Reducing the Complexity of Digital Delta Sigma Modulators using Error Masking

Michael Peter Kennedy FIEEE

Università degli Studi di Pavia
03 December 2009
“Although the number of commercially deployed DDSMs far exceeds that of analog ΔΣ modulators, most of the published ΔΣ modulator analyses apply only to analog ΔΣ modulators.

Interestingly, most of these analyses do not apply or even readily extend to the case of DDSMs”
• Part I
  The ideal DDSM
  Signal processing assumptions
  Application of DDSMs: The oversampled DAC

• Part II
  The real DDSM
  Error masking
  Reduced complexity DDSMs 1: MASH architecture

• Part III
  Reduced complexity DDSMs 2: Nested DDSM architecture
Part I

- The ideal DDSM
- Signal processing assumptions
- Application of DDSMs
• A bandlimited digital signal \( x \) (\( n \) bits) is requantized to a shorter word \( y \) (\( m \) bits)
• The additive quantization noise \( e \) is highpass filtered for later removal by a lowpass filter
• $x$ is bandlimited

• $y$ includes highpass filtered quantization noise

• Quantization noise can be removed by lowpass filtering
The ideal DDSM

- $x$ is bandlimited
- $y$ includes highpass filtered quantization noise
- Quantization noise can be removed by lowpass filtering
• We assume that the Classical Model of Quantization (CMQ) applies...
• $e$ is statistically independent of $x$
• $e$ is uniformly distributed in $[-\Delta/2, +\Delta/2]$
• $e$ is stationary with a flat power spectrum
“CMQ can be applied when the quantizer input traverses several quantization levels between two successive samples”
The ideal DDSM

- The output depends on the signal $x$ and the quantization noise $e$
  \[ Y(z) = STF(z) \cdot X(z) + NTF(z) \cdot E(z) \]

- The signal is unaffected by the DDSM
  \[ STF(z) = 1 \]

- The quantization noise is highpass filtered
  \[ NTF(z) = (1-z^{-1})^l \]
The ideal DDSM

- The output comprises the signal plus the filtered quantization noise

\[ Y(z) = X(z) + E(z) (1 - z^{-1})^l \]
The ideal DDSM

- The noise scales as $f^{2l}$ at low frequencies

$$\left| E(z) (1-z^{-1})^l \right|^2 = \left( \frac{\Delta^2}{12} \right) 2^{2l} \sin^2 \left( \pi f / f_s \right)$$

$$\approx \left( \frac{\Delta^2}{12} \right) \left( 2 \pi f / f_s \right)^{2l} \text{ when } f \ll f_s$$
The ideal DDSM

- $x$ is bandlimited

- $y$ includes highpass filtered quantization noise

- quantization noise can be removed by lowpass filtering
Applications of DDSMs

- Example: Oversampled DAC with noise shaping
Oversampled DAC with noise shaping

- Digital input is oversampled with an interpolation filter
- Oversampled signal is requantized by a DDSM, introducing quantization noise
- Quantization noise is highpass filtered by the DDSM
- Shaped quantization noise is attenuated by the (lowpass) reconstruction filter
Oversampled DAC with noise shaping

Interpolation Filter

\[ u_1[n] \rightarrow \text{IF} \rightarrow u_2[n] \rightarrow \text{DDSM} \rightarrow m \rightarrow u_3[n] \]

Noise-Shaping Loop

\[ u_3[n] \rightarrow \text{DEM} \rightarrow \text{DAC} \rightarrow u_4(t) \rightarrow \text{LPF} \rightarrow u_5(t) \]
Oversampled DAC with noise shaping

\[ u_1, u_2, u_3, u_5 \]

