Audio Engineering Society
Convention Paper
Presented at the 112th Convention
2002 May 10–13 Munich, Germany

Enhanced Sigma Delta Structures for Super Audio CD Applications

Derk Reefman[1], and Erwin Janssen[1]

[1] *Philips Research, Eindhoven, Prof. Holstlaan 4, 5656 AA, Netherlands*

Correspondence should be addressed to Derk Reefman (derk.reefman@philips.com)

ABSTRACT

New Sigma Delta modulators (SDM's) topologies for use in Super Audio CD (SACD) applications are introduced, called Sigma Delta pre-correction (SDPC), which allow the generation of ultra-high quality DSD. Spurious peaks, which are known theoretically to exist in SDM's, are present at levels well below -165 dB, even if undithered. Already a slight amount of dither, will further reduce these signals to levels which are with standard numerical precision undetectable.

INTRODUCTION

The purpose of this paper is to give an overview of some novel SDM topologies, which have in common that they present intrinsic high linearity. The theoretical non-linearity of SDM modulators has been dwelled upon extensively, most notably in [1] and [2]. Though the practical relevance of the non-linearities which occur at an extremely low level in higher order SDM's can be doubted [3], it remains that some non-linearity is present in any SDM. The specific purpose of this paper is to clarify the *limited* importance of linearity in the SDM system. Though the SDM system is theoretically capable of dealing with signals in the frequency range up to 1.4 MHz, practical reality is that this enormous bandwidth is not used for signal representation. The huge bandwidth does, however, offer the possibility to do various kinds of digital signal processing *without* being hindered by the proximity of the Nyquist frequency. We will present pre-correction methods for achieving high linearity in a limited frequency region (for example, 0-80 kHz) while ignoring nonlinearity in the higher frequency bands. We will do this by first discussing the non-linearity in a SDM. Next the pre-correction method is presented. The performance improvement of the pre-correction method will be discussed in a realistic example. Finally conclusions are drawn.

NON-LINEARITY IN A SDM

To present a realistic situation, a spectrum of a SDM that is typically used in SACD applications is presented in Fig. 1. For the purpose of this discussion, this SDM has not been dithered. The input to this SDM has been a 4 kHz sine (-6 dB SACD amplitude). If we are interested in the baseband, extending from 0 to 20 kHz, the relevant distortion products are the 2'nd up to the 4'th component. From inspection of Fig. 1, it can be concluded that the distortion components are all at most -165 dB. The noise floor of this SDM is at -127 dB, resulting in a DR of about 120 dB[1]. It is also instructive to extend the region of interest to the band 0-80 kHz. Obviously, the noise floor is increasing steeply (in the case presented in Fig. 1, this increase is fifth order) causing the maximum Signal-to-Noise Ratio (SNR) to drop to about 90 dB in the band 0-40 kHz, and about 55 dB in the band 0-80 kHz. Any harmonic distortion component, however, is at a level at least below -95 dB. Clearly, any harmonic distortion component that we are dealing with in the broader sense of the audio band, is extremely small, and its importance for the perceptual audio quality can be doubted. In view of the fact that this SDM has not been dithered, it is clear that dithering will even further reduce these numbers. In fact, if this SDM is dithered to its maximum level (where it is just not overloaded) the distortion components in the audio band are all below -180 dB, only observable after 5000 coherent averages, and the components in the broader audio band are below -110 dB.

Still, the total amount of coherent power that is present in the dithered signal is significant. The amount of coherent power can easily be estimated if the actual noise is assumed to have no correlation with the signal. It appears that the total amount of coherent power which is present in Fig. 1, is about -10 dB. It is obvious that this power is mostly above 1 MHz; 99.99% of the coherent power is found in this high frequency area. The exact value of the frequency above which most of the correlated signal is found, is dependent on the signal which is input to the SDM; it will, however, never be very much lower than the quoted 1 MHz. It is beyond doubt, that the origin of these signals in the very high frequency area is in the non-linear behavior of the SDM. Indeed, if a triangular pdf dithered multi-bit quantizer is used in the noiseshaper, the high frequency components disappear. Thus, the coherent signal above 1 MHz can be considered in some sense to be distortion.

