OULU 2012
C 414

ACTA

UNIVERSITATIS OULUENSIS


*Marko Neitola*

# CHARACTERIZING AND MINIMIZING SPURIOUS RESPONSES IN DELTA-SIGMA MODULATORS

*UNIVERSITY OF OULU,*
*FACULTY OF TECHNOLOGY,*
*DEPARTMENT OF ELECTRICAL ENGINEERING*

ACTA UNIVERSITATIS OULUENSIS
C Technica 414

*MARKO NEITOLA*

# CHARACTERIZING AND MINIMIZING SPURIOUS RESPONSES IN DELTA-SIGMA MODULATORS

Academic dissertation to be presented with the assent of the Faculty of Technology of the University of Oulu for public defence in OP-sali (Auditorium L10), Linnanmaa, on 17 February 2012, at 12 noon

UNIVERSITY OF OULU, OULU 2012

Acta Univ. Oul. C 414, 2012

Supervised by
Professor Timo Rahkonen

Reviewed by
Professor Michael Peter Kennedy
Docent Svante Signell

ISBN 978-951-42-9748-9 (Paperback)
ISBN 978-951-42-9749-6 (PDF)

ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)

Cover Design
Raimo Ahonen

JUVENES PRINT
TAMPERE 2012

**Neitola, Marko, Characterizing and minimizing spurious responses in Delta-Sigma modulators**
University of Oulu, Faculty of Technology, Department of Electrical Engineering, P.O. Box 4500, FI-90014 University of Oulu, Finland
*Acta Univ. Oul. C 414, 2012*
Oulu, Finland

***Abstract***

Oversampling data converters based on Delta-Sigma modulation are a popular solution for modern high-resolution applications. In the design of digital-to-analog or analog-to-digital Delta-sigma converters there are common obstacles due to the difficulties on predicting and verifying their performance. Being a highly nonlinear system, a Delta-Sigma modulator's (DSM) quantization noise and therefore the spurious tones are difficult to analyze and predict.

Multi-bit DACs can be used to improve the performance and linearize the behavior of DSMs. However, this will give rise to the need for linearizing the multi-bit DAC. A popular DAC linearization method, data weighted averaging (DWA) shapes the DAC mismatch noise spectrum. There are many variants of DWA, for low-pass and band-pass DSMs. This thesis proposes a generalization which integrates a few published variants into one, broader DWA scheme. The generalization enables expanding the tone-suppression studies into a larger concept.

The performance of one- or multibit DSMs is usually verified by simulations. This thesis proposes a simulation-based qualification (characterization) method that can be used to repeatedly verify and compare the performance of multibit DSM with a DAC mismatch shaping or scrambling scheme.

The last contribution of this thesis is a very simple model for tonal behavior. The model enables accurate prediction of spurious tones from both DSMs and DWA-DACs. The model emulates the tone behavior by its true birth-mechanism: frequency modulation. The proposed prediction model for tone-behavior can be used for developing new tone-cancelation methods. Based on the model, a DWA linearization method is also proposed.

*Keywords:* Delta-Sigma modulation, nonlinear modeling, quantization noise, spurious tones

**Neitola, Marko, Harhatoistojen karaktorisointi ja minimointi Delta-Sigma muuntimissa**
Oulun yliopisto, Teknillinen tiedekunta, Sähkötekniikan osasto, PL 4500, 90014 Oulun yliopisto
*Acta Univ. Oul. C 414, 2012*
Oulu

## *Tiivistelmä*

Delta-Sigma modulaatio on suosituin tekniikka ylinäytteistävissä datan muuntimissa. Riippumatta toteutustarkoituksesta (analogia-digitaali- tai digitaali-analogia-muunnos), Delta-Sigma (DS) modulaatiossa on yleisesti tunnettuja käyttäytymisen ennustamiseen liittyviä ongelmia. Nämä ongelmat ovat peräisin modulaattorin luontaisesta epälineaarisuudesta: DS-muunnin on nimittäin vahvasti epälineaarinen takaisinkytketty systeemi, jonka harhatoistojen ennustaminen ja analysointi on erittäin hankalaa.

Yksibittisestä monibittiseen DS-muuntimeen siirryttäessä muuntimen suorituskyky paranee, ja muuntimen kohinakäyttäytyminen on lineaarisempaa. Tämä kuitenkin kostautuu tarpeena linearisoida DS-muuntimen digitaali-analogia (D/A) muunnin. Tällä hetkellä tunnetuin linearisointimenetelmä on nimeltään DWA (data weighted averaging) algoritmi. Tässä työssä DWA:lle ja sen lukuisille varianteille esitellään eräänlainen yleistys, jonka avulla algoritmia voidaan soveltaa sekä alipäästö- että kaistanpäästö-DS-muuntimelle.

Kuten tunnettua, DS-modulaattorin analyyttinen tarkastelu on raskasta. Yksi- ja monibittisten DS-muuntimien suunnitellun käyttäytymisen varmistaminen tapahtuukin yleensä simulointien avulla. Työssä esitetään simulointiperiaate, jolla voidaan kvalifioida (karakterisoida) monibittinen DS-muunnin. Tarkemmin, kvalifioinnin kohteena on DWA:n kaltaiset D/A -muuntimien linearisointimentelmät. Kyseessä on pyrkimys ennen kaikkea toistettavaan menetelmään, jolla eri menetelmiä voidaan verrata nopeasti ja luotettavasti.

Tämän väitöstyön viimeinen kontribuutio on matemaattinen malli harhatoistojen syntymekanismille. Mallilla sekä DS-muunnoksen että DWA-D/A -muunnokseen liittyvät harhatoistot voidaan ennustaa tarkasti. Harhatoistot mallinnetaan yksinkertaisella havaintoihin perustuvalla FM-modulaatiokaavalla. Syntymekanismin mallinnus mahdollistaa DS-muuntimien ennustettavuuden ja täten auttaa harhatoiston kumoamismenetelmien kehittämistä. Työssä esitetään yksi matemaattisen mallin avulla kehitetty DWA-D/A -muunnoksen linearisointimenetelmä.

*Asiasanat:* Delta-Sigma modulaatio, epälineaarinen mallinnus, harhatoisto, kvantisointikohina

# Acknowledgements

This thesis is based on research carried out at the Electronics Laboratory of the department of Electrical and Information Engineering, University of Oulu. The contributing work in this thesis began early 2009 by toying with the idea of a generalized data weighted averaging algorithm (DWA). Soon it was realized that the model is very convenient in comparing and extending different aspects of DWA shaping with tone-cancelation alternatives. By summer, Janne A. Raappana implemented the generic model with an FPGA realization. The generalized model, along with Jannes's excellent work, triggered further ideas on qualifying DWA for comparative purposes and an actual time-domain spurious-tone model. A final icing on the cake was extending the same spurious tone mechanism to Delta-Sigma modulators.

I would like to express my deepest gratitude to my supervisor, Professor Timo Rahkonen, for his guidance and encouragement. I wish to thank Professors Michael Peter Kennedy and Svante Signell for examining this thesis. I am especially grateful to Professor Kennedy also for proof-reading this thesis.

I would also like to thank all the former and present staff in the Electronics laboratory for pleasant working environment. I am especially grateful to Dr. Maciej Borkowski for his helpful comments and also his excellent contribution in his doctoral thesis.

A warm thank you to my parents Irma and Tuomo for all their support during these years.

Oulu, February 2012 Marko Neitola

# List of abbreviations and symbols

| | |
|---|---|
| A/D | analog-to-digital |
| ADC | analog-to-digital converter |
| BNF | back-and-forth |
| D/A | digital-to-analog |
| DAC | digital-to-analog converter |
| dc | direct current; zero frequency |
| DEM | dynamic element matching |
| DNL | differential nonlinearity error |
| DR | dynamic range |
| DS | Delta-Sigma |
| DSM | Delta-Sigma modulator |
| DWA | data weighted averaging |
| ENOB | effective number of bits |
| ETF | the transfer function from DAC input to DSM output |
| FS | full scale |
| FFT | fast Fourier transform |
| ILA | individual level averaging |
| INL | integral nonlinearity |
| LSB | least significant bit |
| LUT | lookup-table |
| MASH | multistage noise shaping DSM |
| mod | modulus after division |
| MSE | mean squared error |
| NTF | noise transfer function |
| OSR | oversampling ratio |
| pdf | probability distribution function |
| RAM | random access memory |
| rms | root-mean-square |
| SFDR | spurious-free dynamic range |
| SMNR | signal to mismatch noise ratio |
| SNDR | signal to noise and distortion ratio |
| SQNR | signal to quantization noise ratio |
| STF | signal transfer function |

| | |
|---|---|
| $A$ | amplitude |
| $B$ | bits |
| $c_i$ | Fourier coefficients |
| $D$ | delay in generic DWA |
| $E\{x(n)\}$ | expected (mean) value of $x(n)$ |
| $e$, $e_q$ | quantizer noise |
| $e_m$ | unshaped DAC mismatch noise |

| | |
|---|---|
| $f$ | frequency |
| $f_b$ | signal bandwidth for low-pass DSM |
| $f_c$ | center frequency |
| $f_s$ | sampling frequency |
| $i$ | index |
| $k$ | quantizer gain |
| $L$ | DSM order |
| $L_0$ | input path filtering in a DSM loop filter |
| $L_1$ | feedback path filtering in a DSM loop filter |
| m | order (polynomial, harmonic distortion) |
| $M$ | number of samples in a simulation |
| $n$ | sample index |
| $n_q$ | modulation noise |
| $n_m$ | DAC mismatch noise |
| $N$ | number of quantizer steps; commonly the number of unit DACs |
| $N_s$ | number of spurious tones modelled |
| $ptr$ | DWA pointer signal |
| $\underrightarrow{S}$ | power spectral density |
| $\overrightarrow{SEL}$ | selection vector for DEM |
| $seq$ | repeating sequence |
| $t$ | time |
| $T$ | sampling interval |
| $u$ | DSM input signal |
| $v$ | DSM output signal |
| $v_d$ | DAC output signal |
| $w_i$ | weight of unit-DAC element |
| $y$ | DSM quantizer input signal |
| | |
| $\beta_i$ | attenuation parameter for noise contribution model |
| $\Delta$ | quantizer step |
| $\varepsilon$ | error |
| $\phi$ | phase |
| $\sigma$ | standard deviation |
| $\sigma^2$ | variance |
| $\omega$ | angular frequency |

# List of original publications

This thesis consists of an overview and the following four publications:

I Neitola M & Rahkonen T (2010) A Generalized Data-Weighted Averaging Algorithm. IEEE Trans. Circuits and Systems II: Express Briefs 57(2): 115-119.

II Neitola M, Rahkonen T & Raappana J (2010) A Qualification Approach to DAC Mismatch-Shaping Methods. IEEE Trans. Circuits and Systems II: Express Briefs 57(11): 858-862.

III Neitola M & Rahkonen T (2011) Predicting and Avoiding Spurious Tones in a DWA-Mismatch-Shaping DAC. IEEE Trans. Circuits and Systems II: Express Briefs 58(6): 341-345.

IV Neitola M & Rahkonen T (2011) Compact Tone-Behavior Model for Delta-Sigma Modulator. Proc 20th European Conference on Circuit Theory and Design, ECCTD 2011: 261-264.

# Contents

# 1 Introduction

A Delta-Sigma modulator (DSM) [1-3] is a derivative of an interpolative encoder i.e. a feedback system with filtering and a quantizer in its feed-forward path. Contrary to Nyquist-rate data converters, Delta-Sigma (DS) converters are oversampling and have memory. A DS converter trades speed for a simpler realization that has more relaxed accuracy requirements for the analog components. In a data converter application, the quantizer is typically coarse, only 1 to a few bits in word length, which makes a Delta-Sigma (DS) converter very nonlinear and difficult as a subject of exact analysis [4].

In DSM analysis, a classical additive noise model for the quantizer [5] is the most commonly used modeling tool. Unfortunately, its usage is typically for a suggestive prediction of the behavior. In a quasi-linear analysis [6], the signal dependency is typically modelled by nonlinear describing functions of the DSM input, but does not predict the tone frequencies. An exact analysis, in the other hand, is complex and the results difficult to generalize [7]. Nevertheless, the advantages of DS converters surpass the lack of an exact predicting model.

Extending the DSM output word size improves the predictability of the system, because the quantization noise behavior will be more randomized. Also, the performance improves ca 6 decibels per bit [8]. Unfortunately, the increment in word size introduces nonlinearity in the digital-to-analog converter (DAC), which needs to be linearized. Options for linearization without trimming are digital error correction [9] and dynamic element matching (DEM) [10].

This thesis discusses mostly multi-bit DSMs. It consists of four original papers and a hundred and twelve page summary. Some other related papers by the author appear as references.

## 1.1 Overview and contribution

The contribution of this thesis has a lot of weight on multibit DSM's DAC mismatch linearization. This thesis proposes a generalized model for a popular DEM method, namely data weighted averaging (DWA) [11,12]. The generalization clarifies the basic functionality of first-order shaping DWA and expands the existing tone-cancelation methods to band-pass DSMs as well.

Numerous methods to linearize a DWA method evoke a particular problem for scientists with interest on this field: repeatability. As the achieved performance is dependent on the properties of the input signal along with the actual mismatch, a

proper qualification may be an exhaustive and difficult task. This invoked the author's interest in the qualification (characterization) of DEM methods. The proposed qualification approach discusses important issues in reliability and repeatability when benchmarking DEM methods by simulations.

A common problem with many mismatch shaping methods and DSMs is the empirical nature of studying the noise related phenomena. In both cases, the shaped noise can be prone to produce spurious tones. To understand and even cancel unwanted tones, a prediction model for both DSM and DWA is needed. This thesis presents such a time-domain model that is capable of predicting tone behavior with arbitrary stimuli. The model is simple and accurate, and it also reveals the similarity between DSM modulation noise and DWA-related mismatch noise.

The structure of quantization noise is a field that has been studied in numerous publications beginning with the classical paper by Candy and Benjamin [13] from 1981. As the literary review will suggest, there are similarities in how the spurious tones are located for DSM and DWA. The spurious tone models for DWA (Paper III) and DSM (Paper IV) are almost identical and their simplicity and accuracy are quite convincing.

The structure and contribution of this thesis are as follows:

- Chapter 2 concentrates on methods for analyzing DSM quantization noise. The emphasis is on reviewing approaches on how to predict the DSM quantization noise behavior.
- Chapter 3 is an overview of DAC mismatch cancelation methods in multibit DS converters.
- Chapter 4 proposes a DWA generalization (Paper I) that combines several published variants into one concept.
- Chapter 5 proposes the DEM qualification method (Paper II).
- Chapter 6 introduces a simple spurious tone prediction model for DWA (Paper III) and DSMs (Paper IV)

# 2 Methods of Delta-Sigma modulator analysis

## 2.1 Introduction

The purpose of Delta-Sigma modulator analysis methods is to verify stable behavior and predict the performance. Generally, a DSM can be presented as a negative feedback system whose feed-forward path contains linear filtering and a coarse quantizer. The latter can be categorized as a "hard" nonlinearity; a DSM is a highly nonlinear system, which is can be very complicated to predict. Robert M. Gray stated in his classic paper on quantization noise spectra [4]:

> *A longstanding problem with the application of feedback quantization systems has been the difficulty of analyzing and predicting their behavior; the complications of the nonlinear operation are aggravated by its presence within a feedback loop.*

Gray pointed out [4] a general belief that (for 1-bit DSMs) the higher the order of the system, the more accurate the white noise assumption is. A recent dissertation from a former colleague contained a chapter on DS modulators' tonal behavior. This chapter contained a rather pessimistic statement [7]:

> *It should be noted that many works published on the subject, including very recent ones, begin by emphasizing our very limited theoretical understanding of DSMs. Simulation is the basic tool for engineers working with delta-sigma modulation, and this has been used to determine the stability and noise performance of a DSM and to gain insight into LC (limit cycle) behavior.*

Indeed, a simulator is a pivotal tool for determining the stability and performance of DSMs.

Spurious tones affiliated with DSMs are often called pattern noise, limit cycle and idle channel tones. The first two are synonymous and mostly affiliated with dc inputs. A limit cycle is a DS converter output sequence that is indefinitely repeating itself [14]. The spurious tones frequently addressed in this thesis are assumed to result from stable behavior. In [15], it was stated that a stable idle tone has an amplitude that does not change with time, but is a complicated function of the stimulus, as was also concluded in [13].

There are pitfalls in empirical observations from simulations [14]: a system may seem fine with certain initial conditions and stimulus, but may still be unstable. A certain amount of expertise and an extensive amount of simulation is

typically needed. Richard Schreier's empirical study [16] provides groundwork on the subject. Also in [14] (ch. 4), Adams and Schreier described a Cookbook design procedure based on empirical methods and observations. Schreier's contribution to the field of DS converters is immense and his DS toolbox for MATLAB [15] is a well-known DSM design and simulation tool. The DS toolbox was often used in this work.

The use of a multibit quantizer in a DSM improves the performance by ca. 6dB per bit increment [8] and also scrambles the input-dependency of the quantizer noise. For single-bit DSMs, the elimination of idle channel tones requires means of whitening the quantizer noise. These methods are dither and chaos.

A good example of the need for device-level linearization is in audio applications, where the idle tones are easily picked up by the human ear. The order of the DSM noise shaping may also reduce the probability of tones, but in [17] it was demonstrated that even a fourth order DSM can produce spurious tones with sinusoid and dc-input. A pseudo-random dither signal added to the quantizer input is capable of breaking the input signal dependency with a trade-off of increasing the in-band noise floor.

Dither can be added either to quantizer input or the actual stimulus. Latter is more common in the art of digital DSMs, particularly in frequency synthesizers. Pamarti *et al.* [18] presented necessary and sufficient asymptotic conditions in which input LSB-dithering leads to uniform and input-independent quantization noise. Fitzgibbon *et al.* [19] proposed higher-order shaped LSB dithering for digital DSMs, which produces spurious-free quantization noise with dc-inputs.

In [20], it was mentioned that a (pseudo-) random dither signal may impact DSM stability as the effective quantizer gain is reduced. Another approach to device-level linearization a 1st order DSM is to use chaotic DSM [20] i.e. DSM with noise shaping zeroes outside of the unit-circle.

This Chapter discusses the classical modeling approaches: exact and nonlinear analysis in Sect. 2.2, approximate analysis in Sect. 2.3 and simulation in Sect. 2.4.

## 2.2 Exact and nonlinear analysis

Low-order one-bit DS converters are less prone to stability problems but more prone to produce limit cycles. This pattern noise is simple to demonstrate using a first order DSM with dc input signals at various levels. For this case, at rational fractions of the dc input signal, the entire quantization noise will be concentrated

on a few tones.

Candy & Benjamin studied the recurrent 1st order 1-bit DSM patterns [13] for dc-inputs. The characteristic result is a pattern noise graph i.e. the measured values of noise in the baseband against dc-input level. An example of such plot is shown in Fig. 1, with the oversampling ratio (OSR) of 32. The peaks occur adjacent to rational fractions of the dc-input.

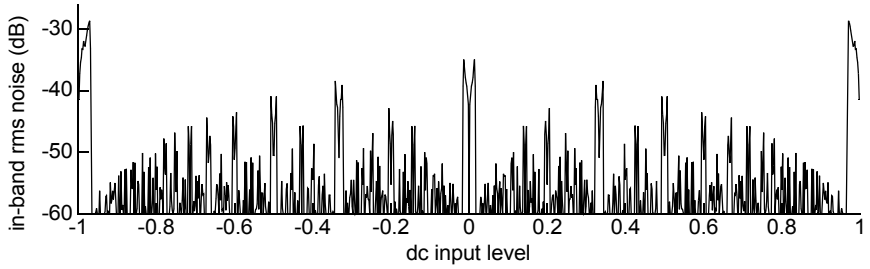



**Fig. 1. The graph of simulated in-band rms modulation noise plotted against the level of dc input.**

Increasing the order of the DS converter tends to scramble the tendency for limit cycles, but not perfectly; predicting tones becomes more difficult. Exact analysis refers to exact solutions to the nonlinear difference equations modeling DSMs.

The quantizer error can be approximated by an input-independent additive white noise model, if the quantizer noise $e$ has the following properties [14]:

1. Statistically independent on the input signal $u$
2. Uniformly distributed in $[-\Delta/2,\Delta/2]$
3. Stationary noise with flat power spectral density.

In exact analysis, the approach of finding the conditions under which the quantization noise can be classified as white, is called the characteristic function method [14]. This approach was criticized in [14] (by R. M. Gray), because it does not fully characterize the general quantization error.

The exact analysis of a DSM is not a popular option. In [15], it was mentioned that the true dynamics of a second and higher order DS converter are intractable, which often leads to using approximate and/or empirical techniques. For the published papers on exact analysis, the generality of the results is often lost; the formulae are rather difficult to apply in practice [7].

In a larger concept, nonlinear analysis methods are namely the methods that study the phenomena that are not amenable to approximate analysis. Farrell and Feely [21] collated the main topics in the field of DSM nonlinear analysis:

– By spectral analysis it was shown that the assumption of white quantization noise was incorrect [4], as there was a strong correlation between the input and the output bit pattern. In his paper on quantization noise spectra, R. M. Gray [4] constructed a Fourier-series model for the quantization noise while retaining the exact nature of his analysis. For a periodic quantizer error (e.g. in Fig. 5) the model of the quantizer error in [4] is a periodic Bessel function containing sinusoids.
– Geometric analysis concentrates on stability analysis with integrator spans. The range of all possible values for the integrator outputs can be determined and the existence and stability of limit cycle behavior can be defined.
– Nonlinear dynamics: analyzing a DSM with a state equation to identify the conditions under which limit cycle behavior arises.

Papers on exact and nonlinear analysis usually discuss basic low-order low-pass DSMs with dc- or single-tone inputs. Some band-pass DSM analyzes are also published, e.g. in [22] a second-order DSM with single-tone stimulus exactly at the DSM center frequency $f_s/4$, where $f_s$ is the sampling frequency.

## 2.3 Approximate[1] solutions

At the core of approximate solutions is the assumption of white quantization noise. The assumption can be valid with a certain choice of DSM and stimuli. For low-order DSMs and dc-inputs, the quantization noise signal can be periodic, leading to a limit cycle with a discrete power spectrum. By increasing the order or using a multibit quantizer, the assumption of white DSM quantizer noise is more justified, but not necessarily guaranteed.

In the aspect of exact analysis, the white noise approximation is not justified mathematically (R. M. Gray in [14]). In [4], Gray concluded that white noise approximations inaccurately predict the spectral nature of the quantizer noise process, which is neither continuous nor white; the amplitude of the quantizer error spectrum depends strongly on the value of the input signal.

[1] Naming convention used in [16]

However, the analysis methods that can be adapted as general-purpose tools are usually based on approximations. This means that the methods are mainly suggestive, perhaps a supporting tool for verifying stability and performance. Linear and quasi-linear analyzes are such tools. The disadvantage of linear and quasi-linear analysis is that the validity of the white noise approximation is system-dependent [15].

### 2.3.1 Linear analysis

In [5] (ref. [14]), Bennett developed the conditions under which the white noise assumption is valid. Bennett's asymptotic conditions for a quantizer model are as follows:

1. The input is in the no-overload region
2. The number of quantizer levels is asymptotically large
3. The quantizer step is asymptotically small
4. The joint probability density function (pdf) of the quantizer input signal at different sample times is smooth.

Unfortunately, these conditions are typically violated in DSMs [4]. Nevertheless, the white noise assumption can be reasonably accurate in some cases and it also enables deriving the approximate signal to quantizer noise (SQNR) estimate.

A linearized DSM model is an initial design tool. Linearizing a DSM topology enables design at a transfer function level, which is pivotal in generating and scaling the DSM coefficients. The linear performance estimate presented later in this Section is a tool for choosing the oversampling ratio and the order of the DSM. For multibit-design, the number of required quantizer levels can also be estimated.

Analyzes based on linearization usually incorporate an additive noise model, which itself is pivotal in explaining the actual noise-shaping properties of a DSM. The additive noise model presents the coarse quantizer as a linear gain $k$ and a quantizer error $e$. With quantizer input $y$ and output $v$, the model is:

$$v = ky+e. \tag{1}$$

The quantizer gain $k$ is a statistical measure of the true quantizer input-output slope. Using the additive noise model it is possible to calculate the frequency domain noise transfer function (NTF) i.e. the transfer function from the additive noise source to the output. For a discrete- or continuous-time NTF, it is possible to

analyze DSM stability by sweeping the values of $k$ in a root-locus plot [23]. The true quantizer gain variation can be obtained from a group of simulations by using an estimation formula presented in Subsection 2.3.2.

