

# Hugo Technology

An introduction into Rob Watts' technology



# About Rob Watts

- Audio chip designer – both analogue and digital
- Consultant to silicon chip manufacturers
- Designer of Chord's digital audio products
- 20 years close association with Chord
- Inventor of Class Z (DDFA) digital power amp technology
- Inventor and owner of:

Pulse Array DAC technology

WTA filter algorithm

- Audiophile

# Sound Perception

- Audio engineering is based upon simple psychoacoustic measurements on the ear's capability – distortion levels, noise perception
- In visual perception it is accepted that only 10% is data from the eyes – the other 90% is image processing by the brain – the auditory system is similar
- What we hear is not the output from the ear – the brain performs auditory processing about which we do not understand
- We perceive individual sounds – this does not come from the ear, but from the brain
- These separated sounds are precisely located in 3D space – again brain processing from auditory cues
- The science of auditory scene analysis (how the brain separates sounds into individual sources) is in its infancy – frankly, we have a poor understanding of how the brain does this processing

# Auditory Scene Analysis

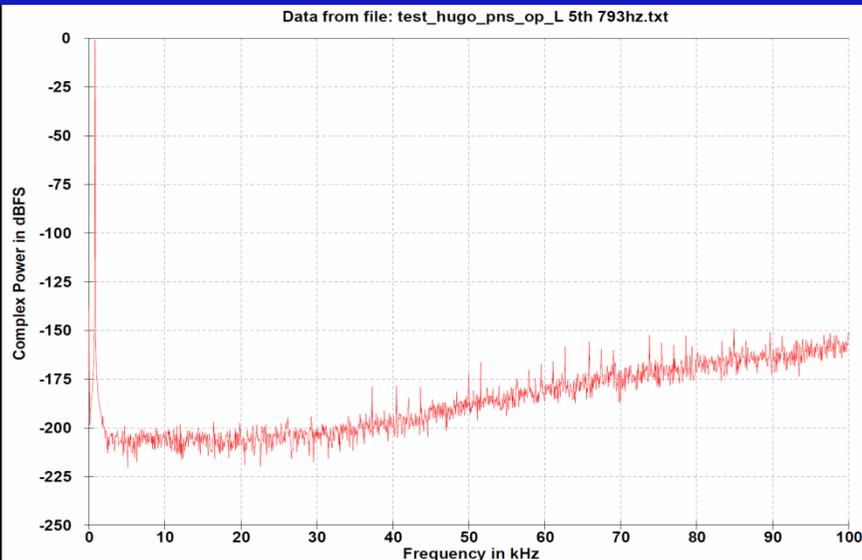
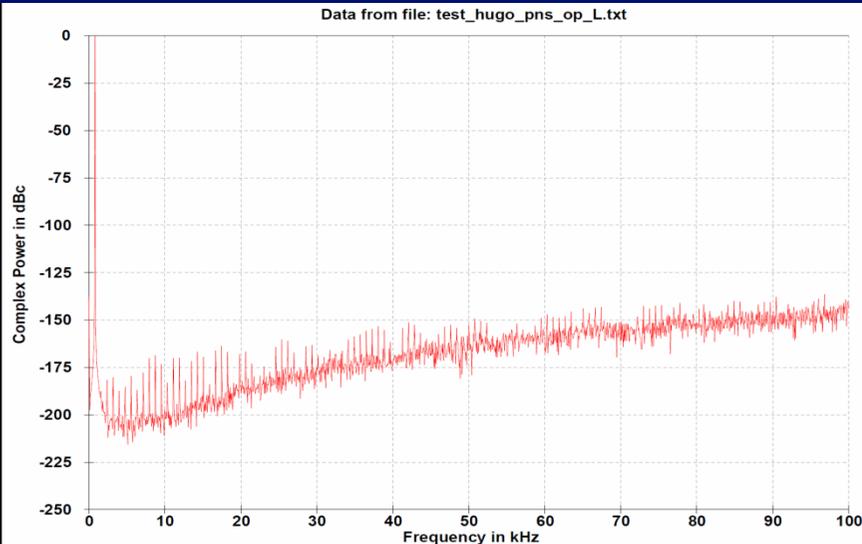


- Imagine this scene in a bar
- You can perceive each instrument as separate entities
- You can understand somebody talking next to you
- You can place each sound to within a foot or two of its actual location in 3D space – height, left right and depth
- If you move back three metres the banjo sounds more distant – move back by 20 metres and you can perceive the music at that depth
- The brain does all this processing and computing in real time
- Science has no detailed understanding on how the human brain does this
- No computer yet designed can do this processing
- And we take it for granted!

