

Chapter 2

Time-interleaved Track and Holds

2.1 Introduction

This chapter describes the time-interleaved Track and Hold (T&H) for the use in a time-interleaved ADC. In Fig. 2.1 an implementation of a time-interleaved ADC is shown consisting of several channels, each with a T&H section and a sub-ADC. The sample-rate of an interleaved ADC is N times the sample-rate of a sub-ADC, with N the number of channels. The main benefit of the time-interleaved architecture is that the overall sample-rate can be very high, while the sub-ADCs only need a moderate sample-rate, enabling a high power efficiency.

For $N = 16$, an example of a corresponding timing diagram is shown in Fig. 2.2. At each falling edge of the master-clock (MCLK), one of the T&Hs goes from track-mode to hold-mode and takes a sample of the input-signal. For a master-clock with sample-rate f_S and a number of channels N , each T&H and sub-ADC has a sample-rate of f_S/N .

Making a time-interleaved ADC involves more than just placing a few non-interleaved ADCs in parallel, since the requirements for a non-interleaved T&H and ADC differ from that of a time-interleaved T&H and sub-ADC: Aspects like offset, gain error and absolute timing, which are usually not an issue for a general-purpose non-interleaved T&H and ADC, are important for a time-interleaved architecture, as will be explained in this chapter. Moreover, the Nyquist frequency of a time-interleaved architecture is N times higher than that of a non-interleaved ADC, so the T&Hs should have a much higher bandwidth and should be able to sample signals with an N times higher frequency.

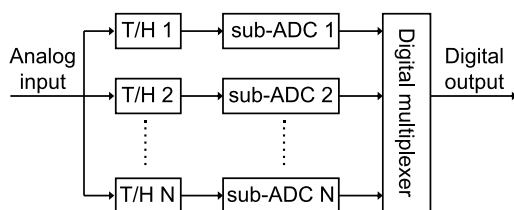
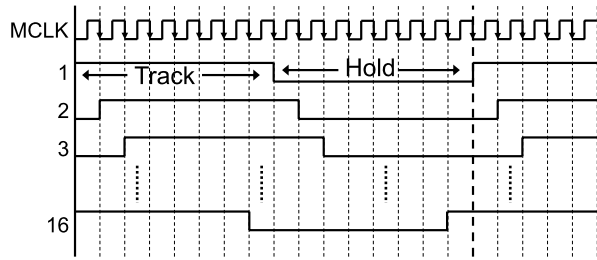


Fig. 2.1 The time-interleaved ADC architecture

Fig. 2.2 Timing diagram of a time-interleaved ADC



In this chapter, aspects associated with high-speed time-interleaved T&Hs will be discussed, starting with matching aspects between channels in Sect. 2.2, followed by the description of time-interleaved T&H architectures in Sect. 2.3. In Sect. 2.4 T&H buffers will be treated, and Sect. 2.5 discusses the use of bottom-plate sampling in a time-interleaved T&H. This is followed by a discussion on the optimum number of channels in Sect. 2.6. The chapter ends with aspects associated with calibration in Sect. 2.7 and jitter requirements in Sect. 2.8.

2.2 Mismatch Between Channels

As stated above, the requirements for a non-interleaved general-purpose ADC for offset, gain and timing (e.g. the delay from the sample-clock to the actual sample moment) are usually not strict. As long as they are constant, they do not affect the Signal-to-Noise-and-Distortion Ratio (SNDR), Integral Non-Linearity (INL) and Differential Non-Linearity (DNL). For a time-interleaved ADC consisting of multiple channels, the situation is different. Differences in e.g. offset, gain or timing/phase between channels cause spurious tones [8, 21]. Offset mismatch causes distortion tones at multiples of f_S/N , while mismatch in gain or timing results in tones at multiples of $f_S/N \pm f_{IN}$. In Fig. 2.3 the spectrum of a reconstructed sinusoid is shown for $N = 8$ and band-limited to $f_S/2 + f_{IN}$, with f_{IN} the frequency of the input signal. The upper part shows the case where only offset mismatch is present and the lower part shows the effect of gain or phase mismatch. The amplitude of the spurious tones depends on the offset/gain distribution of the channels and on the number of channels: for a larger number of channels, the error energy is divided between more tones, so the amplitude per tone decreases.

2.2.1 Origin of Spurious Tones

To give insight in the origin of spurious tones, three graphical examples are given for a time-interleaved T&H with two channels. An analytical analysis is given in [21]. Channel offset is analyzed first. In the upper part of Fig. 2.4, the sampled output of the two channels is shown. From this, a common signal and a difference signal can

Fig. 2.3 Spectrum of a reconstructed sinusoid for a time-interleaved ADC with 8 channels and mismatch in (a) offset and (b) gain or phase

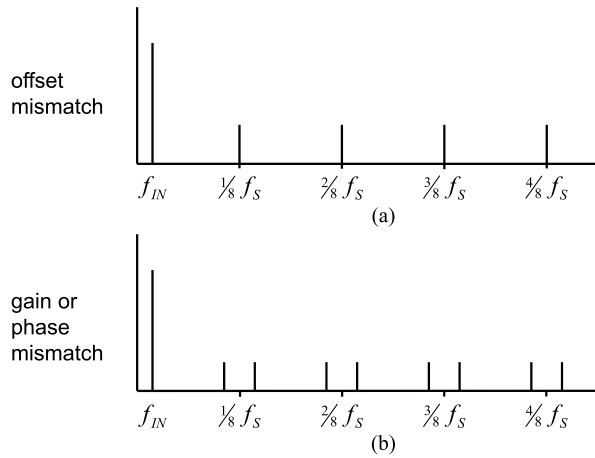
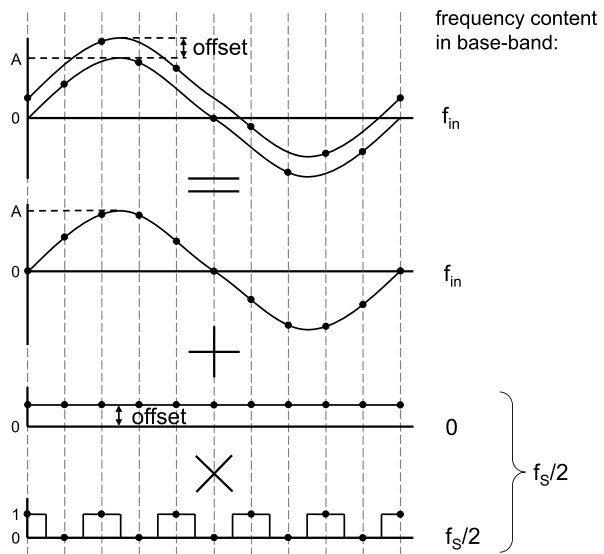


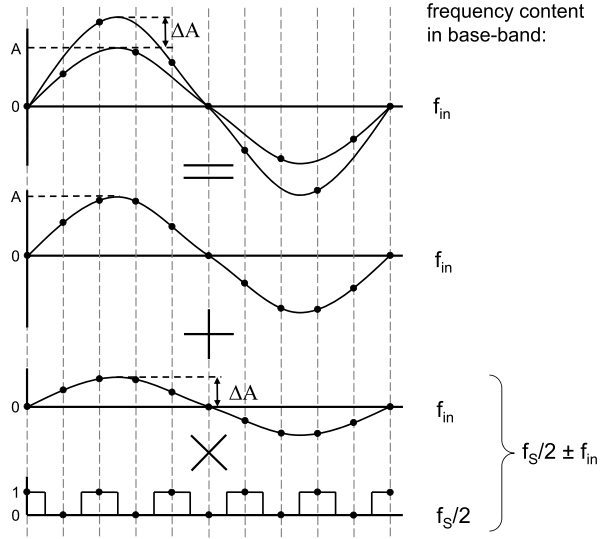
Fig. 2.4 Visualization of frequency content in the case of offset between 2 channels



be derived, see the rest of the figure. The common signal is one of the input signals and the difference signal is the offset multiplied by a square-wave with frequency $f_S/2$. In a spectrum limited to the Nyquist frequency, this results in a tone at $f_S/2$. For a larger number of channels, the situation is similar, however the number of square-waves (tones) is higher [21]. The tones are static and do not depend on the input signal.

In the case of gain mismatch between two channels, the situation is illustrated in Fig. 2.5. The common signal is again one of the input signals and the difference signal is the difference between the input signals multiplied by a square-wave with

Fig. 2.5 Visualization of frequency content in the case of gain mismatch between 2 channels



frequency $f_s/2$. Due to the multiplication, sum and difference frequencies arise:

$$\begin{aligned} \sin\left(2\pi \frac{f_s}{2} t\right) \cdot \sin(2\pi f_{IN} t) &= \frac{1}{2} \sin\left(2\pi \left(\frac{f_s}{2} + f_{IN}\right) t\right) \\ &+ \frac{1}{2} \sin\left(2\pi \left(\frac{f_s}{2} - f_{IN}\right) t\right) \end{aligned} \quad (2.1)$$

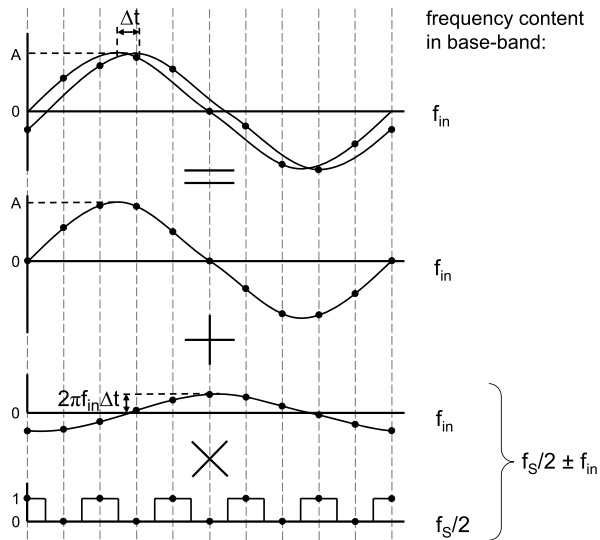
Note that due to aliasing higher-order products of the square-wave¹ result in the same frequencies as the fundamental of the square-wave ($f_s/2$). In summary, gain differences between channels result in scaled copies of the input spectrum around $f_s/2$. And again for more channels, more scaled copies of the input spectrum appear.

Finally, timing-misalignment between channels is considered. The signals are shown in Fig. 2.6 and are similar to the case of gain mismatch, except that the phase of the difference signal is shifted by 90° . The resulting spectrum has the same frequency content. It is easy to understand that phase-differences between channels have the same effects as timing misalignment, since for a sinusoid holds: $\Delta\varphi = 2\pi f_{IN} \Delta t$.

