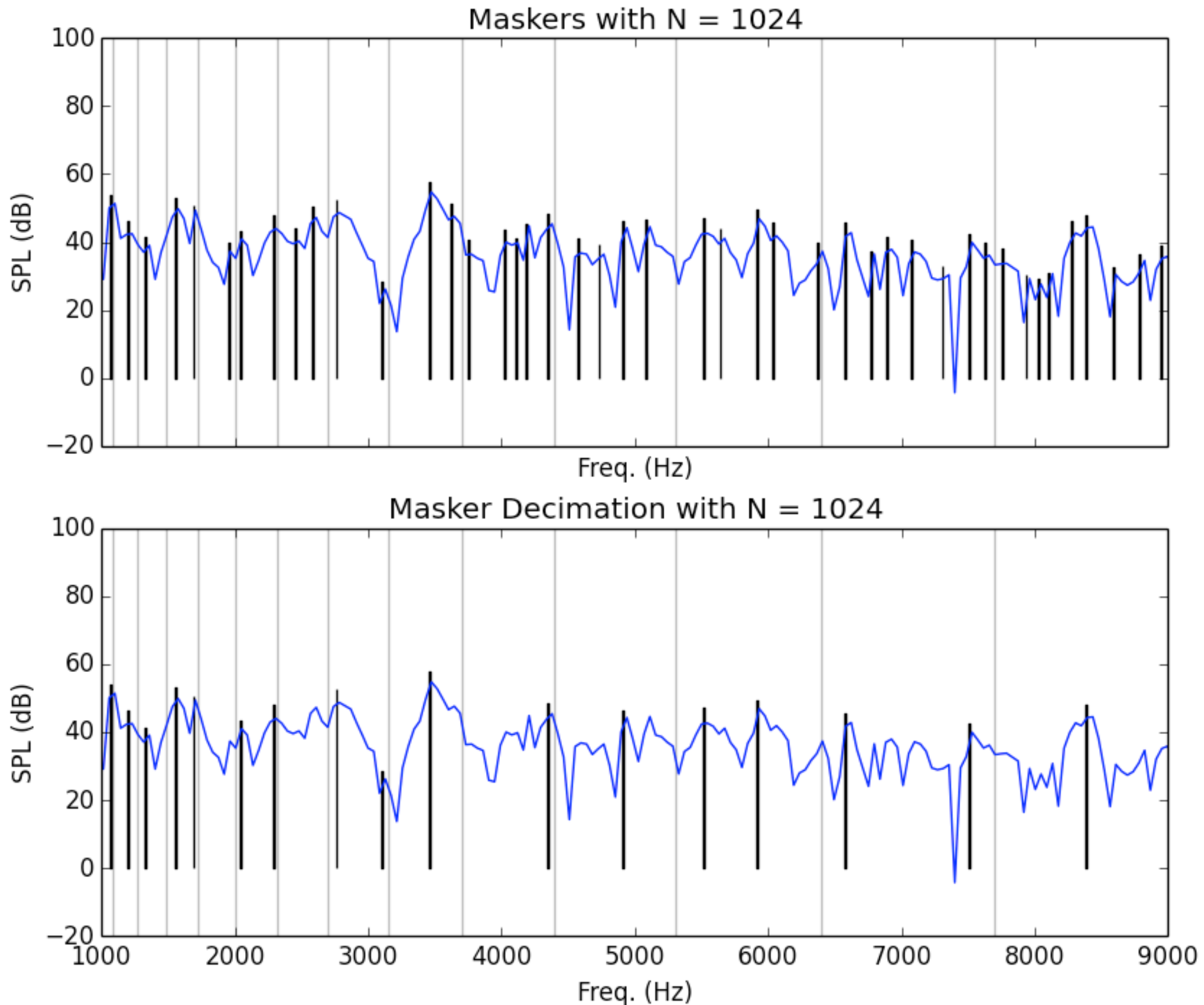


Psychoacoustic Model

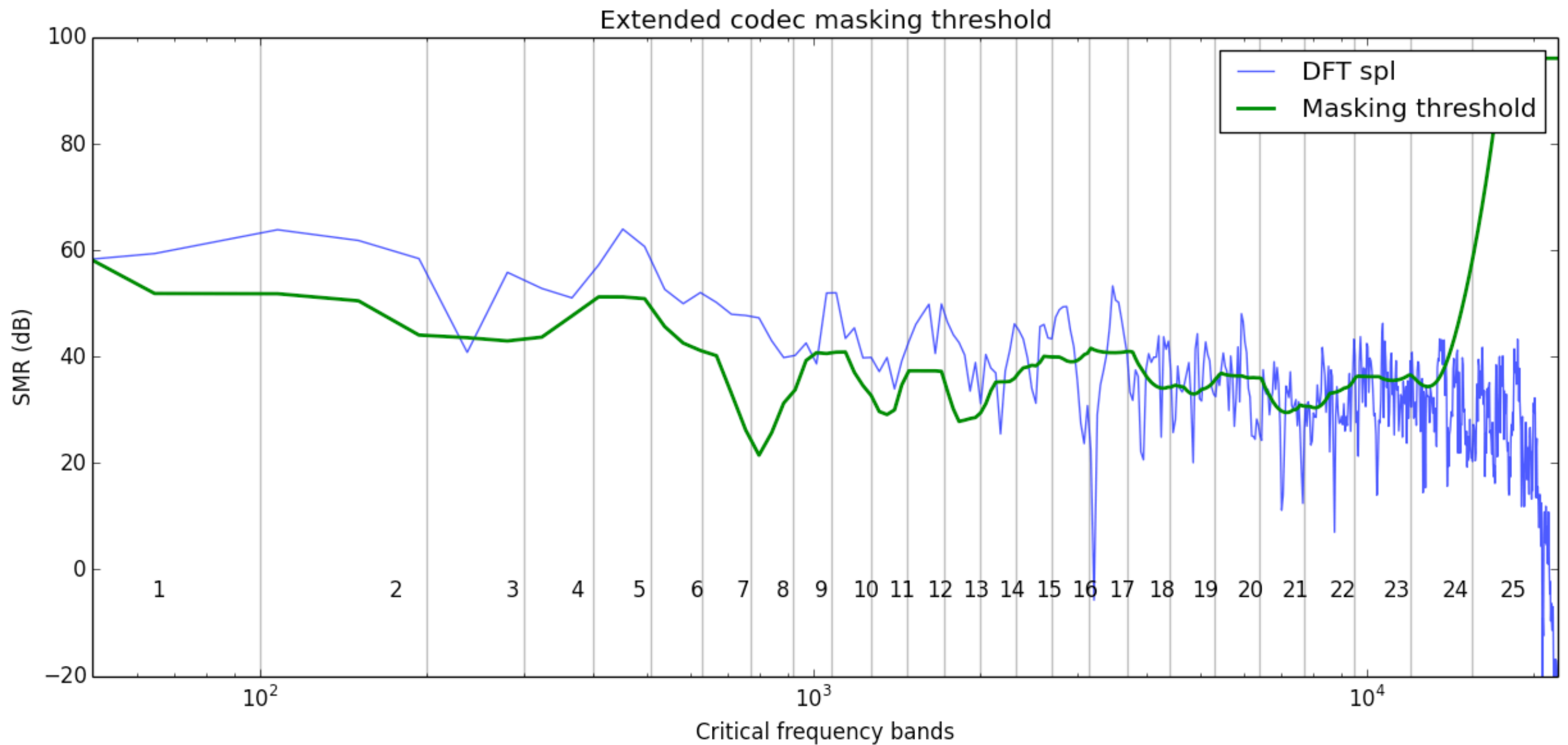
Peak Detection



