



State scaling of continuous-time sigma-delta modulators

C. Zorn¹, T. Brückner², M. Ortmanns², and W. Mathis¹

¹Institute of Theoretical Electrical Engineering, Leibniz University of Hannover, Germany

²Institute of Microelectronics, Ulm University, Germany

Correspondence to: C. Zorn (zorn@tet.uni-hannover.de)

Abstract. In this paper, the common method of scaling the feedback coefficients of continuous time sigma delta modulators in order to stabilize the system is enhanced. The presented approach scales the different states of the system instead of the coefficients. The new corresponding coefficients are then calculated from the solution of the state space description. Therewith, it is possible to tune the maximum out-of-band gain directly in continuous time. In addition, the input amplitude distribution between each quantization level of multi bit sigma-delta modulator can be adapted.

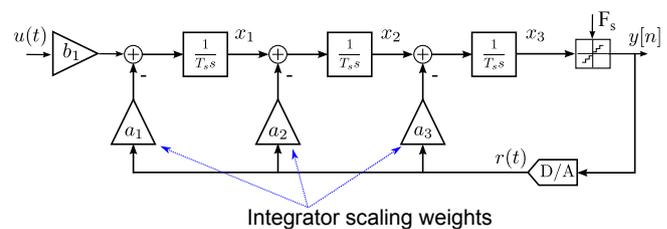


Fig. 1. CT $\Sigma\Delta$ -Modulator (CIFB).

1 Introduction

Analogue to digital (A/D) converters are one of the main parts in the digital home entertainment technology. There are many different approaches and circuits to convert analogue signals into digital signals. It can be differentiated between direct (successive approximation, Flash ADCs etc.) and indirect A/D converters like sigma delta modulators David F.Hoeschele (1994). Sigma delta data converters are compared to simple Nyquist rate or oversampled data converters, feedback systems and thus have the advantage of a closed loop control. Therewith it is possible to describe two different transfer functions. One transfer function is the signal transfer function (STF) and the other is the noise transfer function (NTF). With the NTF it is possible to implement a noise shaping in order to reduce the in-band noise of the modulator. Thereby the noise is shifted to higher frequencies, outside the interesting band. For an improvement of the performance of the modulator, the order of the filter structure or respectively of the NTFs can be raised Schreier (2004) by using a chain of integrators. By this way, the filtering function can be designed more aggressively, whereby the noise shaping of the signal is improved. The problem of a more aggressive filter function is that in addition to the decreased in-band noise, the out of band gain is increased. That can result in an over-

loaded quantizer input or a saturation of the integrators and therewith to a reduced maximum stable input amplitude. To stabilize the system, the NTF is mapped to a structure with weighted feedback path (filter coefficients) as it is shown in Fig. 1. This structure is called *Chain of Integrators with Distributed Feedback* (CIFB). A common method to optimize the system properties, like maximum signal to noise ratio (SNR) or dynamic range (DR), is to scale the feedback coefficients in order to reduce the aggressivity of the NTF. Therefore the simplest way is to scale the feedback value $r(t)$, that causes a scaling of all coefficients with the same factor. There are other methods to scale the coefficients with different factors, which will not be discussed in this paper (for further details see Ortmanns and Gerfers, 2006; Cherry and Snelgrove, 2000). Most of them have in common, that they all use a large number of coefficient variations and simulations in order to adjust the performance of the system. The problem is that the modulator is a non-linear system due to the mixed signal characteristics. The states are not independent from each other and each coefficient has a different strong influence on the system dynamics. In the following a method is presented that does not calculate the states from varied coefficient-sets, instead the new coefficients are calculated directly from the given state. Therefore a mixed signal model of the modulator is introduced that is used to get an analytical description of the maximum states in dependency of the input amplitude

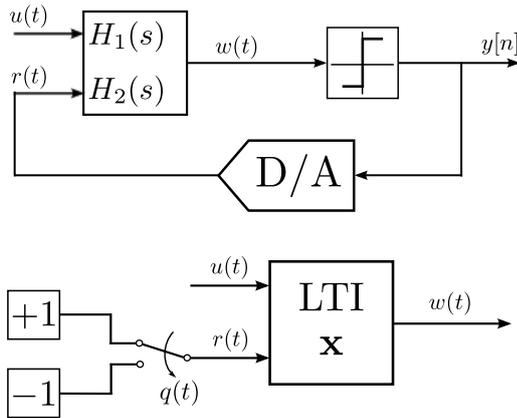


Fig. 2. Switched model.

and the filter coefficients. Finally the method is demonstrated on different example systems to show how the method can be used to tune the system performance.

