MOSFET LF NOISE UNDER LARGE SIGNAL EXCITATION

MEASUREMENT, MODELLING AND APPLICATION

Arnoud van der Wel

Samenstelling promotiecommissie:

Voorzitter:	prof.dr.ir. J. van Amerongen	Universiteit Twente, EWI
Secretaris:	prof.dr.ir. A.J. Mouthaan	Universiteit Twente, EWI
Promotor:	prof.dr.ir. B. Nauta	Universiteit Twente, EWI
Assistent Promotor:	dr.ing. E.A.M. Klumperink	Universiteit Twente, EWI
Referent:	dr. R. Woltjer	Philips Research, Eindhoven
Leden:	prof.dr. H. Wallinga	Universiteit Twente, EWI
	prof.dr.ir. C.H. Slump	Universiteit Twente, EWI
	prof.dr.ir. A.J.P. Theuwissen	TU Delft
	prof.dr. J. Schmitz	Universiteit Twente, EWI

Title: MOSFET LF NOISE UNDER LARGE SIGNAL EXCITATION: MEASUREMENT, MODELLING AND APPLICATION

Author: Arnoud van der Wel

ISBN: 90-365-2173-4

This research was supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.

MOSFET LF NOISE UNDER LARGE SIGNAL EXCITATION

MEASUREMENT, MODELLING AND APPLICATION

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. W.H.M. Zijm,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op 15 april 2005 om 15.00 uur

door

Arnoud Pieter van der Wel
geboren op 31 mei 1974
te Bunschoten

Dit proefschrft is goedgekeurd door:

de promotor prof.dr.ir. B. Nauta

de assistent promotor dr.ing. E.A.M. Klumperink

Contents

List of Symbols

Symbol	Explanation	Units
$\underline{1}$	Heaviside unit step function	–
A	Area of MOSFET (=WL)	m^2
B	Bandwidth	Hz
C	Channel capacity	bits s^{-1}
$C_{\mathrm{ox}}^{\square}$	Oxide capacitance per square meter	Fm^{-2}
C_{xx}	Autocovariance of a stochastic signal $x(t)$	
dc	Duty cycle, between 0 and 1	–
$f(E,t)$	Occupancy function of traps at energy level E and time t	–
f_{c}	$1/f$ noise corner frequency, where $S_{1/\mathrm{f}} = S_{\mathrm{th}}$.	Hz
f_{s}	Sample frequency ($= 1/t_{\mathrm{s}}$)	Hz
g_{m}	Transconductance	AV^{-1}
h	Planck's constant, 6.62×10^{-34}	Js^{-1}
I_{D}	Drain current	A
k	Boltzmann's constant, 1.38×10^{-23}	JK^{-1}
K	Arbitrary constant	–
L	Length of MOSFET	m
m	τ change factor between two bias states of MOSFET (m_{e} and m_{c} are for τ_{e} and τ_{c} respectively)	–

Symbol	Explanation	Units
m_x	Average value of a stochastic signal x	
n	Electron concentration	m^{-3}
N	Noise power	W
N	Total number of free carriers in a sample	–
N_c	Effective density of states in the conduction band	m^{-3}
N_lin	Total number of free carriers in MOSFET in linear region	–
N_sat	Total number of free carriers in MOSFET in saturation	–
$N^\square$	Number of free carriers per unit area	m^{-2}
P	Power	W
q	Elementary charge, 1.6×10^{-19}	C
Q_SS	Charge stored in surface states	C
$r_\mathrm{c}(E,t)$	Capture rate	s^{-1}
$r_\mathrm{e}(E,t)$	Emission rate	s^{-1}
R	Resistance	Ω
R_{xx}	Autocorrelation of a stochastic signal $x(t)$	
R_{RTS}	Autocorrelation function of RTS	
s	Scaling factor from current to future CMOS process generation (> 1)	–
S	Signal power	W
S_x	Power Spectral Density of x	$\mathrm{W\ Hz}^{-1}$
$S_\mathrm{1/f}$	Power Spectral Density of $1/f$ noise	$\mathrm{W\ Hz}^{-1}$
S_th	Power Spectral Density of thermal noise	$\mathrm{W\ Hz}^{-1}$
$S_{\Delta N}$	Power Spectral Density of ΔN	Hz^{-1}
$S^\square_{\Delta N}$	Power Spectral Density of ΔN per unit area	$\mathrm{Hz}^{-1}\mathrm{m}^{-2}$
S_{Q_SS}	Power Spectral Density of Q_SS	$\mathrm{C}^2\mathrm{Hz}^{-1}$
T	Period of a periodic signal	s

