iJDSP – Interactive Illustrations of Speech/Audio Processing Concepts

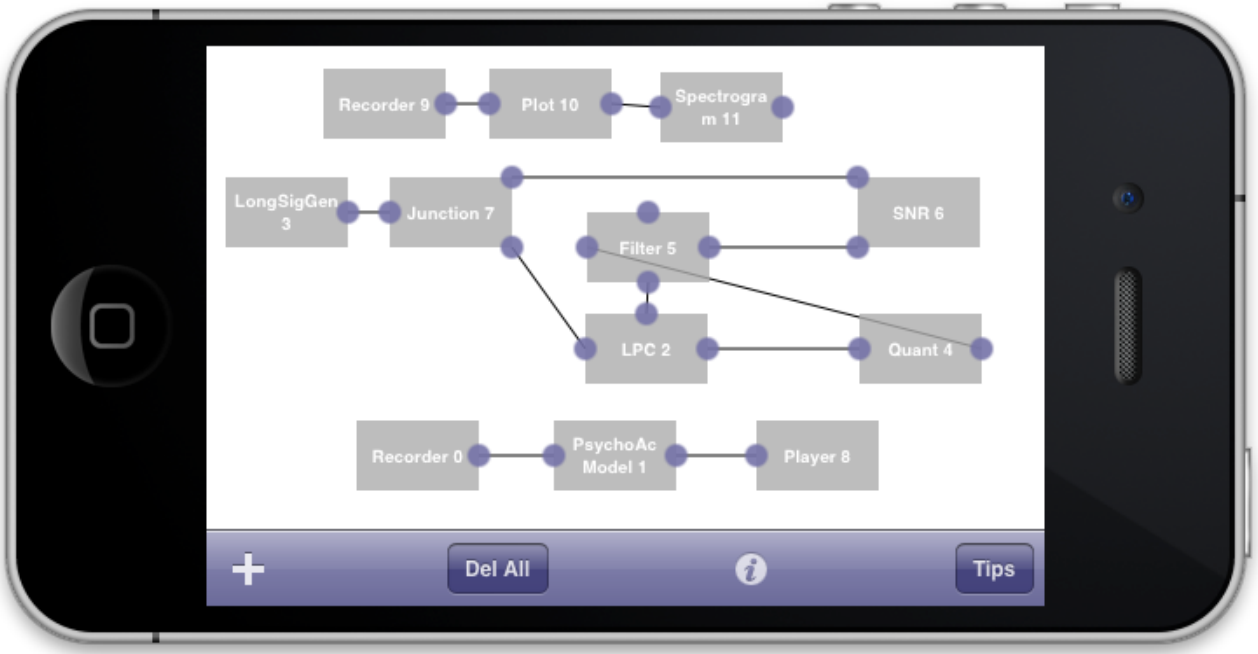
NSF Phase 3 J-DSP Workshop, UCy

Presentation of an Independent Study

By Girish Kalyanasundaram, MS by Thesis in EE

Advisor: Dr. Andreas Spanias, Professor, Electrical Engineering, ASU

Block Diagram Based Learning in iJDSP



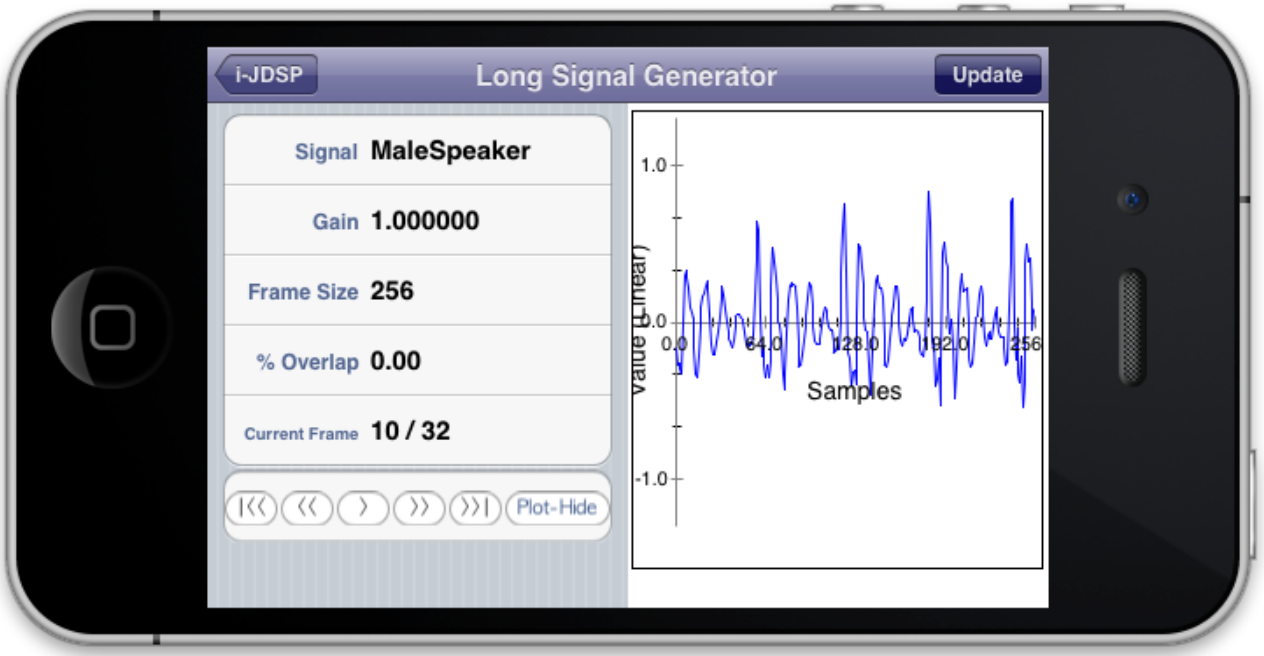
- Effective for constructing basic systems for visualizing speech/audio DSP concepts.

- Requirements:
  - Provision of speech/audio signals
  - Microphone Recording and Playback facility
  - Frame-by-Frame Processing Capability
  - Effective visualization tools

Extension of iJDSP Architecture

- Added framework to create blocks capable of processing long signals.
- These blocks can also interface with conventional blocks processing short signals.
- Frame by frame processing functionality.
- Paradigm designed for creating planned functions in iJDSP – instant setup of complex diagrams.

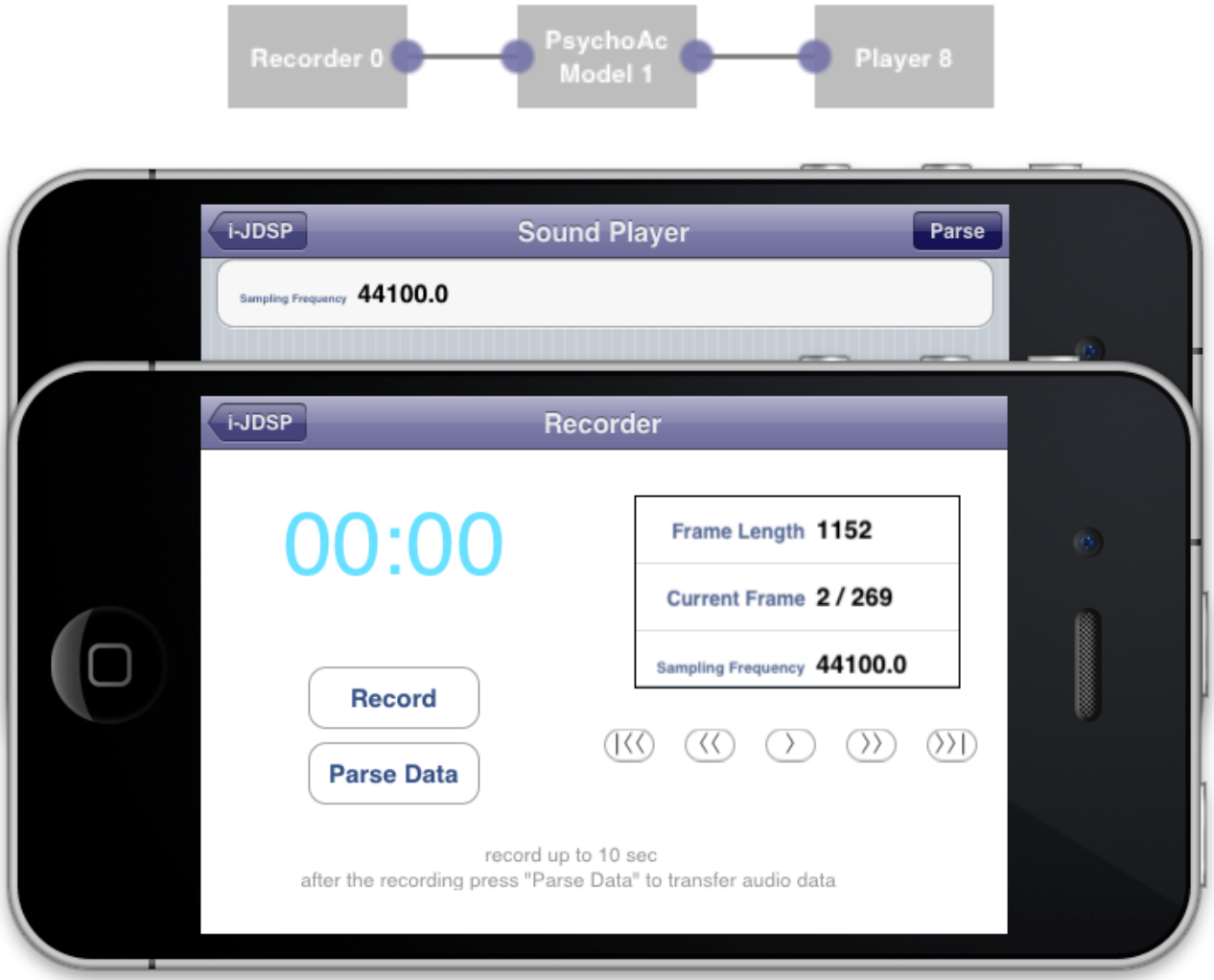
The Long Signal Generator



- Hosts pre-defined speech, music and noise signals.
- Frame-wise traversal and visualization facilitated.