2 Switched system model

A continuous time (CT) $\Sigma\Delta$ -Modulator as shown Fig. 1 can be described as a system with a linear time-invariant (LTI) loop filter and a piecewise linear feedback. The output signal $y[n]$ is given by sampling the quantizer output $w(t)$ periodically with a frequency $F_s = 1/T_s$. For a modulator with single bit quantizer, the feedback signal $r(t)$ is only switched between two different constant signals $r(t) \in \{+1, -1\}$. The loop filter system can be transformed to an equivalent open loop system with the two input signals $r(t)$ and $u(t)$ as depicted in Fig. 2. The input $r(t)$ is switched to $+1$ for $y[n] < 1$ and to -1 for $y[n] \geq 0$ with:

$$y(t) = \text{sign}(w(t)) \begin{cases} 1, & \text{for } w(t) \geq 0 \\ -1, & \text{for } w(t) < 0 \end{cases} \quad (1)$$

It should be noted that the feedback sign of r_i is inverted compared to Fig. 1. The negative sign is already considered in the negative feedback at the integrator inputs. For a system with multi-bit quantizer with B bits, the switching conditions change to

$$r_i = \text{sign} \begin{cases} r_1, & \text{for } w(nT_s) < w_1 \\ r_2, & \text{for } w_1 < w(nT_s) < w_2 \\ \vdots \\ r_L, & \text{for } w(nT_s) > w_{L-1} \end{cases} \quad (2)$$

where the feedback is switched between $L = 2^B$ different feedback values. The loop filter can now be described in the state space as

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B} \begin{bmatrix} u \\ r_i \end{bmatrix} \quad (3)$$

$$w(t) = \mathbf{c}^T \mathbf{x}(t), \quad (4)$$

where $x \in \mathcal{S}$ is the state vector, $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times 2}$ and $\mathbf{c}^T \in \mathbb{R}^{1 \times n}$ are the system matrix, the input matrix and the output vector respectively. The elements of the state vector $\mathbf{x}^T = (x_1, \dots, x_n)$ corresponds to the integrator outputs of an n -th order continuous time filter, where the order n of the filter equals to the number of integrators. With Eq. (2) and the output Eq. (4), there are $L - 1$ different switching conditions

$$w(nT_s) = \mathbf{c}^T \mathbf{x}(nT_s) = w_k \quad (5)$$

for $k = (1, 2, \dots, L - 1)$. For the single-bit modulator the Feedback is switched corresponding to Eq. (1) so that the switching condition is given by

$$w(nT_s) = \mathbf{c}^T \mathbf{x}(nT_s) = 0. \quad (6)$$

The solution $\mathbf{x}(t)$ of the system of differential equations for $t \in [t_i, t_{i+1})$ is given by

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_i)} \mathbf{x}(t_i) + \int_{t_i}^t e^{\mathbf{A}(t-s)} \mathbf{B} \begin{bmatrix} u(s) \\ r_i(x) \end{bmatrix} ds, \quad (7)$$

where t_i and t_{i+1} are two consecutive switching times (Jordan and Smith, 2007; Sun and Ge, 2005). If the system is only considered between two switching points, the feedback value $r_i(x)$ is constant and therewith the system is continuous. Equation (7) can be used to calculate the next switching point $x_{i+1}(t_{i+1})$ starting at $x_i(t_i)$ with $r = r_i(x_i)$.