Apart from those channel mismatches, other differences between channels e.g. differences in linearity or bandwidth will also degrade the performance. Bandwidth mismatch is discussed in the next section. The difference signal caused by linearity mismatch depends on the input signal, like in the case of gain mismatch, and it has therefore the same effect on the output spectrum.

¹In this example, only 2 samples per period are of importance, so the square-wave can also be replaced by a sine-wave.

Fig. 2.6 Visualization of frequency content in the case of timing mismatch between 2 channels



2.2.2 Bandwidth Mismatch

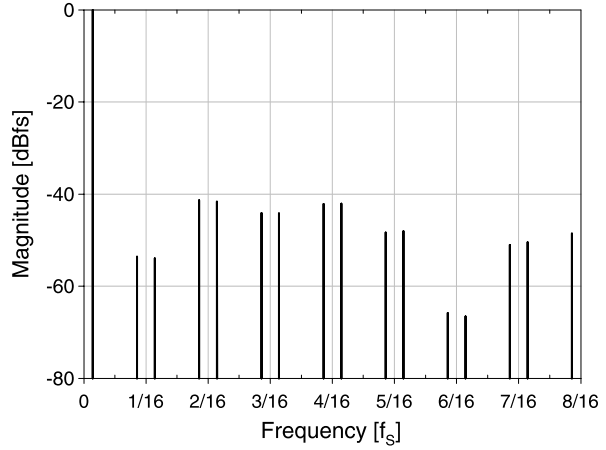
As discussed above, the performance of a time-interleaved ADC is affected by mismatch in offset and gain, and this is not dependent on the signal frequency. For bandwidth mismatch however, the performance degradation is dependent on the signal frequency. In this section, phenomena related to bandwidth mismatch are described and quantified.

Two different channel bandwidth limitations can be distinguished: (1) limitations common to all channels caused by the input resistance and interconnect capacitance and (2) per channel limitations caused by the sample capacitor and the resistance of the sample-switch and that of interconnect.

To get some feeling for the quantitative effects of bandwidth mismatch, a numerical example is given. This requires a few assumptions, which are taken from an implementation presented in Chap. 4. For an overview see Fig. 2.1. In this implementation, the SNR due to kT/C noise [50] needs to be about 10 bits, so for a differential system, a sample-capacitor of 150 fF is sufficient with a signal swing of 0.4 V. With a bandwidth requirement of 1 GHz, the switch resistance should be lower than 1 k Ω . In a 0.13 μm process, a switch with a geometry of 1 $\mu\text{m}/0.13 \mu\text{m}$ has a $\sigma(V_{GS})$ of 13.5 mV in which both V_T spread and spread in β (gain factor) contribute for about the same amount. Under typical bias conditions this leads to a $\sigma(R_{ON})/R_{ON}$ of 3.5%, so $\sigma(R_{ON})$ is 35 Ω .

Mismatch parameters about the interconnect resistance could not be found in literature. Devices like MOSTs ($I_{D,sat}$), active resistors [48] and back-end capacitors have more or less a standard deviation of around 1 % μm and this can also be assumed for interconnect resistance as long as its dimensions are above the minimum feature-size [58]. The interconnect in the example implementation is 170 μm long

Fig. 2.7 Typical simulated spectrum of a reconstructed sinusoid in the presence of bandwidth mismatch, with $N = 16$ and $\sigma(BW)/BW = 3.5\%$



and $0.4 \mu\text{m}$ wide (two times minimum width), and has a nominal resistance of 26Ω . The standard deviation of the resistance is approximately:

$$\sigma(R_{\text{intc}}) = \frac{R_{\text{intc}} \cdot 1\%}{\sqrt{170 \cdot 0.4}} = 0.12\% \cdot 26 \Omega = 0.03 \Omega \quad (2.2)$$

This is much smaller than the standard deviation of the switch resistance, and therefore the variation in interconnect resistance can be neglected.

Capacitor matching in modern CMOS processes is relatively good. The process technology used for the implementation, has capacitors with $\sigma(C_{\text{sample}})/C_{\text{sample}} = 0.03\%$. Taking the three relative standard deviations together leads to a $\sigma(BW)/BW$ of 3.5% , dominated by mismatch in the sample-switch resistance.

To demonstrate the effects of bandwidth mismatch on the spectrum, a 16-channel time-interleaved T&H is simulated in which bandwidth mismatch is the only error and the rest is assumed ideal. In this example $\sigma(BW)/BW = 3.5\%$ and the sample-rate is 1 GHz . The nominal channel bandwidth is also 1 GHz and the signal frequency is close to the sample-rate, such that its alias appears close to zero and the spurious tones appear near multiples of $f_s/16$, resulting in an orderly spectrum. The reconstructed spectrum is shown in Fig. 2.7 and has an SNDR of only 33 dB (or² 5.1 bits), which is much less than required for a 10 bits converter. When the input frequency is reduced to the Nyquist frequency, the SNDR slightly improves to 37 dB (5.8 bits).

In Fig. 2.8 the maximum achievable SNDR is shown as a function of $\sigma(BW)/BW$ for a 16-channel time-interleaved T&H. The input frequency is equal to half the nominal channel bandwidth f_0 , while the SNDR is independent of the sample-rate. For an SNDR of 10 bits and input frequencies up to half the nominal channel bandwidth, $\sigma(BW)/BW$ should be smaller than 0.2% . To improve the matching of R_{ON} from 3.5% to 0.2% by device scaling [38] (increasing both width and length), the

²The relation between the scales in dB and bits is: $\text{dB} \approx 6.02n + 1.76$.

Fig. 2.8 Achievable SNDR as a function of $\sigma(BW)/BW$ with $N = 16$ and $f_{IN} = \frac{1}{2}f_0$ (average SNDR of Monte-Carlo simulation)

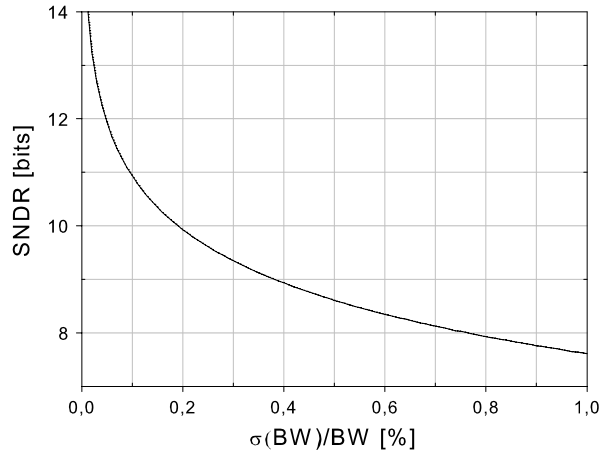
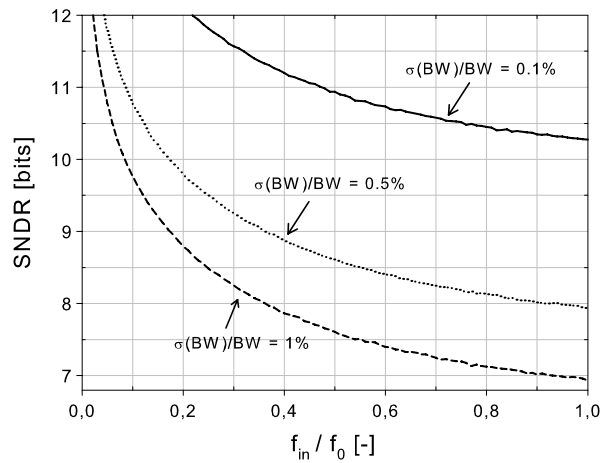


Fig. 2.9 SNDR as a function of the normalized signal frequency for different values of $\sigma(\Delta BW/BW)$



area should be increased $17.5^2 = 306$ times, which would lead to an unacceptably large switch. Scaling of the width only is discussed in the next section.

Performance Improvement by Increasing the Nominal Channel Bandwidth

The amplitude of the spurious tones depends on the ratio of the signal frequency and the nominal channel bandwidth. For a relatively large nominal channel bandwidth, bandwidth mismatch between channels has little impact. To demonstrate this effect, the achievable SNDR as a function of the signal frequency normalized to the nominal channel bandwidth f_0 is shown in Fig. 2.9 for different values of $\sigma(BW)/BW$. The system is equal to the one described above and the normalized frequency is the signal frequency divided by the nominal channel bandwidth.

From the above it becomes clear that instead of improving bandwidth matching, it is also possible to increase the channel bandwidth, such that for the frequencies of interest the gain matching and phase alignment is better. When increasing the width of the sample-switch, both the bandwidth and the matching of the bandwidth improve.³ For example, if a switch of 10/0.13 is used, $\sigma(V_T) = 4$ mV, $\sigma(\beta) = 0.9\%$, $\sigma(R_{ON}) = 1.1\%$ and the bandwidth is about 10 GHz. The achievable SNDR for input signals at 10 GHz is then only about 44 dB or 7 ENOB. However, for signal frequencies up to 1 GHz, an SNDR of 10 bits is achievable, without the need for bandwidth calibration.

Bandwidth Mismatch Split into Resulting Gain and Phase Mismatch

Bandwidth mismatch between channels causes frequency dependent differences in both gain and phase [10]. It is useful to distinguish between these two effects, and therefore a simulation result of a 16-channel time-interleaved T&H is shown in Fig. 2.10, when taking into account: (a) only gain errors due to bandwidth mismatch, (b) only phase errors due to bandwidth mismatch and (c) both errors. On the horizontal axis the normalized signal frequency f_{IN}/f_0 is shown, where f_0 is the nominal channel bandwidth. $\sigma(BW)/BW$ is 1%. For signal frequencies close to the nominal channel bandwidth, the errors caused by (bandwidth mismatch induced) gain mismatch and (bandwidth mismatch induced) phase mismatch degrade SNDR by the same amount. Towards lower frequencies the effect of gain mismatch decreases rapidly (increasing SNDR), while the effect of phase errors only decreases slowly. So, in conclusion, phase errors are dominant for relatively low input frequencies.

2.3 Time-interleaved Track and Hold Architectures

In this section two time-interleaved T&H architectures are discussed: the normal time-interleaved architecture without a frontend sampler and the time-interleaved architecture with a frontend sampler. Optional improvements on both architectures are discussed, and interleaving limits are discussed for both architectures in relation to bandwidth and accuracy. The section ends with a comparison of the architectures.