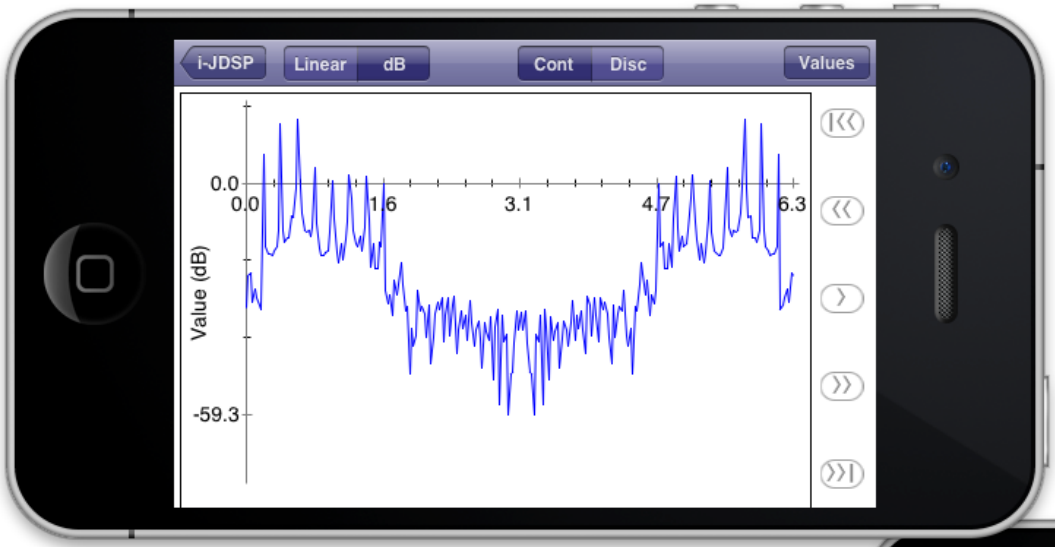
- Configurable Parameters:
  - Frame Size
  - Inter-frame overlap
  - Choice of signal
  - Gain applied to signal

Sound Recording & Rendering



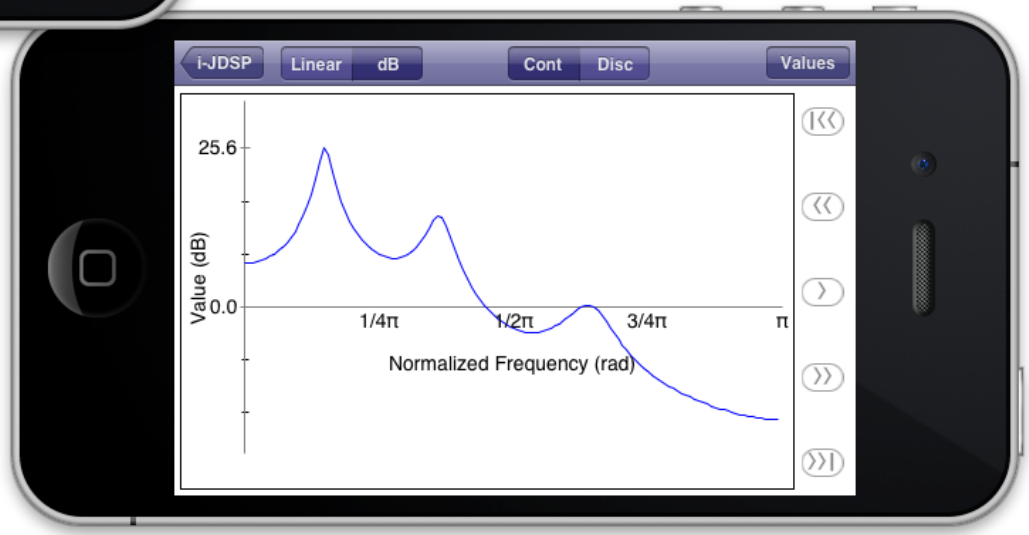
- Records sound at specified sampling rate (8, 16 or 44.1 kHz).
- Frame-wise traversal of signal through playback buttons.
- Sound player aggregates parsed data for rendering.

Frame-By-Frame Visualizations



Plot

Frequency Response

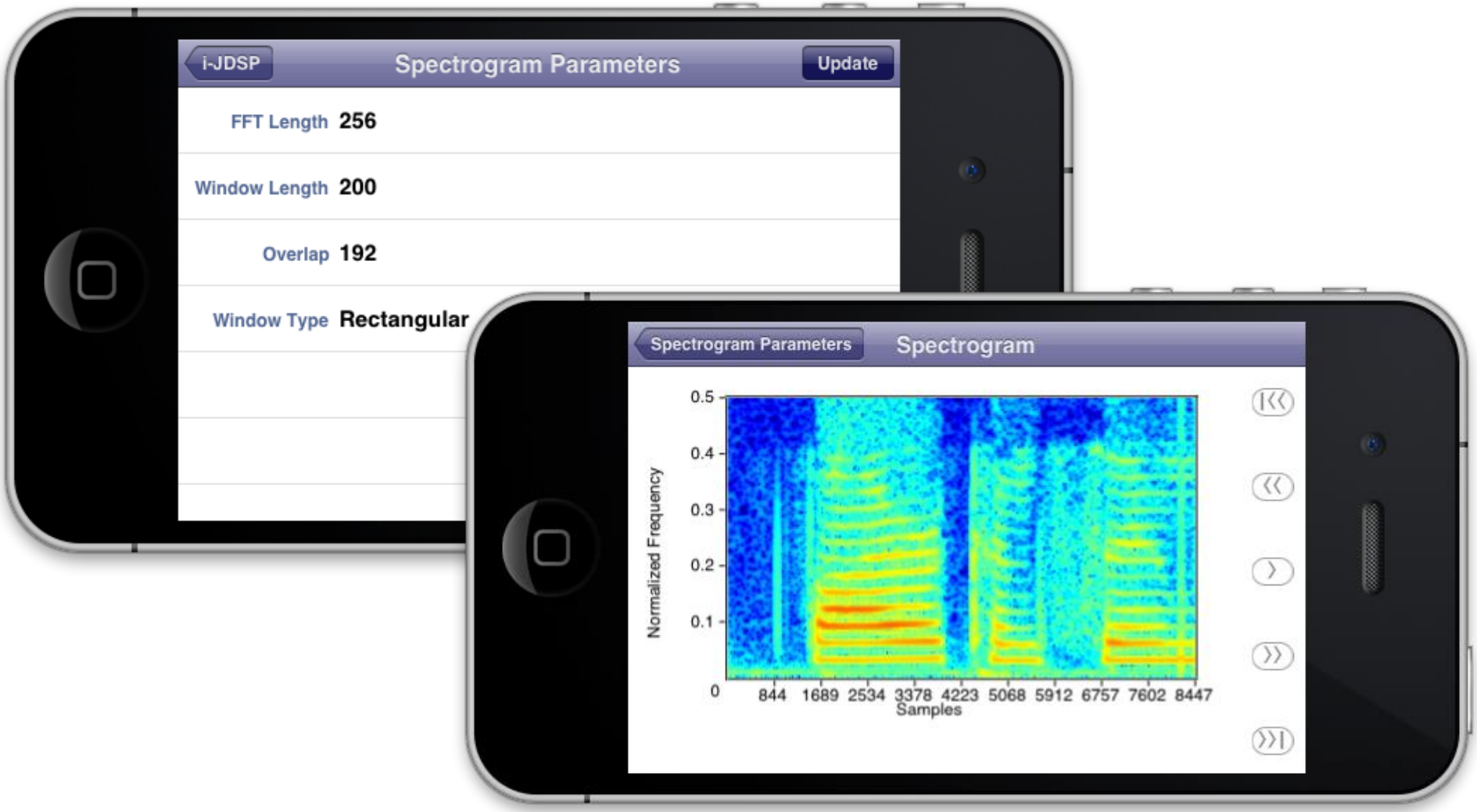


- Playback buttons used to traverse through frames of incoming signal.