Masker Decimation



Masking Threshold



Naive Bit Reservoir

Confronts loss of bit resolution for short blocks

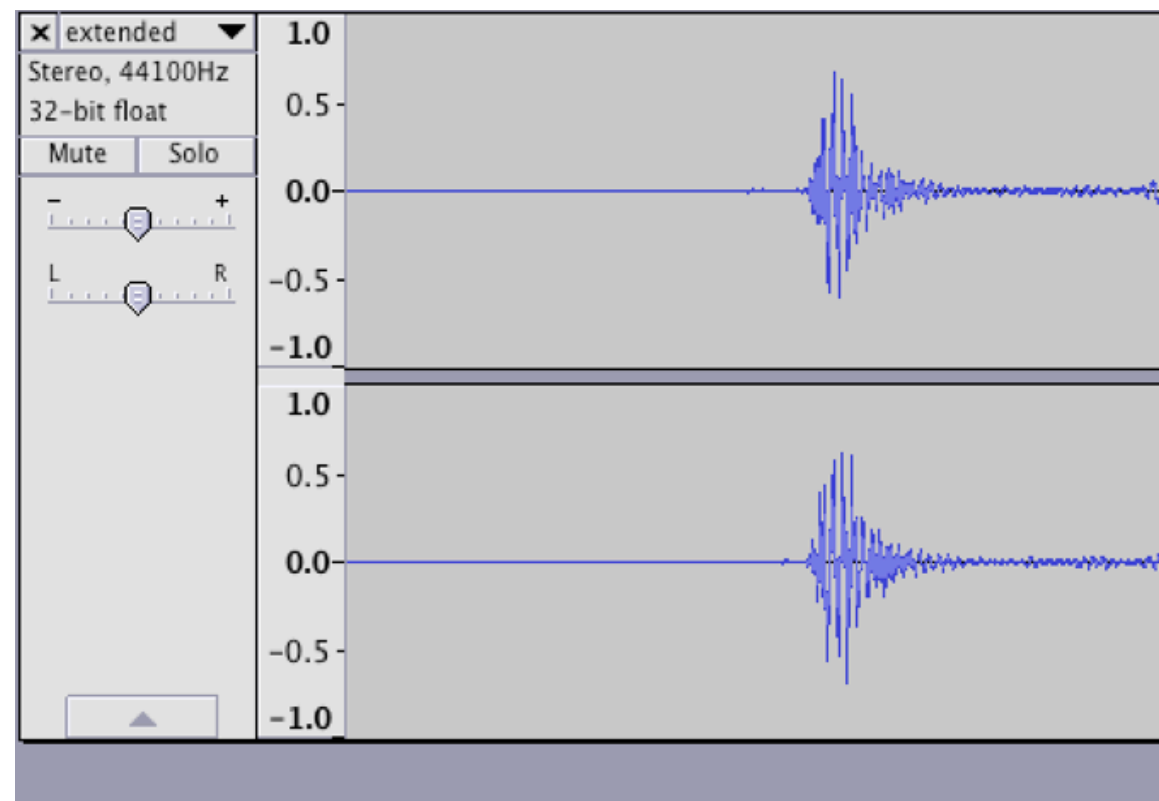