To judge whether these distortion components are harmful, we need to look at the full audio chain which is used to replay DSD in a typical SACD-capable player. Such a configuration is shown in Fig. 2. A typical DAC-chip (see *e.g.* [6] or [5]) contains the first 4 blocks displayed in Fig. 2. The digital filter in the path leading to the n-bit SDM is a crucial part, where most of the HF signal present in the DSD signal can be removed without any compromise. As an example, consider a filter that is designed according to the following criteria: pass-band: 0-100 kHz, flat within 0.01 dB; transition band 100 kHz - 900 kHz; stop band: 900 kHz - 1.4MHz, suppression -100 dB. This leads to a filter with only 22 taps, and thus does not pose any additional constraint in terms of hardware; the filters which are necessary to do proper upsampling from a low sample rate format to the required $m \cdot 64 f_s$, are much more demanding. Also, the digital LPF does not influence the impulse response of DSD [3], as the transition width is extremely large. It is clear, that the application of this filtering will lead to significant suppression of the high frequency components present in the original DSD stream. Still, the signal contains substantial amounts of HF, which is foremost white noise. The signal is then upsampled to a frequency that is used to perform the digital-to-analog conversion on. The SDM will noise-shape this signal into an n-bit signal, where n typically varies between 3 [6] and 5 [5]. It is this signal, which is converted to the analog domain. Due to the noise shaping process, which is intrinsic in modern, high-end DA convertors, and is the sole basis for their very high performance, some additional high frequency noise extending to frequency regimes well above 1 MHz is introduced. This noise is usually removed by an analog low pass filter of first or second order. This filtering is most often passive, and can thus be performed with exceptionally low distortion and intermodulation. In most SACD players, some additional filtering is provided, to reduce the amount of HF noise (which by then, is mostly due to the DSD signal) even further to levels well below -30 dB. It is important to remark, that the HF signal levels at which these additional filters need to operate are quite low due to the digital pre-filtering (which removed a very substantial amount of HF signal causing the total signal power to be substantially less than 1); hence, the linearity of the filters can be quite high and the filtering operation is performed without additional intermodulation products.

This example of a typical SACD signal path shows, that the non-linearity above 1 MHz is not important at all, and does not influence the signal quality. In fact, one can argue that these components are favorable. Because the total power of the SDM output is constant and equals 1, the power which is present in these high frequency tones causes the SNR in the lower frequencies to be higher than anticipated on basis of the linear noise transfer function. Hence, they contribute favorably to the dynamic range of an SDM. This discussion then leads to the question whether it would be possible to linearize a SDM in the important signal band, without bothering about its high frequency behavior.

[1] The SACD reference 0 dB level has been defined as -6 dB with respect to the level in the feedback path

PRE-CORRECTION

In order to have a system which demonstrates in a clear way the effects that we will study in this section, a third-order SDM has been designed. Such low order SDM's are notorious for their relatively bad signal properties [4]. The spectrum of the third order SDM that will be used in the sequel of this paper is shown in Fig. 3.

While this third-order SDM has a dynamic range of about 90 dB, its third harmonic is at a level of -104 dB. While this is still a rather respectable number, it is about 60 dB larger than the distortion component of the SDM shown in the previous section. The higher order harmonic distortion products are significant, too. Also in the broader signal band (0-80 kHz) the distortion components are larger. It should be remarked, that this type of SDM is not recommended for practical use.

When we model the SDM as a non-linear element $\Sigma\Delta$, its transfer characteristic can be written as:

$$\Sigma\Delta(x) = x + \alpha_2 x^2 + \alpha_3 x^3 + \dots \qquad (1)$$

Now, if we could create a signal $s(x)$ according to:

$$s(x) = x - \alpha_2 x^2 - \alpha_3 x^3 - \dots \qquad (2)$$

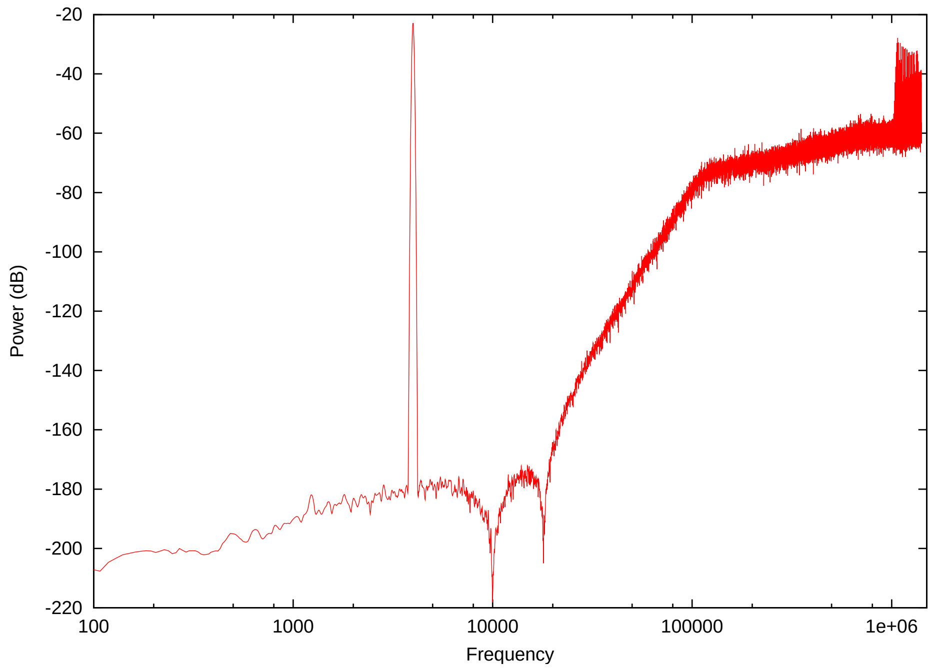

Fig. 1: A noise shaper which is typically used in SACD applications. The spectrum has been coherently averaged 100 times, and this has been repeated 10 times to obtain a power averaged spectrum.

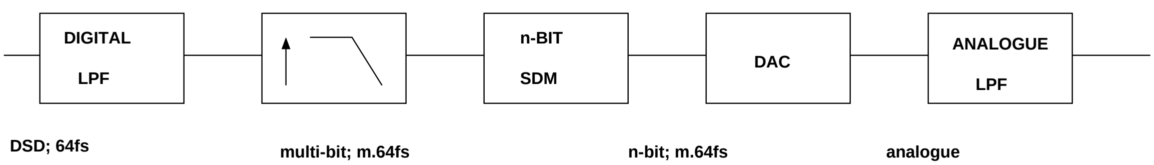

Fig. 2: Example of an audio chain found in an SACD-capable player. The DSD is first low pass filtered in the digital domain, followed by upsampling to m $\cdot$ f_s, typically, 128 or 256 f_s. This high-rate signal is then fed to an n-bit SDM, where n typically varies between 1.5 and 5. Finally, the analog output is passed through an analog low pass filter.

then the resulting output signal $f(v(x))$ would be given by:

$$\Sigma\Delta(s(x)) = x - 2\alpha_2^2 x^3 + \mathcal{O}(x^4) \qquad (3)$$

In other words, the second harmonic distortion component has been completely removed, and the third harmonic component has been substantially reduced (note, that for the low distortions we are dealing with, $\alpha_i \ll 1$). An estimate of the signal $s(x)$ can be obtained using the structure depicted in Fig. 4.

The topology of Fig. 4 operates as follows. The first SDM generates a signal, which is subtracted from the original input signal x. This difference signal v now contains all the distortion components which are generated by the SDM, and the uncorrelated noise which has been added to the signal because of the noise shaping. This signal v is now low-pass filtered in the filter F, which has, for example, a cut-off frequency of 100 kHz. This results in the signal denoted $F(v)$ in Fig. 4. Next, the original input signal x (after the appropri-

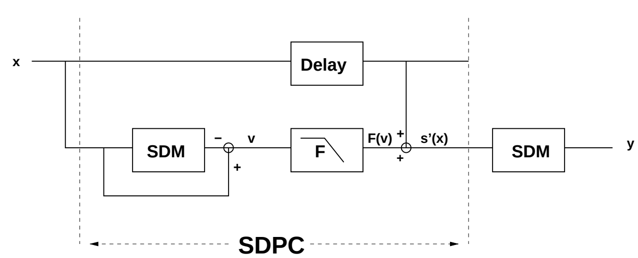

Fig. 4: Basic Sigma Delta Pre-Correction (SDPC) structure.