\[ f_b, f_N, OSR f_N \]
Oversampled DAC with noise shaping

• By using DDMS, the filter rolloff and DAC linearity specifications are relaxed

• Oversampling simplifies the design of the reconstruction filter

• DSM simplifies the DAC linearity problem
Oversampled DAC with noise shaping

Stereo, 24-Bit, 192 kHz Multibit ΣΔ DAC

AD1852*

FEATURES
5 V Stereo Audio DAC System
Accepts 16-Bit/18-Bit/20-Bit/24-Bit Data
Supports 24 Bits, 192 kHz Sample Rate
Accepts a Wide Range of Sample Rates Including:
  - 32 kHz, 44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz, and
  - 192 kHz
Multibit Sigma-Delta Modulator with “Perfect
  Differential Linearity Restoration” for Reduced Idle
Tones and Noise Floor
Data-Directed Scrambling DAC—Least Sensitive to
  Jitter
Differential Output for Optimum Performance
117 dB Signal-to-Noise (Not Muted) at 48 kHz Sample
  Rate (A-Weighted Mono)
114 dB Signal-to-Noise (Not Muted) at 48 kHz Sample
  Rate (A-Weighted Stereo)
117 dB Dynamic Range (Not Muted) at 48 kHz Sample
  Rate (A-Weighted Mono)
114 dB Dynamic Range (Not Muted) at 48 kHz Sample

Flexible Serial Data Port with Right-Justified, Left-
  Justified, I²S-Compatible and DSP Serial Port Modes
28-Lead SSOP Plastic Package

APPLICATIONS
Hi End: DVD, CD, Home Theater Systems, Automotive
  Audio Systems, Sampling Musical Keyboards, Digital
Mixing Consoles, Digital Audio Effects Processors

PRODUCT OVERVIEW
The AD1852 is a complete high performance single-chip stereo
digital audio playback system. It is comprised of a multibit sigma-
delta modulator, digital interpolation filters, and analog output
drive circuitry. Other features include an on-chip stereo attenuator
and mute, programmed through an SPI-compatible serial control
port. The AD1852 is fully compatible with all known DVD
formats including 192 kHz as well as 96 kHz sample frequen-
cies and 24 bits. It also is backwards compatible by supporting
50 μs/15 μs digital de-emphasis intended for “Redbook” compact
discs, as well as de-emphasis at 32 kHz and 48 kHz sample rate.
Oversampled DAC with noise shaping
Part II

- The real DDSM
- Error masking
- Reduced complexity DDSMs 1: MASH architecture
The real DDSM

- The output depends on the signal $x$ and the quantization noise $e$

$$Y(z) = STF(z) X(z) + NTF(z) E(z)$$

where $STF(z) = 1$ and $NTF(z) = (1-z^{-1})^l$

- Consider the simplest case $l=1$

$$Y(z) = X(z) + (1-z^{-1}) E(z)$$
A 1st order DDSM can be implemented using a digital accumulator
The real DDSM

- 1st order Error Feedback Modulator (EFM1)
- $x$ is the input; $y$ (the carry out) is the output; $s$ (the value stored in the register) is the state; $M$ is the modulus
• 1st order Error Feedback Modulator (EFM1)

\[ s[n+1] = (x[n] + s[n]) \mod M \]

\[ y[n] = Q(x[n] + s[n]) \]
The real DDSM

• 1st order Error Feedback Modulator (EFM1)

\[
y[n] = Q(x[n] + s[n])
= Q(x[n] - e[n-1])
= (x[n] - e[n-1]) + e[n]
= x[n] + (e[n] - e[n-1])
\]

\[
x[n] \quad u[n] \quad v[n] \quad Q(\cdot) \quad y[n]
\]

\[
u[n] \quad z^{-1} \quad e[n] \quad M\quad s[n]
\]
The real DDSM

- 1st order Error Feedback Modulator (EFM1)

\[ Y(z) = X(z) + E_1(z) (1-z^{-1})^1 \]
• 1st order Error Feedback Modulator (EFM1)

\[ Y(z) = X(z) + E_1(z) \ (1-z^{-1})^1 \]

(1, N, N, 1) denotes first order, N-bit input \( x \), N-bit quantization error output \( w \), and 1-bit carry output \( y \)
Higher order DDSM: MASH architecture

- Cascade first order stages and add an error cancellation network to realize

\[ Y(z) = X(z) + E_l(z) (1-z^{-1})^l \]

- This lowers the spectral envelope of the quantization noise at low frequencies
• MASH: Idealized power spectra (white noise with 2\textsuperscript{nd} and 3\textsuperscript{rd} order filters)

40 dB/decade

60 dB/decade
All modulators have the same wordlength $N$. 