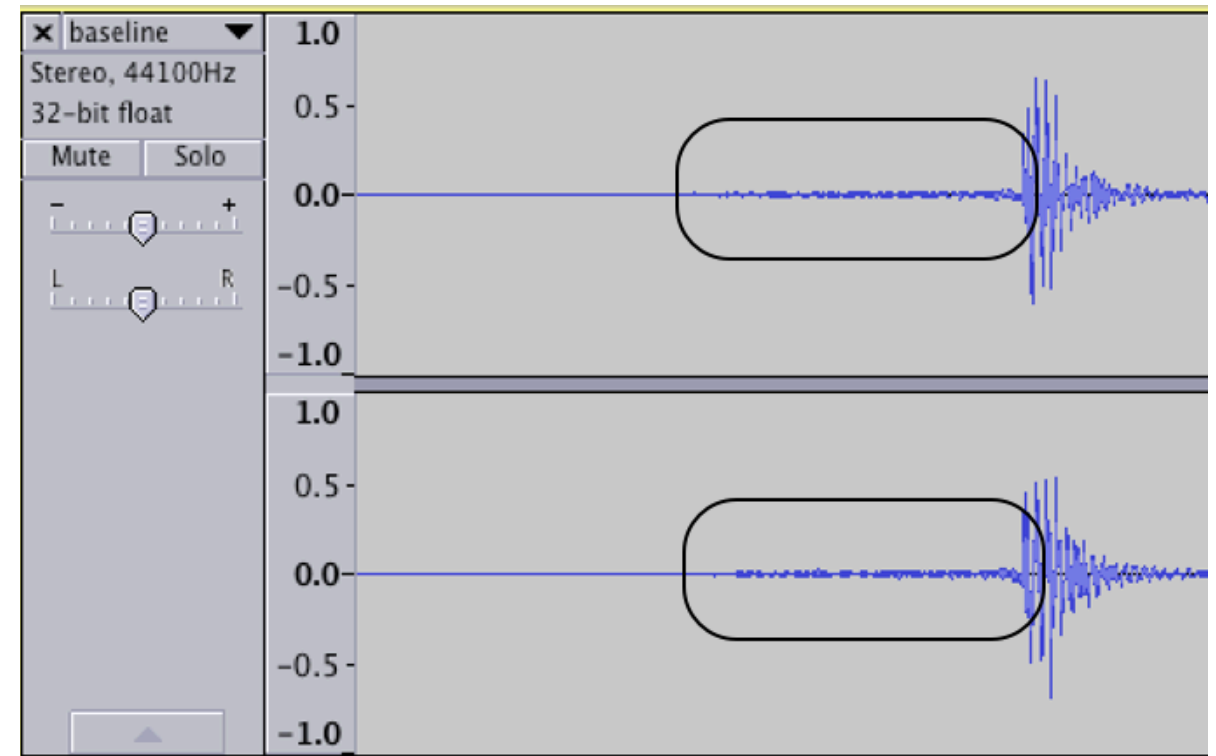
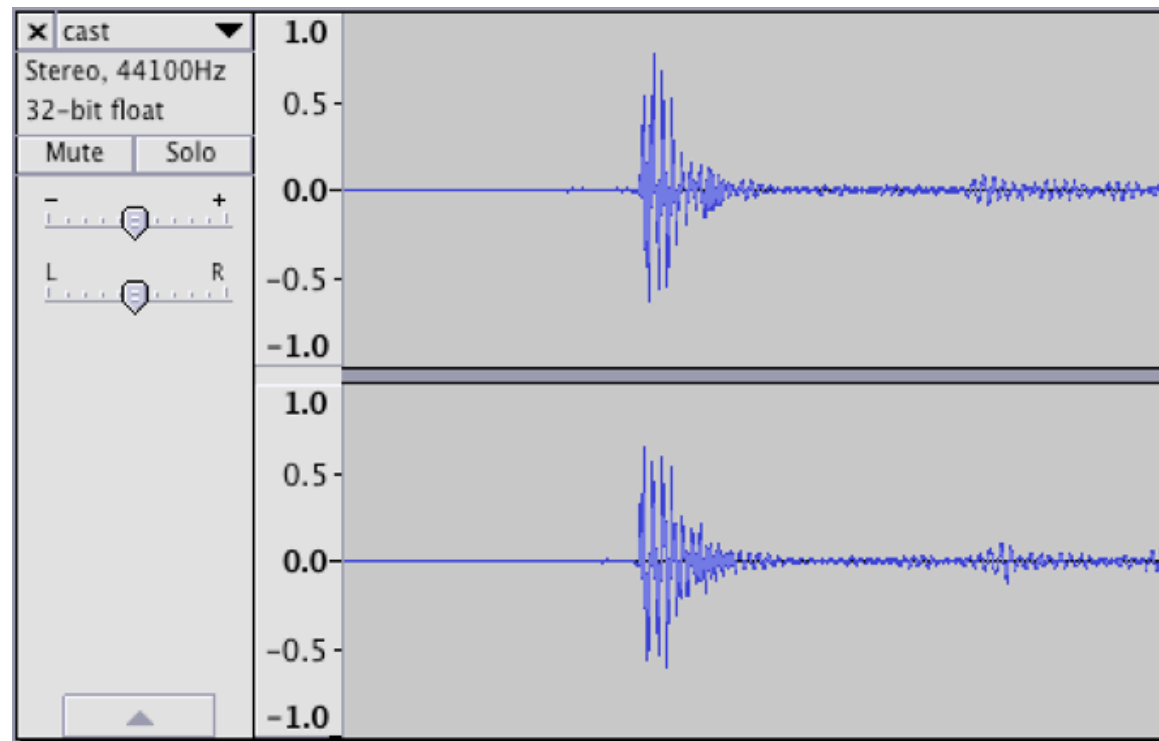
Updated by bitAlloc encoding