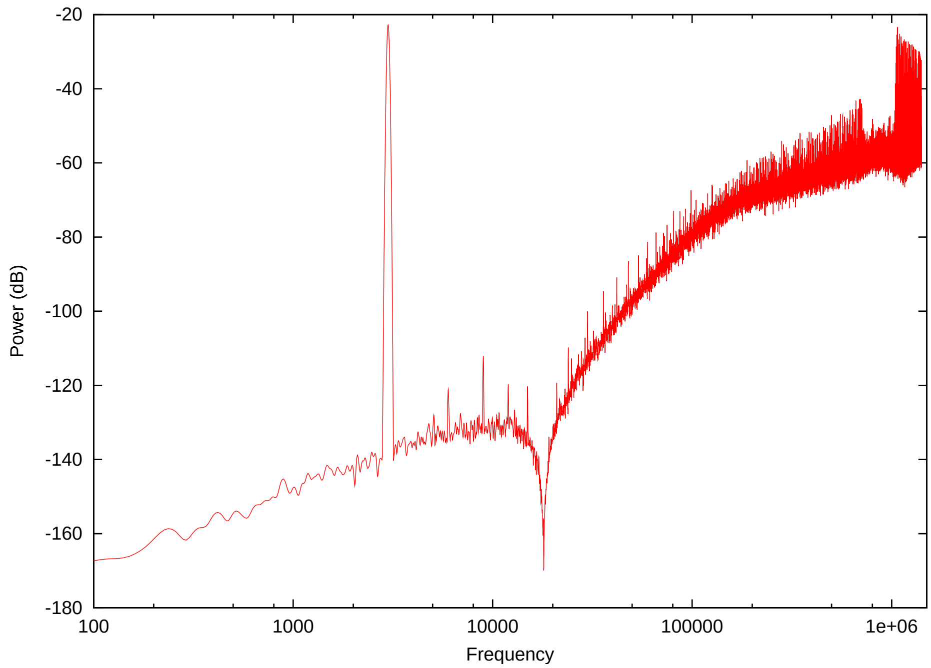

Fig. 3: Spectrum of the third order noiseshaper used in the analysis of the pre-correction technique. The input signal is a 3 kHz sine wave, -6 dB SACD. To obtain this spectrum, a series of 4 coherent averages and 10 power averages has been used.

ate delay to correct for the delay in the filter f) is added to $F(v)$, resulting in the signal $s'(x)$. While the filtering action has removed all HF noise, more in particular, it has removed the strong signals above 1 MHz, it has not removed any noise in the band below 100 kHz. Hence, the signal $s'(x)$ presents only an approximation to the signal $s(x)$ in Eq. 2. The signal $s'(x)$ is than input to a next SDM, which is identical to the SDM used to generate v, resulting in the final output signal y.

To gain some insight in the performance of this algorithm, which we will refer to as Sigma Delta pre-correction (SDPC), we have applied it to the third order SDM displayed in Fig. 3. The spectrum of the resulting signal y is displayed in Fig. 5 in the range 0-100 kHz. The huge suppression of the distortion components is clearly visible. Typically, the distortion has been reduced by about 20 dB. For higher frequencies, the suppression becomes less effective, even though the signal $s'(x)$ contains all distortion components unattenuated in the frequency regime. As always, there is a price to pay for this improvement in THD, which in this case is an increase in the noise floor by 3 dB. This is clear from inspection of Fig. 5, when one realizes that the corrected spectrum has been obtained using twice as many coherent averages which lowers the noise floor by 3 dB, and that the noise floor is identical to the noise floor of the uncorrected spectrum. This also corroborates the fact that this is white noise indeed; if it was correlated, it would result in a more than 3 dB increase. The origin of the increase of the noise floor is the fact that the signal $s'(x)$ still contains the quantization noise present in the low frequency range; the second SDM in the cascade adds its own quantization noise to it. Though not visible in Fig. 5, the high frequency signals above 1 MHz are completely unchanged using the new topology, which is expected on basis of the absence of correction components in the signal $s'(x)$.

SDPC AND DITHER

To appreciate the effect of SDPC, it is also instructive to study the combined action of dither and pre-correction. To that end, we have applied a dither level of 0.1 (the SDM starts overloading at levels of 0.8) to the SDM.

Spectra of the original SDM, and the SDPC spectrum are displayed in Fig. 6. Also in this case, the suppression of the distortion components is at least 22 dB in the band 0-20 kHz; in fact, even after 64 coherent averages, no distortion components can be observed. Note, that distortion has decreased to levels below -135 dB! Hence, the combined action of small amounts of dither, and the pre-correction technique result in extremely low distortion figures. Again, the reduced distortion suppression for higher frequencies is visible; for example in the region above 40 kHz, the suppression is typically only 8-10 dB.