Symbol	Explanation	Units
	Absolute temperature	K
t_s	Sample time ($= 1/f_s$)	s
v_{th}	Thermal velocity of electrons	ms^{-1}
V_{DS}	Drain to Source voltage	V
V_{GT}	Gate overdrive voltage: V_G-V_T	V
$\text{var}(x)$	Variance of a stochastic signal $x(t)$	
W	Width of MOSFET	m
$x(f)$	Frequency domain description of signal	Hz^{-1}
$x(t)$	Time domain description of signal	
α	Coupling coefficient between number and mobility fluctuations used in the Hung LF noise model	–
α_H	Hooge empirical factor	–
β	Asymmetry factor of RTS; 1 for symmetric RTS	–
μ	Mobility of mobile charge carriers	$m^2\ V^{-1}\ s^{-1}$
σ	Conductivity	$CV^{-1}s^{-1}m^{-1}$
σ_{V_T}	Standard deviation of V_T	V
$\sigma(E,x)$	Capture cross section of trap	m^2
τ_c	Mean time before capture event occurs	s
τ_e	Mean time before emission event occurs	s
$\tau_{eff,on}$	Time constant of occupancy change when device is 'on'	s
$\tau_{eff,off}$	Time constant of occupancy change when device is 'off'	s
ω_{0RTS}	RTS corner frequency	$rad\ s^{-1}$
ξ	Channel thermal noise form factor, ≥ 1	–

List of Abbreviations

Abbreviation	Meaning
CDS	Correlated Double Sampling
DUT	Device Under Test
FT	Fourier Transform
IF	Intermediate Frequency
IFT	Inverse Fourier Transform
ISF	Impulse Sensitivity Function
LF	Low Frequency
LO	Local Oscillator
LSE	Large Signal Excitation
PFD	Phase-Frequency Detector
PLL	Phase-Locked Loop
PSD	Power Spectral Density
RF	Radio Frequency
RTS	Random Telegraph Signal
SRH	Shockley-Read-Hall
VCO	Voltage-Controlled Oscillator

Samenvatting

Moderne elektronische schakelingen maken veelal gebruik van CMOS technologie waarin MOSFETs fungeren als actieve elementen. Het gebruik van een MOSFET als versterkend element biedt behalve veel voordelen zoals de lage kosten, de hoge snelheid, en de goede integreerbaarheid ook nadelen, met name de notoire laagfrequente ruis ervan. Het laagfrequent (LF) ruisgedrag van MOSFETs wordt gedomineerd door plaatsgebonden energietoestanden ('traps') aan de Si-SiO_2 grenslaag die een elektron tijdelijk kunnen invangen en daarmee de geleiding van de MOSFET beïnvloeden. In dit proefschrift wordt de LF ruis van MOSFETs bestudeerd onder grootsignaal biascondities.

Onder grootsignaal biascondities vertonen MOSFETs vaak een afname in LF ruis. Bovendien komen grootsignaal biascondities nauw overeen met de manier waarop MOSFETs in echte schakelingen gebruikt worden. Aan drie belangrijke punten wordt in dit proefschrift aandacht besteed. Ten eerste: Hoe kan LF ruis onder grootsignaal biascondities gemeten worden, ten tweede: Hoe komt het dat de ruis vaak afneemt onder grootsignaal biascondities, en tot slot: Is de LF ruisreductie significant en bruikbaar genoeg om toe te passen bij het ontwerp van analoge schakelingen met MOSFETs?

In hoofdstuk 3 wordt ingegaan op het meten van LF ruis onder grootsignaal biascondities. Tevens worden daar de belangrijkste meetresultaten gepresenteerd. Het meten van LF ruis onder grootsignaal condities stelt hoge eisen aan het dynamisch bereik van de meetopstelling omdat de LF ruis veel kleiner is dan de tijdvariante bias. Er worden drie methoden toegepast. Ten eerste kan door differentieel te werken, het bias signaal gemeenschappelijk aangeboden worden, terwijl de ruis differentieel blijft. Op deze wijze wordt het gemeenschappelijke signaal door de meetopstelling onderdrukt, en wordt het mogelijk de ruis te meten. Ten tweede kan de ruis van het bias-signaal gescheiden worden in de tijd, dus er wordt eerst een bias-puls

aangeboden en kort daarna wordt de ruis gemeten, en tot slot kan de ruis en het bias-signaal in frequentie gescheiden worden: de ruis is laagfrequent en het bias-signaal hoogfrequent.