# Designing for high-end audio

- We can't rely on psychoacoustic models – thresholds of audibility of distortion and noise, as they tell us nothing about the brain's processing
- Since we do not know how the brain separates sounds out, how do we know what technical parameters are important to sound quality?
- The solution – lots of carefully designed listening tests to evaluate every aspect of performance – make no assumptions “we can't hear that error”
- Audio chips are generally designed by simple technical performance NOT sound quality
- The solution – design with FPGA's (programmable logic devices) and discrete analogue components as FPGA's allow the ultimate flexibility to maximise performance. Commercial DAC chips have severe subjective and technical limitations

# Example of this unique approach



- Top plot shows the performance of a noise shaper (noise shapers are the heart of a modern DAC) with some distortion and noise – distortion is at -170dB and noise floor at -190dB
- Bottom plot shows same noise shaper but improved – no distortion below 20kHz, noise floor is at -200dB
- On same FPGA, switching between the two options was very audible – top version sounded harder, brighter, with poorer instrument separation, and poorer perception of depth in the soundstage
- Both noise shapers exceed psychoacoustic performance of the ear – but the top noise shaper upsets the brain's ability to process sounds into separate entities.
- Bottom plot is the final noise shaper for Hugo

# Perception - timing

- Timing is an important perceptual cue – the beginning and ending of notes (transients) help the brain to separate sounds out
- The ears have a tapped delay line with neurons firing with differences in timing between the ears – the inter-aural delay network can detect timing differences of the order of 4 $\mu$ S
- Inter-aural delay is used to help localising sounds in space
- The 4 $\mu$ S threshold implies that the brain samples data at 250kHz – much faster than 44.1kHz of CD recordings
- A recent paper in National Physics Review – Human Time Acuity Beats Fourier Uncertainty Principle – shows that timing is extremely important