³The amount of channel charge dump mainly depends on the area of the transistor channel. Since sample switches usually have a large aspect ratio (large W , small L) e.g. 10/0.13, a small absolute variation in W has little impact, while the same absolute variation in L has a much larger impact. The relative mismatch in charge dump is therefore quite independent on W , resulting in an absolute mismatch proportional to W . So, increasing the switch width leads to an increase in the mismatch of the charge dump. For a bootstrapped sample-switch this results in increased offset mismatch. As offset calibration is often required for time-interleaved ADCs, this is not considered to be a problem.

Fig. 2.10 SNDR as a function of the normalized signal frequency with $\sigma(BW)/BW = 1\%$, when taking into account: only gain errors due to bandwidth mismatch, only phase errors due to bandwidth mismatch, both errors

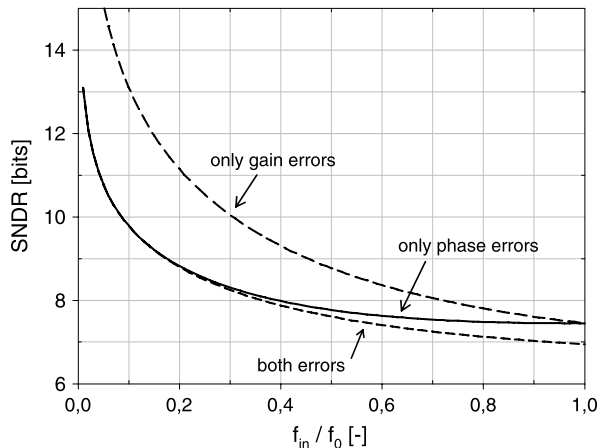


Fig. 2.11 The time-interleaved ADC architecture

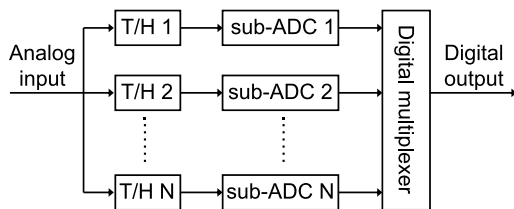
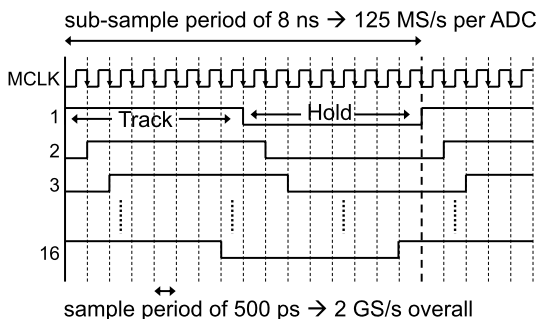


Fig. 2.12 Timing diagram of time-interleaved ADC with $N = 16$

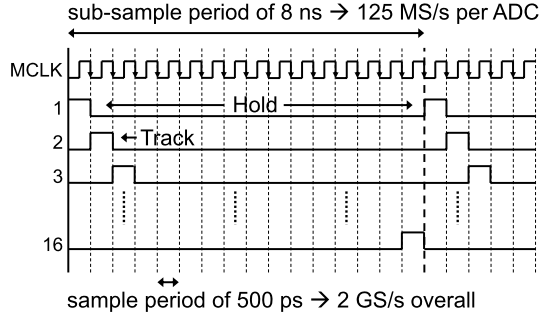


2.3.1 Architecture Without a Frontend Sampler

The most straightforward configuration of a time-interleaved T&H is shown in Fig. 2.11, where each sub-ADC has its own T&H circuit [8, 39]. The corresponding timing diagram is shown in Fig. 2.12. At each falling edge of the master-clock, one of the T&Hs goes from track-mode to hold-mode and takes a sample of the input-signal. This signal is then converted to the digital domain by the ADC in the same channel.

An important consideration is the input capacitance of the time-interleaved T&H. For high speed input signals often transmission lines with on-chip $50\ \Omega$ termination

Fig. 2.13 Timing diagram with track-time of 1 period and $N = 16$



are used to mitigate reflections. The resistance at the input node is therefore 25Ω : 50Ω of the on-chip termination parallel to 50Ω of the external source. With this fixed input resistance, the input capacitance determines the bandwidth.

If the resulting bandwidth is not large enough, an input buffer can be used to increase the bandwidth. The input capacitance of the T&H determines the power consumption of this buffer and for a very large capacitive load it can be unfeasible to drive it with sufficient bandwidth. Also, due to the high demands on this buffer (the combination of a high speed and a large capacitive load), it requires a lot of power. In [40] a front-stage buffer in SiGe technology is used that consumes 1 W of power to drive a 4 pF T&H load. The aim of the research described in this book, is to investigate low power solutions. Therefore, no input buffer is used, instead it is assumed that the T&H is driven by an external 50Ω source.

When the timing diagram of Fig. 2.12 is used, at each moment in time $N/2$ sample-capacitors are connected to the input. The input capacitance can be decreased by reducing the track-time. In Fig. 2.13 the timing diagram is shown for a track-time of one period of the master clock. In this case only one sample-capacitor is connected to the input at a time, lowering the input capacitance and enabling higher number of channels for a given bandwidth.

Now, it is calculated whether one period is sufficiently long to let the voltage on the sample capacitor settle sufficiently close to the input value. This clearly depends on the bandwidth of the sampler, determined by the combination of sample switch and capacitor. For a limited attenuation of the input signal at the Nyquist frequency ($f_S/2$), a sampler bandwidth of two times the Nyquist frequency is assumed, so $BW_{\text{sampler}} = f_S$. The time-constant of the sampler is thus:

$$\tau_{\text{sampler}} = \frac{1}{2\pi f_S} \quad (2.3)$$

There is one period used for settling, so:

$$T_S = \frac{1}{f_S} = (n + 1) \cdot \tau_{\text{sampler}} \cdot \ln(2) \quad (2.4)$$

with n the resolution in bits.⁴ Combining (2.3) and (2.4) yields:

$$n = \frac{2\pi}{\ln(2)} - 1 \approx 8 \quad (2.5)$$

So, in the case the sampler bandwidth is equal to f_S , settling is accurate up to a resolution of 8 bits.

In the previous section it was argued that a large channel bandwidth is advantageous to improve matching. For a bandwidth larger than f_S , one period of tracking is also enough for resolutions of more than 8 bits.

A short track-time implies a long hold-time, which is advantageous in most ADC architectures, as the ADC has more time to do the conversion.

Resetting of the Sample Capacitor

In the above calculation, it was implicitly assumed that when the T&H entered track-mode, an unknown value of the previous sample action was still present on the sample capacitor. To prevent inter-symbol interference, the T&H needs to settle to the full accuracy, as calculated above.

It is possible to use a reset switch, to remove the previous sampled signal from the sample capacitor, before going into track-mode. Assuming the sample process is linear, incomplete settling only leads to attenuation, and inter-symbol interference does not occur.

To exploit the advantage of a reset switch however, the settling time should be constant. So, not only should the track-to-hold moment be defined well, also the hold-to-track moment should be defined well. This increases the complexity of the T&H circuit significantly. As a result, a reset switch should only be used when the T&H settling-time requirement can not be fulfilled easily.

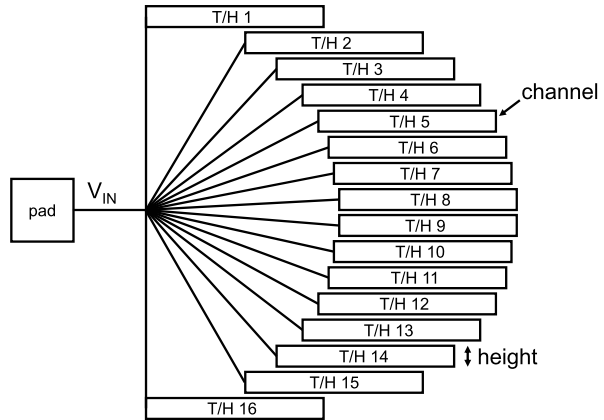
Input Capacitance

Besides one or more sample capacitors, the wiring and the sample-switches also contribute to the input capacitance. For a (half) circular layout, see Fig. 2.14, the capacitance of the sample-switches is proportional to N , while the wiring capacitance is proportional to N^2 : For a fixed channel height, both the number of channels and the wiring length are proportional to N . An approximation for the total input capacitance when using one clock period for tracking is:

$$C_{IN} = C_{\text{sample}} + N \cdot C_{\text{switch}} + N^2 \cdot C_{\text{wire}} \quad (2.6)$$

⁴A derivation is given in Sect. 3.2.2 starting at p. 42. However, a step-size of half the range is assumed there, while here the step-size equals the entire range for an input signal at the Nyquist frequency. Therefore, n needs to be replaced by $n + 1$.

Fig. 2.14 The semi-circular layout for minimizing bandwidth differences



where C_{switch} is the switch capacitance and C_{wire} is the capacitance of a wire of a certain length. The length of the wires depends on the configuration. To give some actual numbers, an example is introduced taken from an actual implementation [24], presented in Chap. 4. In this example a semi-circular layout is used as shown in Fig. 2.14. This layout is chosen in order to make the bandwidth for all channels equal. This could also be accomplished by a full-circle layout, but it has three disadvantages: (1) the distance from the bond-pads to the middle of the circle is longer, resulting in a lower input bandwidth, (2) routing the clock and input signals from the bond-pad to the middle of the circle and shielding these, requires a few metal layers, so less layers are available, and (3) the placement of the ADCs is less practical.

The wire-length depends on the channel height, which has a lower limit for practical reasons. The implementation resulted in T&H blocks of $20\text{ }\mu\text{m}$ by $300\text{ }\mu\text{m}$ and this height is assumed here for determining the wire capacitance.

To get some feeling for the total input capacitance, the values for the capacitances will be approximated and are again taken from the same implementation. C_{sample} has a value of 150 fF , limiting the achievable ENOB to 10.4 bits for a peak-to-peak signal swing of 0.4 V due to kT/C noise. R_{ON} of the switch is $150\text{ }\Omega$, and together with the interconnect resistance of $26\text{ }\Omega$, this results in a bandwidth of 6 GHz . This bandwidth is chosen to be large for reasons explained in Sect. 2.2.2 on p. 9. This value of R_{ON} is reached with an NMOST switch in $0.13\text{ }\mu\text{m}$ CMOS technology for a switch width of $10\text{ }\mu\text{m}$. The total switch capacitance at the source node is 15 fF , assuming the switch is off, since all but one switches are off. The capacitance of a wire depends on its width and length. When using a minimum width of $0.2\text{ }\mu\text{m}$, a 3D EM-field simulation shows a capacitance of 0.12 fF per μm length. Assuming the structure and block-height discussed above results in a C_{wire} of 1.3 fF , which can be used in (2.6).