While the higher harmonics are suppressed less than the lower harmonics, which is shown by Eq. (3), this does not fully explain the reduced suppression. Another origin of this reduced suppression for higher frequencies lies in the fact that the phase characteristic of the SDM used here is not straight for frequencies above 20 kHz. This results in some phase distortion, which is not accounted for in the pre-correction technique according to Fig. 4. To obtain an estimate of the significance of these errors, consider a single harmonic $h(\omega t) = A \sin(\omega t)$, which is positioned around 50 kHz. The

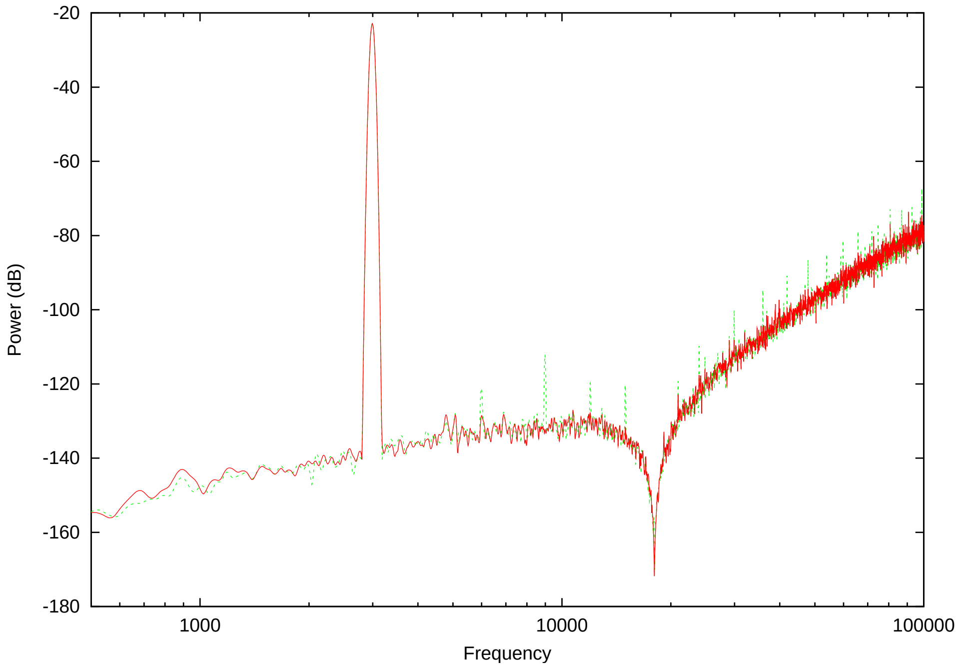

Fig. 5: Spectra of the original SDM (green), and its implementation according to Fig. 4 (red). The spectrum of the original SDM has been obtained using 4 coherent averages and 10 power averages; the other using 8 coherent averages and 10 power averages. The fact that the noise floors of the spectra coincide precisely illustrates the 3 dB loss in SNR due to SDPC.