3 State space analysis

With the Eqs. (3), (5) and the switching conditions (2), the states space is split into 2^B subsets. Each subset belongs to one constant feedback signal r_i and with a range depending on the quantizer step width given by

$$\Delta = \frac{\text{FS}}{2^B - 1}, \quad (8)$$

where FS is the full-scale output range (Ortmanns and Gersfers, 2006).

To simplify the description, the single bit case will be used to analyse the dynamic behaviour of the modulator. The qualitative behaviour of the trajectory can be transferred to each switching surface respectively to each subset of the multi-bit system. The state space of the single bit modulator is split into 2 subsets by the switching condition (6) which can be interpreted as hyperplane in Hesse normal form. These hyperplane will be called switching surface (SWS). Every first

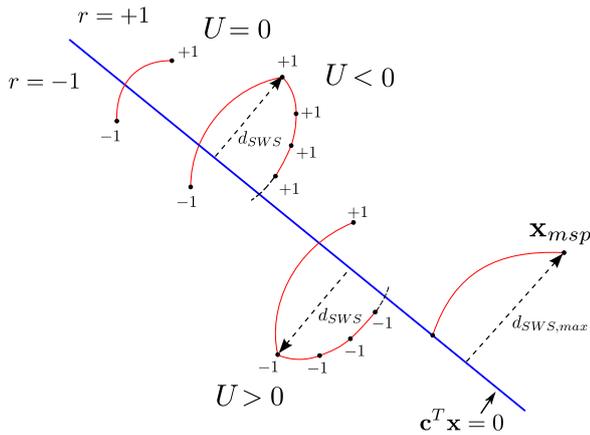


Fig. 3. State space dynamics.

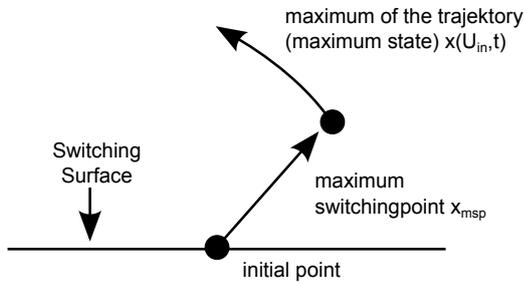


Fig. 4. Maximum switching point.

sampling point after the trajectory crosses the SWS will be called switching point. As shown in Fig. 3 the movement of trajectory is depending on the initial condition and the input amplitude. For a constant amplitude $U = 0$, the ideal modulator output $y[n]$ will be in averages as much $+1$ as -1 . For a constant input amplitude $U < 0$ the output will be more often $+1$ at the sampling time and therewith for $U > 0$ more -1 .

3.1 Maximum switching point

In addition to the trajectory movement, the distance (d_{sws}) from a switching point to the SWS is also depending on the initial point $x(t_i)$ and the input amplitude. For the case that the initial point belongs to the feedback $r_i = -1$ and the next sampling point is a switching point with $r_{i+1} = +1$, the trajectory will cross the SWS. If the distance of the initial point to the SWS is reduced, the time of movement of the trajectory after crossing the SWS with the initial feedback value is increased. That means, by using an initial value directly on the SWS as a *Worst Case Analysis*, the trajectory will move for a whole sampling period in the subspace $S_{r=+1}$ with the initial feedback $r = -1$. This leads to a switching point with maximum distance $d_{sws,max}$ from the switching surface. This switching point will be called maximum switching point (MSP) x_{msp} .

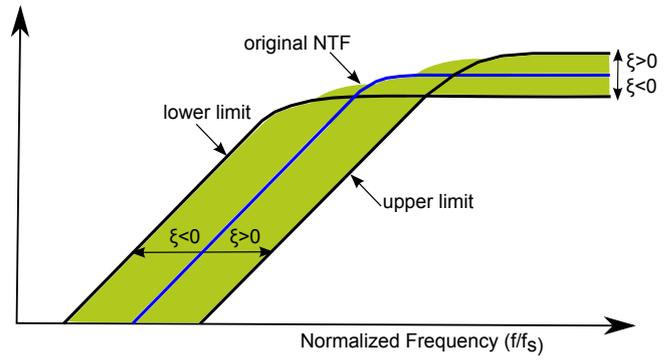


Fig. 5. Design-space.