The linear model presented in Fig. 2 models the quantizer as an additive noise model. This enables the DSM design in transfer function level. These functions are realized by a certain DSM topology (hidden inside the loop filter block in Fig. 2) with coefficients that match the transfer functions. The white noise assumption is not invoked in an initial transfer function design, as the noise is modelled as a separate source. The output of the loop filter in Fig. 2 is given as:

$$Y(z) = L_0(z)U(z) + L_1(z)V(z). \tag{2}$$

The signal and noise transfer functions i.e. *STF* and *NTF* respectively, are defined as:

$$STF(z) = \frac{V(z)}{U(z)}, NTF(z) = \frac{V(z)}{E(z)}. \tag{3}$$

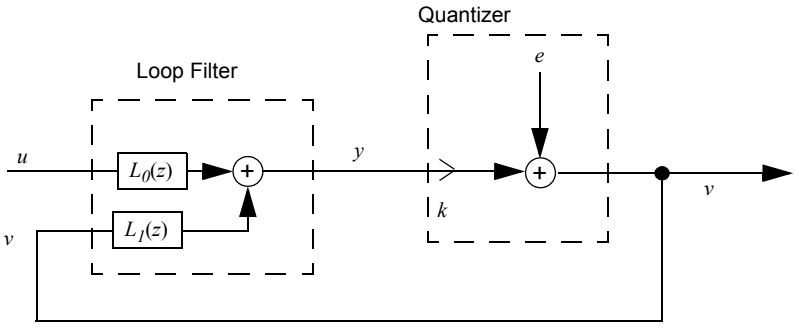



**Fig. 2. DSM additive noise model.**

In terms of signal and noise transfer functions, the loop filters in Fig. 2 are accordingly [14]:

$$\begin{cases} STF(z) = \dfrac{L_0(z)}{1-L_1(z)} \\ NTF(z) = \dfrac{1}{1-L_1(z)} \end{cases} \Rightarrow \begin{cases} L_0(z) = \dfrac{STF(z)}{NTF(z)} \\ L_1(z) = \dfrac{NTF(z)-1}{NTF(z)} \end{cases}. \tag{4}$$

With the quantizer gain $k$=1 in (1) and (2) we get the DSM output as:

$$V(z) = STF(z)U(z) + NTF(z)E(z). \tag{5}$$

Of course, the loop filter has to be realizable. A consequence of the rule of causality (at least one unit delay in the loop) is that the first sample of the *NTF*'s impulse response has to be one [16].

If the quantizer gain $k$ is not defined as unity, *STF* and *NTF* will be reformed as:

$$\begin{cases} STF(k,z) = \dfrac{L_0(z)\cdot k}{1-L_1(z)\cdot k} = \dfrac{k\cdot STF(z)}{k+NTF(z)\cdot(1-k)} \\ NTF(k,z) = \dfrac{1}{1-L_1(z)\cdot k} = \dfrac{NTF(z)}{k+NTF(z)\cdot(1-k)} \end{cases}. \tag{6}$$

The root locus of $NTF(k,z)$ may reveal that the system is unstable at some values of $k$. An indication of instability is when the stable input range of the quantizer is exceeded: a quantizer becomes overloaded and the quantizer's effective gain drops. A typical consequence from instability is a low-frequency limit cycle (for low-pass DSM).

With signal and noise transfer functions, the quantizer input $Y$ becomes:

$$\begin{aligned} Y(z) &= STF(z)U(z) + (NTF(z)-1)E(z), \text{ for } k = 1 \text{ and} \\ Y(k,z) &= \frac{STF(z)U(z) + (NTF(z)-1)E(z)}{k+NTF(z)\cdot(1-k)} \\ &= \frac{STF(k,z)}{k}\cdot U(z) + \frac{NTF(k,z)-1}{k}\cdot E(z), \text{ for general case.} \end{aligned}. \tag{7}$$

With (7), the unstable boundary can be stated as: the gain of the *NTF* can not be too large. For instance, the famous (empirical) Lee's rule-of-thumb [24] (ref. [15])

states, that the maximal out-of-band gain for the *NTF* should be less than 2 for 1-bit DSMs. A larger gain (more aggressive *NTF*) is permitted for multibit DSMs.

In literature, many performance estimates are based on a linear model with noise shaping of [14]:

$$NTF(z) = (1 - z^{-1})^{L}. \tag{8}$$

This is an $L$th order "pure-differentiation" NTF. For (8), the maximal gain ($|NTF(-1)|$) is $2^L$, which is too great for a one-bit DSM with $L>1$. To avoid instability, the loop filter can be modified to accommodate poles $DEN(z)$, making the *NTF* into:

$$NTF(z) = \frac{(1 - z^{-1})^{L}}{DEN(z)}, \tag{9}$$

where the role of $DEN(z)$ is to constrain the maximal gain. A common choice for $DEN(z)$ is using Butterworth poles [16]. The primary choice of poles is made according to the assumed maximal NTF gain i.e., the maximal out-of-band gain. To meet the performance specifications, the zeros of the NTF may also need some adjustment. Complex zero pairs in the signal band attenuate the in-band modulation noise, but complicate the circuit realization.

Fig. 2 hides the topology of the DSM. Some common loop filter topologies are shown in Fig. 3. A topology is a flow-graph presentation of the loop filter with input signals (stimulus and the feedback), one or more integrators (or resonators in the case of band-pass DSM) and some gain coefficients. The integrators "INT" may be either delaying or non-delaying, given that the system remains causal.

The topology in Fig. 3a is the so-called error-feedback topology, commonly used in digital DSMs e.g. for DS-DACs. The filtering is usually realized by a simple FIR-filter: a delay would result in a first-order DSM.

In Fig. 3b and Fig. 3c we have the two common topologies that enable flexible NTF design. The coefficients $\{a_i,b_i,c_i,g_i\}$ and the number of integrators with the chosen topology are designed according to the NTF. These basic topologies contain a chain of integrators either with one feedback and feed-forward summation (Fig. 3b) or with multiple feedbacks (Fig. 3c).

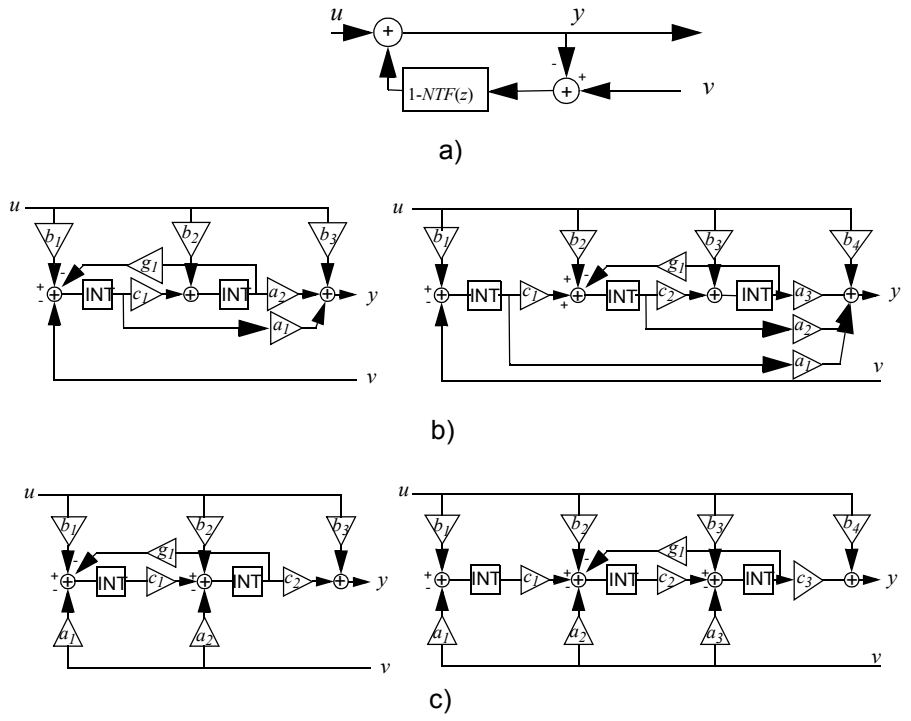



**Fig. 3. DSM loop filter topologies: a) error feedback, b) feed-forward (2nd and 3rd order) and multiple feedback (2nd and 3rd order).**

The topologies in Fig. 3b and Fig. 3c also contain

– additional input paths for designing the STF and
– additional feedbacks (with gains $g_i$) between two integrators for creating NTF zeros.

If a DSM topology has a unity-STF, the chain of integrators (or resonators) in the feed-forward path will only have to process shaped quantization noise. This is easy to conclude [25] by linear analysis by investigating the DSM's negative feedback output i.e. the input for the integrator chain:

$$\begin{aligned} X_{in}(z) &= U(z) - V(z) \\ &= U(z) - STF(z)U(z) - NTF(z)E(z) \\ &= U(z) \cdot (1 - STF) - NTF(z)E(z). \end{aligned} \tag{10}$$

If the *STF* is unity in (10), the first integrator input is the quantization noise $E(z)$ shaped by the NTF. A DSM topology with unity-STF has smaller integrator swings. In [26] this was concluded for the topologies in Fig. 3b and Fig. 3c.

If the basic topologies in Fig. 3b and Fig. 3c had only one input branch to the first integrator, the STF would be referred to as the inherent signal transfer function. The feed-forward topology Fig. 3c has an inherent signal transfer function of [14]:

$$STF(z) = 1 - NTF(z). \tag{11}$$

This STF is inherently flat, i.e. close to unity. Combining (10) and (11) results to a feedback signal of:

$$X_{in}(z) = NTF(z) \cdot (U(z) - E(z)). \tag{12}$$

It was mentioned in [27] that a feed-forward topology exhibits lower distortion than a traditional multi-feedback topology. The integrator swings are typically verified by rigorous simulations. The coefficients realized in the final device are often truncated versions of the original ones. This truncation can be made according to a convenient capacitance ratio (discrete-time DS ADC) or a low-complexity shift and sum operation [15] (digital DSM for DS DAC).

The white noise model applied to Fig. 2 makes is possible to estimate the signal to quantization noise ratio (SQNR). Assuming the quantizer error equally probable and zero-mean between quantizer steps, the quantizer error variance $\sigma_e^2$ becomes:

$$\sigma_e^2 = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} e^2 de = \frac{\Delta^2}{12}, \tag{13}$$

where $\Delta$ is the quantizer step. Assuming uniformly distributed noise power, the 2-sided power spectral density (PSD) is:

$$S_e(f) = \frac{\sigma_e^2}{f_s} = \frac{\Delta^2}{12 f_s}. \tag{14}$$

Without noise shaping, the in-band noise power for low-pass signals is calculated as [28]:

$$P_e = \int_{-f_b}^{f_b} S_e(f)df = \frac{\Delta^2}{12\mathrm{OSR}}, \tag{15}$$

where $f_b$ is the signal bandwidth and OSR is the oversampling ratio. OSR is the ratio between half the sampling frequency and $f_b$. For a DSM, the quantizer noise is filtered by the noise transfer function resulting into modulation noise $n_q$. In the frequency domain, the modulation noise is:

$$N_q(f) = NTF(f) \cdot E(f). \tag{16}$$

The power spectral density of the modulation noise $S_q(f)$ can be presented as:

$$S_q(f) = |NTF(f)|^2 \cdot S_e(f) = |NTF(f)|^2 \cdot \frac{\Delta^2}{12 f_s}, \tag{17}$$

and the in-band shaped noise power becomes:

$$P_q = \int_{-f_b}^{f_b} S_q(f)df = \int_{-f_b}^{f_b} |NTF(f)|^2 \cdot \frac{\Delta^2}{12 f_s} df. \tag{18}$$

The shaped in-band noise power is often approximated by using the pure-differentiation NTF presented in (8) e.g. in [15] and [28]:

$$P_q \approx \sigma_e^2 \cdot \frac{\pi^{2L}}{(2L+1) \cdot OSR^{2L+1}} = \frac{\Delta^2}{12} \cdot \frac{\pi^{2L}}{(2L+1) \cdot OSR^{2L+1}}. \tag{19}$$

The well-known *SQNR* estimate based on pure-differentiation NTF is presented as:

$$SQNR(dB) \approx 10log_{10}\left[\frac{3}{2}\cdot 2^{2B}\cdot\frac{2L+1}{\pi^{2L}}OSR^{2L+1}\right], \tag{20}$$

where $B$ is the quantizer output wordlength in bits. In a more recent publication (2010), Løkken *et al.* [29] presented an SQNR estimate that accommodates the user-defined NTF:

$$SQNR(dB) \approx 10\cdot\log_{10}\left(\frac{\frac{A^2\cdot(N+1)^2}{\pi/(OSR)}}{\frac{1}{3\pi}\cdot\int\limits_{-\pi/(OSR)}^{\pi/(OSR)}|NTF(\omega)|^2 d\omega}\right), \tag{21}$$

where $A$ is the normalized amplitude of the single-tone stimulus and $N$ is the number of quantizer steps.

The white noise approximation is always indicative. The usefulness of the *SQNR* estimate is in comparing the simulation results with the predicted white noise model: it is helpful to have a performance reference.

### 2.3.2 Describing functions

Generally, the quantization noise $e$ is a deterministic function of the input signal [4]. Compared to white noise model, a more accurate quantizer model can be obtained using describing functions. This is the quasi-linear approach which approximates the input signal dependency.

The describing function is an approximation method analyzing and predicting nonlinear behavior. For a specific stimulus (e.g. dc or single-tone) the linear gain $k$ is chosen to minimize the mean squared error (MSE) between the linear approximation and the true response. The quasi-linear approach can be used to predict the occurrence of limit cycles and stability boundaries and it serves as a useful design guide.

For DSMs, the quasi-linear model contains both signal and quantizer noise contributions as does the describing function method. In a classic publication on quasi-linear technique for 1-bit DS converters by Ardalan & Paulos [6], the signal dependency is affiliated with the quantizer gain by nonlinear (describing) functions of the DS converter input. The modulator is split into two linear systems for both signal and noise, assuming that the noise portion has a Gaussian

probability density function. The model provides an accurate estimate at the maximum stable input and maximal SQNR, but is restricted to 1-bit converters and does not predict the spurious tone frequencies.

Ardalan & Paulos [6] proposed an analysis based on modeling the nonlinear quantizer with a linearized gain obtained by minimizing a mean-square-error criterion. This method was used in [6] to derive the regions of stability for higher order modulators, including both parameter and signal dependencies. The linearized gain (i.e. the quantizer gain) for the noise is [6]:

$$k = \frac{cov(y, v)}{var(y)} = \frac{E\{y(n) \cdot v(n)\}}{E\{y(n) \cdot y(n)\}} = \frac{\sum_{n=0}^{N} y(n)v(n)}{\sum_{n=0}^{N} y(n)^2}, \tag{22}$$

where *cov* and *var* are the covariance and variance functions, respectively. For one-bit DSMs, the quantizer output $v(n)$ is the sign of the quantizer input.

A recent (2010) publication by Altinok *et al.* [30] derived stability boundaries with a modified describing function for two-tone stimuli and very high-order 1-bit band-pass DSMs. In their conclusions Altinok *et al.* discussed of the unfortunate fact that the quantization noise is not completely Gaussian: for higher DSM orders, the approximation holds better.

If the number of quantizer levels is larger that two, the quantization noise can often be assumed to be white, given that the stimulus

– exceeds more than two quantizer levels and
– is not slowly varying.

In [7], Borkowski deduced that for the classical model of quantization (more generally, approximate analysis), slowly varying inputs are not well-suited. On the other hand, a spurious response to a slow input is more observable in simulations, which is utilized in the quantization noise model presented in Ch. 6.

## 2.4 Simulation

A reasonable procedure with simulation as a pivotal tool also requires the use of linear approximate tools and some rules-of-thumb. For instance, the quantizer gain

estimate in (22) can be obtained from simulations. With a group of stimuli (dc, single-tone etc.), a reasonable span for $k$ can be used for the root-locus analysis of the noise transfer function.

Empirical rule-of-thumb criteria are a practical tool based on reflecting simulation results with the noise transfer function obtained from the linearized model. None of them are adequate as such [16] (suggestive tool), but simulations are the only reliable method. The most famous rule-of-thumb is Lee's criterion for the maximal allowable out-of-band gain for 1-bit DSMs [24].

Verifying a DSM architecture requires linear analysis followed by simulations. For instance, finding stability bounds require long simulations [14] with various stimuli. Instability manifests itself as long consecutive sets of the overloaded value of the quantizer. This may occur after a long simulation and may also occur with a particular stimulus. Stability should be verified at least with stimuli at $f_c$ (dc-stability in low-pass DSMs) and, to author's experience, single-tones with frequencies across the signal band. Moreover, a basic SQNR versus amplitude simulations should be repeated with a group of single-tone frequencies throughout the band-of-interest.

Especially for 1-bit DSMs, the limited out-of-band gain naturally restricts the noise shaping function. Using multibit DSMs enables a more aggressive noise transfer function i.e. the maximal gain can be increased. For a very aggressive NTF, the SQNR performance may be maximized, but the stable input level can be quite restricted. A DSM with an aggressive NTF is also prone to unexpected behavior that may be found by simulation. A typical NTF usually has the single-tone stable input level limit at 60-80% of the full scale input signal [14].

The fast Fourier transform (FFT) analysis of quantizer noise that contains short-term phenomena is problematic. A long-term windowed FFT can be used to evaluate performance, but it tends to average out some important phenomena than can be non-observable in the spectrum yet can be e.g. audible. This challenges simulation-based empirical observations because the observation of a tone may require short-term estimates of the autocorrelation or power spectral density of the DSM output [14]. The published papers on DSM analysis and dither-like techniques often contain observation-based assumptions on the magnitudes and frequencies of the spurious tone. These assumptions are further discussed in Ch. 6.

### 2.4.1 Case study - an arising limit cycle

He *et al.* [31] studied a double-loop (second-order multiple feedback) DSM topology and concluded conditions under which the quantizer noise can be considered as white. Unfortunately, their result cannot be extended or used of any other topology. Here, it will be demonstrated how a simulator may reveal unexpected behavior.

In Fig. 4, we have a DSM output spectrum split in four consecutive parts. The DSM is of order 2 with four quantizer levels and a dc-input of 5FS/16, where FS is the full scale input amplitude. The NTF was optimized for an aggressive NTF with four quantizer levels and OSR of 16. This optimization was performed by the DS toolbox [15] by using the synthesis procedure developed by Kenney & Carley [32].

At the beginning of the simulation (Fig. 4a and Fig. 4b) the modulation noise looks quite clean, except for a spike at $0.34f_s$. After 16384 samples, the limit cycle begins to appear in the modulation noise spectrum (Fig. 4c) and the spectrum turns completely discrete after 24576 samples (Fig. 4d). As can be seen in Fig. 4, the white noise assumption can be quite volatile.

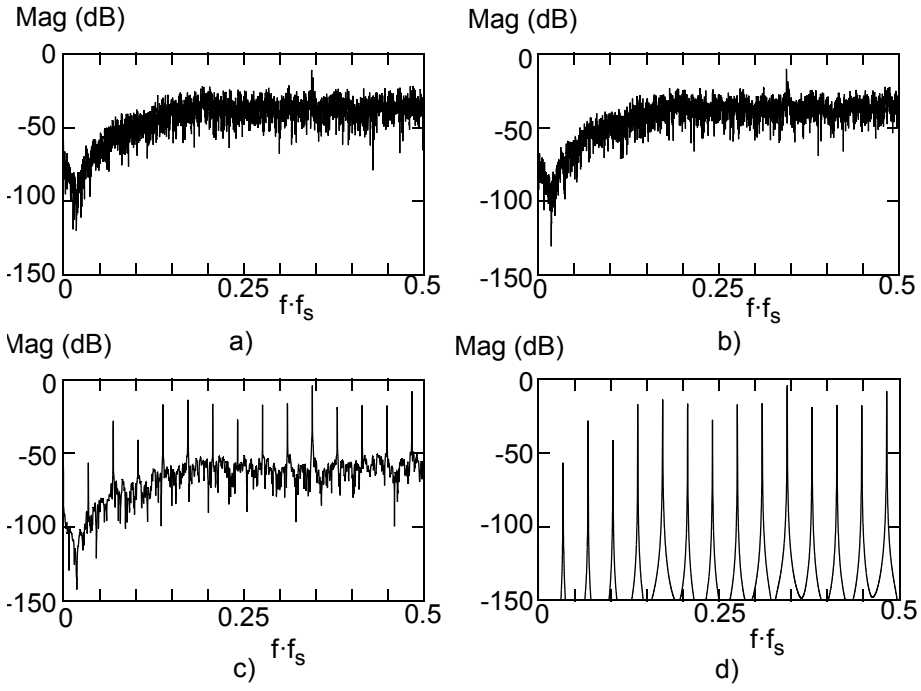



**Fig. 4. A 2nd order DSM dc-response power spectrum divided into four parts: a) samples [1...8192], a) samples [8193...16384], a) samples [16384...24576], a) samples [24577...32768].**

There are four noteworthy points on how prone this particular DSM is to spurious tones:

1. The initial tone at $0.34f_s$ is the strongest tone persistent throughout the simulation. All tone frequencies can be predicted, but the sudden change in Fig. 4c is hardly predictable. Predicting the tone frequency (prior art and the author's contribution) is discussed in Ch. 6.
2. The NTF in this example is aggressive, the input signal level is not allowed to be more than 0.44FS. Moreover, NTF aggressiveness can make the DSM more persistent to dither [33].
3. The NTF zeros are outside of center frequency. This makes the DSM more prone to limit cycle as was also concluded in [33].

4. The nature of the quantizer noise power spectrum depends on NTF zeros in an obscure manner. Integer division dc-inputs (rational dc inputs with respect to full scale) are, however, more likely to produce tonal modulation noise [34].

### *2.4.2 Case study - periodic quantizer error*

In [34], first order 1-bit DSM responses for dc-inputs were studied in the time-domain. Here, the periodicity of the quantizer error depended on whether the dc-input was rational with respect to the quantizer levels of $\pm 1$. Regardless of the rational nature of the input, however, a dc-stimulus always results in discrete spectrum for 1st order DSM. The periodic sequences are the limit cycles. The length of a limit cycle can be determined analytically in the case of digital DSMs with dc inputs [35]. In addition to the stimulus, the limit cycle length depends on the initial conditions of the integrators [14] (chapter 4 by S.R. Norsworthy) as well.

In Fig. 5, we have two quantizer error graphs in the time-domain using two different quantizers: the mid-rise and the mid-tread quantizer. For the mid-rise quantizer, the quantizer steps are placed at {..., -3, -1, 1, 3, ...}and for the mid-tread quantizer: {..., -2, 0, 2, ...}. Here, the quantizer step $\Delta$ is fixed to 2 and the modulator order is 1. In Fig. 5, the dc-input and the amplitude are small compared with $\Delta$: $3\Delta/516$. Small rational dc-input leads to a quantizer error with a long period.

The average of the DSM output data is the same as the average of the stimulus. If the small input signal is in the middle of two quantizer levels, the output varies frequently between the two levels, e.g. $\pm 1$. As the sign of the quantizer input is fed back negatively, the quantizer error in Fig. 5a has to have fluctuation close to $f_s/2$. For the sinusoidal case with the mid-rise quantizer in Fig. 5b, the tones reside near dc and $f_s/2$ as well, but they vary as a function of time.

The point in using two different quantizer types is the following: if the small input is very near the quantizer level, most of the output samples are concentrated at the same level, and few at the adjacent level. The quantizer error for a mid-tread quantizer in Fig. 5a lead to a triangular quantizer error for the mid-tread quantizer in Fig. 5a and its power spectrum would have very low-frequency components. For Fig. 5b, the quantizer error is sinusoidal for the mid-tread quantizer resulting in just one low-frequency tone.

In Fig. 5, the quantizer error graphs for mid-rise and mid-tread quantizer would be swapped if the stimulus were offset by half the quantizer step (given that another quantizer level exists).

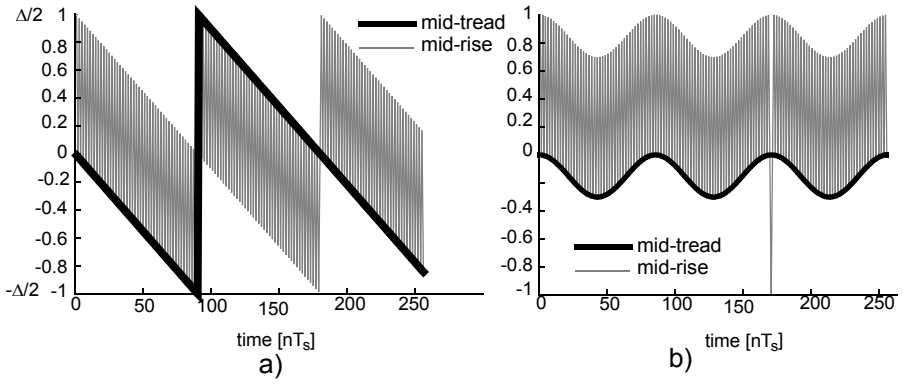



**Fig. 5. Quantizer error response for 1st order DSM with mid-tread and mid-rise quantizer: a) dc-input of 3Δ/516 and b) single-tone input with the amplitude of 3Δ/516.**

## 2.5 Summary

The literature review in this Chapter reveals a need for a DSM model which would allow us to better understand and thus combat the unwanted tones. DSM nonlinearity prohibits the exact analysis for higher order systems, the emphasis is typically is on the approximate and simulation-based analyzes.

The problem of input-dependent quantization noise is heavily alleviated in multi-bit DSMs, but the tonality is still dependent on the stimulus type; a multibit DSM can also have a discrete modulation noise spectrum.

The problem of empiric and linear analysis are also present in linearizing the D/A converter in the case of multi-bit DSMs, and this will be discussed in the next chapter.

# 3 Multibit DAC mismatch error

## 3.1 Introduction

By incrementing the DS converter output word length by one bit, we gain more resolution [8]. Furthermore, more than two quantizer level translates into more scrambled patterns in the DS loop: the correlation between the signal and quantizer noise becomes less clear. In fact, the quantizer gain can usually be assumed unity in the case of multibit DSMs, if all quantizer levels are used. The deviation of quantizer gain $k$ from its nominal value of 1 is small: at most $\Delta/|2 \cdot y(n)|$ [15], where $\Delta$ is the least significant bit (LSB) voltage of the quantizer (the quantizer step) and $y$ is the quantizer input.

As multibit DS converters promise fewer stability problems and better performance, another problem emerges: the D/A converter is no longer linear. For a one-bit DS converter, the DAC is inherently linear as there are just two points in the DAC input-output slope. For a multibit DS ADC or DAC, the static DAC mismatch is not noise-shaped like the quantization noise is. This can be depicted in the linearized scenario for multibit DS ADC and DAC shown in Fig. 6, where $e_q$ and $e_{mm}$ are the additive noise models of quantization and mismatch error.