**Exact 3$^{rd}$ order MASH architecture**
Exact 3\textsuperscript{rd} order MASH architecture

- 3\textsuperscript{rd} order MASH with exact cancellation

\[ Y_1(z) = X(z) + E_1(z) (1-z^{-1}) \]
\[ Y_2(z) = -E_1(z) + E_2(z) (1-z^{-1}) \]
\[ Y_3(z) = -E_2(z) + E_3(z) (1-z^{-1}) \]

\[ Y(z) = Y_1(z) + Y_2(z) (1-z^{-1}) + Y_3(z) (1-z^{-1})^2 \]
\[ = X(z) + E_3(z) (1-z^{-1})^3 \]
Exact 3\textsuperscript{rd} order MASH architecture

- All stages have the \textit{same} wordlength N
- The hardware requirement scales as 3N
- This is overkill…
Later stages have smaller wordlengths: \( L < M < N \).
Reduced complexity 3rd order MASH

- Successive stages have decreasing wordlengths

- The hardware requirement scales less than $3N$

- This is optimal...
• Why does it work?

• Interstage quantizations introduce errors

• But these small errors are “masked” (hidden) by the large error resulting from the last stage
• 3rd order MASH with interstage errors

\[ Y_1(z) = X(z) + E_1(z) (1-z^{-1}) \]
\[ Y_2(z) = -E_1(z) + E_{12}(z) + E_2(z) (1-z^{-1}) \]
\[ Y_3(z) = -E_2(z) + E_{23}(z) + E_3(z) (1-z^{-1}) \]

\[ Y(z) = Y_1(z) + Y_2(z) (1-z^{-1}) + Y_3(z) (1-z^{-1})^2 \]
\[ = X(z) + E_3(z) (1-z^{-1})^3 \]
\[ + E_{12}(z) (1-z^{-1}) + E_{23}(z) (1-z^{-1})^2 \]
The objective of the wordlength selection strategy is to mask the contributions of the interstage quantizers by hiding the spectral components due to $E_{23}$ and $E_{23}$ below $E_3$ or the noise floor, whichever is higher.
Error Masking Strategy

- For a quantized or dithered input, the noise floor is typically dominant.
- We need to hide the components $L_{12}$ and $L_{23}$ below the contributions of $L_3$ and the noise floor $L_{nf}$.

“Corner” frequency
Design Constraints (1)

• In the case of zeroth order shaped LSB dither, the level of the noise floor is:

\[ \mathcal{L}_{n0} = \frac{1}{12 \cdot (2^N)^2} \]

• The largest frequency at which the PSD of the dither is larger than the contribution from \( E_3 \) can be calculated as:

\[
\frac{1}{12} \left| 2 \sin(\pi f_0/f_s) \right|^6 = \frac{1}{12 \cdot 2^{2N}}
\]

which gives

\[ f_0 \approx \frac{1}{\pi \cdot 2^{N/3} \cdot \frac{f_s}{2}} \]

• As \( \mathcal{L}_{12} \) and \( \mathcal{L}_{23} \) are first- and second-order shaped, respectively, we need only check the constraint at the boundary (the “corner frequency” \( f_0 \)).
Design Constraints (2)

• Based on this idea, the constraints can be written as:

\[ L_{12} < L_{3@f_0} \]

\[ L_{23} < L_{3@f_0} \]