Reduces quantization noise

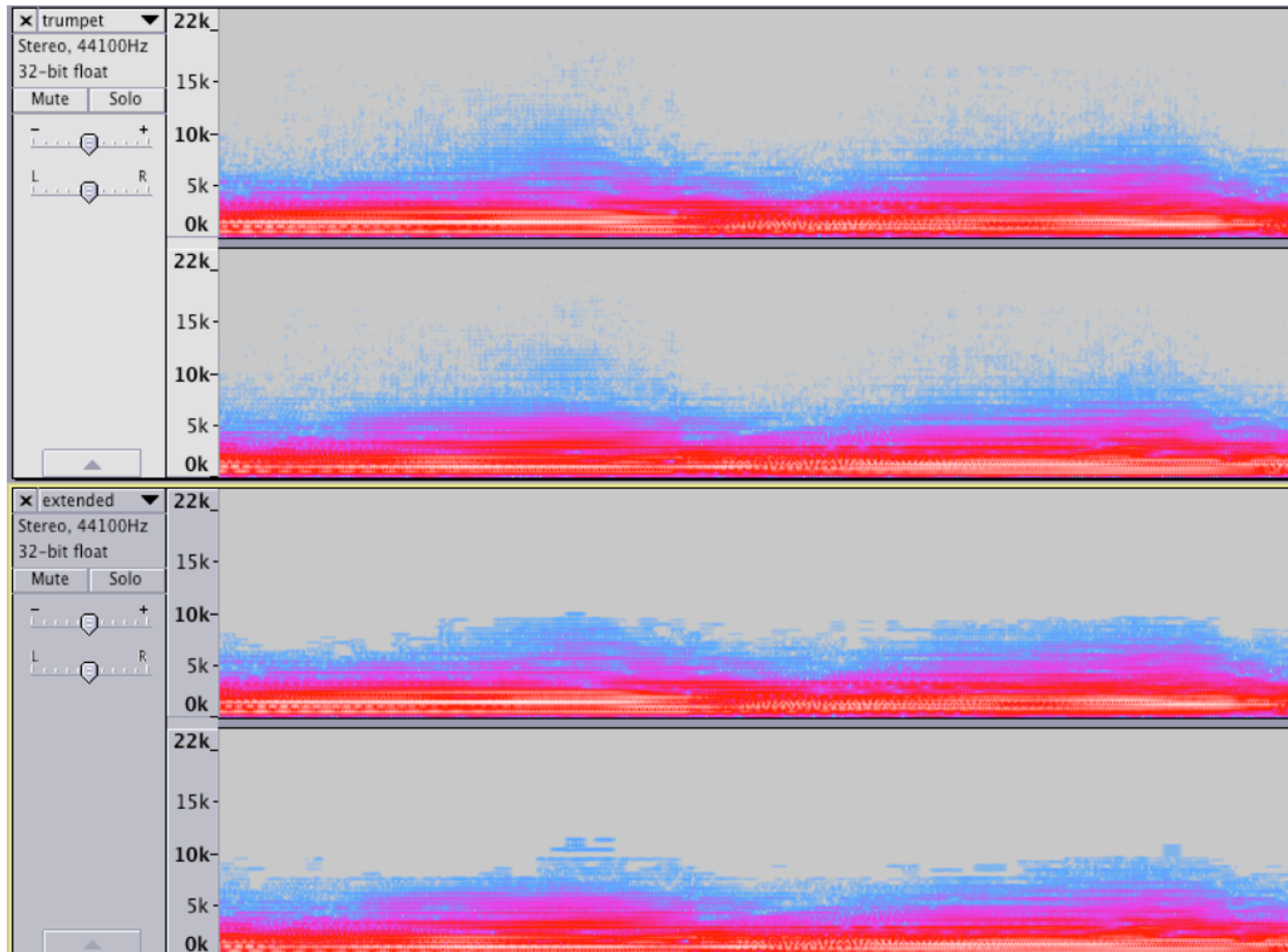
Arbitrary data size per block: no extra header pointers