For DS ADCs (Fig. 6a), the DAC is internal; the mismatch error is injected to the feedback path of the DSM loop. The transfer function from $e_{mm}$ to the output $v$ is basically the same as the inherent signal transfer function. This filtering will not cancel the in-band mismatch noise. In case of a DS DAC (Fig. 6b), the mismatch noise will also remain unshaped, as the DAC is at the output of the digital DS loop's output.

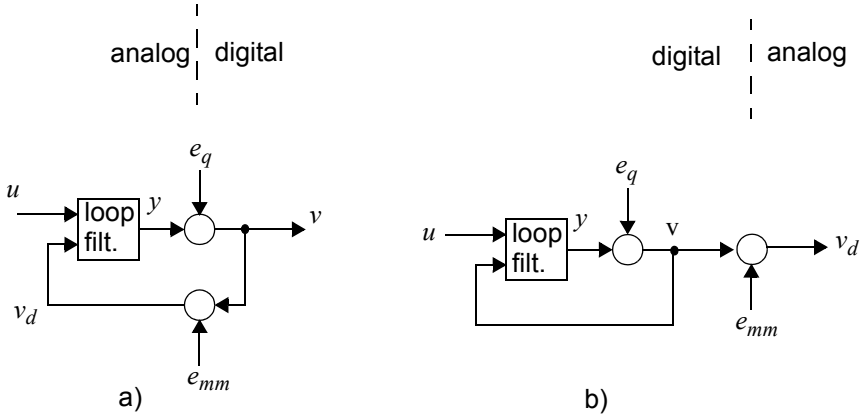



**Fig. 6. Additive noise model for quantizer noise and mismatch noise: a) DS ADC and b) DS DAC.**

For both cases in Fig. 6, there will be in-band spurious tones caused by the DAC mismatch noise; the distortion mechanism remains the same. DAC induced out-of-band tones will be filtered out by the post-filtering (decimation filter or an analog filter), but the in-band DAC-mismatch related noise will not be attenuated.

The DAC linearity is a measure of how accurately the levels in the internal DAC are placed. The ultimate linearity of the DS DAC or DS ADC is no better than the linearity of the DAC [14].

The DAC elements are most commonly realized with current-steering or switched capacitor architectures. Errors either in nominal currents or capacitor dimensions are the sources of unit-DAC mismatches. These errors are typically assumed as static and random. [36]

For a DAC with a resolution of $B$ bits, the LSB voltage can be calculated as:

$$V_{LSB} = \frac{V_{FS}}{2^B}, \tag{23}$$

where $V_{FS}$ is the full-scale output voltage. Therefore, for a $B$ bit DAC, the DAC output levels cannot deviate from their ideal values by more than half of $V_{LSB}$. The linearity of the DAC should be at least the same as the linearity of the DSM. The DAC linearity, i.e. its resolution in bits, is often described by the effective number of bits, ENOB (for a full scale sinusoid):

$$ENOB = \frac{SNDR(dB) - 1{,}76}{6{,}02} bits \quad , \tag{24}$$

where SNDR is the signal to noise and distortion power ratio over the frequency band of interest (dB). For D/A converters used in DS converters, the DAC word size (the actual number of bits) is typically a lot smaller than the ENOB and the linearity demand is not usually achievable without DAC linearization [14].

The organization of this Chapter is as follows. Section 3.2 discusses the problem of static DAC mismatch noise in DSMs and discusses two main approaches on how to minimize this noise. The first one, Digital error correction, is briefly discussed in Sect. 3.3. The focus, however, is on presenting various DEM methods: Sect. 3.4 covers an overview of digital encoders that realize different DEM approaches.

## 3.2 Static DAC mismatch noise

Mismatch errors are considered to be static errors. The unit DAC element selection can be typecast as in [37] either as static selection or dynamic selection. Dynamic (element) selection refers to selecting elements as a function of the present code and of memory of the previous conversions (e.g. an algorithm). In this Section, static selection is discussed; dynamic selection will be discussed in Sect. 3.4.

For a $B$-bit DS converter, the number of quantizer steps $N$ is

$$N = 2^B - 1 \, . \tag{25}$$

The number of quantizer levels is $2^B$ i.e. $N$+1. The focus in this thesis is on thermometer-coded DACs, that contain $N$ nominally equal unit-elements that are cumulated according to the code: the analog DAC output is the sum of nominally equal weights. Another approach would be a binary-weighted DAC, which contains one resistor or current source for each bit of the DAC connected to a summing point.

Assuming unity weights, the true weight of a DAC element $i$ is

$$w_i = 1 + \varepsilon_i, \quad \varepsilon_i = \{1, \ldots, N\} \, , \tag{26}$$

where $\varepsilon_i$ is the differential nonlinearity (DNL) error. Given that the DAC input range is from *0* to *N*, the output of the DAC $v_d$ is a sum of weights (with the notations of Fig. 6):

$$v_d(n) = \begin{cases} \sum_{i=1}^{v(n)} w_i, \text{ if i} > 0 \\ 0, \text{ if v(n)} = 0 \end{cases} . \tag{27}$$

The deviation from a straight DAC input-output graph is specified by the integral nonlinearity (INL) error, i.e. the sum of DNL errors. Therefore, the DAC output $v_d$ is the sum of the digital code *v* and the corresponding INL error:

$$v_d(n) = v(n) + INL(n) . \tag{28}$$

In a well-designed system, the DAC noise *INL(n)* is a zero-mean sequence [38]. To achieve good matching requires careful analog layout techniques, such as the common centroid method [41]. As (27) suggests, the DAC noise is correlated with the input: the DAC input contains a clear deterministic element. For instance, if the DAC input is a pure sinusoid with frequency *f*, we can expect mismatch-related harmonics at 2*f*, 3*f*, etc.

Realizing an N-level DAC based on unit-DAC elements requires encoding the binary DAC input in thermometer code, see Fig. 7. The digital encoder in Fig. 7 contains the thermometer encoder and can also can contain logic for dynamic element selection.

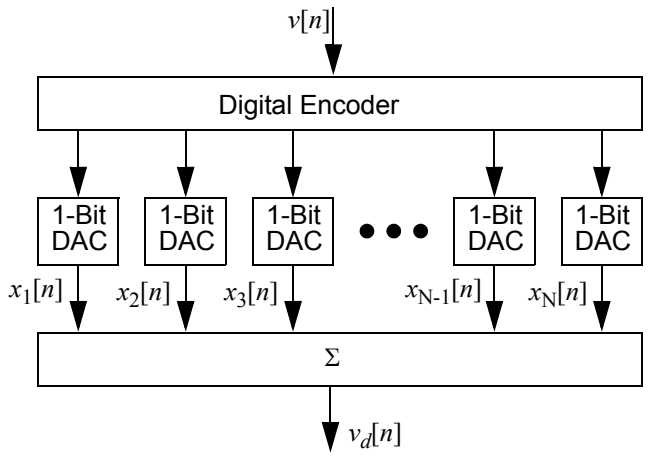



**Fig. 7. The block diagram of an N-level DAC.**

The true shape of the DNL error is not usually known, but the layout technology sets the limits of the expected mismatch quantity. A common way to grade the mismatch error is to use standard deviation [41]:

$$\sigma_{\varepsilon} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\varepsilon_i - E\{\bar{\varepsilon}\})^2}, \tag{29}$$

where $E\{\bar{\varepsilon}\}$ is the expected value of all DNL mismatches in vector $\bar{\varepsilon}$, whose length is $N$. With static selection, to achieve a high-resolution internal DAC requires excellent matching. Unfortunately, the integral nonlinearity (INL) that yields a resolution of more than 12 bits [15] cannot be easily reached without DAC element trimming. To cancel the mismatch noise in the presence of imprecise component matching, the DAC mismatch noise can be attenuated with digital error correction or spectrally shaped with dynamic element matching (DEM).

For a Delta-Sigma ADC, the limiting factor for word size of the feedback-DAC is the internal A/D converter. For word sizes larger than 5-bits, the number of comparators ($>2^5$) may result in unacceptable chip area and power consumption. For DS DAC, the bottleneck is the complexity of the digital circuitry needed for DAC linearization. The DAC complexity can be alleviated by using

– DAC segmentation, which alleviates the size limitation by allocating different DAC weights: the hardware area will be significantly lower in case of the segmented DAC [43].
– Two-step (or multistep) architecture [44], which divides the A/D-conversion into a multibit DS-modulator to two time steps. The resulting feedback from ADC stages has a word length of $B_1$+$B_2$ bits, but the complexity of the ADC stages is proportional only to $B_1$ or $B_2$.
– The encoder complexity can alleviated by spectral DAC error shaping only partially within the full-scale range [42].

### 3.3 Digital error correction

Unlike DEM methods, digital error correction cancels the mismatch noise instead of shaping the spectrum. In the concept of digital error correction, the element selection remains static. The correction is based on measuring the mismatches by the so-called calibration setup.The foreground calibration is a linearization method introduced in the late 80's [9]. This approach relies on first measuring the feedback DAC's INL errors after which the multibit ΔΣ ADC can operate. The term "foreground" [8] refers to a separate mode of operation for digital error measurement i.e. calibration. The background error correction scheme is capable of measuring the DAC errors while the multibit DSM is in operation.

During calibration, the integral nonlinearity (INL) errors are measured by using a single-bit configuration followed by decimation. In the foreground approach, the calibration mode configures the convertor as a single-bit DS ADC, which is inherently insensitive to DAC errors at DC [9]. A digital counter generates all the possible input codes for the DAC.

General correction schemes for a DS ADC and DAC are shown in Fig. 8a and Fig. 8b, respectively. The equivalence in the two models is that for the digital error correction to work, the feedback data in DSM loop and the DSM output should be matched i.e. (nearly) identical.

A digital correction block i.e. the lookup-table (LUT) compensates the unwanted nonlinearity of the N-bit DAC. The data stored in LUT, $\hat{INL}(n)$, is an accurate digital equivalent of the DAC integral nonlinearity errors $INL(n)$. Since the DAC INL for a specific digital input value can be measured, the error signal can be regenerated by using the modulator output as an address to the lookup-table, where all error voltages are stored in digital form. In each clock period, the input word to the LUT selects a $B_2$-bit word ($B_2 >> B_1$) at the RAM output [14].

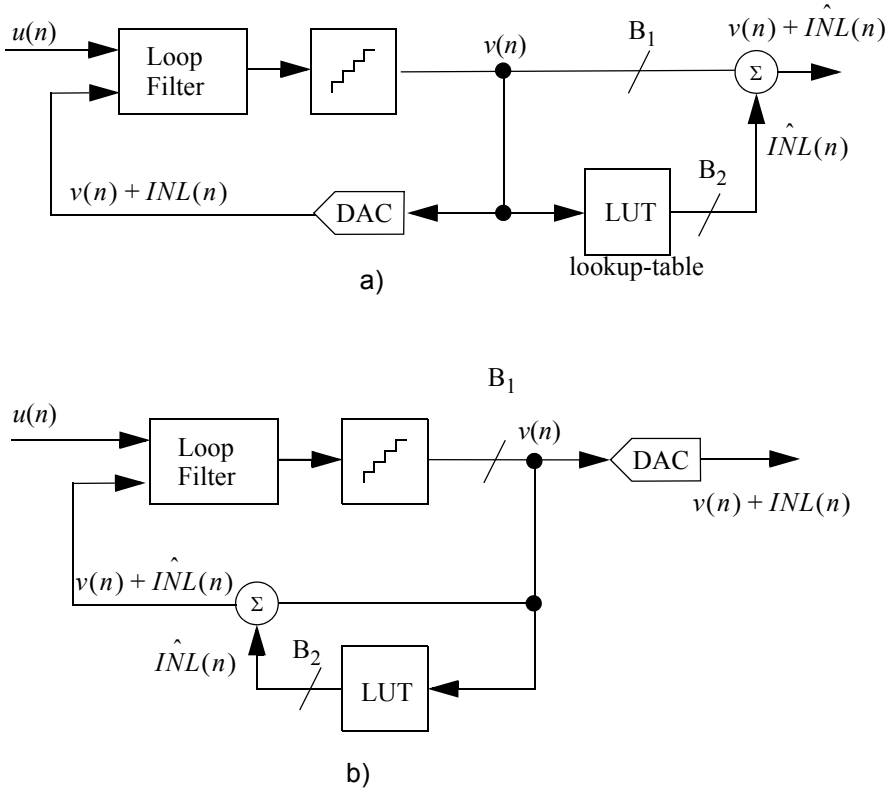



**Fig. 8. Digital error correction for a) DS ADC and b) DS DAC.**

For a current-switching DAC, background error correction was introduced in 1988 [45]. Here, a self-calibration technique was used to realize an array of current sources which are equal to each other within 0.02 percent. The calibration of a MOS (metal–oxide–semiconductor) current source is accomplished by biasing it with a reference current. In [46], the static linearity of a current-mode DAC was obtained by background calibration using an extra DAC element and a 1-bit Delta-Sigma A/D converter.

The idea of background calibration for a single-stage DS ADC is shown in Fig. 9. The idea is to calibrate feedback DAC unit elements successively using a parallel 1-bit shadow DSM. There is one additional unit element in this rotational calibration i.e., while one element is under INL error measurement the others are used by the multibit DS converter.

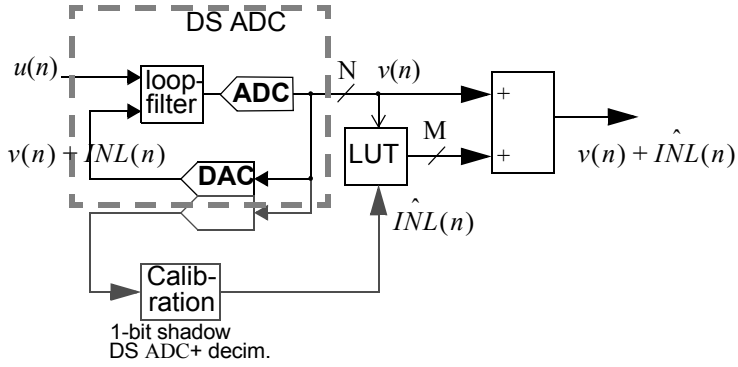



**Fig. 9. The block diagram of the background error correction.**

The element rotation is switched off after calibration, after which the error correction is equivalent to that of the foreground correction. Since there is only one additional unit DAC element the correction speed is nearly the same as for foreground correction. The time required for the calibration (in both correction methods) depends on the required resolution and hence on the order of calibration DSM and the following decimator. The main difference is that the one-bit calibration DSM is now a separate device since the correction is done in parallel. If needed, background correction can also be restarted after initial correction.

In both foreground and background approaches, the lookup-table output may be needed to be filtered (equalized) according to the transfer function from DAC output to DS ADC output [39,40]. This transfer function is notated as ETF in [15] and it is the transfer function from the injected mismatch noise to the DSM output (see Fig. 6). The need for the ETF filtering is dependent on the type of DSM topology. A feed-forward DSM topology presented in Fig. 3b has inherently wideband signal transfer function. As a consequence, the inherent ETF is also wideband and the filtering is not needed in the feed-forward topology.

A more complex background calibration scheme for a multistage noise shaping (MASH) [47-49] DS - ADC was presented by Silva *et al.* in [39]. This correction scheme is based on a pseudo-random thermometer output code scrambling and a correlation algorithm. This approach also contains ETF filtering.

### 3.4 Dynamic element matching

Dynamic matching of unit-DAC elements refers to scrambling the usage of DAC elements so that the signal-dependent deterministic mismatch noise is transformed into a wide-band noise. A basic feature in all dynamic element matching (DEM) methods is that they do not reduce the error - instead the error is moved from the signal band. DEM is a common family name for DAC cancelation methods that average the DAC-induced noise out of the signal band with or without noise shaping. The latter refers to the randomization of the DAC element usage. The majority of the reported DEM methods is however affiliated with the spectral shaping of DAC mismatch noise.

In Fig. 10, we have a DAC input-output scatter plot, where the DAC unit elements are scrambled randomly. Each curve for different unit element placing is non-linear, but the average DAC input transfer curve is close to linear. Indeed, by simply randomizing the DAC element usage, the distortion induced by the DAC mismatch can be averaged into a wide-band noise.

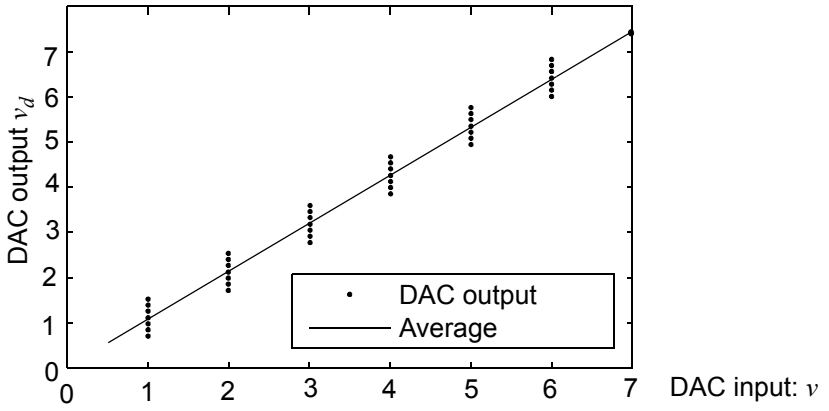



**Fig. 10. DAC input-output scatter graph for element scrambling DAC.**

To realize a noise shaping DEM, there are many choices with their own trade-offs between performance and complexity. For instance, the most common DEM

method called data weighted averaging (DWA) [11],[12], has the least uncomplicated realization. The problem is that DWA does not remove the input data dependency: the mismatch noise is still prone to spurious tones. The order of DEM shaping can be increased, but the complexity trade-off is more expensive in the case of Delta-Sigma modulators.

Behavioral similarities between certain mismatch shaping DEMs and a DSM are quite apparent. Probably the most obvious connection is the context of an arbitrary noise shaping function in [55], which will be discussed in the Sect 3.4.1.

Similarly as for DSM, approximate solution can be applied to DEM. For instance, Nys & Henderson [73] presented a resolution approximation for DWA as:

$$\text{resolution} = log_2\left(\frac{\sqrt{(3N \cdot OSR^3)}}{\pi \cdot \sigma_\varepsilon \cdot \left(1 - \frac{1}{N}\right)}\right) bits, \tag{30}$$

where $\sigma_\varepsilon$ is the standard deviation of the DNL errors and $N$ is the number of unit DACs. Løkken *et al.* [29] estimated the signal to mismatch noise ratio (SMNR) for any DEM method as:

$$SMNR = 10 \cdot \log_{10}\left(\frac{\frac{A^2 \cdot (N+1)}{\pi/OSR}}{\frac{\sigma_\varepsilon}{\pi} \cdot \int_{-\pi/OSR} |H(\omega)|^2 d\omega}\right), \tag{31}$$

where $A$ is the normalized amplitude and $H$ is the mismatch spectral shaping function of the DEM algorithm. Andersson *et al.* [56] condensed a general goal of Dynamic element matching (DEM):

> *The dynamic methods can be used during operation and are continuously compensating for matching errors by manipulating the input signal or the circuit elements using digital circuitry. DEM modifies the distortion terms, hence signal-dependent errors, to become signal-independent. Therefore, by using DEM we try to maximize the SFDR (spurious-free dynamic range).*

Unfortunately DEM methods are not perfect and can be prone to spurious tones. For instance, the DAC-related distortion in data weighted averaging is signal-dependent and there are many approaches on how to modify DWA in order to cancel spurious tones (and thus maximize SFDR). Therefore, approximating the mismatch noise with signal independent noise is merely indicative for DEM methods (as is for DSMs).

This can be illustrated by simulating a DSM with a DWA-DAC, whose SFDR performance is shown Fig. 11. Here, the stimulus is a single-tone with a fixed amplitude and frequency, only the standard deviation of the DNL error $\sigma_\varepsilon$ is swept with four different DNL error shapes. In Fig. 11 there are four SFDR graphs for four different shapes of DNL errors with identical $\sigma_\varepsilon$. Here, the DSM is a 3-bit second order with *OSR* of 32 and with DWA-DAC. The input single-tone amplitude is small in this example, ca 0.02 times the full-scale amplitude. It is apparent, that for a certain value of $\sigma_\varepsilon$, the linearity of the DWA-DAC noise will be below the linearity of the DSM. This breaking point is dependent on the shape of the DNL error [50] as can be seen in Fig. 11. This cannot be predicted by linear approximation.

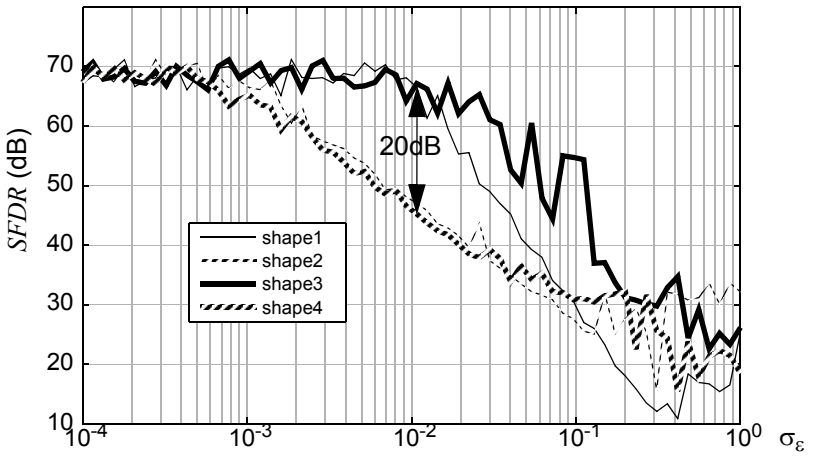



**Fig. 11. SFDR vs. the DNL rms-error with four shapes of DNL errors.**

For the record, the DNL error shapes used in simulation results in Fig. 11 were not generated randomly. Instead, the shapes were generated by using mathematically defined (hence repeatable) shapes. Shapes 1 to 3 are purely sinusoidal shapes with

1 to 3 periods, respectively (see Paper III and Sect. 6.4). Shape 4 is a third order polynomial (see Paper II and Sect. 5.2).

### 3.4.1 Main classes of DEM

In [51], spectral shaping DEM methods were divided into five main classes. These include data weighted averaging (DWA) [11], [12], individual level averaging (ILA) [52], tree structured [53], butterfly shuffler [54] and vector based DEM [55]. A comparison of each DEM-DACs was made in [51] by their complexity, propagation delay, spurious tones and the order of shaping.

#### *Data weighted averaging*

The DWA algorithm and its spectral shaping property are covered in more detail in Chapter 4 (Paper I). In Data weighted averaging, all DAC elements are used equally often by choosing the DAC elements in a rotational (cyclical) manner. The error average is therefore zeroed as soon as all elements are used. The unidirectional way of choosing the elements translates to high-pass spectral shaping with a transfer function of:

$$H(z) = 1 - z^{-1}. \tag{32}$$

As the DAC elements are picked according to the input signal, DWA is prone to unwanted spurious tones. A classic publication on DWA by Baird & Fiez [12] summarizes the spurious tone problem of DWA by the following: “Each bit-increase doubles the elements to cycle through before errors average to zero and thus, the DWA algorithm cannot move distortion as high in frequency”. This limits the number of DAC elements for low oversampling ratios [57], because of the increased in-band tones. The spurious tones for DWA are further discussed in Chapters 5 and 6.

The majority of DWA’s tone cancelation methods are dither-like solutions with slight trade-off between reduced spur-levels and in-band noise (e.g. pseudo-DWA [25] and partitioned DWA [58]). A DWA variant called rotated DWA [59] is particularly interesting in the light of the tone behavior model presented in Ch. 6. Rotational DWA contains normal DWA element swapping algorithm and additionally switches between the DAC usage permutations. Section 6.5 suggests finding only one favorable permutation in DWA.

Chapter 4 (Paper I) integrates the ideas behind a handful of DWA variations that alter the spectral shaping function (32), such as Bi-DWA [60], N-path DWA[61] and $1+z^{-2}$ DWA [62]. This generalization also expands the existing DWA tone cancelation methods for band-pass DSMs.

The reported results for DWA variants are typically quite empirical i.e. based on simulation. Unfortunately, the performance result is dependent on the stimulus as well as the shape of the mismatch. Therefore there is a dire need for a systematic approach to simulation-based DEM-performance reporting. The major issue is in reliability and repeatability and these will be discussed in Ch. 5.

### *Individual level averaging*

The basic concept of individual level averaging (ILA) [52] is that each unit DAC element is used with equal probability for each input code. The algorithm has two variants, but the principle remains the same. The error averaging time is longer than in DWA, which translates to higher in-band noise. This also results in less input signal dependency, i.e. fewer spurious tones will occur at the signal band. In [63], it was concluded that the mismatch noise power spectral density has in-band shaping of 7dB/dec for ILA and 9db/dec for DWA.

For ILA, the implementation complexity increases rapidly as a function of DAC resolution (compared with DWA) [64]. Interestingly, in [64] it was also mentioned that there are compromises between DWA and ILA that strive for a less spur-prone DWA with lower complexity.

### *Butterfly shuffler*

To randomize $N$ possible output elements in a time-varying fashion, the number of possible permutations is $N!$ [14]. For more than 3-bit DACs, the number of possible connections is so large that it may be necessary to select a subset of connections in order to conserve die area [14]. A plausible solution would be to use a butterfly randomizer [65] (ref. [14]), i.e. a series of butterfly networks (such as those used in the FFT architecture) coupling the input to the outputs.

The minimal number of switching stages (that connects any input to the output) is the same as the number of bits $B$ in the internal DAC. More stages results for more possible connections. The control sequences for butterfly switches can be produced by a pseudo-random sequence generator.