Speech/Audio DSP Functions

- Spectrogram
- Linear Predictive Coding (LPC)
- Quantization
- Line Spectral Pairs
- MPEG I Layer 3 Psychoacoustic Model
- Loudness Estimation

Spectrogram

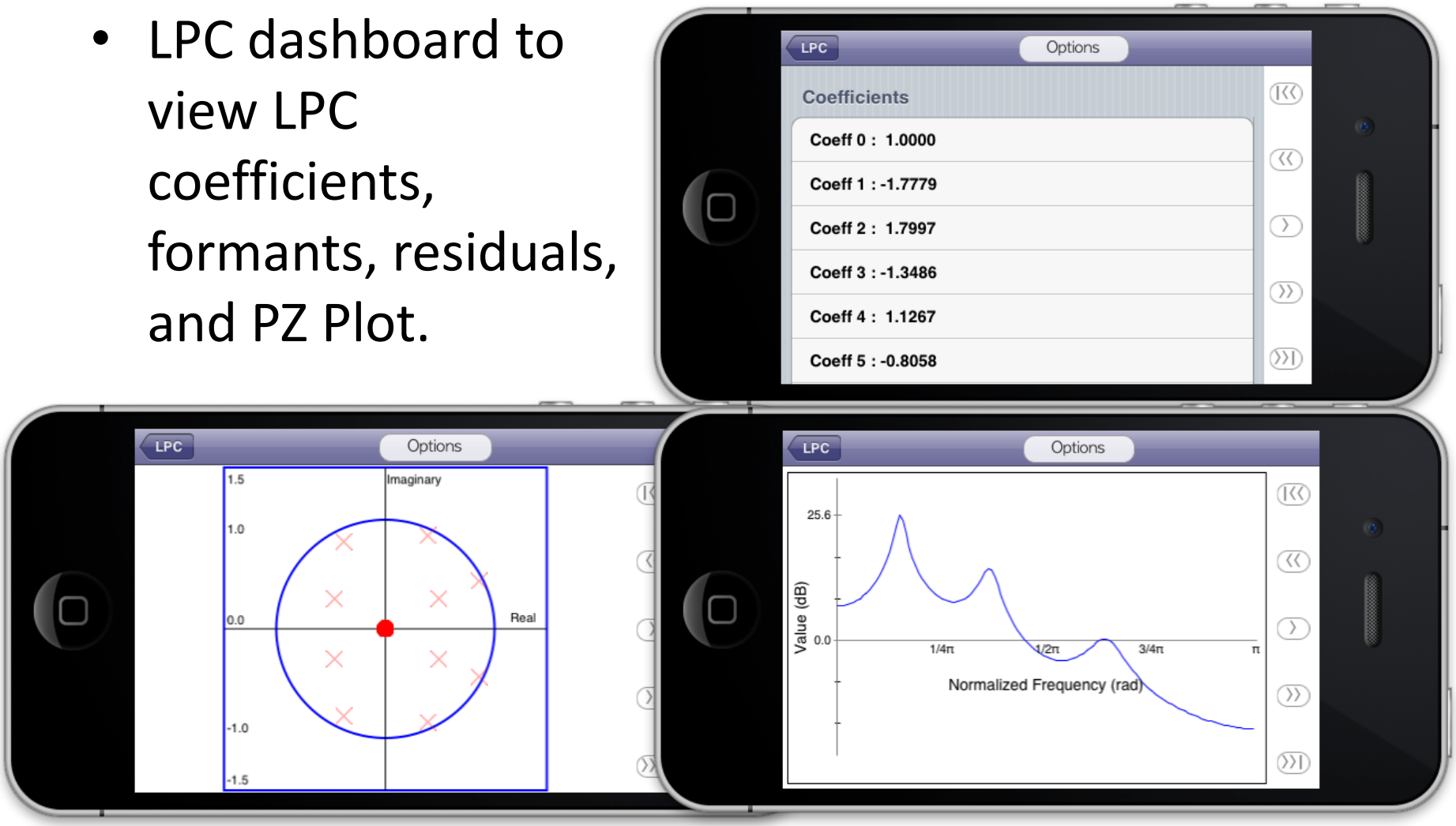


Spectrogram (Contd.)

- Enables visualization of spectrograms of long signals.
- Can study the properties of spectrograms.
- Effect of window length and overlap in time and frequency resolution.
- Window types and effect of taper on resolution.

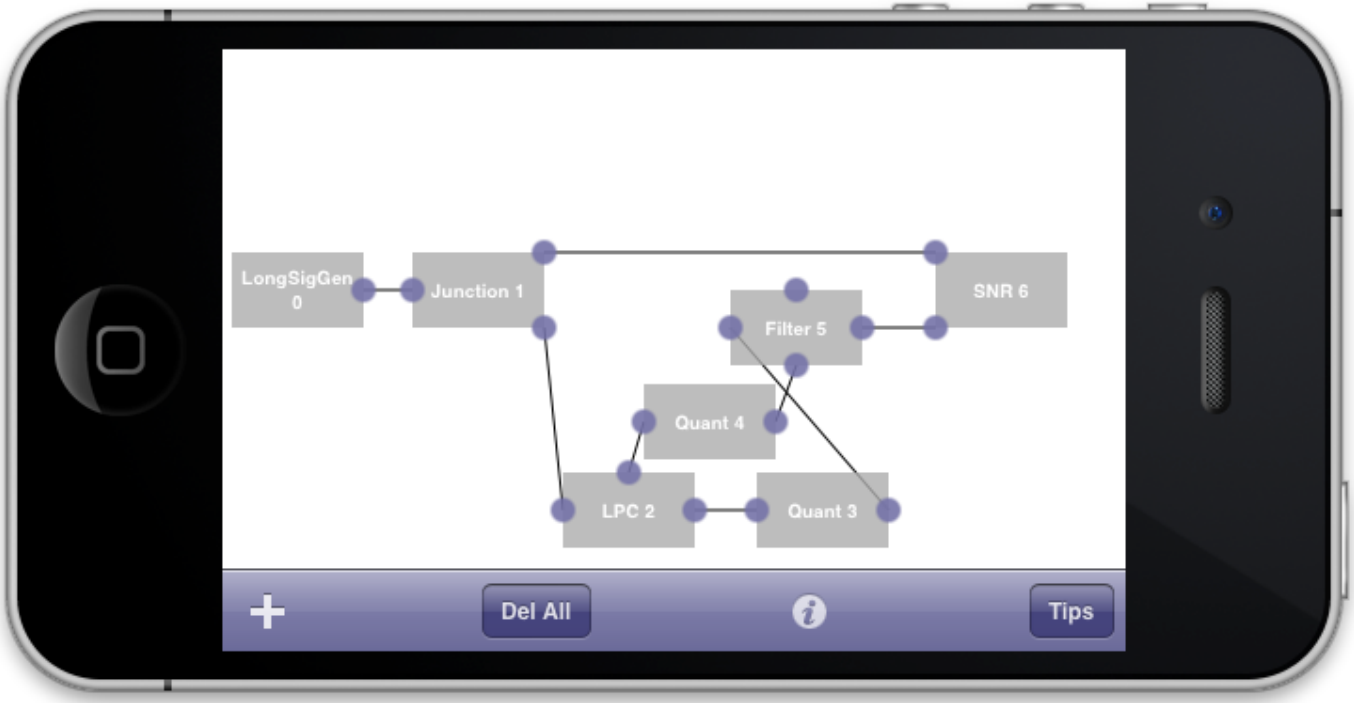
Linear Predictive Coding

- LPC dashboard to view LPC coefficients, formants, residuals, and PZ Plot.



