Advanced Noise-Shaping Data Conversion Techniques

by

Lu Jie

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in the University of Michigan
2021

Doctoral Committee:

Professor Michael Flynn, Chair
Assistant Professor Hun-Seok Kim
Professor Scott Mahlke
Associate Professor Zhengya Zhang

Lu Jie

jielu@umich.edu

ORCID: 0000-0001-5046-3917

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

The rapidly increasing digital device market is creating a huge demand for Data Conversion. In high-resolution data conversion applications, noise-shaping techniques are essential. This thesis first goes through the fundamentals of data conversion and noise-shaping techniques, and then discusses the latest advanced noise-shaping techniques that further boost data converter performance.

Three new noise-shaping techniques (architectures) are introduced. The first technique boosts the order of the noise-shaping system. A system-level approach is used to achieve high-order noise-shaping with enhanced robustness. A prototype noise-shaping SAR ADC is designed and measured. This design is the first 4th-order noise-shaping SAR ADC published.

The second technique effectively increases the bandwidth of noise-shaping data converters using a time-interleaving framework. A prototype time-interleaved noise-shaping SAR ADC is designed and measured to illustrate the technique. This prototype has the highest bandwidth among noise-shaping SAR ADCs to date.

The final technique aims to solve some inherent problems in continuous-time noise-shaping systems. A Continuous Time - Discrete Time (CT-DT) hybrid noise-shaping architecture is introduced, which adopts the advantages of both conventional CT and DT structures. A prototype ADC, showing benefits of tuning-free operation and high tolerance to non-idealities.

Chapter 1 Fundamentals of Noise-Shaping Data Converter

1.1 Concept and Needs of Data Conversion

In electrical signal processing, *Data Conversion* generally refers to the conversion between analog signals and digital signals. Analog signals can be loosely defined as signals continuous in both time-domain and value-domain. Most physical signals, such as voltage, current, temperature, and pressure, are analog signals. Digital signals, on the other hand, are eventually logical signals, which are discrete in time-domain and quantized in value-domain. In most electrical systems, digital signals are represented in two-states, known as binary digits (bits).

The conversion between analog and digital has two possible directions: from analog to digital, known as Analog to Digital (A/D) conversion; and from digital to analog, which is known as Digital to Analog (D/A) conversion. Hence, the system that realizes A/D conversion is called an Analog to Digital Converter (ADC), and one that does the opposite is called a Digital to Analog Converter (DAC). In this thesis, the discussions mainly focus on ADCs, but some techniques and conclusions are also applicable to DACs. To convert an analog signal to a digital one, an ADC generally consists of three functional parts (Fig. 1): 1) A sampler samples the analog signal at given moments, i.e., turns it into a discrete-time signal series; 2) A quantizer then quantizes the sampled signal and describe it with certain levels; The output signal is finally processed by 3) a coder, and is converted into a digital signal with the required coding format.

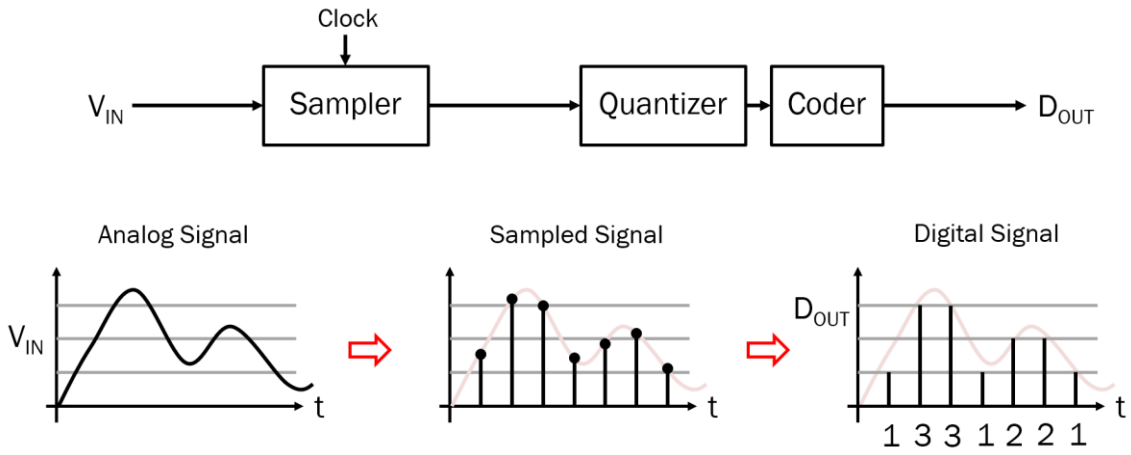


Fig. 1 A general model of an ADC

In recent decades, the requirements for data converters have been growing rapidly, as digitalization has become mainstream for almost all modern electrical systems. Digital systems have many advantages over conventional analog circuits, including high fidelity, high flexibility, and the capability to run sophisticated algorithms. Moreover, modern CMOS integration techniques have enabled low-cost but large-scale digital systems to be built, which are essential for sophisticalted electronic products. However, since real physical world signals are always analog, data converters are the necessary bridges so that digital systems can serve the analog world. Taking modern smartphones as an example, there are usually tens of data converters embedded in these devices, processing signals for audio, radio, video, and all kinds of sensing.

1.2 Main Specifications and Figure of Merit

There are many specifications of data converters, but usually, the most important ones are the speed, accuracy, power, and cost.

There are a few specifications to measure the speed of an ADC. The most commonly used ones are the sampling rate and the analog bandwidth. The sampling rate is the number of samples that an ADC can take and digitize per second, in units of S/s. In most cases, the sampling rate will also be the output digital data rate. According to the Nyquist Sampling Theory, the sampling rate

should be at least twice the signal bandwidth to prevent aliasing. The analog bandwidth is the input frequency range where the ADC gain loss is within 3dB. An ideal ADC should have an infinite analog bandwidth, meaning that signals at any frequency should be converted at a constant gain (even after aliasing). But in general, we consider an ADC as being good enough if the analog bandwidth is half the sampling rate. However, in some sub-sampling ADCs, the analog bandwidth can be much higher than that.

There are two sets of accuracy specifications for ADCs: Static Characteristics and Dynamic Characteristics. Static Characteristics focus on ADC's accuracy for DC inputs, such as offset, gain error, Differential-Non-Linearity (DNL), and Integral-Non-Linearity (INL). Dynamic Characteristics care more about the performance under AC inputs, including Signal-to-Noise Ratio (SNR), Signal-to-Noise-and-Distortion Ratio (SNDR), Spur-Free Dynamic Range (SFDR), and Effective Number of Bits (ENoB). For different applications, different metrics of accuracy are used. But in general, the AC performance is more conservative, as some dynamic effects can degrade the ADC with high-frequency inputs.

Power and cost are two important factors in many on-chip systems. For ADCs, the power is often quantified a energy per sample, because in many cases, the ADC power is proportional to the sampling rate. Cost is generally considered to be related to die area.

To compare two ADCs in overall performance, some Figure of Merits (FoM) are defined to characterize the specifications mentioned above with a single value. Two most commonly used FoMs are defined as:

$$FoM_{,Walden} = \frac{P}{2^{ENoB} \, BW \times 2} \quad \text{Eq. 1}$$

$$FoM_{Schreier} = SNR + 10\log_{10}\left(\frac{BW}{P}\right) \quad \text{Eq. 2}$$

which are given by R. H. Walden [1] and R. Schreier [2], respectively. Here, P is the ADC power consumption in Watts; SNR is in dB; ENoB is in bits, and BW is the effective bandwidth in Hz.

Walden's definition assumes that the ADC's power per sample is proportional to the number of quantization levels, while Schreier's definition assumes that it is proportional to the SNR. In general, Walden's FoM is more reasonable for low-resolution ADCs (ENoB < 10), while Schreier FoM is more reasonable at a higher resolution (ENoB > 10) where the SNR is thermal noise limited (rather than quantization error limited). Under the topic of noise-shaping, this thesis focuses on high-resolution converters, thus FoM_S is used in the following discussions.

A limitation of FoM_S is that it does not consider area (cost). Further, it ignores the robustness of the performance. To cover these, this thesis modifies Schreier and Sauerbrey's FoM [3] to consider both area and robustness:

$$FoM_{MS} = SNR_{worst} + 10\log_{10}\sqrt{\frac{BW}{P\times A\times OSR}} \quad \text{Eq. 3}$$

Here, SNR_{worst} is the worst-case SNR in dB under PVT variation for a certain yield, BW is the effective bandwidth in Hz, P is the total power consumption in Watts, A is the area in mm^2, and OSR is the over-sampling rate (will be explained later). This FoM definition assumes that the error is dominated by thermal noise ($\propto P^{-1}BW$) and kT/C noise ($\propto A^{-1}OSR^{-1}$). Assuming that both types of noise are comparable in a reasonable design, then both ($P^{-1}BW$) and ($A^{-1}OSR^{-1}$) should decrease by 3dB to increase SNR by 3dB.

1.3 Basic Concept of Noise-Shaping

Noise-shaping (NS) is a system-level method aiming at increasing the Signal-to-Noise Ratio (SNR) of data converters. Here, "noise" mainly refers to the quantization errors from the quantization behavior of the converters; And "shaping" means changing the power spectrum distribution of the noise, by some mechanisms discussed later, such that the noise power in the desired frequency band is suppressed. NS is usually associated with another technique: Over-Sampling (OS). An OS data converter runs at a much higher sampling rate than the necessary rate

(i.e., the Nyquist rate, or twice of the maximum signal frequency [4]), where the ratio between is called the Over-Sampling Rate (OSR). OS brings two benefits: 1) it spreads out the quantization noise in the broader frequency range. Therefore, the noise spectrum density is reduced as the total quantization noise power is constant; 2) More importantly, it leaves many "blanks" on the spectrum as the input signal is at relatively low frequencies. And the in-band noise can be moved to the blanks using the NS technique, which further reduces the in-band noise power. Fig. 2 shows the spectrum effects of OS and NS techniques, respectively.

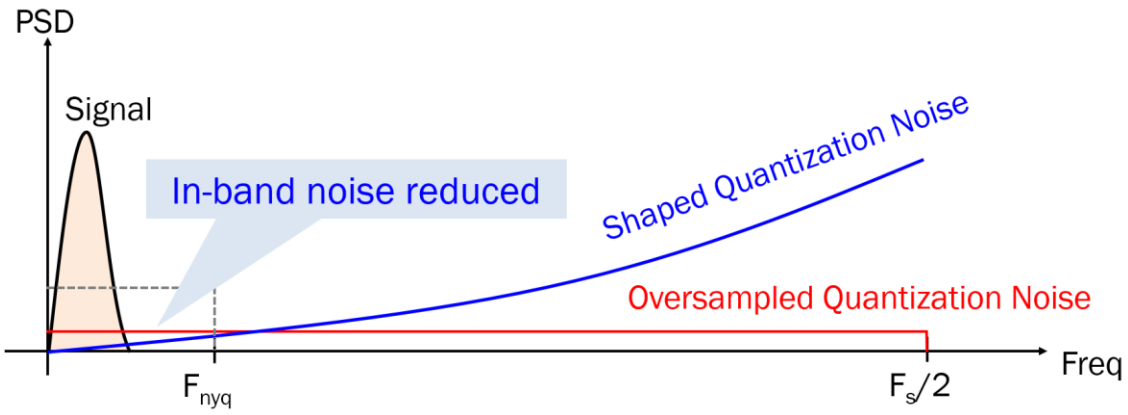


Fig. 2 Spectrum effects of Oversampling and Noise-Shaping

The realization of NS is often by some kind of feedback system. To better understand how these work, Fig. 3 shows an example of a simple NS system with feedback. Here, the quantizer is approximately modeled as an additive white noise source (E_Q). This is a reasonable assumption in most real cases. Two extra blocks, including a loop filter (H) and a feedback DAC, are added to the system and form a simple feedback loop, which essentially shapes the noise.

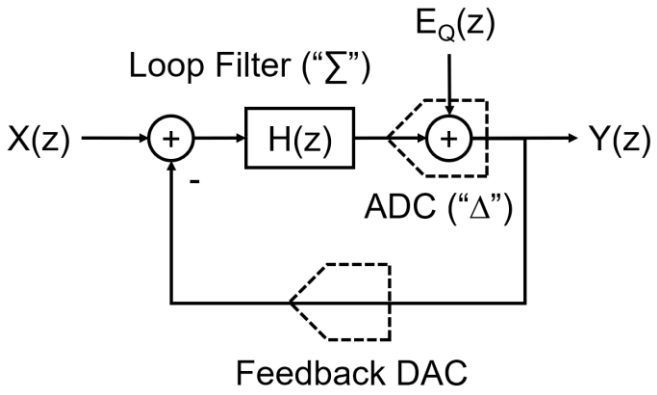


Fig. 3 Signa block diagram of a common NS system.

For simplicity, we consider the case that the input X is already sampled and is discrete in time (i.e., DT signal). Applying feedback theory, we can derive the transfer functions of this system, from the input and the noise source to the final output respectively, given as:

$$Y(z) = X(z)STF(z) + E_Q(z)NTF(z) \quad \text{Eq. 4}$$

$$STF(z) = \frac{H(z)}{1+H(z)} \quad \text{Eq. 5}$$

$$NTF(z) = \frac{1}{1+H(z)} \quad \text{Eq. 6}$$

where the transfer function from input to output is called the Signal Transfer Function (STF), and the transfer function from noise to output is called the Noise Transfer Function (NTF). The results above indicate that the system will pass the input signal while suppressing the noise at the frequency where the loop gain (H) is large. Therefore, we can harness this by making the loop filter high-gain in the wanted band, thereby the in-band noise is effectively reduced, and the SNR is boosted.

The earliest NS data converters have first-order loop filters (i.e., an integrator) and 1-bit quantizers [5]. Therefore, people also name those converters as "Delta-Sigma" (or "Sigma-Delta") converters. In the later developments, higher-order loop filters and multi-bit quantizers are used. But the term Delta-Sigma is still used and can be regarded as an alternative name for NS techniques.

1.4 Common Noise-Shaping Converter Architectures

Although the signal models behind them may be quite similar, NS architectures vary greatly. For NS ADCs, the most conventional architecture is a block-to-block implementation of the signal model, where the loop filter is built with op-amps. Fig. 4 shows a generalized block diagram of the conventional Sigma-Delta (SD) ADC.

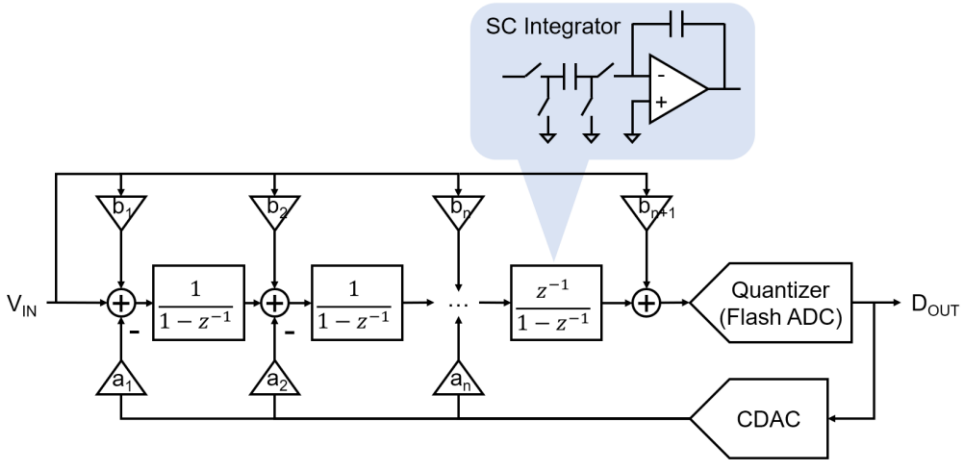


Fig. 4 Sigma-Delta ADC of conventional architecture.

Based on the loop filter's topology, there are a few sub-classes of this architecture (e.g., CIFF, CIFB, and something in between [6]). The loop filter can also work in either discrete-time (DT) mode or continuous-time (CT) mode. However, all of these can be essentially counted as variants of the same architecture. This kind of architecture is good for robustness and accuracy, mainly thanks to the high gain of op-amps which suppress most non-idealities. Nevertheless, as CMOS processes scale down, building high gain op-amps is getting harder. Moreover, in a sense, the op-amp's high gain is highly redundant, which means the design does not fully use the power and area.

On the other hand, recently some alternative architectures exhibit various advantages over the conventional ones. One of the most promising is the Noise-Shaping SAR architecture [7]. As a variant architecture of SAR, NS SAR adopts the advantages of SAR architecture, including low-

power, area-compact, and scaling-friendly circuits. And nearly a decade of development, NS-SAR can now achieve comparable SNR performance to conventional NS architectures. Fig. 5 shows a generalized block diagram of the NS SAR architecture. The main difference between it and a plain SAR ADC is the two extra loop filters inserted (H1 and H2), which process the residue (i.e., quantization errors, E_Q) of the SAR conversion. The two feedback paths form a similar signal model to Fig. 3, and therefore realize NS effects. The NTF of this model can be derived as:

$$\mathrm{NTF_Q} = \frac{1-\mathrm{H1(z)}}{1-\mathrm{H2(z)}} \qquad \text{Eq. 7}$$

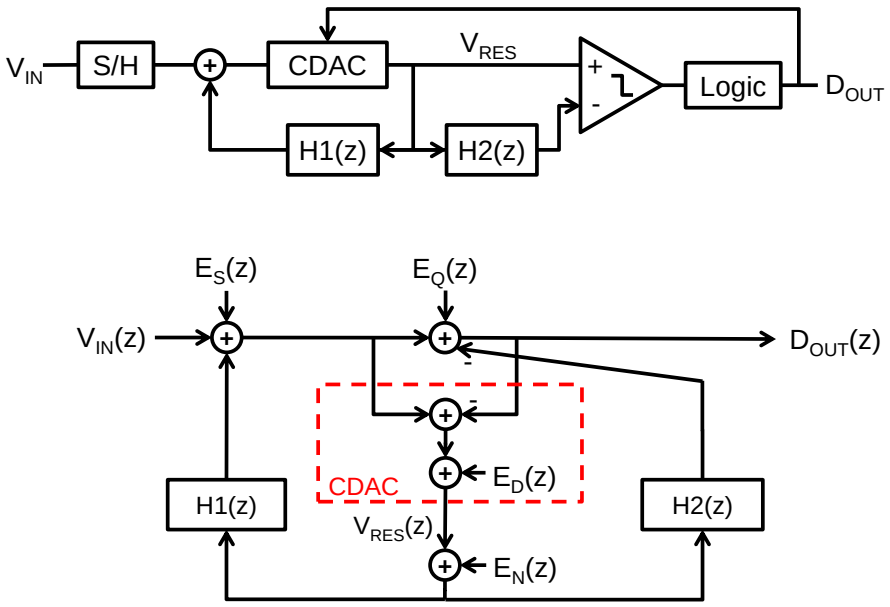


Fig. 5 Generalized behavioral (top) and signal (bottom) model of NS SAR.

Here, the feedback configuration with H1 is also named as Error-Feedback (EF) structure [8], while the configuration with H2 is called Cascaded-Integrator-Feed-Forward (CIFF). Eq. 7 implies that an FIR filter in H1 can place the zeros of NTF, while an FIR filter in H2 can place the poles instead. Generally, the zeros of the NTF is much more critical, so a pure EF based NS SAR is very attractive as it only requires an FIR filter with no integrator.

Note that the Error-Feedback structure can also be applied in the conventional Sigma-Delta architecture. However, it is much harder to extract the residue in conventional architecture. Thus

EF structure is not very practical for them [9]. On contract, a key advantage of the SAR ADC is that the residue can be precisely extracted, as it is naturally presented on the CDAC after each conversion. The availability of the residue enables NS SAR to be more flexible in realizing NS.

Fig. 6 gives an intuitive illustration of NS SAR architecture's advantages by comparing NS-SARs to the conventional Sigma-Delta ADCs (data from the ADC survey [10]). Obviously, NS SAR is much more power-efficient than the Sigma-Delta architecture, and is notably lower in cost (i.e., smaller in the area).

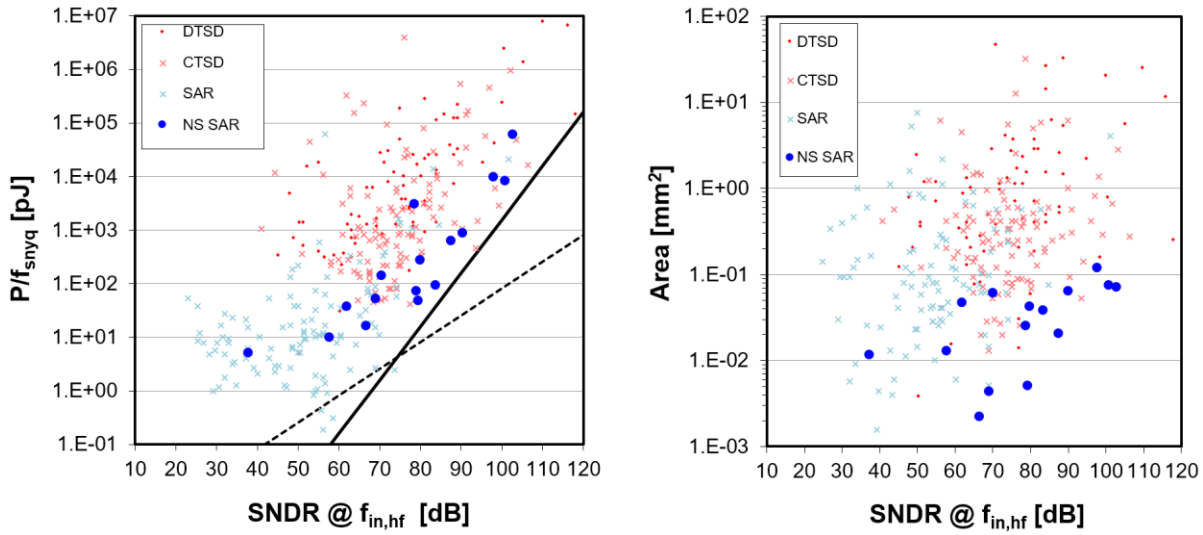


Fig. 6 Power efficiency and area comparison between conventional SD ADC and NS SAR.

Due to the notable advantages of NS SAR, the proposed techniques in this thesis are based (or partially based) on the NS SAR architecture. But some of the techniques are independent of the implementation and are also applicable to the other architectures.

Lastly, there is an interesting thing to mention: Although most NS systems are based on feedback, there are still some exceptional cases where NS can be achieved is open-loop. One of the examples is the Voltage Controlled Oscillator (VCO) based converter [11]. This kind of converter provides an inherent 1st-order NS. The open-loop nature leads to its simplicity in

circuitry and avoids any feedback stability concern. Therefore, it is usually used as a building block (e.g., the quantizer) in some advanced NS converter.

1.5 Fundamental Limitations and Trade-offs

The Noise-Shaping technique has some fundamental limitations. Some of these limitations are due to the nature of feedback systems. A general NS system shown is Fig. 3 is essentially a Single-Input-Single-Output (SISO) feedback system with a single loop, where the noise is recognized as a "disturbance" in classic control theory. By definition, NTF is equivalent to the sensitivity function (S) in this theory. Bode gives a well-known Sensitivity Integral for this kind of feedback system, and B. Wu extends it to the DT case [12]:

$$\int_{-\pi}^{\pi} \ln\left|S\left(e^{j\omega}\right)\right| d\omega = \int_{-\pi}^{\pi} \ln\left|NTF\left(e^{j\omega}\right)\right| d\omega = 2\pi\left(\sum \ln|p_k| - \ln\left|\lim_{z\to\infty} L(z) + 1\right|\right) \quad \text{Eq. 8}$$

where $L(z)$ is the loop filter, and p_k are the unstable poles. In most NS system, the loop filter is designed to be stable and has at least one more pole than zero. Besides, the NTF is even in general. Thus Eq. 8 can be further simplified as:

$$\int_{0}^{\pi} \ln\left|NTF\left(e^{j\omega}\right)\right| d\omega = 0 \quad \text{Eq. 9}$$

The integral above implies the first limitation: the area under the NTF curve, in log scale, is always a constant. Therefore, lowering the NTF at some frequencies inevitably raises the NTF in some other frequencies by the same amount in dB.

The second limitation is related to the maximum slope of the NTF. From filter theory, the roll-off rate of an N-th filter is approximately 20N dB per decade. This rule is also applicable to the slope of an NTF, which means that at least N-th order NTF is required to achieve a slope of 20N dB per decade.

Lastly, the maximum response amplitude of the NTF is also limited. This is because the quantizer in NS system is highly non-linear, and our additive noise model will fail if the input to

the quantizer is too large (which is also known as *quantizer overload*). Once that happens, the behavior of the whole NS system changes, increasing the in-band noise, and in some cases, becoming even unstable. A practical way to prevent this is to limit the NTF response to be lower than a specific threshold. But unfortunately, there isn't a precise (i.e., analytical) conclusion on this threshold. Usually, people use some rules of thumb for a rough design, and then optimize with simulation. One well-known rule of thumb is the (modified) Lee's rule [13]: A NS system (in the form of Fig. 3) with a 1-bit quantizer is likely to be stable if max(|NTF(z)|) < 1.5. And generally, increasing the number of quantizer levels can raise the threshold of the NTF.

The three limitations above directly lead to some design trade-offs in a NS system. More specifically, the in-band SNR improvement, OSR, the NTF order, and the quantizer's resolution are tightly linked in an optimal NS system design. Fig. 7 illustrates such trade-offs intuitively.

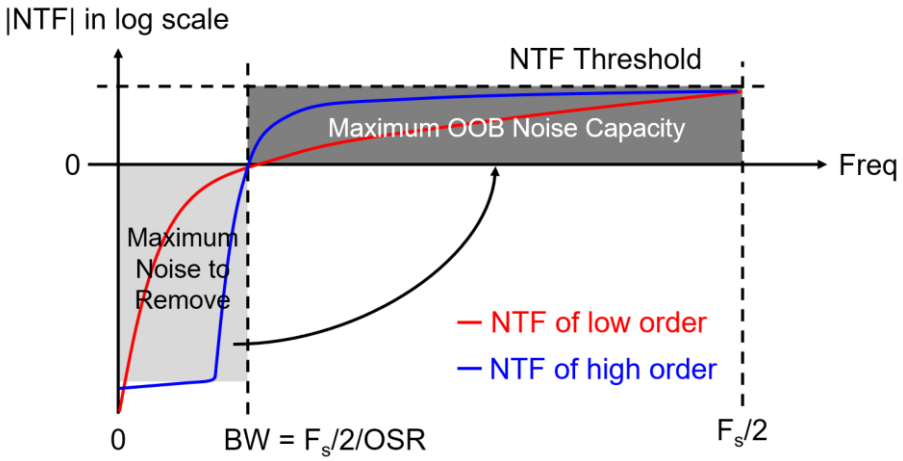


Fig. 7 Trade-offs between in-band SNR, OSR and NTF's order

Due to the NTF threshold and given OSR, there is a limited area in the out-of-band region. And due to the Sensitivity Integral (Eq. 9), this area is equal to the noise area that we can remove in-band. So in an ideal case, where we can implement an infinite order of NTF, the maximum SNR improvement in-band can be calculated as:

$$\Delta SNR_{ideal} = 10\log_{10}(NTF_{thres}) \times (OSR - 1) \qquad \text{Eq. 10}$$

In practice, the order of NTF is finite. In such cases, it is hard to derive a simple expression for the SNR improvement, and usually, this can only be examined by numerical methods. Although some recent work [14] already derives some analytical conclusions about the optimal NTF under given constraints (i.e., NTF threshold, NTF order, and OSR), these optimal NTFs are often heavily rely on the coefficient precision, and thus are not very useful. In a practical design, the NTF is usually chosen to be sub-optimal so that it is less sensitive to the coefficients.

In the following chapters, we will go through three techniques extending the Noise-Shaping. The first technique (Chapter 2) intends to increase the system order, and reduces the cost and sensitivity. The second technique (Chapter 3) aims to improve the bandwidth of noise-shaping converters. The third technique (Chapter 4) further enhances the converter's robustness and practicability, while keeping the speed advantages. Finally in Chapter 5, we introduce some circuit-level techniques to enable better implementations.

Chapter 2 Nested Noise-Shaping Converter

2.1 Different Forms of Noise Transfer Function

We have learned from Section 1.5 that the SNR improvement by NS is related to the NTF and OSR. But it is less obvious that the form of NTF also matters a lot. In this chapter, different forms of NTF are compared, and then the Nested structure is naturally introduced. Later, an implementation of the Nested structure in a NS SAR is given, showing how this technique can be applied in practice. The following discussion is based on DT NS systems, but many conclusions apply to CT NS systems as well.

The NTF of a NS system essentially describes a filter applied to the quantization noise. In the field of filter implementation, there are three common structures for DT filters: Direct Form (type II), Cascade Form, and Parallel Form, as shown in Fig. 8. Similarly, the NTF can also be implemented with one of these three forms. Although all the forms are equivalent in sense of algebra, they are quite different in some practical concerns.

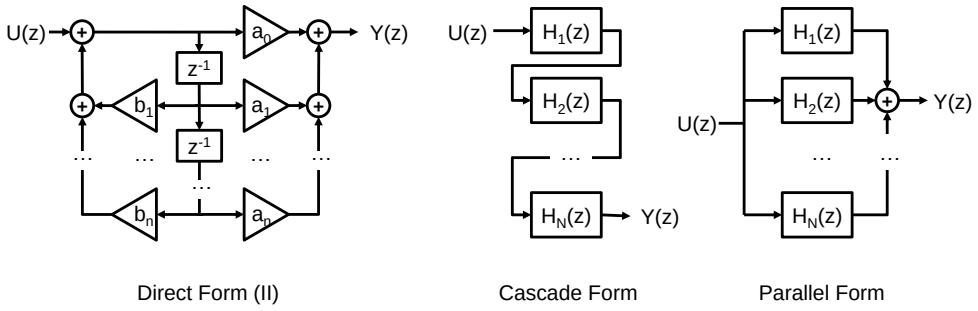


Fig. 8 Different DT filter implementation forms.

The direct form is the most straightforward one, and it can be easily constructed from the filter coefficients. However, the direct form has several inherent disadvantages. It needs the

greatest number of summing blocks and requires relatively large coefficient magnitudes, which increase the cost in terms of power and area. Also, the direct form suffers from high sensitivity to variation in coefficients, which means poor PVT robustness.

The cascade form is the most robust against coefficient variation and requires the smallest coefficients. Unlike the parallel form, the cascade form enables direct, independent control of each zero (pairs), and thus has even lower coefficient sensitivity than the parallel form [15]. Another advantage of the cascade form is that the later stages filter the preceding stages' noise, which reduces the overall in-band noise.

To show the difference between the direct and cascaded forms, a mathematical model of a NS SAR converter is built in this thesis to investigate the theoretical performance. This model considers the variation of NTF coefficients, and evaluates the performance using the modified FoM (see Section 1.2) by a Statistical method. Therefore we can exam the worst-case performance of a NS SAR converter with different orders and different forms of NTF. Fig. 9 shows the result of this modeling with given constraints. The details of the model can be found in Appendix A.

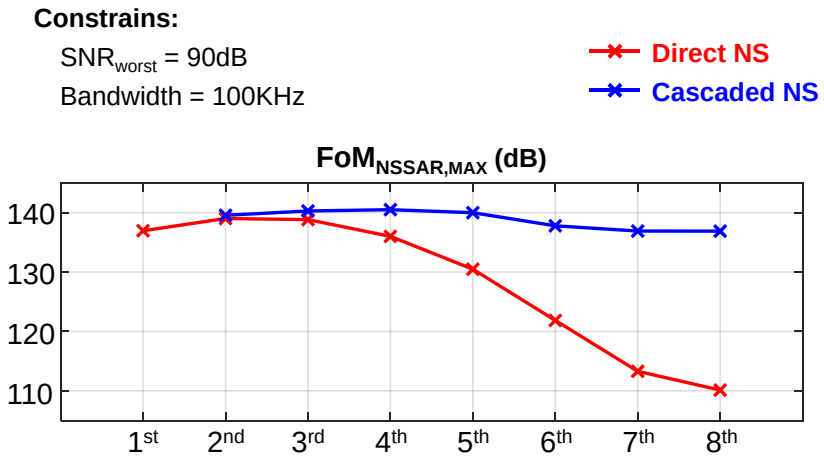


Fig. 9 Maximum FoM versus NTF order for different NTF form (yield = 95%).

