Effects in High sample Rate Audio Material

M.N.Harris[1], R. Kelly[2], D.A.McLeod[2], M.J.Story[2]

A number of new, high sample rate, formats are becoming available for the recording engineer to use. These include 24 bit PCM formats at 88.2 kS/s, 96 kS/s, 176.4 kS/s and 192 kS/s and DSD, the single bit 2.822 MS/s format. Delivery media using some of these formats have been announced or are planned – the SACD has been announced and demonstrated by Sony and Philips, using DSD, and the DVD will support both 96 kS/s and 192 kS/s. Discs based on the DVD format but using a 24 bit 96 kS/s, two channel, format are already on the market in the USA, with consumer equipment available to play them. 88.2 kS/s and 176.4 kS/s do not have release media, although 176.4 kS/s is being used to record and store material for subsequent conversion to both DSD and 96 kS/s.

Recording engineers, and many musicians – particularly in the classical area – are becoming aware that material recorded and edited using these higher sample rates has some attractive qualities. Current theory on how human hearing works has so far been unable to explain the basis for these qualities, but they are none the less easy to demonstrate.

The differences can be easily demonstrated using the equipment set up shown in figure 1 - this is a set up that has credibility in the mixing and mastering professions. It can switch rapidly (A/B comparisons) between the analogue source material and the digital format, and it can change reasonably rapidly (a few seconds) between one sample rate and another while the system is set to "analogue". Changing between DSD and a PCM mode takes a little longer (15 secs or so). Although not perfect, these speeds are sufficient to do some interesting comparisons. The analogue circuitry in the ADC and the DAC are not affected by the format changes – they are both low distortion (-120 dBFs or lower) and have a 3dB bandwidth of between 150 and 200 kHz. Any changes that are heard (from one format to another) are caused by:

either the digital signal processing,
or the presence of high frequency signals in the output affecting the playback system.

This last is a significant issue for DSD, and is addressed later.

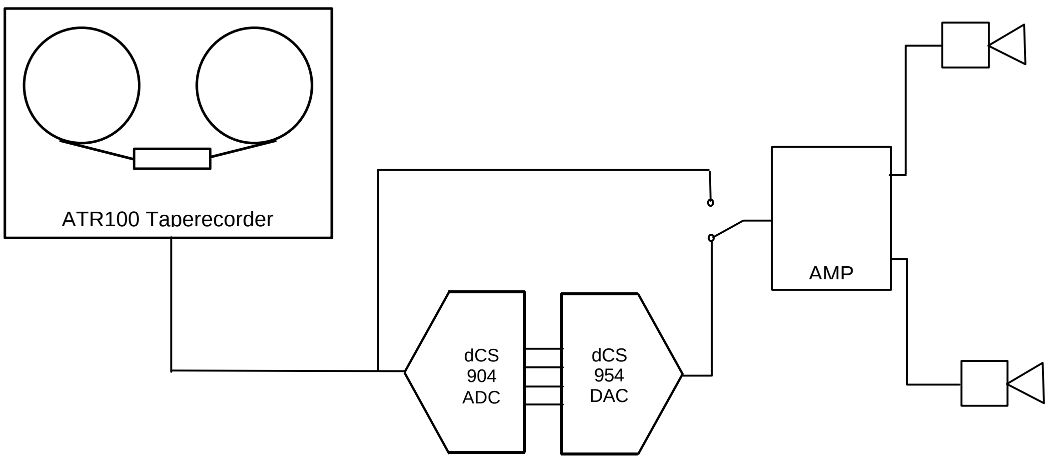


Figure 1 – High Sample Rate Demonstration setup

---

[1] With Chop'em Out mastering studios, London, UK

[2] with *dCS* Ltd, Great Chesterford, UK

The use of a tape source for the analogue feed allows easy repetition of critical phrases. Although there may be some loss of quality in making the original analogue recording, this is small, and the comparison we look for is whether the digital format sounds the same as the analogue tape or not. We do not seek to comment on the quality of the original recording[3].

