Offset, flicker noise, and ways to deal with them

Hanspeter Schmid*†

November 6, 2008

Introduction

After almost one century of research into flicker noise, we still do not know as much about it as we would like to: we do not know enough about its origin, nor do we know everything about its behaviour, nor has the last word about good methods to fight it been spoken: effectively, we still are like Alice standing in front of the rabbit hole, before she enters the Wonderland ...

So the intent of this chapter is to give the reader an idea of what flicker noise is, how it is connected to other low-frequency noise effects, and what today's designers do to fight it. This chapter will just give a broad overview, focusing on concepts and design philosophy, providing just as much mathematics as is strictly necessary. Interested readers will have to follow the literature references to find out details about mathematics and design.

In this chapter, a section on the nature of flicker noise is followed by a section on switched-capacitor techniques and noise sampling. Three more sections deal with the three main techniques used against flicker noise, which are large-scale excitation, chopping, and correlated double sampling. An appendix contains information on how to simulate flicker noise in Matlab, and finally, a short annotaded literature list is given, inviting the reader to find out by herself or himself how deep the rabbit hole really is.

1 What is flicker noise?

Flicker noise, or $1/f$-noise, seems to be so easy to define: it is noise whose power spectral density has the form

$$S(f) = S(1) \cdot \frac{1}{f^x}$$

where x typically is around 1. In most circuits, this means that white noise dominates above a certain frequency, and we will see a behaviour as in Fig. 1.

While this definition looks so simple, it immediately begs the question: does flicker noise really go down all the way to $f = 0$? And what would such behaviour actually mean?

One thing this would mean is that flicker noise would then have infinite power over a finite frequency band, because

$$\int_0^1 S(1) \frac{1}{f^x} \mathrm{d}f \to \infty$$

The problem we are facing with flicker noise is actually rather simple: we are looking at it now in the frequency domain only, without thinking about what integrating from $f = 0$ upwards actually means: It means that we are looking at a process that takes an infinite time to happen, and this is not realistic at all. Looking at spectra is normally very helpful for understanding amplifiers, filters, regulators and the like, but we should never forget that the time domain and the frequency domain are only equivalent mathematically, but in reality, signals are varying in time, and frequency is only an abstract, if helpful, tool we use for our convenience [1].

*Institute of Microelectronics, University of Applied Sciences Northwestern Switzerland (IME/FHNW), hanspeter.schmid@fhnw.ch

†This is the preliminary version of a chapter of the CRC book “Circuits at the Nanoscale — Communications, Imaging and Sensing”, edited by Chris Iniewski, ISBN 978-1-4200-7062-0.

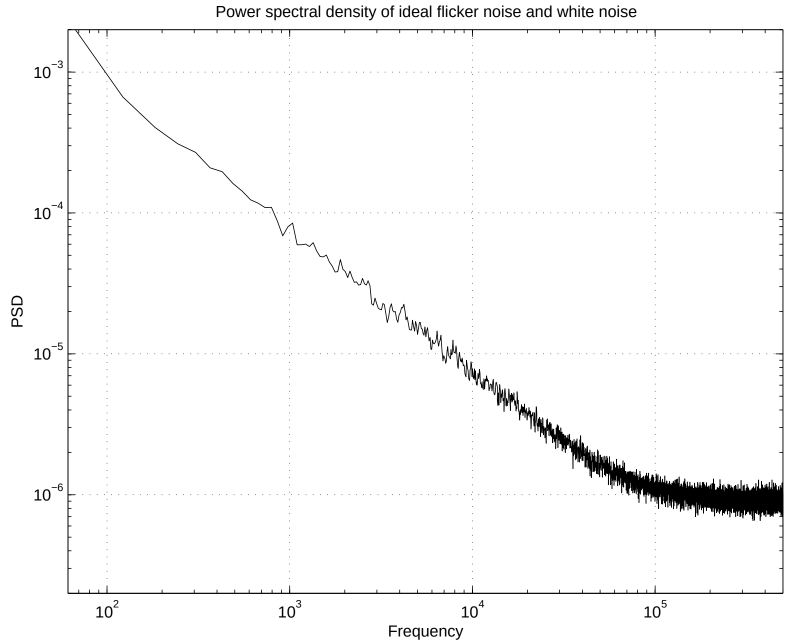

Figure 1: Power spectral density of white noise overlaid by flicker noise.

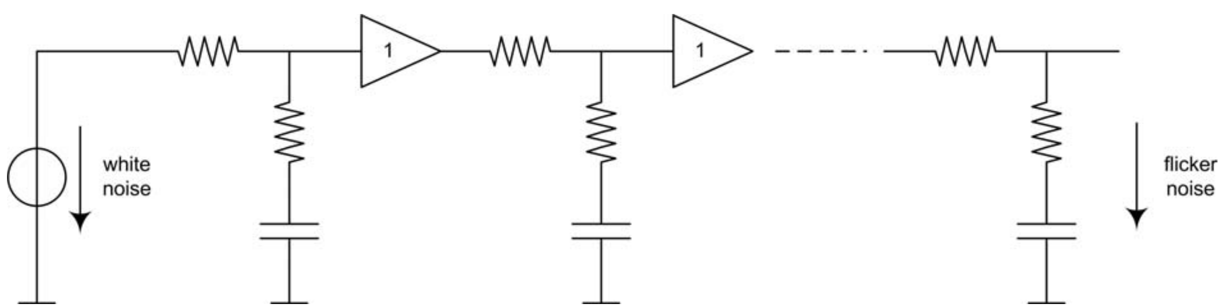

Figure 2: Flicker noise generated from white noise.

1.1 The nature of flicker noise

Looking at processes generating flicker noise in the time domain instead of the frequency domain gives us much more insight into the nature of flicker noise. We have no problems finding flickering systems in nature and science, it seems that flicker noise is the rule rather than the exception. It can be observed in systems like vacuum tubes, diodes, transistors, thin films, quartz oscillators, the average seasonal temperature, the annual amount of rain fall, the rate of traffic flow, the loudness and pitch of music, the pressure in lakes, search engine hits on the Internet, and so on [2, 3].

Keshner showed in 1982 [2] that a system flickers when it has memories whose time constants are distributed evenly over logarithmic time. Therefore, an easy way to produce flicker noise in simulation is to concatenate many stages of first-order filters with one pole and one zero each, and let it filter white noise, as shown in Fig. 2 [2], where four first-order filters are used per decade. The number of filters per decade decide how far the simulated $1/f$ curve deviates from the ideal curve.

The poles and zeros must be spaced evenly on a logarithmic scale. For the simulations shown in this chapter, we have used the spacing shown in Fig. 3, as described in the Appendix.

This system gives the very nice $1/f$ behaviour in Fig. 1, and it is amazing to see that the number of memory blocks needed to make flicker noise is relatively small. According to Bloom [4], MOSFETs show flicker noise behaviour from, e.g., 10^{-8} Hz up to 10^{5} Hz, which would require only 25 memory cells with time constants distributed evenly on a logarithmic scale.

Making simulations with this model of flicker noise, we soon find funny effects. Fig. 4 shows, for example, the variance of the output signal of the circuit in Fig. 2 as a function of time.

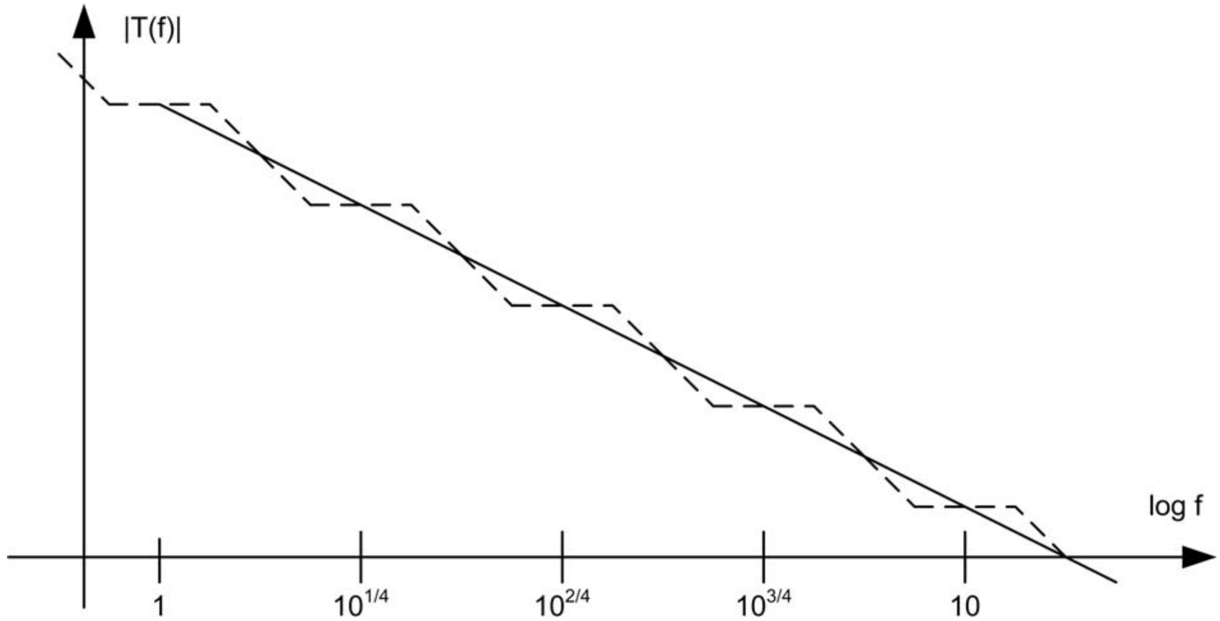

Figure 3: Transfer function of the flicker noise generator.

It is immediately visible that this variance rises with $\log t$! In other words, the random signal we are looking at is not stationary. The theoretical consequences of this have been discussed in [2], and measurements of practical problems coming from this non-stationarity have been shown in [5], so the non-stationarity is not a problem of our model, but an inherent feature of flicker noise: it means that if you have a system with long time constants in its memories, then that system takes a long time to reach its steady state. To return to Bloom [4], the 10^{-8} Hz he mentioned correspond to a time of three years, so normally we will never really see the steady state in MOSFET circuits. However, as long as we do not do correlated double sampling, this does not concern us.