4 State scaling

As depicted in Fig. 4, the distance from the SWS at the maximum switching point has not to be the maximum possible distance of the trajectory. The maximum switching point can be used as new initial point in Eq. (7) to calculate the maximum of the trajectory. In the following it is considered that the system is Lyapunov stable, all states are bounded and that the maximum switching point are as well the points witch maximum distance to the SWS in dependency of a constant input amplitude U . Therewith the Worst case analysis in combination with a system-matrix A in Eq. (3) that is nilpotent, can be used to determine a system of equations that gives an analytical description of the maximum states in dependency of the scaling weights:

$$x_{s,1} + k_1 b_1 + a_1 = x_{msp,1} \quad (9)$$

$$x_{s,1} + x_{s,2} + k_2 b_1 + k_3 a_1 + a_2 = x_{msp,2} \quad (10)$$

$$k_4 x_{s,1} + x_{s,2} + x_{s,3} + k_5 b_1 + k_6 a_1 + k_7 a_2 + a_3 = x_{msp,3} \quad (11)$$

The maximum states from the *Worst Case Analysis* can now be reduced by scaling the maximum switching points with a factor $\xi < 1$.

$$e^{AT_s} x_s(U) + \int e^{As} ds \cdot B \begin{bmatrix} U \\ r_{+/-} \end{bmatrix} = \xi \cdot x_{msp,p/n}(U) \quad (12)$$

This leads to a system of equation with new right-hand side:

$$x_{s,1} + k_1 b_1 + a_1 = \xi x_{msp,1} \quad (13)$$

$$x_{s,1} + x_{s,2} + k_2 b_1 + k_3 a_1 + a_2 = \xi x_{msp,2} \quad (14)$$

$$k_4 x_{s,1} + x_{s,2} + x_{s,3} + k_5 b_1 + k_6 a_1 + k_7 a_2 + a_3 = \xi x_{msp,3} \quad (15)$$

By solving the system of equations, the new coefficients corresponding to the scaled maximum states of the system can be calculated. That means that the continuous time system can be tuned in a design space (see Fig. 5) with the factor ξ without a large number of simulations.

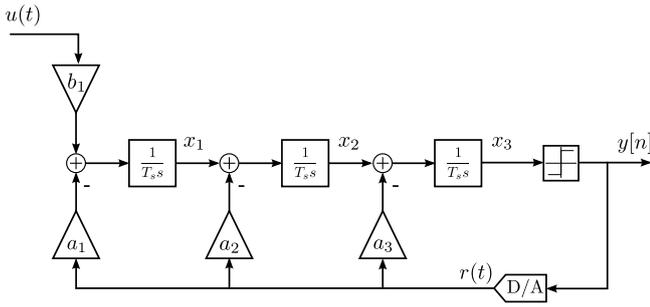


Fig. 6. 3.Order modulator in CIFB-Topology.

5 Example

The following examples from previous works will be shown to illustrate the effects of the state scaling method and give an overview about the possible applications.

5.1 State scaling: max. SNR and max. input amplitude

The first example was presented in Zorn et al. (2011a). Thereby the state scaling method was used to scale the maximum switching point for an input amplitude of $U = 0.8$ with a factor $0 < \xi < 1$ in order to increase the maximum SNR and maximum stable input amplitude. The example system is a 3rd order $\Sigma\Delta$ -Modulator in chain of integrators with distributed feedback topology (CIFB) with an oversampling ratio of $OSR = 128$ and a sampling frequency of $F_s = 320$ kHz. The D/A-converter is a NRZ-DAC Cherry and Snelgrove (2000) with the two output levels $r = +1$ or $r = -1$. The example modulator has no excess loop.