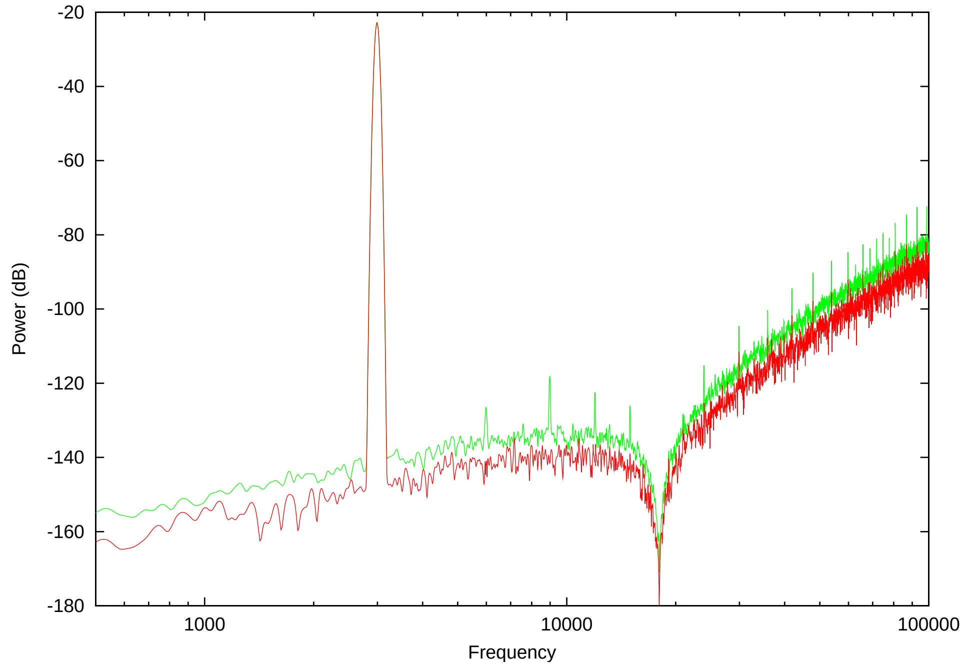

Fig. 6: Spectra of the original (dithered) SDM (green), and its implementation according to Fig. 4 (red) using the same dither. The spectrum of the original SDM has been obtained using 8 coherent averages and 10 power averages; the other using 64 coherent averages and 5 power averages.

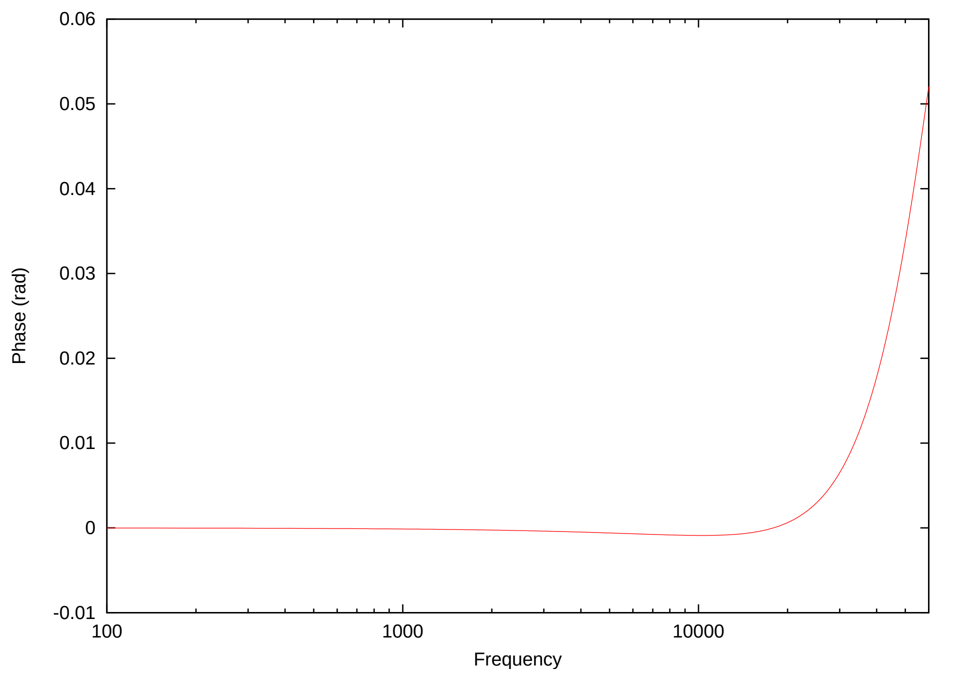

Fig. 7: Phase characteristic of the signal transfer function of the third order SDM used in this paper.

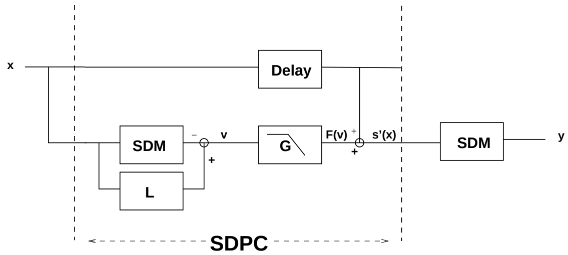

Fig. 8: Improved pre-correction structure. By cascading the Sigma Delta Pre-Correction structure (SDPC) n times, n harmonics can be removed.

absence of phase correction Δ will cause incomplete cancellation of the harmonic; a residual power of $4A^2\Delta^2$ will remain. In this case, this results in a maximum power reduction of the harmonic by only 14 dB.