• Using \( \sin(\pi f_0/f_s) \approx \pi f_0/f_s \) for \( f_0 << f_s \),

we can approximate \( L_{12}, L_{23} \) and \( L_3 \) at low frequencies by:

\[ L_{12} \approx \frac{\Delta_{12}^2}{12} \cdot 2^2(\pi f/f_s)^2, \]

\[ L_{23} \approx \frac{\Delta_{23}^2}{12} \cdot 2^4(\pi f/f_s)^4, \]

\[ L_3 \approx \frac{1}{12} \cdot 2^6(\pi f/f_s)^6. \]
Design Constraints (3)

- Therefore, we can obtain:

\[
\frac{1}{2^{2M}} \cdot \frac{1}{12} \cdot \left( \frac{1}{2^{N/3}} \right)^2 < \frac{1}{12} \cdot \left( \frac{1}{2^{N/3}} \right)^6,
\]

\[
\frac{1}{2^{2L}} \cdot \frac{1}{12} \cdot \left( \frac{1}{2^{N/3}} \right)^4 < \frac{1}{12} \cdot \left( \frac{1}{2^{N/3}} \right)^6,
\]

which reduce to:

\[
M > \frac{2N}{3},
\]

\[
L > \frac{N}{3}.
\]
Hardware Requirements

• The accumulators consume most of the hardware in the DDSM
• The hardware consumption of the accumulators is proportional to their wordlengths
• We estimate the relative hardware consumption (RHC) of our RC MASH DDSM compared with the conventional MASH DDSM as:

\[
RHC_0 \approx \frac{N + \text{ceil}(0.67N) + \text{ceil}(0.33N)}{3N} \\
\approx \frac{2N}{3N} \times 100% ,
\]

• For large \( N \), \( RHC \approx 67\% \).
• The percentage hardware saving of the DDSM plus dither generator is less.
Design Example

- An design example of a dithered 22-bit MASH DDSM is shown. The optimum wordlengths of the first, second and third stages of the RC MASH DDSM are 22, 15, and 8, respectively.

Simulated PSD for the 22-bit DDSM with dither

Simulated PSD for the 22-15-8 bit DDSM with dither
The hardware requirements for a conventional 22-bit MASH 111 DDSM and a RC 22-15-8 DDSM with zeroth-order dither are shown.

The hardware consumption is reported as the number of flip-flops (FFs) and the number of 4-input look-up-tables (LUTs). The total equivalent gate (TEG) is given as well.

<table>
<thead>
<tr>
<th>MASH DDSM with dither</th>
<th>Hardware Consumption</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) 22 bit with zeroth-order dither</td>
<td>149</td>
<td>103</td>
</tr>
<tr>
<td>(2) 22-15-8 bit with zeroth-order dither</td>
<td>111</td>
<td>79</td>
</tr>
<tr>
<td>((2)/(1))%</td>
<td>74%</td>
<td>77%</td>
</tr>
</tbody>
</table>

If we subtract the hardware consumption for the dither block for both DDSMs, which is 324 TEG, our RC MASH DDSM has a relative hardware consumption of 68%, as predicted.
Conclusions

- Error masking can be used to reduce the complexity of digital delta-sigma modulators.

- We have developed an algorithm for calculating the optimum wordlengths for stages in a reduced complexity (RC) MASH 1-1-1-1.

- The RC DDSM can achieve similar spectral performance compared with a conventional DDSM but with up to 20% less hardware.


Part III

- Noise-shaping
- Error masking
- Quantized discrete-time signal
- Oversampling and noise shaping
- Nested DDSM preprocessor
- Nested $l_1$-$l_2$ MASH structures
- Conclusions
Noise-shaping by the DDSM

(a) Power spectrum of the input signal

(b) Power spectrum of the output signal
Key idea: Error Masking

• The number of bits at the quantized input of the DDSM determines the noise floor

• DDSM introduces additional quantization noise but it is shaped toward high frequencies

• By selecting the parameters of the DDSM appropriately, we can hide the in-band quantization noise power from the DDSM below the noise floor of the quantized input