A butterfly network based mismatch shaping swapper was first introduced by Adams and Kwan in [54]. The butterfly shuffler approach can also be used in higher order DEM shaping [66]. The idea is as simple as turning the logic around the filtering function into a truth-table realized as a butterfly network.

The complexity, however, cannot be circumvented in the butterfly approach. The size of the truth table is $2^B$, where $B$ is the required word-size of the filtered output. In [66], a second order mismatch shaping function $1+z^{-2}$ for band-pass DSMs was implemented.

The paper on butterfly shufflers [66] by Haiqing & Schreier presents a model of the structure which shows how a butterfly shuffler can be endowed with "arbitrary" noise-shaping characteristics. A 2nd-order band-pass shuffler was given as an example. Simulations in [66] showed that butterfly shufflers do not reach the performance of the vector-based DEM (see Page 49), but tend to produce smaller harmonics by limiting the element usage patterns.

### *Tree structure and segmentation*

The first-order tree structure DAC configuration has similar complexity to the butterfly network, but offers the advantage that it can be used in conjunction with a simple dithering technique that suppresses spurious tones [53]. A $B$-bit digital encoder consists of $B$ layers of digital devices called switching blocks, each of which splits its input data into two output signals whose sum equals the input. A switching block is driven by a switching sequence, whose properties dictate the possible shaping scheme. A single switching sequence generator resembles a DSM, whose order is the same as the order of shaping.

Second and higher order shaping is always a much more complex solution than the basic first-order shaping DEM. It was mentioned in [53] that a tree-structured DEM can be the least complicated approach in realizing a 2nd order mismatch shaping algorithm.

The tree structure approach with spectral shaping can also incorporate DAC segmentation [67]. Segmentation is a property that truly alleviates the often "prohibitively large" [68] DAC bank using unit-DACs. Fishov *et al.* [68] constructed a first-order tree-structure mismatch noise shaper for a segmented DAC. As in [53], a dithering scheme was proposed as a remedy for spurious tones in [68].

### *Vector-based DEM*

The algorithm behind arbitrary high order DAC mismatch shaping has existed since 1996 [69]. The downside of high-order DEM is the required hardware complexity: it would be preferable to have a DEM scheme that has a minimum amount of hardware and parallel signal processing.

If mismatch noise is dominant, a need to consider high-order shaping arises. This is obvious in the case of low *OSR* and high order DSMs. In this case, digital error correction instead of spectral shaping can be a potential solution. A second-order DWA, i.e. 2DWA [69] is complicated and somewhat impractical as it needs to operate twice as fast as the input signal rate. Partial second-order DWA (P2DWA) [70] circumvents the speed problem with a trade-off having slightly reduced performance, but it seems to involve a lot of parallel computing. The generic vector-based DEM was first described by Schreier & Zhang in [55].

A general vector-based DEM [55] can implement any shaping function. For a general *L*th order mismatch shaper described in [55], the digital encoder is actually a Delta-Sigma modulator topology. The topology of the general mismatch shaper is based on an error feedback DSM structure shown in Fig. 3a. Such a structure is a common DSM topology choice for the digital loop filter of a DS DAC.

In Fig. 12, the mismatch noise shaping transfer function is denoted as *H*. The output of the vector quantizer **sv** is a binary vector, with the number of ones corresponding to the thermometer encoded version of the input *v*[n], i.e. $\Sigma\mathbf{sv}(n) = v(n)$. The bold lines in Fig. 12 indicate *N* parallel data paths i.e. a vector. The address of ones and zeros in **sv** are determined by another vector **sy**(*n*), which contains integer weights denoting the desired usage of an element i.e. priority weight. This requires the use of a sorting algorithm. The desired elements are chosen according to the input data *v* by the vector quantizer (VQ), resulting in the output vector **sv**. The selection error **se** in Fig. 12 is filtered, normalized and fed back to the vector quantizer. The subtraction by the minimum value of filtered data minimizes the internal word size of the vectorized loop.

By linear analysis, the DAC output $v_d$, defined by input *v* and the mismatch shaping transfer function *H*, becomes [55]:

$$V_d(z) = V(z) + H(z)(\mathbf{SE}(z)\cdot\bar{\varepsilon}). \tag{33}$$

Equation (33) illustrates the mismatch error shaping property. In discrete-time domain, the DAC output is extracted from the vector quantizer output as

$$v_d(n) = ((\bar{\varepsilon})^T + 1) \cdot \mathbf{sv}(n). \tag{34}$$

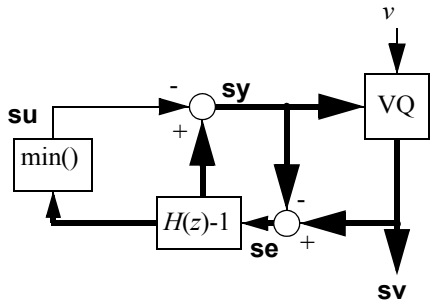



**Fig. 12. Vector-based DEM.**

The vector-based DEM in Fig. 12 has the following undesired properties:

- $N$ parallel signals within a loop, where $N = 2^B$-1
- internal word-sizes that are dependent on the filtering function H
- the system can also be unstable.

The simplest form of vectorized DEM realizes a spectral shaping function H(z) of $1 - z^{-1}$. Examining the element selection vector **sv** reveals a cyclical element selection, a behavior which is characteristic in DWA. This also confirms a clear connection between Delta-Sigma modulation and DWA, which will be further discussed in Ch. 6.

## 3.5 Summary

This Chapter reviewed the main approaches to cancelling DAC mismatch related in-band tones. Digital error correction aims to cancel the mismatch noise and DEM methods convert the narrow-band mismatch noise into wideband noise.

From the practical implementation point-of-view, the trade-off between complexity and performance is crucial in Dynamic element matching. This is the reason why first-order shaping methods with tone cancelation techniques of some kind are widely published, especially DWA.

The next Chapter discusses DWA in more detail, aiming to generalize the DWA algorithm into a more general first-order mismatch noise shaper. Chapter 5

strives to find a proper solution to this behavioral simulation based problem, namely with the theme of qualifying a DWA-variation or an arbitrary DEM method. Comparison may be quite straightforward in a the sense of hardware, but the actual performance qualification may be a surprisingly complicated task.

# 4 Generalized Data Weighted Averaging Shaping Algorithm

## 4.1 Introduction

In this Chapter, the data weighted averaging (DWA) algorithm [11,12] is presented in a generalized form (Paper I). In its basic form, the DWA algorithm realizes a first-order high-pass spectral shaping function of the form:

$$H_{DWA_1}(z) = 1 - z^{-1}. \tag{35}$$

There are publications that have extended the basic high-pass shaping function (35) into band-stop form [60-62]. An extra notch in the shaping function [60,61] decorrelates the error from the input signal, resulting in the cancelation of spurious tones but with a trade-off of an increased mismatch noise floor. A DWA-DAC may also have an extra unit element [74], which can reduce the in-band spurious tones at certain amplitudes. A generalized DWA algorithm described in Paper I integrates the aforementioned DWA versions into one elegantly simple scheme, that realizes first-order shaping with a mismatch noise shaping transfer function of the form:

$$H_{DWA}(z) = 1 \mp z^{-D}, D \in Z_+ . \tag{36}$$

The advantage of having a DWA algorithm that realizes the shaping function (36) is the possibility to study the spurious tones of different DWA shaping functions. For instance, the algorithm makes it easy to test how different shaping functions respond to a particular DWA tone-canceling modification.

To define DWA as expanded in (36) requires some justification. What makes a certain DEM method DWA is that the basic construction contains cyclical element selection realized by a modulo-*N* accumulator, a thermometer encoder and a barrel shifter. In Fig. 13, the implementation scheme for the generalized DWA DAC is shown. Very little is changed in comparison with a traditional high-pass DWA DAC implementation (e.g. the one presented in [38]).

Further justification on the generalization is the constancy of spurious tone behavior when changing from low-pass to band-pass DWA. This will be discussed

in Sect. 4.5 and utilized in the tone prediction model presented in Ch. 6. The new blocks brought by the generalization in Fig. 13 are:

– selection of direction control sequence *seq*
– the multiplier (here: polarity-multiplier)
– delay term $z^{-D}$.

In Fig. 13, [I] the delayed accumulator output *ptr* and the control sequence *seq* dictate the amount and direction of the shift in the barrel-shifter, respectively.

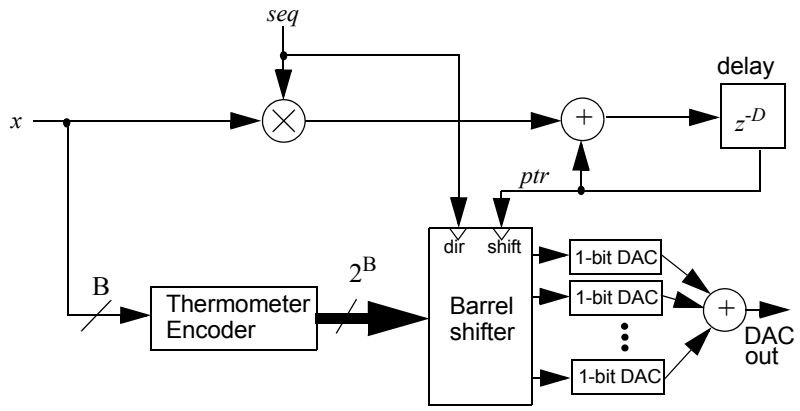



**Fig. 13. The implementation scheme for general DWA algorithm. Revised from [I]. ©[2010] IEEE.**

## 4.2 Generalized behavioral model

For a DWA-DAC with $N$ unit-DACs, the number of levels from the quantizer output is $N$+1. Here, the quantizer output is assumed to be in unsigned integer format. The number of selected elements is from zero to $N$ elements. Modulo-$N$ arithmetic inside the DWA algorithm limits the data range in from 0 to $N$-1.

The z-domain representation of a basic DWA shaping function (35) is explained in many publications, e.g. in [64]. Typically the explanation has been constructed using a pointer function $ptr(n)$ which at sample $n$, points to the DAC unit element from which the DAC selection may proceed. In this work, we are using a different strategy: a selection vector. A selection vector $\overrightarrow{SEL}(n)$ is an address vector that dictates the order in which the DAC elements are chosen. The

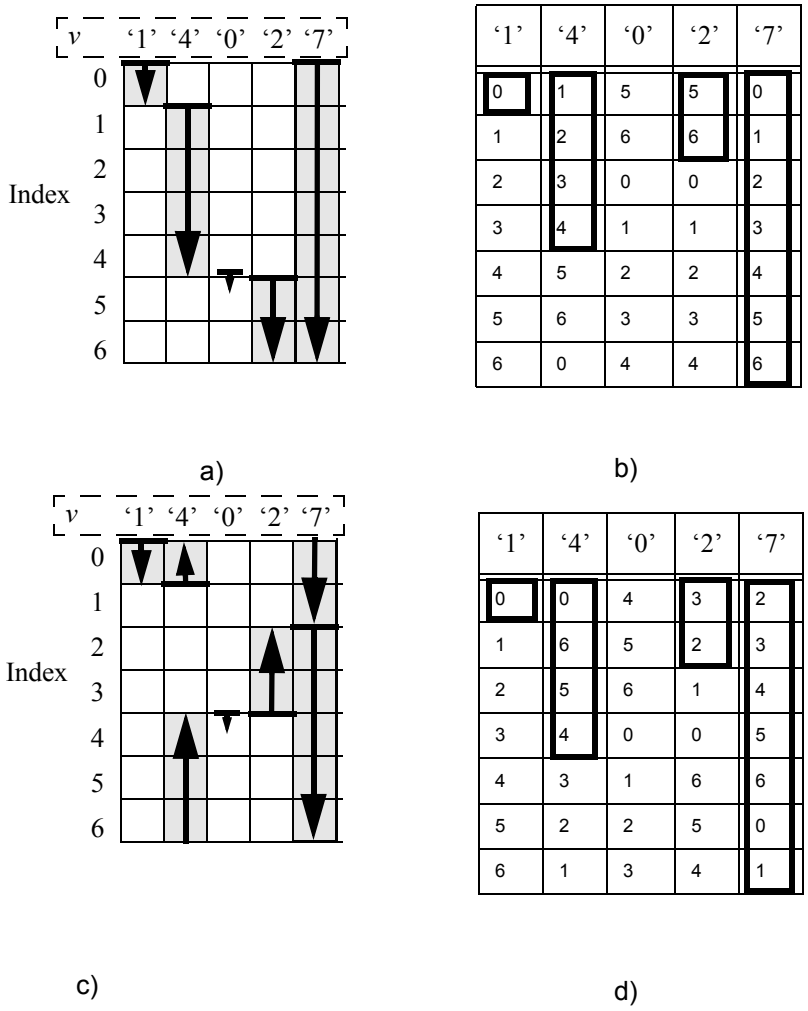
selection order is prioritized i.e. the first index in $\overrightarrow{SEL}(n)$ has the highest priority and the last one the lowest. The connection to more traditional pointer notation $ptr(n)$ is

$$\overrightarrow{SEL}(n) = ptr(n) + \vec{K}, \text{ mod N , where} \tag{37}$$

$$\vec{K} = [0, 1, \ldots, N]^T. \tag{38}$$

In the aspect of element selection, the behavioral difference between the plus- and minus-signs in (36) can be depicted by pointer diagrams. Here, the delay term $D$ is unity. The difference is in the direction of the unit DAC selection pointer. We begin with a pointer diagram of the high-pass shaping function of (35), as shown in Fig. 14a: the vertical numbers denote a zero-based index of 7 unit-DAC elements. Consecutive signal input values $v$ (five samples in all) are '1', '4', '0', '2' and '7'. The beginning of the arrow denotes the actual selection. For the second sample in Fig. 14a, the last index of the previous sample is 0 and the input signal is '4', so the current DAC output is the sum of unit-DACs of indices 1 to 4. The corresponding selection vector $\overrightarrow{SEL}$ can be seen in Fig. 14b, where framed areas highlight the index of chosen unit element.

A pointer behavior for function $1+z^{-1}$ is shown in Fig. 14c. The signal values $v$ are the same '1', '4', '0', '2' and '7'. The element selection pointer changes its direction at every sample instant, resulting in the selection vectors shown in Fig. 14d. This change in direction led to the naming convention used here: back-and-forth (BNF) DWA.




| '1' | '4' | '0' | '2' | '7' |
|---|---|---|---|---|
| 0 | 1 | 5 | 5 | 0 |
| 1 | 2 | 6 | 6 | 1 |
| 2 | 3 | 0 | 0 | 2 |
| 3 | 4 | 1 | 1 | 3 |
| 4 | 5 | 2 | 2 | 4 |
| 5 | 6 | 3 | 3 | 5 |
| 6 | 0 | 4 | 4 | 6 |

| '1' | '4' | '0' | '2' | '7' |
|---|---|---|---|---|
| 0 | 0 | 4 | 3 | 2 |
| 1 | 6 | 5 | 2 | 3 |
| 2 | 5 | 6 | 1 | 4 |
| 3 | 4 | 0 | 0 | 5 |
| 4 | 3 | 1 | 6 | 6 |
| 5 | 2 | 2 | 5 | 0 |
| 6 | 1 | 3 | 4 | 1 |

**Fig. 14. Two DWA pointer functions depicted as pointers and selection vectors: a) normal DWA, selection pointer, b) normal DWA, selection vector, c) BNF-DWA, selection pointer, d) BNF DWA, selection vector. Revised from [I]. ©[2010] IEEE.**

The DWA algorithm $1\text{-}z^{-D}$ can be expressed using the pointer function:

$$ptr(n) = ptr(n-D) + v(n-D), m \bmod \mathrm{N}\,. \tag{39}$$

Using the selection vector $\overrightarrow{SEL}$, (39) can be presented as:

$$\overrightarrow{SEL}(n) = \overrightarrow{SEL}(n-D) + v(n-D), \bmod \mathrm{N}. \tag{40}$$

The reason for using the selection vector is that the generalized model has to accommodate the change in direction. To mathematically explain the selection vector map in Fig. 14c (for $D$=1), the vector needs to be flipped upside-down at every sample synchronously with the direction change. In Fig. 14c, the last element allocated for current DAC output is the first element for the next output. This requires that the previous element selection vector $\overrightarrow{SEL}(n\text{-}1)$ should be flip upside down and be summed sequentially with either $+v(n\text{-}1)$ or $-v(n\text{-}1)$, (modulo-$N$). Hence, the selection vector for BNF-DWA in Fig. 14d is formed as:

$$\overrightarrow{SEL}(n) = \mathrm{flip}\{\overrightarrow{SEL}(n-1)\} + seq(n-1)\cdot v(n-1), \bmod \mathrm{N}, \tag{41}$$

where $flip\{...\}$ operation flips the selection vector upside-down and $seq(n)$ is a polarity multiplicand for input $v$. For the algorithm in (41) (Fig. 14d), $seq(n)$ is {...,1,-1,1,-1,...}. As an example of the agreement between (41) and the example in Fig. 14d, the selection vector for the second code '4' is formed in Table 1. Here, $v(n\text{-}1)$ and $seq(n\text{-}1)$ are both valued 1.

**Table 1. Selection vector calculus for input '4' in Fig. 14d.**

| $\overrightarrow{SEL}(n\text{-}1)$ | flip{$\overrightarrow{SEL}(n\text{-}1)$} | flip{$\overrightarrow{SEL}(n\text{-}1)$}+1 | flip{$\overrightarrow{SEL}(n\text{-}1)$}+1, Modulo-7 |
|---|---|---|---|
| 0 | 6 | 7 | 0 |
| 1 | 5 | 6 | 6 |
| 2 | 4 | 5 | 5 |
| 3 | 3 | 4 | 4 |
| 4 | 2 | 3 | 3 |
| 5 | 1 | 2 | 2 |
| 6 | 0 | 1 | 1 |

To proceed with a more general manner, (41) can be expanded to arbitrary delay $D$:

$$\overrightarrow{SEL}(n) \;=\; \text{flip}\{\overrightarrow{SEL}(n-D)\} + seq(n-D)\cdot v(n-D),\, m\text{od N}\,. \qquad (42)$$

The flip-operation and the polarity-multiplicand *seq* are codependent. Moreover, the sequence in *seq*(*n*) is dependent on the delay term in BNF-DWA.

By choosing one direction i.e. *seq* is $\{1, 1, \ldots\}$ or $\{-1, -1, \ldots\}$ leads to a very basic unidirectional DWA (35). To generalize the definition of the control sequence *seq*, the main rule is that every *D*th sample has the same sign. The general rule for the pointer direction is:

– For $1\text{- z}^{-D}$, every Dth sample in *seq*(*n*) has the same direction (Fig. 14a).
– For $1\text{+z}^{-D}$, every Dth sample in *seq*(*n*) has the opposite direction (Fig. 14c).

Equations (40) and (42) define the functionality of the DWA algorithm as a behavioral model. The model that defines both equations is shown in Fig. 15. Here, the flip operation is performed outside of the loop by summing up the pointer data *ptr* with a constant vector $\vec{K}_2$, which is either $\vec{K}$ or $flip\{\vec{K}\}$.

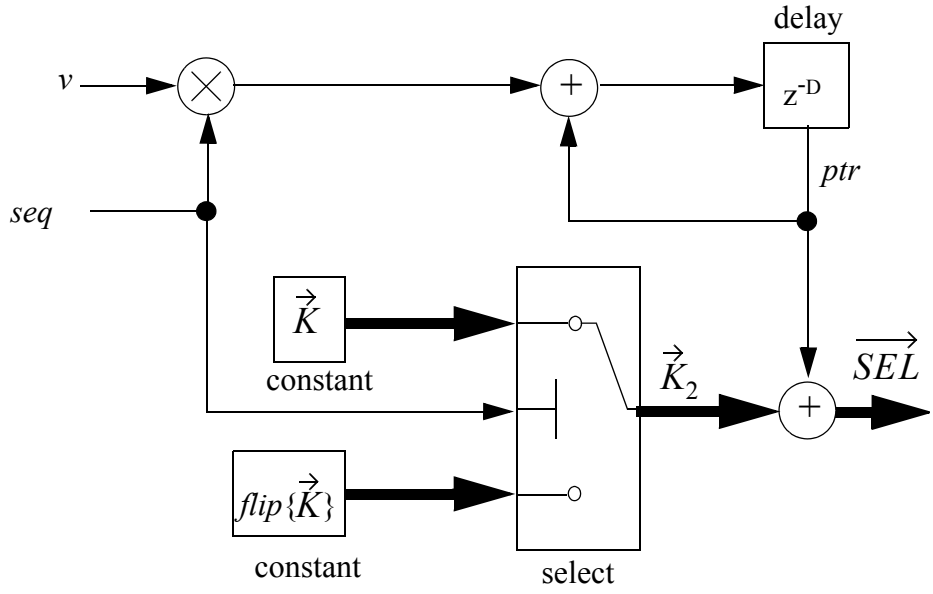



**Fig. 15. General DWA as a non-vectorized loop, where the positive value of *seq* selects *K* and the negative selects *flip*{*K*}. Revised from [I]. ©[2010] IEEE.**

### 4.3 Generalized DWA algorithm and shaping function

Equations (40) and (42) can be expressed as one unified algorithm with the pointer notation:

$$ptr(n) = \pm ptr(n-D) + seq(n-D) \cdot v(n-D) + \text{offset}(n-D), \text{ mod N}. \qquad (43)$$

The parameter *offset* in (43) is non-zero only in the case of BNF-DWA; it is zero for $1-z^{-D}$ DWA. In (43), the '±'-sign combines both DWA modes: plus is for $1-z^{-D}$ and minus is for $1+z^{-D}$. In order to define the spectral shaping transfer function for $1+z^{-D}$, we need to present the flip operation in (42) mathematically. In the general case, given that the vector $\overrightarrow{SEL}$ is modulo-*N* ascending or descending, the flip operation can be presented via negation as:

$$\begin{aligned} flip\{\overrightarrow{SEL}\} &= \overrightarrow{SEL} \cdot (-1) + \text{offset, mod N, where} \\ \text{offset} &= \overrightarrow{SEL}[N - idx_{\overrightarrow{SEL}=0}] \quad \text{and} \end{aligned} \qquad (44)$$

$idx_{\overrightarrow{SEL}=0}$ is the position of the value zero in vector SEL (zero-based indexing used).

The mismatch shaping function can be derived from (43). The input signal is multiplied by the control sequence *seq*(*n*) consisting of plus and/or minus ones. The ± sign of the product and additional offset are additions for a back-and-forth pointer. These properties will assimilate to the linear part of the resulting output signal. The linear and nonlinear part of the DAC ($y_s$ and $y_e$, respectively) output become:

$$\begin{aligned} y_s &= (ptr(n) \mp ptr(n-D)) \cdot E\{\bar{\varepsilon}\} \\ y_e &= INL\{ptr(n)\} \mp INL\{ptr(n-D)\}, \end{aligned} \qquad (45)$$

The output signal is the sum of the linear and nonlinear parts, which in z-domain is:

$$\begin{aligned} Y(z) &= Y_s(z) + Y_e(z) \\ &= (1 \mp z^{-D})PTR(z) \cdot \overline{w} + (1 \mp z^{-D})Z\{INL\{ptr(n)\}\} \end{aligned}, \qquad (46)$$

where $Z\{...\}$ stands for z-transform. Clearly, the nonlinear part $Y_e$ is shaped by the generalized mismatch shaping function of (36).

A transfer function for any spectral shaping DEM is based on a linear model that does not take the input signal dependency into account. Tonality is further discussed in Sect. 4.5 and Chapters 5 and 6.

## 4.4 Integration with previously published methods

As mentioned in Ch. 3, the generalized DWA integrates some previously published methods. Bidirectional DWA (Bi-DWA) [60] offers a tone-free mismatch noise spectrum, with the trade-off of an increased noise floor. The pointer diagram for Bi-DWA presented in [60] revealed that the mismatch shaping transfer function has to be $1-z^{-2}$. This means that the spectrum has notches at dc and $f_s/2$. The extra notch at $f_s/2$ increases the in-band noise floor, but scrambles the spurious tones effectively. The bi-directionality is a choice that is offered by the switching sequence choices presented in Sect. 4.2. However, the change in the direction for Bi-DWA does not offer any advantage. The "legal repeating sequences" presented in Sect. 4.2 offer $2^D$ candidates for *seq*.

N-path DWA ($1-z^{-D}$) by Lindfors *et al.* [61] consists of $D$ time-multiplexed DWA blocks. The result for $D$=2 is equivalent to Bi-DWA and for band-pass DSMs (with $f_c=f_s/4$), $D$=4 can be used. This leads to three notches at dc, $f_s/4$ and $f_s/2$.

Interestingly, Lindfors *et al.* [61] discarded the $1+z^{-2}$ band-pass shaping function, because it would require element inversion and thus cannot be realized with a current-steering DAC. In the generic model, an element inversion is not used thus the digital DWA encoder can be designed to produce the $1+z^{-2}$ shaping. This shaping function of $1+z^{-2}$ is accomplished by the DWA algorithm with back-and-forth element selection rotation published in [62] and it was implemented using a current-steering DAC.

## 4.5 Shaping examples and further concerns

The spectral shaping properties of any DEM method are found by linear analysis. As pointed out in Sect. 3.4, the performance prediction by linear analysis can only be as a suggestive tool.

Lindfors *et al.* [61] pointed out by linear approximation that for 4-path DWA, the in-band noise power is ca 6dB lower than for $1+z^{-2}$ DWA. The extra notches

(at dc and $f_S/2$) in N-path DWA, however, tend to reduce the correlation between the error and input signal: the white noise approximation is better for N-path DWA.