The input capacitances and the input bandwidth can now be derived. In Fig. 2.15 the various capacitances are plotted as function of the number of channels N for: the sample-capacitors, the sample-switches, the wires and the sum of these three. On the right y-axis, the input bandwidth is shown. Above 20 channels, the wiring

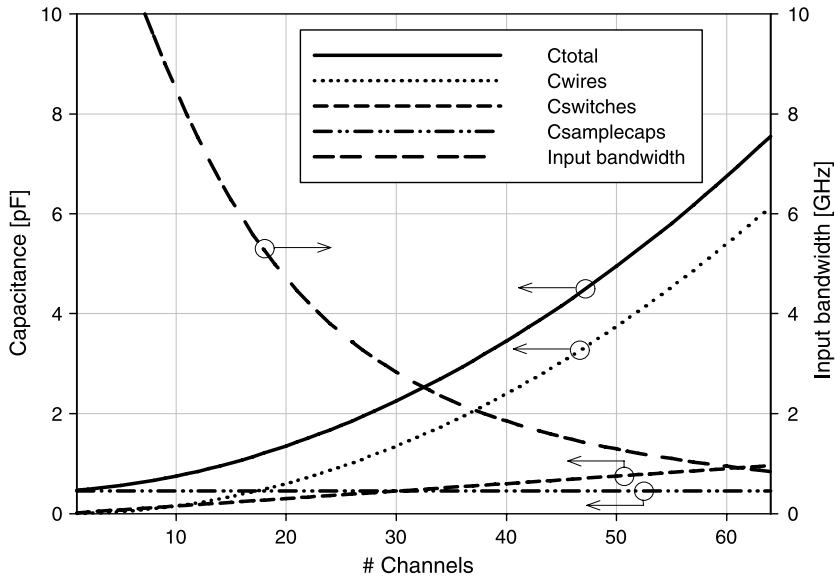
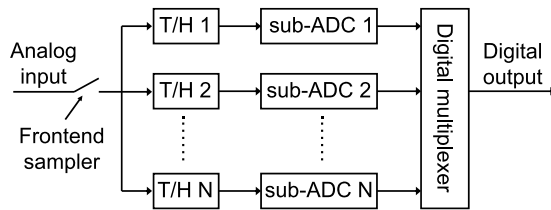


Fig. 2.15 Capacitances and input bandwidth as a function of the number of channels N

Fig. 2.16 Time-interleaved ADC architecture with frontend sampler



dominates the total capacitance. For an input bandwidth of 2 GHz, the number of channels is limited to about 40. For 6 GHz of bandwidth, N should not exceed 16.

2.3.2 Architecture with a Frontend Sampler

In the architecture discussed in the previous section, each channel has its own T&H. Mismatch between these T&Hs results in timing-misalignment of the sample-moments. Depending on the signal frequency, this degrades the SNDR. In this section an architecture is described that does not have this disadvantage.

To avoid timing-misalignment between channels, a frontend sampler (FRS) [16] can be added to the conventional architecture as shown in Fig. 2.16. The essence of this architecture is that the frontend sampler determines all sampling moments, avoiding timing-misalignment.

Fig. 2.17 Timing diagram of a time-interleaved ADC with frontend sampler and $N = 16$

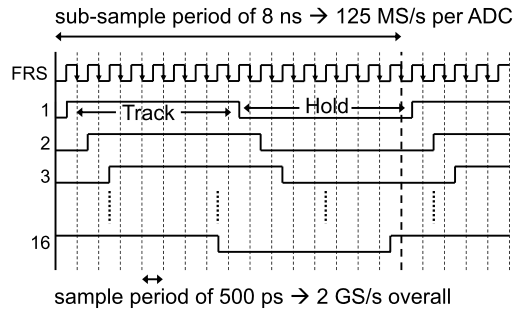
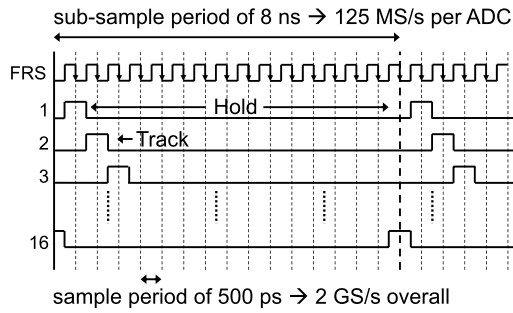


Fig. 2.18 Timing diagram with track-time of 1 period and frontend sampler

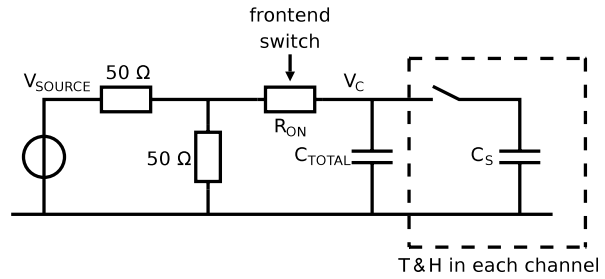


In the timing diagram of Fig. 2.17 the timing of the frontend sampler (FRS) is shown and it equals the master clock of the architecture without a frontend sampler. The timing of the T&Hs is also equal to this architecture with the exception the T&Hs are now delayed by half a clock-period of the master clock. This way the frontend sample switch opens first and determines the sample moment. The timing of the T&H sample-switches is therefore not critical and distortion tones at multiples of $f_S/N \pm f_{IN}$ due to timing-misalignment between channels [8] are avoided.

To reduce the input capacitance and to increase the conversion-time available for the ADC, the track-time can be reduced to one clock period, in the same way as in the conventional architecture as described in the previous section. The resulting timing diagram is shown in Fig. 2.18.

Without a frontend sampler the track-time can be made one or more periods. A disadvantage of the architecture with a frontend sampler is that the track-time is limited to about half a clock-cycle of the master clock, because the frontend sampler has to operate at the full sample-rate. The track-time can be slightly increased by using a clock with a duty-cycle larger than 50%, but it can never reach a full clock period, as the sample-switch in the channel has to be opened while the frontend switch is still open. Ensuring that the clocks are non-overlapping at high sample-rates, takes a significant part of the sample-period.

Fig. 2.19 Schematic for calculation of the bandwidth and settling-time requirements



Input Bandwidth and Settling-time Requirements

The input capacitance as calculated in the previous section is also present in this architecture, however not at the input, but instead after the frontend sample-switch.

The requirements for the frontend switch will now be calculated based on bandwidth and settling-time. The bandwidth can be calculated using the schematic of Fig. 2.19. It is assumed that both the impedance of the signal source and the on-chip termination are $50\ \Omega$ and the resistance of the frontend switch is called R_{ON} . These resistors and capacitance C_{total} cause a first pole.

Each channel contains a T&H switch and a sample capacitor, indicated by the dashed box. These cause a second pole, which for practical implementation would be far away from the first pole, and therefore it is neglected.

So, assuming a first-order system, the bandwidth at node V_C is:

$$BW_{V_C} = \frac{1}{2\pi \cdot R_{eff} \cdot C_{total}} \quad (2.7)$$

with $R_{eff} = 50/2 + R_{ON}$ and C_{total} the total input capacitance of wires, sample switches and a sample capacitor as calculated in the previous section. Figure 2.20 shows the required switch resistance R_{ON} as a function of the number of channels N for various bandwidths. From this graph it becomes clear that for the example of an input bandwidth of 2 GHz and 16 channels, the switch resistance needs to be $50\ \Omega$ or less.

The settling-time requirement is calculated as follows: The signal should settle to the required accuracy within half the sample-period, again assuming a first-order system. For an example sample-rate of 2 GS/s and settling up to an accuracy of $1/2$ LSB at 8 bits level, 6.2τ of settling is required in 250 ps, so τ should be smaller than 40 ps. Using $\tau = R_{eff} \cdot C$, for each N the required R_{ON} can now be calculated. In Fig. 2.21 the required resistance of the frontend switch is plotted as a function of the number of channels for different accuracies.

For e.g. 16 channels and 10 bits of accuracy, the required switch resistance is $8\ \Omega$. For an NMOST in $0.13\ \mu\text{m}$ technology with $V_{GS} = 0.6\ \text{V}$ and $R_{ON} = 8\ \Omega$, the required switch width is $160\ \mu\text{m}$. Such a large sample switch has a large parasitic capacitance of a few hundred fF, which has two consequences: (1) It is hard to drive

Fig. 2.20 Required frontend sample-switch resistance as a function of the number of channels N for various bandwidths

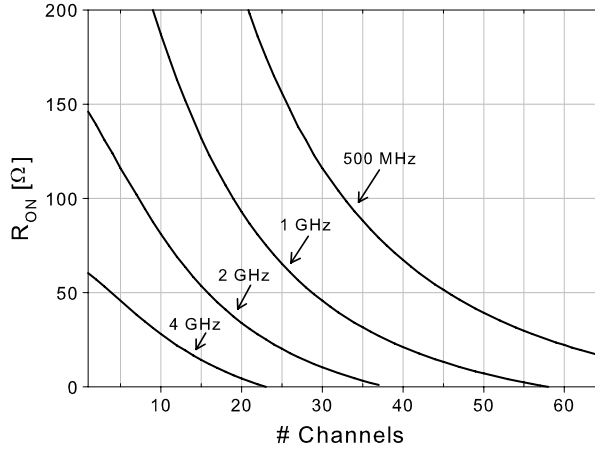
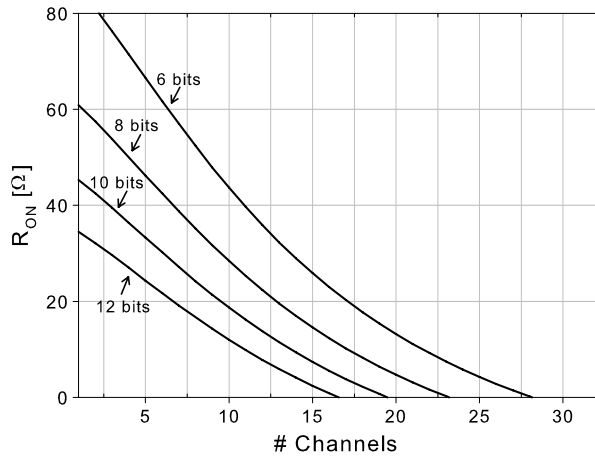


Fig. 2.21 Required frontend sample-switch resistance as a function of the number of channels N for various accuracies



the gate-node with a steep edge to make the sample process close to ideal, and (2) the sample-to-hold step becomes [50] unacceptably large. Under the above assumptions and an accuracy of 10 bits, the number of channels should be limited to about 5 for the frontend sampler architecture.

Note that when the frontend sampler is omitted, the settling time requirements are significantly relaxed: The series resistance of the switch is not there and the settling time can be twice as long, as explained in the previous section.