• The degradation of the SNR is minimized if the contribution of the DDSM quantization noise is masked spectrally below the input noise
Key idea: Error masking

- Quantization noise (of DDSM) contribution is too high
- SNR in signal band is degraded

Diagram:
- Power Spectrum
- Quantization noise (of DDSM)
- Quantization noise (of input)
- "Corner" frequency
- $f_B/(f_s/2)$
Key idea: Error masking

- DDSM contribution is insignificant
- SNR in signal band is not degraded

Power Spectrum

Quantization noise (of DDSM)

Quantization noise (of input)

“Corner” frequency

$\frac{f_B}{(f_s/2)}$
• Consider signal $x$ with bandwidth $f_B$ sampled at $f_s$ and quantized using a $B$-bit quantizer with quantization interval $\Delta$

$$\Delta = \frac{X_{FS}}{M}$$

$$x_B = x + e_B$$

$$-\frac{\Delta}{2} \leq e_B \leq +\frac{\Delta}{2}$$
Classical Model of Quantization (CMQ)

$$p(e_B) = \begin{cases} \frac{1}{\Delta}, & -\frac{\Delta}{2} \leq e_B \leq +\frac{\Delta}{2} \\ 0, & \text{otherwise} \end{cases}$$

$$e_{B_{\text{rms}}} = \int_{-\Delta/2}^{+\Delta/2} p(e_B)e_B^2 \, de_B$$

$$= \frac{\Delta^2}{12}$$
Quantized Discrete-Time Signal

- In-band signal and quantization noise power are given by

\[
N_{e_B}(f_B) = \frac{\Delta^2}{12\text{OSR}} \quad S_x(f_B) = \frac{2^{2B} \Delta^2}{8}
\]

- Signal-to-noise ratio (SNR) defined by

\[
\text{SNR} = \frac{S_x(f_B)}{N_{e_B}(f_B)} = 2^{2B} \left(\frac{12}{8}\right) \text{OSR}
\]
Quantized Discrete-Time Signal

- In dB

\[ \text{SNR}_{dB} = 6.02B + 1.76 + 3.01 \log_2(\text{OSR}) \]

- The effective number of bits (ENoB) of a signal with \( \text{SNR}_{dB} \) is defined by

\[ \text{ENoB} = \frac{\text{SNR}_{dB} - 1.76}{6.02} \]

\[ \text{ENoB} = B + 0.5 \log_2(\text{OSR}) \]

- ENoB increases by 0.5 bits for every factor of two increase in the OSR
First Order Error Feedback Modulator

\[ x[n] \in [-b, +b] \]

\[ y[n] = Q(v[n]) = \begin{cases} 
+ b, & v[n] \geq 0 \\
- b, & v[n] < 0 
\end{cases} \]

\[ Y(z) = X(z) + (1 - z^{-1}) E(z) \]
Estimation of SNR by simulation

- Extract sinusoid using DFT on time series at signal frequency
- Subtract sinusoid from time series, leaving noise components
- Calculate FFT of signal and noise
- Integrate power spectra up to the number of samples corresponding to the desired bandwidth
$Y(z) = X(z) + (1 - z^{-1})^l E_l(z)$
Consider signal $x$ with bandwidth $f_B$ sampled at $f_s$ and quantized using a $B$-bit quantizer producing output $x_B$ which is applied to an $l^\text{th}$ order DDSM

$$Y(z) = \text{STF}(z) X_B(z) + \text{STF}(z) E_B(z) + \text{NTF}(z) E_Q(z)$$

In a typical $l^\text{th}$ order MASH DDSM

$$\text{NTF}(z) = (1 - z^{-1})^l$$

Substituting $z = e^{j2\pi f/f_s}$

$$|Y(e^{j2\pi f/f_s})|^2 = |X(e^{j2\pi f/f_s})|^2 + |E_B(e^{j2\pi f/f_s})|^2$$
$$+ |1 - e^{-j2\pi f/f_s}|^{2l} |E_Q(e^{j2\pi f/f_s})/f_s|^2$$
• The signal and noise power in the frequency band \([-f_B, f_B]\) are given by

\[
S_y(f_B) = \int_{-f_B}^{f_B} |X(f)|^2 df = S_x(f_B)
\]

\[
N_y(f_B) = \int_{-f_B}^{f_B} \left( |E_B(f)|^2 + |\text{NTF}(f)|^2 |E_Q(f)|^2 \right) df = N_{e_B}(f_B) + N_{e_Q}(f_B)
\]