We notice that for a certain SNR and BW (90dB and 100kHz in this example), the optimal FoM for the ADC with a direct-form NTF occurs for 2nd-order, while the peak for the cascade

form is with a 4th-order NTF. We also notice that a better FoM is possible with higher-order noise-shaping for the cascade form NTF, while direct form cannot provide as good a performance.

2.2 Realizing Cascaded NTF by Nested Noise-Shaping

A. Cascading Loop Filter

A straightforward "cascading" of an NS system might directly cascade the loop filter. Fig. 10 shows an example of cascading the loop filter in an Error-Feedback NS system. The cascade of FIR_1 and FIR_2 implements the loop filter.

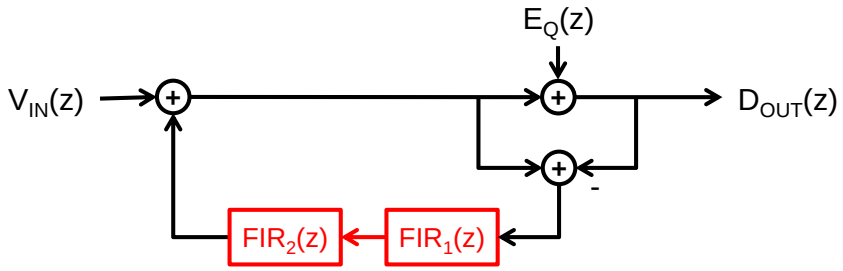


Fig. 10 EF NS system with loop filter in cascaded FIR form.

Although such an implementation can reduce the magnitude of the filter coefficients, it does not preserve the ability to independently place the zeros of the NTF, which is an essential advantage of cascade form. The NTF for E_Q is:

$$NTF_Q(z) = 1 - FIR_1(z) \cdot FIR_2(z) \qquad \text{Eq. 11}$$

We notice that the zeros of the NTF in Eq. 11 are not directly related to FIR_1 or FIR_2. In other words, cascading the loop filter does not provide a cascade form of the NTF. This problem not only increases design complexity, but also worsens the robustness of the NTF. For a 4th order NTF in the form of Eq. 11, the SNR degradation with coefficient variation can be 20dB larger than an NTF with a true cascade form. We provide a more detailed numerical comparison in the next section.

B. Nested Structure

Some existing work on CT SD ADCs with noise-shaped quantizers [16][17] suggests a practical alternatives to using a cascaded loop filter. Fig. 11 illustrates the signal model of this kind of SD ADC.

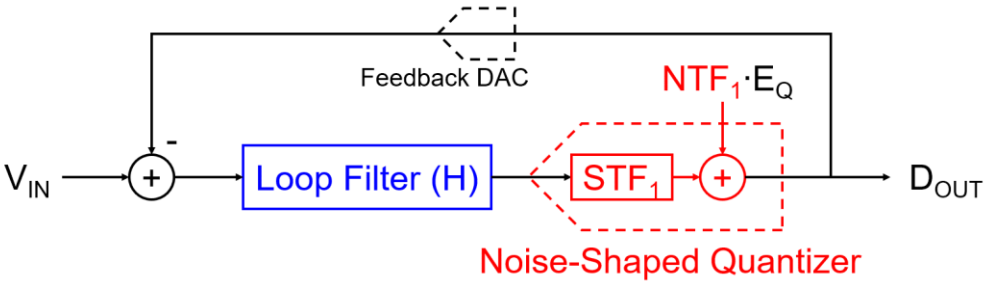


Fig. 11 SD ADC with noise-shaped quantizer.

In these SD ADCs, an outer modulator loop is built around an inner NS system (i.e., the Noise-Shaped Quantizer). In classic control theory, this configuration is also recognized as the Nested Structure. Notice that the overall NTF of this system is in the cascade form of the two individual NTFs:

$$NTF(z) = NTF_1(z) \cdot NTF_2(z) = \frac{NTF_1(z)}{1+H(z)STF_1(z)} \quad \text{Eq. 12}$$

Although most of these SD works use the nested structure for higher NTF order, this configuration can provide other benefits of cascade form, especially improvement of robustness. Theoretically, any NS converter can be used as the noise-shaped quantizer. But in practice, we would like to make the signal transfer function of the quantizer (i.e., STF_1) to be all pass, so that the design of the outer NS loop is more straightforward. Conventionally, the most commonly used noise-shaped quantizer is an error-feedback-based NS converter, also known as the Noise Coupling technique [18]. Fig. 12 shows the signal model of a Noise Coupled SD ADC.

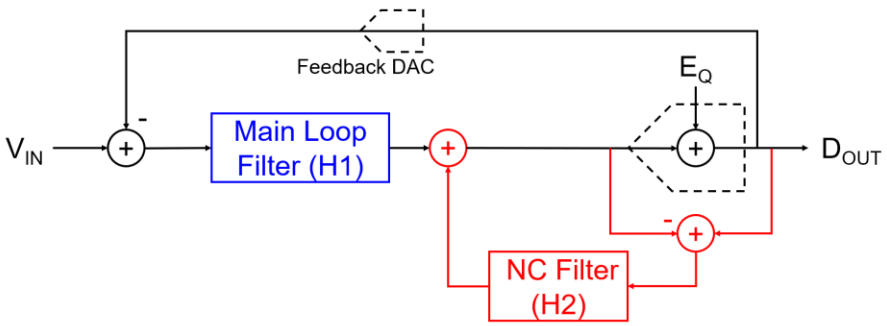


Fig. 12 SD ADC with Noise Coupling (EF based NS quantizer).

As mentioned in Chapter 1, EF based NS system is much simpler in its loop filter design. Still, due to the difficulty of residue extraction, EF structure is never used to form the outer NS loop in conventional designs. However, with the help of SAR architecture, residue extraction is much easier. Besides, the tolerance to residue extraction error can be greatly relaxed if the nested structure is induced. Thus, we find that a pure EF based nested NS system is practical, and can be even simpler than conventional designs. Fig. 13 gives the model of the EF NS system in the nested structure.

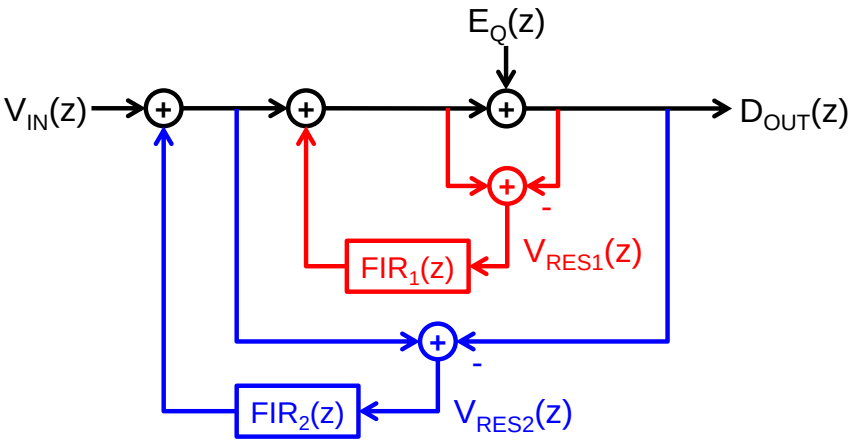


Fig. 13 Nested Error-Feedback Noise-Shaping System.

In this NS system, the overall NTF is in cascade form, and the sub-loop-filters independently control the zeros:

$$NTF(z) = \big(1 - FIR_1(z)\big)\big(1 - FIR_2(z)\big) \qquad \text{Eq. 13}$$

As mentioned, it enables a high-order NTF with small coefficients. But more importantly, this independent control dramatically reduces the sensitivity of the NTF to coefficient variation. To better illustrate this, Fig. 14 shows the calculated performance distributions for four different NTFs for a 1% 1-sigma coefficient variation. The horizontal axis represents the SNR improvement from noise-shaping, assuming an OSR of 10. The NTF of the proposed nested structure provides a much better overall SNR, even accounting for manufacturing variation. Notice that, the direct-form loop filter and cascaded-form loop filter (i.e., Fig. 10) of 4th-order NTF perform even worse than 2nd order one. Therefore, the proposed cascaded architecture is crucial for reliable high SNR operation.

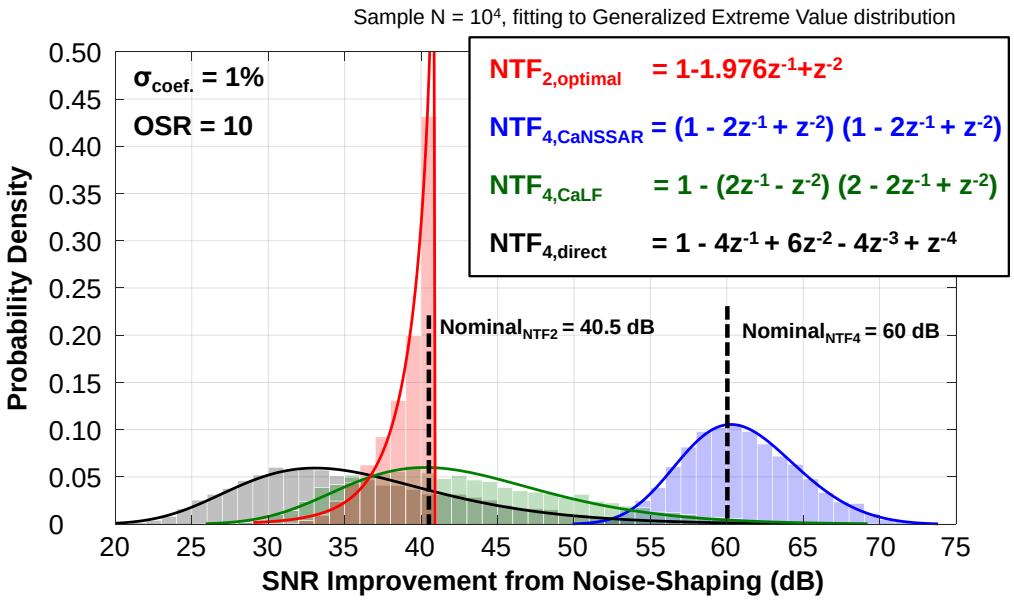


Fig. 14 Performance distributions for four different NTFs for a 1% 1-sigma coefficient variation.

The nested structure extends the advantages of EF at the circuit-level, as well. The outer loop shapes the noise from the inner loop and the noise from the FIR1 filter. This shaping greatly relaxes the noise requirements for FIR1, and thus reduces power and die area.

2.3 A Design Example of Nested Noise-Shaping SAR ADC

In this example, we apply the Nested Structure to a NS SAR ADC and observe the advantages. The prototype was taped-out and measured, and the result was published in [19].

A. Architecture

For simplicity, the NS SAR in this example is based on pure EF structure, which means there is no CIFF path (see section 1.4). The goal is to implement the signal model shown in Fig. 13.

As we know, the SAR conversion naturally converges the comparator's input to near zero. In other words, after each round of conversion, we will get the residue (E_Q) at the comparator's input. At the same time, notice that in Fig. 13, the input to FIR2 (V_{RES2}) is derived as:

$$V_{RES2}(z) = V_{RES1}(z) - FIR_1(z) \cdot V_{RES1}(z) = FIR_1(z) \cdot E_Q(z) - E_Q(z) \qquad \text{Eq. 14}$$

This suggests a straightforward architecture of a Nested NS SAR, where we can use an adder to synthesis the V_{RES2} from E_Q. Fig. 15 shows the block diagram of this method.

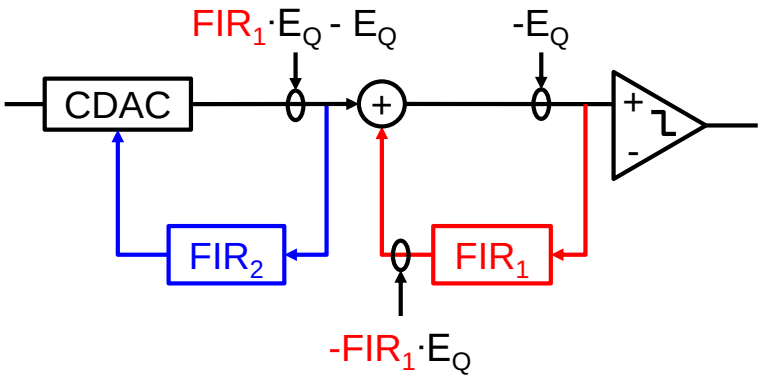


Fig. 15 A straightforward architecture of the Nested NS SAR.

Certainly, we can use an active adder to implement Fig. 15, but an active adder consumes extra power, brings in extra noise, and may even worsen linearity. Passive adder is thus attractive. In this example, we introduce a serial capacitor [20] to realize the summation, as shown in Fig. 16.

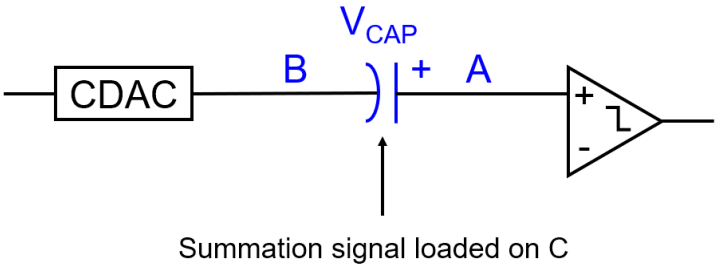


Fig. 16 Realizing passive summation by a serial capacitor.

It is evident that the voltage at node B is the voltage at node A subtracted by the voltage across the capacitor. Such summation needs only a single capacitor, and the summation gain is independent of the capacitance. This makes it insensitive to any possible variation. The FIR filters should be active to charge the capacitor, but this is acceptable as it only processes a sampled signal (i.e., DT signal). We will cover the FIR filter in the later section.

The design aims to provide 90dB SNR over a 100kHz bandwidth. For simplicity, we choose a classic 4th order NTF ($(1-z^{-1})^4$) such that it can be easily decomposed into the product of two sub-NTF for cascading. OSR is set to 10x, which corresponding to a 2MHz sampling rate. Such a NS configuration leave plenty margin for circuit variations.

B. Implementation

Fig. 17 shows the complete schematic of the prototype Nested NS SAR ADC. Besides the nested EF structure, the remainder of the implementation is relatively straightforward. The SAR core uses an 8-bit CDAC, a Strong-ARM dynamic comparator, and asynchronous SAR logic [21]. The input sampling switch is boot-strapped, while the other switches are simple transmission gates. Chopping in the main signal path suppresses flicker noise.

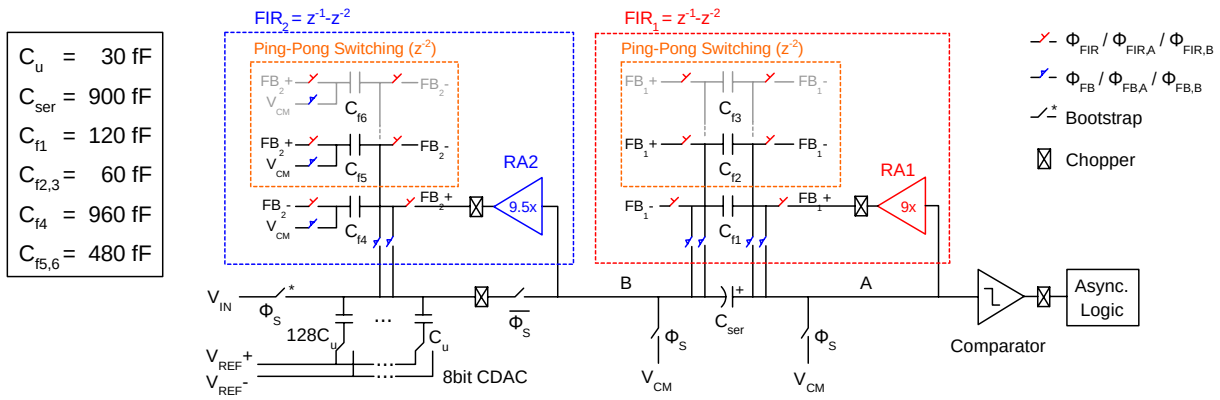


Fig. 17 Schematic of the example Nested NS SAR ADC.

A serial capacitor (C_{ser} in Fig. 17) cascades the noise-shaping stages and implements the inner loop feedback. C_{ser} is reset before each round of feedback to clear any residual charge from the previous cycle. Feedback for the inner stage is through simple change-sharing of the output of FIR_1 with C_{ser}. In the subsequent conversion, the voltage on the top plate of C_{ser} (V_A) is naturally equal to the voltage on the bottom plate of C_{ser} (V_B), plus the voltage across C_{ser} (V_{ser}):

$$V_A = V_B + V_{ser} = V_B + \alpha V_{FIR1} \qquad \text{Eq. 15}$$

where, V_{FIR1} is the output of FIR_1 and α represents the loss due to charge-sharing. As mentioned, this summation requires only one single capacitor and is entirely passive. And because the outer stage relaxes the noise requirements for the inner stage, C_{ser} can be a relatively small capacitance. The power and area overheads are therefore low. On the other hand, the large capacitance of CDAC suppresses the noise from FIR_2, which is the dominant noise of the system. In effect, this reuse of the CDAC capacitance for FIR_2 noise reduction reduces the area further.

A residue amplifier (RA) driving the input of each FIR filter compensates for the charge-sharing loss (α) and maintains the FIR coefficients. Cross-differential sampling, as in [20], provides an extra gain of 2, which relaxes residue amplifiers' gain requirement and provides common-mode suppression (Fig. 18). Charge-injection from the switches is negligible because of the fully differential operation, and also because of the relatively large capacitance of FIR filters.

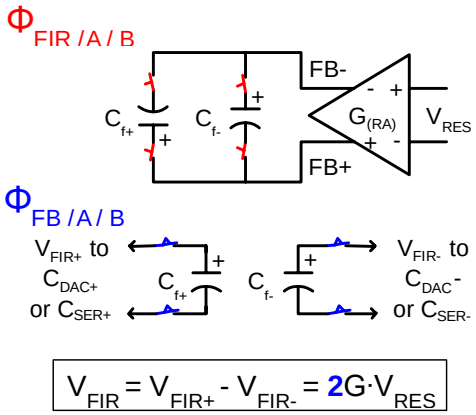


Fig. 18 Cross-differential sampling of FIR capacitors.

The FIR filters in this design use ping-pong switching to realize two cycles of delay. C_{f2} and C_{f3} alternatively sample the output of RA1 (i.e., FB_1) and transfer their sampled charges to C_{ser} two cycles later, as Fig. 19 shows. Such an implementation simplifies the switching logic as there is only a single charge-sharing step in each conversion cycle.

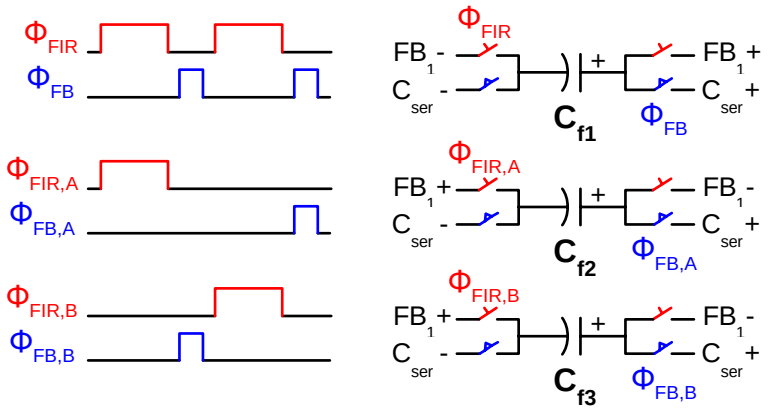


Fig. 19 FIR_1 switches timing. FIR_2 operates with same timing.

FIR_2 operates similarly to FIR_1 and injects its output to the CDAC by charge-sharing [8]. This charge-sharing inevitably attenuates the sampled input voltage on CDAC and thus reduces the overall SNR. A high RA gain is required to mitigate the attenuation entirely, but this increases the design complexity and sensitivity. Therefore, we choose an RA with around 10x gain, leading to a tolerable SNR degradation of 1.5dB (16%).

Conversion occurs in four steps, as described in Fig. 20. The overall timing is simple and little more complex than in a conventional SAR ADC.

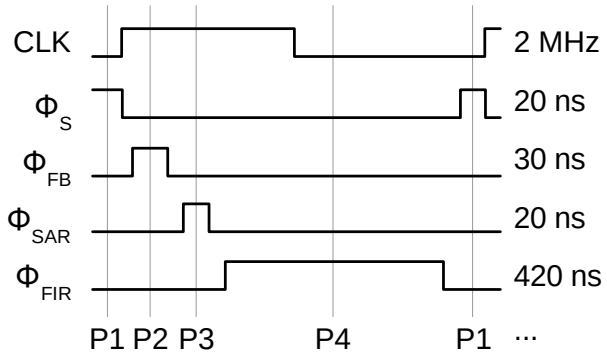


Fig. 20 Operation timing.

In the first step (P1), CDAC samples the input signal. At the same time, the series capacitor C_{ser} discharges (reset). Feedback occurs in the second step (P2). The FIR filters inject their outputs onto C_{ser} and CDAC by charge-sharing. The third step (P3) proceeds in the same way as conversion in a conventional SAR ADC. The fourth and final step (P4) is residue sampling. In this step, the two RAs amplify the residues on CDAC and the top plate of Cser. After that, the two FIR filters sample the amplified outputs using the cross-differential connection for an additional gain of 2.

This operation tolerates linear parasitic capacitance at the plates of C_{ser} and $C_{f1\text{-}6}$ because these capacitances only cause a linear gain error, and Nested NS SAR is insensitive to gain error. However, any nonlinear parasitic capacitance on nodes A and B, such as from the input transistors of the RAs, causes distortion and degrades SNDR. To reduce this distortion, we induce the Parasitic Pre-Charging method that reset nodes A and B to V_{CM} during P1. After P1, the parasitic capacitances at A and B share some of the charge from CDAC and C_{ser}. However, since V_A and V_B converge to virtual ground during SAR conversion, the charge shared by these two parasitic capacitors eventually returns, mostly eliminating the nonlinear error. More details on the Parasitic Pre-Charging technique are discussed in section 5.1.

As mentioned, an advantage of the cascaded architecture is that it suppresses the noise of the inner loop. This greatly relaxes the noise requirements for RA1, allowing it to be a simple low-power gm-R amplifier (Fig. 21). However, the noise of RA2 remains critical, because its noise is not shaped and dominates the overall ADC noise. To solve this problem, we design a gm-R amplifier with the Two-Phase Settling technique to suppress noise and improve efficiency. The details on the Two-Phase Settling technique will be discussed in section 5.2.

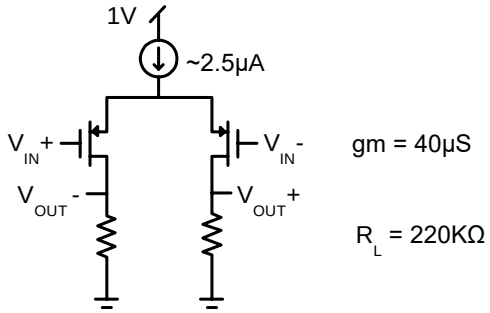


Fig. 21 Schematic of RA1

As shown in Fig. 22, RA2 is a gm-R structure with a push-pull structure to double the transconductance and improve efficiency. The CMFB resistors (R_L) are the dominant resistive load. Cascoding reduces the influence of channel-modulation on gain and linearity. The output resistance is configured by shorting the output series resistors (R_O). And thus, the main pole location can be changed during settling. The noise of R_O is suppressed by the gain of RA2 and is thus negligible.

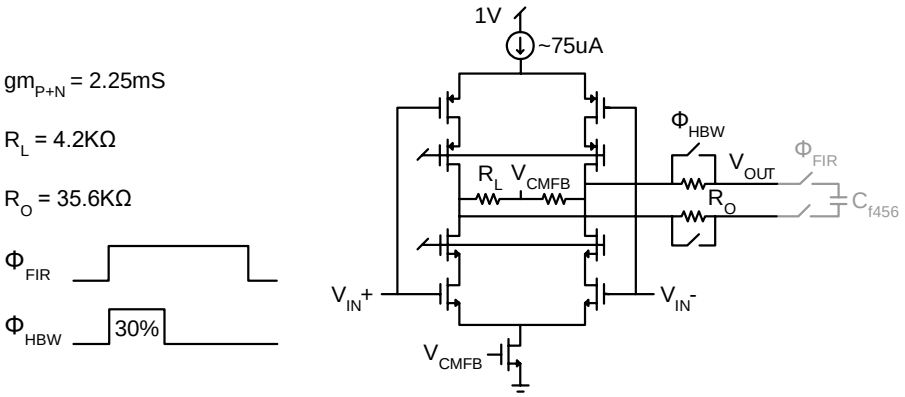


Fig. 22 Schematic of RA2.

Table 1 compares the simulated noise performance of RA2 for different settling configurations. The simulations are for the same RA2 design and differ only in R_O and the switch operation. In the conventional settling case, an external R_O is added to satisfy 6-τ settling. The simulations show that 2-phase settling reduces noise by 33%.

Table 1 Noise Performance of RA2

		Conventional	Two-Phase	
			P1	P2
gm		2.25 mS		
R_L		4.2 KΩ		
C_L		2.88 pF		
R_O		7 KΩ [+]	shorted	35.6 KΩ
T_S		420 ns	120 ns	300 ns
main pole (p) [++]		2.2 MHz	6.6 MHz	0.7 MHz
time constant (τ) [++]		71 ns	28 ns	230 ns
$v_{n,in}^2$ [+++]	gm noise only	$(15.0\ \mu V)^2$	$(10.2\ \mu V)^2$ **[54% reduction]**	
	all noise enabled	$(17.6\ \mu V)^2$	$(14.4\ \mu V)^2$ **[33% reduction]**	

[+] R_O is added to satisfy 6-τ settling

[++] from post-layout AC simulation

[+++] statistical result from post-layout transient simulation with noise (1000 cases)

C. Silicon Results

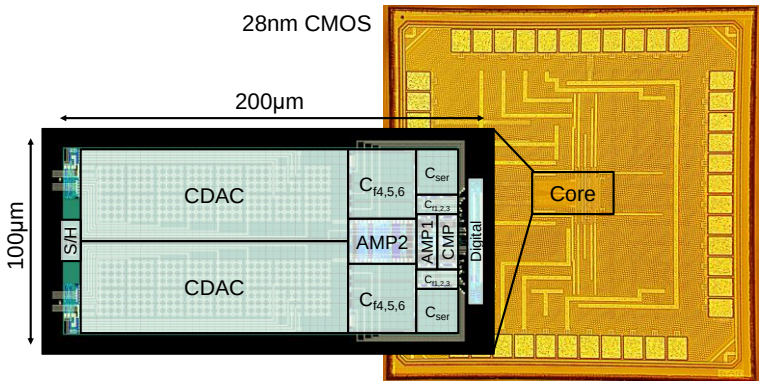


Fig. 23 Die photo.

The prototype design is fabricated in 28nm CMOS and measures 100µm by 200µm (Fig. 23). The majority of the chip area is occupied by the CDAC, while the noise-shaping FIR filters account for only 21% of the total area.

In single tone testing with 19.3kHz full-scale input signal (1.05Vp differential), the measured peak SNDR over a 100kHz bandwidth is 87.6dB. A 4th-order-shaped noise floor with an 80dB per decade slope is evident in the output spectrum, as shown in Fig. 24. Here, CDAC mismatch is canceled with an off-chip calibration method. Details about this calibration are described in Appendix B.

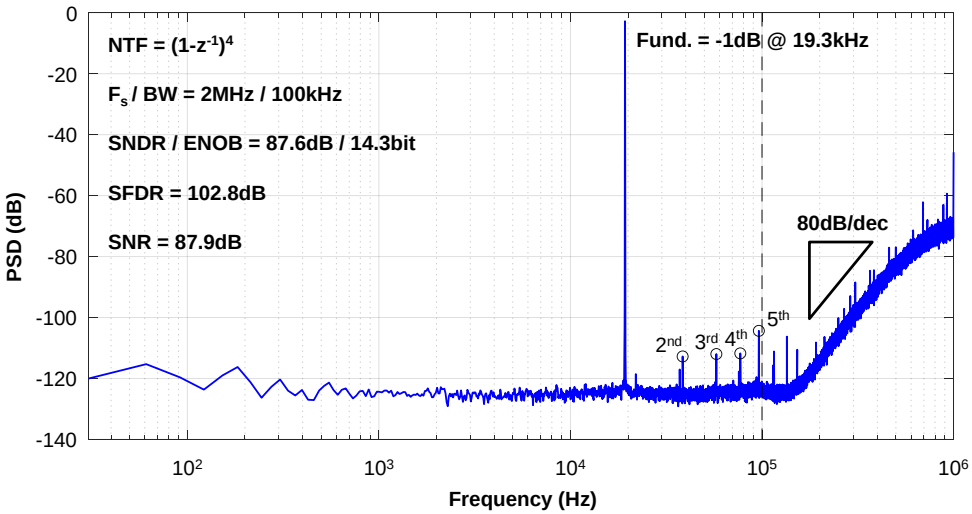


Fig. 24 FFT for single-tone test (64K samples Hann window, 16x averaging).

For comparison, we perform the same test with different configurations (Fig. 25). SNDR is improved by 25.3dB when the outer loop (FIR_2) is enabled. The inner loop (FIR_1) provides an additional 4.5 dB SNDR improvement. The mild SNDR improvement from the inner loop is due to a conservative design strategy that makes the shaped quantization error much smaller than the thermal noise. Therefore, there is a large margin for NTF variation so that quantization error does not overwhelm thermal noise. However, a more aggressive design strategy with a lower OSR and quantizer resolution can achieve the same SNR but with less NTF margin. The low-cost increased noise-shaping order provided by Nested NS SAR architecture gives the designer extra freedom and can either improve performance or robustness.

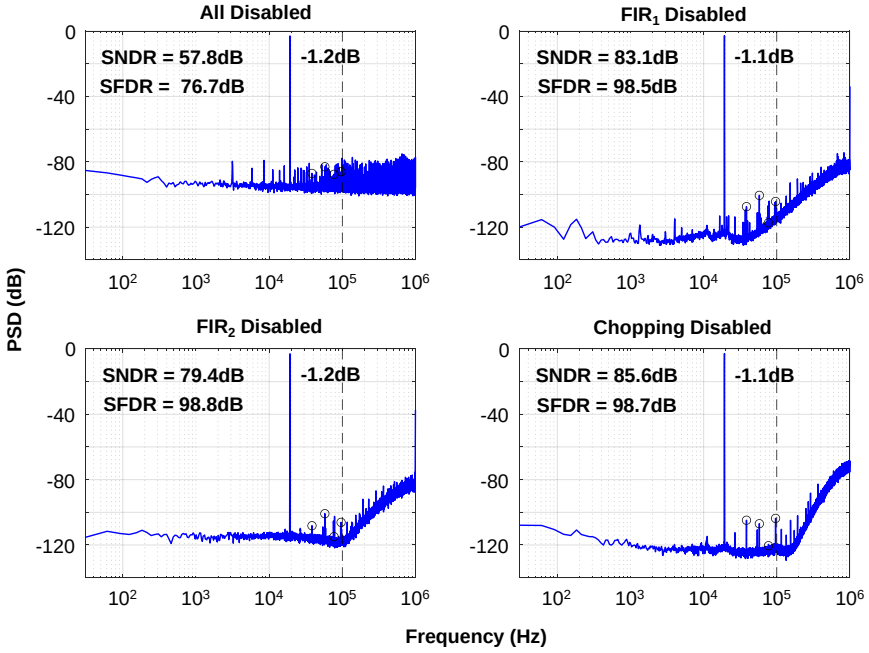


Fig. 25 Single-tone tests for different configurations.