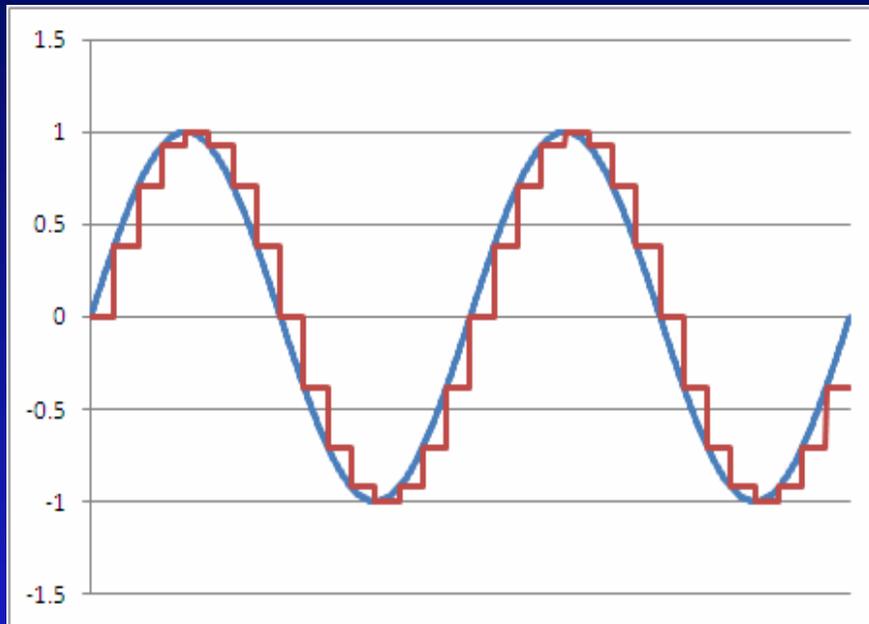
# Timing with Digital Audio

- We know the ears can differentiate 4 $\mu$ S timing differences
- We also know that timing is an important for the brain's processing of audio
- How is the timing of the original transients preserved if CD is sampling at 22 $\mu$ S?
  - The interpolation filter (an FIR filter that has a line of taps multiplying coefficients to delayed data) recovers the original amplitude and timing information of the recording
  - This filter re-creates the missing bits between samples
  - If you look at the original Whittaker-Shannon sampling theory, then for a bandwidth limited signal, if you use an infinite tap length FIR filter then the "missing bits" will be perfectly reconstructed
  - The FIR filter has a  $\text{sinc}(x)/x$  response – if you use taps that have 16 bit coefficient accuracy, you need about 1,000,000 taps for an 8 times filter!
  - Practical filters have limited tap length – a few hundred maximum
  - These conventional filters do not properly reconstruct the original timing of transients

# Timing – WTA filter

- Given that 1,000,000 taps is not currently possible, can we maximise timing performance with different algorithms?
  - Yes the algorithm does have an impact on performance
  - The WTA filter algorithm has been carefully developed with many listening tests
  - The WTA algorithm is much closer to an ideal Whittaker-Shannon interpolation filter than any other interpolation filter used in audio
- But even with the WTA algorithm, increasing the tap length gives a substantial improvement in sound quality
- Long tap length WTA filters have the following sound quality benefits:
  - Much better stereo image (as the inter-aural timing delay is used by the brain for location)
  - Much deeper bass (bass perception depends upon the starting and finishing transients of the bass note)
  - Instrument separation and focus is much better (distorted timing damages the brains ability to separate sounds out)
  - Fast rhythms are easier to perceive; you can hear individual piano keys rather than a jumble of notes
- Hugo has a huge 26,368 tap length WTA filter. It uses 16 208MHz DSP cores in parallel to create this filter

# 2048FS filters



The graph to the left shows a 12kHz sine wave filtered to 8FS (red plot) against filtering at 2048FS (blue plot). The 8FS filtered output – as found in most audio DAC's - does not look like the original sine wave, but the 2048FS looks perfect

The 8FS output has a number of problems – the big step changes overloads the analogue sections creating more HF distortion; the step changes also make the DAC more sensitive to clock jitter – this has the result of more noise floor modulation. That is more noise depending upon the signal – in this case it depends upon the rate of change of the signal. The ear/brain is extremely sensitive to noise floor modulation

# Noise floor modulation

- Noise floor modulation occurs when noise increases/decreases depending upon the music signal.
- The ear/brain is extremely sensitive to this problem as it interferes with the brain's ability to separate sounds into individual entities
- Listening tests have shown sensitivity to noise floor modulation well below levels that are measureable
- Noise floor modulation make the sound hard, bright and aggressive; it degrades instrument separation and focus; reducing noise floor modulation improves sense of focus, smoothness and refinement – it sounds much more natural

# Hugo and noise floor modulation

- The DAC architecture has a large influence on noise floor modulation – pulse array DAC's have innately very low levels of modulation
- The reference power supply to the DAC is crucial, this is very low noise and low impedance, with individual references per channel
- RF noise is a major problem as it inter-modulates with the analogue electronics causing noise floor modulation – extensive RF filtering is employed, together with steps to reduce the analogue sensitivity
- Quad layer ground planes are used, so that ground induced noise and distortion is eliminated
- DSD sources have large HF noise – this is digitally filtered by -110dB at 100kHz giving more natural sound quality for DSD
- Jitter is a big source of noise floor modulation – incoming jitter is eliminated by a digital phase lock loop (DPLL).