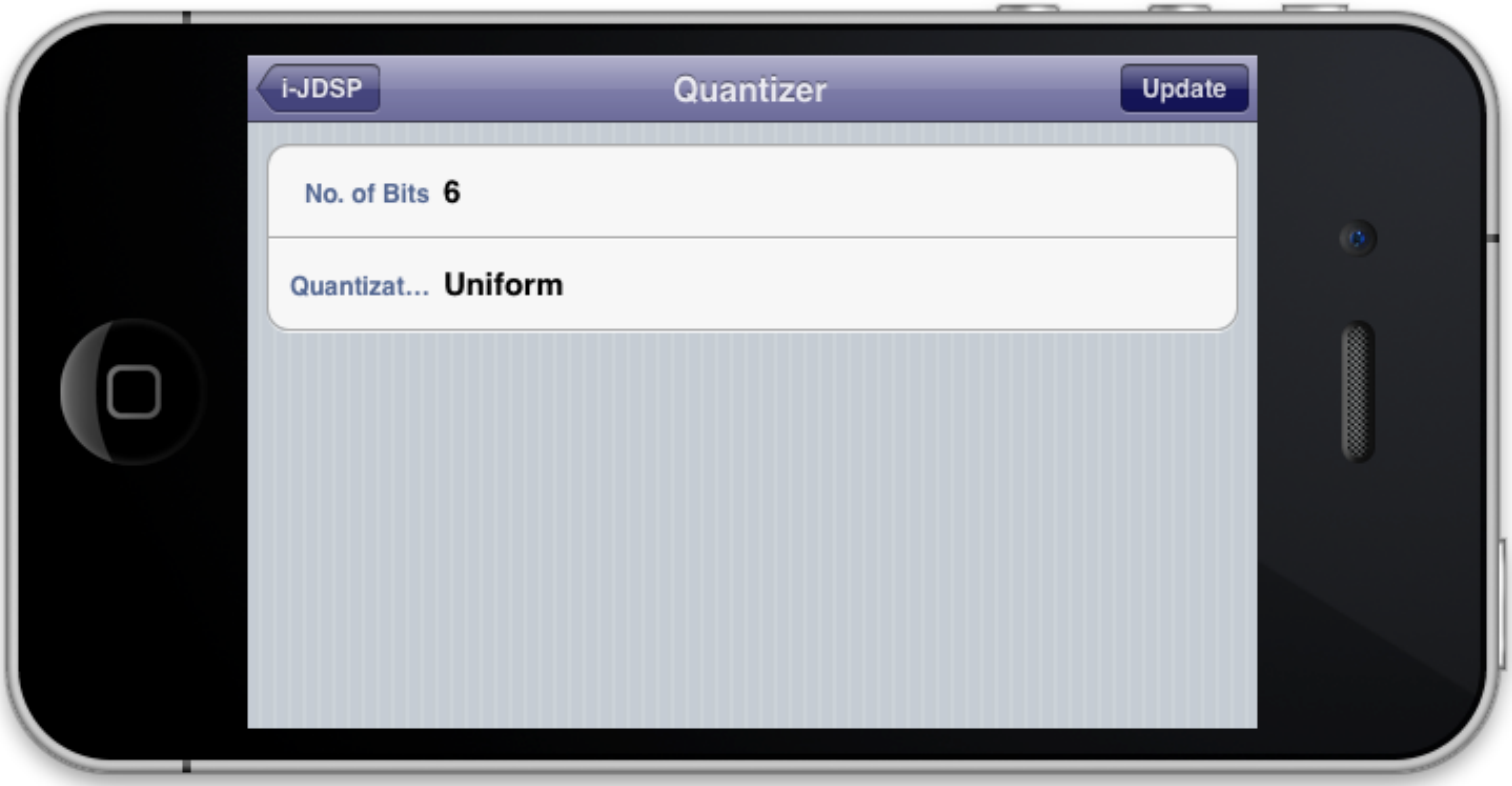
LPC & Quantization

- Observation of quantization effects on LPC coefficients and residual.
- Observation of overall SNR of signal.



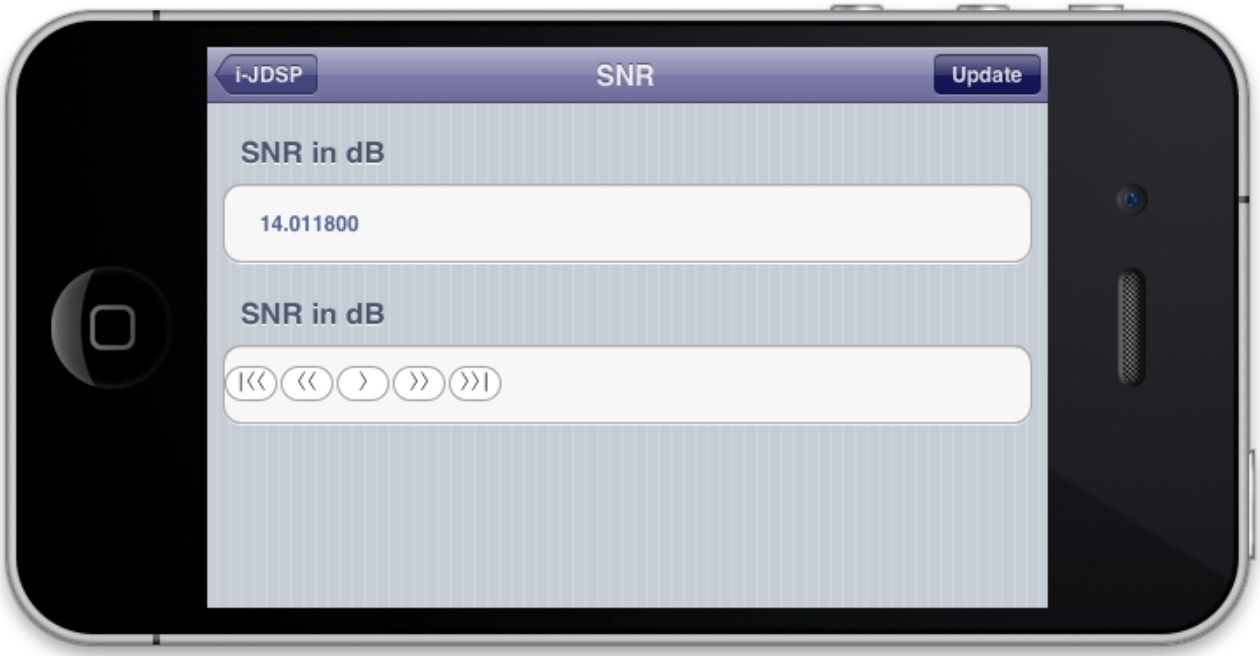
Quantizer Block

- Quantizer block can be use in quantizing LPC coefficients.



LPC & Quantization – SNR

- The SNR block has been equipped with capability to aggregate SNR for entire signals.

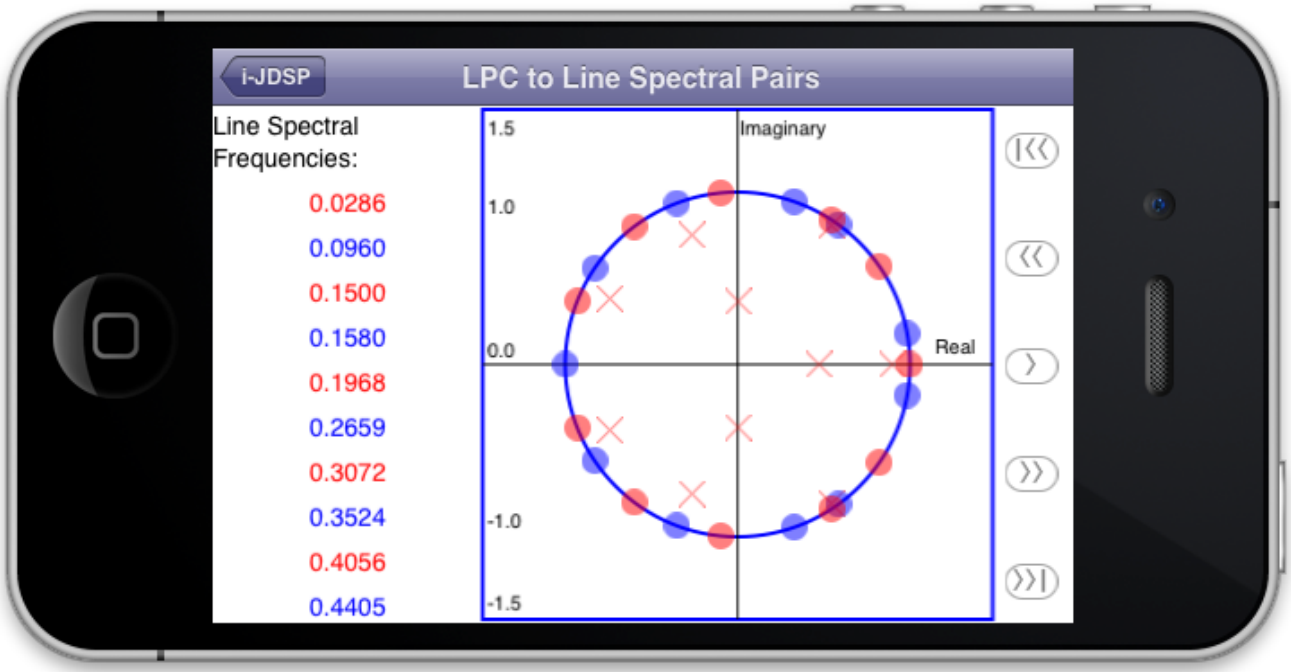


Line Spectral Pairs (LSPs)

- Block to display LSPs and Line spectral frequencies of input LPCs.
- LPC – LSP Placement Demo.
- Block demonstrating quantization effects of LPC.