Fig. 26 shows the measured spectrum for a two-tone test. The measured IMD3 is -83dB for a near-maximum input frequency. Fig. 27 shows the measured performance for different input amplitudes and frequencies. These measurements indicate consistent performance over the entire input frequency range. The measured dynamic range is 89dB.

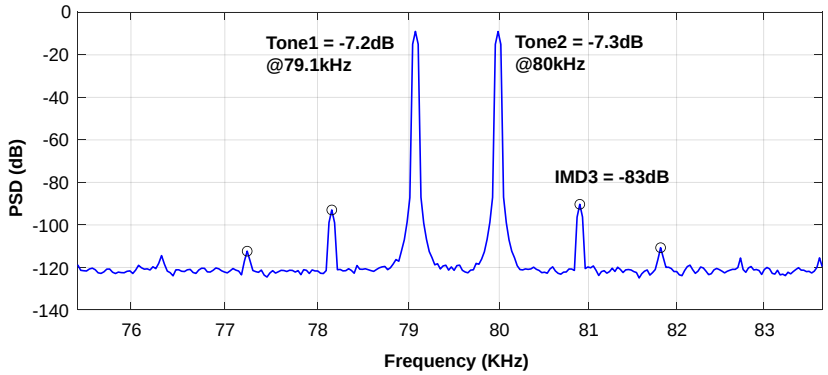


Fig. 26 Two-tone testing result.

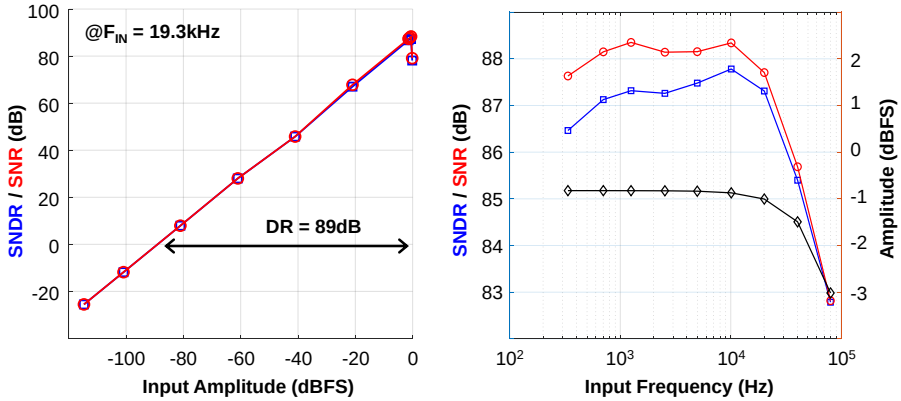


Fig. 27 Measured performance versus input amplitude and frequency.

Five different devices are measured to demonstrate the PVT robustness of the design. Without any PVT calibration, the measured SNR variation across the devices is only 2dB. As shown in Fig. 28, the measured variation in SNR is within 3dB over ±10% change in analog supply voltage and a 0-70ºC temperature range.

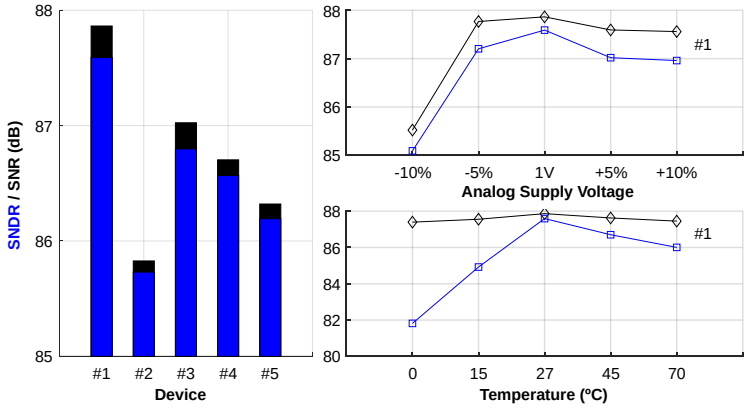


Fig. 28 Measured performance versus input amplitude and frequency.

Fig. 29 shows a breakdown of the measured power and area. The ADC consumes 120µW at 2MS/s. Only 2.2% of power and 6% of area is taken by the inner NS loop, which contributes <1% of the total noise. Therefore, the cost of the increased order is negligible.

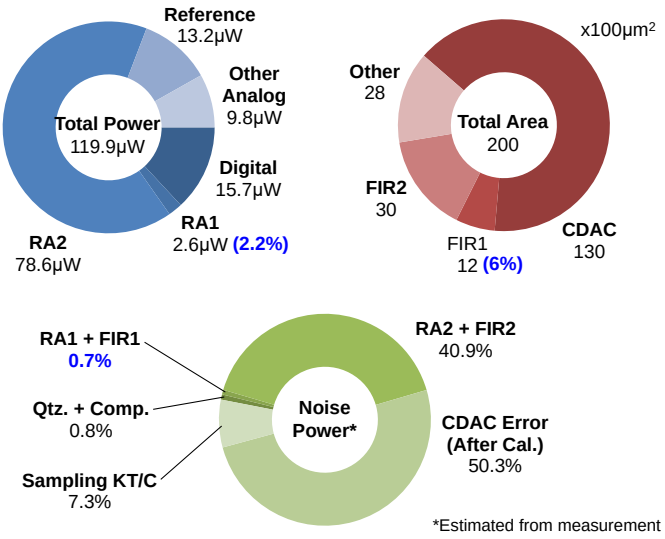


Fig. 29 Power, area and noise (estimated) breakdown.

Finally, Table 2 compares this design with some other state-of-the-art sub-MHz ADCs. The prototype exhibits higher order noise-shaping than conventional noise-shaping SAR ADCs. The overall efficiency, as measured with the Schreier FoM, is comparable to that of the most efficient conventional noise-shaping SAR ADCs, but our implementation is free of dynamic amplifiers and is PVT stable on system-level. Compared to Sigma-Delta ADCs, the prototype is more than 7x smaller in area, and thus far lower in cost. The prototype Nested NS SAR ADC demonstrates the highest NTF order and highest SNDR among NS SAR ADCs over 100kHz bandwidth to date (Feb. 2020), and provides a low-cost solution for high-resolution sub-MHz analog-to-digital conversion.

Table 2 Comparison Table

	This design	ISSCC 2020 X. Tang [22]	ISSCC 2020 J. Liu [23]	ISSCC 2018 S. Li [8]	VLSI 2017 W. Guo [24]	ISSCC 2018 P. Vogelmann [25]	VLSI 2018 C. Lee [26]
Architecture	**CaNS-SAR**	NS-SAR	NS-SAR	NS-SAR	NS-SAR	DT-SD	CT-SD
NTF Order	**4**	2	2	2	2	3	3
Amplifier	**2-Phase Settling**	Dynamic, Closed-Loop	Passive	Dynamic	Passive	Op-Amp	Op-Amp
Process (nm)	**28**	40	40	40	40	180	65
Area (mm^2) [+]	**0.02**	0.037	0.06	0.024	0.04	0.363	0.14
Supply (V)	**1**	1.1	1.1	1.1	1.1	3	1.2
Power (μW) [+]	**120**	107	67	84	143	1100	68
Fs (MHz)	**2**	10	2	10	8.4	30	6.14
OSR	**10**	8	25	8	16	150	128
BW (KHz)	**100**	625	40	625	263	100	24
SNDR (dB)	**87.6**[++]	83.8[++]	90.5	79[++]	80	86.6	94.1
SFDR (dB)	**102.8**[++]	94.3[++]	102.2	89[++]	-	101.3	107
DR (dB)	**89**	85.5	94.3	80.5	-	91.5	98.2
FoM_S (dB)	**176.8**	181.5	178.2	177.7	172.6	166.2	179.6

[+]Excludes power and area of CDAC calibration

[++]With CDAC mismatch calibration

Chapter 3 Time-Interleaved Noise-Shaping Converter

3.1 Basic Concept of Time-Interleaving

Time-Interleaving (TI) is a commonly used technique to speed-up the system. The key idea of TI is to have multiply identical copies of a system and run them in parallel. With proper input and output multiplexing, the overall interleaved system can achieve higher throughput. Fig. 30 shows a generalized block diagram of time-interleaved ADCs.

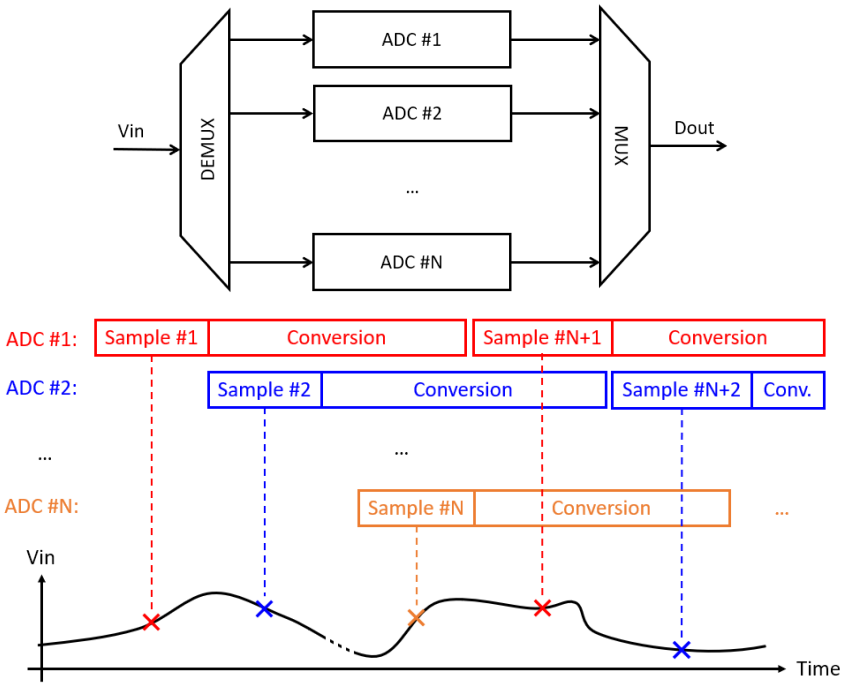


Fig. 30 Time-interleaved ADC (top) and timing diagram (bottom).

The overall sampling rate (i.e., throughput) of a time-interleaved ADC is given by:

$$F_{S,tot} = N \times F_{S,ch} \qquad \text{Eq. 16}$$

where $F_{S,tot}$ is the overall sampling rate of the interleaved ADC, N is the number of channels and $F_{S,ch}$ is the sampling rate of each sub-channel ADC. Therefore, TI is a direct trade-off between speed and hardware resources (power and cost, etc.). But it cannot improve the system's efficiency alone, that is, the throughput per channel keeps the same. Another thing to mention is that the multiplexer is the bottleneck in a TI system. As for TI ADC, the sampler has to run at full rate, and its performance determines the upper bound of the overall system's performance.

3.2 Interleaving of Noise-Shaping System

A. Direct Interleaving

Although TI is a possible way of mitigating the reduced bandwidth of NS ADCs, the combination of TI and NS is challenging in practice. The difficulty is not only in the circuit-level implementation, but is also fundamental at the system level. This system-level difficulty is related to the extra feedback delay in a TI system: When we attempt to interleave multiple ADCs containing feedback, the effective feedback delay changes. Thus, the overall system does not preserve the noise-shaping transfer function of the individual NS system. To better explain this, Fig. 31 shows a direct attempt at interleaving multiple conventional NS ADCs.

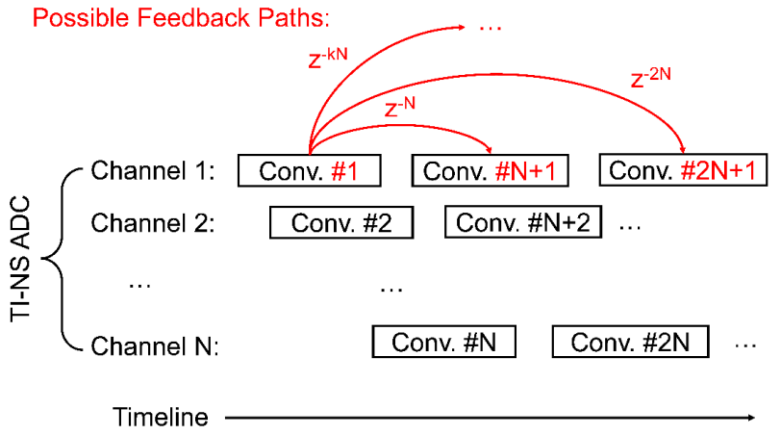


Fig. 31 Direct interleaving of NS ADCs

Feedback is within each NS ADC, however due to the interleaving, the actual feedback delay is a N times the interleaved-ADC sampling period. This means all unit delays, z^{-1}, in the individual NTF are replaced with z^{-N}. Therefore the equivalent NTF of this TI-NS ADC becomes:

$$NTF_{TI}(z) = NTF_{single}(z^N) \qquad \text{Eq. 17}$$

Fig. 32 shows the resulting NTF. Since the overall NTF is now in a repeating pattern, we can hardly synthesize the desired single-notch noise-shaping characteristic.

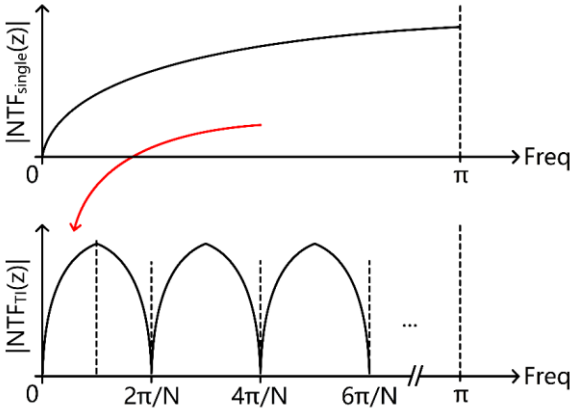


Fig. 32 Comparison of the original NTF for single NS SAR ADC (top) and the repeating pattern of the interleaved NTF (bottom).

B. Inter-channel Feedback and Causality Restrictions

Knowing the root cause of the problem, one may wonder if we can re-arrange the feedback path to have the correct delay. One obvious way is to feed the signal from one channel to the others, i.e., Inter-Channel Feedback, which seems to be a promising solution (Fig. 33). With inter-channel feedback, the effective feedback signal delay can be as small as a single sampling delay, so that the NTF can be similar to a classical noise-shaping ADC.

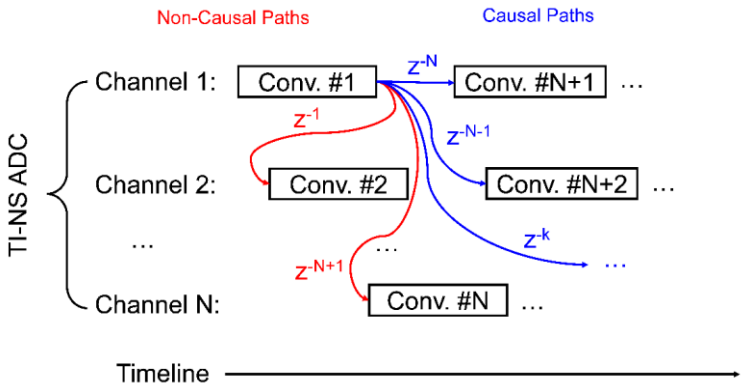


Fig. 33 Inter-channel feedback and the possible non-causal paths.

Nevertheless, the TI-NS ADC shown in Fig. 33 is not physically realizable, as some of the feedback signal paths are non-causal. Inter-channel feedback paths traveling earlier in time can never be implemented. If we attempt to make these inter-channel feedbacks causal, then each channel must finish conversion before the next channel starts, as shown in Fig. 34. However, imposing this requirement eliminates the benefit of interleaving as there is no longer an overlap between channels to improve conversion throughput.

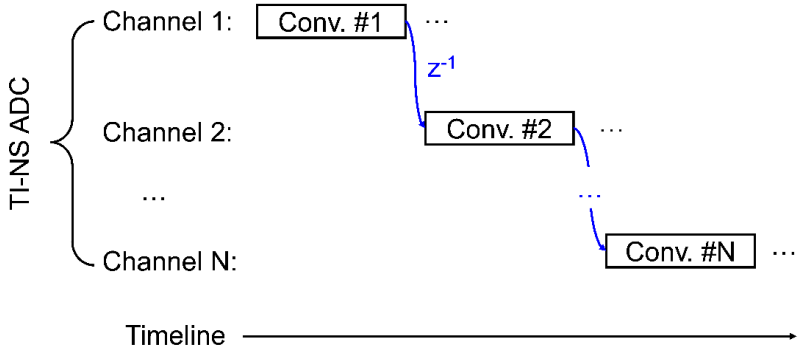


Fig. 34 This causal inter-channel feedback loses the throughput advantages of interleaving.

C. Midway Feedback and Multi-Phase Conversion

As discussed above, it is difficult to retain the benefits of interleaving and maintain flexibility in the transfer function. Although this may seem discouraging, we can still draw two useful conclusions from Fig. 33 and Fig. 34. First, for a feedback path with a longer delay, that is, traveling from a channel to another channel further separated in the interleaving sequence, the

causality restriction is relaxed. For example, if a TI-NS system does not have any feedback from a channel to an adjacent channel (i.e., there is no z^{-1} term in the transfer function), then there can still be some overlap in the conversions, as shown in Fig. 35. Therefore, if we can decompose the system into subsystems each with different delay, each subsystem can retain some benefits of interleaving (except for the subsystem with a delay of z^{-1}).

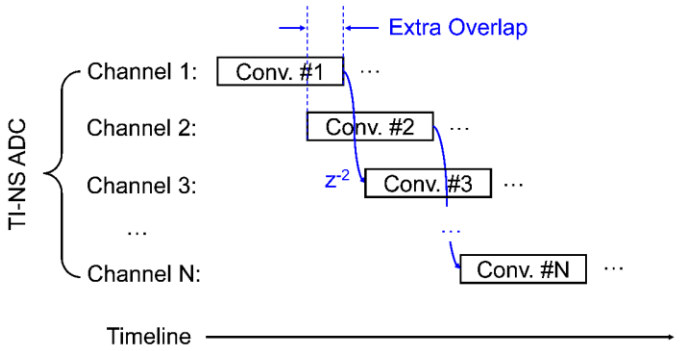


Fig. 35 Timing diagram for a time-interleaved noise-shaping ADC with no single interleaving period (z^{-1}) delay.

Second, the discussion above is based on the analysis of the feedback system. However, in a practical ADC system, many other actions introduce delay, including sampling, signal settling, and logic delay, etc. All these events are not restricted by the causality considerations that we mention, and therefore can be overlapped (i.e., interleaved) for higher ADC throughput.

Inspired by these observations, we introduce Midway Feedback and Multi-Phase Conversion to implement a realizable TI-NS system. Fig. 36 shows the concept of Midway Feedback with a timing diagram. In this configuration, the feedbacks get into the channels at the midway of their running. And to make this work, the conversion process for each channel is decomposed into multiple phases, so each phase performs only part of the conversion (i.e., Multi-Phase Conversion). Besides, the feedback path is also decomposed into multiple feedback paths with different delays, so that each sub-path feeds between different phases of different channels, enabling the maximum overlap.

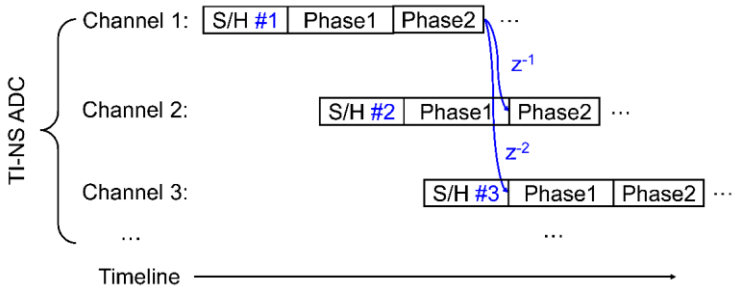


Fig. 36 Midway Feedback with Multi-Phase Conversion.

D. Non-Idealities of Interleaving

A significant drawback of time-interleaving is the degradation in accuracy due to the mismatch between channels. Typically, three kinds of mismatch dominate: offset mismatch, gain mismatch, and sampling skew. Offset mismatch causes input-independent tones at $\frac{F_S}{N}$ and its multiples, where F_S is the overall sampling rate and N is the number of channels. Gain mismatch and sampling skew cause modulation around $\frac{F_S}{N}$ and its multiples [27], as shown in Fig. 37.

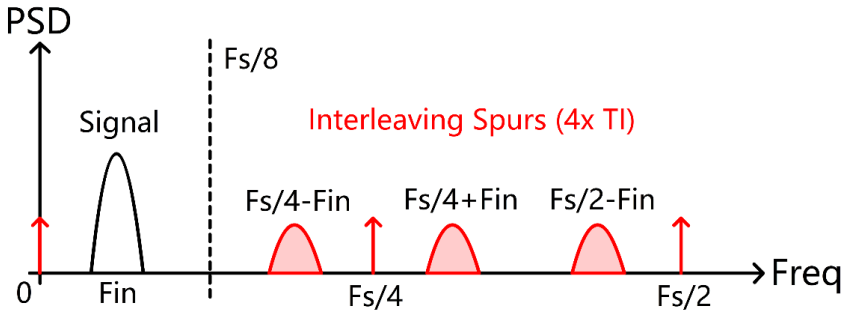


Fig. 37 Output spectrum due to channel mismatch (4x TI as example).

For interleaved Nyquist-rate ADCs, these artifacts fall in the band of interest, and therefore the performance of traditional TI ADCs is susceptible to mismatch. However, for a TI-NS ADC, the band of interest is reduced by the OSR. We notice in Fig. 37 that, as the artifacts are only located at frequencies around $\frac{F_S}{N}$, it is possible to limit the bandwidth of interest so that the interleaving artifacts all fall out-of-band. More specifically, assuming the signal band is from DC to BW, the lowest possible artifact is located at $\frac{F_S}{N} - BW$. Thus, if we limit $BW < \frac{F_S}{N} - BW$ or

$OSR = \frac{F_S}{2BW} > N$, then, the artifacts fall out-of-band. In this way, TI-NS ADC can suppress the mismatch problem naturally. This is a significant advantage because, generally, TI ADCs need complicated and power-hungry calibration to mitigate channel mismatch impacts. Although out-of-band blockers can still be mixed-down into the signal band, in this case, a pre-filter can be used to suppress the out-of-band blockers. An advantage is that this filter can be simpler than the anti-aliasing filter in a Nyquist rate ADC, as the modulation by channel mismatch is relatively weak.

3.3 A Design Example of Interleaved Noise-Shaping SAR ADC

In this example, we apply the Midway Feedback and Multi-Phase Conversion techniques in a TI-NS SAR ADC and demonstrate the advantages. The prototype was taped-out and measured, and the result is published in [28].

A. Architecture

As mentioned in section 3.2, a prerequisite for Midway Feedback is the decomposition of the original conversion into Multi-Phase Conversion. This might not be simple for some ADC architectures (e.g., Flash), but fortunately, it is easy for a SAR ADC. As we know, the SAR conversion naturally contains multiple cycles. Therefore, we can simply group these cycles into phases, as shown in Fig. 38.

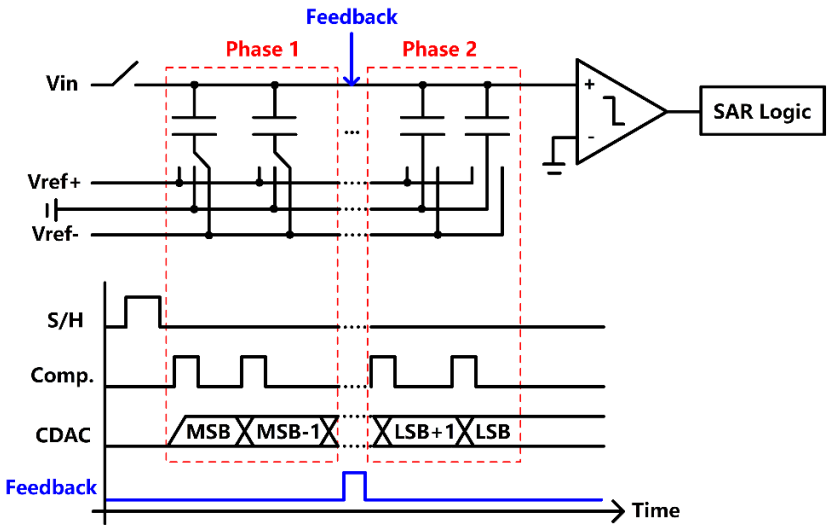


Fig. 38 Multi-phase division of SAR ADC (2-phase example).

Such a division is only in a logical sense, as each phase of the conversion is still performed with the same physical circuitry. Similar to a pipeline ADC with an inter-stage residue gain of 1, each phase partially digitizes the signal and passes a residue to the next phase. Although each conversion phase generates its independent quantization noise, this noise is digitized in subsequent phases and passed to the digital output. Eventually, the quantization noise from all phases, except the last, cancels at the output. Furthermore, as long as the ADC is not overloaded, any signal injected onto the CDAC, irrespective of the conversion phase, is all digitized and passed to the digital output, as shown in Fig. 39.

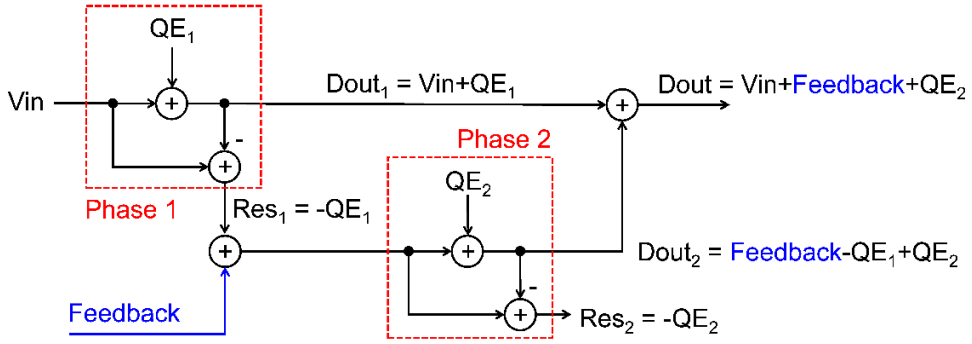


Fig. 39 Channel signal model with injected midway feedback signal.

The feedback summation position in this configuration naturally supports the Error-Feedback structure, which is promising as we already discussed the advantage of EF structure in section 1.4. Therefore, we implement the TI-NS SAR with Midway Error-Feedback in this design, as shown in Fig. 40.

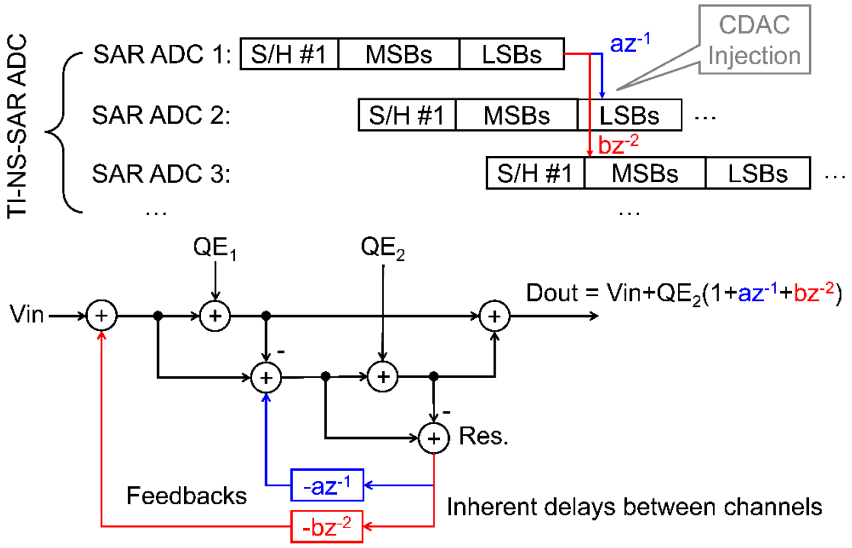


Fig. 40 Midway Error Feedback and its signal model (2nd order case as example).

The combination of CDAC injection and TI indicates an interesting implementation of the FIR filter. Recall that for Midway Feedback, we decompose the feedback path into multiple different feedbacks with different delays. Thus, we can simply sum different delayed paths with appropriate weights to form any desired FIR filter. In other words, we can make use of the inherent delay between different channels in a TI ADC to significantly simplify the overall architecture.

Although midway feedback in a TINS-SAR is elegant and straightforward, it is susceptible to overload, especially during the later quantization phases. We simply model each quantization phase as an ideal quantizer with purely additive quantization noise. However, this model fails when the input signal is larger than the quantizer's maximum quantization range. In a multi-phase SAR, since there is no gain between phases, the quantization range shrinks along with the successive-approximation steps, as Fig. 41 shows. Therefore the overloading condition, $QE + Feedback >$

$V_{in,max}$, can easily occur. In particular, for the final phase, the conversion range is as small as a few LSBs, so even a small injected feedback signal can cause overload.

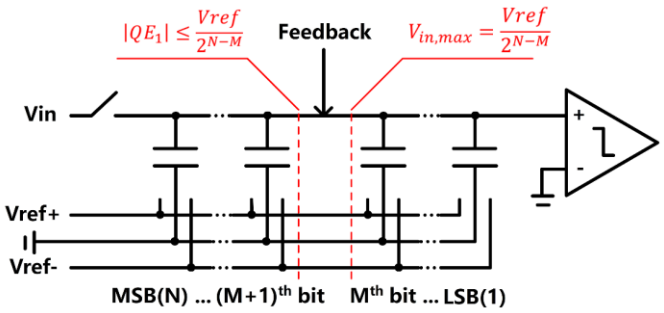


Fig. 41 Potential overload caused by feedback injection.

Once overload occurs, the quantizer's digital output can no longer be regarded as the summation of the input and noise, and the assumption of quantization noise cancellation no longer holds. As a result, the overall noise-shaping performance badly degrades, and the system can even become unstable. Worse, for a high-order FIR filter, the coefficients are large, which means the injected feedback signal is amplified, making overload even more likely.

We introduce two modifications to solve the overload problem. Firstly, we add redundant bits [29] to each quantization phase. The extra redundant decision bits provide an extended input range to tolerate errors made by previous decisions. This extended signal range also helps with the overload problem, as shown in Fig. 42. Although additional redundancy bits can further prevent overload, the tradeoff is reduced overall sampling rate. In practice, we add redundant bits mainly to the last phase as it has a more limited conversion range and is easier to overload.

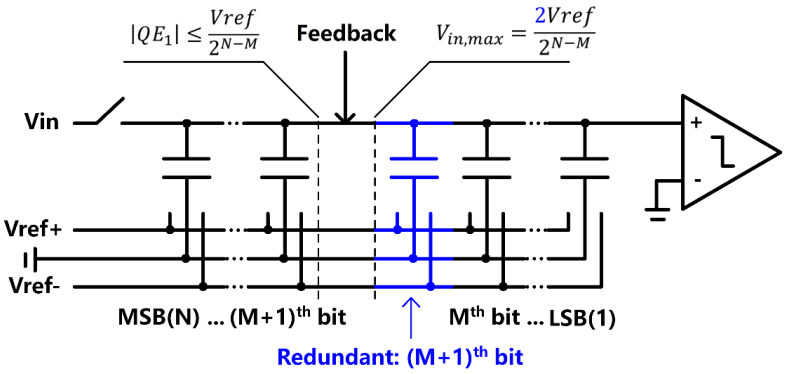


Fig. 42 A redundant bit enlarges the quantization range.