Increasing the Input Bandwidth

The main disadvantage of a frontend sampler is the decrease in bandwidth, due to the large capacitance of the wires and switches after the sampler. This capacitance can be decreased by using additional switches. An example is shown in Fig. 2.22,

Fig. 2.22 Architecture with a frontend sampler and additional switches to increase the bandwidth

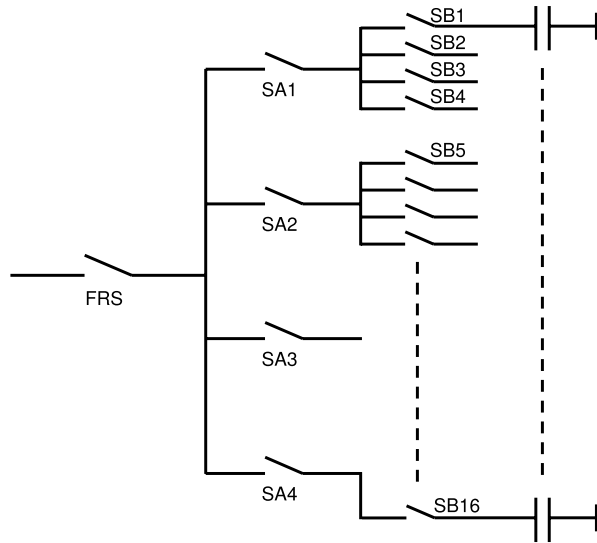
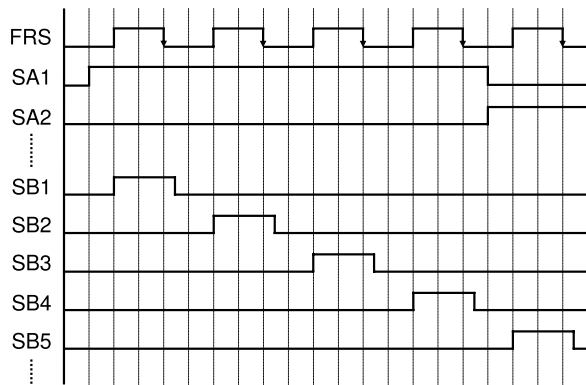


Fig. 2.23 Timing diagram of architecture with frontend sampler and additional switches



where 4 additional switches (SA1–SA4) are placed between the frontend sampler (FRS) and the T&H switches in the channels (SB1–SB16). The additional switches will also increase the resistance, however the net effect on the bandwidth can still be positive.

The corresponding timing diagram is shown in Fig. 2.23 and the operation is as follows: Suppose switches FRS, SA1 and SB1 are conducting, such that the first channel is in track mode. Then, FRS opens first and determines the sample moment. Next, SB1 opens and fixes the charge on the sample capacitor. After this, FRS and SB1 close and the second channel is in track-mode. After a sample period, FRS opens again followed by SB2 and so on.

When SB4 is opened (after FRS is opened), also SA1 is opened and SA2 is closed, such that the next four channels can take samples of the input signal.

The A-switches can be opened, after the B-switches are opened. Since the charge on the sample-capacitor is then already fixed, charge injection of the A-switches does not degrade the performance.

The advantage of this architecture is that timing misalignment is avoided, and the bandwidth is larger than without using additional switches. A disadvantage of this architecture is that the (in this example) four quarters of the circuit will have bandwidth mismatch due to spread in the A switches (w.r.t. R_{ON} and $C_{parasitic}$) and the capacitance of the wire and the B switches. This can limit the performance or require bandwidth calibration.

2.3.3 Conclusions on Architectures

In conclusion it can be said that the use of a frontend sampler has the advantage of good alignment of sampling moments between different channels, but that the product of bandwidth and accuracy is limited. To achieve moderate accuracy (8–10 bits) together with a high bandwidth (>1 GHz), the use of a frontend sampler is not viable under the assumptions made. The use of additional switches between the FRS and the T&H switches increases the bandwidth and accuracy, but can cause bandwidth differences between the groups of T&Hs sharing an intermediate switch.

2.4 Track and Hold Buffers

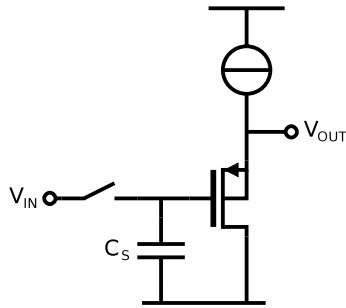
This section handles buffers for the use in a time-interleaved T&H. In a time-interleaved ADC multiple sub-ADCs operate in parallel, resulting in an N times higher sample-rate, with N the number of channels. If the ratio of the maximum input frequency and the sample-rate (e.g. $1/2$ for Nyquist operation) is kept constant, the T&Hs need to operate with N times higher input frequencies than if used non-interleaved.

To achieve good linearity, closed loop configurations using feedback are commonly used in T&Hs for medium signal frequencies [2, 37, 64]. These configurations are however not suitable for high-frequency input signals: the gain-bandwidth product is limited, so for high frequencies the gain is limited and the feedback mechanism for correction of imperfections is less effective, resulting in reduced linearity at higher input frequencies. For signal frequencies in the gigahertz range, closed loop configurations are not considered feasible in the target process technology.

Open loop configurations offer a higher bandwidth at the cost of accuracy and linearity. A bandwidth of e.g. 1 GHz is easily achievable with a configuration with a source-follower buffer,⁵ see Fig. 2.24. This configuration suffers from two problems: (1) distortion introduced by the sample process, for which solutions are presented in Sect. 4.3.1 and (2) distortion caused by the buffer, which is discussed in this section.

⁵This buffer does use feedback, but it is only local.

Fig. 2.24 High-speed, open-loop T&H configuration



The problem of buffer distortion is tackled in three steps. First, even-order distortion is discussed, second the dominant distortion mechanism of a conventional buffer is solved, and third the effect of a capacitive load is treated.

2.4.1 Even-order Distortion

Depending on the blocks in front of the ADC, a (quasi) differential implementation of the T&H can reduce even-order harmonics by a large amount. The actual reduction depends on the matching of the halves of the circuit. In general, matching of up to about 1% is realistic, resulting in a decrease of even-order distortion products with 40 dB.

The non-suppressed odd-order distortion products will now dominate the total distortion.

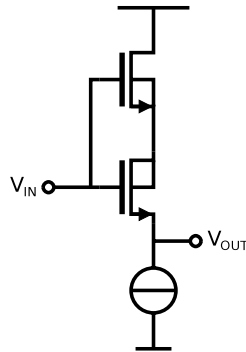
2.4.2 Buffer Distortion

Higher-order harmonics and most inter-modulation products can be canceled in the same way as second-order distortion, although additional phase-shifted versions of the input signal are required [29]. For example, by using three input signals, with phase-shifts of 120 degrees in between, all but the 4th, 7th, 10th, etc. order harmonics are canceled.⁶ It is however hard to generate signals with a constant phase-shift over a wide bandwidth, and therefore it is not considered to be a feasible option for a broadband ADC.

Therefore, other solutions are considered. Consider the source follower buffer of Fig. 2.24. The bulk of the PMOS is tied to its source, to mitigate the non-linear body effect. Assuming an ideal current source, the small signal transfer function is

⁶Increasing the number of phase shifted input signals does not only remove harmonics. For example, when going from 2 to 3 signals, some even-order harmonics appear again.

Fig. 2.25 The cascode source follower



given by:

$$V_{OUT} = V_{IN} \frac{1}{1 + \frac{1}{g_m(V_{IN}) \cdot r_{out}(V_{IN})}} \quad (2.8)$$

with g_m the transconductance of the transistor and r_{out} the output resistance of the transistor. Both g_m and r_{out} are functions of the drain-source voltage V_{DS} due to channel length modulation, and as V_{DS} depends on the input voltage, they are functions of the input voltage. So, when the input voltage varies, the transfer function varies, and the output signal becomes distorted.

In modern sub-micron CMOS processes, the non-linearity of the output resistance is the dominant source of distortion in the configuration of Fig. 2.24. To get a high bandwidth, the length of the transistor must be small, so the absolute value of the output resistance is small. If the intrinsic gain ($g_m \cdot r_{out}$) is small and nonlinear, the output signal is significantly distorted.

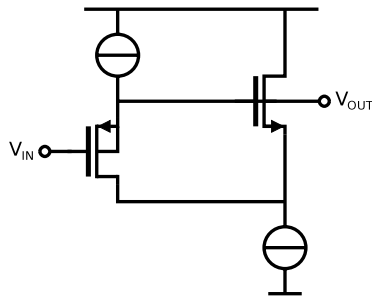
The best way to increase the linearity is to decrease the variation of the drain-source voltage. An example of a circuit where this is implemented is the cascode switch source follower [17, 18], shown in Fig. 2.25. A disadvantage of this implementation is the increased input capacitance. Moreover, the upper transistor needs to have a much smaller threshold voltage than the lower transistor to keep the lower transistor in saturation. This can be accomplished by scaling the transistors, which can be disadvantageous for other circuit aspects like speed or it can be accomplished by using a process option such as the low- V_T option [18], which requires additional process steps.

The schematic of a new unity-gain buffer is shown in Fig. 2.26 [26]. It is in fact a P-type source-follower (SF), with an additional N-type SF aiming to keep the drain-source voltage of the PMOS transistor constant.

The second SF decreases the variation in V_{DS} of the PMOST, such that the effective output resistance of the PMOST is increased and that the gain and linearity of the buffer are increased. This is explained in the next paragraphs.

The second SF transistor needs to have a short channel length to achieve a large SF bandwidth, and its bulk is connected to ground, since this is required by most standard CMOS processes. Due to the small output resistance and the body-effect,

Fig. 2.26 The schematic of a new unity gain buffer



the voltage gain of the 2nd SF buffer is only around 0.9. The signal swing over the source-drain of the first SF transistor is therefore only $0.1V_{OUT}$ instead of V_{OUT} . Consequently, there will flow 10 times less current in the output resistance, and its effective resistance is increased by the same factor. The g_m of the first SF transistor is unchanged, so the intrinsic gain ($g_m \cdot r_{out}$) is increased by a factor of 10 as well, and the voltage gain of the buffer will be closer to 1.

For the linearity the following holds: Suppose the output resistance is described by the following equation:

$$r_{out} = a + bV_{DS} + cV_{DS}^2 + dV_{DS}^3$$

Compared to a conventional SF, V_{DS} is 10 times less (-20 dB), the second-order distortion component cV_{DS}^2 is reduced by 40 dB (100 times) and the third-order distortion component dV_{DS}^3 is reduced by 60 dB (1000 times).