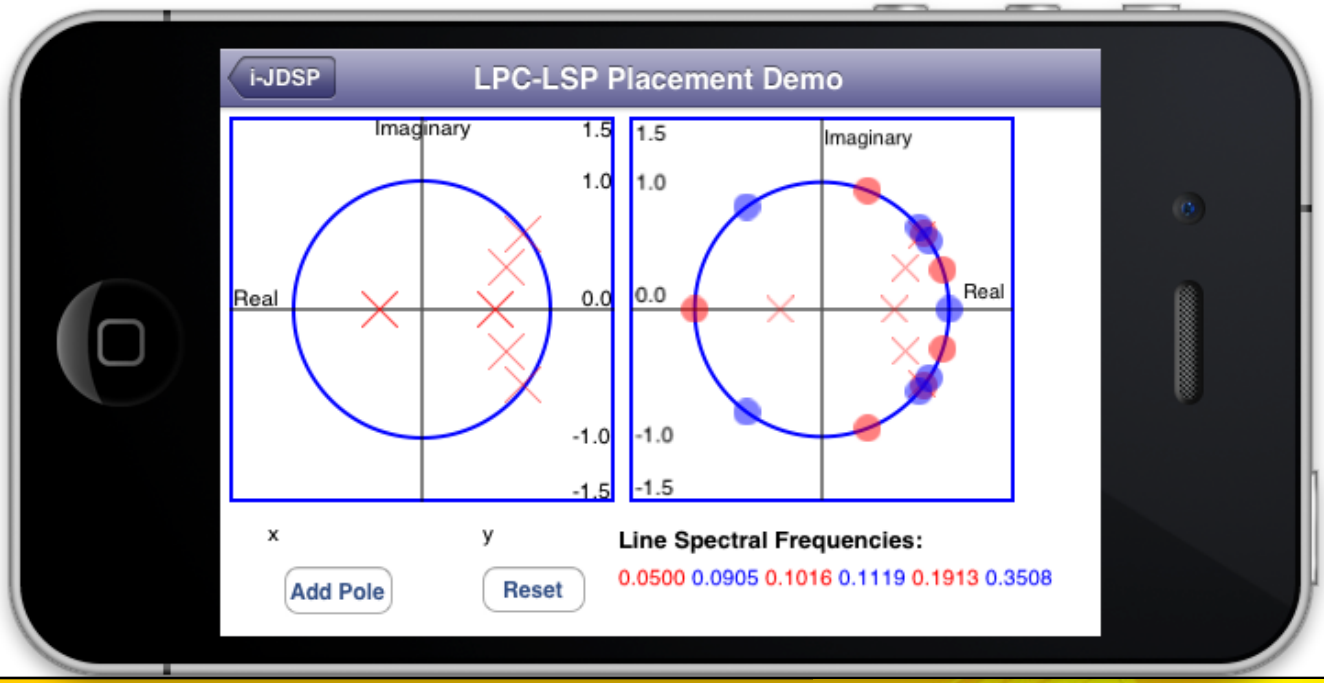
LSP Block

- Displays the LSP Filters Pole Zeros Plots and the line spectral frequency values.



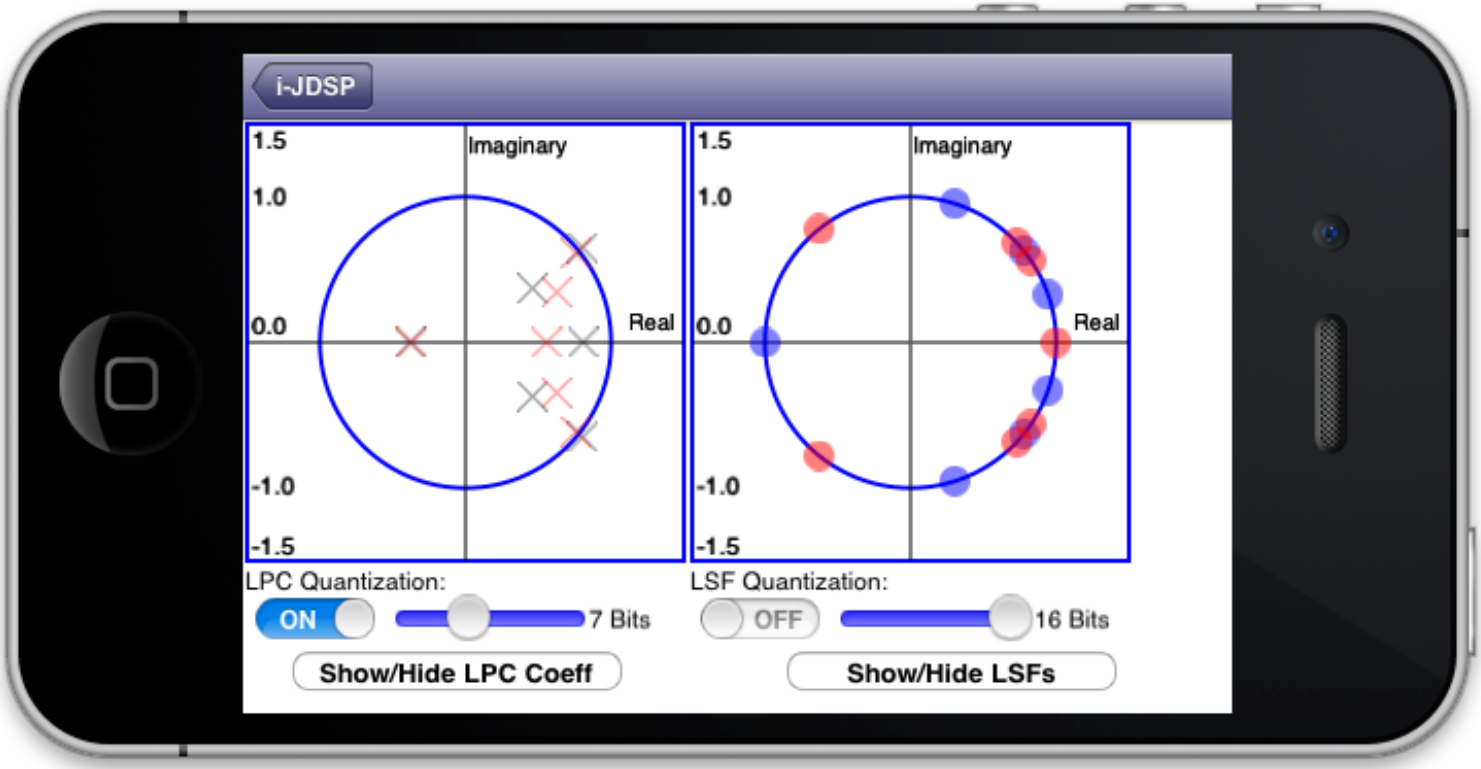
LPC-LSP Demo

- LPC poles can be placed and the LSPs will be shown dynamically.
- Relation between the LPC pole locations and the LSP poles properties can be studied.



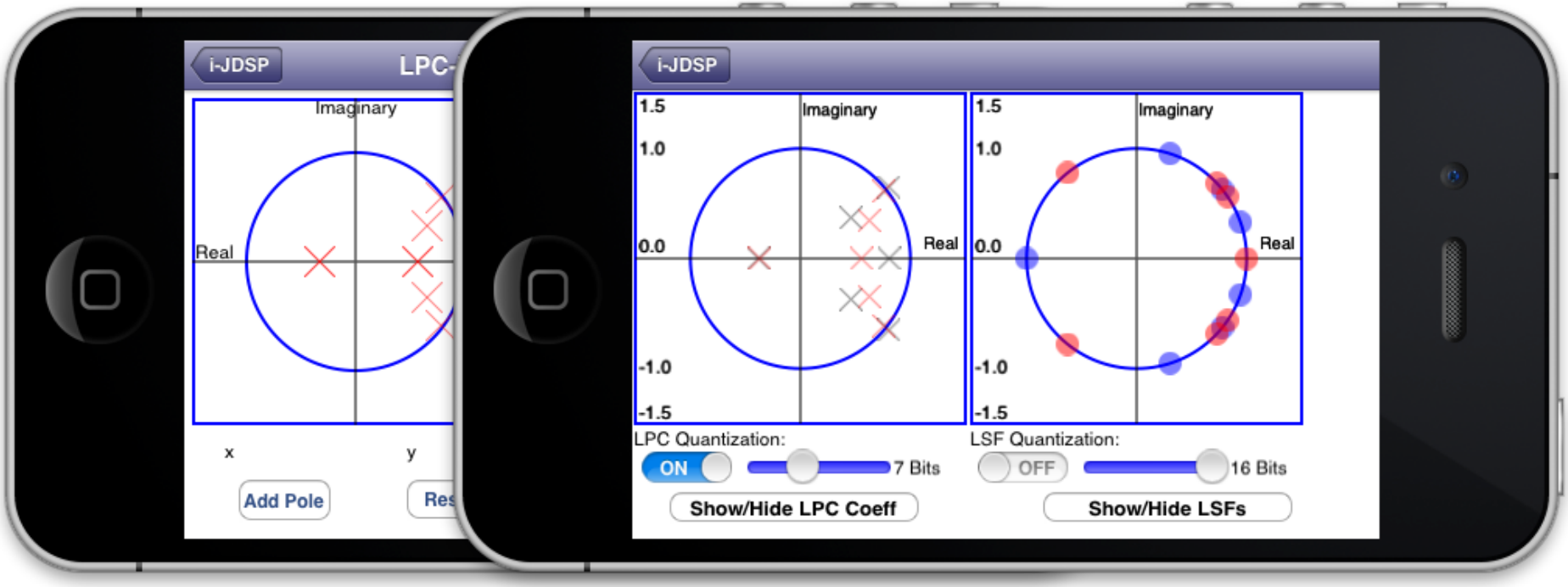
LPC-LSP Quantizer

- This block can demonstrate that LSF quantization prevents unstable poles better than LPC quantization.