Secondly, we reduce the coefficients of the FIR filter to limit the amplitude of the feedback signals. As mentioned, the coefficients of a high-order FIR filter are generally large, especially for an aggressive NTF such as $(1-z^{-1})^N$. However, for NS ADCs with low OSR as 4x, a mild NTF with relatively small coefficients still delivers near-optimum SNR. Fig. 43 compares the performance for a conventional NTF and for a mild NTF to illustrate this idea. In this prototype, we use $(1-0.5z^{-1})^4$ since the 0.5 coefficient is easily implemented in layout with a ratio of 2. Another advantage is that a mild NTF is much more tolerant to coefficient variation. This is an important advantage as the poles' and zeros' positions always vary in real circuits due to mismatch and gain variations. An aggressive NTF degrades rapidly in the presence of small variation even without considering overload. A quantitative comparison of the NTFs is given in the next section.

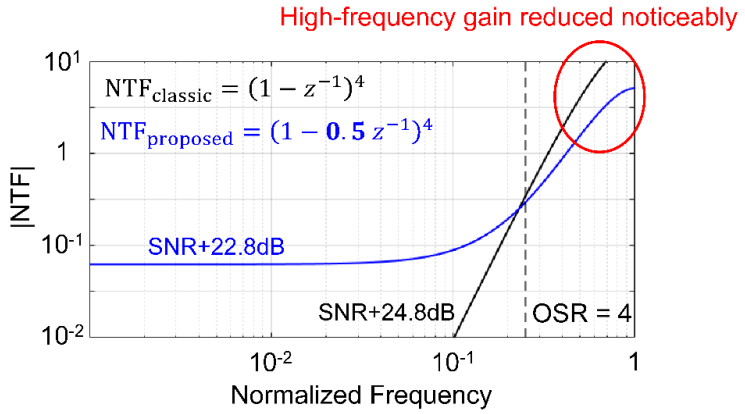


Fig. 43 Comparison of an aggressive NTF and a mild NTF.

Finally, the prototype TINS-SAR ADC targets a 400M/s sampling rate with four-way interleaving. It employs 4th order EF-based noise-shaping, as shown in Fig. 44. Each channel of the TINS-SAR performs 16 conversion cycles grouped in 4 phases. 6 bits of redundancy are added to the 10-bit binary digitization to eliminate overload based on model simulation results. As mentioned, the NTF is relaxed to $(1-0.5z^{-1})^4$ to help further prevent overload. With 4x OSR, this is enough to provide 70dB SNR over 50MHz bandwidth.

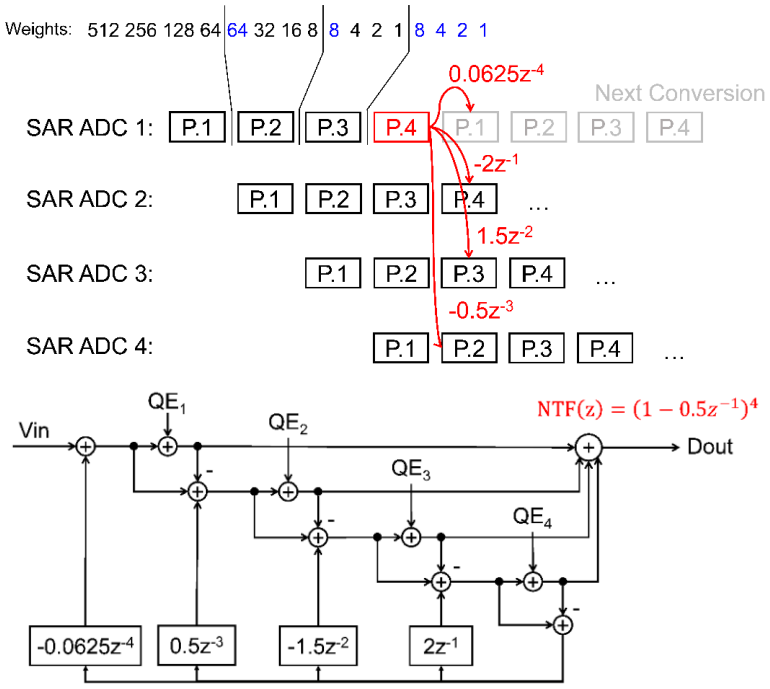


Fig. 44 The example design's timing diagram (top) and signal model (bottom).

B. Implementation

In this design example, we introduce a summing pre-amplifier to accomplish CDAC injection, as shown in Fig. 45.

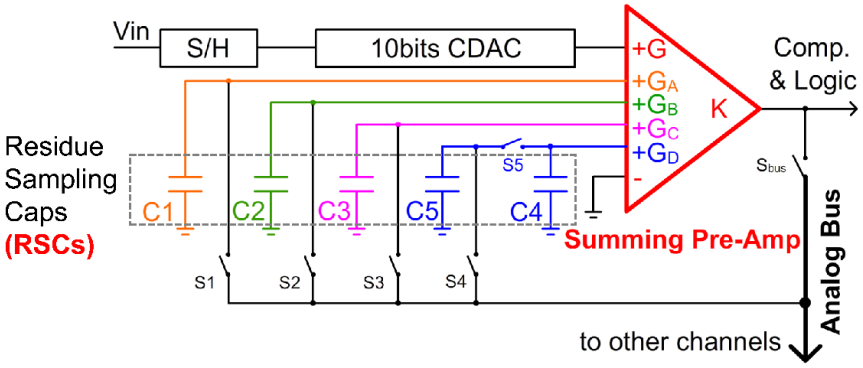


Fig. 45 Summing pre-amplifier implementation of EF.

In each channel, a summing pre-amplifier drives the comparator. The pre-amplifier output also provides feedback to all the interleaved channels through a single shared analog bus. The pre-amplifier is a multi-input low-gain differential amplifier, with its inputs connected to the CDAC and to four Residue Sampling Capacitors (RSCs, C1-C4 in Fig. 45), which hold the feedback signals from each channel. The pre-amplifier sums and weights the CDAC voltage as well as the stored feedback values. Thus, the output of the pre-amplifier is equivalent to CDAC output voltage after feedback injection. From another perspective, the pre-amplifier and the CDAC form a *Virtual CDAC* that realizes feedback injection (Fig. 46).

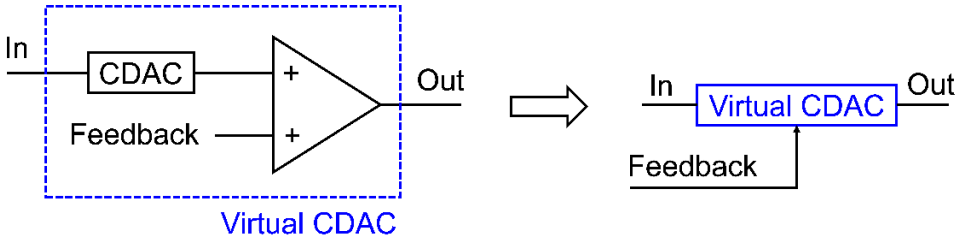


Fig. 46 The pre-amplifier and CDAC form a Virtual CDAC that realizes an equivalent feedback injection.

There are significant benefits to using a pre-amplifier. First, the charge on the CDAC is not contaminated during the feedback operation, and the cancelation of quantization error is therefore preserved. Second, the pre-amplifier provides good isolation between the comparator and both the CDAC and the RSCs (which hold the feedback voltages), thereby reducing comparator kickback

concerns. Third, the pre-amplifier can be realized as a simple single-stage open-loop amplifier due to the low required gain. A single-stage open-loop amplifier is smaller and more power-efficient than a high-gain amplifier, and is also easier to be implemented in modern CMOS processes. Furthermore, unlike a dynamic amplifier, such an amplifier does not require accurate timing.

In the pre-amplifier, multiple differential pairs with ratioed sizes drive a cross-coupled diode load, as shown in Fig. 47. The different input pair sizes implement the various feedback-gain coefficients (i.e., G, G_A – G_D), shown in Fig. 45. The amplifier provides the required gain (~10x considering all inputs) and a stable common-mode output voltage. Since the input of the amplifier (i.e., the residue) is only a few mV, such an open-loop design's linearity is sufficient for the target SNDR.

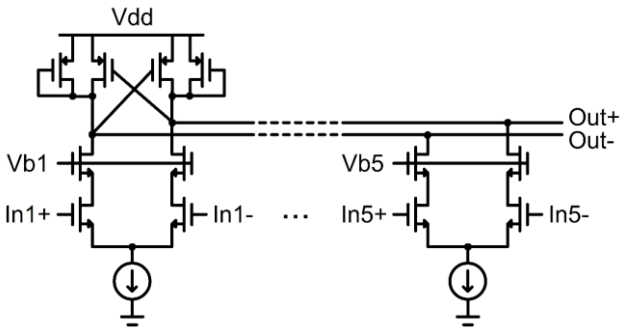


Fig. 47 Summing pre-amplifier schematic.

At any time, only one channel is completing a conversion and generating a residue to feedback. A shared bus (i.e., Analog Bus in Fig. 45) passes feedback between channels, keeping the wiring implementation simple. In addition to the inter-channel sampling capacitors (C1-C3), an auxiliary capacitor (C5) samples the residue of the channel itself. After C5 samples the pre-amplifier output, this sampled value is passed to the corresponding residue storage capacitor (C4) by simple charge-sharing. The attenuation due to charge-sharing is simply considered to be part of the feedback coefficient.

Fig. 48 shows the prototype's actual timing sequence, where phases 1 and 2 are approximately aligned to phases 3 and 4, respectively. Thus, some of the feedbacks are actually injected in earlier phases than is shown in Fig. 44. Note that Fig. 48 only illustrates the timing in the nominal case and is subject to possible skew, which will be addressed in the next section. The advantage of such a design is that more time can be assigned to the feedback phase, so that the bandwidth of the pre-amplifier can be smaller and the noise minimized. Similarly, the sampling phases are also lengthened to improve the linearity of the bootstrap switch.

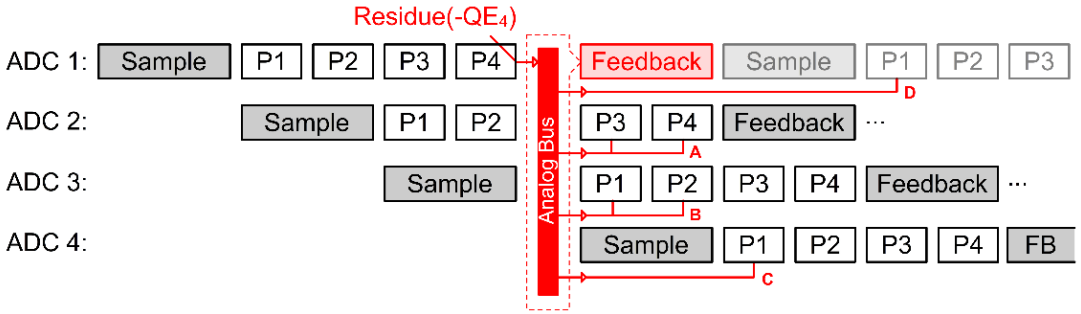


Fig. 48 Actual ADC timing sequence.

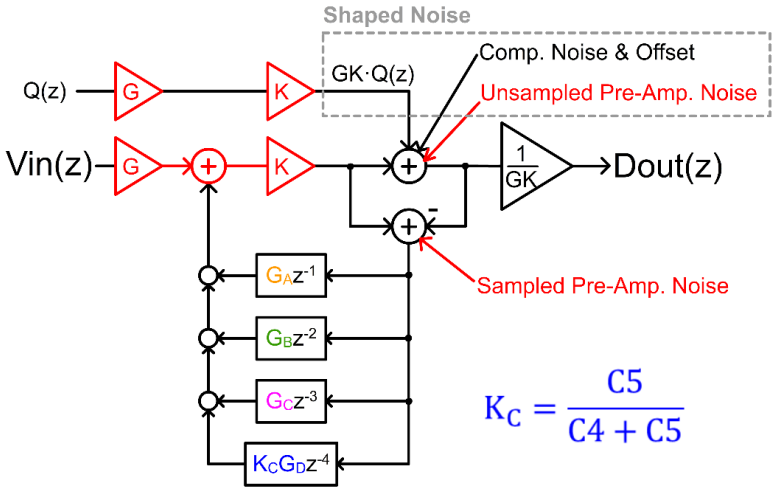
C. Addressing Non-Idealities

In this section, we cover a few possible non-idealities in this design, and suggest solutions. Some of the problems are also further discussed in section 4.3A, and corresponding solutions are provided.

i- Noise

The summing pre-amplifier's noise is a main source of the total noise, but fortunately, it is partially shaped by the NTF and can be well controlled. Fig. 49 shows a signal model that considers pre-amplifier noise. The noise of the pre-amplifier contributes differently to conversion and feedback. During the conversion phase, the pre-amplifier's noise (Unsampled Noise in Fig. 49) can be referred to the output and combined with the comparator input-referred noise. In this case, the

pre-amplifier and comparator's noise are shaped by the NTF and thus have a negligible effect on performance.



$$NTF(z) = 1 - K(G_A z^{-1} + G_B z^{-2} + G_C z^{-3} + K_C G_D z^{-4})$$

Fig. 49 Signal model considering noise and gain variation.

During the feedback phase, while the pre-amplifier is generating a residue, the pre-amplifier's noise is sampled onto the RSCs and added to the feedback summation node. Therefore, this noise contribution is determined by the Signal Transfer Function (STF) rather than NTF and is not shaped. Fortunately, during the feedback phase, the required signal-settling speed is much slower than during the conversion phase. Therefore the bandwidth of the pre-amplifier can be significantly reduced to limit thermal noise. Fortunately, the loading of RSCs can easily restrict the bandwidth, and therefore this part of pre-amplifier noise contribution is well controlled.

ii- PVT variation

PVT variation is another concern with open-loop amplifiers. Although the offset of the pre-amplifier does not affect performance, as it is similar to the comparator offset and only causes out-of-band mismatch tones, another concern is that the gain is inaccurate and sensitive to variation. However, since the pre-amplifier inputs are built with layout-matched differential pairs and drive

the same load, the gain ratio between the different inputs is quite precise and independent of PVT variation. The signal model in Fig. 49 decomposes the gain of the pre-amplifier into gain ratio and common gain to explain this advantage. G and G_A-G_D represent the nominal gains of different inputs, which are matched by layout techniques. K, nominally equal to 1, represents the common-gain variation from PVT. In modern CMOS processes, it is relatively easy to get 1% or better gain matching (i.e., G and G_A-G_D accuracy), while the common-gain variation, K, is roughly 10% or even higher. Interestingly, while our proposed NTF strongly depends on the gain ratios, it only weakly depends on the common gain. This advantage is clearly shown by the behavioral-level Monte Carlo simulations reported in Fig. 50. These simulations use a behavioral TINS-SAR model and vary the NTF coefficients and channel mismatch. With a 10% RMS gaussian common gain variation, the proposed NTF provides a robust SNDR improvement, while the conventional NTF suffers from considerable performance degradation and the risk of instability. Therefore, the performance of the pre-amplifier-based approach is robust enough under PVT variation to be free of calibration, even though the amplifier itself may have a large absolute gain variation. This feature also relieves the amplifier's settling requirements, as any settling error can simply be considered a common-gain reduction.

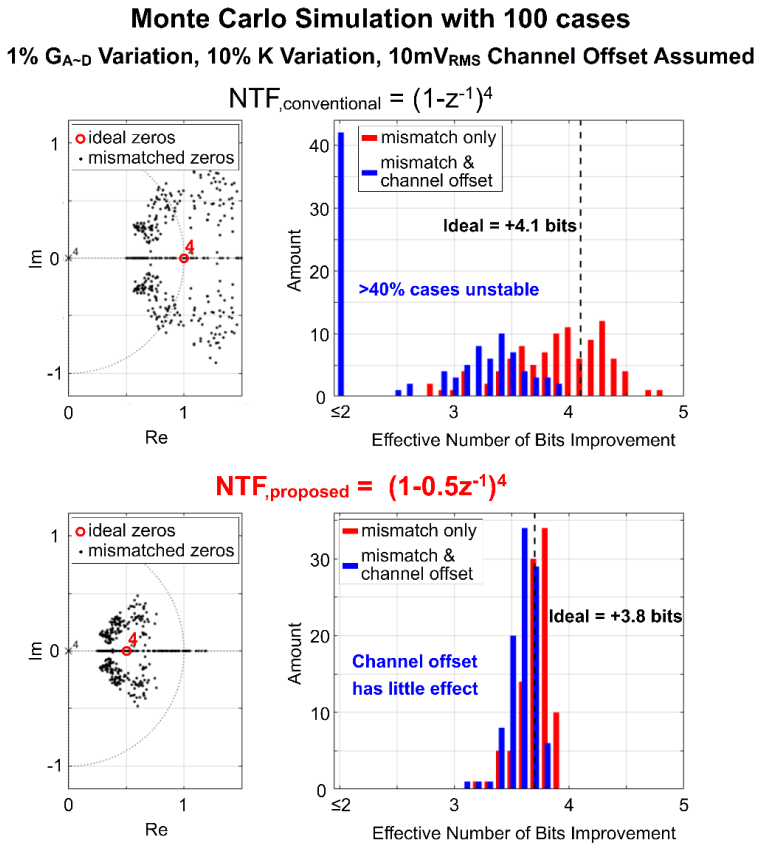


Fig. 50 Signal model considering noise and gain variation.

iii- Timing

A possible concern with asynchronous logic in the TINS-SAR architecture is the difficulty of aligning feedback timing between channels. Although in our analysis, we assume that the conversion phases of different channels are perfectly aligned in time, and that feedback happens right at the end of the last phase of each channel (Fig. 44), such strict timing is not necessary in practice. As mentioned, the division of phases is a grouping of the conversion cycles in a logical sense. Therefore, there is no need to have an equal division of phases as any grouping is valid for the analysis. Suppose conversion phases of different channels are misaligned due to mismatches in the asynchronous logic or the input-dependent delay of the comparator. In that case, the only change from the signal point of view is the magnitude of quantization error of each phase (exclude

the last phases), as shown in Fig. 51. Since the quantization errors of the former phases are eventually canceled at the output, such a timing misalignment does not affect the overall performance, as long as the timing skew is not large enough to trigger overload. In measurements, we did not observe any overload by timing issues.

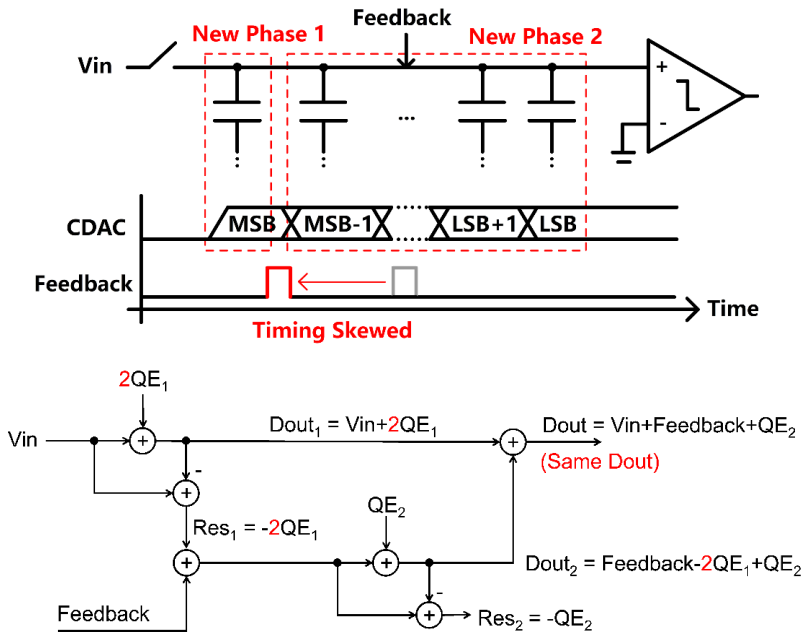


Fig. 51 Feedback timing skew has no effect on the system behavior.

iv- Linearity

Noise shaping effectively reduces the noise floor due to quantization error and comparator noise. Nevertheless, noise-shaping does not improve ADC's linearity, which can then limit performance in high-resolution applications. We apply three linearity enhancements, shown in Fig. 52, to help maintain the accuracy advantage of the TINS-SAR design. First, Dynamic Element Matching (DEM, [30]) shuffles the 4 MSBs of the CDAC, reducing distortion caused by CDAC mismatch. Second, the non-linear charge injection of the bootstrap switch is another source of distortion. To mitigate this, the CDAC switches connected to the bottom-plates are disconnected

shortly before the bootstrap switch opens (i.e. during Φ_{BPS}), which equivalently realizes Bottom-Plate Sampling (BPS) to prevent non-linear charge injection. Lastly, to deals with the sizeable non-linear capacitance of the pre-amplifier input (i.e. C_{par}, ~50fF), we apply the Parasitic Pre-Charging technique with two extra switches (S_{short} and $S_{connect}$). C_n is placed across $S_{connect}$ to limit its noise bandwidth. Details on the Parasitic Pre-Charging technique are discussed in section 5.1.

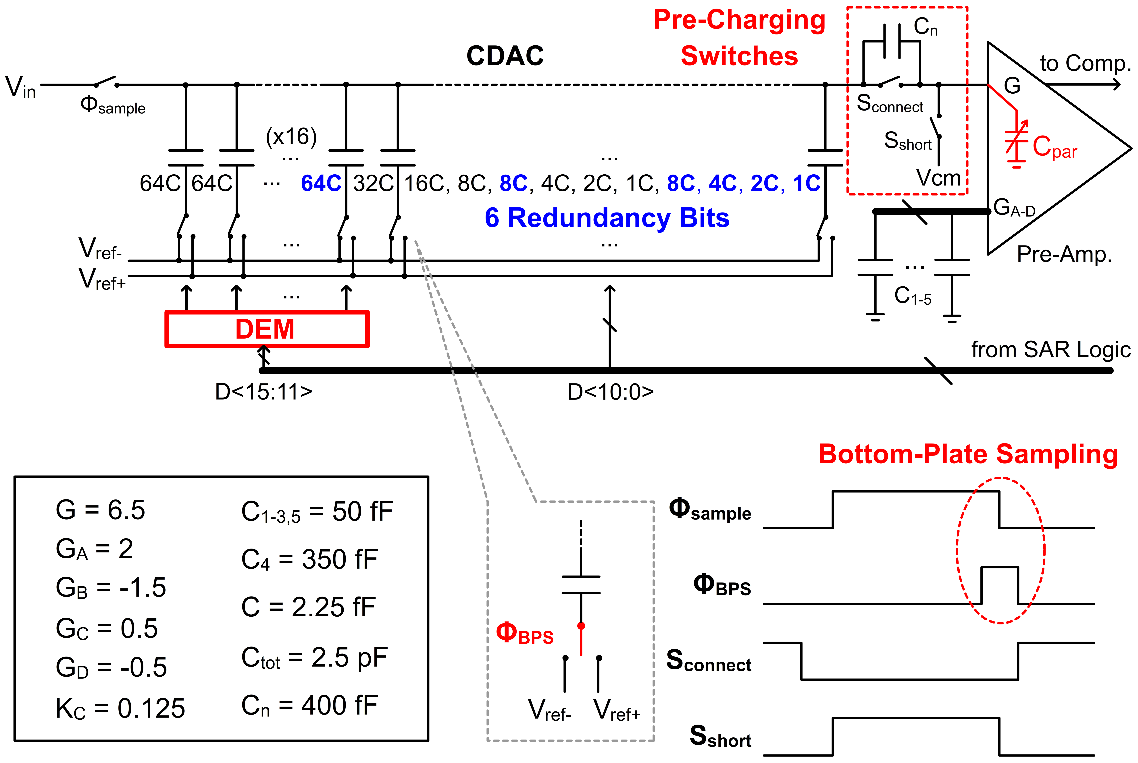


Fig. 52 DEM, BPS and Parasitic Pre-Charging technique enhance linearity.

D. Silicon Results

The prototype TINS-SAR ADC is fabricated in 40nm CMOS and has an active area of 0.06mm^2. Fig. 53 shows a die photo.

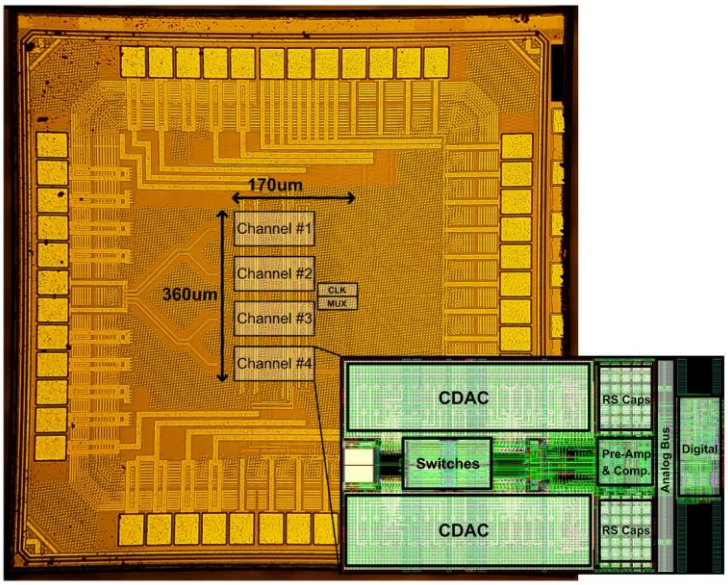


Fig. 53 Die photo and layout zoomed in.

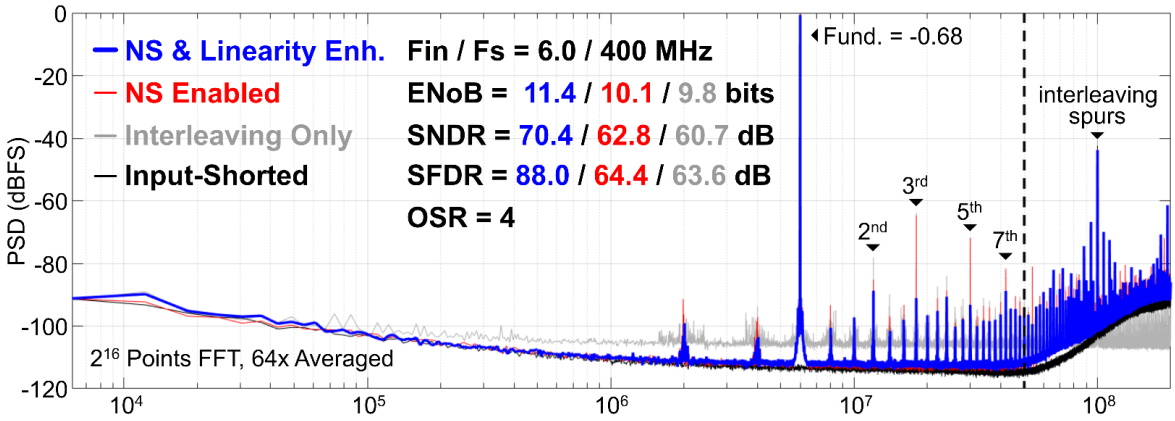


Fig. 54 Measured output spectrum for different configurations.

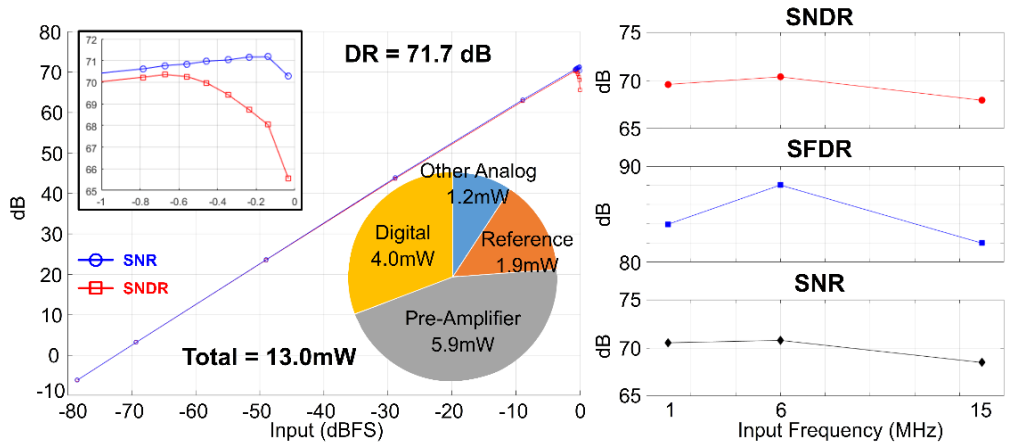


Fig. 55 Measured SNR/SNDR versus input power, power breakdown (left), and performance versus input frequency (right).

Fig. 54 shows the single-tone FFT at 400MS/s, indicating peak SNDR and SFDR of 70.4dB and 88.0dB, respectively. The rise in the noise floor at low frequency is mainly due to the pre-amplifier's flicker noise and has a negligible contribution to the overall ADC resolution as the total bandwidth is high. Fig. 54 also compares the performance when the noise-shaping and the linearity enhancements are enabled and disabled. The ADC consumes 13mW from a 1V supply while running at 400MS/s, where 1.9mW, 5.9mW, 1.2mW, and 4.0mW are dissipated by the reference, the pre-amplifier, the other analog circuitry, and the digital circuitry, respectively, resulting in a Schreier FoM of 166.3dB. The measured performance vs. input amplitude and input frequency are presented in Fig. 55, showing a dynamic range (DR) of 71.7dB.

To demonstrate the PVT robustness of the proposed design, we evaluate the performance of 10 different devices, without calibration. The results are shown in Fig. 56, where the average SNDR and SFDR exceed 69dB and 84dB, respectively. With a ±10% variation in the pre-amplifier supply voltage, the measured SNDR varies by less than 0.3dB. SNDR varies by less than 1.4dB for a 0-70°C variation in temperature.

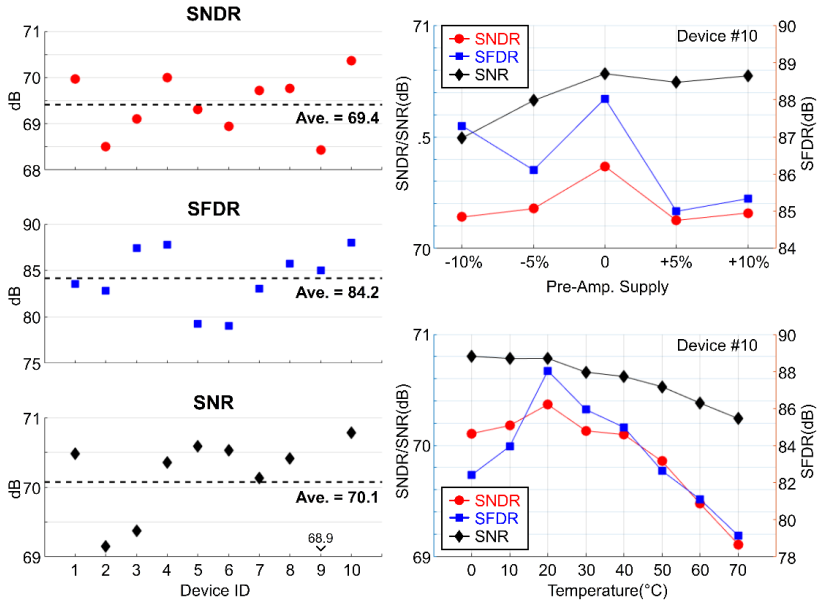


Fig. 56 Measured performance of 10 devices (left), and measured performance under supply voltage, and temperature variation (right).

Table 3 provides a performance summary and compares with state-of-the-art, high-bandwidth NS ADCs, highlighting the advantages of bandwidth, high NTF order, and the accuracy of the proposed TINS-SAR architecture compared to conventional NS-SARs. This prototype design attains the highest bandwidth among NS-SAR ADCs and approaches the performance of state-of-the-art CT-SD ADCs.