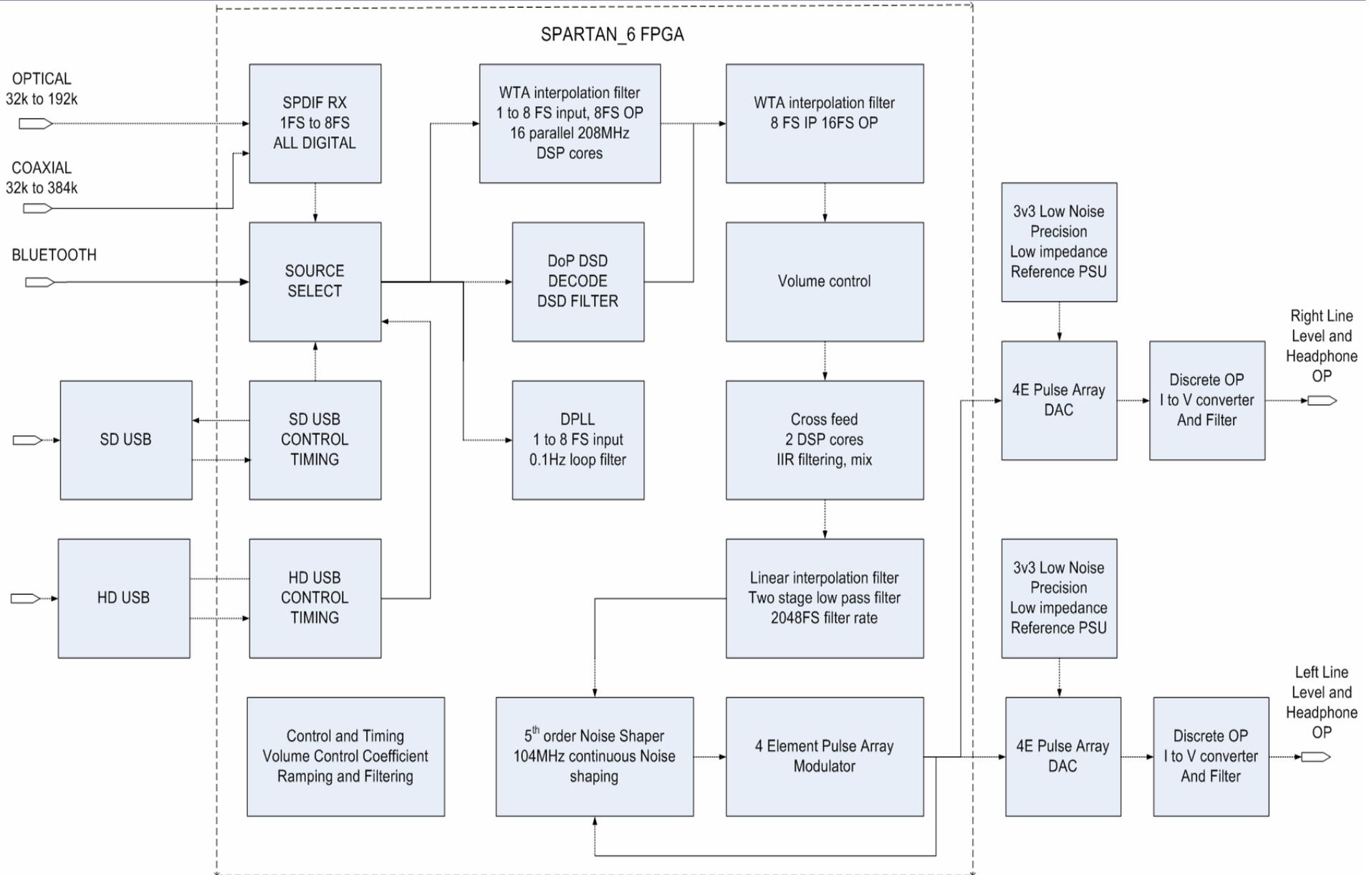
# Digital phase lock loop (DPLL)

- The DPLL, to eliminate incoming jitter, and to prevent low frequency jitter causing skirts on fundamentals, has a very low frequency 0.1Hz cut-off frequency
- It has rapid lock acquisition time of 1mS
- Hugo has a three stage 2048FS filter in order to reduce noise floor modulation

# Headphone drive

- Hugo has a discrete OP stage integrated into the DAC output amplifier and filter
- The OP stage is full Class A (with 300 ohm loads)
- It is capable of 5v RMS and peak output currents of 0.5A
- The OP stage is very low output impedance of 0.075 ohms
- Digital cross-feed, with 3 settings, is implemented digitally at 16FS with two 48 bit DSP cores. The cross-feed uses an analogue type IIR filtering

# Hugo Block Diagram



4e Pulse array discrete DAC, together with dual low noise, low impedance reference PSU

FPGA –handles SPDIF decoding, USB timing, DPLL,WTA filtering, DSD decoding and filtering, volume control, cross-feed, control, noise shaping and DAC