Results

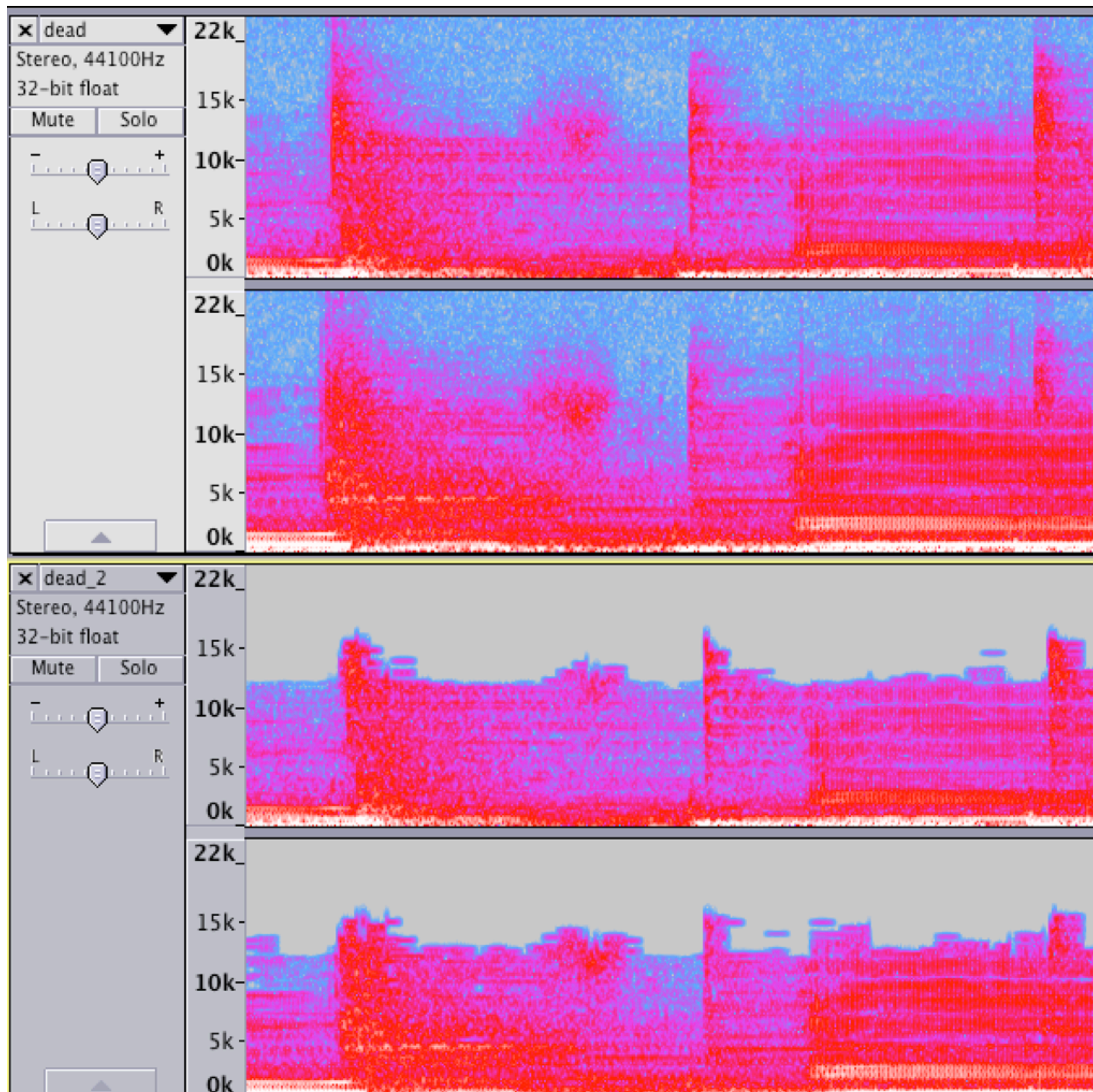
Pre-Echo Cancellation



Spectral Redundancies



SQAM Trumpet Sample



"Box of Rain" by The Grateful Dead

Compression Ratios

Sample Name	Coded File Size	Original Size	Compression Ratio
SQAM Castanets	295,129 bytes	1,590,021	5.4 : 1
SQAM Trumpet	570,092 bytes	3,146,652	5.5 : 1
SQAM Harpsichord	545,698 bytes	2,891,112	5.3 : 1
SQAM German Male Speaker	536,857 bytes	2,949,076	5.5 : 1