1.2 Memory in systems

Each of the systems mentioned above have memory of some sort. For example, it is described in [3] that the number of vowels in words like 'aargh' and 'loooove' and the number of hits (the frequency) when these words are entered as search terms in Internet search engines are related by an $1/f^x$ law. The absolute numbers are different for each word, but the exponent x only depends on the nature of the memory, which is: when a person sees a word like 'looooove' on a web page, that person may feel inclined to write 'love' with even more o's in an attempt to express stronger feelings. So in this case the memory are Internet pages interacting with users' memories, and x is the same for all words.

Lakes also show flicker noise; in this case the behaviour is close to $1/f^{5/3}$ for every lake in the world, only the magnitude is different. What happens there is that the Coriolis force (from earth rotation) causes whirls of big dimensions; these whirls transfer their energy to smaller whirls, and so on, until their energy is dissipated at molecular level. This cascade of whirls is not very much unlike the filter cascade shown in Fig. 2, and Kolmogorow showed long ago that simply having such a cascade of whirls already determines the exponent 5/3, but again not the magnitude of the flicker noise.

1.3 Memories in MOSFETs and other electronic devices

Almost every electronic device shows some flicker noise: vacuum tubes, resistors, diodes, BJTs and MOSFETs; but in MOSFETs, the magnitude is by far the largest. The reason for this is that there are several different effects causing flicker noise in electronic devices, in every case with $1/f^x$ and $x \approx 1$, but these effects can be divided into volume effects and surface effects [6].

The two main volume effects are *Bremsstrahlung* and carrier scattering. *Bremsstrahlung* is a German word used in quantum mechanics that roughly means "deceleration radiation". Whenever an electron is accelerated, it will emit low-frequency *Bremsstrahlung*, and will be slowed down by its own

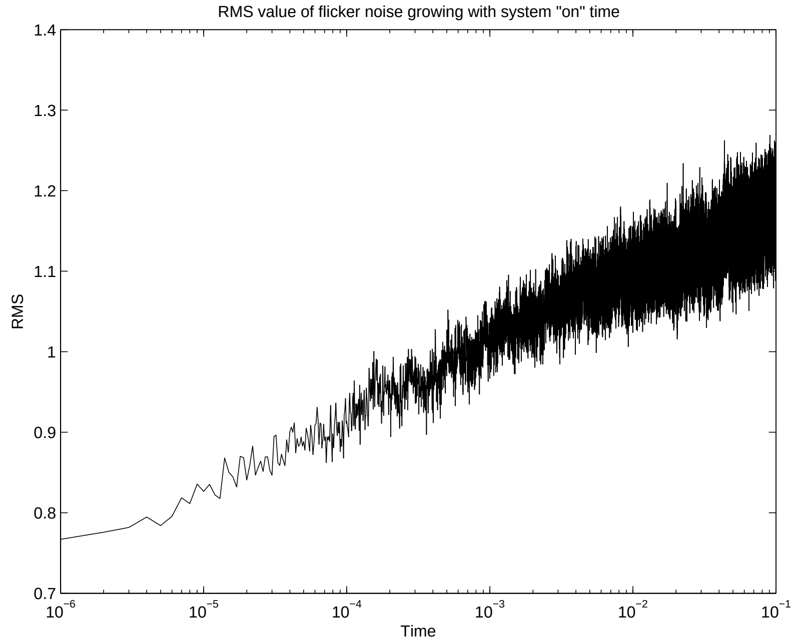

Figure 4: Variance of flicker noise as a function of system "on" time.

Bremsstrahlung, as will other electrons in its vicinity. Thus we again have low-frequency energy and a cascade that remembers it, giving $1/f$ noise. This is the main source of $1/f$ noise in vacuum tubes.

The second volume effect is scattering, when electrons are scattered at the silicon lattice, or at impurities in the material, or by acoustical or optical phonons, and so on. In all cases, the scattering will interact with the lattice, generating phonons, which will later cause more scattering, and again we will have 1/f noise. This is the dominant source in most solid-state devices.

The effect that dominates in MOSFETs, though, is something quite different: in MOSFETs, electrons tunnel from traps in the oxide to the gate and the conducting channel, and vice versa. If there is only one single trap (which may indeed happen in minimum-size deep-sub-micron transistors), then this causes a power spectral density of the drain current

$$S(f) \approx \frac{\tau}{1 + \omega^2 \tau^2}$$

with a certain trap time constant τ. This is $1/f^2$ behaviour, as white noise fed through a one-pole low-pass filter would give, but due to the quantum nature of the electron trapping, this noise signal will only have two current levels. Such noise is called "random telegraph noise" [5]. Now what happens if we have several traps? It can be shown that the time constant for a trap at a distance z from the interface is

$$\tau = \tau_0 \exp\left(\frac{10^{10}}{\mathrm{m}} \cdot z\right) \tag{1}$$

for some process-dependent time constant τ_0, so if traps are uniformly distributed over $z = 0 \ldots z_{\mathrm{g}}$, we will have memories with time constants that are uniformly distributed over a logarithmic scale, as in Figs. 2 and 3! The difference is that we drive the filter in these figures with Gaussian white noise instead of a two-level signal with white frequency characteristic. We also see from (1) that even for a gate with thickness $z_{\mathrm{g}} = 1$ nm, the time constants of the flicker noise are spread over more than three orders of magnitude.

Experiments with large-scale excitation of MOSFETs – where part of the memory is deleted and therefore flicker noise is reduced intrinsically – show that flickering occurs even when the transistor is switched off completely. It can then just not be measured directly, because what we can measure are just the effects caused by electron trapping: electrons tunnelling in and out of traps will cause both carrier number fluctuations and also fluctuations of the carrier mobility μ [5], which in turn make the drain current of the MOSFET flicker. This is also reflected in one of the widely used simple flicker noise models of the MOSFET,

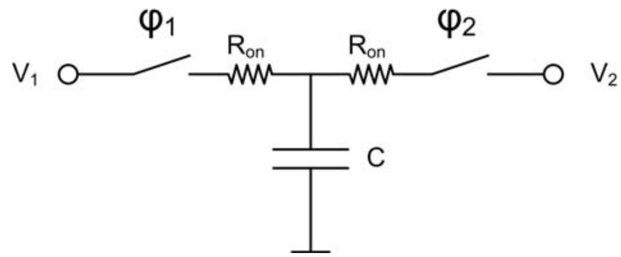

Figure 5: Switched-capacitor resistor.

$$\overline{V_g}^2 = \frac{K}{WLC_{\text{ox}}f}$$

where K and C_{ox} are technology parameters, and W and L the transistor dimensions: this formula does not depend on the bias conditions of the device, meaning it does not depend on whether any current flows through the MOSFET.

1.4 Memory and correlation

Turning back to the mathematics of flicker noise: the Fourier transform of a power spectral density is the autocorrelation function, which, for $1/f^x$ noise, is [2]

$$R(\tau) \sim |\tau|^{x-1}$$

So for $x = 1$, $R(\tau)$ is constant, meaning the present value of the flicker noise signal correlates very well with *all other values* of the same signal, and so flicker noise can be removed effectively with techniques that operate on correlated samples of the flickering signals (e.g., correlated double sampling).

1.5 Flicker noise is offset extended in frequency

If we extend our view of flicker noise down to $f = 0$, we look at an error signal that is constant in time: offset. While this is not mathematically inspiring, it still means something in practice: most techniques removing flicker noise will also cancel offset, and vice versa.

1.6 Techniques to reduce flicker noise

Considering all that has been said until here, we end up with three techniques to fight flicker noise:

- Knowing that flicker noise comes from memory, we attempt to reset this memory. This is known as *large-signal excitation (LSE).*
- Knowing that flicker noise has a flat autocorrelation function, we attempt to remove it by subtracting two correlated samples. This is known as *correlated double sampling (CDS).*
- Knowing that flicker noise is a low-frequency effect, we attempt to modulate it into a frequency band outside the signal band. This is known as *chopping.*

Except for chopping, these techniques only work on sampled signals, so we must first have a look at switched-capacitor techniques and noise sampling.

2 Switched-Capacitor Techniques

Fig. 5 shows a very simple switched-capacitor circuit. The two switches are closed during the clock phases ϕ_1 and ϕ_2, respectively, and the two clock signals do not overlap, such that the two switches are never closed simultaneously.

When ϕ_1 is closed, the capacitor is charged to V_1, storing the charge $Q = C \cdot V_1$. When ϕ_2 is closed, $Q = C \cdot V_2$. Therefore, in every clock cycle, the charge $\Delta Q = C \cdot (V_1 - V_2)$ is transferred. The mean current

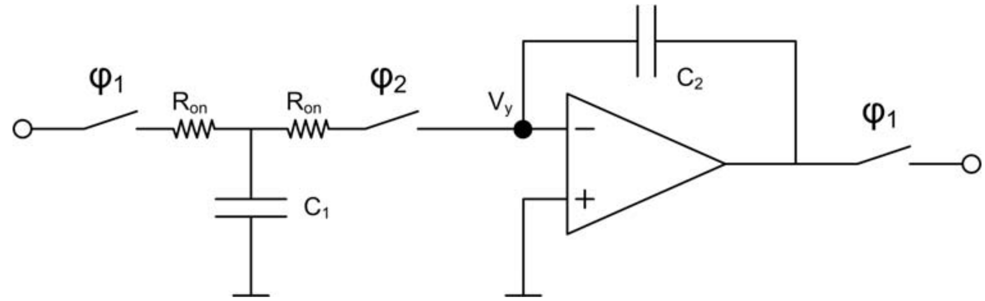

Figure 6: Switched-capacitor integrator.

through this circuit is the $I_{12} = \Delta Q/T_{\text{clk}} = f_{\text{clk}} \cdot C \cdot (V_1 - V2)$, so we have a resistor with equivalent resistance $R_{\text{eq}} = 1/(f_{\text{clk}}C)$.

The interesting thing about SC filters is that they become *much* faster with technology scaling. This can be shown as follows [7]: For good settling, we require $T_{\text{clk}}/2 > 5R_{\text{on}}C$, where R_{on} is the on-resistance of the switches. So we want

$$f_{\text{clk}} < \frac{1}{10R_{\text{on}}C} \tag{2}$$

The on-resistance of a MOSFET switch is

$$R_{\text{on}} = \frac{1}{\mu C_{\text{ox}}\frac{W}{L}V_{\text{eff}}} \tag{3}$$

where μ is the carrier mobility, C_{ox} the gate oxide capacitance density, W/L the width over the length, and V_{eff} the gate overdrive voltage.