Note that this only holds for the linearity of the output resistance and does not imply that the distortion of the complete buffer is reduced by these amounts. Other distortion components (such as limited output resistance of the current sources) will now dominate the distortion.

Input Capacitance

It is important that the input capacitance of the T&H buffer is low and linear, to avoid distortion at the input of the buffer for high-frequency input signals. The new buffer has less non-linear input capacitance than a conventional or the cascoded source-follower, as described in the following: In the conventional source-follower, the gate-source capacitance is effectively lowered thanks to the Miller effect:

$$C_{eff} = (1 - A_V) \cdot C_{real}$$

with C_{eff} the effective capacitance when looking into the gate, A_V the voltage gain between the gate and the source and C_{real} the real gate-source capacitance. For a source-follower, the gain A_V is close to 1 and the effective capacitance is only a small fraction of the real capacitance. This is true for both the gate-source and the gate-bulk capacitance, assuming the bulk is connected to the source. What remains

is the gate-drain capacitance and this is the dominant input capacitance for both the conventional and the cascoded source-follower.

In the new unity-gain buffer, the drain terminal of the input transistor also tracks the input signal. The gate-drain capacitance is therefore mitigated as well, resulting in a very small input capacitance.

2.4.3 Distortion at High Frequencies with a Capacitive Load

If a buffer, implemented as a switch source follower (or similar) is loaded with a capacitance (e.g. an ADC), the current through the input transistor of the buffer varies when the capacitance is charged or discharged, see Fig. 2.28 with switch S2 closed. If the bias current is not constant, the gate-source voltage V_{GS} of the input transistor is not constant and the output will be distorted.

Before going into detail, the difference between non-interleaved and time-interleaved T&Hs is described first, as this has impact on the requirements for settling-time and bandwidth. For a non-interleaved (NI) T&H and a buffer with first-order settling behavior, the bandwidth requirement for the buffer with respect to settling is⁷:

$$BW_{NI, \text{settle}} > \frac{(n + 1) \cdot \ln(2) \cdot 2 \cdot f_s}{2\pi} \quad (2.9)$$

with n the resolution in bits, f_s the sample-rate and assuming half the sample-period for settling. This equation is derived in the next chapter. For the example of $n = 10$, the resulting bandwidth requirement yields: $BW_{NI, \text{settle}} > 4.9 f_{\text{Nyquist}}$. An input buffer with this bandwidth even tracks input signals at the Nyquist frequency closely.

For a time-interleaved T&H the bandwidth requirement for settling is relaxed by the interleaving factor (number of channels). The bandwidth requirement for a time-interleaved T&H is:

$$BW_{INT, \text{settle}} > \frac{(n + 1) \cdot \ln(2) \cdot 2 \cdot f_s}{2\pi \cdot N} \quad (2.10)$$

with N the interleaving factor and again assuming half the sample-period for settling. For the example of $n = 10$ and an interleaving factor of 16, the bandwidth requirement is: $BW_{INT, \text{settle}} > 0.3 f_{\text{Nyquist}}$.

If a buffer with minimal bandwidth for settling is used to save power, the buffer output no longer tracks input signals at the Nyquist frequency, but a large attenuation and phase-shift is present, and the problem as shown in Fig. 2.27 arises: During tracking, the buffer output V_{BUF} cannot follow the input signal $V_{T\&H}$ and at the sample moment (t_{SAMPLE}), the output signal V_{BUF} is not yet fully settled. After the sample moment, the buffer output V_{BUF} will slowly settle to its final value. During

⁷Compare (2.3) and see footnote 4 on p. 15.

Fig. 2.27 Sampling a high-speed input signal with limited buffer bandwidth

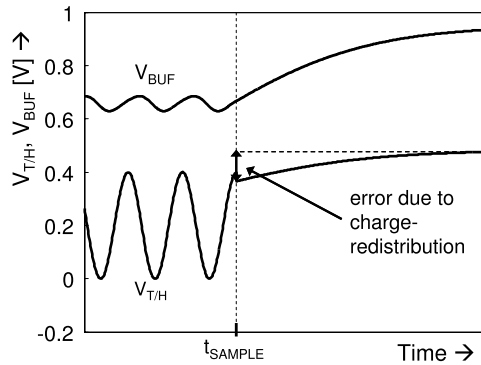
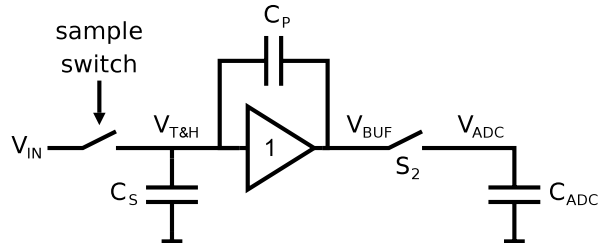


Fig. 2.28 T&H configuration with additional switch S₂, which is open during tracking to increase the buffer bandwidth and avoid distortion



this settling, charge-redistribution between (1) the non-linear parasitic capacitance C_P between the input and output of the buffer and (2) the sample capacitor C_S , causes distortion of the voltage on the sample capacitor $V_{T\&H}$ and the buffer output V_{BUF} , as indicated in the figure.

To avoid distortion, the buffer bandwidth could be increased, but this increases its power consumption significantly. Moreover, up-scaling of the buffer is limited, as this also increases the nonlinear input capacitance of the buffer, which requires more drive-power and introduces distortion at the input of the buffer. Up-scaling is therefore always a compromise between the required bandwidth on one side, and linearity, power and available drive on the other side.

To overcome this compromise, switch S₂ is introduced between the buffer output and the input capacitance of the ADC as shown in Fig. 2.28 [26]. In track-mode this switch is open and the load capacitance of the buffer is small. Hence the buffer bandwidth is high and output V_{BUF} can now follow the input $V_{T\&H}$ closely, as shown in Fig. 2.29. In this case, the distortion due to charge redistribution is mitigated, without decreasing the linearity or increasing the power consumption.

When the ADC is connected at $t = t_{\text{SWITCH}}$, the buffer output will first make a step to the value of the previous sample, still present on the ADC input capacitance. Then the buffer will charge the ADC load to the new sample value. charge redistribution after $t = t_{\text{SWITCH}}$ causes a signal dependent step in $V_{T\&H}$, marked by S. This seems to cause distortion, however as V_{BUF} settles to its final value, the process of charge redistribution is reversed and $V_{T\&H}$ returns to its initial, undistorted value. This is thanks to charge conservation at the capacitor plates connected to the input node of the amplifier.

Fig. 2.29 Sampling a high-speed input signal with enhanced buffer bandwidth in track-mode, thanks to disconnected capacitive load

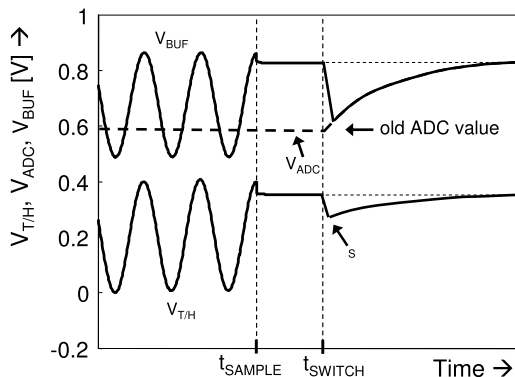
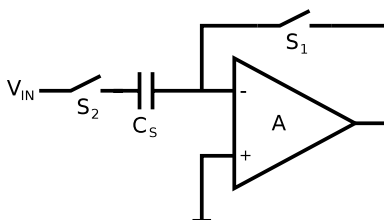


Fig. 2.30 Schematic of a T&H using bottom-plate sampling



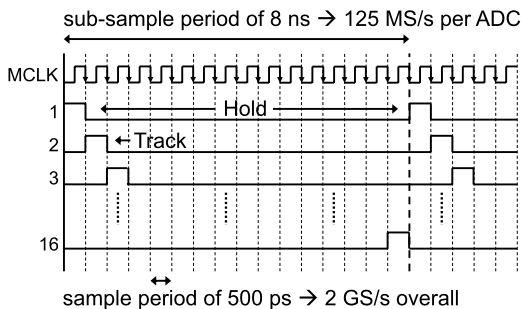
In conclusion: in an interleaving architecture the settling time can be relatively long. If the buffer has a large capacitive load, its bandwidth can be reduced to save power. However, this causes distortion. Now, by disconnecting the load during tracking, the distortion is avoided and the buffer bandwidth can remain reduced and power is saved.

2.5 Bottom-plate Sampling in a Time-interleaved ADC

Pipeline ADCs are broadly used and especially the voltage-mode type using an opamp for residue amplification is popular [2, 37, 64]. To make the sample process linear, these converters use bottom-plate sampling. This section considers the use of bottom-plate sampling in a time-interleaved ADC.

Bottom-plate sampling works as follows: The opamp creates a virtual ground node and the input signal is tracked on the sample capacitor as shown in Fig. 2.30 with switches S_1 and S_2 closed. When switching to hold-mode, S_1 is opened first and fixes the charge on capacitor C_S . Ideally this operation does not cause distortion and the operation is independent on the input signal, as the drain and source potentials of switch S_1 are at virtual ground potential. After this, switch S_2 is opened, which does not affect the amount of charge on C_S , since the other side of the capacitor is floating. This is called bottom-plate sampling, because the actual sampling takes place at the bottom-side, the (virtual) ground-side, of the capacitor.

Fig. 2.31 Timing diagram with a track-time of 1 period



The virtual ground is created by the opamp, which has a limited closed-loop bandwidth. When combining a number of such sampling structures in a time-interleaved ADC without a frontend sampler, two problems arise: The first is that it is difficult to reduce the track-time significantly. The opamp requires a certain amount of time to restore the virtual ground potential in case the input signal differs from the previous sample, and most time is required for signals close to the Nyquist frequency. The timing scheme with a track-time of 1 period shown in Fig. 2.31 and described in Sect. 2.3.1 is then not viable. For a longer track-time, more sample-capacitors are connected to the input at a time, limiting the input bandwidth.

The second problem is that the virtual-ground cannot be maintained for high-frequency input signals. Except for over-sampling it is therefore not useful to deploy the described bottom-plate sampling technique in a time-interleaved T&H without a frontend sampler.

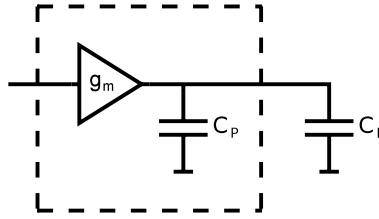
The use of bottom-plate sampling in a time-interleaved T&H with a frontend sampler is considered next. When using a frontend sampler, it turns off before the switch in the T&H, see Figs. 2.16 and 2.17. So, the signal conducting switch turns off first, which is contrary to the clocking scheme for bottom-plate sampling as explained above. Bottom-plate sampling can therefore not be used in combination with a frontend sampler.