ITU-R Test Results

Double-blind, triple stimulus tests with 6 CCRMA Listeners

Castanets

-1.57 → 0.03

+1.54

German

-1.90 → -0.25

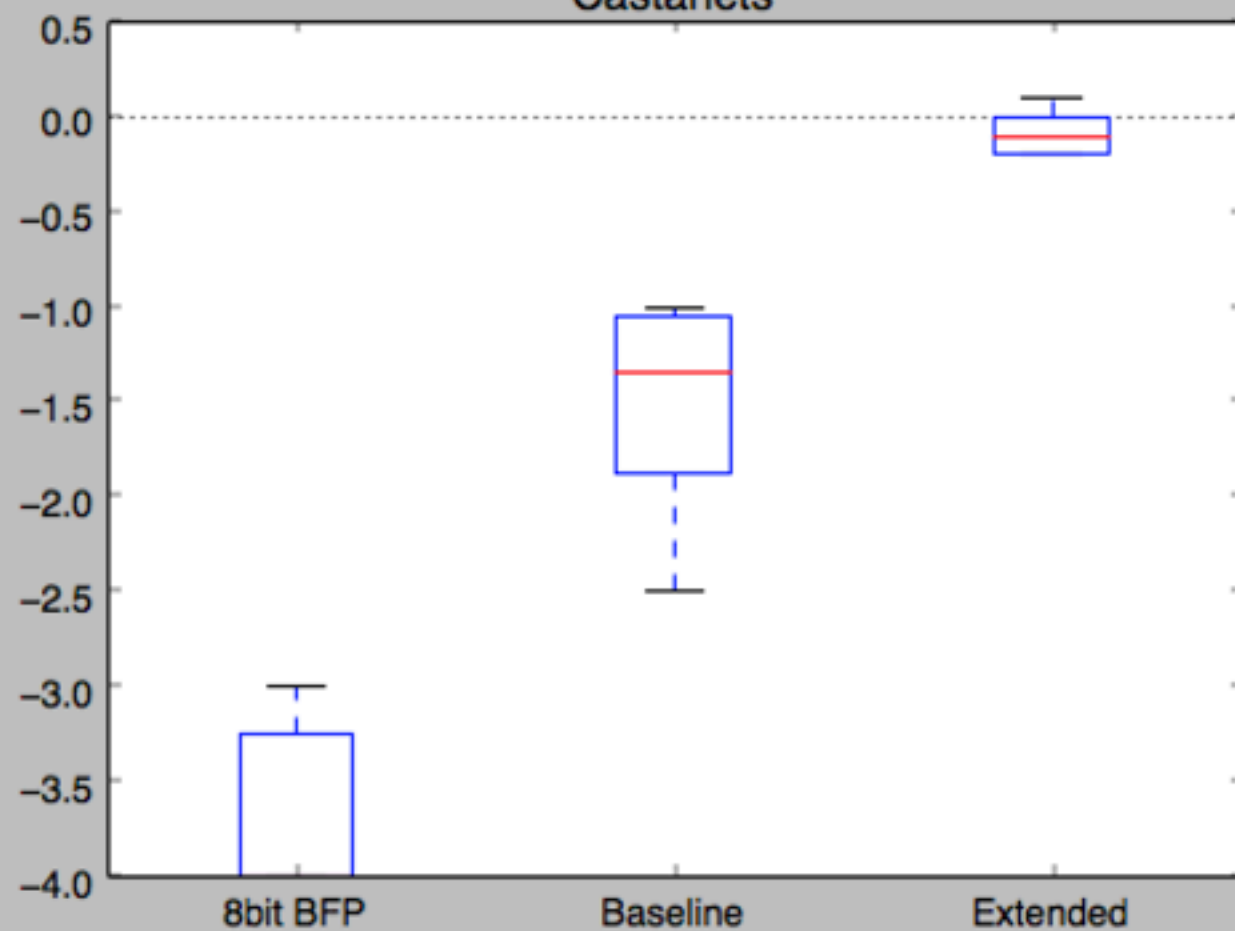
+1.65

Harpsichord

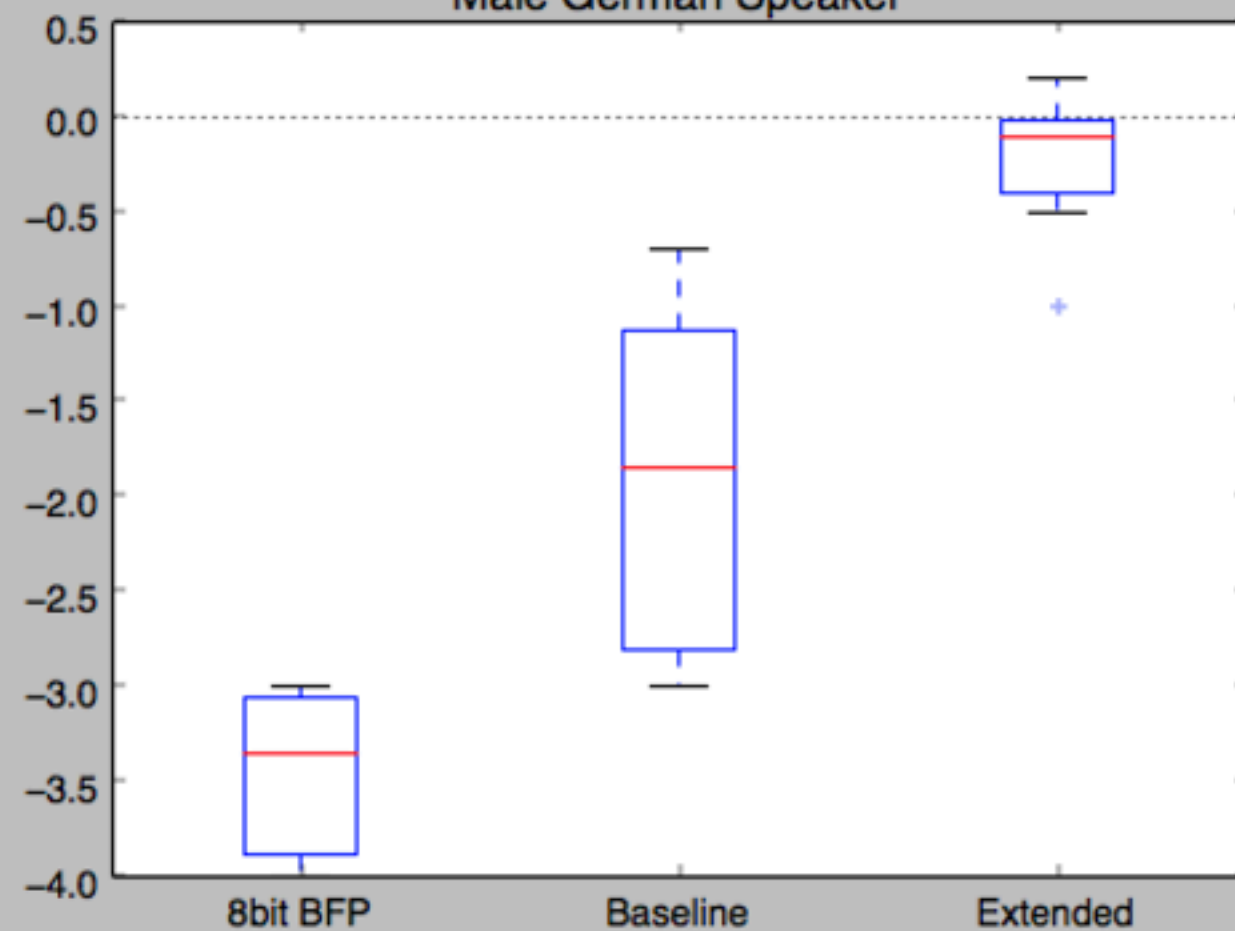