Met behulp van deze meetmethoden worden metingen gepresenteerd aan verschillende MOSFETs. Waar grote MOSFETs systematisch een afname van de LF ruis vertonen in grootsignaal bedrijf, is het gedrag van kleine MOSFETs met een oppervlak van ver onder de $1\mu m^2$ veel minder voorspelbaar. Deze transistoren vertonen al in de stationaire toestand veel variatie in hun LF ruis, en bovendien is niet met zekerheid van te voren vast te stellen hoe ze zullen reageren op grootsignaal biascondities. Meestal neemt de LF ruis af, maar soms neemt deze ook toe. Dit komt omdat het gedrag van deze transistoren door slechts één of enkele 'traps' gedomineerd wordt. Zowel n- als p-kanaals MOSFETs vertonen soortgelijk gedrag, hetgeen er op duidt dat het ruisgenererende proces hetzelfde is. Tevens wordt gemeten dat het periodiek variëren van zowel de 'gate' als de 'source' spanning van de MOSFET een vergelijkbaar effect heeft op de LF ruis ervan.

In hoofdstuk 4 wordt een verklaring gepresenteerd voor de meetresultaten aan de hand van de al eerder genoemde 'traps', waarvan het gedrag wordt beschreven door de Shockley-Read-Hall theorie. Aan de hand van deze theorie, toegepast onder grootsignaal biascondities, wordt aangetoond dat de ruis onder grootsignaal biascondities wordt gedomineerd door traps die zich dichter bij het midden van de bandgap bevinden. Aangezien de trapdichtheid in het midden van de bandgap vaak lager is dan in de buurt van de geleidingsband, kan zo worden uitgelegd dat de ruis gemiddeld gezien afneemt in grootsignaal bedrijf. Om de voorspellingen van het model te toetsen wordt een simulator beschreven die in staat is het gedrag van de transistoren nauwkeurig te reproduceren. Ook transitief ruisgedrag is goed te begrijpen op basis van het model en de simulator.

Als laatste wordt in ***hoofdstuk 5*** aandacht besteed aan de vraag of grootsignaal excitatie voor circuitontwerpers praktisch toepasbaar is om LF ruis van MOSFETs tegen te gaan. In tijdcontinue schakelingen blijkt het mogelijk grootsignaal excitatie als ruisreductietechniek te gebruiken, maar moet vergeleken worden met een modulerende structuur die de LF ruis geheel verwijdert. Het gebruik van grootsignaal excitatie in dit soort situaties is dan ook niet zonder meer aan te bevelen. In tijddiscrete schakelingen is het eveneens mogelijk grootsignaal excitatie nuttig toe te passen als ruisreductie techniek, temeer daar de schakeling tussen de samplemomenten uitgeschakeld kan worden, en er zodoende geen noemenswaardige bezwaren aan het gebruik ervan kleven. Bij 'correlated double sampling' structuren moet er

voor gezorgd worden dat de biasgeschiedenis van beide samplemomenten identiek is, daar de ruis anders eerder toe dan af zal nemen. Ook hoogfrequent schakelingen hebben last van LF ruis door ‘opconversie’ van de ruis. Aangezien de LF ruisreductie niet gevoelig blijkt te zijn voor de frequentie waarmee de MOSFET aan en uit gezet wordt, kan in deze schakelingen ook nuttig gebruik gemaakt worden van grootsignaal excitatie als ruisreductie techniek.

Hoewel ruisreductie door grootsignaal excitatie weliswaar een meetbare afname van de LF ruis van MOSFETs oplevert, is de winst beperkt en afhankelijk van het proces in kwestie. Om grootsignaal excitatie ten behoeve van ruisreductie nuttig toe te kunnen passen moet eerst een voldoende nauwkeurig stationair LF ruismodel beschikbaar zijn. Vervolgens moet het LF ruisgedrag van het proces in kwestie onder grootsignaal condities gekarakteriseerd worden.