Format	Observation
44.1 kS/s, 16 bit	(current format)
96 kS/s, 24 bit	• much less "busy signal" break up – very good quality • better separation of reverberation/room acoustics from instrument output • bass better balanced • percussion (particularly cymbals) better • some stereo image formation
192 kS/s, 24 bit	• no "busy signal" break up – excellent quality • very good separation of reverberation/room acoustics from instrument output • **but** - bass can appear light and slightly out of time • **and** - stereo image can be strong but widened (1.5 times?)
DSD	• detailed comparisons not yet performed on enough systems, but well liked by (classical) artists after sessions • no "busy signal" break up • very good separation of reverberation and room acoustics • no observations on bass so far • strong stereo image formation, no observations on width so far

The two highlighted effects with 192 kS/s are interesting. The bandwidth of 192 kS/s far exceeds the normal bandwidth attributed to human hearing, even using gentle roll off filters, so it is hard to believe it is related to frequency response. The stereo image widening is a very strong effect, observed by virtually all listeners, when a comparison with the analogue source is available. The bass problem is commented on by experienced listeners, and shows up particularly well on multi-mic'd rock music – it can enable 192 kS/s material to be identified in absence of the analogue source material for comparison. The bass response of the ADC and DAC units in question is flat down to 1 Hz, so there is no question of it being due to bass roll off there. We began to wonder:

- were the effects related? and
- were they related to pre-ringing in the digital filtering?

To investigate these effects further, we set up an ADC and DAC pair with near minimum phase filters[4], and subjected these to listening tests. The filters were asymmetrical FIR types, rather than IIRs, to avoid truncation effects. Such filters tend to be longer than symmetrical FIRs or IIRs, and consequently we were rather more DSP power limited than we had wanted. Because of this, the experiment was restricted to establishing whether or not there was a relationship, and whether or not minimum phase was important. We did not seek to produce commercially saleable filters. The frequency responses of the individual filters used, and the impulse responses of the combinations tested are given below.

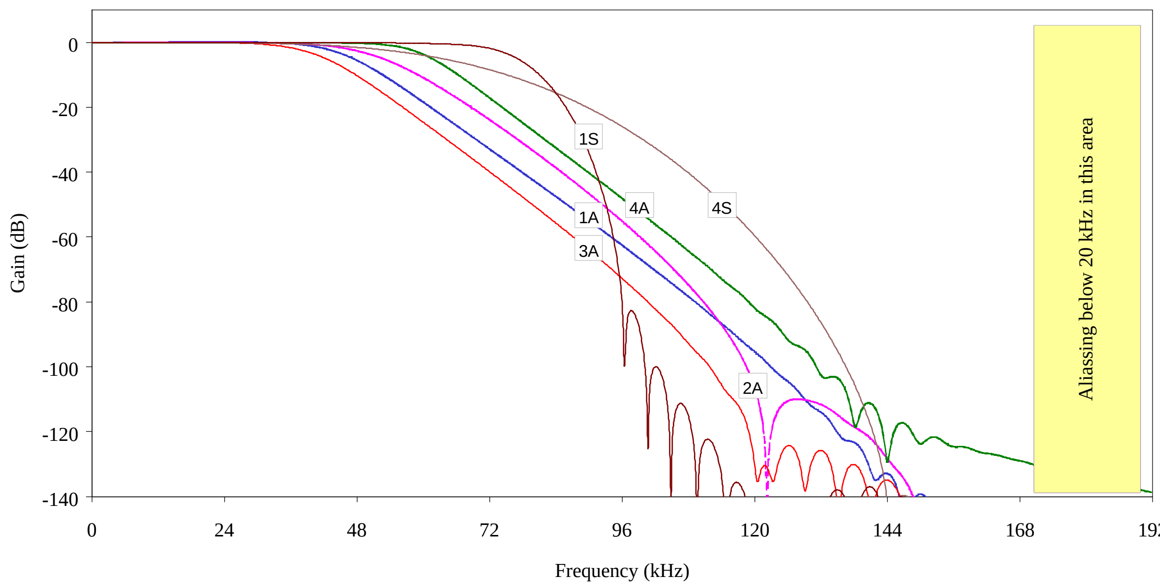
Each unit (ADC or DAC) had four filters. Six combinations of asymmetrical filters (not all combinations) and two symmetrical filter combinations were tested on a high quality system. The combined impulse response of the final digital filters of the eight combinations tested is shown.

---

[3] although we have been lucky in that we have been able to use some very good tapes.

[4] Minimum phase filters have no pre-ringing before a transient. Linear phase filters ring equally before and after a transient.