An improved pre-correction technique is therefore displayed in Fig. 8. In this diagram, the phase error introduced by the SDM, is corrected for by the filter L. Another improvement can be obtained by cascading the structures displayed in Figs. 4 and 8. In a non-cascaded structure, the cancellation of lower order terms, causes the generation of higher order terms, albeit of much lower amplitude, as can be concluded from Eq. (3). These new, higher order terms, can in turn be cancelled in exactly the same way as the lower order ones were cancelled, resulting in cascading the structure in Fig. 8.

PERFORMANCE OF A REALISTIC SDM WITH SDPC

To end with a realistic situation, and to show how SDPC also suppresses DC tones, a standard fifth order SDM has been designed, with a SNR of 118 dB over 0-20 kHz.

As illustrated in Fig. 9, harmonic distortion levels of this SDM in the phase-corrected SDPC structure are reduced to well below -185 dB if undithered, which amounts to an improvement of about 35 dB compared to 20 dB improvement with the standard SDPC. If the SDM is slightly dithered, the distortion levels drop to much deeper levels, which numerically appeared to be unaccessible (*i.e.*, below -220 dB). Also, distortion levels at higher frequencies are reduced more compared to the standard SDPC algorithm. As with the uncorrected SDM, the SNR in the baseband (0-20 kHz) is slightly reduced from 118 dB to about 115 dB (no dithering) or 114 dB (with dithering).

The effects of a DC input to the SDPC system are illustrated in Fig. 10. As input to this system, a DC value of 1./1024 has been applied, which results in a tone around 5.5 kHz. The SDM has not been dithered.

In the spectrum of Fig. 10, a tone can be observed with an amplitude of about -145 dB. Application of the pre-correction algorithm, in its basic form, reduces this amplitude to about -165 dB. If a small amount of dithering (RPDF with amplitude 0.05) is applied, which is much less than the maximum allowed amount of dither (0.4 RPDF), the amplitude of the tone cannot be observed after 256 coherent averages, indicating that the tone is at least less than about -175 dB. Also application of the improved SDPC results in values for spurious signals that are not easily accessible numerically.

CONCLUSIONS

The new techniques introduced in this paper, coined Sigma Delta pre-correction (SDPC) techniques, have been shown to be able to reduce distortion levels by a significant amount of 20 dB. Even the simplest pre-correction technique is successful for both low and high order sigma delta modulators. As a result, spurious signals in realistic SDM configurations reaching about 115 dB of SNR, are at an extremely low level of -165 dB in the baseband, and at levels of about -80 dB at 100 kHz. Using more sophisticated SDPC techniques, these levels can be reduced even further. For example, with a fifth order modulator, spurious signals drop to below -185 db in the baseband upon using a phase-corrected SDPC technique.

The high frequency content of a SDM signal is left unaltered, containing significant amounts of correlated components. It has been clarified, however, that in practical situations these HF components are harmless.