• We have that

\[
N_{e_B}(f_B) = \frac{\Delta^2}{12\text{OSR}}
\]
The DDSM quantization noise power in the signal band is given by

\[ N_{eQ}(f_B) = \frac{2}{f_s} \int_0^{f_B} \left| 1 - e^{j 2\pi f / f_s} \right|^{2l} e_Q^{2} \text{rms} \, df \]

If \( f_B << f_s \), then \( \sin(\pi f / f_s) \approx (\pi f / f_s) \) and simplifying yields

\[ N_{eQ}(f_B) = \frac{\Delta_Q^2}{12} \left( \frac{\pi^{2l}}{2l + 1} \right) \left( \frac{1}{\text{OSR}} \right)^{2l+1} \]
The SNR at the output of the DDSM is given by

\[
\text{SNR} = \frac{S_x(f_B)}{N_e_B(f_B) + N_e_Q(f_B)}
\]

Substituting appropriate values and simplifying yields

\[
\text{SNR} = 2^{2B} \left( \frac{12}{8} \right) \frac{1}{\text{OSR}} \left[ 1 + \frac{\Delta_Q^2}{\Delta^2} \left( \frac{1}{2l+1} \right) \left( \frac{\pi}{\text{OSR}} \right)^{2l} \right]
\]

DDSM has negligible impact on the SNR if

\[
\frac{\Delta_Q^2}{\Delta^2} \left( \frac{1}{2l+1} \right) \left( \frac{\pi}{\text{OSR}} \right)^{2l} \ll 1
\]
• Define $\Delta E\text{NoB}$ as the reduction in $E\text{NoB}$ due to noise-shaping
MASH 1-1-1 (DDSM3) Output Spectrum

- OSR=256
- l=3
- B=20
- Theoretically $\Delta E\text{NoB}=0.31$
- From simulation $\Delta E\text{NoB}=0.30$
Observation 1

- The DDSM reduces the ENoB compared to oversampling alone!
Observation 2

- ...but it also reduces the number of output bits... that’s why we’re using it!
• Is there another way to get the same $\Delta E\text{N}_0\text{B}$ and the same number of output bits, but with less hardware?
Sample target specification

- 20-bit input
- OSR=256
- 3rd order DDSM
- ENoB=24
Consider signal $x_{B1}$ with bandwidth $f_B$ sampled at $f_s$ and quantized using a $B_1$-bit quantizer producing output $x_{B2}$.

$x_{B1}$ is split

- Lower $B_{11}$ bits applied to $l^\text{th}$ order DDSM
- Upper $B_1-B_{11}$ bits combined with DDSM output to form $x_{B2}$
• Compare $ENoB$ of $x_{B1}$ with $x_{B2}$

• In the Z-domain, we can write

$$X_{B1}(z) = X(z) + E_{B1}(z)$$
$$X_{B2}(z) = X(z) + E_{B1}(z) + E_{\text{nested DDSM}}(z)$$

• Signal and noise power given by

$$S_x(f_B) = \frac{2^{2B_1} \Delta^2}{8}$$

$$N_{x_{B2}}(f_B) = N_{eB1}(f_B) + N_{e\text{nested DDSM}}(f_B)$$
• Quantization noise floor is the same and given by

\[ N_{eB1}(f_B) = \frac{\Delta^2}{12\text{OSR}} \]