Table 3 Comparison Table

	This design	ISSCC'18 S. Li [8]	VLSI'17 Y. Lin [31]	ISSCC'12 J. Fred. [7]	ISSCC'18 T. He [32]	VLSI'16 A. Jain [33]
Architecture	TINS-SAR	NS-SAR	NS-SAR	NS-SAR	CT-SD	CT-SD
Calibration Free	√	X	√	√	X	X
Technology (nm)	40	40	14	65	28	65
Area (mm^2)	0.061	0.024	0.0043	0.0462	0.25	0.07
Supply Voltage (V)	1	1.1	1	1.2	1.16/1.5	1.4
Power (mW)	13.0	0.084	2.4	0.8	64.3	13.3
Sampling Rate (Ms)	400	10	300	90	2000	6000
OSR	4	8	6	4	20	50
Bandwidth (MHz)	50	0.625	25	11	50	60
NTF Order	4	2	1	1	4	4
SNDR (dB)	70.4	79	69.1	62.1	79.8	67.6
SFDR (dB)	88.0	89	78	72.5	95.2	77.4
DR (dB)	71.7	80.5	72	-	82.8	76
FoM_S (dB)	166.3	178	169.3	163.3	168.7	164.1
FoM_W (fJ/c-step)	48.1	9	20.6	35.8	80.5	56.5

Chapter 4 CT-DT Hybrid Noise-Shaping Converter

4.1 Continuous-Time Noise-Shaping System: Advantages and Problems

The NS systems we discussed so far are all Discrete-Time (DT). A DT NS system is straight forward in design as it can be easily analyzed with a DT signal model. However, DT NS converters have some inherent drawbacks, especially when they come to high-bandwidth applications. The first difficulty is the implementation of the DT loop filter. In most cases, the DT filters are realized with SC circuits, in which high-speed amplifiers are needed to settle the capacitor voltage to a specific accuracy. This means that the bandwidth of those amplifiers needs to be at least a few times higher than the sampling rate. Recall that we are usually doing oversampling in a NS system, which means that the bandwidth of the amplifier should be even tens of times higher than the required signal bandwidth in practice! Therefore, high bandwidth DT NS ADCs are quite inefficient, if not even practical.

Besides, the input sampling of DT NS ADC can also be problematic at a high sampling rate. Usually, the sampling is done with a capacitor storing the input voltage. But this implies that the input source has to charge the sampling capacitor at enough accuracy within one sampling period. Such high-speed, high-accuracy charging or discharging of a capacitor generally requires a power-hungry driver, and the power of the can be even higher than the ADC itself!

Continuous-Time (CT) NS is an excellent solution to these two problems. Fig. 57 shows the signal model of a typical CT NS ADC. The main difference between it to the DT counterpart is that its loop filter is a continuous-time one, and sampling happens in the quantizer instead of at

the input. The feedback DAC converts the DT signal back to CT with some specific DAC waveform.

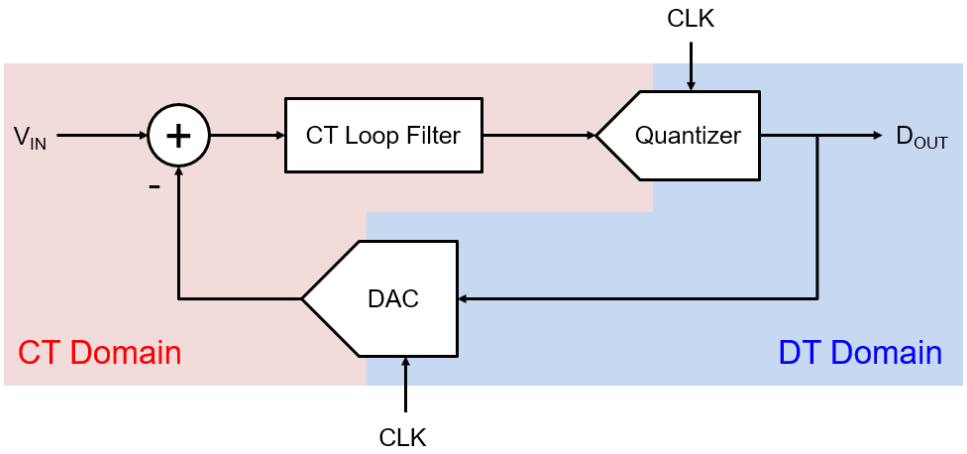


Fig. 57 Signal model of a common CT NS ADC

A general method of analyzing and designing a CT NS system is by mapping the CT system to a DT equivalent, or vice-versa. Fig. 58 illustrates a CT loop and a DT NS system, respectively.

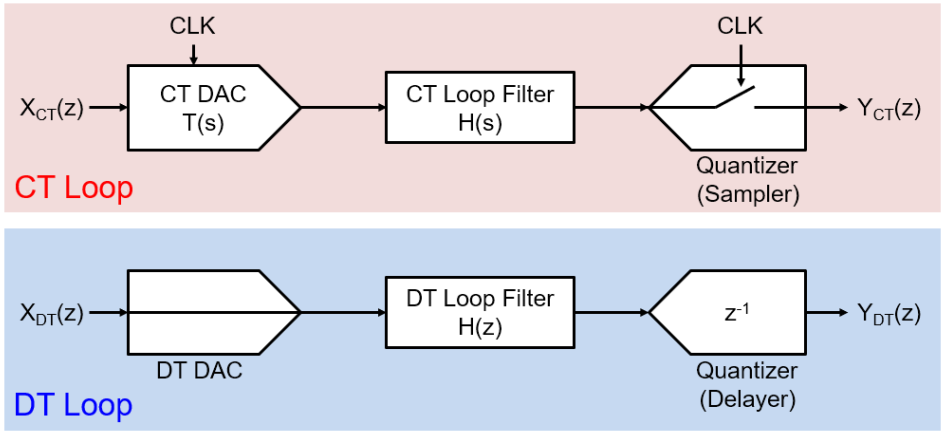


Fig. 58 CT Loop and its equivalent DT Loop

We say the CT loop is equivalent to the DT one if:

$$H_{L,CT}(z) = \frac{Y_{CT}(z)}{X_{CT}(z)} = H_{L,DT}(z) = \frac{Y_{DT}(z)}{X_{DT}(z)} \quad \text{Eq. 18}$$

In practice, designers usually first work on a DT loop design as the analysis is easier. After that, the DT loop can be mapped to a CT version with a carefully selected DAC waveform (i.e., T(s)). There are a few methods to realize such a mapping [34], and the commonly used one is the

impulse invariance method. There are also many software/toolboxes that can perform the mapping, such as Matlab.

CT NS well solves two problems of DT NS. Firstly, the bandwidth of the CT loop filter is the same as the required signal bandwidth. This greatly reduces the speed requirement of the amplifiers in the loop filter. The reduction in amplifier bandwidth also effectively reduces the total induced active noise, which is, in a sense, saving power.

Secondly, as the sampler is moved past the loop filter and is not presented at the input port, the driving load on the input source is also much relaxed. In practice, the loop filter is usually implemented with "active RC" structures, which provide a pure and linear resistive impedance at the input. Such a resistive impedance is generally easy to be drive, and can even provide impedance matching in some RF applications.

Lastly, as the input signal has to get through the loop filter first before it is sampled, the loop filter naturally provides an anti-aliasing function, and usually, no extra anti-aliasing filter is required.

Regardless of the advantages mentioned above, there are also some tricky problems induced by the CT structure: 1) Unlike a DT filter built with SC circuits, the CT filters' coefficients are usually dependent on RC time-constant, whose PVT stability is generally poor. Therefore for most CT NS ADC, designers have to fine-tune the RC time-constant by some digital configuration methods (e.g., switching capacitor banks). Besides, 2) the DAC in a CT NS system is very sensitive to timing variation (i.e., jitter and skew). This is because any timing changes can induce Inter-Symbol-Interference (ISI) to the DAC output, which can be regarded as a source of error added to the ADC's input. Although some methods exist to alleviate the ISI, such as using an RZ waveform or an FIR DAC, the timing sensitivity is still inevitably higher than for the DT counterpart.

Moreover, 3) there are some delays in the CT loop due to quantization and signal transmission, which are not included in the model of Fig. 57. These delays, named Excessive Loop Delay (ELD), change the loop-transfer function, induce extra poles, and often cause unwanted peaking in the STF and NTF [35]. To compensate the ELD, extra feedback or feedforward paths are needed, increasing the system's cost. If the ELD changes due to PVT, the compensation degrades too. Lastly, 4) the CT loop filter is designed for a constant sampling rate, which means a CT NS converter cannot adapt to different clock rates without changing the loop filter. This prohibits CT NS converters from being used in some variable sampling rate applications.

4.2 CT-DT Hybrid Noise-Shaping System

As mentioned, there are certainly many existing solutions to the drawbacks of CT NS converters. But in this section, we are going to solve those problems more fundamentally by introducing a new NS architecture: CT-DT Hybrid Noise-Shaping.

A. Basic Concepts

To introduce the new architecture, we first redraw the CT NS converter's signal model, as shown in Fig. 59.

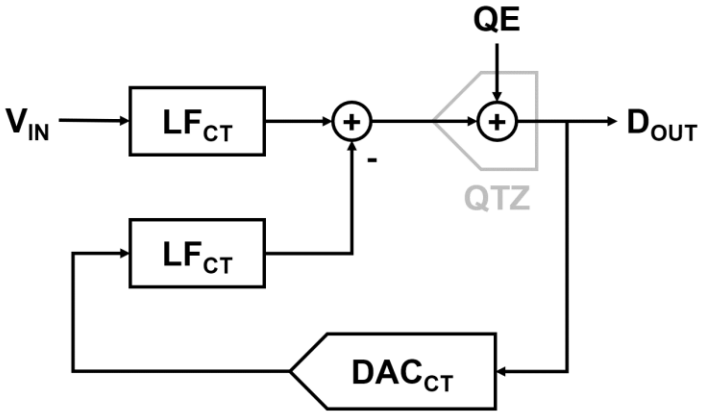


Fig. 59 A transformed signal model of CT NS converter.

In this transformed model, the loop filter is divided into two: One in the forward and the other in the feedback loop. Notice that most of the CT NS advantages, such as the anti-aliasing and easy-driven feature, come from the forward filter. On the other hand, we also notice that the disadvantages of the CT NS all come from the CT loop filter. Although the CT loop filter is relaxed in bandwidth requirements, the CT nature directly make it Fs incompatible and jitter sensitive. Therefore, we first replace the CT loop filter with a DT one, as shown in Fig. 60.

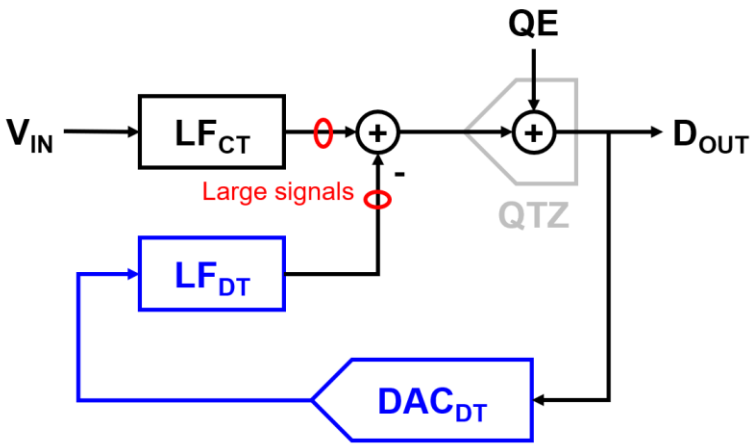


Fig. 60 Replacing the CT loop filter with a DT counterpart.

If a DT filter, together with a DT DAC, is used, all the disadvantages of CT NS converters are eliminated. Although this method looks promising, it is actually problematic. Recall that this is just an equivalent model. If we implement such a model, the signal swing before the adder will be huge due to the high gain in the loop filter and may saturate the circuit. Besides, the amplifier's excessive bandwidth requirement inside the DT filter comes back as a challenge again.

Fortunately, there is a simple solution to this. We can extract the gain from the two filters and move it to after the adder, as shown in Fig. 61. The DC gain of the filters becomes unity, and the signal swing at the filter output can be safely within the headroom. Besides, the amplifier in the gain block can be a continuous-time one, and does not need to settle fast.

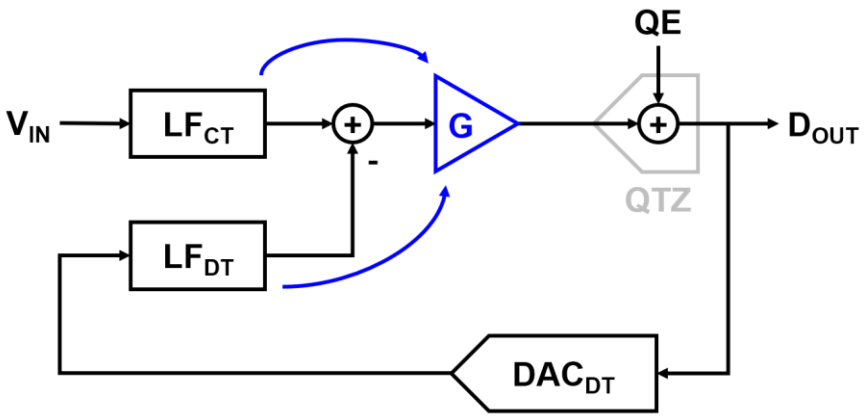


Fig. 61 Relocate the gain of the filters.

However, this modification leads to another problem: the noise of the amplifier. Notice that the amplifier's noise is directly quantized without passing through the loop filter, so there will be "excessive" noise from the amplifier that gets aliased in-band and undermines the SNR.

To solve this, we can further insert a secondary CT loop filter behind the gain block, so that the noise from the amplifier is restricted. To keep dominance of the DT loop filter, the poles of this secondary filter have to be away from the DT filter ones. Since we only care about the in-band noise, even if the amplifier's noise bandwidth is higher than the loop bandwidth, it will not be problematic as long as the noise does not alias back in-band.

Fig. 62 shows the final block diagram of the new CT-DT hybrid NS architecture, which combines the advantages of both CT and DT DSM: The entire forward-path is kept in CT domain, thus it still provides anti-aliasing and easy-driving features; The amplifier is CT and does not need to settle fast; Noise is not a concern with the help of the secondary CT filter (LF_{CT2}). The dominant pole (in LF_{SC}) and feedback DAC (CDAC) are SC circuits. Therefore they do not need tuning and are insensitive to timing. The CDAC provides better matching than a current DAC as well.

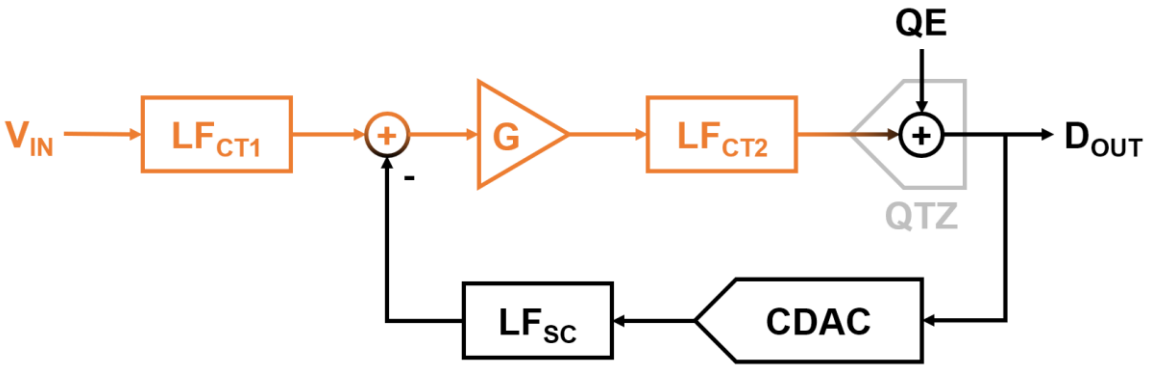


Fig. 62 Proposed CT-DT hybrid NS converter.

B. System Trade-offs

Although the extra noise-limiting CT loop filter is subdominant, it inevitably changes the loop characteristic. Therefore, the loop design of this architecture requires careful optimization in practice. Intuitively, the DT filter's bandwidth has to be lower than the CT filter's to keep dominance. But if the CT filter's bandwidth is set too high, there will be too much amplifier noise and lower performance. So there is a trade-off between robustness and noise efficiency, as shown in Fig. 63.

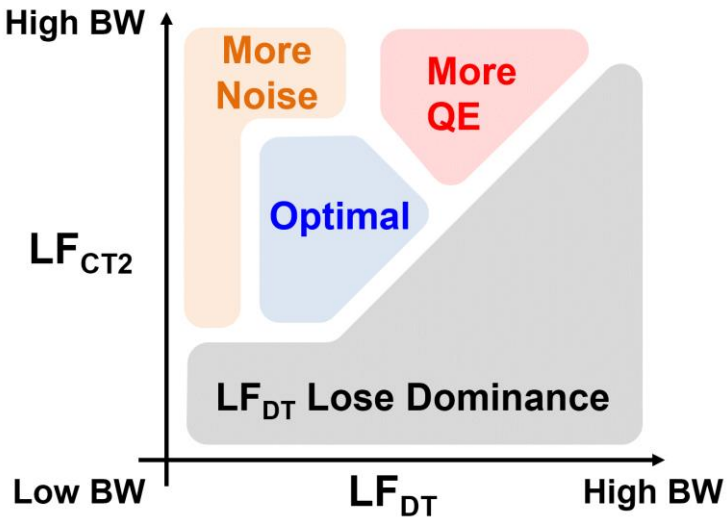


Fig. 63 Effects on the CT and DT loop filter's bandwidth.

To better understand this, we model an example in which both the DT and CT filter are 1st order. The two contour plots in Fig. 64 show the relative noise strength from amplifier and quantization. In this case, modeling shows that the two filters' optimal pole locations are round 0.75 and 0.45, respectively. Certainly, the conclusion will be different for different NS configurations in practice, but essentially we aim to find the region where the noise is small, and the CT filter variation is tolerable.

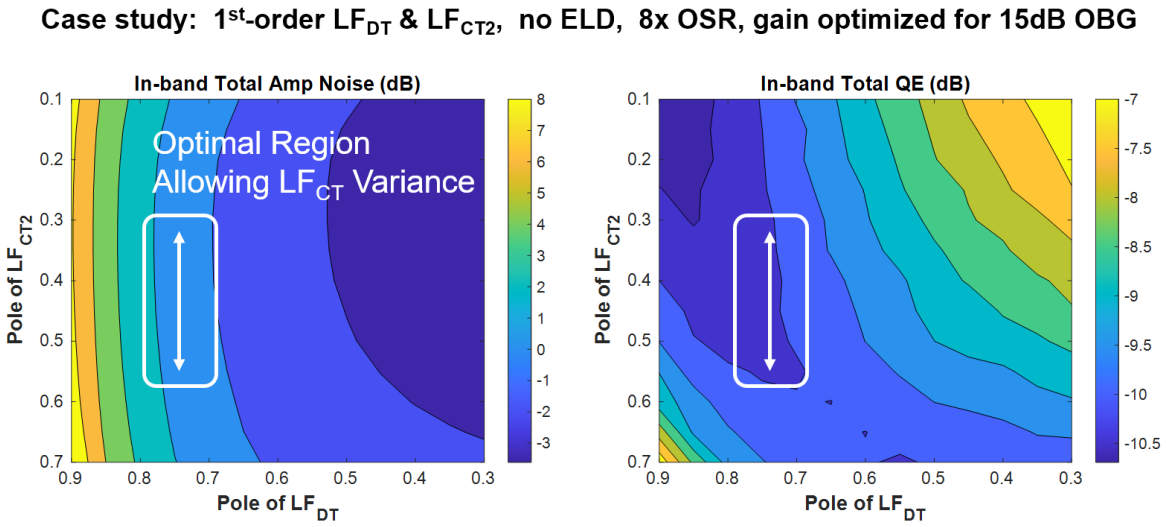


Fig. 64 An example of loop design trade-off.

C. Concerns for High Order Noise-Shaping

Another underlying concern of this architecture is its extensibility to high order. The difficulties are from a few aspects: Firstly, it is hard to implement a high-order DT loop filter without an amplifier (i.e., passive) - Recall that, as we hope to place the amplifier in CT domain, the dominant DT loop filter should not contain any amplifier. Secondly, the amplifier noise will is more difficult to control with a high order loop. This is because the bandwidth of the noise limiting CT filter has to be high enough to avoid affecting the dominant poles, but obviously, this will let more amplifier noise pass. Lastly, the ELD is not negligible in a high-order loop and has to be

compensated. The compensation complicates the loop design further. Fig. 65 highlights these difficulties.

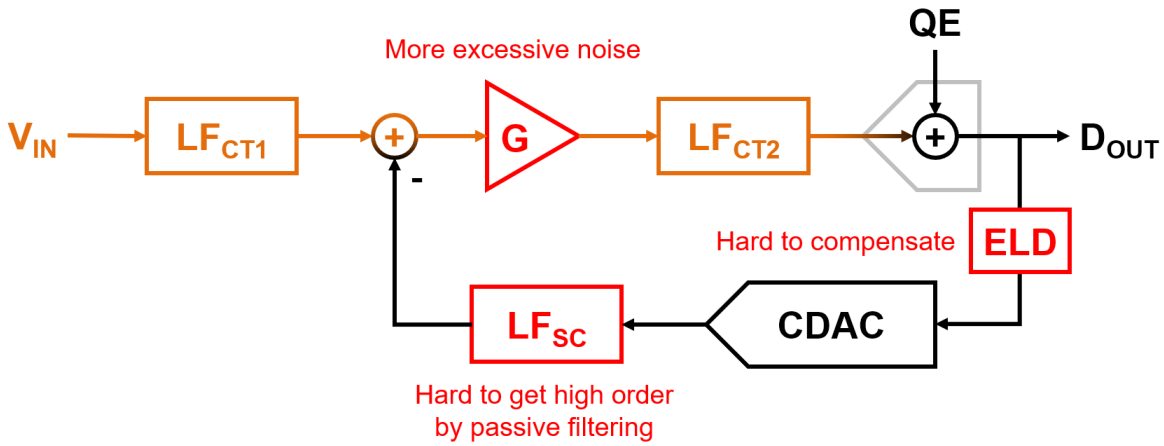


Fig. 65 Potential problems in high order hybrid NS system.

But fortunately, the order limitation of this architecture is not necessarily problematic, as there is another way to achieve high order NS: using Nested Structure (Chapter 2). Applying the Nested Structure is a very promising strategy for the hybrid NS converter because it increases the system order and decouples the noise transfer function. The latter improves the robustness of the system as well. As for the quantizer's architecture, both the conventional noise coupling structure and VCO based quantizer are feasible choices, but they are limited in order and resolution. On the other hand, the emerging noise-shaping SAR is an even better candidate, as they can provide both high-order and high resolution, and in our case, we are looking for a high-order quantizer to increase the system order. The speed limitation of NS SAR can be relaxed by the Time-Interleaving technique (Chapter 3). In the following section, we introduce an improved TINS SAR design to further resolve NS SAR's speed limitation.

4.3 A Design Example of Hybrid Noise-Shaping Converter

In this example, we design a high dynamic-range, high bandwidth converter using the CT-DT Hybrid NS architecture and discuss its advantages. An improved version of the TINS SAR converter is also introduced as the nested quantizer. The prototype was taped-out and measured, and the results were published on [36].

A. Improved TINS SAR Quantizer

As mentioned in section 4.2, it is a good strategy to implement the hybrid NS converter with the nested structure, i.e., using a noise-shaped quantizer to increase the system order. NS SAR is a promising architecture choice for this quantizer as it can provide high-order. Since the hybrid NS converter aims for high bandwidth, we hope that the quantizer can run at a high sampling rate. The time-interleaving technique mentioned in Chapter 3 is an excellent tool, but the 400M/s sampling rate of the example TINS SAR converter is still a little low. (Many CT SD ADCs are running at Giga samples per second) Thus in this section, we first push the TINS SAR architecture for even higher speed.

For discussion convenience, we redraw the timing diagram of the TINS SAR in Fig. 66.

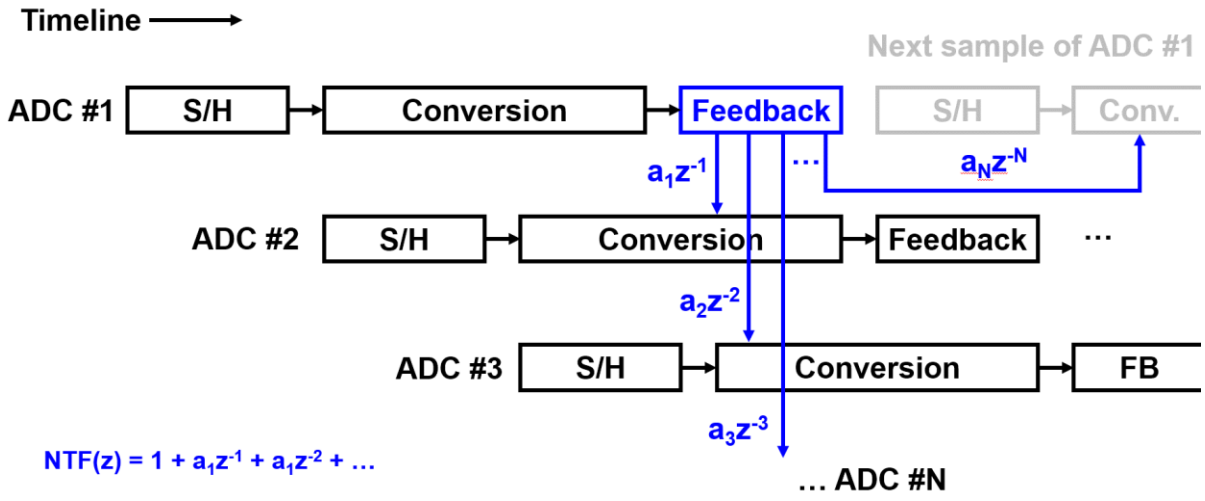


Fig. 66 Timing diagram of the TINS SAR in section 3.3.

In this timing design, the critical path is the z^{-1} path that feeds the residue from one channel to the next, as marked in red in Fig. 67. To convert this feed-in signal without overloading, at least 3~4 SAR cycles are needed. This prevents the further overlapping of the adjacent channels, and thus limits the sampling rate.

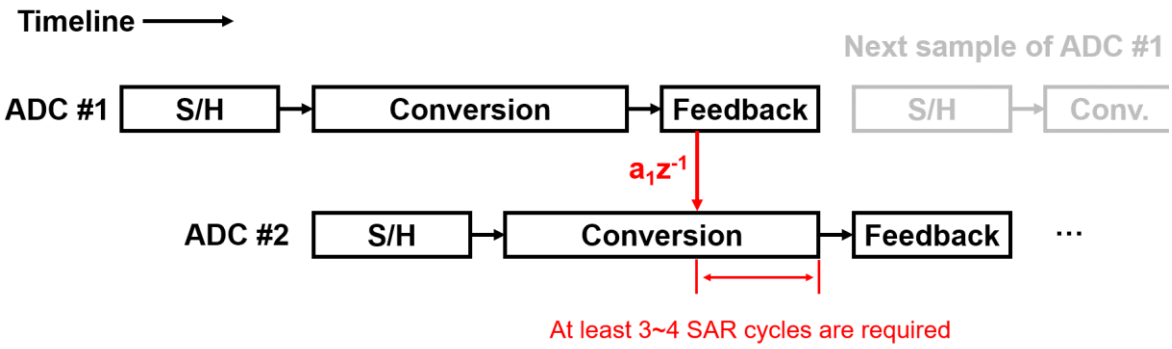


Fig. 67 The critical path in TINS SAR.

To speed up the system, we consider eliminating this critical path, or equivalently, eliminating the z^{-1} term in the NTF. Fortunately, this can be easily accomplished by a simple z to z^2 transformation, which is a special case of Filter Frequency Transformation [37], as shown in Fig. 68. For a DT transfer function, if we substitute every z with z^2, all the zeros and poles around DC will be transformed to a quarter of the sampling rate. This means that if we apply the transformation to a lowpass NTF, we will get a bandpass NTF without any z^{-1} term.

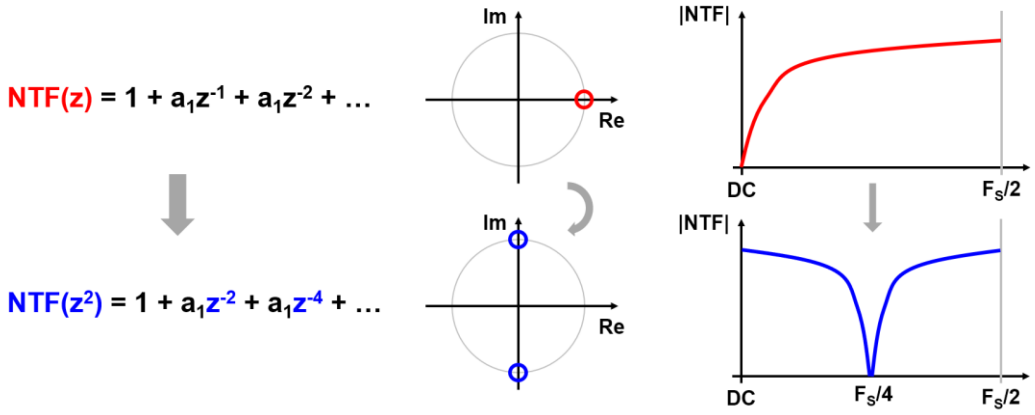


Fig. 68 z to z^2 transformation.

Therefore, we introduce a bandpass TINS SAR without the z^{-1} critical path, as shown in Fig. 69. This doubles the overlapping between the channels. The extended conversion time also prevents overloading, allowing an aggressive NTF to be used. It also reduces the redundancy bits needed, so the overall conversion cycles are reduced to further improve the sampling rate.

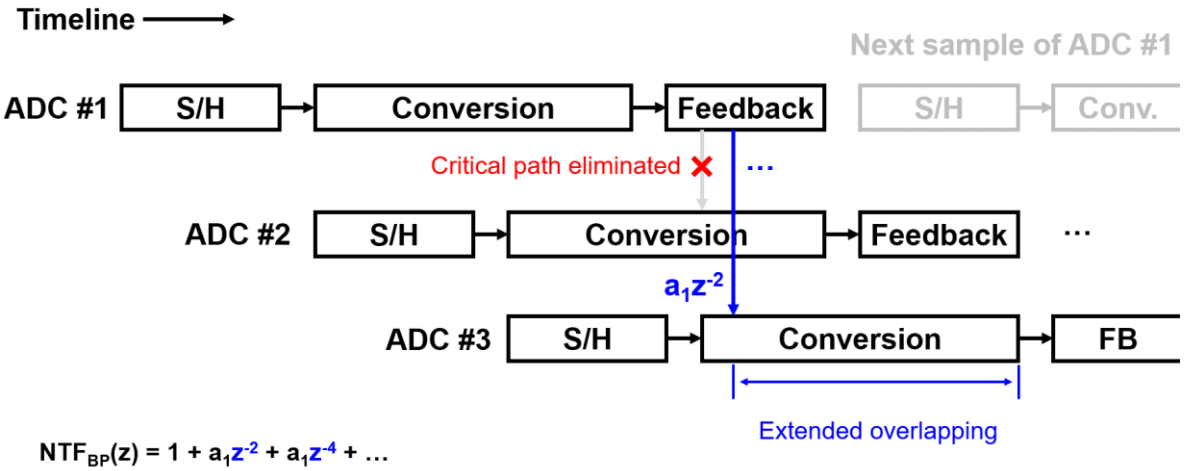


Fig. 69 Modified TINS SAR with bandpass NTF.

As mentioned in section 3.2D, interleaving artifacts is not a concern for the lowpass design because the artifacts are all at high frequencies and are out-of-band by oversampling. But for a bandpass system, a part of the artifacts will be mixed to in-band, as shown in Fig. 70. Although the fixed tone from offset mismatch may be easily removed in digital, the images mixed by gain mismatch are problematic.