In addition, we know that when a switch is opened, approximately half of the channel charge $Q_{\text{ch}} = -WLC_{\text{ox}}V_{\text{eff}}$ will go into the capacitor and cause a voltage error

$$|\Delta V| = \frac{|Q_{\text{ch}}|}{2C} = \frac{WLC_{\text{ox}}\,|V_{\text{eff}}|}{2C}$$

So the C we have to use for a certain switch and some given $|\Delta V|_{\text{max}}$ is

$$C = \frac{WLC_{\text{ox}}\,|V_{\text{eff}}|}{2\,|\Delta V|_{\text{max}}} \tag{4}$$

Replacing R_{on} in (2) according to (3) and C according to (4) gives a very simple result:

$$f_{\text{clk}} < \frac{\mu\,|\Delta V|_{\text{max}}}{5L^2} \tag{5}$$

$|\Delta V|_{\text{max}}$ depends on the maximum signal and therefore on V_{dd}. The product $\mu\,|\Delta V|_{\text{max}}$ does not change a lot as technology scales, so, to the first order, (5) means that the maximum speed of SC circuits scales as does the number of transistor per area, which means that Moore's law is also valid for the speed of SC circuits.

The main advantage of SC techniques can be shown with Fig. 6. This is an integrator with time constant

$$\tau = R_{\text{eq}}C_2 = \frac{C_1}{C_2} \cdot \frac{1}{f_{\text{clk}}}$$

So we have a time constant derived from a ratio of capacitors, which can be made precise to within less than one percent, and a clock frequency, which is even more precise.

2.1 Sampled noise in SC circuits

This great advantage is paid with more aliasing, though. The precise calculation is quite difficult even for the simple circuit in Fig. 6 — see [8] for details — because, at the output, one simultaneously sees direct noise from the op-amp as well as sampled noise from the earlier stages. Fortunately, aliased broad-band noise often dominates, and a simplified analysis can be made.

What noise sampling means can be shown using the very simple circuit in Fig. 5. When ϕ_1 closes, and we wait for the system to reach the thermal equilibrium, then the energy stored in the capacitor is $\frac{1}{2}CV_c^2$. Similarly, the noise energy coming from a noise voltage $\overline{V}_{c,\text{rms}}$ is $\frac{1}{2}C\overline{V}^2_{c,\text{rms}}$. We also know from thermodynamics that the energy in a system with one degree of freedom is $\frac{1}{2}kT$, so it directly follows that the variance of the thermal noise is

$$\frac{1}{2}C\overline{V}^2_{c,\text{rms}} = \frac{1}{2}kT \qquad \Longrightarrow \qquad \overline{V}^2_{c,\text{rms}} = \frac{kT}{C} \tag{6}$$

This can also be shown in a different way: the noise caused by R_on is $\overline{V}_r^2 = 4kTR_\text{on}$, and the bandwidth of the filter consisting of R_on and C is $1/R_\text{on}C$. Integrating the filter's noise over the bandwidth will again give the result in (6).

So, essentially, as long as R_on is low enough such that the circuit in Fig. 5 reaches equilibrium at the end of the clock phase, the integrated noise power depends on C only. To the first order, this noise is white noise. So what goes into the node V_y of Fig. 6 is essentially sampled white noise with a power spectral density (PSD) of

$$S_n(f) = \frac{kT}{Cf_\text{clk}} \qquad \text{for } -\frac{1}{2}f_\text{clk} \le f \le \frac{1}{2}f_\text{clk}$$

We also have to look at sampled white noise. Assume that the inputs of the circuit in Figs. 5 and 6 are driven by a pre-amplifier producing white noise up to a noise bandwidth f_nbw that is related to the amplifier bandwidth, so that its single-sided PSD is approximately

$$S_a(f) = \frac{\overline{V}^2_\text{amp,rms}}{f_\text{nbw}} \qquad \text{for } 0 \le f \le f_\text{nbw}$$

The square root of the level of this noise PSD would be in the unit $\text{nV}/\sqrt{\text{Hz}}$ value often found in op-amp data sheets. Since the amplifier must be fast enough to settle well within one clock period, we normally have $f_\text{nbw} \gg f_\text{clk}$ and therefore the noise is aliased. Through aliasing, the noise is compressed from a range $0 \ldots f_\text{nbw}$ to a range $-\frac{1}{2}f_\text{clk} \ldots \frac{1}{2}f_\text{clk}$, so the aliased noise is scaled up:

$$S_{a,\text{aliased}}(f) = \frac{f_\text{nbw}}{f_\text{clk}} \frac{\overline{V}^2_\text{amp,rms}}{f_\text{nbw}} \qquad \text{for } -\frac{1}{2}f_\text{clk} \le f \le \frac{1}{2}f_\text{clk}$$

or, if we use single-sided spectra for the sampled signals,

$$S_{a,\text{aliased}}(f) = 2 \cdot \frac{f_\text{nbw}}{f_\text{clk}} \frac{\overline{V}^2_\text{amp,rms}}{f_\text{nbw}} \qquad \text{for } 0 \le f \le \frac{1}{2}f_\text{clk}$$

This means: sampling 10-MHz-wide white noise at 1 MHz gives *twenty* times higher noise power. In Fig. 6, this noise is then integrated by the SC integrator.

With this way of thinking, we can identify all noise sources, calculate their noise transfer functions to the output of the circuit, and add all contributions. [8] shows this using Fig. 5 as an example. A general method using matrix equations and including white noise, flicker noise and amplifier noise, was presented in [9]. [10] describes the simplified noise analysis of choppers and correlated double samplers; this will be discussed again briefly in the following sections of this chapter.

Fortunately, in SC applications that do not attempt to cancel flicker noise, sampled white noise normally dominates, which makes an analysis simpler. To illustrate this, the lower curve in Fig. 7 is a (sampled) signal with a white-noise and a flicker-noise component. The flicker noise corner frequency is at approximately 1/5 of the signal bandwidth. If this signal is under-sampled ten times, the upper curve results, with the same flicker noise, but ten times more white noise, so the flicker noise corner frequency still is at approximately 1/5 of the signal bandwidth. So sampling generally *reduces* the flicker noise corner frequency.

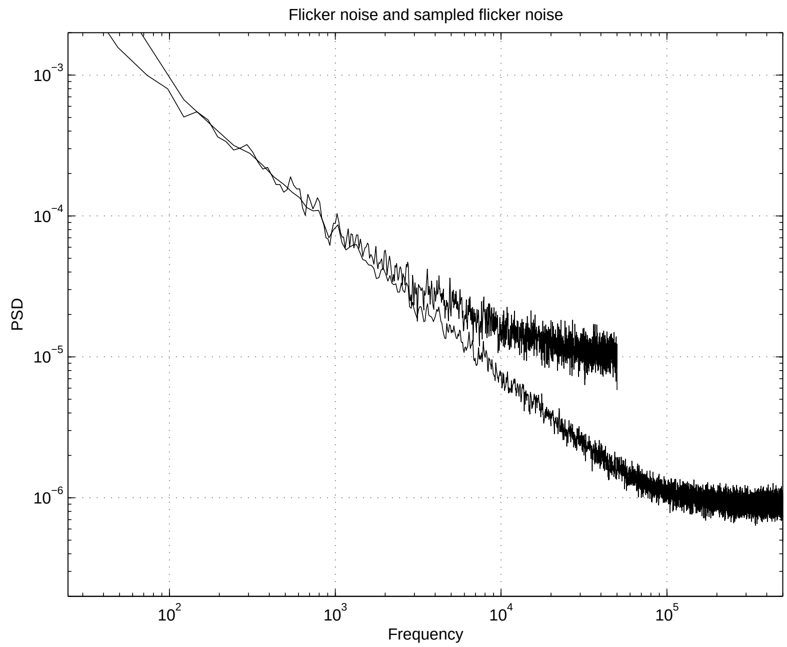

Figure 7: PSD of white noise overlaid by flicker noise, sampled with 1MHz and 100kHz.

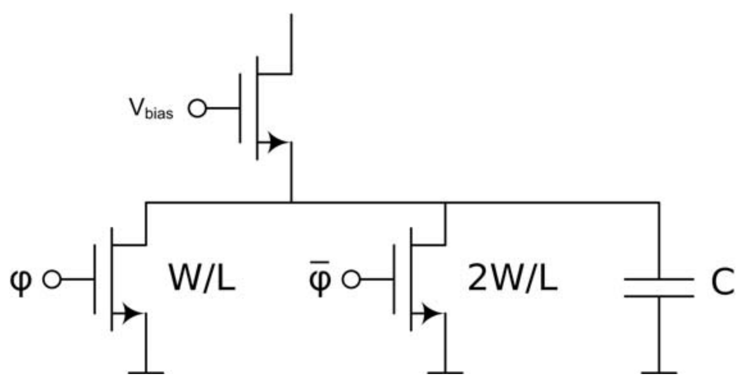

Figure 8: Switched current source.

3 Bias switching and large-scale excitation (LSE)

Figure 8 shows a switched current source. If this circuit is operated with a variable-duty-cycle clock ϕ and its inverse $\overline{\phi}$, then the current can be tuned by a factor of two. It has been observed that for duty cycles between 0% and 100%, this circuit is much less noisy than the circuit simulator predicts [5]. The reason for this is that switching a transistor off deletes some of its flickering memory by kicking some of the trapped electrons out of their traps.

Fig. 9 shows another Matlab simulation in which the memory of the flicker noise is deleted almost completely once every 10 μs. The flicker noise disappears almost completely in this example; normally, some flicker noise remains at low frequencies because it is not possible to delete all of the memory. This effect can be calculated [5], but not simulated; there is as yet no circuit simulator that takes flicker noise memory effects into account. However, there are already many applications other than Fig. 9 in which LSE is used.

For example, [11] presents an op-amp with a switched input differential pair as in Fig. 10. The two transistors are used alternatively; the clock switches the unused one off, deleting its flicker noise memory. This will of course introduce spikes in the output voltage at multiples of f_{clk}, but it also reduces the flicker noise of the op-amp. In [11] the measured noise at low frequencies was reduced by 5 dB.