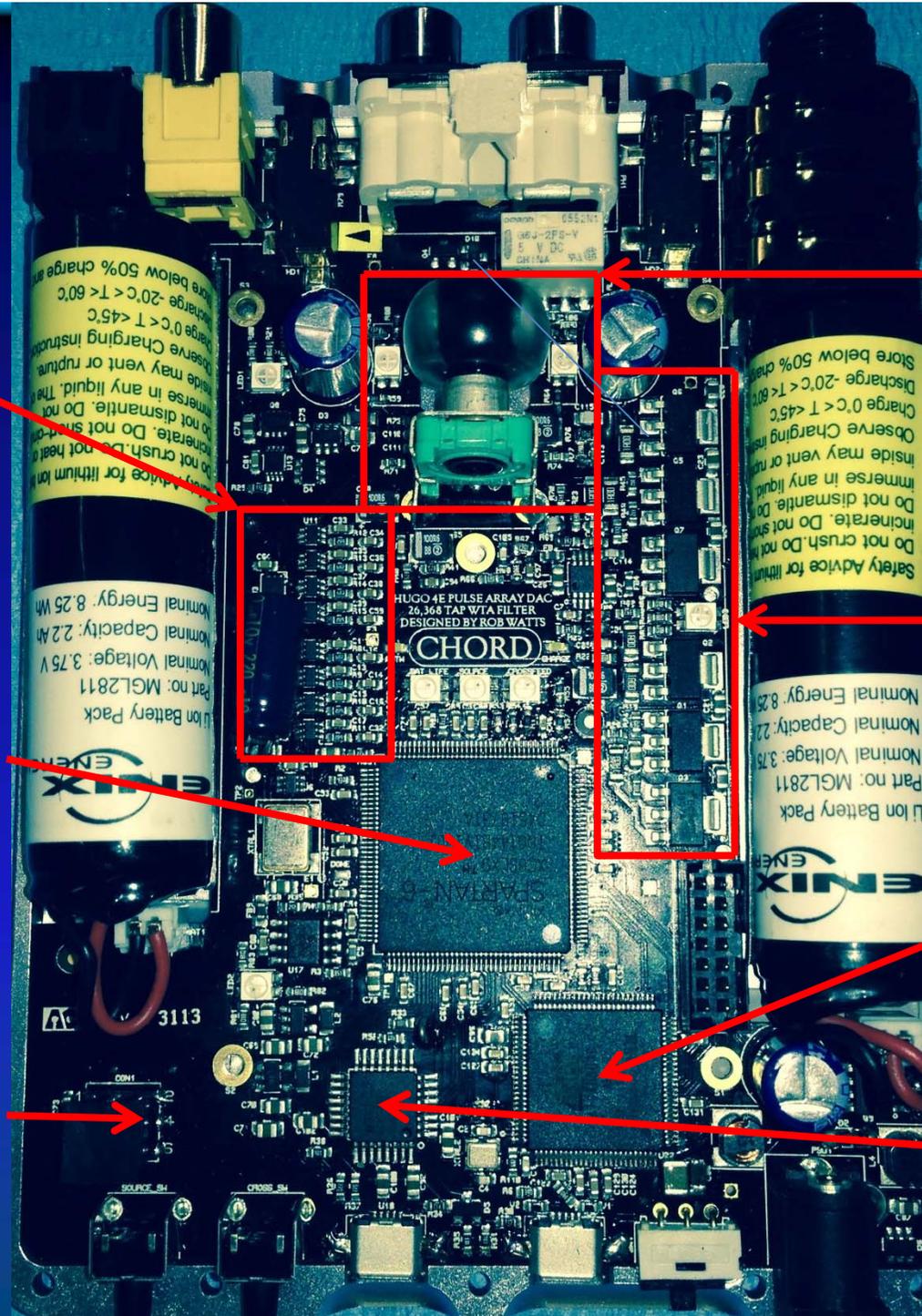
A2DP and APTX CSR Bluetooth module (underneath)

Digital volume control

Discrete Class A OP stage – ultra low distortion 0.0004% including DAC distortion

HD USB decoding

SD USB decoding



# And finally...

Thank-you for reading this presentation.



Copyright Rob Watts 2014