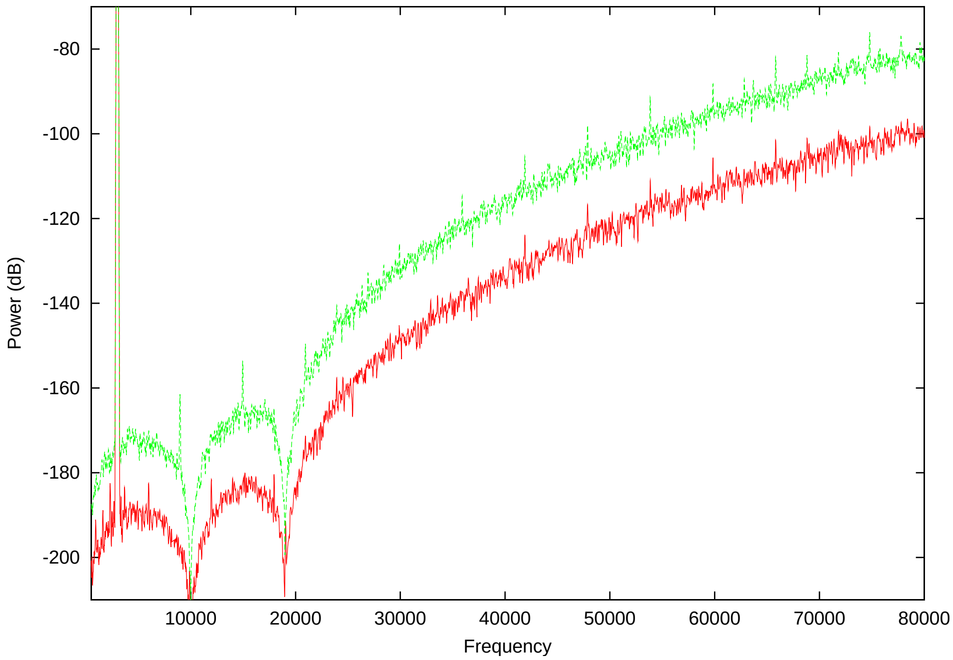

Fig. 9: Fifth order SDM, with a 3 kHz input of -6 dB. The uncorrected spectrum (green) has been obtained after 16 coherent and 10 power averages; the corrected spectrum (red) after 2048 coherent and 10 power averages.

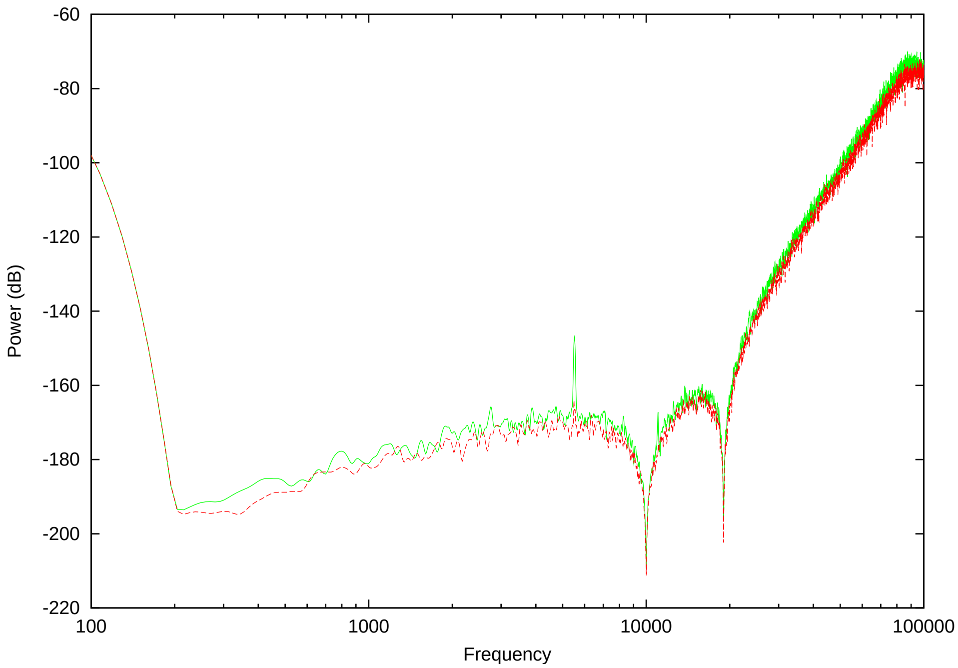

Fig. 10: Fifth order SDM, with a DC input of 1/1024. The uncorrected spectrum (green) has been obtained after 4 coherent and 10 power averages; the corrected spectrum (red) after 32 coherent and 10 power averages.

REFERENCES

[1] Lipshitz, S.P., and Vanderkooy, J. 'Why professional 1-Bit Sigma Delta Conversion is a bad idea', presented at the 109'th AES, Los Angeles (2000), preprint 5188.

[2] Lipshitz, S.P., and Vanderkooy, J. 'Why 1-Bit Sigma Delta Conversion is unsuitable for high quality Applications', presented at the 110'th AES convention (2001), Amsterdam. Preprint 5395.

[3] Reefman, D., and Nuijten, P.A.C.M., 'Why Direct Stream Digital is the best choice as a digital audio format', presented at the 110'th AES convention (2001), Amsterdam. Preprint 5396.

[4] Norsworthy, S.R., Schreier, R., and Temes, G.C. 'Delta-Sigma Data Convertors', IEEE Press (1996).

[5] Nakao, S., Terasaw, H., Aoyagi, F., Terada, N., and Hamasaki, T. 'A 117dB D-Range Current-mode Multi-bit Audio DAC for PCM and DSD Audio Playback', presented at the 109'th AES, Los Angeles (2000), preprint 5190.

[6] Adams, B., Nguyen, K., and Sweetland, K. 'A 116 dB SNR Multi-bit noise shaping DAC with 192 kHz sample rate', presented at the 106'th AES, Munich (1999), preprint 4963.