-1.62 → -0.88

+0.77

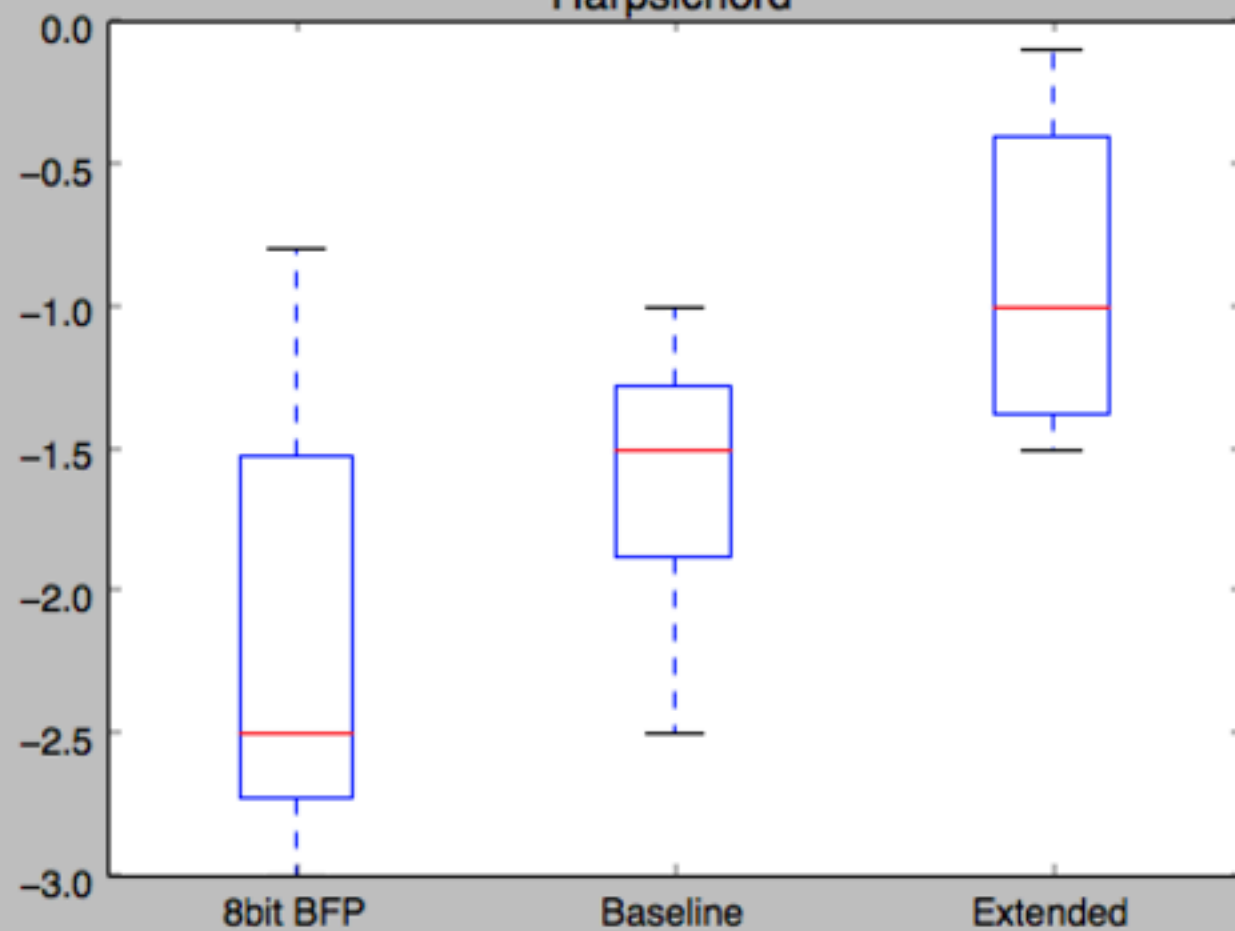
Castanets



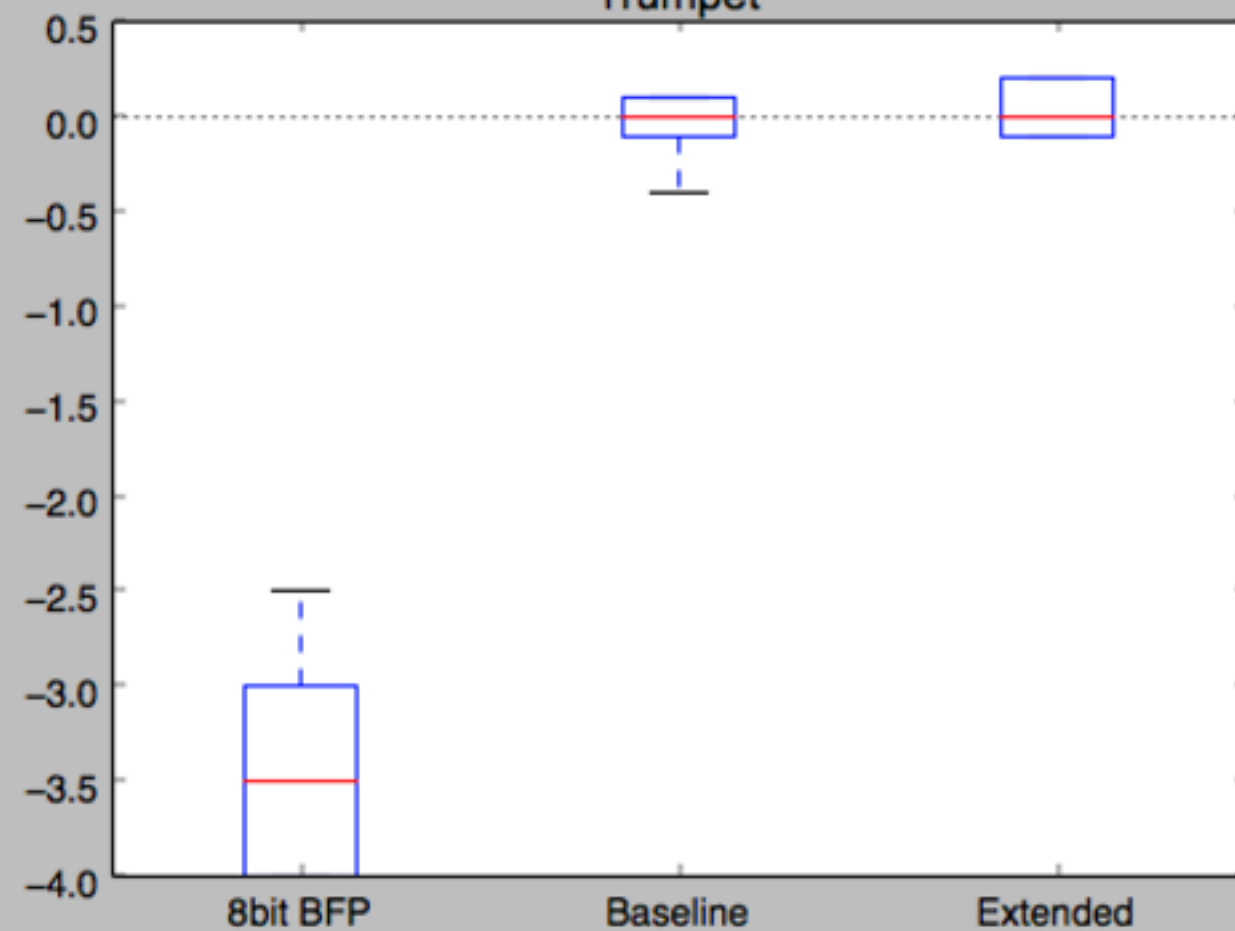
Male German Speaker



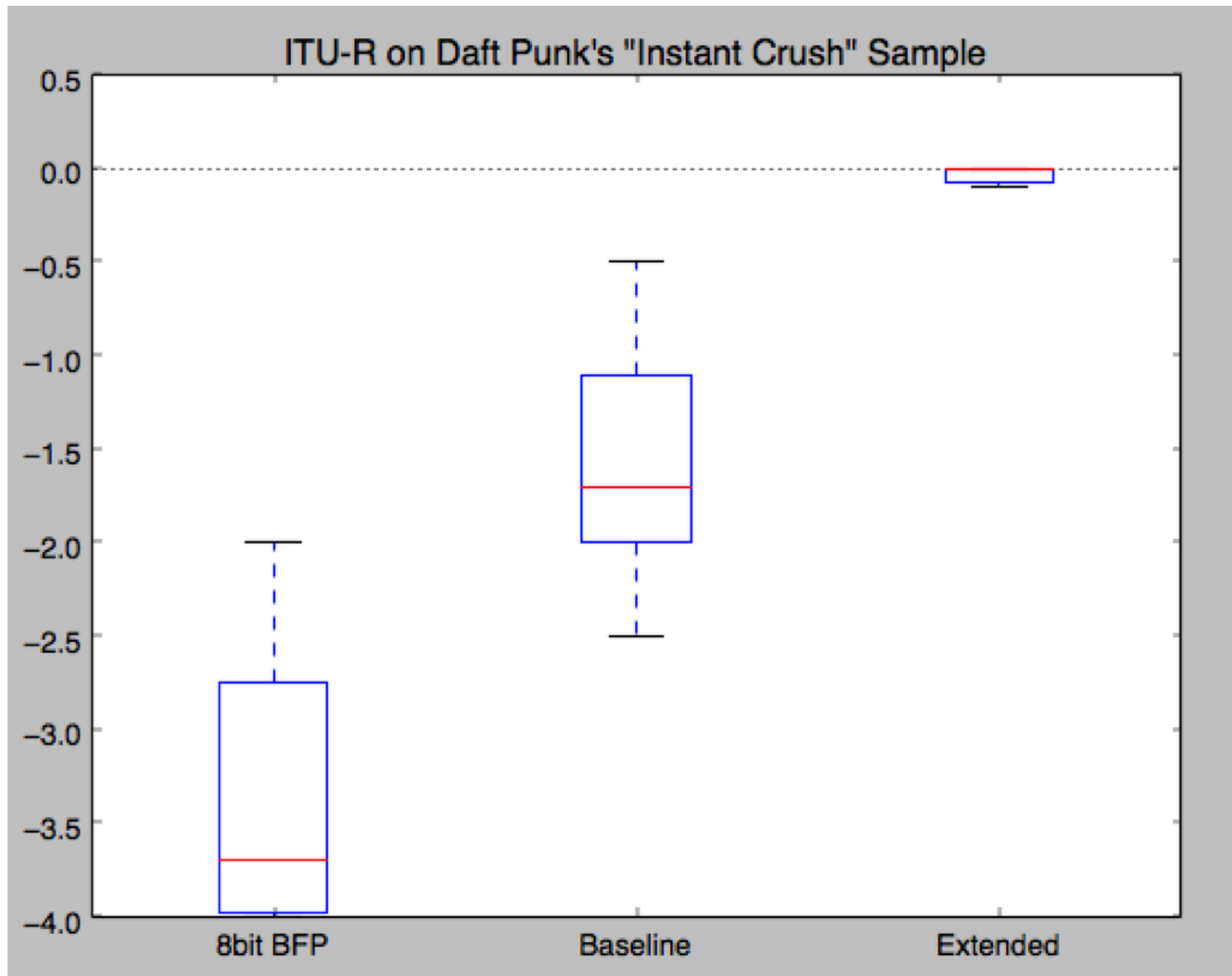
Harpichord



Trumpet



Buy It, Use It, Break It, Fix It



Tested Popular Songs

“Box of Rain” by The Grateful Dead

“Immigrant Song” by Led Zeppelin

“Blackbird” by The Beatles

Future Work

Temporal Noise Shaping

Adaptive onset detection threshold

Lossless Compression

Fin.