The desired discrete time (DT) noise transfer function $NTF(z)$ is obtained by means of the Schreier-Toolbox Schreier and Temes (2005) in MATLAB and the corresponding open loop noise transfer function $H_{noise}(z)$ of the NTF is transformed to its CT equivalent $H_{noise}(s)$ using the method proposed in Loeda et al. (2006). In order to realize a symmetric input, the coefficients are set to $a_1 = b_1$. As described in Sect. 3.1 an initial point on the SWS is needed according to the *Worst Case Analysis* to calculate the maximum state of the trajectory. For the input amplitude $U = 0.8$ the initial point on the SWS $\mathbf{x}_s = [0.1579, 0.2895, 0]^T$ is obtained for the state scaling method. By means of Eq. (7) the MSP for the feedback $r = -1$ is calculated to:

$$\mathbf{x}_{msp,n} = \begin{bmatrix} x_{msp,n,1} \\ x_{msp,n,2} \\ x_{msp,n,3} \end{bmatrix} = \begin{bmatrix} 0.2387 \\ 0.7335 \\ 1.1756 \end{bmatrix}. \quad (16)$$

The loop filter coefficients were calculated using Eqs. (12) to (15) for different $\xi = 1 \dots 0.7$. For each ξ , the following system of linear equations has to be solved for a_1 , a_2 and a_3 :

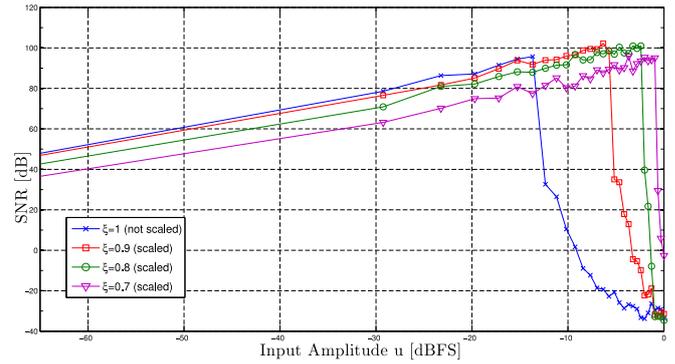


Fig. 7. SNR in dependency of the sine wave input amplitude u for $\xi = 0.1, 0.9, 0.8, 0.7$.

Table 1. Scaled coefficients.

ξ	0.95	0.90	0.85	0.80	0.75	0.70
a_1	0.0524	0.0471	0.0417	0.0364	0.0310	0.0257
a_2	0.1265	0.1087	0.0910	0.0733	0.0556	0.0378
a_3	0.3912	0.3657	0.3403	0.3148	0.2893	0.2638

$$0.1579 + 1.8a_1 = \xi \cdot x_{msp,n,1} \quad (17)$$

$$0.147 + 0.9a_1 + a_2 = \xi \cdot x_{msp,n,2} \quad (18)$$

$$0.36845 + 0.3a_1 + \frac{1}{2}a_2 + a_3 = \xi \cdot x_{msp,n,3} \quad (19)$$

The results of a simulation of the SNR for a sine wave input with increasing amplitude is depicted in Fig. 7 for various scalings.

5.2 State scaling: STF

The second example Zorn et al. (2011b) shows the effect of the state scaling method on the STF of a CT sigma delta modulator and how the method can be used to reduce the peaking of the signal transfer function. The initial NTF_2 is more aggressive than the NTF in example 5.1 with an oversampling ratio $OSR = 64$ and a sampling frequency of $F_s = 320$ kHz. The unscaled coefficients of the NTF_2 ($\xi = 1$, $a_1 = b_1 = 0.0578$, $a_2 = 0.1442$ and $a_3 = 0.4167$) are calculated by simulation under the aspect of a high SNR.

The optimization method was applied to this system as described before. The calculated coefficients are listed in Table 1.