To solve this, the simplest solution is using an odd interleaving number. The reason behind this is to avoid mixing with Fs/2, which brings the images to in-band. Fig. 70 shows a comparison between using 4x and 5x interleaving. In the case of odd channel interleaving, all artifacts are out of band as long as the OSR is greater or equal to the number of channels.

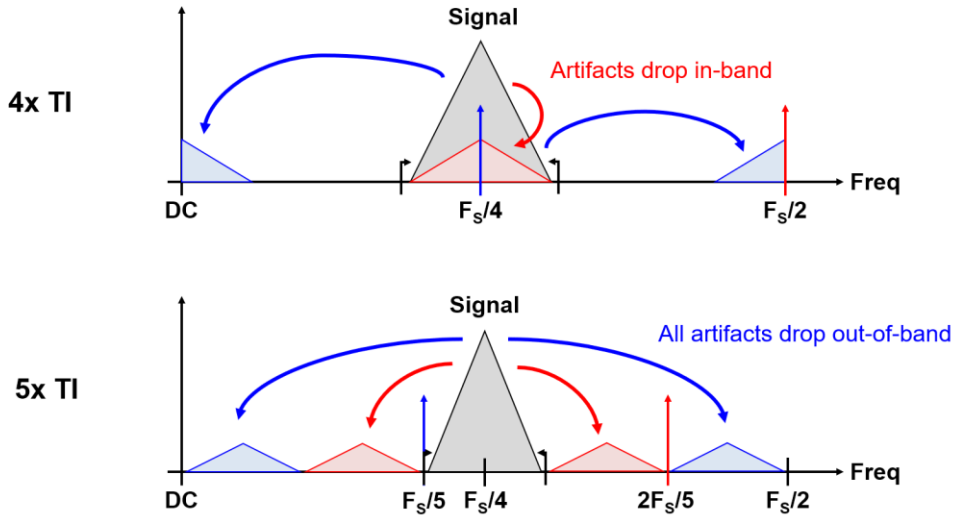


Fig. 70 Interleaving artifacts under bandpass input with different TI configurations.

The TINS SAR implementation in section 3.3 is mainly based on a summing pre-amplifier between the CDAC and comparator. This pre-amp sums and transfers the feedback signal to the other channels. And meanwhile, a group of capacitors accepts and samples the feedbacks from other channels. Fig. 71 shows a simplified schematic of the implementation.

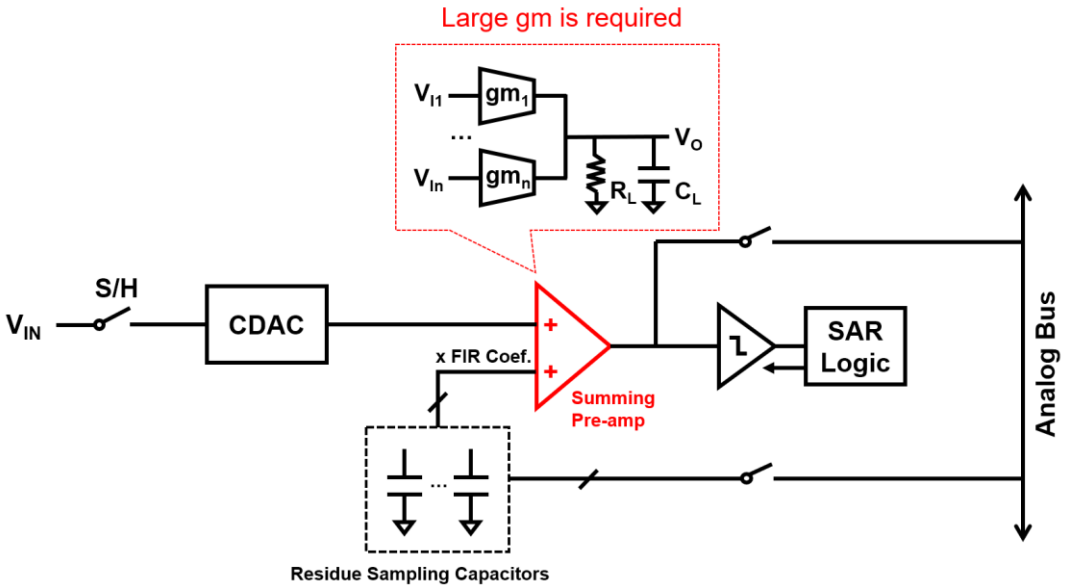


Fig. 71 The large gm requirement for the summing pre-amp in TINS SAR.

However this implementation has a serious drawback: The pre-amplifier is in a simple gm-R structure, and the gm needs to be large, and therefore burns a lot of power. On one side, this large gm is for noise, as this preamp's noise cannot be shaped. Fortunately, the TINS SAR quantizer's noise is not a big concern in this design, as the outer NS loop can suppress it. But on the other hand, this large gm is also for high bandwidth, as it has to settle the analog bus (a large capacitive load) in a very short time. This becomes even more challenging when we want to increase the sampling rate in this design further.

To solve this, we introduce a new circuit block named Dual Mode Summer (Fig. 72), which is a replacement for both the preamp and the comparator. The input stage of this summer is still a bunch of parallel gm cells, but there are two groups of circuitries loading on it for different operation mode.

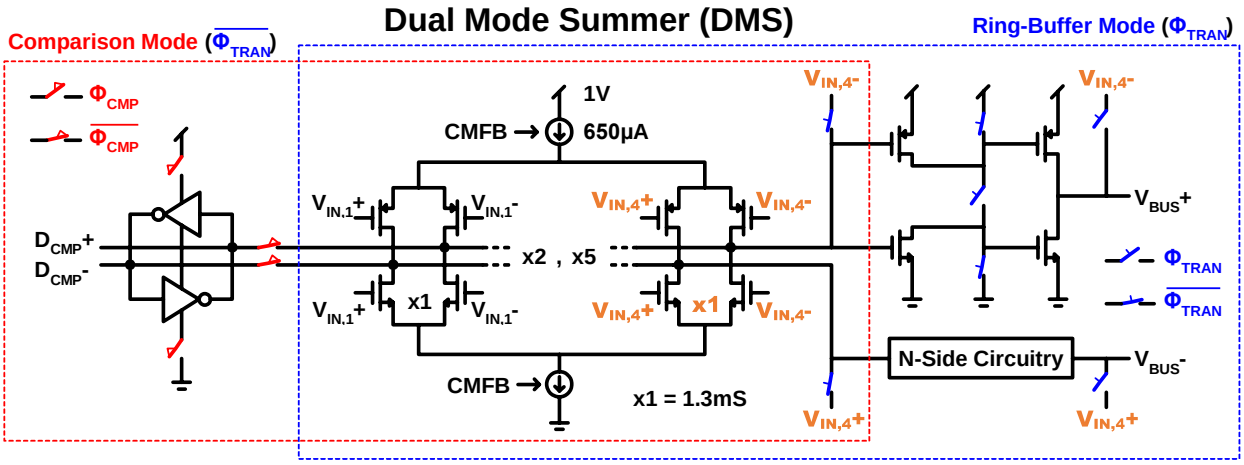


Fig. 72 Schematic of the Dual Mode Summer

The first mode is the comparison mode (Φ_{TRAN}=0). In this mode, the fourth input pair (connected to V_{IN4}) is self-connected, forming a low impedance load. Thus the input stage becomes a low gain but fast summing preamp and drives the high-speed latch behind for comparison. This mode is used for normal SAR conversion.

The second mode is more interesting, and is called the ring buffer mode (Φ_{TRAN}=1). In this mode, two cascaded inverters form a Ring Buffer (see section 5.3). Compared to the previous gm-R structure, such a design has an accurate gain that is well defined by the gm ratio:

$$V_{BUS} \approx -\frac{gm_1 V_{IN1}+gm_2 V_{IN2}+gm_3 V_{IN3}}{gm_4} \quad \text{Eq. 19}$$

And more importantly, the input pair is now decoupled from the loading cap. Therefore a much smaller gm can be used to save power.

Except for the Dual Mode Summer, the rest of this TINS SAR design is similar to the one in section 3.3, as shown in Fig. 73. But thanks to the extra conversion time by the bandpass configuration, only 1-bit of redundancy (i.e., the MSB) is required. Lastly, the input sampling switches act as an up-mixer when sampling the output of the loop filter. This is further explained in the next section.

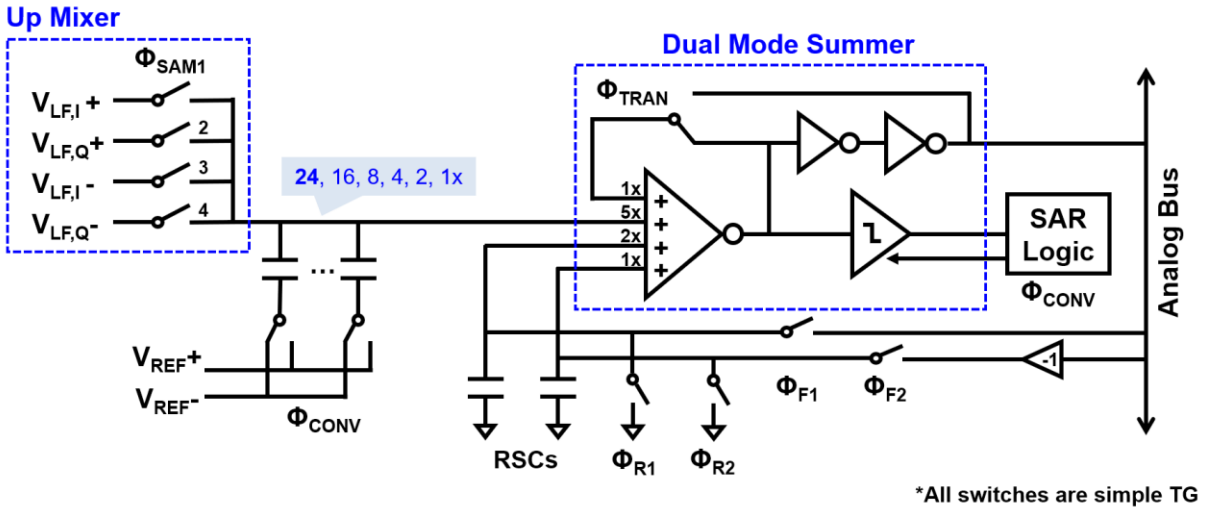


Fig. 73 Schematic of the improved TINS SAR design.

Fig. 74 shows the timing details of the TINS SAR. The design is 5x interleaved. Each channel is also divided into five different phases: The first cycle is for input sampling; The second

and third phases are for SAR conversion; Then the fourth phase is for residue transfer; And finally, the last phase is for sampling feedback. Besides, another round of feedback is getting in at the beginning of the SAR conversion (i.e., the second phase). These two feedback paths result in a 4th order bandpass NTF as shown. Lastly, the residue sampling caps are also reset before sampling the feedback to prevent memory effects.

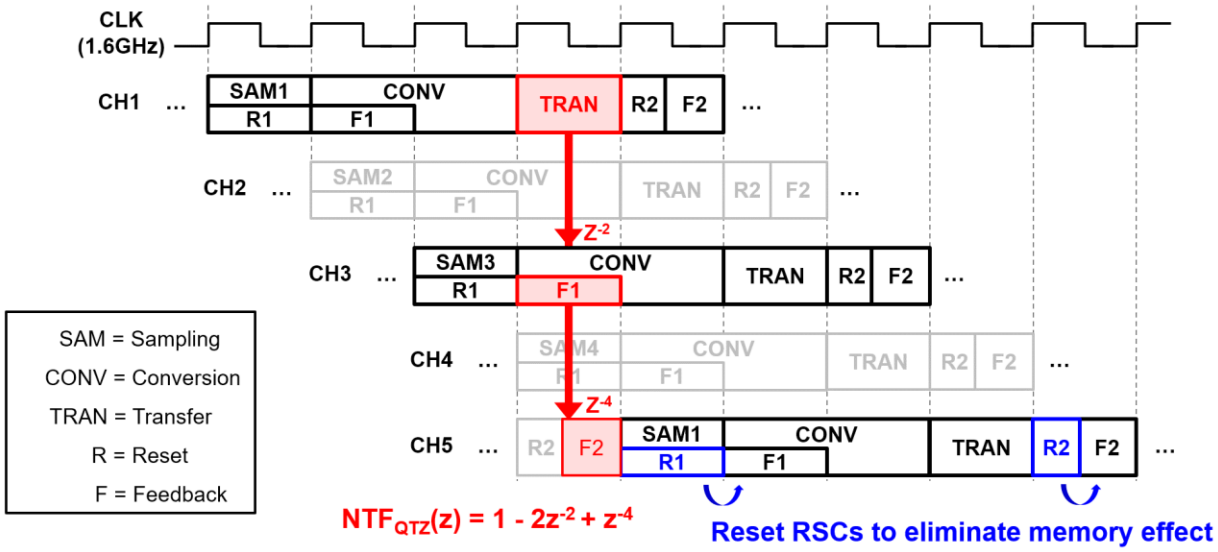


Fig. 74 Timing diagram of the TINS SAR.

B. Architecture

To operate with a bandpass quantizer, we built the hybrid loop DSM in a quadrature form [38], which accepts I and Q inputs. Fig. 75 shows the overall architecture of the design.

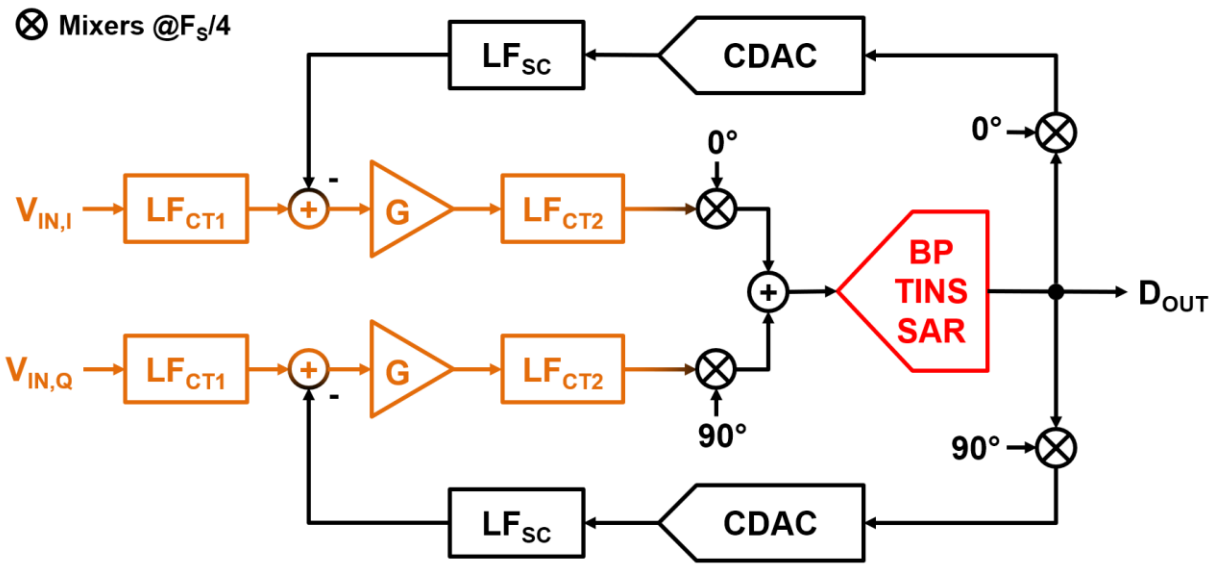


Fig. 75 Architecture of the example CT-DT Hybrid NS converter.

In this design, a pair of mixers up-mix the loop filters' output to Fs/4 before sending them to the bandpass quantizer. The output of the quantizer is down-mixed again before closing the feedback loop. Quadrature conversion is useful for many applications, such as wireless communication and radar. Compared to using two stand-alone I and Q ADCs, this configuration is more compact and low cost. Furthermore, handling quadrature signals in a single converter also helps matching I and Q channels.

C. Implementation

Fig. 76 shows a schematic of the prototype converter. Table 4 lists the main design parameters. The hybrid NS loop is formed with a 1st order DT filter and a 2nd order CT one. The bandpass noise-shaping SAR is a 6-bit 4th NS SAR running at a 1.6GHz sampling rate. The input CT filter is built with a simple RC network. This filter nearly matches the DT loop filter, so that the STF is flat in-band. Besides, it also provides some anti-aliasing and an easy-driven input structure.

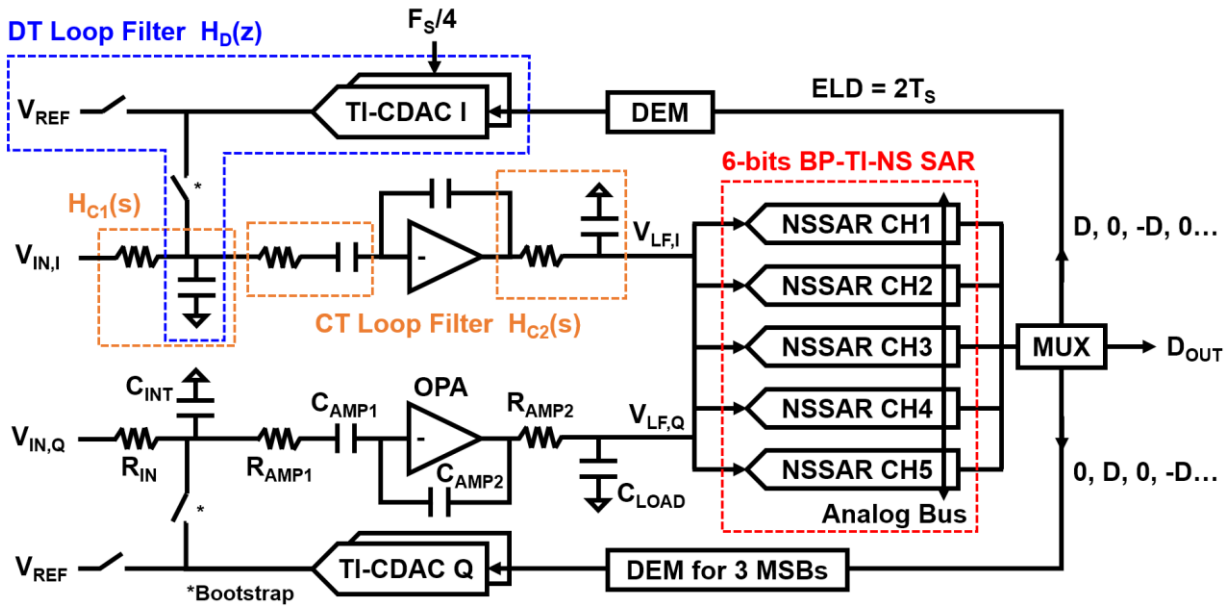


Fig. 76 Schematic of the example CT-DT Hybrid NS ADC.

Table 4 Main Design Parameters

R_{IN}	400 Ω	C_{INT}	5.9 pF
R_{AMP1}	500 Ω	R_{AMP2}	100 Ω
C_{AMP1}	2.9 pF	C_{AMP2}	360 fF
C_{DAC}	1 pF	C_{LOAD}	1.9 pF
C_U	1.5 fF	$C_{RS1,2}$	100 fF
G	8	G_{DAC}	0.8
		G_{QTZ}	

A closed-loop amplifier provides the loop gain, and the gain is accurately defined by capacitor ratio (C_{AMP1}/C_{AMP2}). The op-amp is a 3-stage feedforward architecture for high gain and high bandwidth (Fig. 77). The subdominant CT loop filter is built with two RC poles ($R_{AMP1}C_{AMP1}$ and $R_{AMP2}C_{LOAD}$) and separately placed around the amplifier. The extra pre-filtering by R_{AMP1} and C_{AMP1} is to block out the high-frequency switching noise and protects the amplifier from distorting.

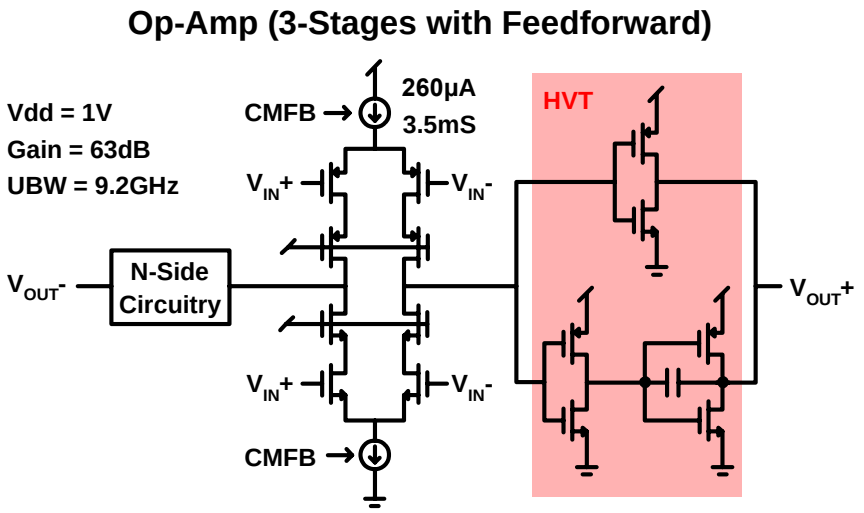


Fig. 77 Schematic of the op-amp in this design.

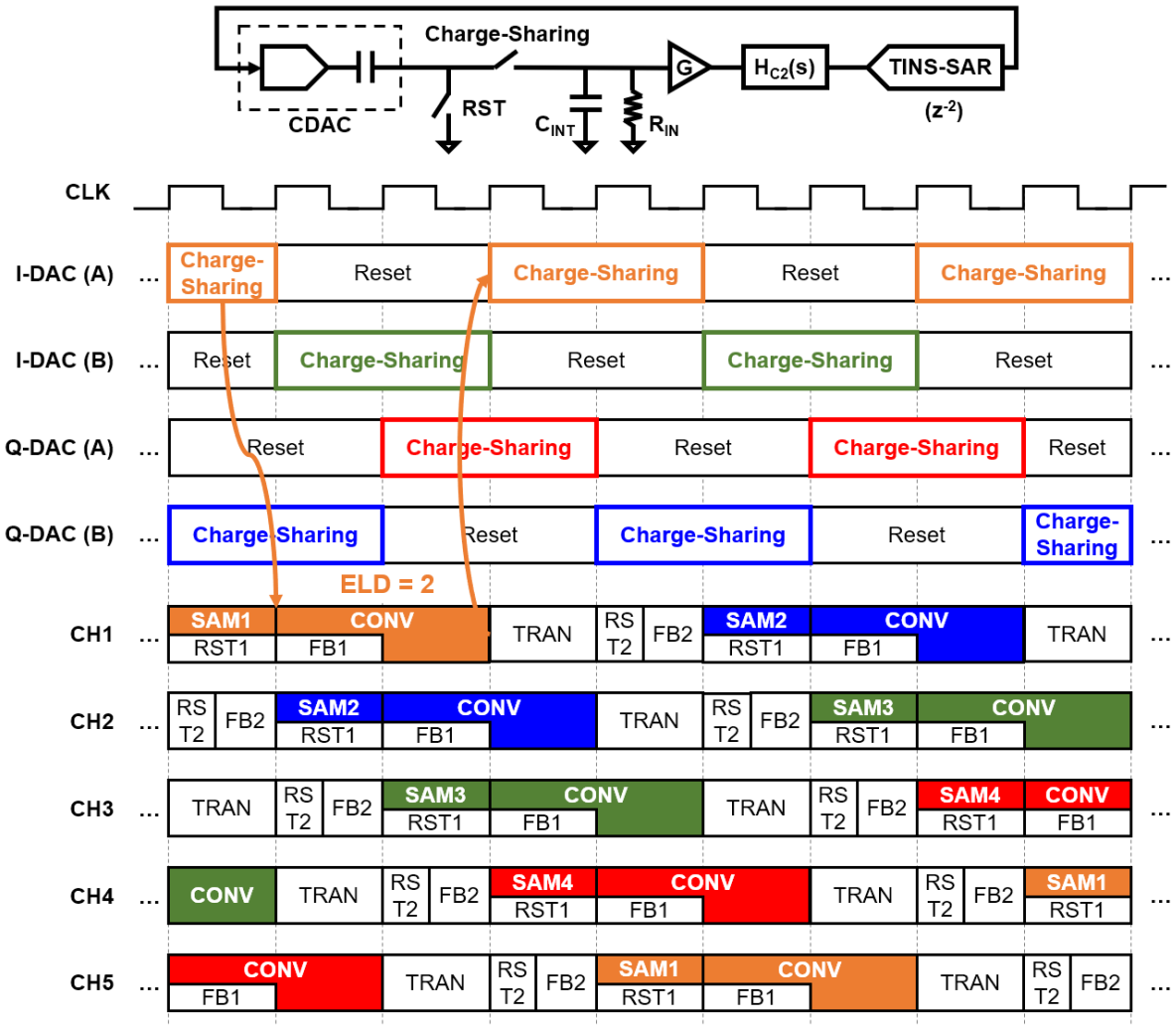


Fig. 78 Operation timing details.

A 2x interleaved CDAC charges an integration cap (C_{INT}) to perform the first-order filtering (i.e., the dominant DT pole). Bootstrap switches are used to improve linearity. Fig. 78 shows the operation timing details of CDAC and quantizer. In every two clock cycles, one of the CDAC is discharged to reset, and the other one gets connected to the integration cap for charge-sharing. The quantizer samples in the later cycle of the CDAC charge-sharing and finishes conversion in 2 cycles, which induce an ELD of 2 cycles.

Fig. 79 shows the zeros and poles location of the NTF and the loop loci are marked. From feedback theory, the open-loop poles turn into the zeros of the NTF, and the poles themselves are moved to the closed-loop locations along the root loci. The zeros and poles in Fig. 79 are before up-mixing for simplicity. And they will be transformed to Fs/4 in the actual converter. Notice that the zeros from the TINS SAR quantizer (marked in red) are independent of the NS loop, which means the NTF is decoupled.

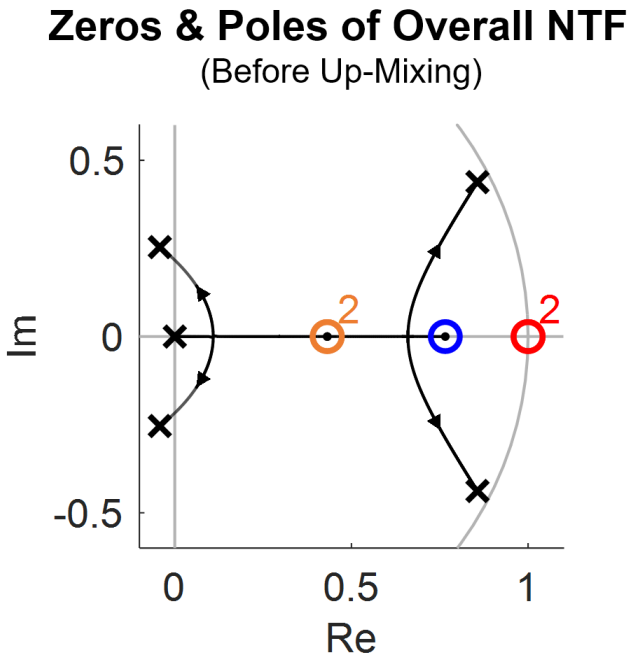


Fig. 79 Zeros and poles location of the NTF (before up mixing).

D. Silicon Results

The prototype converter is fabricated in 28nm CMOS, and is 0.09mm^2 in area. Fig. 80 shows a die photo. As we can see, the majority of area is taken by the capacitors in the loop filter, while the TINS SAR quantizer is compact.

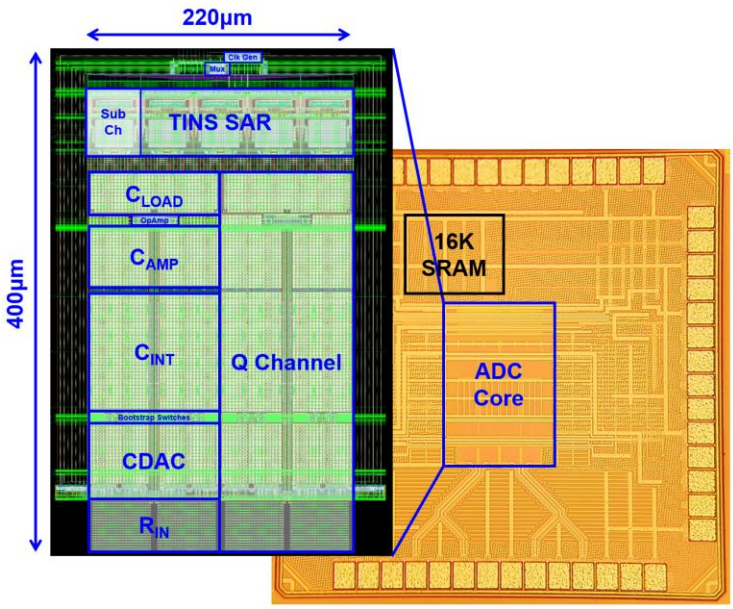


Fig. 80 Die photo and layout zoomed in.

Fig. 81 shows the noise floor of the converter when its input is shorted. Different NS configurations are compared, and we can clearly see the contribution of the two noise shaping systems. The interleaving artifacts show up as tones on the spectrum, but they are all out-of-band. The only artifact in-band is the tone at Fs/4 caused by the loop filter's offset. However, this tone can be easily removed in the digital domain, so we remove it for clarity.

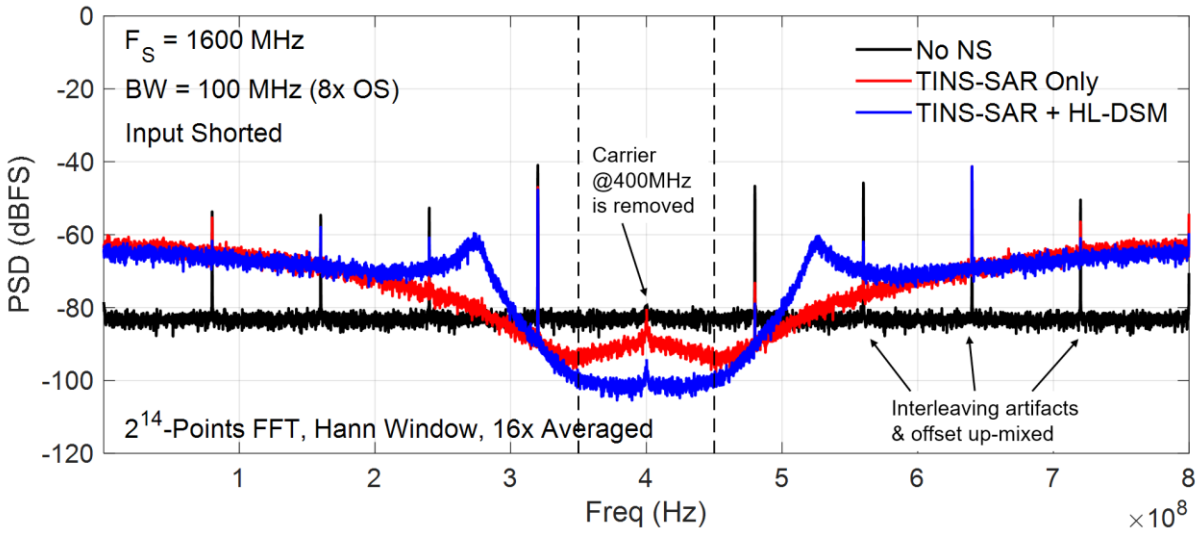


Fig. 81 Input shorted noise floor under with NS configuration.

Fig. 82 shows the spectrum of single-tone testing. The prototype achieves a peak SNDR of 67.5dB over a 100MHz bandwidth. As we can see, the noise floor well matches the designed NTF. Note that this is by the design's inherent robustness, and no tuning or calibration is applied. The image tone is due to the gain mismatch between the I and Q channel. But as a quadrature converter, this image can be easily fixed in the digital baseband, so it is not counted as an error.