Another place where such memory effects are observed are oscillators. In oscillators, transistor flicker

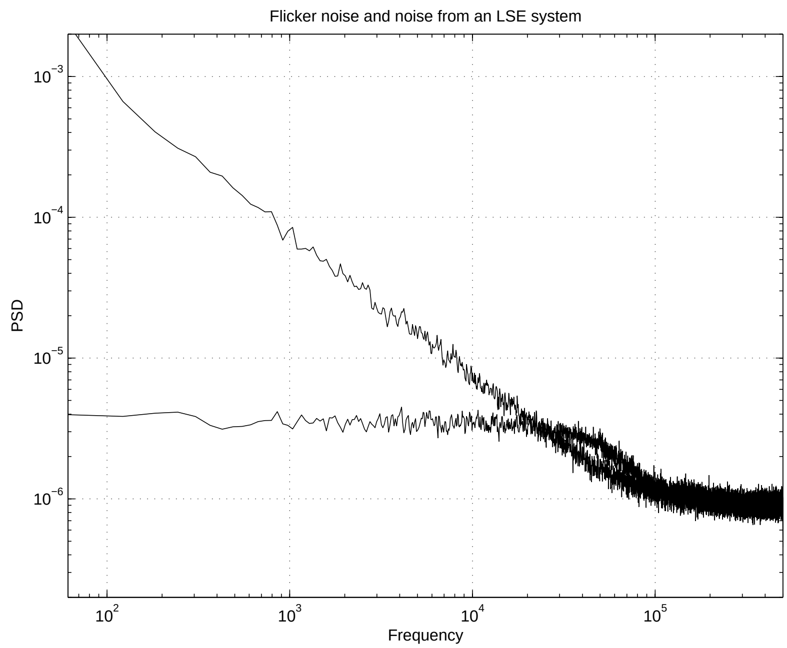

Figure 9: .

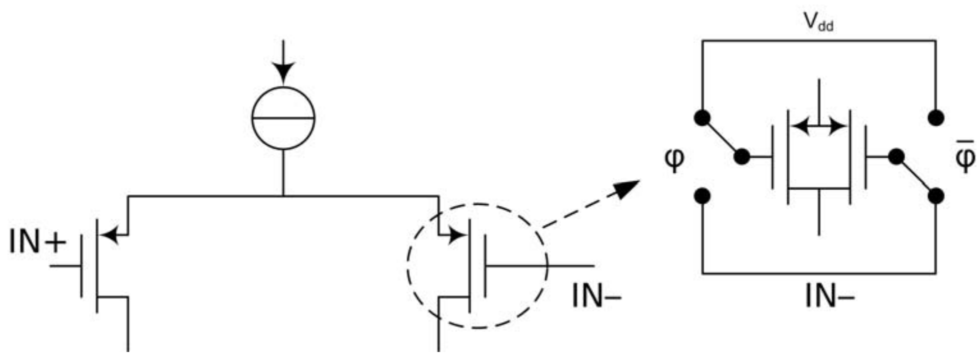

Figure 10: Switched differential pair.

noise will cause low-frequency phase noise, which is narrow-band noise around the oscillator centre frequency that is not less paradox in nature than flicker noise itself [12]. Periodically switching off MOSFETs in oscillators should reduce such low-f phase noise because it reduces flicker noise. This has been shown experimentally both for CMOS ring oscillators, where the measured phase noise often is lower than simulated [13], and for RF LC oscillators, where flicker noise can be reduced by using two alternatively switched tail transistors, similar to what has been done in Fig. 10 [14].

Figure 11 shows a pixel of an image sensor [5]. In this circuit, the photo diode accumulates charge while it is exposed to light. To read out, M1 is switched on, charging the floating diffusion to a high potential. This voltage is read out by activating M3, "row select". In a second step, the readout transistor between the wells is activated, transferring the photo charge to the floating diffusion. Then a second readout is made. The difference of the two measurements is formed, removing offset and also flicker noise. Flicker noise in this circuit comes mainly from M2, and it is possible to reduce the intrinsic flicker noise of M2 by resetting it after each read-out through pulling the column bus. This, however, can be a bad idea, as will be discussed in the section on correlated double sampling.

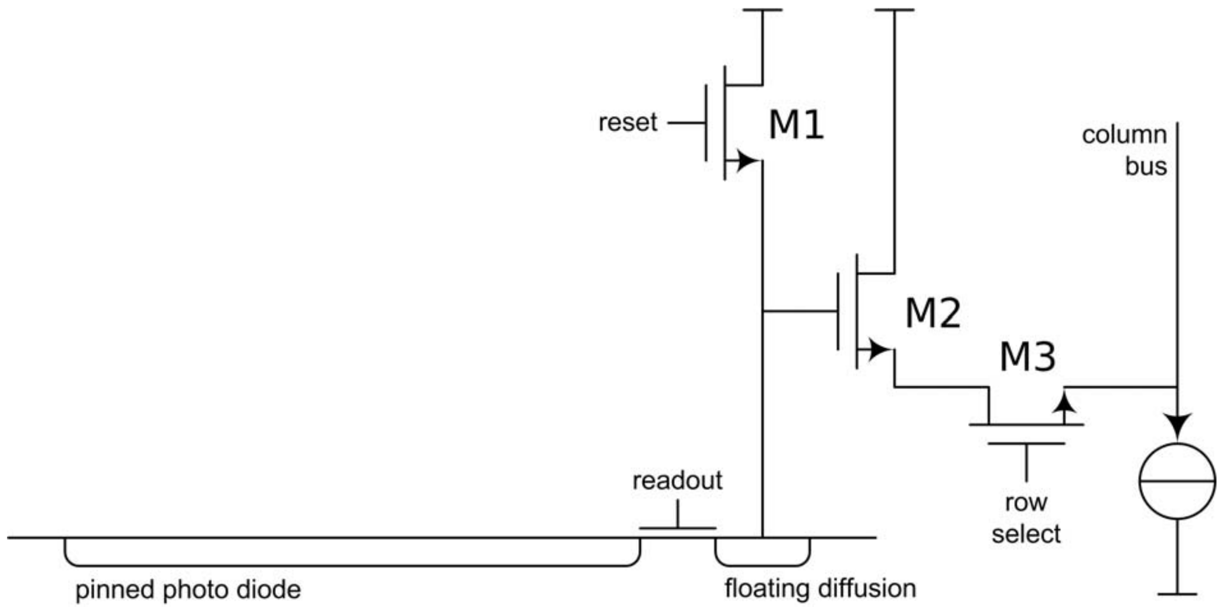

Figure 11: Image sensor pixel.

4 Chopping

Chopping is one of two fundamentally different ways to remove flicker noise from the signal. Chopping can be done whenever it is possible to feed the signal through the flickering amplifier with different signs in every other clock period. This chopping operation can then be reversed at the output, after the amplifier, as shown in Fig. 12.

Essentially this system *modulates* the input signal up to the frequency f_{chop}, and also $3f_{\text{chop}}$, $5f_{\text{chop}}$, and so on. Then the signal goes through the amplifier, where it picks up flicker noise and also offset. After the amplifier, the signal is modulated back to the base band, but at the same time, the flicker noise and the offset are modulated up to the multiples of f_{chop}. So, as long as f_{chop} is far enough above the signal band, the signal is not disturbed by flicker noise [10].

The formulae for the chopped noise spectrum can be found in [10], but Fig. 13 shows that the relations between the amplifier output noise and the spectrum after the second multiplier in Fig. 12 are really simple: below f_{chop}, the noise is white and on the level of the amplifier output noise at frequency f_{chop}. This makes it advisable to choose the chopper frequency f_{chop} at the $1/f$-noise corner frequency, or higher.

Note that chopping is just a modulation, it does not involve sampling! So while it is possible to use chopping in a sampled-data system, it is just as well possible to use it in a continuous-time system, where it will not do any noise aliasing.

It is equally important to note that chopping does not remove offset and flicker noise. For example, if the amplifier has an input offset of 1 mV, a gain of 100, no input signal, and $f_{\text{chop}} = 10\,\text{kHz}$, then its output will be a rectangular signal with frequency f_{chop} and a magnitude of $200\,\text{mV}_{\text{pp}}$! This means that when a signal is present, that signal will be added to this huge rectangular wave, and may well saturate the following stages, which is why most chopping systems have low-pass filters after the second chopper.

4.1 Conventional chopper amplifier

Fig. 14 shows a conventional amplifier. Although we draw a multiplier in Fig. 12, the chopper section is very simple to realise, all that is needed are four switches that cross the lines of the balanced amplifier during ϕ_2, or do not cross them during ϕ_1 [10]. The design constraints on such a system are:

- f_{chop} should be higher than the $1/f$-noise corner frequency and must be at least twice the signal band's upper frequency, f_{sig}.
- The amplifier will process the signal in the frequency band $f_{\text{chop}} \pm f_{\text{sig}}$, so it must work well and with sufficient slew rate in this frequency range.

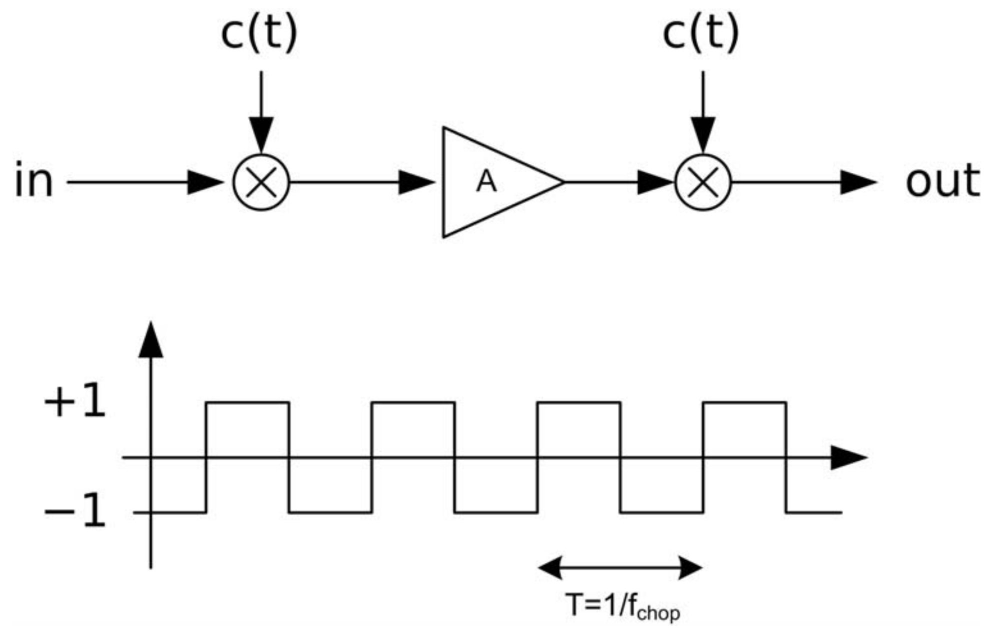

Figure 12: The principle of chopping.