Table 2 contains simulation results from three different DS DACs with three different forms of DWA shaping, but the same unit-DAC DNL errors. The performance merits in Table 2 are SNDR (signal to noise and distortion) and SFDR (spurious-free dynamic range). The mismatch was given in a simulator by assigning a third order polynomial form of DNL error with $\sigma_\varepsilon$ of 0.01 (1%). This shape assignment is further discussed in the next Chapter.

The corresponding mismatch noise spectra are shown in Fig. 16. The simulations in Fig. 16b and Fig. 16c are both from a band-pass DS DAC, but the spectral shaping function is different in each. In Fig. 16b, we have the BNF-DWA shaping [62] and in Fig. 16c the configuration is the 4-path DWA [61] that has extra notches at dc and $f_S/2$.

**Table 2. Performance simulations of different 3-bit DS DACs with DWA.**

| Figure | DSM order, $f_c$ | OSR | SNDR | SFDR | DWA transfer function | seq(n) |
|---|---|---|---|---|---|---|
| 16a | 2, dc | 32 | 69dB | 82dB | $(1-z^{-1})$ | {1,1,1,1,...} |
| 16b | 4, $f_S/4$ | 32 | 70dB | 85dB | $(1+z^{-2})$ | {1,1,-1,-1,...} |
| 16c | 4, $f_S/4$ | 32 | 64dB | 81dB | $(1-z^{-4})$ | {1,1,1,1,...} |

What is apparent in Fig. 16a (low-pass DS DAC) and Fig. 16b (band-pass DS DAC) is that the DWA algorithm is prone to spurious tones in both cases. As expected, the extra notches in Fig. 16c smoothen the spectrum quite nicely and the overall in-band noise is elevated by ca. 6dB (see Table 2).

The first concern is that in Table 2 the spurious level in Fig. 16c is 4dB lower than in Fig. 16b. Does that mean that the 4-path version does not improve the SFDR results (at the expense of the elevated noise floor) as may be expected? In reality, changing the shape of DNL error, input signal amplitude and frequency may result in a different conclusion. Chapter 5 proposes a qualification approach that takes these three factors into account.

The second concern can be seen by comparing Fig. 16a and Fig. 16b. The cluster of spurious tones near $f_S/2$ seem to be very similar: Fig. 16b is mirroring the

low-pass version at the center frequency $f_s/4$. The frequency range of spurious tones seems to match and this was achieved by normalizing the input single-tone frequency $f$ by the corresponding conversion:

$$f \rightarrow \frac{f}{2} + \frac{f_s}{4}. \tag{47}$$

Also, for the spurious tone cluster to match between Fig. 16a and Fig. 16b, the DNL error has to be identical.

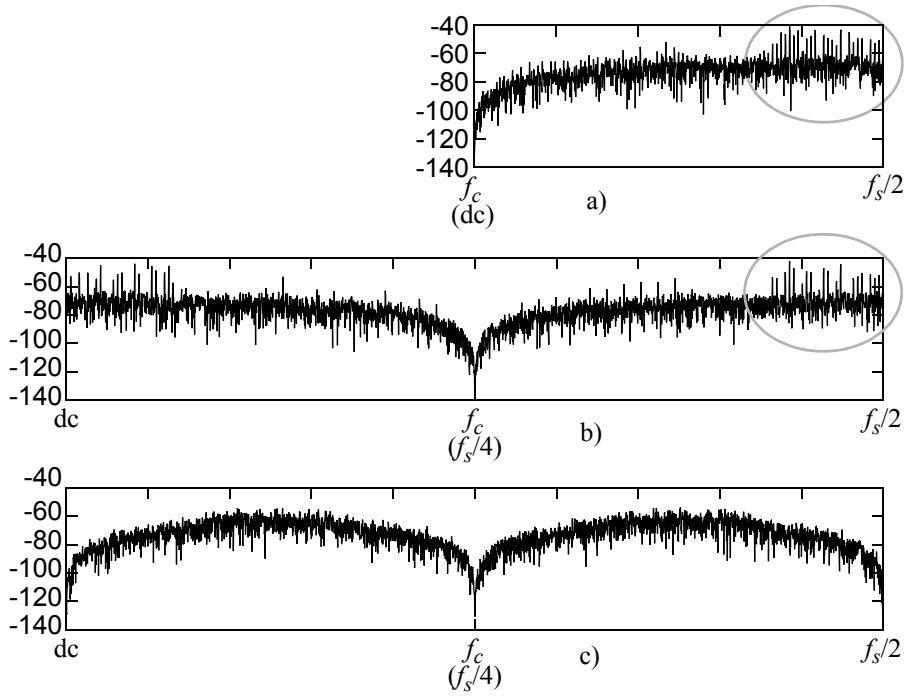



**Fig. 16. The mismatch noise magnitude spectrum for a) $1-z^{-1}$ DWA, b) $1+z^{-2}$ DWA and c) $1-z^{-4}$ (4-path DWA).**

Incremental DWA (IDWA) [74] is a curious DWA variant which will be revisited in the last two Chapters. The basic idea for IDWA is to increment the number of unit DACs e.g. by one. In other words, the number of unit-DACs ($N$) is larger than the number of quantizer steps in DSM.

Using IDWA will not change the spectral shaping properties or the actual DWA algorithm. Instead, IDWA shifts the spurious tone frequencies from their

typical location. For low-pass DSM with DWA shaping of $1-z^{-1}$ and a certain shape of DNL error, a cluster of spurious tones may be located near dc. Incrementing the number of unit elements by one, the frequency range of this cluster of tones is conveniently shifted by $f_s/N$. IDWA was studied in the original papers II and III. It was concluded in Paper II, that IDWA can perform better, but only for low-pass DS converters.

As an example, in Fig. 17, we have the magnitude spectrums for a second order low-pass DS DAC with an oversampling ratio of 64. The amplitude is 0.126 per full scale and the frequency is $f_s \cdot 249/2^{16}$. For this particular choice of order, OSR, amplitude and frequency, IDWA shifted the spurious tone out of the band of interest. In Fig. 17, the SFDR result is ca. 20 dB better for IDWA.

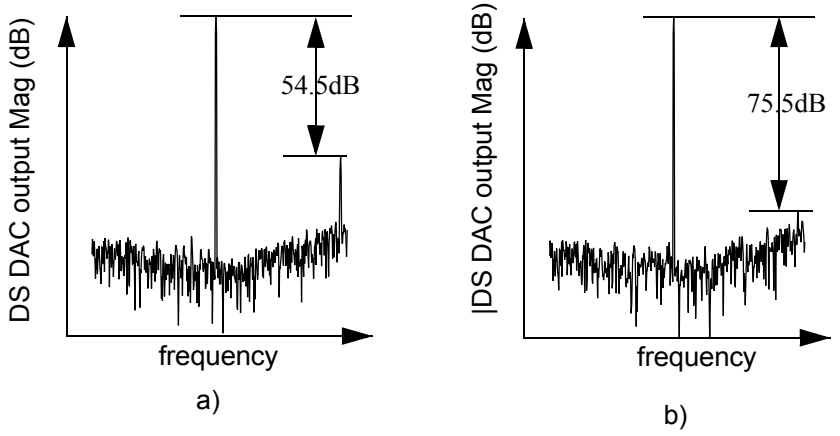



**Fig. 17. The in-band magnitude spectrum for 3-bit low-pass DS DAC output with OSR of 64: a) DWA and b) IDWA.**

## 4.6 Summary

The extension of the DWA algorithm via (36) is justified by the fact that the realization has the basic components of a high-pass DWA (35). The extended concept of DWA results in the interesting integration of DWA variants such as Bi-DWA [60], N-path DWA [61], back-and-forth ($1+z^{-2}$) DWA [62] and also IDWA [74]. The generalization may also expand the possibilities of existing DWA tone-cancelling variants to be used in band-pass DSMs. The model has been published

at the Matlab Central File Excange[2] under the title “Data Weighted Averaging for Simulink”.

The distortion that manifests itself as spurious tones is linked to the shape of the differential nonlinearity (DNL) error. The true DNL error variation can be coarsely predicted by the layout process, but as the number of quantizer levels is typically low in DSMs, the error shape is random. To qualify and compare different DWA variants or DEM methods in general, the randomness of the mismatch can be troublesome: the simulation test bench should be reproducible and reliable. The next Chapter will propose a solution for qualifying a DEM method.

[2] http://www.mathworks.com/matlabcentral/fileexchange/?term=authorid:17648

# 5 Reliable DEM qualification

## 5.1 Introduction

As in the case of DSMs, the appearance of the spurious tones in the DEM mismatch noise can be dependent on the DSM stimulus. A typical test stimulus in simulations is a single-tone, whose amplitude and frequency affect the spurious tones related to both DSM and DAC. As the shape of the DAC mismatch is not exactly known, this results in problems related to reliability, repeatability and simulation time.

The subject of qualification covered in Paper II, can be expanded to practically any DEM method. Here, however, the examples are biased toward data weighted averaging (DWA) for which a proper qualification method is direly needed. The reason for this is the vast number of tone-cancelation methods proposed for DWA.

As demonstrated in Sect. 3.4, the simulation-based approach can be troublesome: a performance result from one single-tone simulation may not provide a reliable conclusion. However, simulation is the pivotal tool in characterizing the performance, like in the case of DSMs.

There are two performance merits discussed in this Chapter: SNDR (signal to noise and distortion) and SFDR (spurious-free dynamic range). The former includes the integral of all in-band noise content and the latter includes the highest spurious tone in the signal band. SNDR or SFDR curves are typically shown as a function of logarithmic amplitude range. Especially for the DWA tone-cancelation variants, this approach is inadequate: there are pairs of single-tone stimulus' amplitude and frequency that can be significantly detrimental to the performance [50].

In a typical simulation test bench, the mismatch (DNL error) has a constant rms-magnitude. Here, it is pointed out that the shape of the unit-DAC DNL error can be a factor of significance. A DEM method can be insensitive to the shape of the mismatch, but this is certainly not the case in DWA and other low-order shaping schemes.

To rely on a single randomly selected DNL error vector is by no means a reliable approach. Relying on a large group of randomly generated DNL error shapes is statistically more reliable, but this approach can be very time-consuming and typically non-repeatable. Of course, a seed number can be provided for random error vectors but this is hardly practical for a large number of error vectors (e.g. in a publication).

Testing a DEM method may become extremely time-consuming with the stimulus amplitude and frequency, the shape of the mismatch with some extra parameters affiliated with the algorithm. Comparing the simulation results of competing techniques in a reliable fashion would require them to be tested with the similar test bench as well.

A solution to the issue of repeatability proposed here is to use a extensive range of amplitudes and frequencies, along with a small group of repeatable unit-DAC DNL error vector shapes.

## 5.2 DNL error test polynomials

In a simulation test bench, a polynomial DNL-error curve can be selected as opposed to a large group of randomly generated DNL error sets [50]. This speeds up the evaluation of a specific DEM method, but a skeptic would say that a large group of randomly generated DNL-errors is more reliable. Some information about the error distribution can be available, but the actual shape cannot be predicted. For instance, one can expect random, linear ($\chi^1$) or quadratic ($\chi^2$) trend in the DNL error after typical common-centroid [41] layout technique. Possible error skewness ($\chi^1$, $\chi^2$, etc.) does not remove the random portion of a single unit DAC error. However, a small group of different polynomial shaped DNLs provides a repeatability and a small set of errors in DEM evaluation.

In general, the shape of a DNL test-curve can be other than a polynomial. Repeatability is the main issue here. For instance, a sinusoidal DNL error shape reveals a fundamental property of DWA-related tones and this will be discussed in the next Chapter.

To construct an *m*th order polynomial curve, we begin by defining a normalized column vector $\overline{\psi}$ also denoted as $(\overline{\psi}_i)$. The vector $\overline{\psi}$ is linearly spaced from +1 to -1 and its length is the same as the number of DAC elements. The vector $\overline{\psi}$ is then raised element-wise to *m*th power:

$$\bar{\varepsilon}_1 = (\psi_i^m) . \tag{48}$$

The next step is to scale the vector $\bar{\varepsilon}_1$ curve and the scaling is typically made according to a user-defined standard deviation value $\sigma_\varepsilon$:

$$\bar{\varepsilon}_2 = \sigma_\varepsilon \cdot \frac{\bar{\varepsilon}_1}{\mathrm{stdev}(\bar{\varepsilon}_1)} \quad , \tag{49}$$

where stdev($\bar{\varepsilon}_1$) is the standard deviation of $\bar{\varepsilon}_1$. Finally, the mean value of the curve is normalized to zero and the outcome is the DNL-error curve used in the behavioral model:

$$\bar{\varepsilon} = \bar{\varepsilon}_2 - \text{mean}(\bar{\varepsilon}_2), \tag{50}$$

where mean($\bar{\varepsilon}_2$) is the average of $\bar{\varepsilon}_2$.

For a group of SNDR or SFDR values from $J$ DNL error curves, the average performance and noise related values are calculated by the $J$-sized group of decibel values, using the following equation:

$$X_{dBA} = 10 \cdot \log_{10}\left(\frac{1}{J} \cdot \sum_{i=1}^{J} 10^{X_i/10}\right), \tag{51}$$

where ($X_i$) contains a group of decibel-valued samples and $X_{dBA}$ is the power averaged decibel value. In our simulations, $J$ equals 4 (four polynomials).

To qualify a DEM method using the repeatable test shapes, it is imperative that the shapes will not result in an overly optimistic performance results compared to the results form a large group of random errors.

It has been verified that the average results from 100 randomly generated DNL sets studied in [50] are slightly more optimistic in comparison with results from a set of 4 polynomial DNL shapes. These shapes are justified in Sect. 5.4 by using the performance degradation merit described next.

### 5.3 DEM performance degradation merits

The reference point for the performance results is the SNDR and SFDR result with the same amplitude-frequency points with zero-DNL error. A zero-DNL performance reference can be justified by the fact that the spurious tone caused by the DSM is always present in DS-DACs and ADCs.

s the stimulus amplitude and frequency are swept, the magnitude of maximal in-band spurious tone and the in-band cumulative noise will change. In Fig. 18, $\Delta S$ and $\Delta N$ indicate the amount of variation (dB) in maximum in-band spurious tone and the in-band cumulative noise, respectively.

To merit the performance, we also need to merit the deterioration of both SNDR and SFDR. These results are compared with their zero-DNL counterpart and the deterioration is denoted as $\Delta SFDR$ and $\Delta SNDR$. These are defined as:

$$\Delta SFDR = SFDR_{ideal} - SFDR \text{ and} \tag{52}$$

$$\Delta SNDR = SNDR_{ideal} - SNDR, \tag{53}$$

where the subscript *ideal* is the result unaffected by the unit DAC. For a simulation model of a DS DAC, the ideal result can be obtained from the DAC input (digital DS modulator output). For a DS ADC simulation model, the *ideal* result can be obtained from the feedback DAC's output (given that the DSM remains stable with a given stimulus and mismatch error).

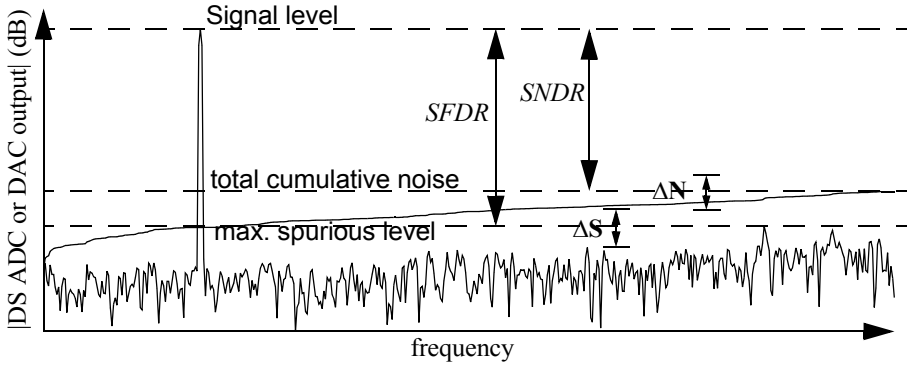



**Fig. 18. An example of the signal band of a DAC output spectrum. Revised from [II]. ©[2010] IEEE.**

In a broader sense, it would be convenient to have some kind of measure of how prone an arbitrary DEM method is to spurious tones. This is partly possible by observing the DAC element usage patterns. This was also studied in Paper II.

A selection (priority) vector concept presented in Ch. 4 can be expanded to other DEM methods. The priority vector $\overrightarrow{SEL}$ contains the address of the highest priority unit-element in its first element, and the lowest priority address in its last element. By investigating an array of successive index numbers of the highest priority element, repeating element usage patterns can be found. The repeating patterns may be an indication of spurious tones without guessing the shape of the DNL error.

A simpler way of finding patterns is to observe the spread of the top usage priority provided by the DEM algorithm. If the highest priority element usage does

not spread evenly (amongst all elements), it is likely that there will be spurious tones.

The characterization based on element usage presented in Paper II correlates with the signal dependency of the DWA algorithm but unfortunately does not predict the performance. The key on breaking the inherent tendency to in-band spurious tones is in the shape of the DNL error, and this will be discussed in the next chapter.

## 5.4 DEM fingerprints

As mentioned, characterization based on single-tone simulations should be obtained from a two-dimensional amplitude and frequency sweep. A sweep contains equally spaced frequencies throughout the signal band and an extensive range of logarithmically spaced amplitudes that are within the stable input signal range.

A graph called as a fingerprint is based on tabulated values of any performance degradation merit $\Delta S$, $\Delta N$, $\Delta SFDR$ or $\Delta SNDR$ as a function of amplitude and frequency. It is presented as a contour plot from which a discrepancy in a certain amplitude-frequency pair can be spotted. At a glance, it looks like a spectrogram, but it contains performance information from multiple simulations. In the examples presented in Paper II, there are 4 shapes of DNL-curves, 25 amplitudes and 16 signal frequencies, yielding 1600 simulations per contour plot.

To somehow rationalize a choice of a handful of polynomial DNL-curves, we tested [II] the performance degradation with polynomial DNL errors of orders $m = \{1, 2, 3, 4\}$. The model is a second order, 3-bit low-pass DS DAC with OSR of 32.

From the contour-plot in Paper II, it was apparent that for odd orders, the contour-plots are nearly identical. The same conclusion can be made between the even order fingerprints. Because of the similarity between odd- and even-order contour-plots, the polynomial orders should be chosen between even and odd, e.g. four curves of between orders 2 and 3. A non-integer polynomial order $m$ requires to modify (48) into:

$$\bar{\varepsilon}_1 = Real(\psi_i^m), \tag{54}$$

where $Real(...)$ is the real part of the result. The length of $\bar{\varepsilon}_1$ is the same as the number of DAC elements. In the cases studied in [II], it was 7, indicating a three-

bit DAC. Non-scaled shapes of the four sets of polynomial DNL shapes used in Paper II are shown in Fig. 19.

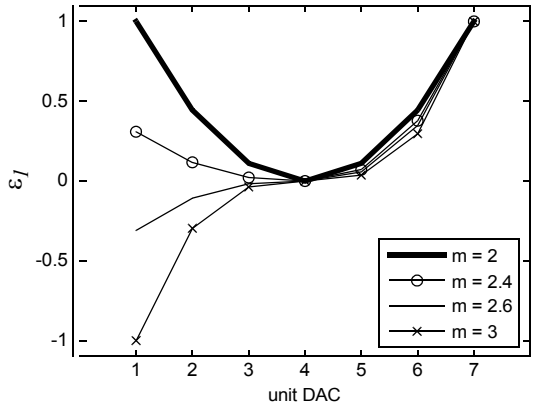



**Fig. 19. Four DNL shapes used in [II].**

The next Chapter discusses the mechanism for generating spurious tones. In the light of this theme, it is appropriate to highlight some supportive findings in Paper II. For a $1\text{-}z^{-1}$ (low-pass) and a $1\text{+}z^{-2}$ (band-pass) DWA, it was found that an amplitude level of the strongest in-band tones is approximately:

$$A_{worst} = \frac{FS}{2OSR}, \tag{55}$$

where $FS$ is the full scale input signal amplitude. In [II] it was verified that the amplitude level of the worst-case spurious tones decreases as the OSR increases.

In Sect. 4.5, the Incremental DWA (IDWA) [74] was briefly introduced. IDWA was also studied in Paper II and it was found that an 8-level IDWA (one extra unit-DAC) different yet no that severe "worst-case" amplitude from the normal, 7-level DWA (55). As mentioned in Sect. 4.5, IDWA shifts the expected frequency range of spurious tone out of the band of interest. At small amplitudes this is favorable, but a slight degradation was observed around $FS{\cdot}0.2$.

The mechanism for generating spurious tones discussed in Ch. 6 explains how DWA is dependent on the instantaneous stimulus magnitude. The model presented in Ch. 6 also incorporates IDWA.

## 5.5 Summary

The proposed graphical qualification method, or fingerprinting, provides a characterization and benchmarking for DAC mismatch shaping methods. The fingerprinting approach is a transparent way to find performance merits that can easily be hidden by e.g. simulating with fixed amplitude or frequency. Using a polynomial-shaped DNL error enables repeatable behavioral simulations and effectively minimizes the number of required simulations.

The trade-off between complexity and performance is crucial in the spectral shaping of DAC mismatch noise. This is the reason why first-order shaping methods with tone cancelation techniques of some kind are widely published, especially for DWA. The proposed qualification method is an attempt to bring consensus on how different DEM methods can be benchmarked.

Chapter 6 is the culmination of this work, it proposes a DWA tone mechanism that can predict the DAC mismatch spurious tone contribution to any stimulus and DNL error shape. By knowing how the shape of the DNL produces the in-band spurious tones, a new DWA tone cancelation approach has been discovered.

# 6 Model of DSM and DWA-Related Spurious Tones

## 6.1 Introduction

Understanding the mechanism by which spurious tone frequencies move as a function of stimulus is desirable for the purpose of defining a tone cancelation methodology. For stimuli other than dc, the tones indeed wander, making the spurious tones in DSM quantization noise or DWA-DAC mismatch noise very difficult to predict.

The nature of DSM quantization noise is troublesome. Limit cycle or spurious tones can appear with a specific stimulus even if the DSM order is high. For instance, a high-order DSM may have limit cycles that can be surprisingly persistent to dither [33]. The use of multibit DSM will reduce the DSM-related tones, but a D/A converter with DEM can also be tonal.

Section 6.2 contains behavioral observations that have led to the proposed models. Section 6.3 contains a literature review of previously published results on predicting the DSM- and DWA-related idle channel tones. The similarity between DSM and DWA-DAC tone behavior has been acknowledged in prior publications.

The uniform model proposed here supports the observations from previous publications and in Sect. 6.4 expands into a simple the uniform model for both DSM and DWA-DAC related spurious tones. Section 6.5 contains case-studies on tone estimation.

## 6.2 Observations on tone behavior

Contrary to Gray's spectral analysis approach [4], the model presented in Sect. 6.4 is a time-domain model that can be excited with an arbitrary stimulus. The quantization noise model in Sect. 6.4 is, however, in line with Gray's Bessel function model. The proposed periodic quantizer model

– excited by a dc input (for low-pass DSM) results in a group of tones with fixed frequencies and

– excited by a non-dc input (e.g. single-tone) results in frequency modulation spectrum with tone frequencies that vary as a function of time.

The proposed tone estimation model for both DSM and DWA were found by simulation results with a slowly changing input signal. With inputs like dc, ramp and certain single-tones, the spurious tones in the modulation noise become more observable. The choice of stimulus also contributes to the whiteness of the quantization noise, as was mentioned in Ch. 2.

The initial insight for the proposed model was discovered by simulating a group of DSMs with some special group of stimuli and observing the modulation or mismatch noise in a spectrogram view. As the spectrogram is based on multiple discrete Fourier transforms, the stimuli were carefully chosen so that the spurious tones are observable. In the spectrogram analyzes, the data of interest is

– the DSM output data that contains the signal and modulation noise (Subsection 6.2.1) or

– the DWA-shaped DAC mismatch noise (Subsection 6.2.2).

The vertical direction in a spectrogram illustrates how the spectrum varies with the simulation time. By increasing the dc-magnitude or single-tone amplitude linearly, the observable tones can be studied as a function of stimulus magnitude.

The frequency of a single-tone stimulus $f$ can be presented as:

$$f = \frac{J \cdot f_s}{M} + f_c, J = \{0, 1, 2, \ldots\}, \tag{56}$$

where $M$ is the number of samples in a simulation. To emphasize observability, $J$ has to be small. This also applies to band-pass DSMs, where the center frequency $f_c$ is typically $f_s/4$. Obviously, $J = 0$ in (56) refers to dc input to low-pass DSMs.

Interestingly, for band-pass DSMs, $J = 0$ in (56) results in dc-like tone behavior. To generalize a the concept of dc-input or "slowly varying input" for band-pass DSMs as well, the statement in Subsection 2.3.2 should be rephrased as: "signals close to the DSM center frequency are not suited for the classical model of quantization".

### *6.2.1 DSM modulation noise*

In Fig. 20, we have four simulation results as spectrograms of the DSM output. A spectrogram consists of a number of short-time Fourier transform magnitudes depicting how the magnitude response changes as a function of time. The result is

presented as a surface plot. In each example of Fig. 20, the number of samples $M$ (56) is $2^{18}$ and the FFT length is $2^9$. The quantizer step $\Delta$ is 2. The lightest shades of the color map in Fig. 20 indicate the highest magnitude in decibels. The magnitude variation in the spectrograms is ca. 100 dB.

In Fig. 20, the input level or amplitude increases linearly with time, which therefore replaces the vertical time-axis. This enables observing the spurious tones as a function of instantaneous input level of the stimulus.