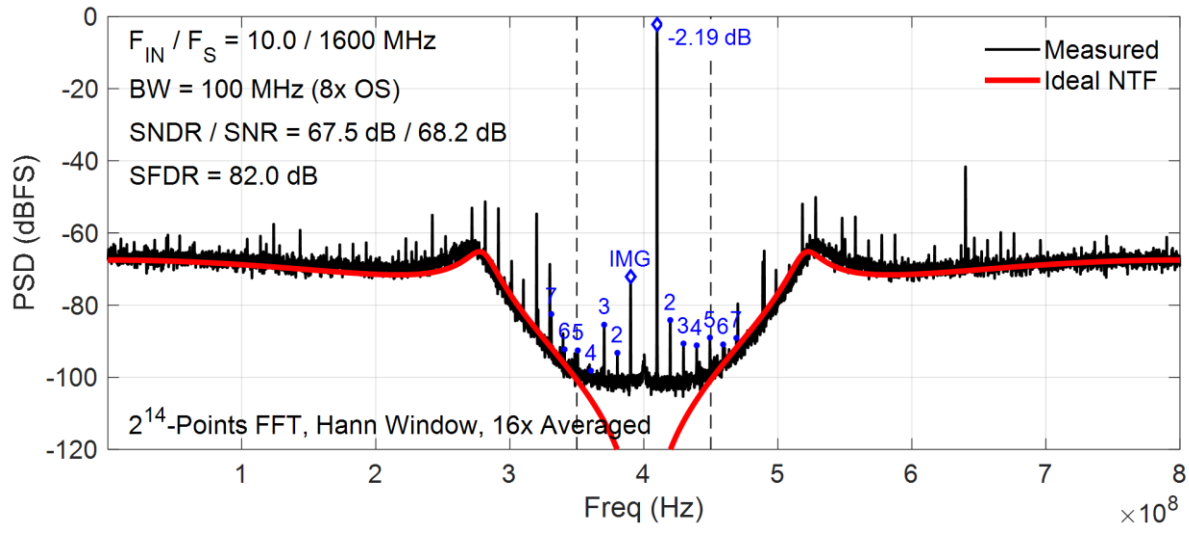


Fig. 82 Single-tone testing result.

Fig. 83 shows the spectrum of two-tones testing. An IMD3 of -79dB is measured at near the maximum input frequency (50MHz), further proving the design's linearity.

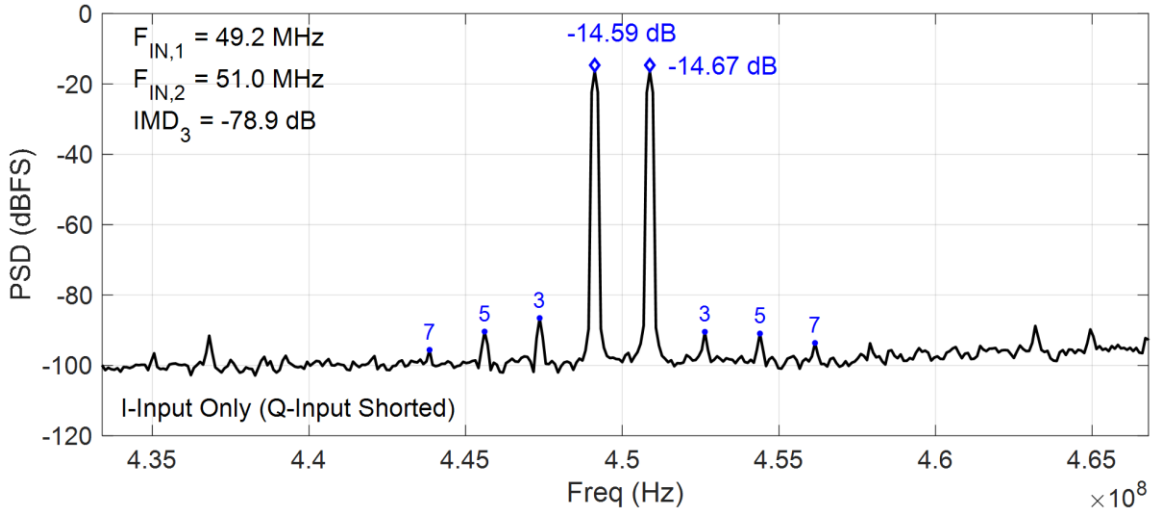


Fig. 83 Two-tone testing result.

The left plot on Fig. 84 shows the measured performance over different input amplitudes, where the dynamic range is measured to be 69dB. The right plots on Fig. 84 show the frequency response and noise floor over input frequencies. The measured signal transfer function is flat in-band and well-matched the design. And the noise performance is also well consistent over the whole frequency range.

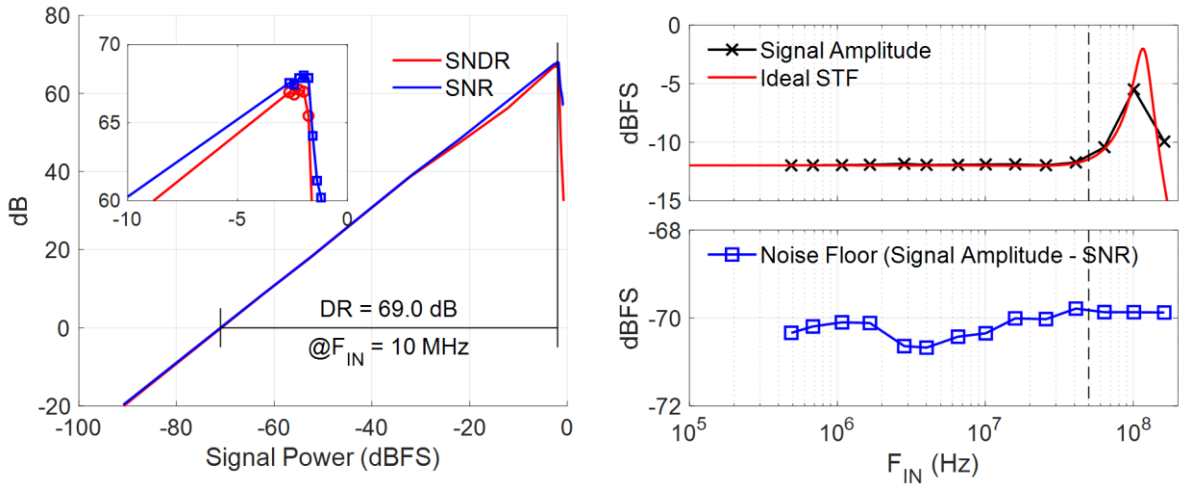


Fig. 84 Measured performance over input magnitudes (left) and frequencies (right).

As mentioned, an essential feature of the hybrid NS system is its robustness to clock variations. To illustrate this, we sweep the clock frequency and observe a consistent noise performance across a few octaves, as shown in Fig. 85 (left). We also try injecting jitter to the DAC clock, and the observed SNR degradation is notably better than the CT DSM, even of the simulations (Fig. 85, right).

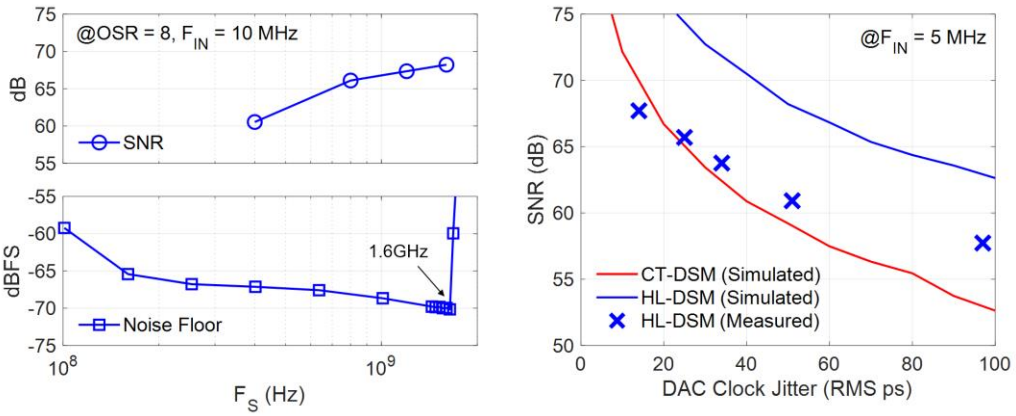


Fig. 85 Measured performance over sampling rate (left) and with DAC jitter (right).

To further exam the robustness, we measured 18 different devices over analog supply variation, as shown in Fig. 86. The resulting 1-sigma range of SNR is less than 3dB. Fig. 86 also shows the power breakdown of the converter running at a 1.6GHz sampling rate, where the total power consumption is 13.4 mW.

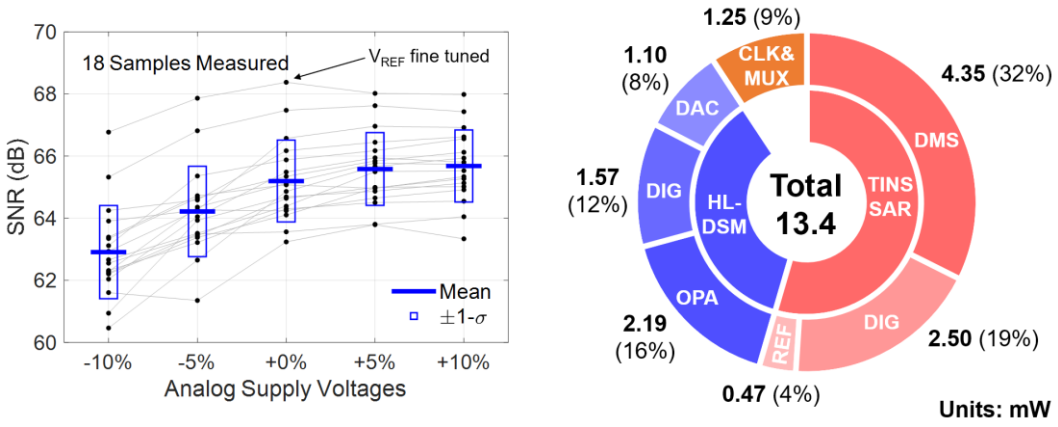


Fig. 86 Multi-device measurement over analog supply voltage variation (left), and power breakdown (right).

Finally, Table 5 compares this converter with other state-of-the-art converters. Compared to the conventional CT SD ADCs, our architecture is free of calibration and tuning, and is insensitive to jitter and ISI. But it is still easy-driven and with inherent anti-aliasing, which is lacking in the standalone NS SAR converters. The bandwidth and efficiency are still comparable to CT SD converters.

Table 5 Comparison Table

	This design	ISSCC 2017 S. Huang [39]	ISSCC 2019 W. Wang [40]	ISSCC 2018 T. He [32]	ISSCC 2019 L. Jie [28]
Architecture	**HL-DSM w/ TINS-QTZ**	CT-DSM w/ VCO-QTZ	CT-DSM	CT-DSM	TINS-SAR
Calibration & Tuning	**None**	DAC Mis. + RC	DAC Mis. + RC	DAC ISI + RC	None
Jitter & ISI Sensitivity	**Low**	High	High	Calibrated	Lowest
Easy-Driven & Anti-Aliasing	**Yes**	Yes	Yes	Yes	No
Process (nm)	**28**	16	28	28	40
Area (mm^2)	**0.088**	0.217	0.019	0.25	0.06
Power (mW)	**13.4**	54	16.3	64.3	13.0
Supply (V)	**1**	1 / 1.35 / 1.5	1.1 / 1.5	1.16 / 1.5	1
Fs (GHz)	**1.6**	2.15	2	2	0.4
NTF Order	**6 (I+Q)**	4	4	4	4
BW (MHz)	**100 (I+Q)**	125	100	50	50
SNDR (dB)	**67.5**	71.9	72.6	79.8	70.4
DR (dB)	**69.0**	74.8	76.3	82.8	71.7
FOM_S (dB)	**166.2**	165.5	170.5	168.7	166.3
FOM_W (fJ/c.step)	**34.6**	67.2	23.4	80.5	48.1

Chapter 5 Advanced Circuit-Level Techniques for Data Converters

In this chapter, we go through some circuit-level techniques that support the implementation of advanced NS data converters. All of these techniques are original from the design examples in this thesis.

5.1 Parasitic Pre-Charging

Parasitics are always a big concern in switch-capacitor (SC) circuits, especially the parasitic capacitance. This is because the parasitic capacitance can induce extra charge, which can be regarded as a kind of error in SC circuits. This problem can be even worse if the parasitic cap is nonlinear: Nonlinear error is much harder to remove or compensate in later processing. But unfortunately, many common parasitic caps are highly nonlinear, such as the gate capacitor of a transistor and a parasitic diode's capacitance.

Conventionally, this problem can be addressed by bottom-plate sampling. The key behind this is to make sure that the parasitic cap always returns to the same state, such that it won induce error charge at the end of each cycle. But bottom-plate sampling solves the problem at the cost of increasing circuit complexity. Such increment of complexity is especially notable in SAR converters, as their CDAC array requires multiple bootstrap switches to implement bottom-plate sampling. Fig. 87 gives a comparison of SAR converter with top-plate sampling and bottom-plate sampling.

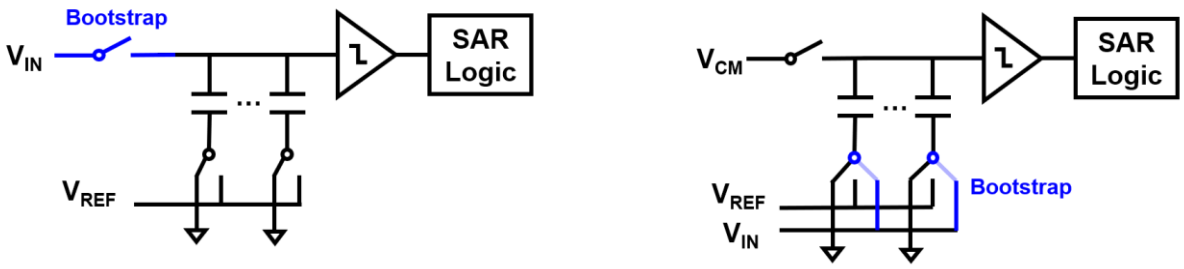


Fig. 87 SAR converter with top-plate sampling (left) and bottom-plate sampling (right)

As we can see, in a bottom-plate sampling SAR converter, every single cap in the CDAC requires a standalone sampling switch (usually a bootstrap switch). This can be costly in the sense of chip area and wire routing. On the other hand, top-plate sampling only requires one sampling switch regardless of the scale of CDAC. The simplicity of top-plate sampling also brings advantages in speed and power, which is attractive for some high bandwidth applications.

Indeed, for many data conversion applications, linear errors (e.g., gain error and offset) are acceptable. This implies that we don't necessarily need to eliminate all parasitic effects, but we just need to address those with non-linearity. Back to the SAR converter example, there are usually two sources of parasitic capacitance on the CDAC top plate: the wiring parasitic and the transistor parasitic, as shown in Fig. 88.

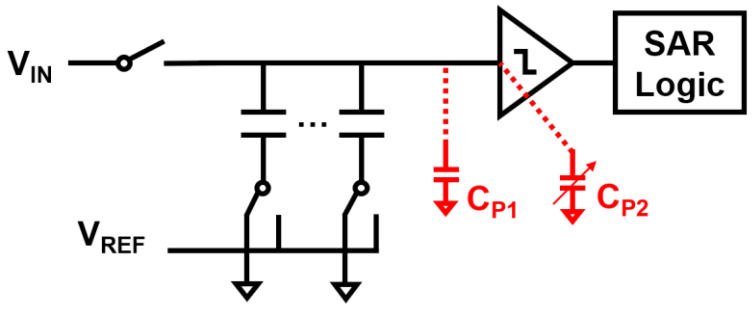


Fig. 88 The two kinds of parasitic caps in top-plate sampling SAR converter.

Here, C_{P1} is the wiring parasitic and is typically linear, while C_{P2} is the nonlinear parasitic from the comparator's input transistors (or any other active blocks connected to the CDAC). Based

on our previous discussion, C_{P1} will only induce gain error and is usually acceptable. But we hope to cancel out the nonlinear charge that brings by C_{P2}.

Following the essence of bottom-plate sampling, we introduce a method to preset the state of C_{P2} at the initial of each cycle, named Parasitic Pre-Charging. Fig. 89 shows the basic concept of this technique.

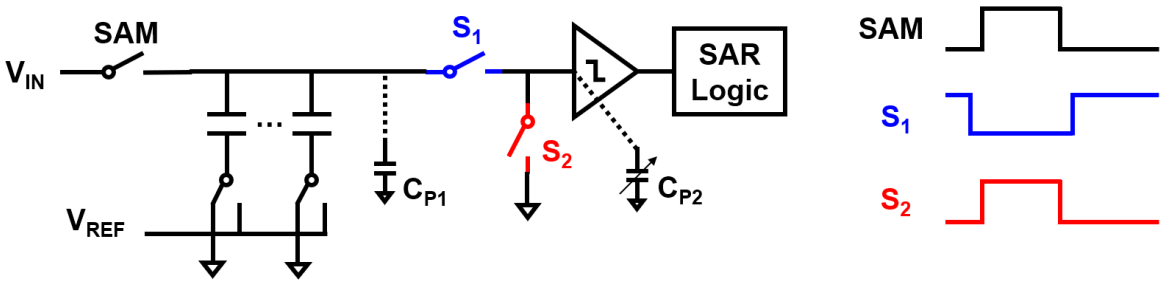


Fig. 89 Basic concept of Parasitic Pre-Charging technique.

The key of Parasitic Pre-Charging is to charge the nonlinear parasitic cap to its final voltage at the very beginning. As we know, in a SAR converter, the voltage at the input of the comparator always converges to around zero (or the virtual ground). This means that the final voltage on the parasitic cap will be very close to zero, and we just need to charge (reset) the cap to zero voltage before each sampling. To realize this, switch S_1 and S_2 are induced. Before sampling, S_1 is open, disconnecting the CDAC and the comparator (i.e., the C_{P2}). This allows S_2 to reset the C_{P2}. And after sampling, S_1 is closed while S_2 is open, connecting the CDAC to the comparator again. So C_{P2} will no longer induce extra charge to the CDAC at the end of conversion, and thus the nonlinearity is canceled.

A potential drawback of this technique is that the switch S_1 induces a new noise source, which can be significant if C_{P2} is small and the noise has a wide bandwidth. To reduce this noise, using a large size S_1 to reduce its on-resistance is a possible solution. But an over-sized S_1 may induce another significant parasitic again. Thus, we introduce another solution: placing an extra

capacitor (C_n) across switch S_1 to limit its noise bandwidth, as shown in Fig. 90. This capacitor also helps with the settling in case of a fast SAR conversion. The extra charge sampled on this capacitor is not a concern, as it will be eventually discharged by S_1 and will not get into the CDAC.

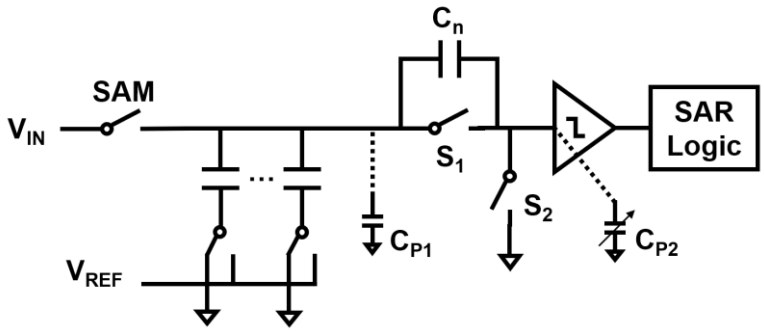


Fig. 90 Adding a noise limiting capacitor.

5.2 Dynamic Bandwidth Amplification

Most of the data converter architectures rely on Discrete-Time (DT) amplification. For example, the pipeline converter requires a residue amplifier between stages; Most DT NS converters need active SC circuits to build the loop filter, including the DT amplifier. In high-resolution converters, the noise of the amplifier usually dominates the SNR. We always hope to reduce the amplifier's noise at a given power consumption. In this section, we discuss the noise and timing-robustness of DT amplifiers, and introduce an advanced DT amplification technique named Dynamic Bandwidth Amplification.

A DT amp is essentially an amplifier that works on sampled signals Fig, and its output is also sampled after a given time (i.e., amplification time).

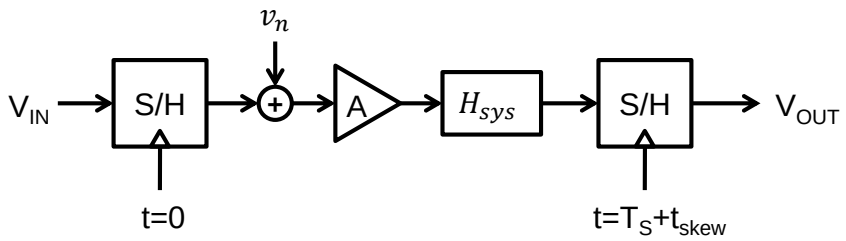


Fig. 91 Generalized signal model of DT amplifier.

Here, A is the nominal gain of the amplifier, and H_{sys} is a linear system that models the settling behavior. A continuous-time white noise source, $v_n(t)$, models the additive noise of the amplifier. In most cases, the amplifier transconductance dominates v_n and determines the amplifier power consumption. T_S and t_{skew} are the sampling period and timing skew, respectively. Two critical specifications of a DT amplifier, the relative gain variation (g) and input-referred noise ($v_{n,i}$), can be derived from the model:

$$g = \frac{a(T_S+t_{skew})}{a(T_S)} \quad \text{Eq. 20}$$

$$v_{n,i} = \frac{\int_0^{T_S} v_n(t)h(T_S-t)dt}{a(T_S)} \quad \text{Eq. 21}$$

where, $h(t)$ and $a(t)$ are the impulse and step response of H_{sys}, respectively.

i- Conventional DT amplifiers

Conventional DT amplifiers can be classified into Settling Based Amplifiers (SBAs) and Integration Based Amplifiers (IBAs). IBAs are also known as Dynamic Amplifiers[1] (DAs). Fig. 92 illustrates the two types and also provides the corresponding H_{sys}.

[1] "Dynamic amplifier" refers to open-loop gm-C amplifiers. Closed-loop dynamic amplifiers, such as [22], are essentially SBAs with variable bandwidth.

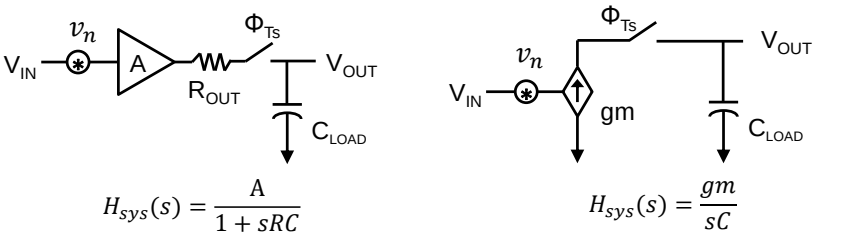


Fig. 92 Conventional DT amplifiers.

Although the two types of amplifier appear to be very different, the H_{sys} for both is a time-invariant, single-pole system, i.e. $H_{sys}(s) = \frac{K}{s-p}$. We define two metrics, $\mathcal{R}_n$ and $\mathcal{R}_g$, to measure the noise performance and gain-timing sensitivity, respectively:

$$\mathcal{R}_n \triangleq \frac{noise\ of\ DT\ amp.}{noise\ of\ IBA} = \frac{v_{n,i}^2}{\left(\frac{v_n^2}{\Delta f}\frac{0.5}{T_S}\right)} = \frac{1}{2}pT_S \cdot \frac{e^{pT_S}+1}{e^{pT_S}-1} \quad \text{Eq. 22}$$

$$\mathcal{R}_g \triangleq \frac{sensitivity\ of\ DT\ amp.}{sensitivity\ of\ IBA} = \frac{\mathrm{d}(g)}{\mathrm{d}\left(\frac{t_{skew}}{T_S}\right)} = pT_S \cdot \frac{e^{pT_S}}{e^{pT_S}-1} \quad \text{Eq. 23}$$

where $p < 0$ for LHP pole. Physically, $\mathcal{R}_n$ compares the input-referred noise of a DT amplifier to an IBA, and $\mathcal{R}_g$ compares the gain-timing sensitivity of a DT amplifier to an IBA.

Fig. 93 plots the two metrics as functions of p for a DT amplifier. We see that there is a tradeoff between $\mathcal{R}_n$ and $\mathcal{R}_g$, which means we cannot simultaneously improve both noise and robustness by changing the pole of H_{sys}.

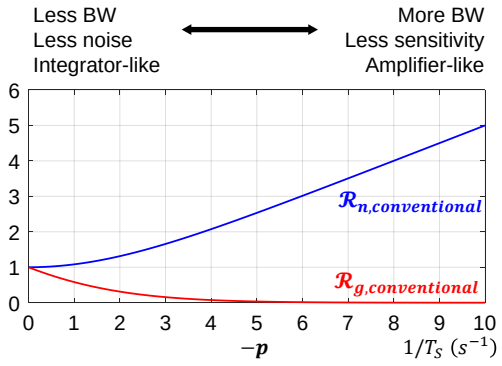


Fig. 93 Noise-robustness tradeoff for a conventional DT amplifier.

ii- DT amplifier with Multi-Phase Settling

We can break the noise-robustness tradeoff of conventional DT amplifiers by using a time-variant H_{sys} (i.e., Dynamic Bandwidth). Here we introduce a simple realization of Dynamic Bandwidth Amplification named as Multi-Phase Settling. In a Multi-Phase Settling amplifier, H_{sys} is still a single-pole system, but we change the pole location in different phases. As an example, Fig. 94 shows the amplification progress for a two-phase settling amplifier.

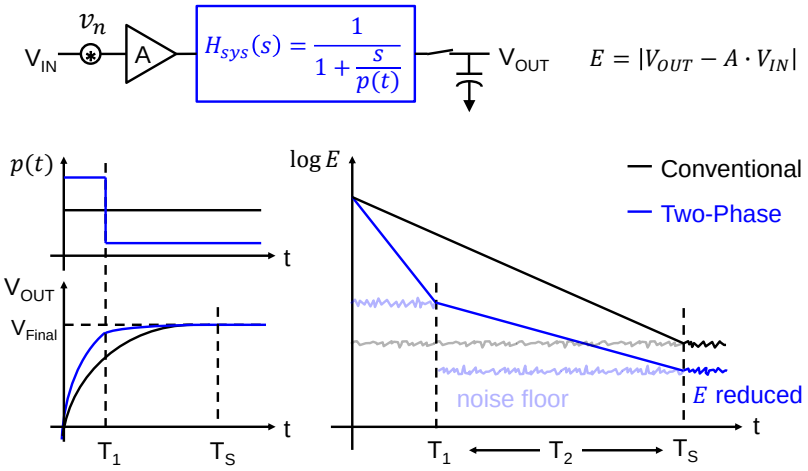


Fig. 94 Amplification progress (step response) of two-phase settling amplifier.

Here, H_{sys} has a large bandwidth in the first phase but also induces more noise. During the second phase, the bandwidth is reduced, suppressing noise. Intuitively, since the second phase's bandwidth is smaller than with a conventional SBA, the overall noise is reduced. An essential advantage is that gain-timing sensitivity is much lower compared to a DA. The two metrics of a two-phase settling amplifier (i.e., $\mathcal{R}_{n,2P}$ and $\mathcal{R}_{g,2P}$) are:

$$\mathcal{R}_{n,2P} = \frac{1}{2}T_S \cdot \frac{-p_1(1-e^{2p_1T_1})\cdot e^{2p_2T_2} - p_2(1-e^{2p_2T_2})}{(1-e^{(p_1T_1+p_2T_2)})^2} \quad \text{Eq. 24}$$

$$\mathcal{R}_{g,2P} = -(p_1T_1 + p_2T_2)\frac{e^{(p_1T_1+p_2T_2)}}{1-e^{(p_1T_1+p_2T_2)}} \quad \text{Eq. 25}$$

Fig. 95 plots these metrics versus phase-division ratio, which supports the conclusion that two-phase settling can improve both noise efficiency and gain-timing robustness.

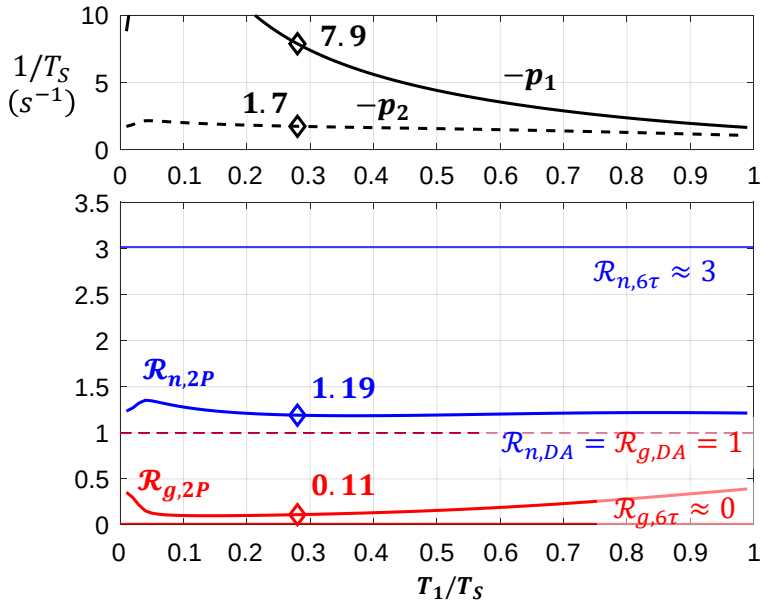


Fig. 95 Optimal pole locations (top) and metrics (bottom) of two-phase settling amplifier versus phase-division ratio.

In the figure, the x-axis is the fraction of T_S that the amplifier operates in phase 1 (i.e., T_1). The pole locations are calculated by minimizing $\mathcal{R}_{n,2P} + \mathcal{R}_{g,2P}$ for each value of T_1. When compared to a conventional SBA with 6τ settling, two-phase settling reduces noise ($\mathcal{R}_n$) by about 60%. This implies 60% less power on gm transistors is required to achieve the same noise level. Although the noise with two-phase settling (and optimal phase division) is still a little higher (18%) than with a DA, the gain sensitivity ($\mathcal{R}_g$) is 9x smaller. In a practical design, different weightings can be applied to $\mathcal{R}_{n,2P}$ and $\mathcal{R}_{g,2P}$ to emphasize noise or sensitivity depending on the application.

Adding more settling phases can further reduce noise and gain sensitivity, as Fig. 96 shows. With enough phase divisions, the theoretical noise approaches that of a DA, while the gain robustness remains as good as a conventional SBA. But two-phase settling is still the most practical choice among multi-phase designs, as it provides good performance with a negligible increase in complexity.

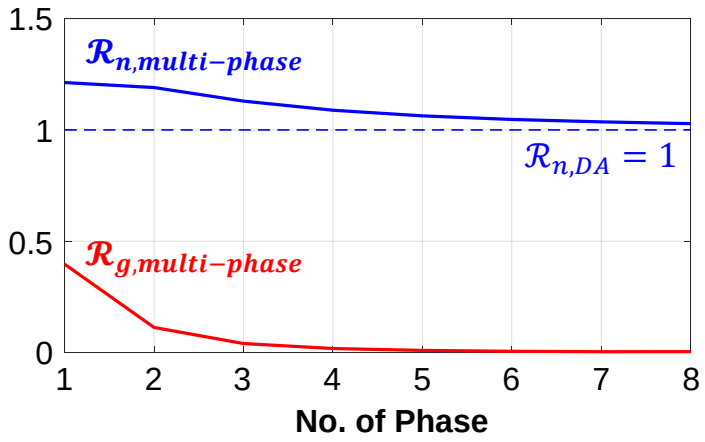


Fig. 96 Theoretical (optimal) performance versus number of settling phases.