- It is advisable to remove the energy of the chopped signal after the second chopper using a low-pass filter with passband up to f_{sig} and stop band below f_{chop}.
- The switches must be designed such that they result in as little charge injection as possible (see the section on switched-capacitor circuits); such charge injection will cause residual offset.

One way to reduce residual offset due to charge injection is shown in Fig. 15. In this amplifier, the inner chopper is designed at a frequency above the $1/f$ corner frequency, thus moving $1/f$ noise out of the signal band. A second outer chopper can then operate on a frequency below the $1/f$ corner frequency, it will remove the residual offset of the inner chopper, and will cause a low residual offset itself, because it operates at a low frequency. A 100 nV-offset nested chopper amplifier was reported in [15]. Note that in such amplifiers, f_{sig} must be lower than half of the *lower* chopper frequency.

Very good results can also obtained with tackling the residual offset at its source, for example by staggering the clock edges of the second chopper in Fig. 14 slightly behind the edges of the first chopper, leaving a small time gap in which the error pulses of the first chopper can die away [16].

4.2 Multi-path chopper amplifiers

Nevertheless, in all these examples, the chopper frequency must be above twice the maximum signal frequency. This limitation can be overcome by building a multi-path amplifier, as in Fig. 16.

If g_{m4} is chosen such that both the DC gain of the lower path and its unity-gain frequency are much lower than those of the upper path, a situation as in Fig. 17 occurs: the transfer functions of the two paths will cross at the frequency f_{cross}; below this frequency, the lower path will dominate the op-amp's behaviour; above f_{cross}, the upper path.

So it becomes possible to replace the lower path by a chopper amplifier as in Fig. 14, and operate it on a very low chopper frequency. [17] presents a chopper amplifier that has 1 μV offset, $f_{\text{chop}} = 4$ kHz, and a unity-gain frequency of 1.3 MHz with 50 pF load. This amplifier has more residual offset than the ones in [15] and [16], but the upper signal frequency of 1.3 MHz is large compared to the 5.6 kHz of [16] and huge compared to the 8 Hz (sic!) of [15]. This shows that the main frequency limitation of chopper amplifiers can be overcome, although with considerable circuit design effort.

4.3 Chopping in sampled-data systems

Finally, chopping can also be used in sampled-data systems. For example, Fig. 18 shows the cross section of a MEMS acceleration sensor and a block diagram of the read-out electronics.

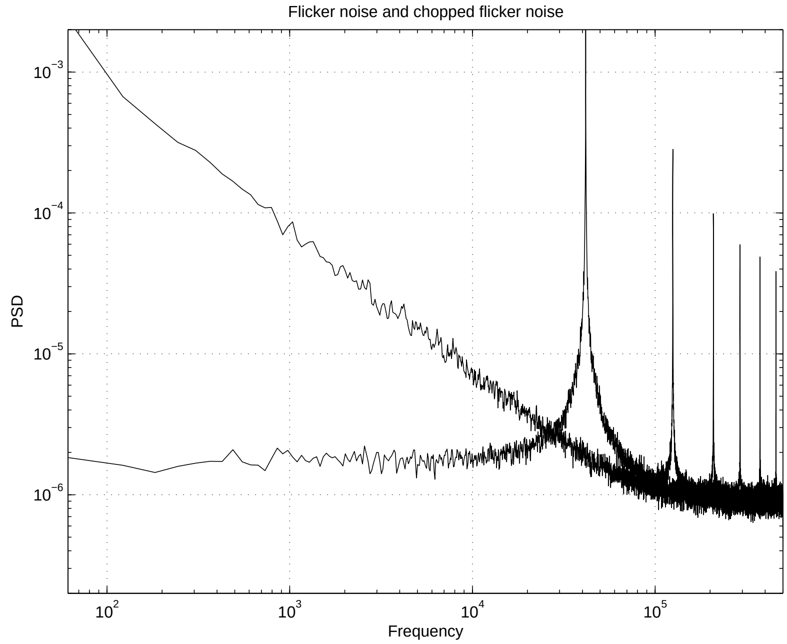

Figure 13: Flicker noise and chopped flicker noise.

The sensor is capacitive, with two rigid plates at the top and the bottom, and one plate that hangs in free space, attached by a spring, in the centre. When accelerated, the centre plate will move up or down, resulting in a different distribution of the capacitances towards the top plate and bottom plate. Since this is a linear electrical system, the position can be read out by measuring V_{centre} while either setting $V_{\text{top}} = V_{DD}$, $V_{\text{bottom}} = V_{SS}$; or by setting $V_{\text{top}} = V_{SS}$, $V_{\text{bottom}} = V_{DD}$. This will give the same value with opposite sign, which can be read out by a switched-capacitor low-noise amplifier. So doing the two possibilities alternatively amounts to chopping at the input of the amplifier (LNA).

If the offset and flicker noise is not too big in such a system, the output of the LNA can be digitised and the second chopper can be a simple digital sign change on the sampled value. However, if the offset or flicker noise are so big that the analog stages after the LNA are saturated, then it is necessary to add an analog second chopper and a filter after the LNA as in Fig. 14.

5 Correlated Double Sampling (CDS) and Auto-Zero techniques

The third idea to deal with flicker noise is to remove it after it has occurred. Techniques doing this are called "auto-zeroing" or "correlated double sampling". Both are fundamentally the same, what is done is to first sample without a signal (i.e., only the offset), and then sample again with a signal, and subtract the two values.

The effect on offset, ideally, is that it is removed, because the offset of the sampling amplifier will be the same for both samples. Flicker noise will mostly be removed, because two samples of a flicker noise process correlate well (see Sec. 1 and the discussion of the autocorrelation function of $1/f$ noise). White noise, however, does not correlate with earlier samples of itself, so the power of the white noise of the amplifier will simply be doubled.

This can be seen well in Fig. 19, which shows the spectrum of a process with flicker noise and white noise (bottom); the same process sampled, having ten times as much white noise; and double sampled, with twenty times as much white noise, but no flicker noise.

Fig. 19 shows CDS performed on a signal that had already been sampled. Sampling a continuous-time signal gives different results. We will now look at the white-noise and the flicker-noise contributions independently. For white noise whose bandwidth $B = \frac{\pi}{2} f_c$ is much larger than the input sampling frequency $2f_s$, the spectrum after CDS is [10]

$$S_{\text{CDS,white}} \approx \left(\pi \frac{f_c}{2f_s} - 1 \right) S_0 \text{sinc}^2 \left(\frac{\pi f}{2f_s} \right)$$

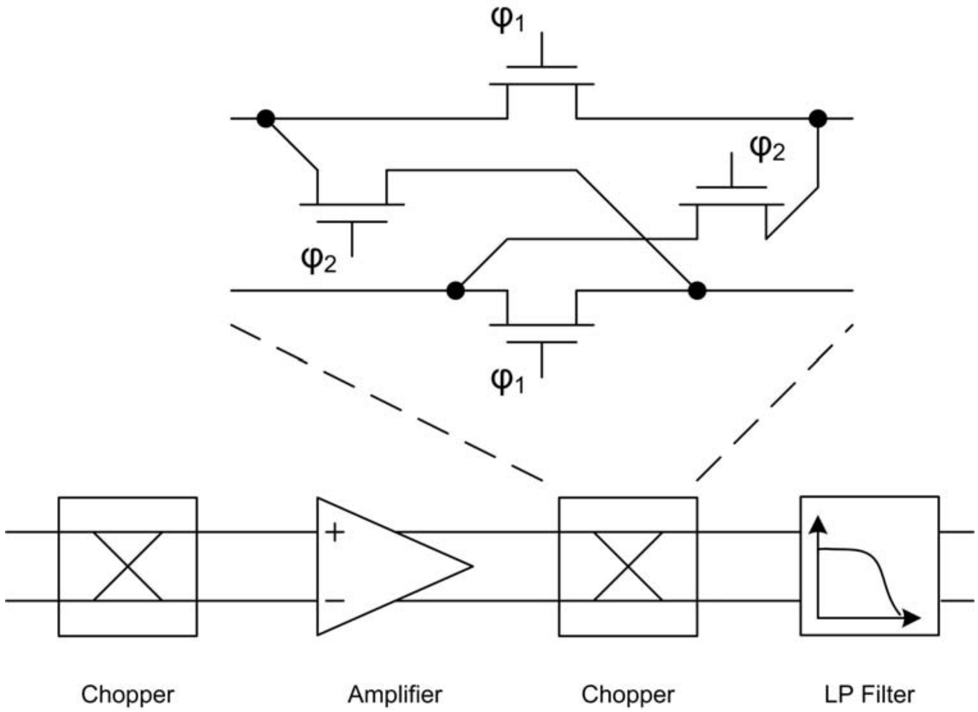

Figure 14: Chopper amplifier.

where f_c is the corner frequency of the white noise and S_0 is the DC noise level. Note that here we choose f_s to be the sampling frequency at the output of the CDS block, after two samples have been subtracted.

Similarly, the flicker noise at low frequencies will not disappear completely; a fold-over component will dominate at low frequencies:

$$S_{\text{fold},1/f} \approx \frac{S_0 f_{1/f}}{f_s}\left[1 + \ln\left(\frac{1}{3}\frac{f_c}{f_s}\right)\right]\text{sinc}^2\left(\frac{\pi f}{2 f_s}\right)$$

where $f_{1/f}$ is the corner frequency of the $1/f$ noise. The shape of the two spectra is exactly the same, and the different factors in front of the sinc function mean that as long as the flicker noise corner frequency $f_{1/f}$ is sufficiently far below the sampling frequency, aliased white noise will dominate the behaviour at low frequencies.

Fig. 20 shows the sum of these aliased components superimposed on Fig. 19. The effect of using CDS on a continuous-time signal is that while a simple calculations as in Sec. 2 or a simulation with sampled signals estimate the total noise correctly, both *underestimate* the low-f noise by a factor of $\frac{\pi}{2} = 1.57 = 4\,\text{dB}$! On the other hand, they overestimate HF noise at the upper end of the frequency band, where almost only white noise of power S_0 will be seen in reality.