To use opamp-based pipeline converters in a time-interleaved ADC, alternative sampling techniques have to be used such as in [16], where a separate T&H is used in front of the pipeline converters to perform the actual sampling. Consequently, the advantage of linearity of bottom-plate sampling is lost.

2.6 Number of Channels

In this section, aspects determining the number of channels of a time-interleaved ADC are discussed. The relation between the number of channels and the input bandwidth was already treated: In Sect. 2.3.1 it was argued that for the architecture without a frontend sampler, the input bandwidth depends on the number of channels, and in Sect. 2.3.2 it became clear that the number of channels determines the bandwidth together with R_{ON} of the frontend sample-switch.

Fig. 2.32 Transconductor with parasitic capacitance C_P and load capacitor C_L



If a frontend sampler is used, the achievable resolution decreases with the number of channels. Without a frontend sampler, this is much less an issue, as the settling time is significantly relaxed.

The input bandwidth can be increased by using a buffer with a low output impedance in front of the T&H instead of driving it with a $50\ \Omega$ source. Due to the high demands on this buffer (low impedance, high speed and high linearity) it requires a lot of power⁸ [40].

2.6.1 Sub-ADCs

Another important factor for determining the number of channels is the specification of the sub-ADCs. The sample-rate of a sub-ADC is: $f_{S,\text{subADC}} = f_S/N$ with f_S the sample-rate of the time-interleaved ADC and N the number of channels. For a lower number of channels the sample-rate of the sub-ADCs needs to be higher. The sample-rate of a non-interleaved medium resolution ADC (8–12 bits) is practically limited to a few hundred MS/s, as becomes clear from overviews of published ADCs [32, 60]. Besides a maximum sample-rate there is also a trend visible suggesting that beyond a certain sample-rate the power consumption increases more than proportional with the sample-rate. This can be explained as follows [53]: Suppose a transconductor with transconductance g_m and parasitic capacitance C_P , is charging a capacitive load C_L , see Fig. 2.32. The accompanying time-constant is:

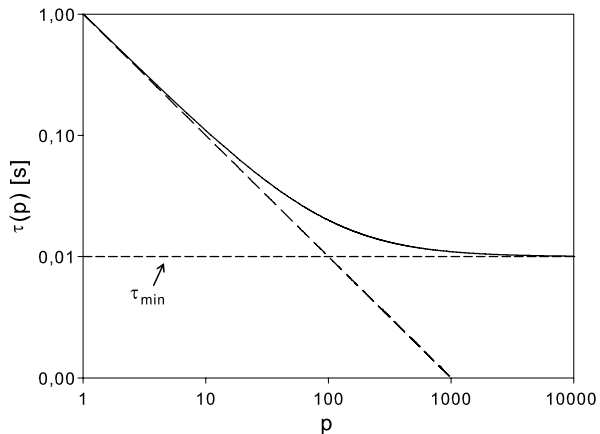
$$\tau = \frac{C_P + C_L}{g_m} \quad (2.11)$$

This buffer is used in an ADC and needs to charge the load capacitor within a certain accuracy in a clock period. Suppose $C_P \ll C_L$. In this case the time-constant is relatively large and the sample-rate limited. For a higher sample-rate, the time-constant needs to decrease, which can be accomplished by putting multiple buffers in parallel. The new time-constant is given by:

$$\tau(p) = \frac{p \cdot C_P + C_L}{p \cdot g_m} \quad (2.12)$$

⁸This is in contrast with the buffers after the T&H switch, as these have relaxed speed requirements due to the interleaving [26].

Fig. 2.33 The time-constant τ as a function of the number of parallel buffers p , with $C_P = C_L/100$. Both axes contain only relative numbers and are not related to a physical quantity



with p the number of parallel buffers. For small values of p , the time-constant decreases linearly with p , but when the value of $p \cdot C_P$ approaches C_L , the decrease of the time-constant becomes less than linear and for very large values of p , the time-constant even becomes independent of p . In this case:

$$p \cdot C_P \gg C_L \quad (2.13)$$

and so:

$$\tau = \tau_{\min} = \frac{C_P}{g_m} \quad (2.14)$$

This is graphically shown in Fig. 2.33, where C_P is chosen $C_L/100$. The values on both axes are just relative numbers and are not related to a physical number of buffers or transistor sizes.

From this, it can be concluded that starting from a low sample-rate, the power increases proportional with the sample-rate, while for higher sample-rates, the power increases more than linear with the sample-rate. The power efficiency therefore decreases for high sample-rates.

Dependency on Resolution

Another trend is that the sample-rate decreases with increasing resolution. For most architectures this is easy to explain. For example in a pipeline converter: for a higher resolution, the opamps require a longer settling time, so the achievable sample-rate is lower. In a successive approximation (SA) ADC, a higher resolution means more steps per conversion and more settling time per step, both lowering the achievable sample-rate. Finally, in flash architecture, more resolution implicates more com-

parators in parallel and a higher accuracy per comparator. This results in more capacitance⁹ and longer settling times, both degrading the achievable sample-rate.

2.6.2 Guidelines

The optimum number of channels is thus a trade-off between T&H architecture, bandwidth, resolution, power consumption and sub-ADC architecture. Moreover, it also depends on the process technology. Most of these trade-offs are hard to quantify and depend on a lot of variables and implementation details. It is therefore not possible to derive the exact optimum number of channels, however some guidelines can be given:

- In literature, ADCs with a medium resolution and good power efficiency can be found with sample-rates up to about 50–150 MS/s. $f_s/100$ MS/s could therefore serve as a starting point for the number of channels. So, e.g. 20 channels for a 2 GS/s time-interleaved ADC.
- A higher ADC resolution requires longer settling times and more conversion steps, limiting the maximum sub-ADC sample-rate. If the required resolution is relatively low (6–7 bits), the number of channels can be a bit lower, and when the resolution is relatively high (10–12 bits) the number of channels should be increased.
- Newer technologies (e.g. 65/45 nm) offer more speed than older technologies (e.g. 0.18 μm) [57] and allow a higher sample-rate of the sub-ADCs, and therefore the optimum number of channels is slightly lower for newer technologies.
- If a high input bandwidth of the T&H is required the number of channels should be reduced. This will probably increase the power consumption of the sub-ADCs.
- The use of a frontend sampler reduces the input bandwidth considerably. If a high bandwidth is required and timing-alignment can be made sufficiently accurate, do not use it.
- The optimum number of channels is relatively flat, adding or removing a few channels has no major impact.

2.7 Calibration

Mismatch causes errors like offset, and gain and timing differences. Calibration can be used to compensate for these errors. In this section, different kinds of calibrations are discussed together with their implementations.

Calibration of a circuit can be split-up into several parts, see Fig. 2.34. From left to right, the following blocks can be found: a test-signal generation block, the circuit which is calibrated, an error detection block and a correction block.

⁹Depending on the architecture, the capacitance increases between 4 and 8 times more per bit.

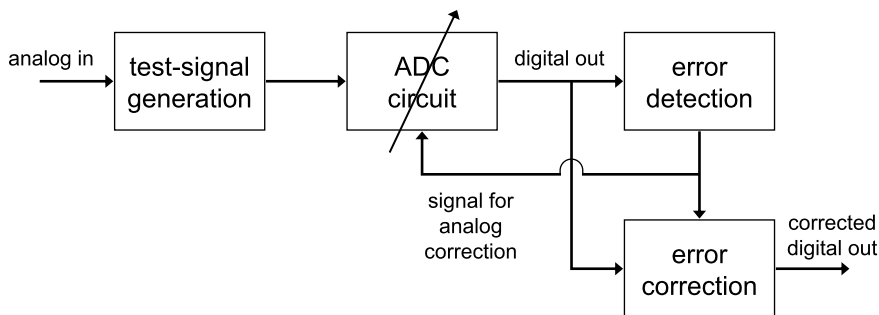


Fig. 2.34 Overview of blocks in a calibration system

The blocks are now discussed, starting with the test-signal generation block. Depending on the calibration, this may have a relatively easy function of shorting the inputs, swapping the input signals or generating a noncritical DC test-signal, or it may have a more sophisticated function, like generating accurate high-speed test-signals.

In order to perform calibration of a circuit, the error should be measured first and usually some signal processing and memory is needed for the calibration, therefore, the result of the measurement should preferably be in a digital format. As a T&H is normally followed by an ADC, this ADC can also serve as measurement device for the T&H. Moreover, the combination of T&H and ADC can be calibrated as one system. Therefore, the circuit to be calibrated is assumed to be a combination of T&H and ADC.

The detection block detects the errors and controls the signal generation block and correction blocks. Several techniques exist for performing a calibration. In foreground calibration, the ADC is not usable during calibration and specific input-signals are applied. This kind of calibration can be performed at start-up, after warm-up or periodically if the application allows for this.

Also, several background calibration techniques exist and the ADC is usable during this kind of (usually continuously running) calibration. Depending on the application, the normal input signal can be used to determine the errors or pseudo random data is added to the input (or the differential input signals can be interchanged) to distinguish between ADC errors and the input signal.

The correction of errors can be done both in the analog and in the digital domain. An advantage of digital correction is that the correction is fully transparent: e.g. an addition of 1 LSB is always exactly an addition of 1 LSB. A disadvantage is that the digital correction can consume a significant amount of power, especially in the case of more complex operations such as needed for bandwidth or timing correction.

Analog correction has the advantage that it can be implemented with little or no additional power consumption, that it is possible to do sub-LSB corrections (e.g. subtract 0.3 LSB from the input signal, or decrease the gain with 1%) without increasing the number of output bits and that the input range is not decreased. Moreover, it can undo errors made in the analog domain rather than correct for the

consequence of errors afterwards, as in digital correction. In the case of e.g. timing misalignment, it is usually not possible to reconstruct the original input from a distorted signal.

2.7.1 Offset Calibration

Mismatch of the sample-switches causes channel-to-channel variation of clock feed-through and channel-charge injection, which leads to offset between channels. If a buffer is used in a channel it can also cause offset. Reducing buffer offsets by circuit scaling to a small enough level (for example $\frac{1}{2}$ LSB) can, even for moderate resolutions, be unfeasible [24] as the input capacitance becomes very large and limits the input bandwidth unacceptably.

Both the detection and the correction of offset errors is relatively easy. In the case of foreground calibration, the differential inputs can simply be shorted to detect the channel offsets and for background calibration the average signal level over a long period can be determined, assuming the input signal is DC free. Correction of the error in both the analog [25] and digital domain is feasible.