The unscaled NTF_2 has large STF_2 -peaking for the assumed topology with only one feed-in coefficient as presented in Fig. 8. However, it can be seen that scaling the coefficients implies a reduction of the STF_2 's peaking as shown in Fig. 8. Decreasing the coefficients by a factor of

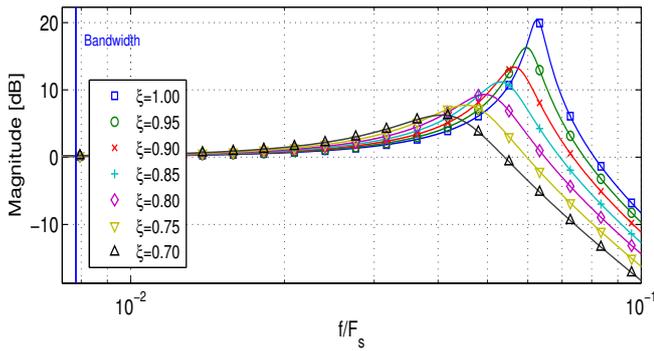


Fig. 8. STF_2 for various scaling factors.

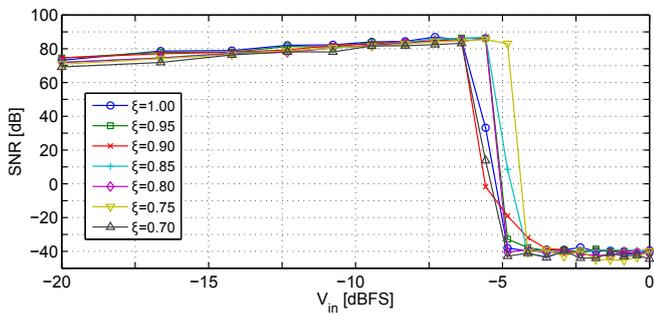


Fig. 9. Dynamic range for various scaling factors (NTF_2).

$\xi = 0.70$ attenuates the overshoot about 14 dB while the dynamic range remains almost unaffected (cf. Fig. 9).

6 Conclusions

In this paper a switched system model was presented that can be used to get an analytical description of the maximum states in dependency of the input amplitude and the filter coefficients. Based on this model the state scaling method was shown. Therewith it is possible to change the states of the system in order to tune the performance of the modulator directly in continuous time and to calculate the corresponding coefficients without a large number of simulations.

Acknowledgements. The authors would like to thank the German Research Foundation (DFG) for the financial support.

References

Cherry, J. A. and Snelgrove, W. M.: Continuous-Time Delta-Sigma Modulators for High-Speed A/D Conversion – Theory, Practice and Fundamental Performance Limits, Kluwer Academic Publishers, 2000.

David F. Hoeschele, J.: Analog-to-Digital and Digital-to-Analog Conversion techniques, John Wiley & Sons, second edn., 1994.

Jordan, D. W. and Smith, P.: Nonlinear Ordinary Differential Equations, Oxford University Press, 2007.

Loeda, S., Reekie, H. M., and Mulgrew, B.: On the Design of High-Performance Wide-Band Continuous-Time Sigma-Delta Converters Using Numerical Optimization, IEEE Transactions on Circuits and Systems I: Regular Papers, 53, 802–810, doi:10.1109/TCSI.2005.860116, 2006.

Ortmanns, M. and Gerfers, F.: Continuous-Time Sigma-Delta A/D Conversion Fundamentals, Performance Limits and Robust Implementations, Springer Series in Advanced Microelectronics, Springer Berlin Heidelberg, 2006.

Schreier, R.: The Delta-Sigma Toolbox Version 7.1, <http://www.mathworks.com>, 2004.

Schreier, R. and Temes, G. C.: Understanding Delta-Sigma Data Converters, IEEE Press, 2005.

Sun, Z. and Ge, S. S.: Switched Linear Systems – Control and Design, Springer-Verlag, 2005.

Zorn, C., Stegemann, S., Brückner, T., Ortmanns, M., and Mathis, W.: A Novel Optimization Method for CT Sigma-Delta-Modulators Using a Switched System Model, in: IEEE International Symposium on Circuits and Systems (ISCAS 2011), 1848–1851, 2011a.

Zorn, C., Widemann, C., Brückner, T., Ortmanns, M., and Mathis, W.: STF Optimization of 1-Bit CT SD Modulators Based on Scaled Loop Filter Coefficients, 2011b.