LPC-LSP Quantizer (Contd.)

- The block can be used in combination with the LPC-LSP placement demo block to achieve this.



LPC-LSP Quantizer (Contd.)

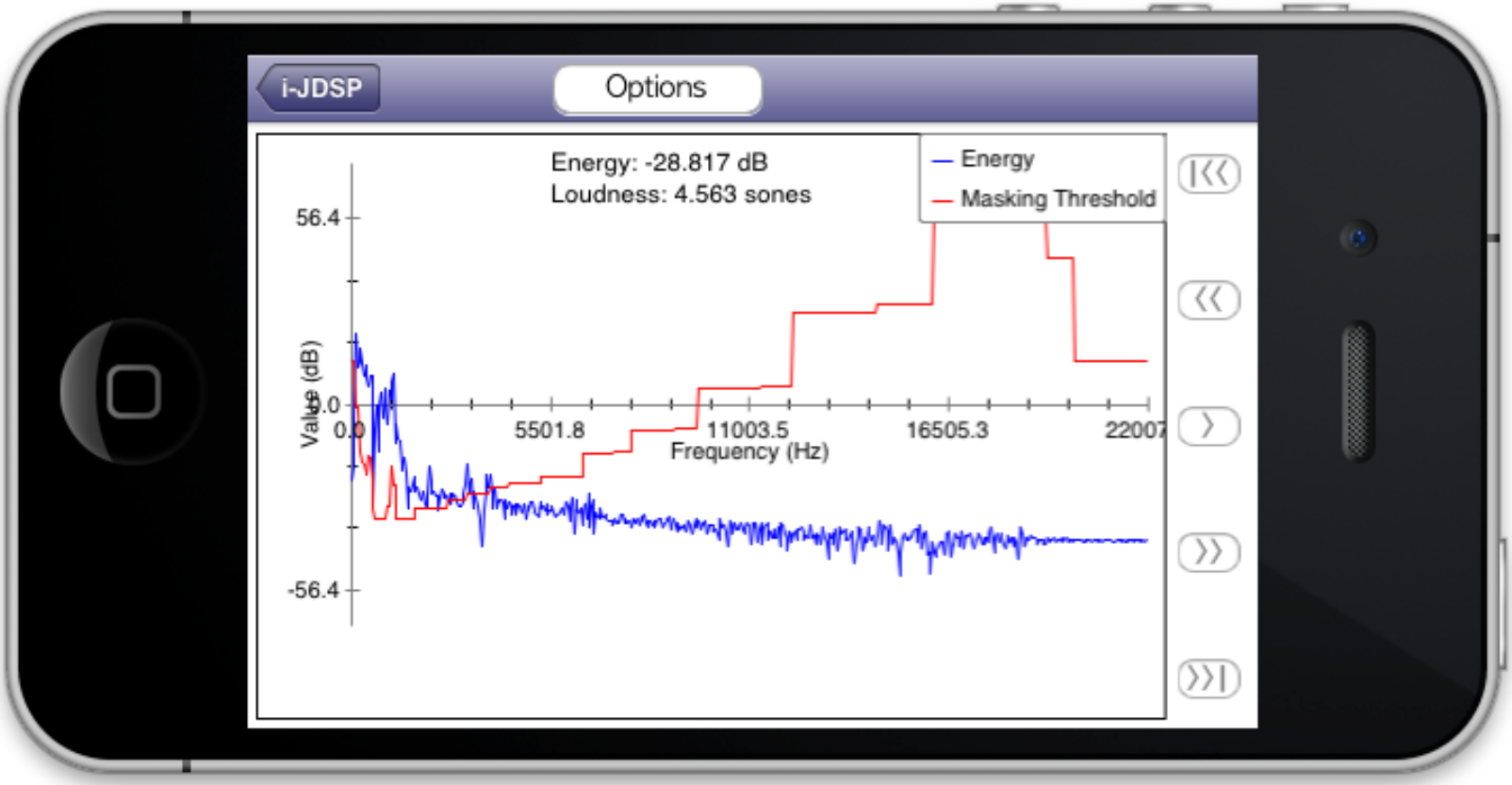
- Displays to view LPC poles and LSF Zeros

Psychoacoustic Model

- The psychoacoustic model block implements the MPEG I Layer 3 psychoacoustic model .
- A modified form of the ISO standard C source code, with improved efficiency for the psychoacoustic model is adopted.
- Visualizations shown:
  - Global Masking Threshold (JND curve)
  - Psychoacoustically peak-picked spectrum
  - Comparison between original and psychoacoustically peak-picked signal

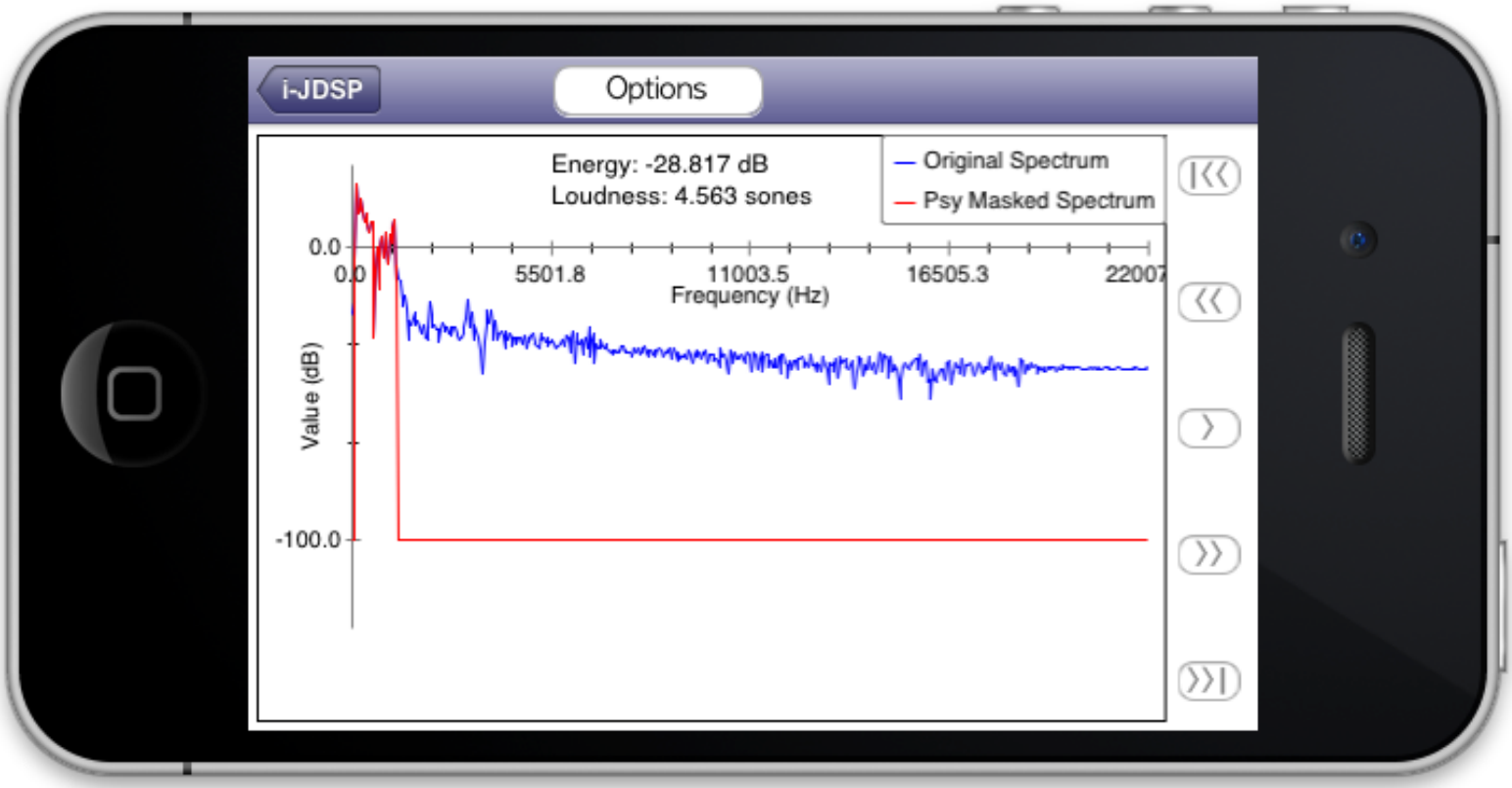
Psychoacoustic Model Interface

- JND curve.