All examples were generated using the second order DSM. The order for band-pass examples in Fig. 20c and Fig. 20d are also 2, but the shaping order is one. These spectrograms were generated using the mid-rise quantizer type (Subsection 2.4.2). In each example of Fig. 20, the number of samples $M$ in (56) is $2^{18}$ and the quantizer step $\Delta$ is 2.

In Fig. 20a, a low-pass DSM output data spectrogram for a ramp input (i.e. a linearly increasing dc level) is shown. The trajectory of the spurious tones clearly change as a function of the input level. In Fig. 20b, the spectrogram for sinusoid input as a function of increasing amplitude is shown. The envelopes of the spurious tones are the same as the trajectories in Fig. 20a. In Fig. 20a and Fig. 20b, the frequency parameter $J$ is 3 in (56).

For a band-pass DSM, the spurious tones can also be either trajectories or envelopes. If the frequency is exactly at the band center $f_c$ ($f_c = f_s/4$), the spurious tones are trajectories similar with low-pass DSM: between dc and $f_c$ and also mirrored between $f_c$ and $f_s/2$. In Fig. 20d, the single-tone stimulus' frequency is slightly more than $f_c$, which results in envelopes that follow the trajectories in Fig. 20c. In Fig. 20d, the frequency parameter $J$ is also 3 (56).

At a linear transfer function level, transforming the NTF from low-pass to band-pass ($f_c$: dc$\rightarrow f_s/4$) can be done by first order DSM N-path transformation ($z \rightarrow -z^2$) [14]. By comparing the upper and lower spectrograms in Fig. 20, the transformation clearly preserves the tone behavior.

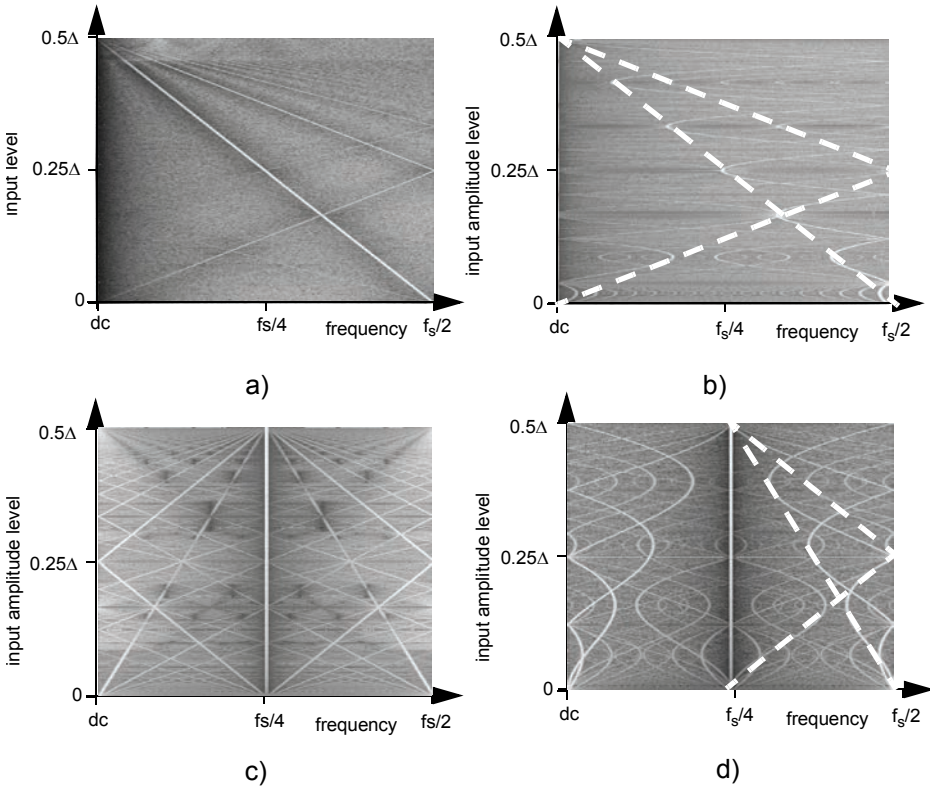



**Fig. 20. DSM output spectrograms: a 1st order low-pass with a) ramp input and b) sinusoid input, and a 2nd order band-pass with sinusoid input c) at $f_c$ and d) near $f_c$. Revised from [IV]. ©[2011] IEEE.**

As mentioned, the quantizer type for the spectrograms Fig. 20 is mid-rise. Changing the quantizer type to mid-tread changes the spur trajectories or envelopes. The shape of corresponding spectrograms for mid-tread would look the same as in Fig. 20 *except* they would be upside-down. The reason for this is the signal and amplitude levels in Fig. 20 move from between a quantizer level towards another quantizer level. For a mid-rise quantizer, the situation is the opposite. This is in-line with quantization noise observations made in Subsection 2.4.2.

For the DSM tone frequencies, changing the quantizer type has similar effects as offsetting the stimulus by Δ/2. However, Δ/2 offset may not be practical due to DSM output's limited low word length; quantizer type is more appropriate factor.

### *6.2.2 DWA mismatch noise*

In [14], Carley *et al.* pointed that for constant inputs, DWA mismatch noise frequencies will be placed at the frequency $f_s/N$ and its harmonics. This is a coarse assumption that can be can be expanded with a simple behavioral simulation.

The input signal to the DWA has to be DS-modulated data. If the input data is a single-tone only truncated to $N$+1 levels, the spurious tones will be clustered at frequencies $f_s \cdot i/N$, where $i$ = {1,2,3,...}. This is not realistic as the DSM modulation noise should be mainly out-of-band. A DS-modulated scenario can be seen in the DAC mismatch noise spectrogram plot in Fig. 21, where a slowly increasing dc-stimulus (ramp) has been Delta-sigma modulated by a second order 3-bit ($N$=7) low-pass DSM with $1\text{-}z^{-1}$ DWA. As the signal level increases, the are three moving tones with a linear slope, two of them starting from $f_s/2$ and one starting from dc. As the faintest slope approaches dc, it will be aliased back to positive frequencies.

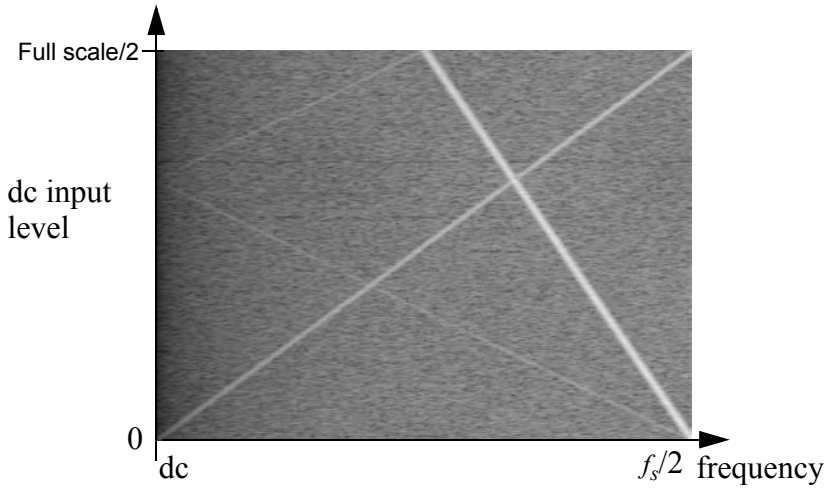



**Fig. 21. The spectrogram of DWA DAC mismatch noise for DS-modulated (3-bit) ramp input signal.**

The mismatch noise spectrogram in Fig. 21 looks much like the lower half of the spectrogram in Fig. 20a. It will be shown that almost the same model corresponds to the mismatch noise with first-order DWA shaping.

For DWA, the mechanism on how the trajectories are born has very little to do with the order and the OSR of the DS modulator. The dominant factors are (in addition to the stimulus): DWA type (low-pass, band-pass), the number of unit DAC elements along with the shape and quantity of the DAC mismatch error.

It will be shown that the shape of the DNL error is the key factor in the appearance of spurious tone. This information can be utilized to cancel the in-band spurious tones.

## 6.3 Prior publications on spurious tone behavior

### *6.3.1 DSM spurious tones*

Candy & Benjamin [13] proposed a mathematical expression for the dc-response for the first order DSM output $v$ that lie in the half sample band of frequencies can be expressed as:

$$v(t) = u + 2\sum_{i} \frac{\sin(\pi i u)}{\pi i} \cos\left(\frac{2\pi t \cdot (i \cdot u, \bmod 1)}{T}\right), \tag{57}$$

where $u$ is the dc-input level and $T$ is the sampling interval.

Ledzius and Irwin [72] demonstrated that the spurious tone is related to the dc-stimulus level, from which Dunn and Sandler [76] presented their estimation on the spurious tone frequencies near dc. Spurious tone frequencies near $f_s/2$ ($f_s$ is the sampling frequency) were anticipated by Risbo [75] as well as Norsworthy [14].

Gray's classic paper on quantization noise spectra [4] defined the following: for a first-order LP DSM and dc inputs the quantization error spectrum is discrete, with tones at

$$f_{tone} = \left(i\left(\frac{1}{2} \pm \frac{u_{dc}}{\Delta}\right), \bmod 1\right) \cdot f_s \text{ , for i = 1,2,... ,} \tag{58}$$

and the amplitude of these tones is

$$A_{spur} = \frac{1}{(2\pi i)^2} \,. \tag{59}$$

De la Rosa *et al.* [77] presented an extension of (58) into second order band-pass DSMs with single-tone stimulus exactly at the center of the signal band ($f_c = f_s/4$) and the amplitude is (notated here for the sake of similarity) $u_{dc}$:

$$f_{tone} = \begin{matrix} f_s\left(i\left[\frac{1}{4} \pm \frac{u_{dc}}{2\Delta}\right], \text{mod } 1\right), \text{ and} \\ f_s\left(i\left[\frac{1}{2} \pm \frac{u_{dc}}{2\Delta}\right], \text{mod } 1\right) \end{matrix} \quad \text{for i = 1, 2, ...} \tag{60}$$

Equations (58) and (60) are restricted to stimuli exactly at $f_c$. For second order band-pass DSM, the stimulus exactly at the band center results in limit cycle (with steady frequencies). This setup is equivalent with a first order low-pass DSM with dc-inputs.

Single-tone stimuli outside of $f_c$ will result in the tone frequencies drifting as a function of the stimulus level. These tones have the following properties:

1. The frequency span of the drifting (envelope or trajectory) is proportional to the signal amplitude.
2. The drifting velocity is proportional to the stimulus' frequency content.
3. The quantizer type (mid-rise or mid-tread) also has great significance.

In the publications mentioned in this Subsection, the quantizer type is mid-rise. The time-domain model presented in Sect. 6.4 takes the quantizer type into account.

### *6.3.2 DWA-related mismatch noise spurious tones*

The number of published DWA tone suppression methods is quite large, e.g. [25, 60-62, 74]. For the publications that actually interpret the mechanism how (untreated) DWA works, there are quite few eligible publications. Chen and Leung [57] reported the following on DWA tone behavior using observations from the FFT magnitude response of the mismatch noise:

1. There is an amplitude dependence of the spurious tones.
2. Relatively low OpAmp dc gain affects the spurious tones: dc gain should be over-designed.

3. For high OSR, dither would be a solution for DWA-related idle tones.
4. The number of DAC elements ($N$) should be less than 2OSR/($m$+1), where $m$ is the order of the harmonic distortion component.

Point 4 means that for low OSR and a large number of DAC elements, DWA mismatch noise is detrimental. Therefore, for low OSR it can be expected that the tones are more likely to drift into the signal band.

In Sect. 5.4, the approximate amplitude level of the in-band tones was given by (55). Doubling the DSM output word size doubles the elements to cycle through before the errors average to zero [12]. As regards of observable tones, doubling the word size does not change the expected tone frequencies, but increases the number of new spurious tones two-fold. This will be discussed in Subsection 6.2.2.

Chen & Kuo [74] deduced that the mismatch noise frequency spectrum is composed of discrete components of frequencies

$$f_{tone} = \frac{r}{N} \cdot f_s \cdot i, \quad \text{i=1, 2, 3, ...} \tag{61}$$

where $r$ is the value of the greatest common denominator between the number of elements $N$ and the DAC input code $v$. Being a reasonable accurate prediction, Chen and Kuo presented a clever spurious tone suppression method (incremental DWA, or IDWA), which does not alter the DWA algorithm. In IDWA, $N$ is larger than the number of DSM quantizer steps. IDWA does not change the spectral shaping properties, but manipulates the DWA-related spurious tones. An IDWA-DAC with $N_1$ extra elements shifts the tone frequencies in (61) [74] as follows:

$$\begin{aligned} f_{tone} &= \frac{r}{N+N_1} \cdot f_s \cdot i \text{ , for mid-tread quantizer} \\ f_{tone} &= \frac{r}{2(N+N_1)} \cdot f_s \cdot i \text{ , for mid-rise quantizer.} \end{aligned} \tag{62}$$

In the next Section, time-domain models for estimating the DSM modulation noise and DWA mismatch noise are presented. The spurious tones for both cases follow a simple equation that resembles frequency modulation. Nys & Henderson observed the similarity in [73] by stating:

> *For values of DC level close to 1/2, 1/3, 2/3 ,1/4, 3/4, ... of the range, a low order noise component of high power is folded back into the baseband, and*

*the equivalent resolution is decreased, in a similar way as for the quantization noise with first order Sigma-Delta modulation.*

Nys & Henderson [73] also presented a time-domain estimator for DWA-DAC mismatch noise. Unlike the model presented here, the model in [73] is only for dc-inputs and accommodates the DNL error variance (not the actual shape).

## 6.4 Noise models for DSM and DWA

The frequency prediction formulae in Sect. 6.3 output a group of frequencies. The model presented in this Chapter is a time-domain model that can be used to study the DSM modulation noise and DWA mismatch noise with an arbitrary stimulus $u$. The proposed model is also applicable to band-pass DSMs and $1+z^{-2}$ DWA.

The time-domain model for both DSM quantization noise and DWA mismatch noise is denoted by $f(u(n))$ in Fig. 22. As will be shown, the model has the same basis for both cases. The model also acquires information about the quantizer type and step $\Delta$ and, for DWA, the DAC DNL error vector. The center frequency $f_c$ in Fig. 22 affects the stimulus and the model (low-pass or band-pass mode).

The output of the modeling function is the unshaped noise $\hat{e}_q$ for the DSM and $\hat{e}_{mm}$ for DWA. The modulation noise and mismatch noise estimates $\hat{n}_q$ and $\hat{n}_{mm}$ are obtained by filtering the unshaped estimates by the corresponding shaping transfer function.

One of the reasons the proposed model has not been presented before may be the difficulty in observing the spurious tones (as was discussed in Ch. 2): FFT analysis tends to hide tones that wander over a band. Observing spectrograms with slowly changing inputs lead to an accurate model that is also applicable to stimuli that are more difficult to observe.

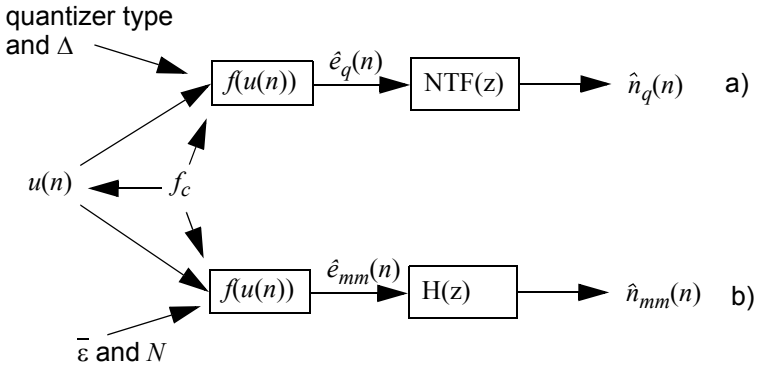



**Fig. 22. The noise contribution modeling approach as a simple flow graph: a) DSM modulation noise and b) DWA-DAC mismatch noise.**

As concluded in Papers III and IV, the basis of both tone-behavior models is frequency modulation (FM). In FM, the modulating baseband data signal is denoted $x_m(t)$ and a sinusoidal carrier is defined as

$$x_c(t) = A_c \cos(2\pi f_c t), \tag{63}$$

where $f_c$ is the base frequency of the carrier and $A_c$ is the amplitude. The modulator combines the carrier with the baseband data signal to produce the transmitted signal:

$$y(t) = A_c \cos\left(2\pi \cdot \left(f_c t + f_\Delta \int_0^t x_m(\tau) d\tau\right)\right), \tag{64}$$

where the sum $f_c$+$f_\Delta$ is the instantaneous frequency of the tone.

The tone mechanism was found by studying slowly changing input signals $u$ for low-pass configured DSM and $1\text{-}z^{-1}$ DWA. For a slow ramp input $u$, it was found that the instantaneous input level $u_{dc}$ results in spurs at instantaneous frequencies of

$$f_\Delta = \frac{i \cdot u_{dc}}{X} f_s, i = \{1, 2, 3, \ldots\}, \tag{65}$$

where $X$ is the quantization interval $\Delta$ (DSM) or the number of quantization steps $N$ (DWA). This same group of tone trajectories can be seen in both the DSM (X=$\Delta$) output spectrogram and the DWA-DAC mismatch noise spectrogram ($X$=$N$). The frequencies in (65) always alias back to the band from dc to $f_S/2$, making the slope either $\pm\, i/\Delta$ or $\pm\, i/N$.

The tone estimator, presented as a discrete-time equation (normalized sample time) is given by (66). The integral required in FM (64) is represented by convolution (operation $\otimes$) with function $h$.

$$\hat{e}_{qm}(n) \;=\; \beta_i \cdot \cos\left(2\pi \cdot \left(\frac{i}{X} \cdot (u(n) \otimes h(n))\right) + \phi\right), \tag{66}$$

where

- $\hat{e}_{qm}(n)$ is the non-shaped estimate for quantization or mismatch noise
- $\beta_i$ is an attenuation coefficient
- $i = \{1, 2, 3, ..\}$
- $X$ is the quantization interval $\Delta$ (DSM) or the number of quantization steps $N$ (DWA)
- $h$ is the impulse response of an integrator (low-pass model) or resonator (band-pass mode)
- $u$ is the input signal
- $\phi$ is an experimental phase term, which can help to find an accurate in-band noise estimate for a DSM with single-tone stimulus.

Equation (66) is completely observation-based and will be refined for both DSM and DWA. The applicable DSM center frequencies $f_c$ for the estimator are either dc or $f_S/4$. For a band-pass DSM or DWA, $f_S/4$ is a typical choice for the center frequency. The agreement between basic FM equation and the tone-prediction model are listed in Table 4.

**Table 3. The equivalent parameters in FM and the proposed noise model.**

| FM (64) | Noise model (66) |
|---|---|
| $A_c$ - carrier amplitude | $\beta_i$ - attenuation factor |
| $f_\Delta$<br>instantaneous frequency<br>(minus the constant carrier freq.) | $f_s \cdot u \cdot i/X$,<br>where $i = \{1,2,3,...N_S\}$,<br>X is: $\Delta$ (quantizer step) for DSM<br>or $N$ for DWA |
| $x_m$ - baseband data signal | $u$ - DSM input signal |
| Integral | discrete convolution |
| $\int_0^t x_m(\tau)d\tau$ | $u(n) \otimes h(n)$, |

Looking at a spectrogram as a function of a linearly increasing ramp level reveals several trajectories that follow the slope of (65) and are all aliased to the band from dc to $f_s/2$. This can be observed in Fig. 20 and Fig. 21. The modulation or mismatch noise is, in fact, the sum of several tones at different values of $i$. This culminates in the uniform model to a sum of spurious tones:

$$\hat{e}_{qm}(n) = \sum_{i=1}^{N_s} \beta_i \cdot \cos\left(2\pi \cdot \left(\frac{i}{X} \cdot (u(n) \otimes h(n))\right) + \phi\right) \quad , \tag{67}$$

where $N_s$ is the number of estimated spurs.

Note that the estimation result is not spectrally shaped. To compare the magnitudes of the simulated and predicted spurious tones, the predicted noise estimate (67) has to be filtered by the corresponding high-pass or band-stop noise transfer function. This function is either the NTF (for DSM) or H (for DWA mismatch shaping transfer function).

### *6.4.1 Estimator parameters for DSM*

For DSMs, the spurious tone estimate accommodates an offset parameter *of* and a phase parameter $\phi$:

$$\hat{e}_q(n) = \sum_{i=1}^{N_s} \beta_i \cdot \cos\left(2\pi \cdot \left(\frac{i}{\Delta} \cdot ((u(n) + of) \otimes h(n))\right) + \phi\right) \quad , \tag{68}$$

where *of* is an additional offset that depends on the type of quantizer:

$$of = \begin{cases} 0, \text{ for mid-tread quantizer} \\ \Delta/2, \text{ for mid-rise quantizer} \end{cases} . \tag{69}$$

A practical value of spurious tone magnitude $\beta_i$ has been found for first order DSM by studying the in-band noise for dc-stimuli with levels between ±0.5Δ. By comparing the simulated pattern noise graphs [13] with estimation, the matching value for $\beta_i$ was found:

$$\beta_i = \frac{4}{3\pi \cdot i^L} , \tag{70}$$

where $L$ is the order of the DSM. Obviously, the number of tones ($N_s$) needed to be modeled depends on the order $L$ in (68). For $L = 1$, the attenuation is the smallest, so $N_s$ needs to be large.

Basically, a proper number for $N_s$ would be the number of samples in a simulation, but this is not necessary due to the steep attenuation for increasing $i$. In simulations it was found that $N_s$ of 100 is sufficient for $L = 1$. This is the point where increasing $N_s$ e.g. 10-fold does not increase the estimated in-band noise contribution with dc or single-tone stimulus. For a large value of $L$, a sufficient value for i was determined experimentally by fixing the minimal difference as follows:

$$\beta_{i-1} - \beta_i \leq 4.23 \cdot 10^{-5} . \tag{71}$$

The numerical maximum in (71) was obtained for $L = 1$ and $N_s = 100$. The resulting values for L = {2,3,4} are $N_s$ = {28,13,7}. Finding a proper value $\beta_i$ for $L > 1$ is somewhat paradoxical, because $\beta_i$ models the magnitude of the $i^{\text{th}}$ **spurious tone**. The whiteness of the DSM modulation noise power spectral density depends on the stimulus, the NTF, and the number of quantizer steps. For $L > 1$, (71) may not always be a good prediction of the quantizer noise. This is further discussed in Subsection 6.5.3.

The model is compatible with bandpass DSM with center frequency of $f_s/4$. The main difference between (68) and the model presented in Paper IV is the presence of center frequency $f_c$ in the latter. Nevertheless, the models are the same; in this compendium $f_c$ is included in the phase term $\phi$. Section 6.5 contains case-studies on band-pass DSM.

### 6.4.2 Estimator parameters for DWA

As noted in [73], the mismatch noise from DWA-DAC is similar to DSM related modulation noise with first-order shaping. Indeed, the estimate filtered with the first-order shaping function results in a very accurate estimate of the true mismatch noise.

For the general function of the tone estimator in (67) only the attenuation parameter and the number of tones have to be specified for DWA. The spurious tone magnitude $\beta_i$ was found in a similar manner as for the DSM: namely in-band mismatch contribution from a dc-sweep. The magnitudes are:

$$\beta_i = \frac{c_i}{i \cdot \pi} , \tag{72}$$

where the coefficient $c_i$ refers to expressing the DNL error by its Fourier coefficients: the magnitudes for the different periodicity of the DNL error vector are defined by $c_i$ (see Paper III). Here, the indexing in coefficient $c_i$ is zero-based: the first value $c_0$ relates to the expected value of $\bar{\varepsilon}$, which is disregarded in the model (expected to be zero).

The model for DWA spurious tones is:

$$\hat{e}_m(n) = \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \beta_i \cdot \cos\left(2\pi \cdot \left(\frac{i}{N} \cdot (u(n) \otimes h(n))\right) + \phi\right) . \tag{73}$$

The model assumes that the DSM output data is in unsigned integer format. If an input signal $u$ is normalized to ±1 full scale, it will be scaled as

$$u \rightarrow N/2 \cdot (u + 1). \tag{74}$$

As can be seen in (73), the number of tones created by the DWA-DAC is limited by the number of DAC elements $N$, namely:

$$N_s = \left\lfloor \frac{N}{2} \right\rfloor . \tag{75}$$

The number of tones in (75) were deduced by studying the periodicity of the DNL error vector. In Paper III, it was found that the periodicity of the DNL error vector has a straightforward impact on the spurious tone frequency allocation. This is most easily studied by assigning a sinusoidal DNL error vector shape with $i$ periods. Sinusoidal shape results to Fourier coefficients $c_i$ with only one non-zero coefficient with the values $i$ limited by (75). In a DWA-DAC, a sinusoidal DNL error shape results in only one spur trajectory or envelope, depending on $i$.

For the sinusoidal DNL error vector shape, there are exactly $\lfloor N/2 \rfloor$ (67) alternatives for DNL error periodicity. For a sinusoidal DNL shape with $i$ periods, it takes $N_i$ samples for the errors to average out, where

$$N_i = \left\lfloor \frac{N}{i} \right\rfloor , i = \left\{ 1, 2, \ldots, \left\lfloor \frac{N}{2} \right\rfloor \right\}. \tag{76}$$

A realistic DNL shape results in all non-zero coefficients $c_i$ and $N_i=N$. In Paper III, it was concluded that for dc inputs and single-tones with low amplitudes, the spurious tones will be generated at or near frequencies $0.5 f_s \cdot i$ (for the $1-z^{-1}$ DWA). These tones will alias between dc and $f_s/2$, resulting in DSM-like tone behavior.