5.3 Ring Buffer

As mentioned in the last section, high-performance DT amplifiers are among the most critical blocks for many converter architectures. And there are usually two strategies for DT amplification: open-loop and closed-loop. Generally speaking, open-loop amplifiers are simpler to design, lower in cost, and are usually more efficient. But they are inevitably sensitive to circuit variations (PVT), which in return increase the cost for calibration or trimming. In contrast, closed loop amplifiers are much more robust, because the high gain in their feedback loop suppresses the variations. The biggest challenge for closed-loop amplifiers is the low intrinsic gain of advanced node CMOS. This problem is getting even worse as the supply voltage shrinks, limiting many circuit structures, like cascodes. Closed-loop amplifiers will eventually lose their advantages without the support of enough loop gain.

Recently, an emerging amplifier structure, named Ring Amplifier [41], provides a promising option for closed loop amplification. Fig. 97 illustrates a basic schematic of the Ring Amplifier.

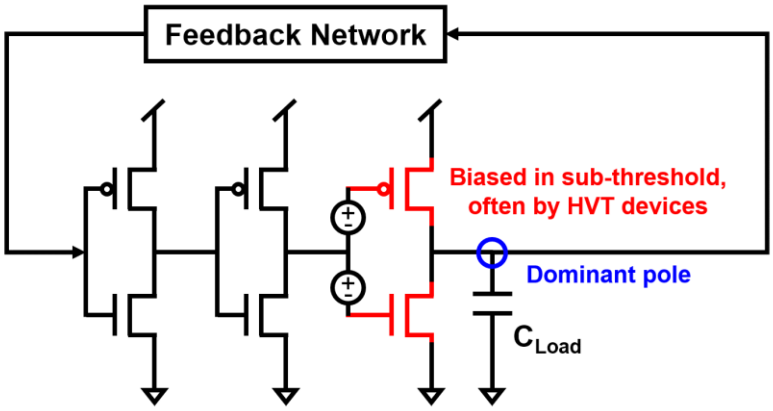


Fig. 97 Basic schematic of Ring Amplifier.

Essentially, the Ring Amplifier is a 3-stage, class-AB op-amp, but it has a few unique features: 1) Each stage (especially the 2nd and 3rd stage) is built with an inverter-like structure, with only self-biasing, such that it can work on a low supply voltage. 2) The output stage is biased at class-C state (sub-threshold), thus it provides a very high output resistance if no output DC load is mounted. This high output resistance forms a low-frequency pole when driving a (not too small) capacitor load, and this pole is designed to be the dominant pole to stabilize the feedback loop. 3) the output stage consumes a very low quiescent current because of the biasing, but it can deliver high current at slewing. Indeed, a properly designed Ring Amp slews most of the time, and only performs linear settling at the very end of the amplification cycle. This makes the Ring Amp highly efficient, not only because of the low quiescent-to-output ratio, but also because it naturally features the dynamic bandwidth feature mentioned in section 5.2. Compared to the Multi-Phase Settling technique, Ring Amp is more elegant as it changes bandwidth automatically.

However, there are still some challenges when using Ring Amp in practice. One problem is the restricted feedback configuration. Indeed, there is almost only one way to configure Ring Amp as a fixed gain amplifier, as shown in Fig. 98.

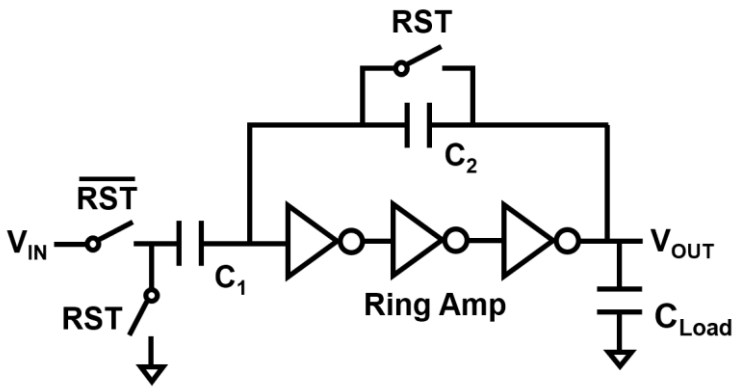


Fig. 98 Classic SC gain block by Ring Amplifier.

Although this is a very well-known SC amplifier configuration, it has a few drawbacks: 1) It presents a capacitor load (C_1) to the input, which can be problematic in some cases. Especially, this capacitor can be considerably large for KT/C noise requirement. 2) Unlike conventional op-amp, Ring Amp is very poor at driving resistive load, so the DC point of this circuit cannot be biased with resistors. Instead, the circuit has to be reset by every cycle. This might not seem a problem as the whole system is DT, after all. But it does indtroduce an extra reset phase, so the available amplification time is reduced, leading to higher bandwidth requirement (i.e., more power). Reset switches also increase the area and routing. Besides, the capacitor C_1 and C_2 take extra chip area, but they are also not necessary in principle.

To solve these problems, we introduce a light-weight variant of Ring Amp, named Ring Buffer (Fig. 99). The Ring Buffer is designed for DT small signal amplification. It can work as a fixed-gain amplifier alone without any extra components. Moreover, it provides a high impedance input, enabling easy deployment in a system.

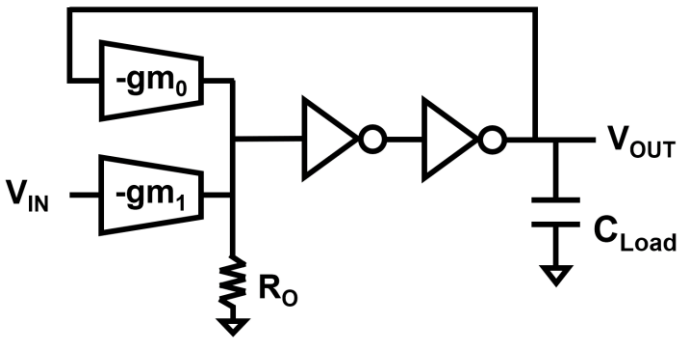


Fig. 99 Schematic of Ring Buffer.

The basic principle of Ring Buffer is the same as Ring Amp. But its input stage is a multi-input gm summer. This input stage realizes the feedback summation in the current domain, and the matching of input transistors guarantees the summation ratio. In advance node CMOS, the matching of transistors is acceptable for many applications.

The Ring Buffer's output is connected back to one of the input stage (gm_0) and closes a feedback loop. Due to the high gain of the 3-stage structure, the Ro of the input stage is negligible, and the nominal closed-loop gain of Ring Buffer is given as:

$$A = \frac{gm_1}{gm_0} \qquad \text{Eq. 26}$$

The Ring Buffer provides great driving ability to the loading cap. Meanwhile, it presents high impedance (with tiny parasitic caps) on the input side. Such features are very similar to an ideal buffer, thus our name. Compared to the gm-R amplifier mentioned in section 5.2, the Ring Buffer has an accurate gain which is well defined by the gm ratio. More importantly, the primary gm (input stage) is decoupled from the loading cap, so a smaller gm may be used to save power.

A possible concern on the Ring Buffer is its linearity, as it relies on the open-loop linearity of the gm transistor. However as mentioned, the Ring Buffer is aimed at small signal amplification, where linearity may not be a concern. Section 3.3 gives a good example of deploying a Ring Buffer in a data converter.

Chapter 6 Conclusions

In Chapter 1, we first review the basics of data conversion and go through the main specifications. A new modified version of FoM including the area is introduced for a more comprehensive comparison. We then review the fundamentals of noise-shaping, including some popular noise-shaping converter architectures, and discuss their limitations.

In Chapter 2, the discussion concerns nested noise-shaping. We first review the three implementation forms for digital filters, and apply those to the NTF of noise-shaping converters. The Nested structure is then introduced to realize the NTF in cascade form, which brings the advantages of high order and robustness. Finally, we go through a prototype Nested NS SAR ADC as an example of nested noise-shaping. The measurement results illustrate the benefits of this technique.

In Chapter 3, we move our discussion to the time-interleaving techniques. A brief review of time-interleaving is given. Following that, we discuss the combination of time-interleaving and noise-shaping, analyze the inherent difficulties of such a combination, and introduce two solutions: Midway Feedback and Multi-Phase Conversion. Non-idealities of interleaving and their solutions are also discussed. At the end of this chapter, we go through another design example, an Interleaved NS SAR ADC, to better understand how to apply time-interleaving in practice. The silicon results prove the effectiveness of time-interleaving in boosting the bandwidth of NS data converters.

In Chapter 4, we further step in the continuous-time noise-shaping field. The pros and cons of CT NS converter are first reviewed. A CT-DT hybrid NS ADC is then introduced as an ultimate solution. We also examine the trade-offs and practical concerns of this new type of NS system. Lastly, we walk through the third prototype converter with the new techniques. An improved version of TINS SAR converter is also introduced as the nested quantizer in this prototype, which achieves higher speed and efficiency. The final measurement results support our conclusions about the new architecture's advantages.

Finally in Chapter 5, we detailly discuss three circuit-level innovations collected from the design examples. These circuit-level techniques further improve the NS converter's performance from different aspects, including linearity and amplifier efficiency.

The main innovations of this thesis include:

- The first nested (cascaded) NS SAR ADC, realizing robust high-order (section 2.3). This is the first 4th-order NS SAR with aggressive NTF, and it is the first >85dB SNDR NS SAR over 100kHz.
- The first time-interleaved NS SAR ADC (section 3.3). This is the first 4th-order NS SAR (with mild NTF), and it is the first NS SAR achieving 50MHz bandwidth.
- The first tuning-free NS ADC with inherent anti-aliasing (section 4.3).
- The parasitic pre-charging technique mitigating nonlinear parasitic cap in top-plate sampling SAR ADC (section 5.1).
- The multi-phase settling technique for DT amplification (section 5.2).
- The Ring Buffer (section 5.3).

Appendices

A. Assumptions for the Noise-Shaping SAR Model Used in Section 2.1

The NTF for the direct and cascade forms are:

$$NTF_{direct} = (1 - z^{-1})^n = 1 + \sum_{k=1}^{n} a_{n,k} z^{-k} \quad \text{Eq. 27}$$

$$NTF_{cascade} = (1 - z^{-1})^{\left\lfloor \frac{n}{2} \right\rfloor} \cdot (1 - z^{-1})^{\left\lceil \frac{n}{2} \right\rceil}$$

$$= \left(1 + \sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} a_{\left\lfloor \frac{n}{2} \right\rfloor, k} z^{-k}\right)\left(1 + \sum_{k=1}^{\left\lceil \frac{n}{2} \right\rceil} a_{\left\lceil \frac{n}{2} \right\rceil, k} z^{-k}\right) \quad \text{Eq. 28}$$

where n is the order of NTF, and $a_{n,k}$ are the binomial coefficients with variation, i.e.,:

$$a_{n,k} = (-1)^k \cdot \binom{n}{k} + \mathcal{N}(0, \sigma_{n,k}^2) \quad \text{Eq. 29}$$

SNR is evaluated as:

$$SNR = \frac{K_1}{\frac{E_S^2}{OSR} + \frac{E_Q^2}{\pi}\int_0^{\frac{\pi}{OSR}} |NTF(jw)|^2 dw + \frac{E_N^2}{\pi}\int_0^{\frac{\pi}{OSR}} |1 - NTF(jw)|^2 dw} \quad \text{Eq. 30}$$

where K_1 is a constant related to the reference voltage. E_S^2, E_Q^2 and E_N^2 are the noise (error) from sampling, quantization, and dominant thermal noise, respectively. CDAC mismatch is ignored in this model. The worst-case SNR, i.e., SNR_{worst}, is calculated by:

$$\mathcal{P}(SNR > SNR_{worst}) = yield \quad \text{Eq. 31}$$

based on the Monte-Carlo method and distribution fitting.

Power is the core power plus filter power:

$$P = P_{core} + P_{filter} = BW \cdot OSR\left(\frac{K_2}{\sqrt{E_Q^2}} + \frac{K_3}{E_N^2}\right) \quad \text{Eq. 32}$$

where K_2 and K_3 are related to FoM_W and the process.

Area is evaluated as the CDAC area plus filter area:

$$A = A_{CDAC} + A_{filter} = \frac{K_4}{E_S^2}(1 + K_5) \quad \text{Eq. 33}$$

where K_4 and K_5 are related to capacitance density and the ratio between CDAC and filter caps.

Lastly, the variation of NTF coefficients is related to the area of filter cap:

$$\sigma_{n,k}^2 = K_6^2 \cdot \frac{2^n}{A_{filter}} \cdot \left|\overline{a_{n,k}}\right| \quad \text{Eq. 34}$$

where K_6^2 is the variance of a unit area of capacitance, and $\frac{2^n}{A_{filter}}$ is the reciprocal of FIR filter unit capacitor's area. (2^n is the sum of all FIR coefficients)

The results in Fig. 9 are from solving an optimization problem:

maximize	FoM_{NSSAR}	
variable	$\{E_Q, E_S, E_N, OSR\}$	
subject to	$SNR_{worst} = 90$	(dB)
	$\frac{2^{-4}}{\sqrt{12}} \le E_Q \le \frac{2^{-11}}{\sqrt{12}}$	(V)
	$10^{-5} \le E_S, E_N \le 10^{-2}$	(V)
	$4 \le OSR \le 512$	

with the following presets:

$$K_1 = 0.125, \quad K_2 = 2.9 \times 10^{-14}, \quad K_3 = 1.3 \times 10^{-20}, K_4 = 1 \times 10^{-11}, \quad K_5 = 0.1,$$

$$K_6 = 2.5\% \ (per\ \mu m^2), \quad BW = 100K, \quad yield = 95\%$$

where K_{1-6} are based on process data and information from the ADC survey [10]. Since the objective function is stochastic, the optimization is done by first running the Surrogate Algorithm on the full parameter space, followed by running a localized Pattern Search Algorithm. The optimization is repeated 10 times, and the result with minimal OSR is picked as the final result.

B. LUT Based Linearity Post-Calibration for Noise-Shaping Converter

A large part of ADC's nonlinearity comes from the discrepancy of its DC characteristic curve to an ideal one. Fig. 100 shows an example of such nonlinearity. This kind of nonlinearity is independent of the input signal, and thus is possible to be canceled out by digital post-processing.

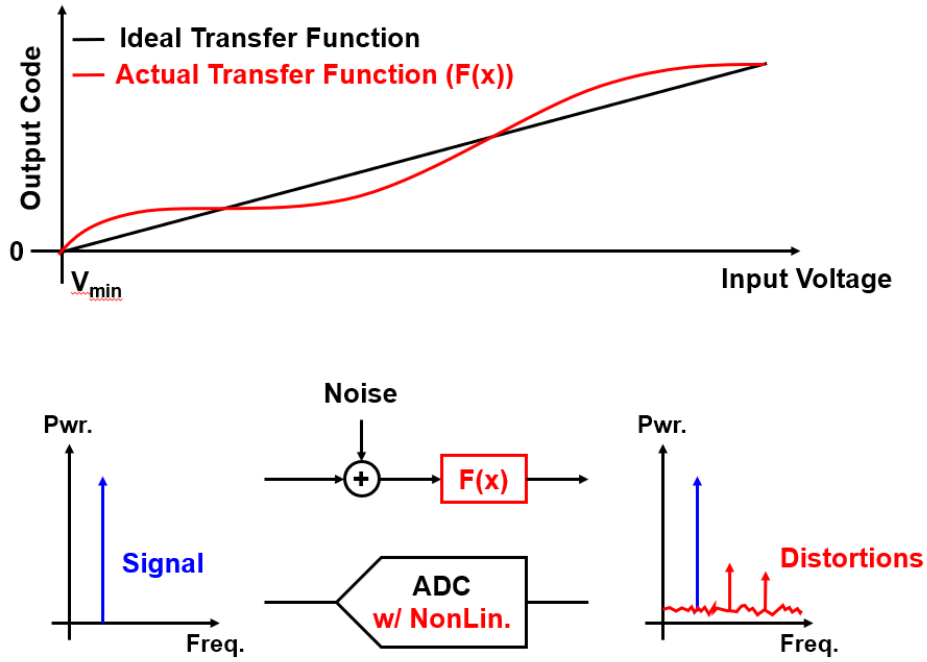


Fig. 100 The nonlinearity of ADC's characteristic curve (top) and the distortions induced (bottom)

A Look-Up Table (LUT) can generate an inverse characteristic function and re-map the ADC's output code to its correct value, as illustrated in Fig. 101. In an ideal case, the LUT elements should be with infinite accuracy so that they can completely cancel out the nonlinearity from ADC. But in practice, the elements are with finite precision. To provide a desire post-calibration performance, a rule of thumb is to build the LUT with at least N bits of binary digits, where N equal to the desired ENoB+1.

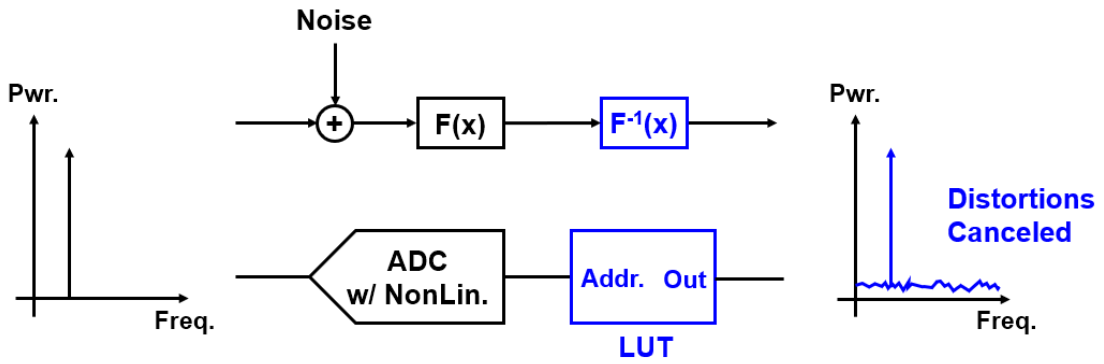


Fig. 101 LUT based post-calibration cancels out the distortions caused by ADC's nonlinearity.

There are many different methods to generate the LUT for post-calibration. One simple but effective way is the LSM algorithm, described as follows: the device under calibration runs as normal operation, and a high-quality, property-known sinewave is inputted to the device. Since the sinewave is completely known, we can calculate the ideal quantized output of this sinewave signal. Therefore, the LUT should map the real device output to this ideal one, i.e., assign the element addressed at the actual output code with the ideal code. In practice, the sinewave should be low-frequency to make phase-misalignment and AC nonlinearity negligible. Besides, the generation process is usually run with many conversion cycles. The values assigned to each LUT element are often sufficiently filtered so that the affection of noise is minimized. Fig. 102 shows the block diagram of the LUT generation method above.

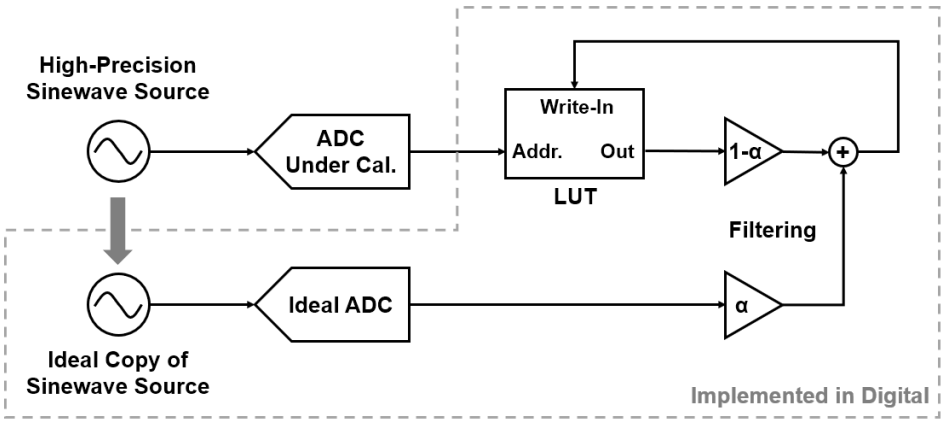


Fig. 102 A practical LUT generation method based on sinewave input testing.

However, there will be an issue if we directly apply the LUT calibration method to a NS converter: the out-of-band quantization noise will be mixed back to in-band by the LUT, hurting the in-band SNR. To solve this problem, we introduce a modified LUT calibration method, as shown in Fig. 103.

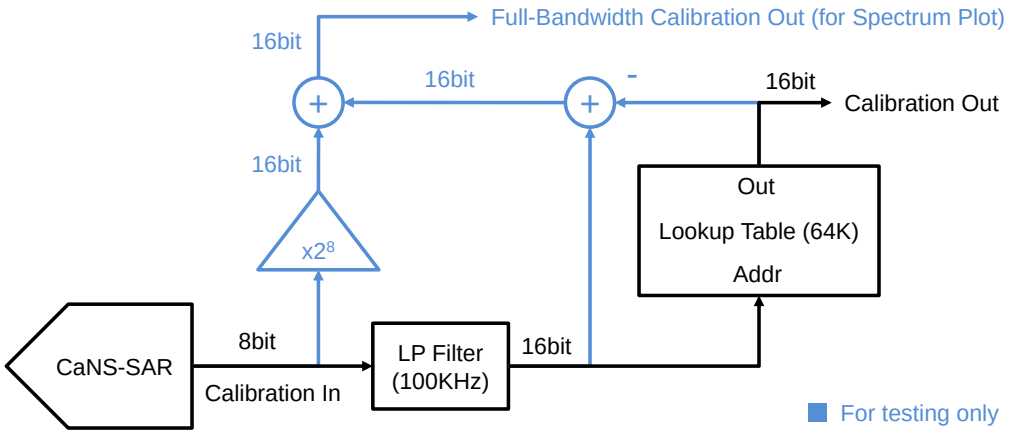


Fig. 103 LUT based calibration system (applied on the NS SAR in section 2.3 as example)

The key is to use a low-pass filter to filter the ADC's output (i.e., Calibration In). The filter bandwidth matches the ADC bandwidth and provides enough attenuation so that out-of-band quantization errors become negligible. The truncated outputs (16-bit here as an example) of the filter provide the address input to the LUT. Such that the LUT will not down-mix any out-of-band quantization noise.

Bibliography

[1] R. H. Walden, "Analog-to-digital converter survey and analysis," in IEEE Journal on Selected Areas in Communications, vol. 17, no. 4, pp. 539-550, April 1999.

[2] S. Pavan, R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters, 2nd ed., Hoboken, New Jersey: John Wiley & Sons, Inc., 2017, pp. 19-21.

[3] J. Sauerbrey, T. Tille, D. Schmitt-Landsiedel and R. Thewes, "A 0.7-V MOSFET-only switched-opamp ΣΔ modulator in standard digital CMOS technology," in IEEE Journal of Solid-State Circuits, vol. 37, no. 12, pp. 1662-1669, Dec. 2002.

[4] H. Nyquist, "Certain Factors Affecting Telegraph Speed," Bell System Technical Journal, vol. 3, April 1924, pp. 324–346.

[5] H. Inose, Y. Yasuda and J. Murakami, "A Telemetering System by Code Modulation - Δ-Σ Modulation," in IRE Transactions on Space Electronics and Telemetry, vol. SET-8, no. 3, pp. 204-209, Sept. 1962.

[6] S. Pavan, R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters, 2nd ed., Hoboken, New Jersey: John Wiley & Sons, Inc., 2017, pp. 104-113.

[7] J. A. Fredenburg and M. P. Flynn, "A 90-MS/s 11-MHz-Bandwidth 62-dB SNDR Noise-Shaping SAR ADC," in IEEE Journal of Solid-State Circuits, vol. 47, no. 12, pp. 2898-2904, Dec. 2012.

[8] S. Li, B. Qiao, M. Gandara, D. Z. Pan and N. Sun, "A 13-ENOB Second-Order Noise-Shaping SAR ADC Realizing Optimized NTF Zeros Using the Error-Feedback Structure," in IEEE Journal of Solid-State Circuits, vol. 53, no. 12, pp. 3484-3496, Dec. 2018

[9] S. Pavan, R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters, 2nd ed., Hoboken, New Jersey: John Wiley & Sons, Inc., 2017, pp. 60.

[10] B. Murmann, "ADC Performance Survey 1997-2020," [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html.

[11] M. Hovin, A. Olsen, T. S. Lande, and C. Toumazou, "Delta-sigma modulators using frequency-modulated intermediate values," IEEE J. SolidState Circuits, vol. 32, no. 1, pp. 13–22, 1997.

[12] B. F. Wu and E. A. Jonckheere, "A simplified approach to Bode's theorem for continuous-time and discrete-time systems," in IEEE Transactions on Automatic Control, vol. 37, no. 11, pp. 1797-1802, Nov. 1992.

[13] W. Lee, A Novel Higher Order Interpolative Modulator Topology for High Resolution Oversampling A/D Converters. Ph.D. dissertation, Massachusetts Institute of Technology, 1987.

[14] S. Kidambi, "On the Design of Optimal Noise Transfer Functions for Delta–Sigma Modulators," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 66, no. 3, pp. 392-396, March 2019.

[15] J. G. Proakis, D. G. Manolakis, “Analysis of Sensitivity to Quantization of Filter Coefficients” in Digital Signal Processing, 3rd ed., Upper Saddle River, NJ, USA: Prentice-Hall, 1996, pp. 569-578

[16] T. Kim, C. Han and N. Maghari, "An 11.4mW 80.4dB-SNDR 15MHz-BW CT delta-sigma modulator using 6b double-noise-shaped quantizer," 2017 IEEE International Solid-State Circuits Conference, San Francisco, CA, 2017, pp. 468-469

[17] T. Lo, C. Weng, H. Hsieh, Y. Shu and P. Chiu, "An 8x-OSR 25MHz-BW 79.4dB/74dB DR/SNDR CT Δ σ Modulator Using 7b Linearized Segmented DACs with Digital Noise-Coupling-Compensation Filter in 7nm FinFET CMOS," 2019 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 2019, pp. 334-336

[18] Lee K., Temes G.C. (2011) Noise-Coupled Delta-Sigma ADCS. In: Casier H., Steyaert M., van Roermund A. (eds) Analog Circuit Design. Springer, Dordrecht.

[19] L. Jie, B. Zheng, H. Chen, R. Wang and M. P. Flynn, "A 4th-Order Cascaded-Noise-Shaping SAR ADC with 88dB SNDR Over 100kHz Bandwidth," 2020 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 2020, pp. 160-162

[20] Y. Lin, C. Lin, S. Tsou, C. Tsai and C. Lu, "A 40MHz-BW 320MS/s Passive Noise-Shaping SAR ADC With Passive Signal-Residue Summation in 14nm FinFET," 2019 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 2019, pp. 330-332

[21] S. M. Chen and R. W. Brodersen, "A 6-bit 600-MS/s 5.3-mW Asynchronous ADC in 0.13um CMOS," in IEEE Journal of Solid-State Circuits, vol. 41, no. 12, pp. 2669-2680, Dec. 2006.

[22] X. Tang et al., "A 13.5b-ENOB Second-Order Noise-Shaping SAR with PVT-Robust Closed-Loop Dynamic Amplifier," 2020 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 2020, pp. 162-164.

[23] J. Liu, X. Wang, Z. Gao, M. Zhan, X. Tang and N. Sun, "A 40kHz-BW 90dB-SNDR Noise-Shaping SAR with 4× Passive Gain and 2nd-Order Mismatch Error Shaping," 2020 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 2020, pp. 158-160

[24] W. Guo, H. Zhuang and N. Sun, "A 13b-ENOB 173dB-FoM 2nd-order NS SAR ADC with passive integrators," 2017 Symposium on VLSI Circuits, Kyoto, 2017, pp. C236-C237

[25] P. Vogelmann, M. Haas and M. Ortmanns, "A 1.1mW 200kS/s incremental ΔΣ ADC with a DR of 91.5dB using integrator slicing for dynamic power reduction," 2018 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 236-238.

[26] C. Lee, M. H. Jang and Y. Chae, "A 1.2V 68μW 98.2DB-DR Audio Continuous-Time Delta-Sigma Modulator," 2018 IEEE Symposium on VLSI Circuits, Honolulu, HI, 2018, pp. 199-200.

[27] N. Kurosawa, K. Maruyama, H. Kobayashi, H. Sugawara and K. Kobayashi, "Explicit formula for channel mismatch effects in time-interleaved ADC systems," Proceedings of the

17th IEEE Instrumentation and Measurement Technology Conference, Baltimore, MD, USA, 2000, pp. 763-768 vol. 2.

[28] L. Jie, B. Zheng and M. P. Flynn, "A Calibration-Free Time-Interleaved Fourth-Order Noise-Shaping SAR ADC," in IEEE Journal of Solid-State Circuits, vol. 54, no. 12, pp. 3386-3395, Dec. 2019.

[29] M. Hesener, T. Eicher, A. Hanneberg, D. Herbison, F. Kuttner and H. Wenske, "A 14b 40MS/s Redundant SAR ADC with 480MHz Clock in 0.13pm CMOS," 2007 IEEE International Solid-State Circuits Conference. Digest of Technical Papers, San Francisco, CA, 2007, pp. 248-600.

[30] H. T. Jensen and I. Galton, "A low-complexity dynamic element matching DAC for direct digital synthesis," in IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 45, no. 1, pp. 13-27, Jan. 1998, doi: 10.1109/82.659453.

[31] Y. Lin, C. Tsai, S. Tsou, R. Chu and C. Lu, "A 2.4-mW 25-MHz BW 300-MS/s passive noise shaping SAR ADC with noise quantizer technique in 14-nm CMOS," 2017 Symposium on VLSI Circuits, Kyoto, 2017, pp. C234-C235.

[32] T. He, M. Ashburn, S. Ho, Y. Zhang and G. Temes, "A 50MHZ-BW continuous-time ΔΣ ADC with dynamic error correction achieving 79.8dB SNDR and 95.2dB SFDR," 2018 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 230-232.

[33] A. Jain and S. Pavan, "A 13.3 mW 60 MHz bandwidth, 76 dB DR 6 GS/s CTΔΣM with time interleaved FIR feedback," 2016 IEEE Symposium on VLSI Circuits (VLSI-Circuits), Honolulu, HI, 2016, doi: 10.1109/VLSIC.2016.7573466.

[34] S. Pavan, R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters, 2nd ed., Hoboken, New Jersey: John Wiley & Sons, Inc., 2017, pp. 233-234.

[35] S. Pavan, R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters, 2nd ed., Hoboken, New Jersey: John Wiley & Sons, Inc., 2017, pp. 259-270.

[36] L. Jie, H. Chen, B. Zheng and M. P. Flynn, "A 100MHz-BW 68dB-SNDR Tuning-Free Hybrid-Loop DSM with an Interleaved Bandpass Noise-Shaping SAR Quantizer," 2021 IEEE International Solid-State Circuits Conference. (publish pending)

[37] A. Krukowski and I. Kale, Frequency Transformations, in DSP System Design, Boston, MA: Springer, 2004.

[38] S. A. Jantzi, K. W. Martin and A. S. Sedra, "Quadrature bandpass ΔΣ modulation for digital radio," in IEEE Journal of Solid-State Circuits, vol. 32, no. 12, pp. 1935-1950, Dec. 1997.

[39] S. Huang, N. Egan, D. Kesharwani, F. Opteynde and M. Ashburn, "28.3 A 125MHz-BW 71.9dB-SNDR VCO-based CT ΔΣ ADC with segmented phase-domain ELD compensation in 16nm CMOS," 2017 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2017, pp. 470-471.

[40] W. Wang, C. Chan, Y. Zhu and R. P. Martins, "20.7 A 72.6dB-SNDR 100MHz-BW 16.36mW CTDSM with Preliminary Sampling and Quantization Scheme in Backend Subranging QTZ," 2019 IEEE International Solid- State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2019, pp. 340-342.

[41] B. Hershberg, S. Weaver, K. Sobue, S. Takeuchi, K. Hamashita and U. Moon, "Ring amplifiers for switched-capacitor circuits," 2012 IEEE International Solid-State Circuits Conference, San Francisco, CA, 2012, pp. 460-462.