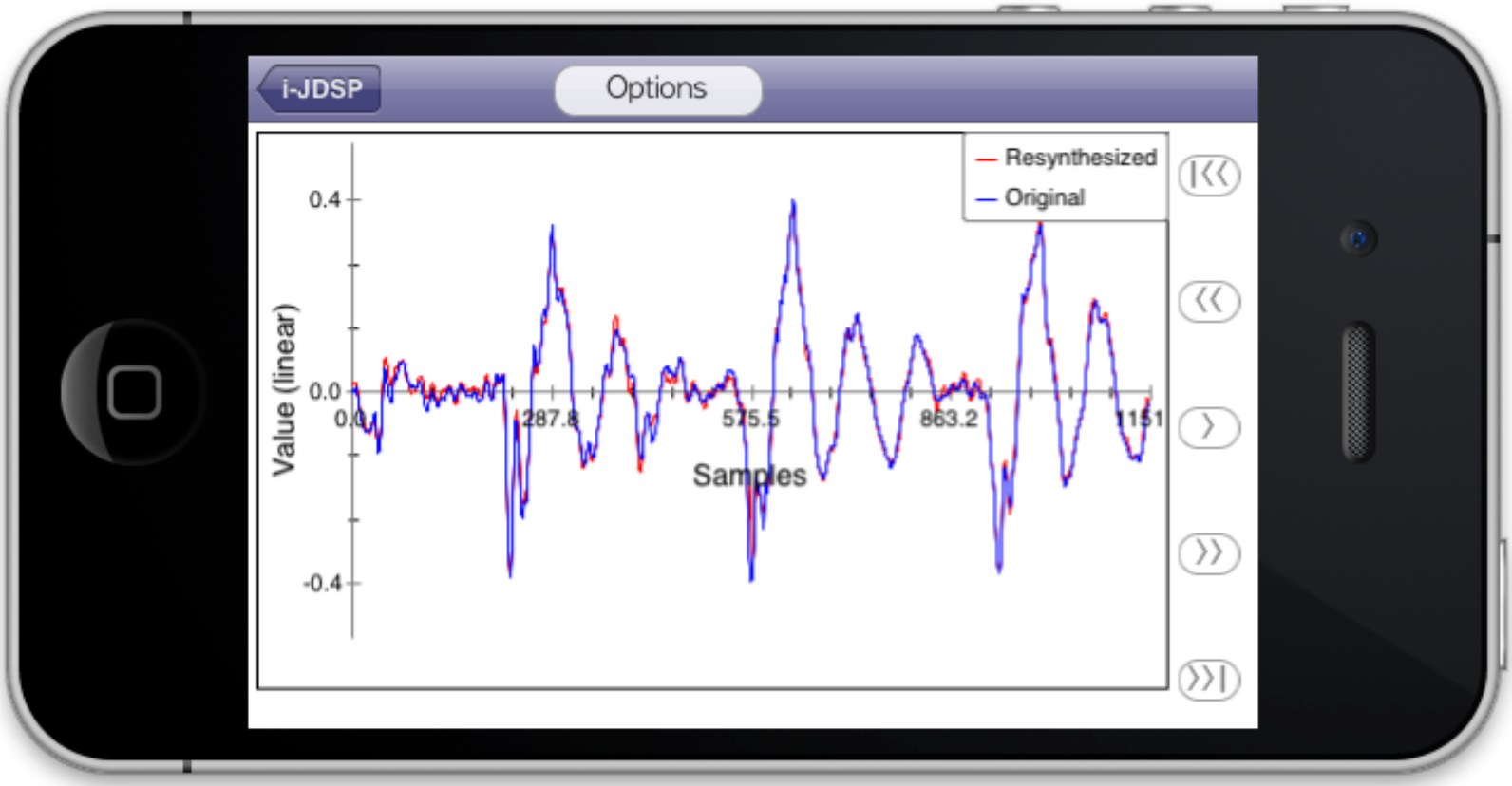
Psychoacoustic Model Interface

- Psychoacoustically masked spectrum.



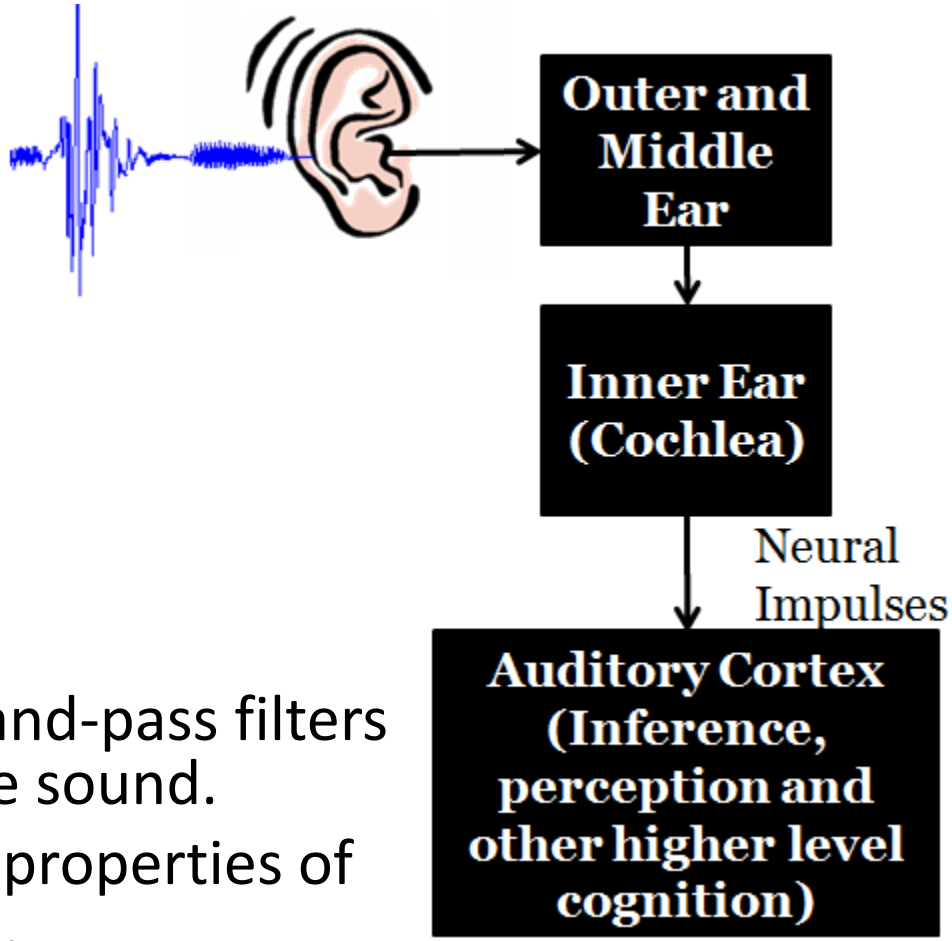
Psychoacoustic Model Interface

- Original vs. peak-picked signal.



Loudness

- A measure of perceived intensity.
- A 'psychophysical' phenomenon.
- A non-linear quantity.
- Reflects properties of the auditory system.



Human Auditory System:

- The ear acts as a bank of band-pass filters detecting frequencies in the sound.
- Perceptual models capture properties of the human auditory system.

The Units of Loudness

The 'Phon' scale

- Derived relative to a 1 kHz sine wave reference.
- A signal has loudness level of X phons when a 1 kHz sinusoid of X dB SPL is perceived to be equally loud.

The 'Sone' Scale

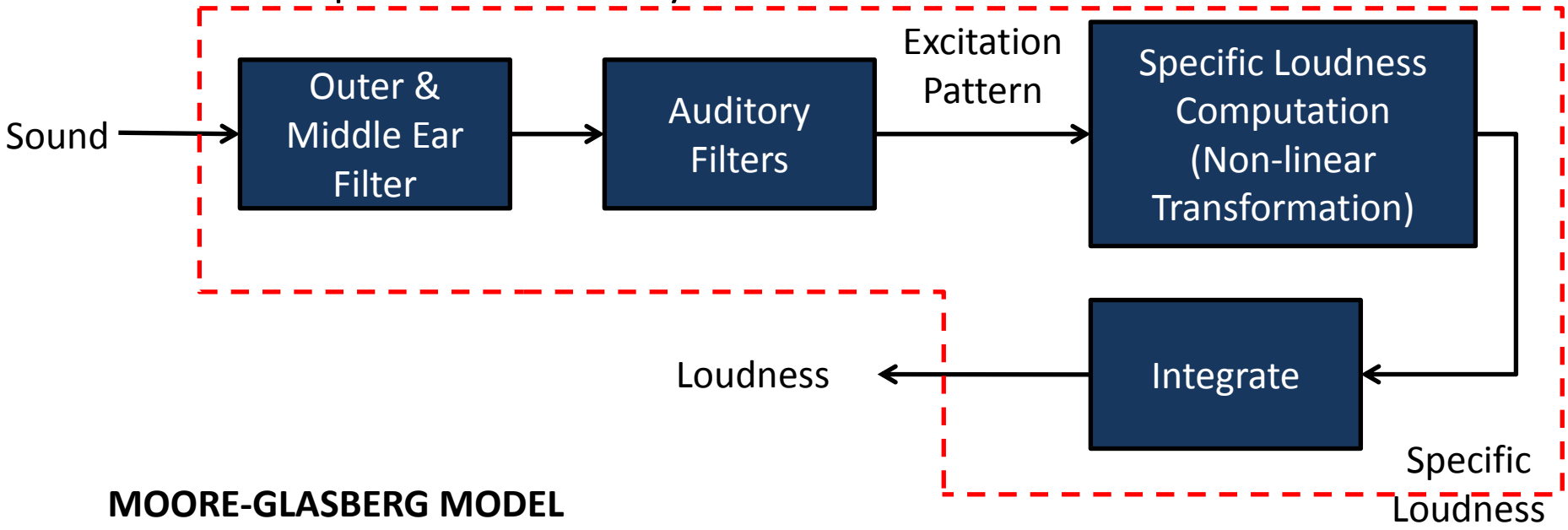
- For a signal with loudness level X phons, loudness in sones is:

$$L = 2^{\frac{X-40}{10}}$$

- A signal with L sones is twice as loud as a signal of loudness L/2 sones.