5.1 Switched-capacitor comparator with CDS

Correlated double sampling is not very difficult to implement, and it is used in many applications. Our first example, Fig. 21, is a comparator that can be used in Flash A/D converters [18]. The operation of this comparator is simple: in phase ϕ_1 (Fig. 22 left), the input voltage V_{in} is sampled onto the capacitor C. Because of the closed negative feedback loop, the negative input of the amplifier settles to the offset voltage V_{os}, so C is charged to the voltage $V_{\text{in}} - V_{\text{os}}$. In the phase ϕ_2 (Fig. 22 right), the voltage $V_{\text{in}} - V_{\text{os}}$ between the negative input and V_{gnd} is compared to the new value of V_{os}, so the comparator actually tests whether

$$V_{\text{os}}|_{\phi_2} - \left(V_{\text{os}}|_{\phi_1} - V_{\text{in}}|_{\phi_1}\right) > 0$$

The difference $V_{\text{os}}|_{\phi_2} - V_{\text{os}}|_{\phi_1}$ is formed; this is correlated double sampling that removes offset and a lot of $1/f$ noise as explained above.

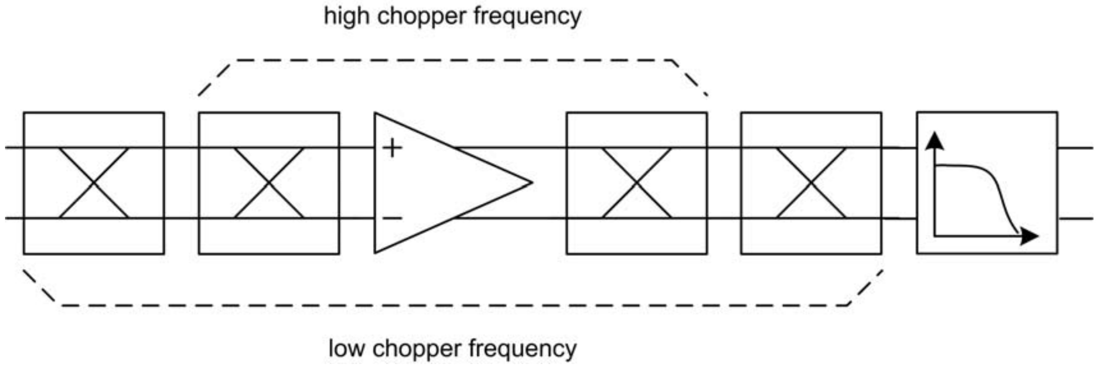

Figure 15: Nested chopper amplifier.

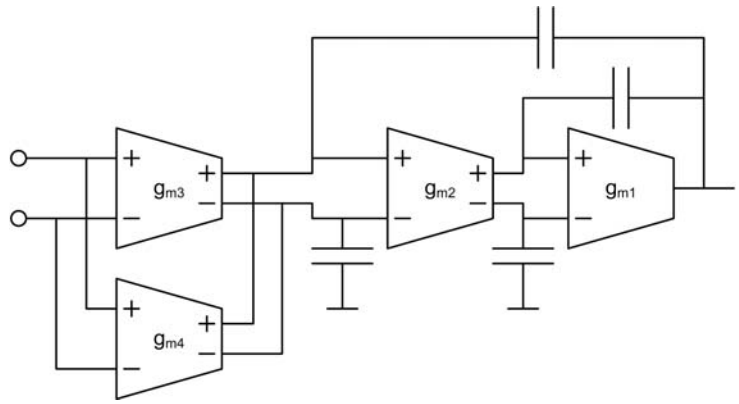

Figure 16: Multi-path amplifier.

As with all switched-capacitor circuits, the main difficulties of this circuit are parasitic charges injected when switches open. Apart from that, such a system can remove so much offset that the comparator can even be a simple CMOS inverter, as shown in Fig. 23 [18].

For a high-resolution comparator, an inverter will not have sufficient gain in the transition region, so an op-amp must be used, for example a Miller op-amp. The problem there is that during ϕ_1, the amplifier must be stable in the feedback loop, while during ϕ_2 it just has to be as fast as possible. A switchable compensation as shown in Fig. 24 will take care of this, and with proper scaling of the switch transistor, this switch, while on, will introduce a compensating zero in the Miller amplifier (c.f. [7]).

5.2 Switched-capacitor amplifier with CDS

The comparator in Fig. 21 can readily be modified to give an SC amplifier [10] by adding a capacitor that is switched into the signal path in phase ϕ_2, as shown in Fig. 25.

Then, in ϕ_1, C_1 will be charged to $V_{\text{in}} - V_{\text{os}}$, and C_2 to $-V_{\text{os}}$; in ϕ_2 the difference is formed and the resulting output voltage $V_{\text{out}} = V_{\text{in}} \cdot C_1/C_2$ with the offset and flicker noise removed.

While this circuit works well in practice, it has two problems: first, ϕ_1 and ϕ_2 must not overlap. This means that while neither is active, the amplifier is in an open-loop configuration, and care must be taken that the output does not jump to a supply rail during that time and pushes the op-amp into a state from which it takes long to recover. Second, during ϕ_1, the output is always V_{os}, so the output jumps forth an back between the signal voltage and the small voltage V_{os}, and thus the amplifier needs to have a high slew rate. [10] gives a good overview on SC amplifiers in which the amplifier needs to have only a modest slew rate.

5.3 Correlated double sampling in sampled systems

In Sec. 4, we introduced an example of an acceleration sensor — in Fig. 18 — and discussed chopping. This system can easily be transformed into a CDS system: the sensor is operated with the same sequence

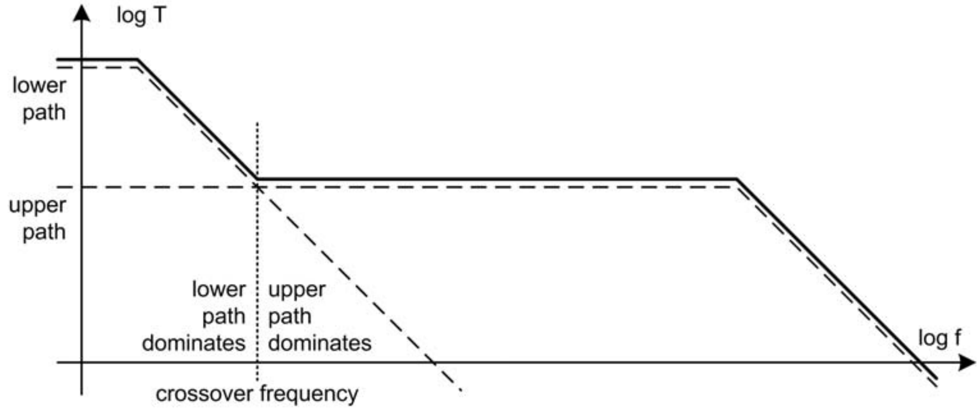

Figure 17: Open-loop transfer function of the multi-path amplifier.

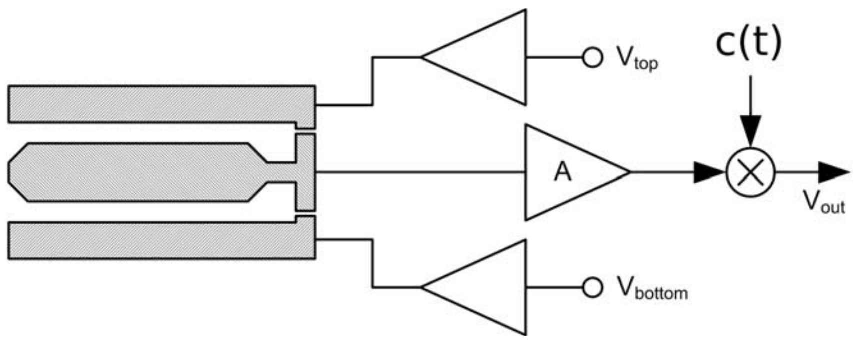

Figure 18: Acceleration sensor with SC LNA of gain A.

as in Sec. 4, but instead of only changing the sign of every second sample, we then also form the difference of two consecutive samples.

The advantage of this system is that offset and flicker noise are removed — and not only modulated out of the signal band — so amplifiers and A/D converters after the CDS stage are not in danger of being saturated. The clear disadvantages are that white noise is doubled, and also that now two input samples are needed to provide one output sample. The latter means that either the time available for sampling has to be cut in two pieces, requiring faster amplification than in the chopper system, or that two circuits forming differences are operated in parallel, one making $V[2n+1]-V[2n]$, the other making $V[2n+2]-V[2n+1]$.

5.4 Correlated double sampling combined with LSE

In Fig. 11 we showed a simple photo diode readout circuit, in which large-scale-excitation was used to reduce the intrinsic noise of the readout transistor. Simultaneously, correlated double sampling is also used.

The problem is that doing both at the same time can give *more* flicker noise instead of less flicker noise [5]. LSE resets the memory of the transistors, so after they return into their operating point, the memory starts to fill up again, and the variance of the flicker noise will start to increase as shown in Fig. 4. So if the two samples used for CDS are taken at two times when the variance of the flicker noise is very different, the two samples do not really correlate and CDS can increase the flicker noise instead of cancelling it. This is extremely difficult to simulate, but has been shown by measurement in [5].

This means in general: it is not enough that the transistor biasing is the same at both sample times, the *history* of the biasing also needs to be as similar as possible at both sampling instants.