2.7.2 Gain Calibration

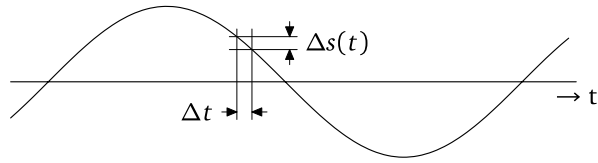
Gain differences between channels can have the same origin as offset errors. Foreground detection of gain errors is not complicated: after offset correction, a non-critical DC input signal can be applied to determine the channel gain. Analog correction of e.g. the gain of a buffer is feasible and does not have to cost power. An example is [25], where a digital adjustable impedance is placed between the two halves of a pseudo differential buffer. Compared to this, digital correction will always cost some power needed for the digital processing.

For some applications it can be beneficial to perform calibration based on the measured temperature or supply voltage. This is possible if e.g. the gain is dependent on the temperature and this dependency can be determined accurately. Also in this case, the adjustment can be performed in the analog domain.

2.7.3 Timing Calibration

Calibration of timing [40] is complicated. In the case of foreground calibration, high-frequency test signals are required and the detection of timing differences requires sophisticated algorithms. The correction of timing errors is also non-trivial and flexibility in timing tends to increase the power consumption and/or jitter of the clock signal. Background calibration often relies on spectral characteristics of

Fig. 2.35 An input signal as a function of time and the effect of sampling jitter



the input signal and usually involves additional hardware [19]. So, although timing calibration is possible, it is better to make the timing alignment accurate by itself, such that calibration can be avoided.

2.7.4 Bandwidth Calibration

Calibration of per-channel bandwidth is not trivial, as the detection of bandwidth differences is not straightforward: it requires high-frequency test-signals and complicated detection algorithms. Adjustment of the bandwidth could be another problem. It is therefore advantageous if bandwidth matching is accurate by design. This can be accomplished by making the channel bandwidth larger than strictly required, so that phase and gain match well within the band of interest. More information can be found in Sect. 2.2.2 on p. 9.

2.8 Jitter Requirement on the Sample-clock

Jitter in the sample-clock means uncertainty in the exact sample moment. In Fig. 2.35, it is graphically shown that a deviation from the ideal sample-moment leads to an error in the sampled signal. For a full-scale sinusoid with frequency f_{IN} , the signal-to-noise ratio is given by:

$$SNR_{\Delta t} = \frac{1}{\sigma(\Delta t) \cdot 2\pi \cdot f_{\text{IN}}} \quad (2.15)$$

with $\sigma(\Delta t)$ the RMS value of the timing jitter. Figure 2.36 shows this equation graphically, with $\sigma(\Delta t)$ as a parameter. For e.g. an input signal of 1 GHz, and a state-of-the-art timing jitter of 0.5 ps RMS, the resulting SNR is 50 dB or 8 ENOB.

Depending on the application, this definition may be too strict and may lead to over-design. In for example wireline communication systems, channels have increased insertion loss at high signal frequencies [16]; a maximum signal amplitude at the maximum (Nyquist) frequency, does not occur.

In [14] an example is given which compares the variance on the phase-skew between channels when assuming either (1) a sinusoid at the maximum frequency or (2) white noise filtered with an ideal filter with a cutoff frequency equal to the maximum frequency. The requirement on the phase-skew variance is a factor three lower for the filtered white noise, independent of the number of channels. Random

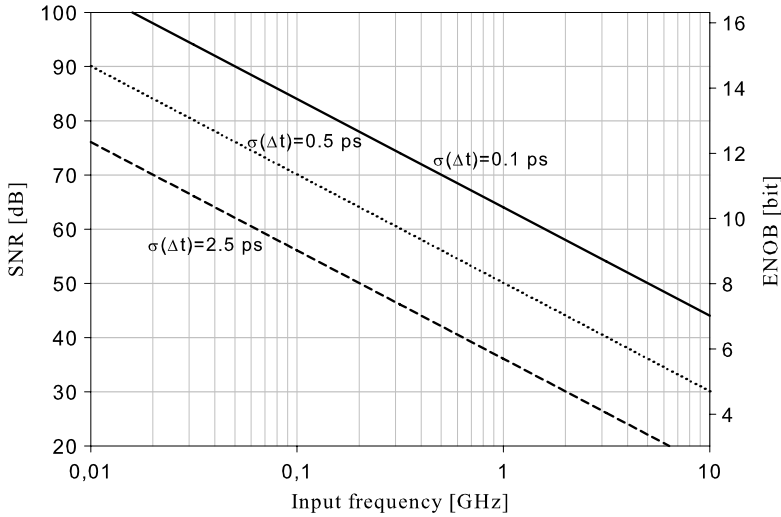


Fig. 2.36 Signal-to-noise ratio and ENOB as a function of the input frequency with $\sigma(\Delta t)$ as parameter

clock jitter can be considered as phase skew for an infinite number of channels, and this conclusion therefore also holds for random clock jitter. So, assuming a sinusoid at the maximum input frequency leads to an overestimate of the required jitter variance.

Also for a software (-defined) radio receiver the above definition can be far too strict. It can be shown that [5]:

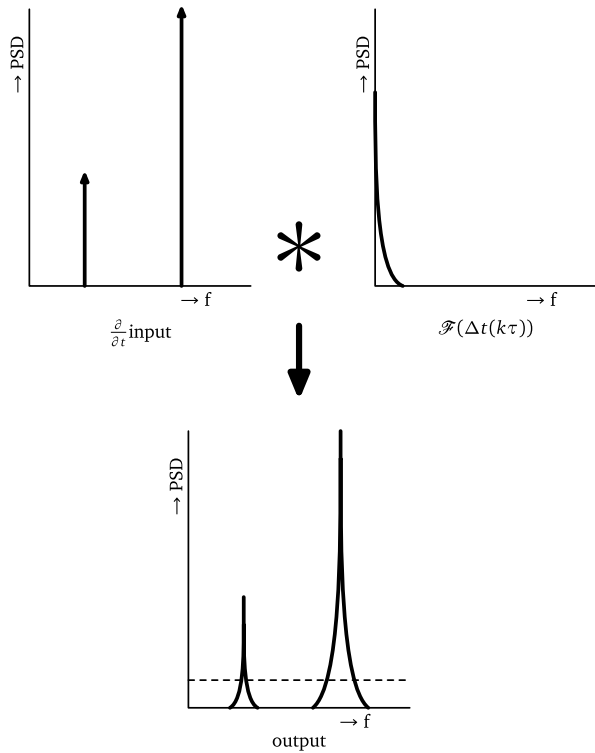
$$\mathcal{F}(\Delta s_{\tau}(k)) \approx \mathcal{F}(\Delta t(k\tau)) * \mathcal{F}\left(\left.\frac{\partial}{\partial t}s(t)\right|_{k\tau}\right) \quad (2.16)$$

where \mathcal{F} denotes the DTFT, $*$ denotes convolution, $\Delta s_{\tau}(k)$ is the error in the sampled signal and $\Delta t(k\tau)$ is the sampling time error. When the sampling clock is derived from a synthesizer containing a VCO, $\Delta t(t)$ can be assumed to have a f^{-2} power spectrum outside the synthesizer loop bandwidth [41]. Due to its f^{-2} nature, most energy in the error of the sampling clock is at low frequencies. Assuming such a spectrum, (2.16) is illustrated in Fig. 2.37.

Knowing that in the frequency domain this is convoluted with the derivative of the input signal leads to the following [4]:

- The jitter spectrum is convoluted with the input spectrum, and therefore the jitter-induced error is concentrated around the input frequencies.
- Input signals with higher power are surrounded by more jitter-induced error in the output than input signals with lower power.
- Input signals of higher frequencies are surrounded by more jitter-induced error in the output than signals at lower frequencies.

Fig. 2.37 Illustration of (2.16). The convolution of the input spectrum (*upper left*) with the spectrum of $\Delta t(k\tau)$ (*upper right*) gives the output spectrum of the ADC (*bottom*)



So, if a small input signal needs to be received in the presence of a strong out-of-band interferer, the jitter requirement should be based on the small input signal (with possibly also a lower frequency) rather than on the strong interferer. This can relax the jitter specification by two orders of magnitude [4].

2.9 Summary and Conclusions

In this chapter the time-interleaved Track and Hold architecture was discussed. Mismatch between channels, like differences in offset, gain and timing, degrade the performance. In the case of bandwidth mismatch it is shown that it is advantageous to have a large bandwidth, such that for the frequencies of interest the effect of the mismatch is mitigated.

Two T&H architectures were discussed, one with a frontend sampler and one without. The use of a frontend sampler has the advantages of good timing alignment between channels, the resistance of the switch is however a problem: it limits both the input bandwidth and the achievable resolution. The input bandwidth and accuracy can be improved by placing additional switches, which decrease the capacitance after the frontend sampler.

A buffer is introduced which has a large bandwidth and improved linearity compared to a switch source follower. Driving a large capacitive load with a large bandwidth consumes a lot of power. The time-interleaved architecture offers the possibility to decrease the buffer bandwidth, as there is a relatively long time available for settling of the buffer. However, the reduced bandwidth causes distortion. To avoid this, the following solution is introduced: a switch is inserted between the buffer and the capacitive load, such that the buffer bandwidth is increased and the distortion is mitigated. This solution can save a significant amount of power.

Some guidelines are given for determining the number of channels of a time-interleaved ADC. This depends on resolution, bandwidth, technology, and whether a frontend sampler is used or not. For the target specifications and process technology, and considering the current state-of-the-art, a sub-ADC sample-rate of about 100 MS/s can serve as a starting point, as literature shows that power efficient sub-ADCs with this sample-rate are currently feasible.

The topic of calibration is discussed and it turns out that offset and gain calibration is relatively easy to implement, while the implementation of timing and bandwidth calibration is much harder. One should therefore strive for good timing-alignment and bandwidth matching in the frequency band of interest.

Finally, the jitter requirements on the sample-clock are discussed. The traditional requirement assumes a full-scale sinusoid at the maximum frequency. For many applications this requirement is too strict and leads to over-design.

Regarding the desired implementation, the following design choices are derived:

- Use the architecture without a frontend sampler
- The channel bandwidth should be large compared to f_{IN}
- Use the presented T&H buffer
- In track-mode, disconnect the capacitive load from the buffer
- The number of channels should be around 20
- Timing alignment should be made accurate by design
- Use analog calibration to correct for offset and gain errors

Time-interleaved Analog-to-Digital Converters

Louwsma, S.; van Tuijl, E.D.; Nauta, B.

2011, XVI, 136 p., Hardcover

ISBN: 978-90-481-9715-6