Loudness Estimation Methods

- *Neural Excitations (Auditory representations)*
  - Compute neural activity for stimuli

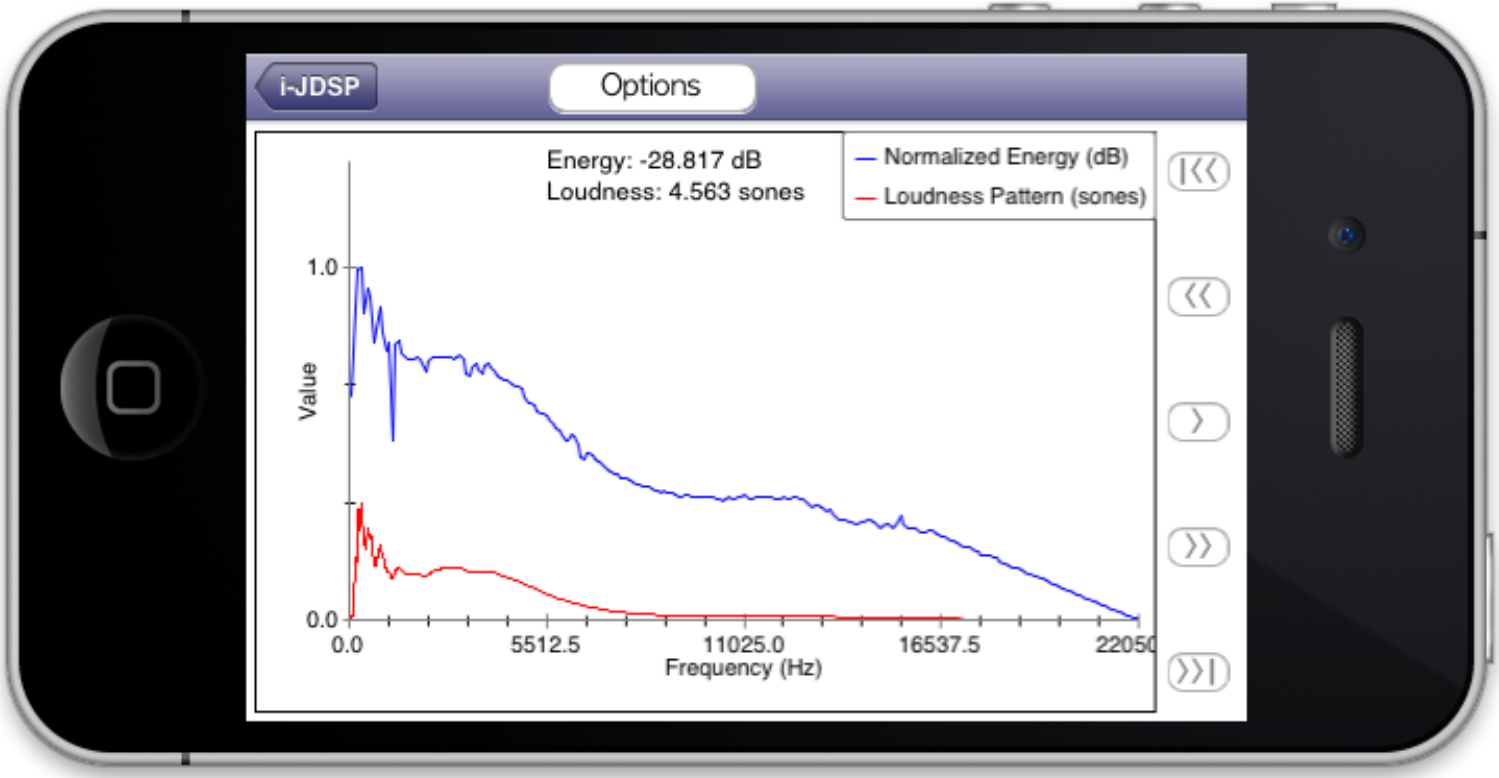


MOORE-GLASBERG MODEL

- Approved by ANSI as a new loudness standard.
- Satisfactory performance for several kinds of spectra.

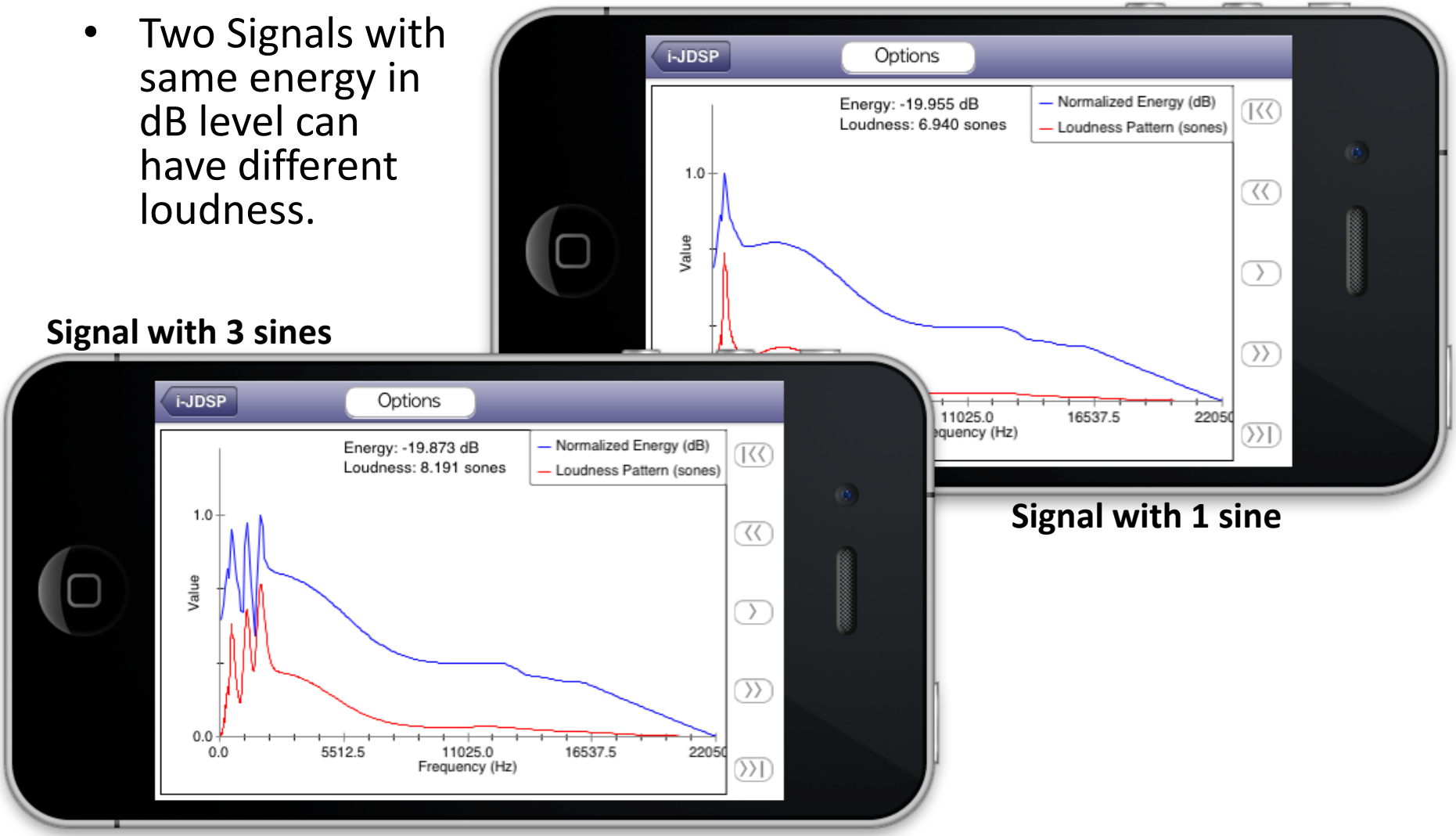
Loudness in Psychoacoustic Model Block

- Spectrum, Specific Loudness Pattern and Total Loudness Displayed in Psychoacoustic Model Block.



Loudness Illustration

- Two Signals with same energy in dB level can have different loudness.



Signal with 3 sines

Signal with 1 sine

THANK YOU

Loudness Estimation Methods

- *A, B, C Weighting*
  - Filters spectrum to compute loudness (derived from ELC)
- *Steven's Law*
  - A power law for loudness of signal with intensity I:

$$L = kI^{0.3}$$

Equal Loudness Contours (ELC) *

ReplayGain: A loudness leveler using ELC. Spotify™ uses ReplayGain.

* http://www.music.princeton.edu/~newton/teaching/music3/2007/fletchermunson.html