In [12] it was pointed out that as the word size (or the number of unit-DACs) increases, it takes longer for the errors to average to zero and the DWA algorithm cannot move the distortion in higher frequencies. Here, it was pointed that the number of DWA-DAC related tones is half of $N$ (75). For increasing $N$, the congestion of tones will inevitably contribute to the performance. This is also in-line with the conclusions by Chen and Leung [57] presented in Subsection 6.3.2: for DWA, $N$ cannot be high for low OSR.

Again, there is a difference in how the effect of center frequency is presented in (73) and the original Paper III: in this compendium $f_c$ is included in the phase term $\phi$. Section 6.5 contains a bandpass DWA-DAC example with shaping function of $1+z^{-2}$.

## 6.5 Case studies on tone estimation

### 6.5.1 In-band tone minimization for DWA

The periodicity of the DNL error vector shape determines the Fourier coefficient $c_i$ in (67). By this observation, it was deduced in Paper III that by finding a proper DNL error shape, the (DWA-related) in-band spurious tones can be minimized. Permuting a detrimental DNL-vector into a different shape can result in nearly spurious-free in-band noise. The key point in the permuting approach is to minimize $c_i$ for even $i$, as these coefficients are tied to tones around the center frequency $f_c$ (dc in low-pass DSM)

Paper III did not propose any algorithm or a method for the permutation, only a brute force random permutation was used instead. The permutation can be done either initially or repeatedly to cope with varying DNL error. This may result in complicated hardware, but the permutation approach has two important merits:

– It is based on studying the accurate time-domain model DWA.

– As groundwork, the permutation approach is a DWA tone cancelation method that cancels the tones without increasing the in-band noise power.

The prediction model accommodates incremental DWA (IDWA) [74] as well. As mentioned, using extra unit-DACs will shift the frequency around which the tones fluctuate. This frequency shift by (62) is observable by similar spectrogram plots, as in Fig. 20. This was presented in Paper III.

IDWA is clearly an alternative to the permutation approach but should be carefully studied for a given $N$ and its increment. This is because increasing the number of unit DACs increases the number of spurious tones. IDWA with the permutation method has not been tested by the author, but it may be worthwhile to compare it with regular DWA with the permutation.

### 6.5.2 Pattern noise

A stable limit cycle is a result of a repeating pattern in a DSM feedback loop with a dc-input. The first order DSM-like behavior of a DWA results from the fact that the mismatch noise also has limit cycle. The pattern noise graph is a result of sweeping dc-levels and measuring the in-band cumulative (or rms [13]) noise or maximal spur-levels. The tone estimation model (67) is very accurate in estimating

the pattern noise magnitudes. The phase modulation term ($\phi$) is not needed for dc-inputs because this would only add a constant phase in the estimation.

### *Estimating DSM Pattern Noise*

In Fig. 23a we have a case of a first order low-pass DSM with an OSR of 32. The quantizer type is mid-rise and two quantizer levels are used. The measure is taken from the cumulative in-band quantization error power (dB) with the stimulus full scale levels normalized to $\pm\Delta/2$ (normalized here to $\pm1$).

The corresponding estimate is shown in Fig. 23b. With 100 estimated spurs, the spikes in the estimated in-band cumulative noise are less than one decibel from the simulated value.

In spectrograms, parameter $i$ in (68) referred to a certain tone ($i = 1$ pointing to the strongest spur). As the number of spurs is controllable in the prediction, it is possible to connect the index number $i$ in (68) to a specific detrimental input dc-level in Fig. 23b. In a classic publication, Candy and Benjamin [13] pointed out that the worst-case dc-inputs for a first order DSM are the simplest rational numbers (normalized to the full scale). With the prediction model, the connection between rational dc-inputs and spurious tone index $i = \{1 \text{ to } 8\}$ has been summarized in Table 4. Here the full scale input and quantizer levels are both at $\pm1$.

**Table 4. The connection between pattern noise and the index *i* in (68).**

| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| a peak near rational dc-level of: | ±1 | 0 | ±1/3 | ±2/4 | ±1/5, ±3/5 | ±2/6, ±4/6 | ±1/7, ±3/7, ±5/7 | ±2/8, ±4/8, ±6/8 |

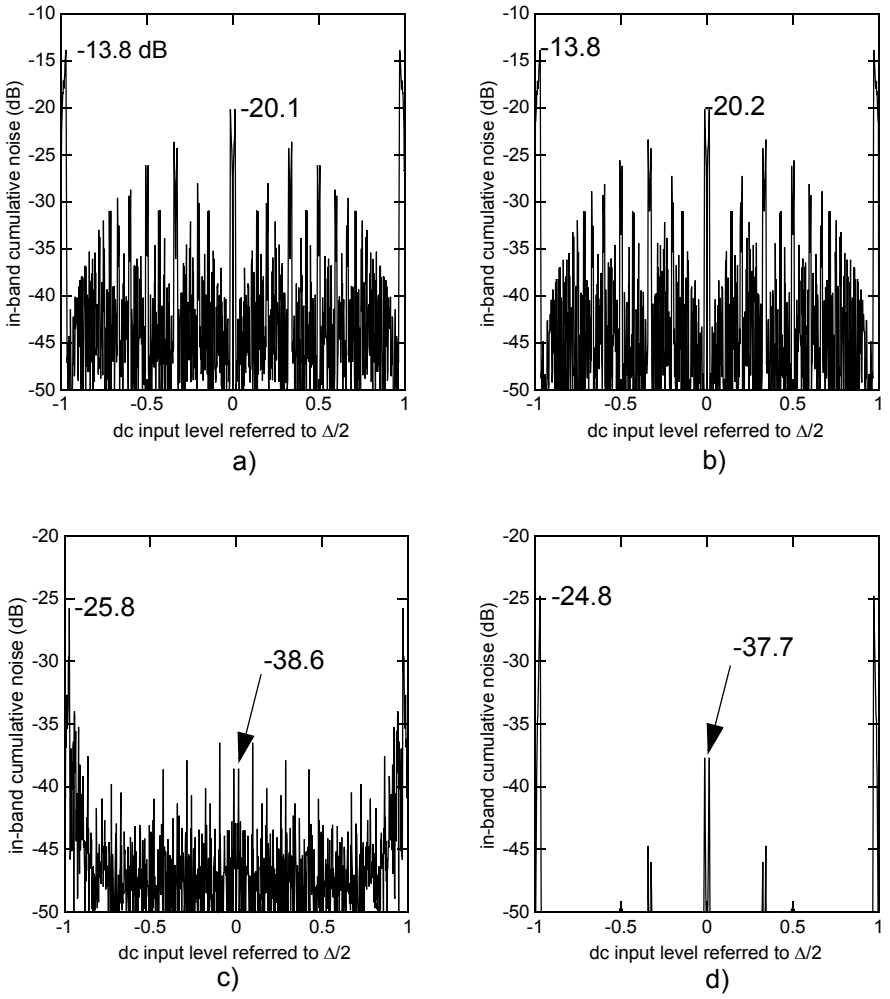



**Fig. 23. Pattern noise graph for 1-bit low-pass DSM: a) *L*=1, simulation and b) *L*=1, estimate ($N_s$ = 100), c) *L*=2, simulation and d) *L*=2 estimate ($N_s$ = 28). Revised from [IV]. ©[2011] IEEE.**

From Table 4 it can be deduced that given an arbitrary $i$, the corresponding spurious tone will contribute to increased in-band noise dc-inputs in the vicinity of:

$$\frac{\pm j}{i} \text{ , where } j = \begin{cases} \{(i-2),(i-4),...,1\} \text{ for odd i and} \\ \{(i-2),(i-4),...,0\} \text{ for even i} \end{cases} \tag{77}$$

Equation (77) assumes that the input $u$ is normalized between ±1. If there are more than two quantizer levels and $N$>0, the pattern noise graphs (true and estimated) follow the same shape as for two quantizer levels.

As mentioned, for higher order DSMs, the more whitened nature of the quantizer noise sets an obstacle to noise estimation. Fig. 23c shows the pattern noise graph for a second order DSM. The estimate in Fig. 23d finds the major peaks near ±1 and zero, but not the numerous other spikes simulated in Fig. 23c. This problem will be revisited and discussed in Subsection 6.5.3.

### *Estimating DWA-Related Pattern Noise*

Using dc-stimuli, it is possible to recreate the pattern noise graph for the mismatch noise as well. Here, it will be shown that for a high number of unit-DACs, the pattern noise graph will resemble the ones from first order 1-bit DSM.

In Fig. 24a, we have a mismatch pattern noise graph for a system containing a second order low-pass-DSM with 7 unit-DACs. Here, the DNL standard deviation was 0.01 and the DNL shape was set to a first order polynomial. The estimation in Fig. 24b is quite accurate, but only for the three largest pairs of spikes. The number of spurious tones modeled in (73) is also three.

The less-dominant portion of the mismatch noise not modeled by (73) is affiliated with the DSM. For instance, if the multibit DSM has its NTF designed for 1-bit DSM, the excess noise may also have more observable tones (or dc-related spikes) than stated in (75). This is a marginal scenario, as the NTF out-of-band gain is usually customized to the number of DSM quantizer levels.

The equivalence between DWA-DAC and a first-order 1-bit DSM becomes more apparent, if the value of $N$ is increased. In Fig. 24c, $N$ is 255 (8-bit DSM output) and the estimated pattern noise graph in Fig. 24d is almost identical: the estimated spikes are less than 1 dB from the simulated values.

The connection with the one-bit DSM pattern noise graph is apparent. Given an arbitrary $i$, the corresponding spurious tone will contribute to increased in-band noise with dc-inputs in the vicinity of:

$$N\left(\frac{i \pm j}{i}\right) \quad , j = \begin{cases} \{(i-2),(i-4),...,1\} \text{ for odd i and} \\ \{(i-2),(i-4),...,0\} \text{ for even i} \end{cases} . \tag{78}$$

Equation (78) assumes that the input is normalized between 0 and $N$; the same range as the quantizer.

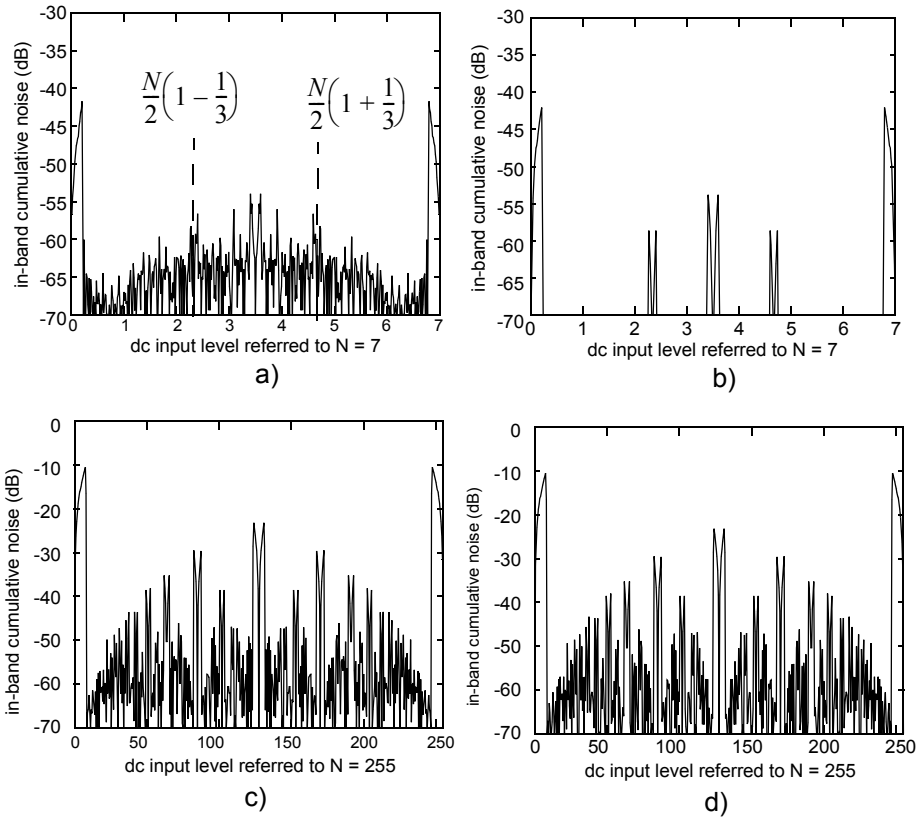



**Fig. 24. A mismatch pattern noise graph: a) 3-bit DWA, simulation and b) 3-bit DWA, estimation, c) 8-bit DWA, simulation and d) 8-bit DWA, estimation.**

### *Band-pass DSM*

For a second order band-pass DSM with an N-path transformation ($z \rightarrow -z^2$), the pattern graph estimation is just as accurate as it is for first order DSM and DWA. The difference is that for a band-pass DSM, the corresponding input is a single-tone exactly at the center frequency of $f_s/4$ (47). With bandpass DSM and this stimulus, a constant phase modulation term $\phi$ in (68) and (73) of the value $2\pi/4$ was needed to properly match the simulated results.

The pattern noise graphs for band-pass DSM can be seen in Fig. 25. In Fig. 25a, we have the cumulative in-band DSM modulation noise as a function of amplitude along with the estimate by (68) in Fig. 25b. In Fig. 25c the graph is for cumulative in-band mismatch noise for a 4th order DSM attached to a DWA-DAC. As in the previous example, the DNL noise shape was set on a first-order polynomial. The corresponding estimate by (73) can be seen in Fig. 25d. As the shaping order is one, $L$ (72) in was set to one as well.

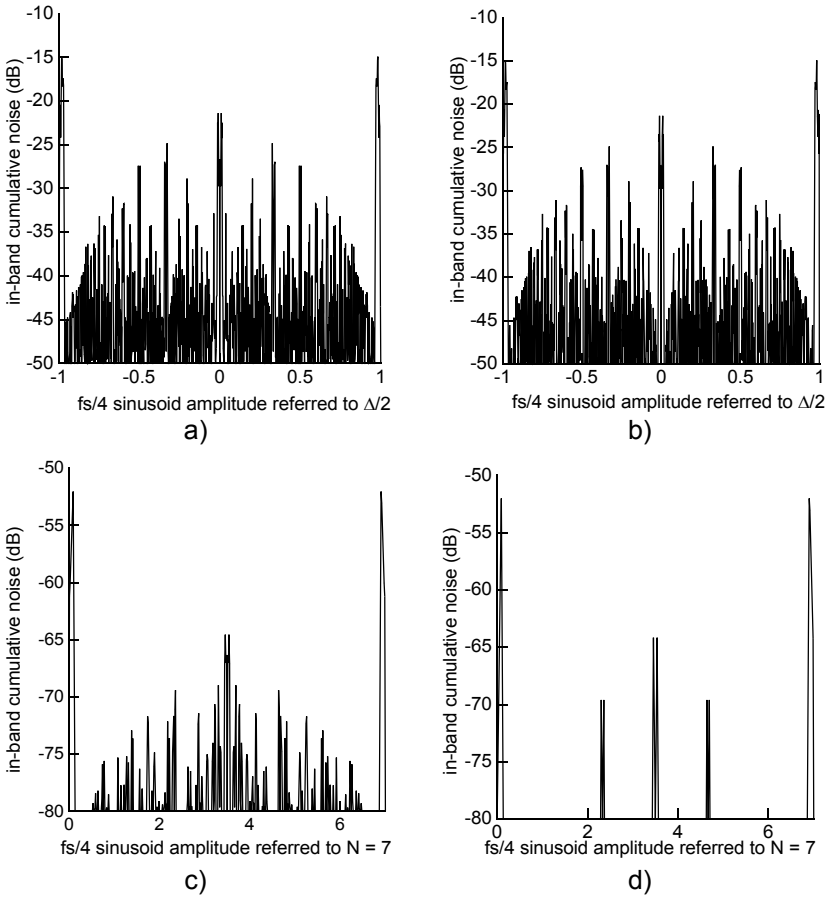



**Fig. 25. Pattern noise graphs for band-pass DSM: a) modulation noise for 1-bit DSM, simulated, b) modulation noise for 1-bit DSM, estimated, c) mismatch noise for 3-bit DWA, simulated and d) mismatch noise for 3-bit DWA, estimated.**

### *6.5.3 Performance estimates for DSM and DWA*

To estimate SNDR/SFDR versus single-tone amplitude, the quantizer noise estimate in (68) will be filtered by the corresponding NTF and summed with the stimulus. Initially, for a first order DSM, it was found that the estimate was not quite spot-on. Increasing the number of estimated spurs from 100 to 1000 did not change the estimation results. The experimental phase term $\phi$ in (68) was found to

be the key solution. Instinctively, this term has to be attached to both the stimulus and the DSM. The quantizer gain is a parameter that fits this description.

For this purpose, the quantizer gain estimate is generated with one simulation by exciting a DSM with a slowly changing input signal. Being a statistical quantity, the quantizer gain is generated e.g. once every 2048 samples. After covering the whole input range, the quantizer gain $k$ as a function of instantaneous input level (Fig. 26) is then fitted to a 6th order polynomial $k_p$. The order 6 was chosen as a compromise between polynomial simplicity and small residual error.

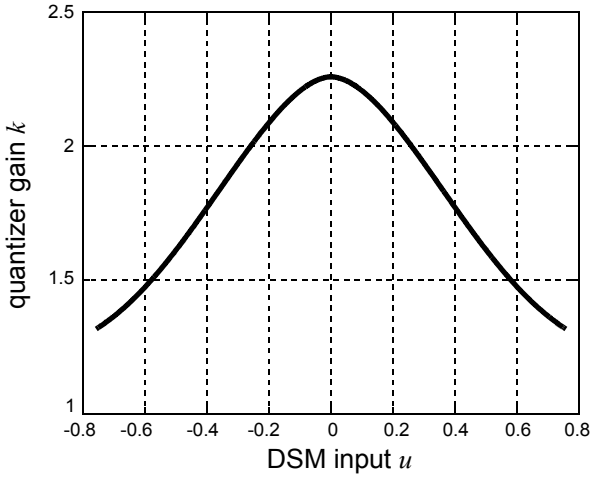



**Fig. 26. The quantizer gain as a function of instantaneous DSM input *u*.**

The polynomial fit $k_p$ was then used to obtain a phase modulation term $\phi$ in (68) as:

$$\phi(n) = (1\text{-}k_p(n))\cdot 2\pi. \tag{79}$$

For instance, $k_p$ varying from 1.3 to 2.3 causes a phase shift of $2\pi$. Using the estimate (79) in (68), the tone prediction model is very precise for a first order DSM. An example is shown in Fig. 27, where we have two SNDR and two SFDR graphs along with their estimates. The graphs in Fig. 27 are for single-tone inputs with the same amplitudes and frequencies but 90° phase difference. These stimuli are here named as "input 1" and "input 2". Here, the 1-bit DSM has an OSR of 32 and the frequency is ca $f_b/16$, where $f_b$ is the band of interest.

The subtle difference in stimuli causes clearly observable difference in-band cumulative noise and maximal spurious level. This difference is remarkably well-predicted by the time domain estimator (68) with phase modulation (79). The single-tone frequency is defined by (56) with parameters $J$ and $M$ as 127 and $2^{17}$, respectively.

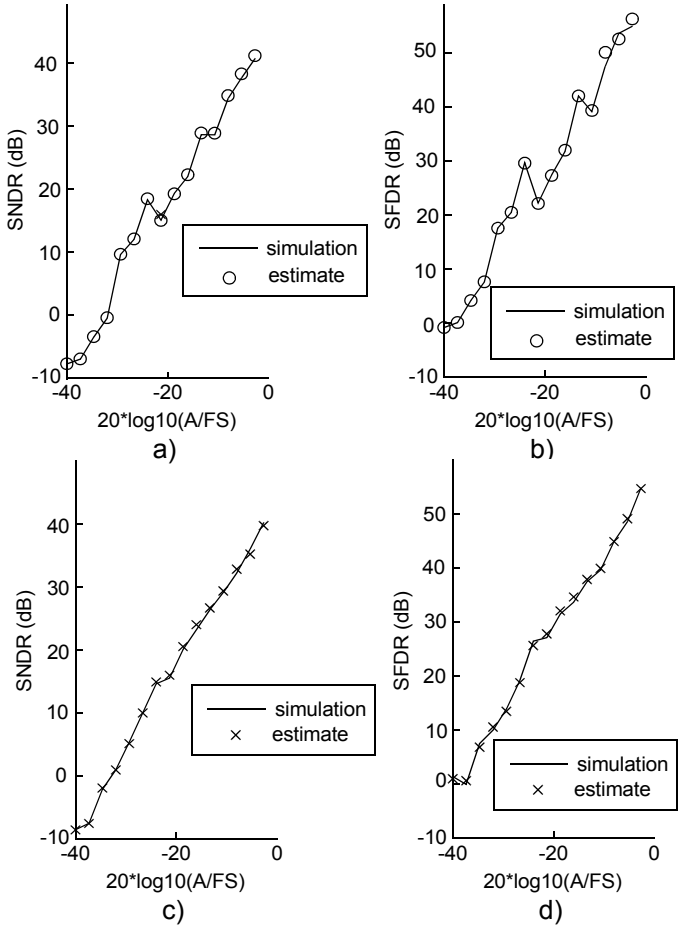



**Fig. 27. Performance graphs for 1st order, 1-bit DSM: a) SNDR for input 1, b) SFDR for input 1, c) SNDR for input 2 and d) SFDR for input 2. Revised from [IV]. ©[2011] IEEE.**

With the aforementioned parameters, the graphs for "input 1" are clearly non-monotonic. This kind of non-monotonic curve was also presented in the paper by Candy & Benjamin [13]. For single-tone inputs they concluded the following:

*Complete graphs of the resolution would be tedious to calculate; they are better obtained by simulation.*

The proposed time-domain model enables an accurate computer-aided numerical solution to the noise and spur contents in the band-of interest. Using this model, an analytical solution would be very complicated, especially with the phase term $\phi$. In the end, the solution would be totally stimulus-dependent.

The prediction results in Fig. 27 are promising. However, the phase modulation term in (79) is an adequate add-on only for first order DSMs (referring to the shaping order). To expand the estimation accuracy to higher order DSMs, it is required that the quantizer noise estimate has to be somehow scrambled according to the order of the DSM. Some scrambling ideas were tested, e.g. by iteratively generating more spurious tones by assigning the previous spur estimate as an input to (68). This method tends to mimic the way the quantizer noise is mixed with the stimulus in the DSM feedback loop. So far, an accurate method has not been found. Perhaps the phase modulation term modified for higher order DSMs is the solution.

For a second order band-pass DSM with N-path transformation ($z \rightarrow -z^2$), the tone estimation should be accurate just as it was for a first order DSM. The problem here is to determine the quantizer gain as the input signal is near $f_s/4$. A reasonably well in-band spurious tone prediction results were obtained by using the value of $\phi$ for low-pass case (with the quantizer gain model) and modifying the values by following transformation:

$$\phi(n) \rightarrow 2\phi(n) + \frac{2\pi}{4}. \tag{80}$$

The results are shown in Fig. 28. As a reference, a linear fit is also included in Fig. 28. Here, the single-tone frequency is defined by (56): $\{J,M\} = \{127, 2^{17}\}$. Again, $L$ in the attenuation factor (70) was set to one.

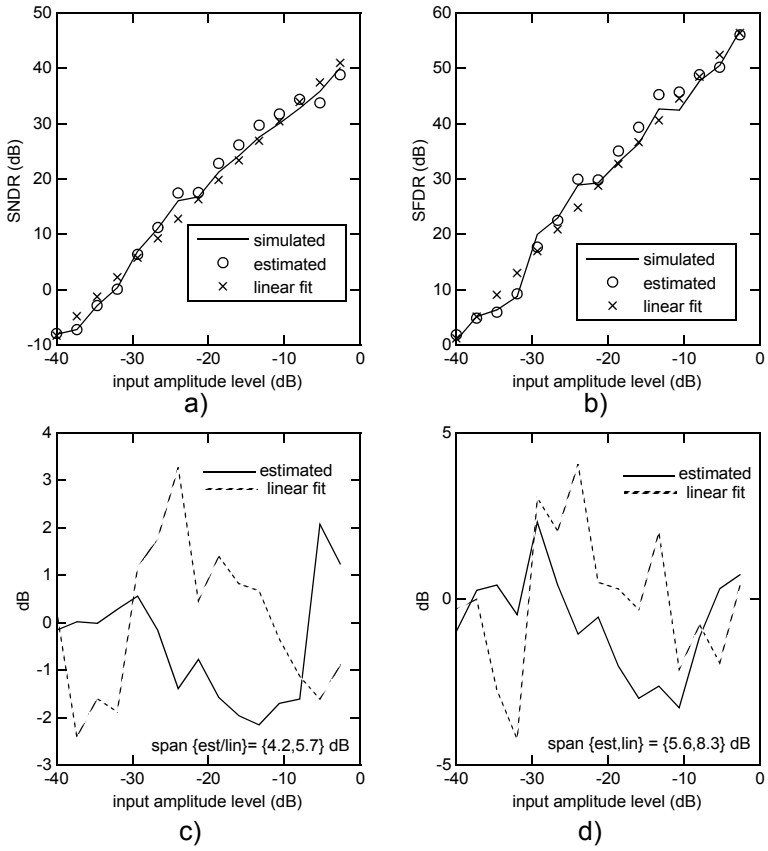



**Fig. 28. Simulated, estimated and fitted performance graphs for 1-bit 2nd order band-pass DSM: a) SNDR b) SFDR, and the residual errors c) SNDR estimate residual and d) SFDR estimate residual. Revised from [IV]. ©[2011] IEEE.**

For DWA, the performance prediction with a single-tone stimulus also works quite well. Here, we have a second order low-pass DSM connected to a DWA with 7 unit DACs. The DNL shape is linear (first order polynomial) with σ of 1%. Again, the single-tone frequency is defined by (56): $\{J,M\}= \{127, 2^{17}\}$.