• Quantization noise power due to nested DDSM given by

\[ N_{e\text{nested DDSM}}(f_B) \approx \frac{\Delta^2 Q}{12\text{OSR}} \left( \frac{2^{2B_{11}}}{2l_1 + 1} \right) \left( \frac{\pi}{\text{OSR}} \right)^{2n_1} \]

• Signal-to-noise ratio given by

\[
\text{SNR} = 2^{2B_1} \left( \frac{12}{8} \right) \text{OSR} \left[ \frac{1}{1 + \left( \frac{2^{2B_{11}}}{2l_1 + 1} \right) \left( \frac{\pi}{\text{OSR}} \right)^{2l_1}} \right]
\]

• Nested DDSM has negligible effect on SNR (and ENoB) if

\[
\left( \frac{2^{2B_{11}}}{2l_1 + 1} \right) \left( \frac{\pi}{\text{OSR}} \right)^{2l_1} \ll 1
\]
**ENoB of oversampled input** $x_{B1}$

- OSR=256
- B=20
- ENoB=24

**Power Spectrum**

- **Quantization noise (of input)**

![Power Spectrum Graph]

- Power (dB)
- Frequency ($f/(f_s/2)$)
- Frequency ($f_B/(f_s/2)$)
Quantization noise (of input)
Quantization noise (of nested DDSM)

- OSR=256
- $l_1=1$
- $B_1=20$
- $B_{11}=8$
- $B_2=13$
- $\text{ENoB}=23.28$
ENoB of $x_{B2}$ with 2\textsuperscript{nd} order nested DDSM

- OSR=256
- $l_1=2$
- $B_1=20$
- $B_{11}=8$
- $B_{2}=14$
- ENoB=24

Power Spectrum

Quantization noise (of nested DDSM)

Quantization noise (of input)
Motivation

- Now apply the output of the nested $l_1$th order DDSM preprocessor to the input of a conventional $l_2$th order DDSM
Motivation

- The hardware of the DDSM scales as $l_1B_{11} + l_2(B_2 + l_1)$ instead of $l_2B_1$

- Can we choose $l_1$ and $B_{11}$ such that $l_1B_{11} + l_2(B_2 + l_1) \leq l_2B_1$?
Motivation

• The output of a DDSM contains the input in addition to shaped quantization noise. For a third order DDSM, we write

\[ Y(z) = X(z) + (1 - z^{-1})^3 E_Q(z) \]

• Suppose that extra noise corrupts the input and it is made by a white term or by terms shaped by a higher order filter

\[ X'(z) = X(z) + E_{Q_0}(z) + (1 - z^{-1})E_{Q_1}(z) + (1 - z^{-1})^2 E_{Q_2}(z) \]

• Integration over the signal band of the additional noise terms degrades the SNR. This degradation is negligible if the extra noise is small compared to main shaped \( E_Q(z) \) term

• Goal is to obtain architectures with less hardware and negligible increase in quantization noise
Nested DDSM Architecture

- Nested DDSM divides the input into multiple segments

- For third-order MASH 111 DDSM we divide the input into two segments

\[ N = N_{MSB} + N_{LSB} \]

- \( N \)-bit input can be written as

\[ X = X_{MSB} \cdot 2^{N_{LSB}} + X_{LSB} \]

- Weight of \( X_{LSB} \) is lower by a factor \( 2^{N_{LSB}} \) than the full scale and consequently so is the quantization noise!
Nested 1-3 DDSM

- Total in-band quantization noise power given by
  \[ N_i = \frac{\Delta_Q^2}{12} \left( \frac{1}{OSR} \right)^{2l+1} \frac{\pi^{2l}}{2l + 1} \]

- For 1-3 nested DDSM, we have
  \[ N_0 \approx \frac{\Delta^2}{12} \cdot \frac{1}{OSR}, \]
  \[ N_1 \approx \frac{\Delta_Q^2}{12} \cdot \left( \frac{1}{2^{N_{MSB}}} \right)^2 \cdot \left( \frac{1}{OSR} \right)^3 \cdot \frac{\pi^2}{3}, \]
  \[ N_3 \approx \frac{\Delta_Q^2}{12} \cdot \left( \frac{1}{OSR} \right)^7 \cdot \frac{\pi^6}{7}. \]