Symmetric (S) and Asymmetric (A) ADC Filters Used - Frequency Responses
Fs is 192 kS/s



The summarised listening results are given in the table below:

Filter Combination		Comments on Bass	Comments on Image
Combination 1 (ADC 1A, DAC 1A)	Asymmetrical	Lacking weight	widened
Combination 2 (ADC 1S, DAC 1S)	Symmetrical	Lacking weight	widened
Combination 3 (ADC 4A, DAC 4A)	Asymmetrical	Good	negligible shift
Combination 4 (ADC 4S, DAC 4S)	Symmetrical	Good	width okay, moved forward?
Combination 5 (ADC 4A, DAC 2A)	Asymmetrical	Light	a bit widened
Combination 6 (ADC 4A, DAC 3A)	Asymmetrical	Good	little shift
Combination 7 (ADC 3A, DAC 2A)	Asymmetrical	-	moved forward
Combination 8 (ADC 2A, DAC 2A)	Asymmetrical	OK	little shift

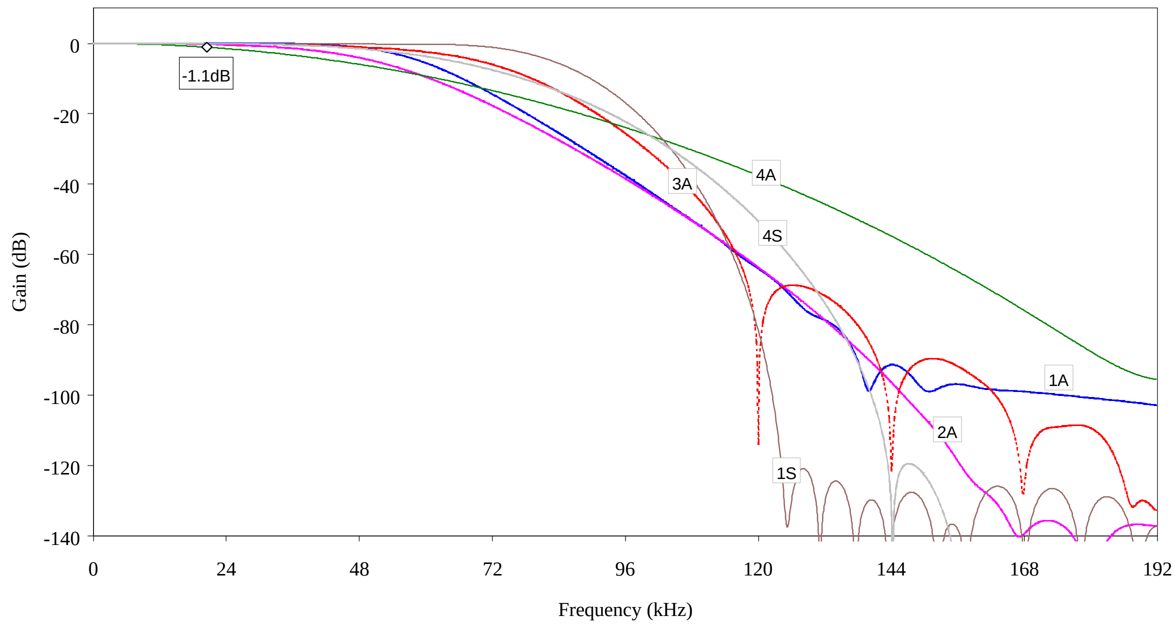
From the results above, and looking at the combination impulse responses, we conclude the following:

- the widening of the stereo image is related to the perceived problems in the bass (from correlations in the table) and
- filter impulse or transient response may be significant in correct image formation, and also proper bass perception.

DSD

Listening tests with DSD are a little more complex than PCM formats. DSD inherently has a high level of “out of band” noise, and this can cause significant problems in some playback systems. The noise can be reduced by low pass filtering from somewhere between (depending on the system) 30 kHz and 60 kHz, but then this filter becomes a part of the system. These hf reduction filters have not so far received much attention – DSD users so far prefer to find systems that are hf tolerant, and minimise the extra filtering required.

Symmetric (S) and Asymmetric (A) DAC Filters - Frequency Responses
Fs is 192 kS/s



When played back on a system with adequate hf tolerance, the results from DSD are very good. We have not so far done enough comparative listening tests (DSD vs PCM or DSD vs Analogue) to comment usefully on the differences. We have done some testing of different filters in the noise shaping loop in a DSD modulator. However, even these results are dependent on the monitoring system and useful comments cannot be made at this stage.

Combined ADC/DAC Impulse Responses
Fs is 192 kS/s