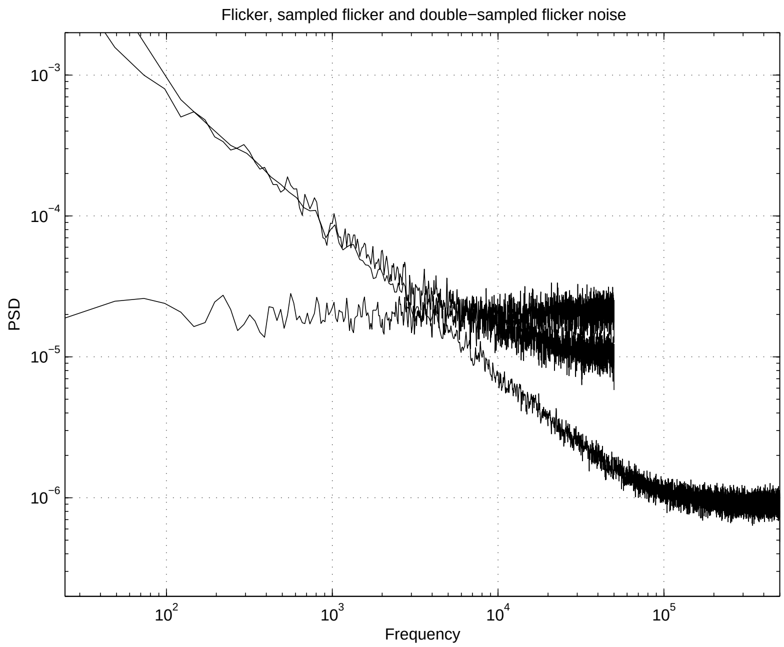

Figure 19: Flicker noise subjected to sampling and correlated double sampling.

6 Conclusion

Not all three methods to fight flicker noise can be used in every system. Large-scale excitation is mostly used — or happens by itself — in sensor circuits with low transistor count, and in oscillators: it cannot be simulated, and calculating it is also difficult. Correlated double sampling is mostly used in systems that process sampled data, or are designed to sample data. Chopping is mostly used in continuous-time systems.

This chapter has given an introduction into all three techniques, together with a description of the nature of flicker noise, and of noise sampling. The literature in the References section was chosen carefully to give the interested reader starting points for going deeper into different aspects of flicker noise; the four main papers to read would be [1] for the mathematics of flicker noise, [6, 5] for its physics, and [10, 5] for cicruit solutions.

Appendix

Fig. 26 shows the Matlab/Simulink model use to make the simulations for this book chapter. Fig. 1 was made with the following three scripts:

flickr˙fig01.m

```
%
% Hanspeter Schmid, June 2007
%
% Draw the flicker noise / white noise spectrum
%
clear

poles_and_zeros

sim('flickr_gen')

save data_fig01
```

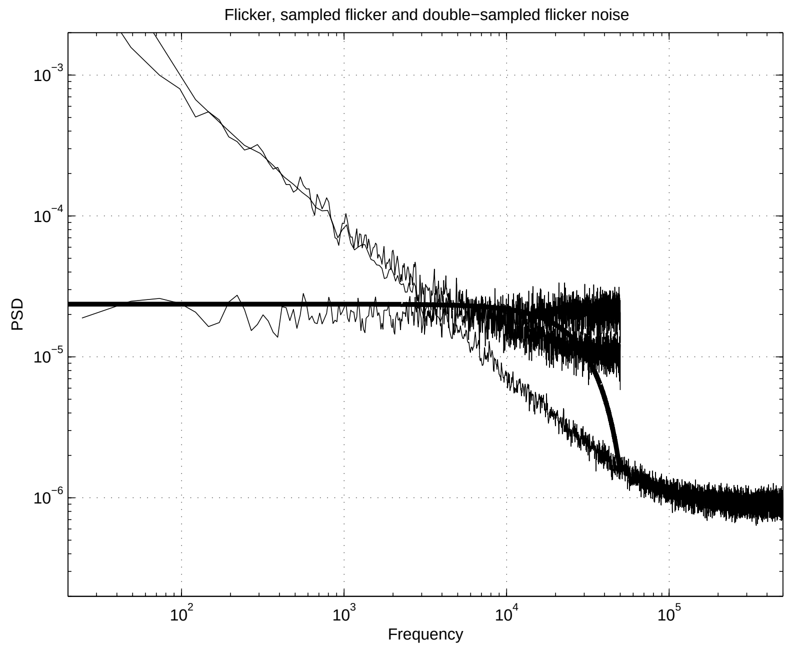

Figure 20: Continuous-time flicker noise subjected to correlated double sampling

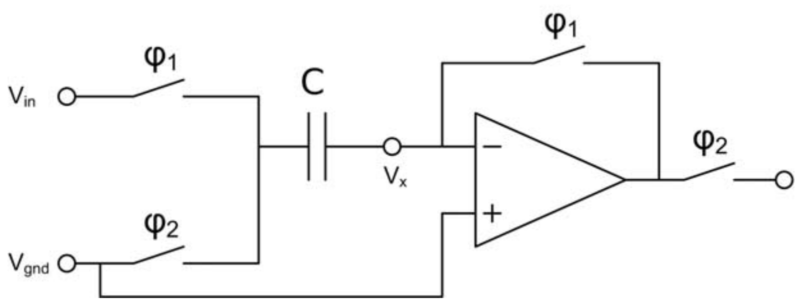

Figure 21: Comparator with CDS.

poles˙and˙zeros.m

```
%% Poles and Zeros for the flicker noise generator
divideDecade=8;
fMax=4e5;
iMax=floor(log10(fMax)*divideDecade)

Vpoles = [];
Vzeros = [];

tSim=1;
rSeed=26649;

for k=1:4:iMax
    Vpoles=[Vpoles -10^((k+1)/divideDecade)];
    Vzeros=[Vzeros -10^((k+3)/divideDecade)];
end

format compact
Vpoles
```

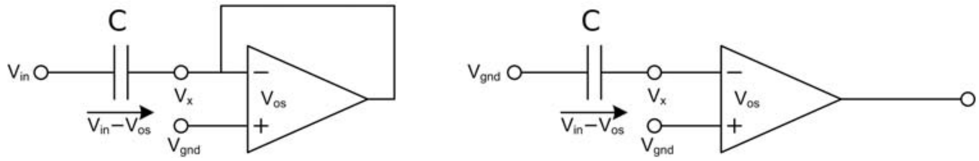

Figure 22: The CDS comparator in phases ϕ_1 (left) and ϕ_2 (right).

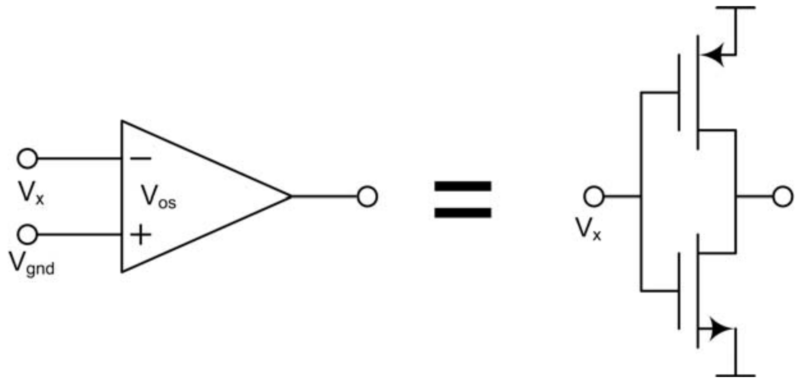

Figure 23: Inverter-based comparator with CDS.

```
Vzeros
Vgain=1/10^(1/divideDecade);
```

plot˙fig01.m

```
load data_fig01

[Pxx,w]=pwelch(flickr.signals.values(1:1e6+1),hann(2^14),2^13,2^14,1e6);
loglog(w,Pxx,'r');
axis([min(w) max(w) 2e-7 2e-3])
grid
grid minor

xlabel('Frequency')
ylabel('PSD')
title('Power spectral density of ideal flicker noise and white noise')

print -deps2 matlab_fig01.eps
```

Biography

Hanspeter Schmid (hanspeter@schmid-werren.ch) received the diploma in electrical engineering in 1994, the post-graduate degree in information technologies in 1999, and the degree Doctor of Technical Sciences in 2000, all from the Swiss Federal Institute of Technology (ETH Zrich), Switzerland.

He joined the Signal and Information Processing Laboratory of the ETH Zürich as a teaching assistant in 1994 and worked there as a lecturer and research assistant in the field of analog integrated filters. From 2000–2005 he was an Analog-IC Designer with Bernafon AG, Switzerland, where he was part of a design team who developed a new mixed-signal IC platform for a new generation of hearing aids that do wireless ear-to-ear communication. Now he is a Research Fellow and Industry Consultant at the Institute of Microelectronics (http://www.ime.technik.fhnw.ch/) of the University of Applied Sciences Northwestern Switzerland, and a senior lecturer at ETH Zrich (Analog Signal Processing and Filtering).

Hanspeter's research interest lie in low-power high-speed circuits and systems, mostly sensor systems, and sigma-delta systems. Hanspeter Schmid presently is the chair elect of the Analog Signal Processing

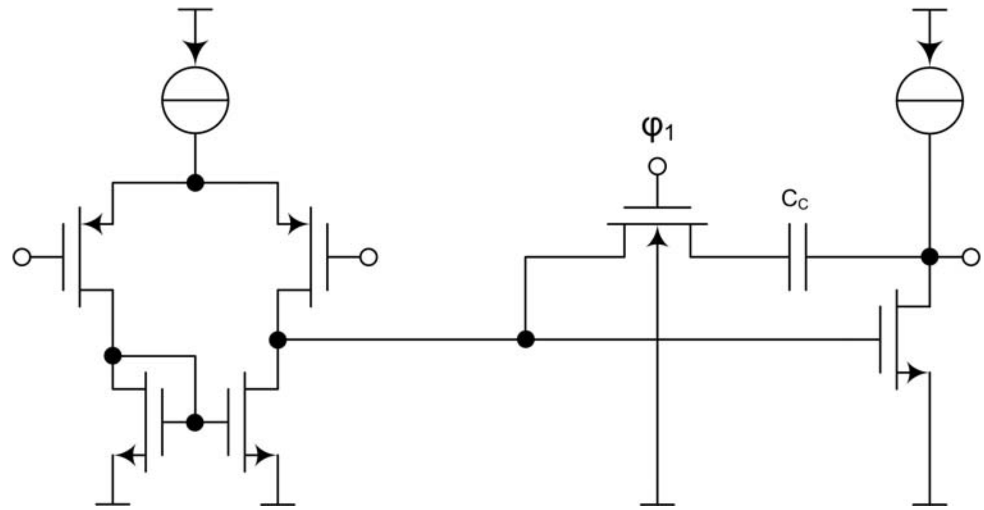

Figure 24: Miller amplifier with switchable compensation.