The similarity with a first order 1-bit DSM suggests that the phase modulation term for the DSM is required. As for the DSM, the quantizer gain polynomial (needed for the phase modulation term ϕ) was used for DWA. The quantizer gain is not an obtainable parameter for the DWA system, so a first order 1-bit DSM quantizer gain polynomial was used.

In Fig. 29, only the mismatch noise contribution was taken into account,. The term $SFDR_{DWA}$ in Fig. 29b is the signal to maximal in-band DAC-mismatch spurious tone power ratio. As can be seen in Fig. 29b, the model is very accurate in predicting the maximal in-band spurious tone.

Here, the time domain model with $N = 7$ only accommodates three dominant spurious tones (75) and thus omits the DSM related excess mismatch noise. This was also apparent in dc-tests (Fig. 24a). Therefore, the signal-to-mismatch-noise ratio (SMNR) estimate is slightly too optimistic, as shown in Fig. 29a (especially at larger amplitudes).

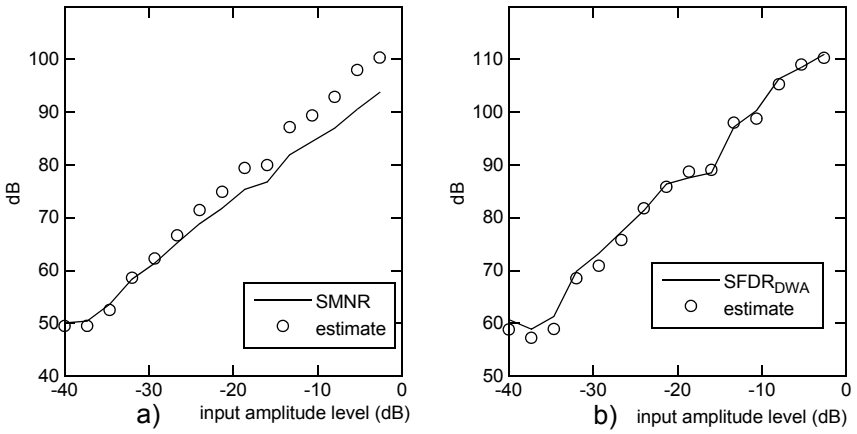



**Fig. 29. Mismatch noise performance merits: a) SMNR and b) $SFDR_{DWA}$.**

The proper scaling for the input to the polynomial is as follows. If an input signal $u$ is normalized to ±1 full scale, the scaling of $u$ for the quantizer gain polynomial $k_p$ is:

$$u \rightarrow (u + 1)/4. \tag{81}$$

The reasoning for the scaling in (81) is as follows. The input to the DWA algorithm is in signed integer format: a sinusoid (plus modulation noise) with offset and maximum magnitude equal to $N \cdot 0.5$. The scaled result by (81) is the stimulus, whose maximum span is normalized by $N$: between 0 and 0.5.

As a conclusion to this Section, to estimate the performance of DSM or DWA very accurately requires phase modulation term $\phi$. The solution for $\phi$ (79) has been determined heuristically. This solution, (including the band-pass and DWA adaptations) is quite interesting and requires further analytical rationalization.

## 6.6 Summary

This Chapter introduced a frequency modulation based tone prediction model for both Delta-Sigma modulators and DWA-DAC. In Papers III-IV, the model is used for tone prediction. This is true, but in a wider sense the model attempts to explain the quantization noise or unshaped mismatch noise.

The common model for both systems reveals a profound parallelism between first order 1-bit DSM and *N*-level DWA-DAC. This duality has been reported previously, and it can be discovered e.g. by dc-simulations.

For first order DSMs (and second order band-pass DSMs), the quantizer noise prediction model is very accurate. Excellent accuracy for non-dc stimuli requires the use of a phase modulation term, which is related to the instantaneous quantizer gain of a first order low-pass DSM. This signal-dependent modulation term is extendable to second order band-pass DSM and the generalized DWA.

For DSMs in general, the expected tone frequencies and magnitudes can be predicted for any DSM order by the proposed model. The problem of predicting quantizer noise whiteness, however, still remains. For higher DSM orders, the model needs a signal-dependent parameter that would properly scramble the quantizer noise.

The time-domain model reveals that the DWA-DAC related spurious tones can be manipulated with the shape of the DAC mismatch error. The even-order Fourier coefficients of the DAC's differential nonlinearity error vector are the key factors in estimating the magnitudes of these tones. The errors are usually unknown, but by permuting the DAC-elements a favorable error shape can be found. The permutation approach would be a DWA tone cancelation variant that does not have the trade-off between the in-band spurious tones and the noise floor. The realization is discussed in Sect. 7.3.

Both time-domain models are published at the Matlab Central File Excange[3], under the title "Delta Sigma converter spurious tone predictor".

[3] http://www.mathworks.com/matlabcentral/fileexchange/?term=authorid:17648

# 7 Conclusions

## 7.1 Summary

The complexity of spectral shaping DEM logic increases quite rapidly as the order of shaping is increased: parallel computing may be needed. This is why first-order shaping is the most common solution in practical applications. Realizations of DEM methods are prone to spurious tones and therefore dither-like solutions are quite popular amongst them. The most popular DEM method, DWA, is one of the main subjects of this thesis.

In Chapter 4 and Paper I, the traditional data weighted averaging technique was generalized to include the shaping function $1 \pm z^{-D}$. This is justified by the fact that such a generic model is realizable and maintains the simple barrel-shifter based operation for the DWA algorithm. This generalization integrates some former publications on band-pass DWA variations. In the author's view, the work on generalizing DWA is helpful in understanding the variations in DWA spectral shaping and hopefully in creating new tone-suppression methods.

The qualification approach presented in Chapter 5 and Paper II proposed a simulation-based methodology for evaluating DAC mismatch shaping logic. The advantage of the approach is that different methods are more easily comparable with repeatable and extensive simulation setups. An important point in the qualification approach is the discussion on the number and types of parameter sweeps that are required to obtain a reliable result.

The spurious-tone models from both DSM and DWA were proposed in Chapter 6 and Papers III-IV. The time-domain models are similar in their behavior and they also predict the spurious tones quite accurately with any stimulus. For DWA, the tones are correlated with the input signal just as for DSM, but the shape of the mismatch (DNL error) is also significant. So far, the noise models are very accurate for first-order noise shaping (DSM or DWA) functions $1-z^{-1}$ and $1+z^{-2}$.

## 7.2 Discussion

The generalization of the DWA algorithm and the qualification method, along with the time-domain mismatch noise model, provide a convenient set of design tools. These enable studying existing DWA variants in a broader perspective, also for both low-pass and band-pass DSMs.

For qualifying and comparing DAC mismatch shaping methods in necessary scope, the thesis proposed a reasonably rapid and easily repeatable behavioral simulation scheme. The point of this work was to clarify (without pointing fingers to any publication) that to qualify a DEM method is not possible with one fixed amplitude, frequency or mismatch shape. The challenge addressed in this work was in the random shape of the static mismatch. Other DAC non-idealities, such as mismatch based timing errors [78] have been omitted.

The time-domain noise contribution model enables predicting the tonality with an arbitrary stimulus. A good example of utilizing the model is the proposed idea of DWA tone cancelation by unit-DAC DNL-shape permutations.

The thesis also omitted a wide range of behavioral challenges that are apparent in DS ADCs, where the imprecise parameters for the analog part should be taken into account. This would require a DSM model with the presence of at least thermal noise, OpAmp nonlinearities and clock jitter. An example of an impressive simulator model was provided for SIMULINK by Malcovati, who is also the co-author of the related publication [79]. The author's contribution to this field was a lookup-table based integrator settling error model that can be modeled e.g. in SIMULINK [80].

## 7.3 Future work

Tone prediction from the time-domain model needs to be refined to cover different types of DSM. The following age-old question need to be reflected upon with the model: How to predict the true nature of the quantization noise? Whether the quantization noise has a white power spectral density or not depends mainly on

– the stimulus
– the noise shaping properties of the DSM.

Although this work did not find a sufficiently accurate quantizer noise model for higher order DSMs, the quantizer and mismatch noise models are capable of predicting the spurious tone frequencies with basically any stimulus. An important finding was utilizing the phase modulation term $\phi$ in the prediction. This term is based on the estimation of instantaneous quantizer gain. So far, $\phi$ requires analytical rationalization on why and how it works. The phase term $\phi$ could be key to developing an accurate time-domain quantizer noise model for higher order

DSMs (or DEM methods). This would be the long-sought model that is able to describe DSM quantizer noise behavior.

The generic DWA was realized on a Xilinx FPGA board in 2009 by Janne A. Raappana and tested with analytical DNL shapes (not a real DAC). For the generalized form DWA-DAC algorithm, it was found that the shape of the mismatch DNL error is essential regardless of the low-pass or band-pass configuration. With the accurate tone predictor, the tone minimization based on the permutation method was discovered found quite easily. To realize such a permutation requires a study of whether it should be done initially or repeatedly to cope with varying DNL error. The permutation itself can be realized by a simple butterfly network.

The DEM characterization (referred to "qualification" in Paper II) is a contribution that proposes a unified approach to benchmarking the performance of DAC mismatch shaping methods. A convenient evolution would be to find an optimal small set of DNL error shapes (or even one single shape).

# References


1. Cutler CC (1960) Transmission System Employing Quantization. U.S. Patent No. 2 927 962.
2. Spang H III & Schultheiss P (1962) Reduction of quantizing noise by use of feedback. IEEE Transactions on Communications 10(4): 373–380.
3. Inose H & Yasuda Y (1963) A unity bit coding method by negative feedback. Proc IEEE 51(11): 1524–1535.
4. Gray RM (1990) Quantization Noise Spectra. IEEE Transactions on Information Theory 36(6): 1220–1244.
5. Bennett WR (1948) Spectra of quantized signals. Bell System Technical Journal 27: 446–471.
6. Ardalan SH & Paulos JJ (1987) Analysis of nonlinear behavior in delta-sigma modulators'. IEEE Transactions on Circuits Systems 34: 593–603.
7. Borkowski MJ (2008) Digital Δ–Σ modulation: variable modulus and tonal behaviour in a fixed-point digital environment. Ph.D. Thesis, Oulu University Press.
8. Maloberti F (2007) Data converters. Springer, the Netherlands.
9. Larson LE, Cataltepe T & Temes GC (1988) Multi-bit oversampled S-D A/D converter with digital error correction. Electronics Letters 24: 1051–1052.
10. Van De Plassche RJ (1976) Dynamic element matching for high-accuracy monolithic D/A converters. IEEE Journal of Solid-State Circuits 11(6): 795–800.
11. Jackson HS (1993) Circuit and Method for Cancelling Nonlinearity Error Associated with Component Value mismatches in a Data Converter. U.S. Patent 5 221 926.
12. Baird RT & Fiez TS (1995) Linearity enhancement of multibit Sigma Delta A/D and D/A converters using data weighted averaging. IEEE Journal of Circuits and Systems II 42: 753–762.
13. Candy JC & Benjamin OJ (1981) The Structure of Quantization Noise from Sigma-Delta Modulation. IEEE Transactions on Communications 29(9): 1316–1323.
14. Norsworthy SR, Schreier R & Temes GC (1997) Delta-Sigma Data Converters. IEEE Press, New York.
15. Schreier R & Temes GC (2005) Understanding Delta-Sigma Data Converters. 0471465852. Wiley-IEEE Press, First edition.
16. Schreier R (1993) An Empirical Study of High-Order Single-Bit Delta-Sigma Modulators. IEEE Transactions on Circuits and Systems-II: analog and digital signal processing 40(8): 461–466.
17. Dunn C & Sandler M (1995) Efficient linearisation of sigma-delta modulators using single-bit dither. Electronics Letters 31: 941–942.

18. Pamarti S, Welz, J & Galton I (2007) Statistics of the Quantization Noise in 1-Bit Dithered Single-Quantizer Digital Delta–Sigma Modulators. IEEE Transactions on Circuits and Systems I 54(3): 492–503.
19. Fitzgibbon B, O'Neill K, Grannell A, Horgan C, Zhipeng Y, Hosseini K & Kennedy MP (2009) A spur-free MASH digital delta-sigma modulator with higher order shaped dither. Proc ECCTD 2009, European Conference on Circuit Theory and Design: 723–726.
20. Schreier R (1994) On the use of chaos to reduce idle-channel tones in delta-sigma modulators. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 41: 539–547.
21. Farrell R & Feely O (1997) Application of non-linear analysis techniques to sigma-delta modulators. Proc Signals Systems and Chaos (Ref. No. 1997/393), IEE Colloquium on: 4/1–4/6.
22. Chan T-Y & Bibyk B (1999) Exact analysis of second order bandpass delta-sigma modulator with sinusoidal inputs. Proc IEEE International Symposium on Circuits and Systems ISCAS 2: 372–375.
23. Ritoniemi T, Karema H & Tenhunen H (1990) Design of stable high order 1-bit sigma-delta modulators. Proc IEEE International Symposium on Circuits and Systems ISCAS 4: 3267–3270.
24. Lee WL (1987) A novel higher order interpolative modulator topology for high resolution oversampling A/D converters. Master's Thesis, Massachusetts's Institute of Technology, Cambridge, MA.
25. Hamoui AA & Martin K (2002) Linearity enhancement of multibit Delta-Sigma modulators using pseudo data-weighted averaging. Proc IEEE International Symposium on Circuits and Systems ISCAS 3: 285–288.
26. Steensgaard J (1998) Nonlinearities in SC Delta-Sigma A/D Convertes. Proc IEEE International Conference on Electronics, Circuits and Systems 1: 355–358.
27. Silva J, Moon U, Steensgaard J & Temes GC (2001) Wideband low-distortion delta-sigma ADC topology. Electronics Letters 37(12): 737–738.
28. de la Rosa JM (2011) Sigma-Delta Modulators: Tutorial Overview, Design Guide, and State-of-the-Art Survey. IEEE Transactions on Circuits and Systems I: Regular papers 58(1): 1–21.
29. Løkken I, Vinje A, Hernes B & Sæther T (2010) Review and advances in delta-sigma DAC error estimation based on additive noise modeling. Springer Analog Integrated Circuits and Signal Processing 62: 179–191.

30. Altinok GD, Al-Janabi M & Kale I (2010) Stability Analysis of Band-Pass Sigma-Delta Modulators for Single and Dual Tone Sinusoidal Inputs. IEEE Transactions on Instrumentation & Measurement 60(5): 1546–1554.
31. He N, Kuhlmann N & Buzo A (1990) Double-loop sigma-delta modulation with dc input. IEEE Transactions on Communications 38(4): 487–495.
32. Kenney JG and Carley LR (1993) CLANS: a high-level synthesis tool for high resolution data converters. Proc ICCAD-88, IEEE International Conference on Computer-Aided Design: 496–499.
33. Reefman D, Reiss J & Janssen E (2005) Description of Limit Cycles in Sigma-Delta Modulators. IEEE Transactions on Circuits and Systems I: Regular Papers, 52(6): 1211–1223.
34. Schreier R (1996) Oversampling Delta-Sigma Data Converters. Oregon State University, Course material ECE 627.
35. Hosseini K & Kennedy MP (2007) Calculation of sequence lengths in MASH 1-1-1 digital delta sigma modulators with a constant input. Proc PRIME 2007. Ph.D. Research in Microelectronics and Electronics Conference 2007: 309–312.
36. Bach E (1999) Multibit oversampling D/A converters using dynamic element matching methods. Siemens AG, Semiconductors Group. ESPRIT project: SYSCON, K.U. Leuven, ESAT-MICAS.
37. Nys OJAP & Henderson RK(1997) A 19-bit low-power multibit sigma-delta ADC based on data weighted averaging. IEEE Journal of Solid-State Circuits 32(7): 933–942.
38. Welz J & Galton I (2002) Necessary and sufficient conditions for mismatch shaping in a general class of multibit DACs. IEEE Journal of Circuits and Systems II 49(12): 748 – 759.
39. Silva J, Wang X, Kiss P, Moon U & Temes GC (2002) Digital techniques for improved ΔΣ data conversion. Proc IEEE Custom Integrated Circuits Conference: 183–190.
40. Neitola M & Rahkonen T (2003) Fully Digital DAC Mismatch Error Cancellation Schemes for Multibit Delta-Sigma A/D Converters. Proc Finnish Signal Processing Symposium, FINSIG, Tampere: 249–253.
41. Hastings A (2006) The art of analog layout. Second edition. Pearson Prentice-Hall, New Jersey.
42. Fujimori I, Nogi A & Sugimoto T (2000) A Multibit Delta-Sigma Audio DAC with 120-dB Dynamic Range. IEEE Journal of Solid-State Circuits 35(8): 1066–1073.
43. Adams R & Nguyen KQ (1998) A 113-dB SNR Oversampling DAC with Segmented Noise-Shaped Scrambling. IEEE Journal of Solid State Circuits 33(12): 1871–1878.

44. Lindfors S & Halonen K (2000) Two-step quantization in multibit ΣΔ-modulators. Proc IEEE International Conference on Circuits and Systems, ISCAS 2: 25–28.
45. Groeneveld DWJ, Schouwenaars HJ, Termeer HAH & Bastiaansen CAA (1989) A self-calibration technique for monolithic high-resolution D/A converters. IEEE Journal of Solid-State Circuits 24: 1517–1522.
46. Bugeja AR & Song B-S (2000) A self-trimming 14-b 100-MS/s CMOS DAC. IEEE Journal of Solid-State Circuits 35: 1841–1852.
47. Leslie TC & Singh B (1990) An improved sigma-delta modulator architechture. Proc IEEE International Symposium on Circuits and Systems 1: 372–375.
48. Candy JC & Huynh A (1986) Double integration for digital-to-analog conversion. IEEE Transactions on Communications 26(12): 77–81.
49. Hayashi T, Inabe Y, Uchimura K & Kimura T (1986) A multistage delta-sigma modulator without double integration loop. Proc IEEE International Solid-State Circuits Conference, Digest of Technical Papers, ISSCC 1986, Digest of Technical Papers: 182–183.
50. Neitola M & Rahkonen T (2009) A systematic approach for testing a dac mismatch shaping method. Proc European Conference on Circuit Theory and Design ECCTD: 453–456.
51. Weltz J & Galton I (2001) Simplified logic for first-order and second-order -shaping digital-to-analog converters. IEEE Transactions on Circuits and Systems II, Analog and digital signal processing 48(11): 1047–1027.
52. Leung BH & Sutarja S (1992) Multi-bit sigma-delta A/D converter incorporating a novel class of dynamic element matching techniques. IEEE Transactions on Circuits Systems II 39: 35–51.
53. Galton I (1997) Spectral Shaping of Circuit Errors in Digital-to-Analog Converters. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing 44(10): 808–817.
54. Adams RW & Kwan TW (1995) Data-directed Scrambler for Multi-bit Noise-shaping D/A Converters. U.S. Patent Number 5404142, April 4 1995 (filed Aug. 1993).
55. Schreier R & Zhang B (1995) Noise-shaped multibit D/A convertor employing unit elements. Electronics Letters 31(20): 1712–1713.
56. Andersson NU, Andersson KO, Vesterbacka M & Wikner JJ (2001) Proc NORCHIP Conference, Kista, Sweden: 155–160.
57. Chen F & Leung B (1998) Some Observations on Tone Behavior in Data Weighted Averaging. Proc IEEE International Symposium on Circuits and Systems, ISCAS 1: 500–503.

58. Vleugels K, Rabii S & Wooley BA, (2001) A 2.5-V sigma-delta modulator for broadband communication applications. IEEE Journal of Solid-State Circuits 36: 1887–1899.
59. Radke RE, Eshraghi A & Fiez T (2000) A 14-Bit Current-Mode ΣΔ DAC Based Upon Rotated Data Weighted Averaging. IEEE Journal of Solid-State Circuits 35(8): 1074 – 1084.
60. Fujimori I, Longo L, Hairapetian A, Seiyama K & Kosic S (200) A 90-dB SNR 2.5-MHz Output-Rate ADC Using Cascaded Multibit Delta-Sigma Modulation at 8 x Oversampling Ratio. IEEE Journal of Circuit Solid-State Circuits 35: 1820–1828.
61. Lindfors S, Ööpik P & Halonen K (1997) N-Path Dynamic Element Matching for Multibit Bandpass Delta-Sigma Modulators. IEEE International Journal of Circuit theory and Applications 25: 335–346.
62. Shui T, Schreier R & F. Hudson (1999) Mismatch shaping for a Current-Mode Multibit Delta-Sigma DAC. IEEE Journal of Circuit Solid-State Circuits 34: 331–338.
63. Geerts Y & Steyaert M (1999) Guidelines for implementation of CMOS multibit oversampling modulators. ESPRIT project: SYSCON, K.U. Leuven, ESAT-MICAS.
64. Lindfors S (2000) CMOS Baseband Integrated Circuit techniques for Radio Receivers. Ph.D. Thesis, Helsinki University of Technology.
65. Kenney J & Carley LR (1988) A 16-bit 4th order noise-shaping D/A converter. Proc IEEE Custom Integrated Circuits Conference 21(7): 1–2.
66. Haiqing L & Schreier R (1999) A bandpass -shaped multi-bit ΣΔ switched-capacitor DAC using butterfly shuffler. Proc IEEE International Solid-State Circuits Conference, 1999. Digest of Technical Papers: 58–59.
67. Chan KL & Galton I (2006) A 14b 100MS/s DAC with Fully Segmented Dynamic Element Matching. Proc IEEE International Solid-State Circuits Conference: 2390-2399.
68. Fishov A, Siragusa E, Welz J, Fogleman E & Galton I (2002) Segmented mismatch-shaping D/A conversion. Proc IEEE International Symposium on Circuits and Systems, ISCAS 4: 679–682.
69. Henderson RK & Nys OJAP (1996) Dynamic element matching techniques with arbitrary noise shaping function. Proc IEEE International Symposium on Circuits and Systems, ISCAS 1: 293–296.
70. Gong X-M (2000) A 120 dB multi-bit SC audio DAC with second-order noise shaping. Proc IEEE International Solid-State Circuits Conference: 344–346.

71. Aghdam EN, Benabes P & Abbasszadeh J (2009) Completely first order and tone free partitioned data weighted averaging technique used in a multibit delta sigma modulator. Proc ECCTD 2009, European Conference on Circuit Theory and Design: 53–56.
72. LedziusR & Irwin J (1993) The basis and architecture for the reduction of tones in a sigma-delta dac, IEEE Transactions on Circuits Systems II 40: 429–439.
73. Nys OJAP & Henderson RK (1996) An analysis of dynamic element matching techniques in sigma-delta modulation. Proc IEEE International Symposium on Circuits and Systems 1: 231–233.
74. Chen K-D, & Kuo T-H (1999) An Improved Technique for Reducing Baseband Tones in Sigma-Delta Modulators Employing Data Weighted Averaging Algorithm Without Adding Dither. IEEE Transactions on Circuits and Systems II 46(1): 63–68.
75. Risbo L (1995) On the design of tone-free sigma-delta modulators. IEEE Transactions on Circuits Systems II 42(1): 52–56.
76. Dunn C & Sandler M (1995) Linearizing Sigma-Delta Modulators using Dither and Chaos. Proc IEEE International Symposium on Circuits and Systems, ISCAS: 625–628.
77. de la Rosa JM, Perez-Verdu B, Medeiro F, del Rio R & Rodriguez-Vazquez A (2001) Analysis and Experimental Characterization of Idle tones in 2nd-Order Bandpass SD Modulators - a 0.8um CMOS Switched-Current - Case Study. Proc IEEE International Symposium on Circuits and Systems, ISCAS 4: 774–777.
78. Doris K, van Roermund A & Leenaerts D (2002) Mismatch-based timing errors in current steering DACs. Proc International Symposium on Circuits and Systems, ISCAS 1: 977-980.
79. Fornasari A, Malcovati P & Maloberti F (2005) Improved Modeling of Sigma-Delta Modulator Non-Idealities in SIMULINK. Proc IEEE International Symposium on Circuits and Systems, ISCAS 2005: 5982–5985.
80. Neitola M & Rahkonen T (2010) Lookup-Table Based Settling Error Modeling in SIMULINK. Proc Scientific Computing in Electrical Engineering - SCEE 2008, Springer-Verlag Berling Heidelberg: Part II, edited by Roos J & Costa LRJ: 241–248.

# Original publications

I Neitola M & Rahkonen T (2010) A Generalized Data-Weighted Averaging Algorithm. IEEE Trans. Circuits and Systems II: Express Briefs 57(2): 115-119.

II Neitola M, Rahkonen T & Raappana J (2010) A Qualification Approach to DAC Mismatch-Shaping Methods. IEEE Trans. Circuits and Systems II: Express Briefs 57(11): 858-862.

III Neitola M & Rahkonen T (2011) Predicting and Avoiding Spurious Tones in a DWA-Mismatch-Shaping DAC. IEEE Trans. Circuits and Systems II: Express Briefs 58(6): 341-345.

IV Neitola M & Rahkonen T (2011) Compact Tone-Behavior Model for Delta-Sigma Modulator. Proc 20th European Conference on Circuit Theory and Design, ECCTD 2011: 261-264.



Original publications are not included in the electrical version of the dissertation.

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/

UNIVERSITY OF OULU P.O.B. 7500 FI-90014 UNIVERSITY OF OULU FINLAND

# ACTA UNIVERSITATIS OULUENSIS



ISBN 978-951-42-9748-9 (Paperback)
ISBN 978-951-42-9749-6 (PDF)
ISSN 0355-3213 (Print)
ISSN 1796-2226 (Online)

UNIVERSITY of OULU
OULUN YLIOPISTO