- Assume a sinusoidal input with amplitude 1V, quantized to \( N \) bits; \( \Delta = 2/2^N, \Delta_Q = 2 \)
• In-band quantization noise powers can be rewritten as

\[ N_0 \approx \left( \frac{1}{2^{N_{MSB}+N_{LSB}}} \right)^2 \cdot \frac{1}{3OSR}, \]
\[ N_1 \approx \left( \frac{1}{2^{N_{MSB}}} \right)^2 \cdot \left( \frac{1}{3OSR^3} \right) \cdot \frac{\pi^2}{3}, \]
\[ N_3 \approx \left( \frac{1}{3OSR^7} \right) \cdot \frac{\pi^6}{7}. \]

• SNR of conventional and 1-3 nested architecture given by

\[ SNR_{CONV} = \frac{1/2}{N_0 + N_3} \quad SNR_{13} = \frac{1/2}{N_0 + N_1 + N_3} \]

• Reduction in SNR and ENOdB given by

\[ \Delta SNR_{13} = 10 \log \left( \frac{N_0 + N_1 + N_3}{N_0 + N_3} \right) \]
\[ \Delta ENoB_{13} = \frac{10}{6.02} \log \left( \frac{N_0 + N_1 + N_3}{N_0 + N_3} \right) \]
PSD of DDSM3 and nested 1-3 DDSM

- OSR=256
- N_{MSB}=12
- N_{LSB}=8
- ΔENoB_{13}=0.83
- PSDs almost identical!
- 13 bits at input of DDSM3
- From simulation, ΔENoB_{13}=1.01

Power Spectrum

Quantization noise (of input)

Nested DDSM1
• For nested 2-3 DDSM, we have

\[ N_2 \approx \frac{\Delta_Q^2}{12} \cdot (\frac{1}{2^{N_{MSB}}})^2 \cdot \left( \frac{1}{OSR} \right)^5 \cdot \frac{\pi^4}{5} \]

\[ \approx \left( \frac{1}{2^{N_{MSB}}} \right)^2 \cdot \left( \frac{1}{3OSR^5} \right) \cdot \frac{\pi^4}{5} \]

• SNR of conventional and nested 2-3 architecture given by

\[ SNR_{CONV} = \frac{1/2}{N_0 + N_3} \quad SNR_{23} = \frac{1/2}{N_0 + N_2 + N_3} \]

• Reduction in SNR and ENoB given by

\[ \Delta SNR_{23} = 10 \log \left( \frac{N_0 + N_2 + N_3}{N_0 + N_3} \right) \]

\[ \Delta ENoB_{23} = \frac{10}{6.02} \log \left( \frac{N_0 + N_2 + N_3}{N_0 + N_3} \right) \]
PSD of DDSM3 and nested 2-3 DDSM

- OSR=256
- $N_{MSB}=6$
- $N_{LSB}=14$
- $\Delta E\text{NoB}_{23}=0.42$
- PSD almost identical!
- 8 bits at input of DDSM3
- From simulation $\Delta E\text{NoB}_{23}=0.38$
PSD of DDSM3 and nested 2-3 DDSM

- OSR=256
- $N_{MSB}=12$
- $N_{LSB}=8$
- $\Delta E\text{NoB}_{23}=0.001$

- PSD almost identical!
- 14 bits at input of DDSM3
- From simulation $\Delta E\text{NoB}_{23}=0.03$
Conclusions

- Quantization errors introduced by the nested DDSM preprocessor can be masked

- Optimized design methodology is under development

- We expect savings in hardware and power

- Shorter words will potentially give a speed advantage too
References


References


Acknowledgements

• FIRB Italian National Research Program RBAP06L4S5
• Science Foundation Ireland Grants 02/IN1/I45 and 08/IN1/I1854