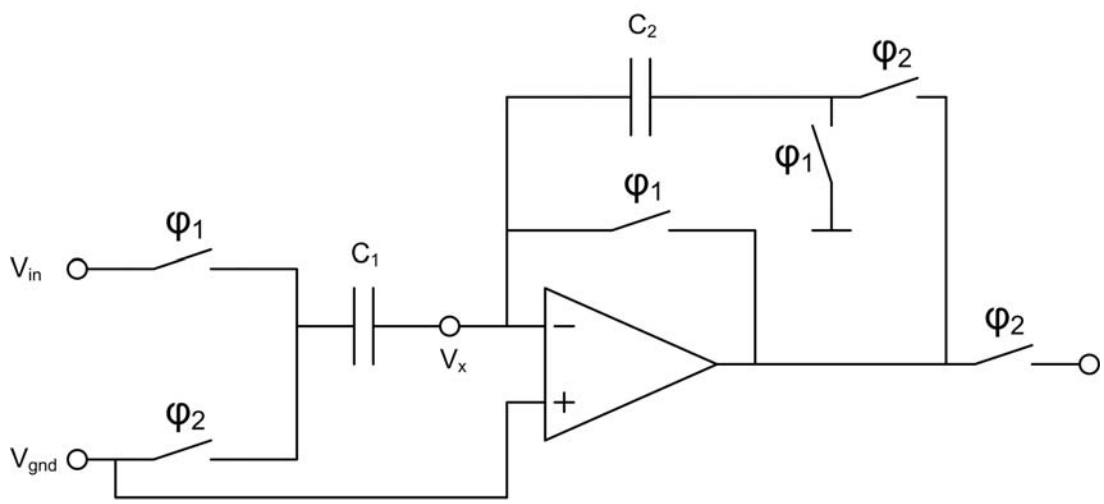

Figure 25: SC amplifier with correlated double sampling.

Technical Committee of the IEEE CAS Society, an Associate Editor of the IEEE Transactions on Circuits and Systems, Part I, and he serves as a reviewer for several journals and conferences. His publications can be found on `http://www.schmid-werren.ch/hanspeter/publications/`.

References

[1] D. Slepian, "On bandwidth," *Proceedings of the IEEE*, vol. 64, no. 3, pp. 292–300, Mar. 1976.
Slepian shows in this paper, which is the paper version of a Shannon Lecture, that time is real and frequency is not. He does this with a philosophicval discussion of the role of mathematical models in the exact sciences, and gives a new (mathematical) version of the 2WT theorem.

[2] M. Keshner, "$1/f$ noise," *Proceedings of the IEEE*, vol. 70, no. 3, pp. 212–218, Mar. 1982.
This paper describes the main properties of $1/f$ noise (memory, autocorrelation function, etc.) and also shows how flicker noise can be derived from white noise with a relatively simple filter. The main impact of this paper are its mathematical-philosophical implications.

[3] H. Schmid, "Aaargh! I just loooove flicker noise," *IEEE Circuits and Systems Magazine*, no. 1, pp. 32–35, Jan. 2007.

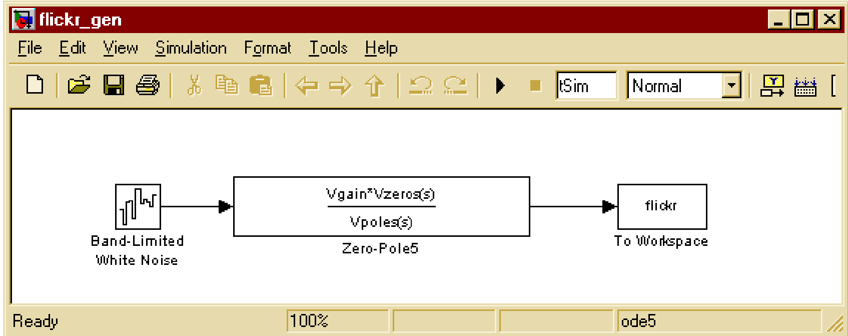

Figure 26: Simulink model generating flicker noise.

This is a lightly written column on flicker noise that aims at telling the reader qualitatively what flicker noise is. It also shows how even internet search engines flicker.

[4] I. Bloom and Y. Nemirovsky, "$1/f$ noise reduction of metal-oxide-semiconductor transistors by cycling from inversion to accumulation," *Applied Physics Letters*, vol. 58, no. 15, pp. 1664–1666, Apr. 1991.

This is the first paper that experimentally demonstrated a reduction of flicker noise by using large-scale excitation of MOSFETs. It contains an intuitive explanation of why this works, but no mathematics.

[5] A. P. van der Wel, E. A. M. Klumperink, J. S. Kolhatkar, E. Hoekstra, M. F. Snoeij, M. Cora Salm, H. Wallinga, and B. Nauta, "Low-frequency noise phenomena in switched MOSFETs," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 3, pp. 540–550, Mar. 2007.

Apart from a very good summary of the sources of flicker noise and random-telegraph noise in MOSFETs, and a precise explanation of how flicker noise can be reduced with large-scale excitation, the authors also give measurements showing that combining LSE with correlated double sampling can be a very bad idea.

[6] A. van der Ziel, "Unified presentation of $1/f$ noise in electronic devices: Fundamental $1/f$ noise sources," *Proceedings of the IEEE*, vol. 76, no. 3, pp. 233–258, Mar. 1988.

This paper gives an in-depth mathematical and practical discussion of all surface and volume flicker noise sourcves in semiconductors and other devices. It also distinguishes fundamental and non-fundamental flicker noise sources.

[7] D. A. Johns and K. Martin, *Analog Integrated Circuit Design*. New York: John Wiley & Sons, 1997.

One of the best books for teaching analog IC design, because many things are described qualitatively and intuitively, as well as with formulas.

[8] C.-A. Gobet and A. Knob, "Noise analysis of switched capacitor networks," *IEEE Transactions on Circuits and Systems*, vol. 30, no. 1, pp. 37–43, Jan. 1983.

This paper shows how the noise in a simple SC integrator can be analyzed; it includes both continuous-time noise and sampled noise at the output, shows the frequency shaping due to the sample-and-hold process, and discusses dominating noise effects.

[9] L. Tóth, I. Yusim, and K. Suyama, "Noise analysis of ideal switched-capacitor networks," *IEEE Transactions on Circuits and Systems–I*, vol. 46, no. 3, pp. 349–363, Mar. 1999.

This paper shows how analyses such as in [8] can be made with matrices, hence making some sort of automation possible, or simplifying the use of symbolic analysis tools such as Mathematica.

[10] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization," *Proceedings of the IEEE*, vol. 84, no. 11, pp. 1584–1615, Nov. 1996.

The OpAmp imperfections this paper decribes are finite gain, offset, and flicker noise. It is a very good overview paper, with some mathematics, many examples, and a good list of references. Many good graphics that can be used for teaching are presented. The figures are distributed funnily over the text, though, and it is worthwile to use a marker to mark the first occurence of every figure reference in the text.

[11] J. Koh, D. Schmitt-Landsiedel, R. Thewes, and R. Brederlow, "A complementary switched MOSFET architecture for the $1/f$ noise reduction in linear analog CMOS ICs," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 6, pp. 1352–1361, June 2007.

Koh et. al. show how large-scale excitation can successfully be applied to a differential pair. The paper gives both a thorough mathematical analysis and measurements of an implemenation in 120-nm CMOS.

[12] F. M. Gardner, "Can analog PLLs hold lock? A paradox explored," *IEEE Circuits and Systems Magazine*, no. 3, pp. 46–52, July 2007.

This column discusses the paradox that, in theory, analog PLLs cannot lock in the presence of $1/f^3$ phase noise, but do of course lock in practice. The discussion is closely related to the RMS of flicker noise that rises with time, because the source of the $1/f^3$ phase noise is flicker noise.

[13] S. L. J. Gierkink, E. A. M. Klumperink, A. P. van der Wel, G. Hoogzaad, E. van Tuijl, and B. Nauta, "Intrinsic $1/f$ device noise reduction and its effect on phase noise in CMOS ring oscillators," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 7, pp. 1022–1025, July 1999.

A ring oscillator built from HEF4007 MOS ICs is shown to have 8 dB lower $1/f^3$ phase noise than expected from theory and simulation. It is shown that this phase noise depends on the MOS gate voltages, and that the obserbed reduction can be attributed to the large-scale excitation of the transistors used in the ring oscillator.

[14] C. C. Boon, M. A. Do, K. S. Yeo, J. G. Ma, and X. L. Zhange, "RF CMOS low-phase-noise LC oscillator through memory reduction tail transistor," *IEEE Transactions on Circuits and Systems–II*, vol. 51, no. 2, pp. 85–90, Feb. 2004.

This paper shows that the flicker-noise-reduction effect observed in [13] for ring oscillators also occurs in LC oscillators when the tail transistors are subject to large-scale excitation.

[15] A. Bakker, K. Thiele, and J. H. Huijsing, "A CMOS nested-chopper instrumentation amplifier with 100-nV offset," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 12, pp. 1877–1883, Dec. 2000.

This paper shows how very low offset can be achieved in an amplifier, at the expense of bandwidth, by using an outer slow chopper and an inner fast chopper. The outer chopper then reduces the residual offset, while the inner chopper removes $1/f$ noise.

[16] Q. Huang and C. Menolfi, "A 200nV offset 6.5nV/$\sqrt{\text{Hz}}$ noise psd 5.6kHz chopper instrumentation amplifier in 1μm digital CMOS," in *Proceedings of the IEEE International Solid-State Circuits Conference*, (San Francisco), p. 23.3, Feb. 2001.

In the chopper instrumentation amplifier presented in this paper, a simple and effective way to minimize charge injection in a chopper amplifier is shown. It relies on proper timing of the switch control signals.

[17] J. F. Witte, K. A. A. Makinwa, and J. H. Huijsing, "A CMOS chopper offset-stabilized opamp," in *Proceedings of the European Solid-State Circuits Conference*, (Montreux, Switzerland), pp. 360–363, Sept. 2006.

The amplifier presented in this paper has two forward paths: a slow, chopped path with high gain, and a fast, not chopped path with lower gain. Using this methods, the authors build an amplifier that has the good offset performance of a chopper amplifier, and the speed of a non-chopped amplifier.

[18] R. Gregorian, *Introduction to CMOS Op-Amps and Comparators*. New York: John Wiley, 1999.

A very good book on operational amplifiers an comparators, with a detailed view on their application in sampled-data systems and data converters.