Chapter 1

Introduction

1.1 Noise is everywhere

Whenever people communicate, whether by smoke signals, waving their hands, shouting, using a satellite phone or an optical fiber, noise limits how well they can communicate and how much effort is needed to get the message across. If you are shouting to someone on the other side of a busy road, the noise of the traffic makes it difficult to communicate. You will either have to raise your voice, speak more slowly or repeat yourself if you want to be understood.

Electronic circuits are limited by noise in much the same way. Going against my firm conviction that equations do not belong in a introduction, I will present to you what is perhaps the single most important equation in the field of information theory. Indeed, this is the equation that *defines* information theory and that is therefore at the foundation of our modern information society. It was proposed by Claude Shannon in 1948 [58][1]:

$$C = B\log_2\left(1+\frac{S}{N}\right) \tag{1.1}$$

This equation features prominently in the first few pages of any text on information or communication theory, so it is not necessary to discuss it in any depth here. The *form* of the equation is recognizable, however. It formalizes what was said in the first paragraph, namely that how well you

[1] It is fortunate that Shannon chose to publish his work at all, since he was reportedly none too impressed with it, and published it only after repeated urging from colleagues.

can communicate (C) depends on how much effort you spend (S) and what the noise level is (N).

Noise then, if you like, is the foundation of the foundation.

Not only in information theory does noise occupy a central position, but also in physics. Any dissipative element exhibits thermal noise. This includes (but is not limited to): wires, resistors, transistors, diodes, inductors, capacitors, transformers, MOSFETs, and opamps. It is a safe bet that if it can be soldered, requires batteries or can be plugged into the mains, it exhibits thermal noise.

Thermal noise is not the only type of noise, however. Many components exhibit excess noise at low frequencies. The term 'excess' indicates 'more than expected from the theory of thermal noise'. This excess noise is variously referred to as 'flicker noise' or 'LF noise', and it includes such phenomena as $1/f$ and Random Telegraph Signal (RTS) noise.

The low frequency (LF) noise of MOSFETs is particularly interesting: a modern MOSFET is so small that the effects of single electrons can be seen. One example of such noise is in figure 1.1.

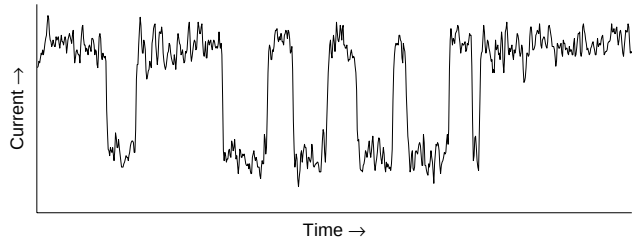


Figure 1.1: LF noise of a 0.18 μm^2 MOSFET. The abrupt jumps are caused by single electrons. Superimposed on this are other types of noise.

It is remarkable that single electron effects can be observed so easily: the experiment takes place at room temperature, and apart from one minimal-area MOSFET (the CPU in your computer contains tens of millions of these), you only need a preamplifier (€1.50) and an oscilloscope.

LF noise in MOSFETs is not only fascinating for the detailed view it provides into the chaotic interior of the MOSFET, but it also gets a lot of attention because it forms a significant limitation to the practical use of the MOSFET in an electronic circuit.

1.2 LF noise and analog circuit design

Design of analog CMOS circuits is a complex tradeoff between resources and goals. The aim is typically to realize a high speed, high accuracy design, with minimal expenditure of supply power and chip area. In this context, an optimal design is one that expends minimal resources to just achieve the specified goals. If specifications are not met the design is unsuccessful and if specifications are exceeded, or if the specifications are met whilst expending more than the minimally required resources, the design is not optimal. In the search for design optimality, it is extremely useful to know where the physical limits are so one can strive to reach them, and at the same time ensure that the physically impossible is not attempted.

In CMOS, the smallest signal that can be reliably processed depends on the noise of the MOSFET. The MOSFET's dominant sources of noise are thermal noise due to the dissipative character of the conducting channel and LF noise.

LF noise in MOSFETs has been studied for many decades. There is ample data indicating the presence of a bulk $1/f$ noise source in homogenous semiconductor samples [26]. It it reasonable to assume this is also present in a MOSFET, however MOSFET LF noise behaviour is dominated by the silicon surface, at the interface between the silicon and the gate oxide. The Si-SiO_2 interface is imperfect for at least two reasons. First, the crystalline structure of Si does not fit the amorphous SiO_2 and this leads to so-called 'dangling bonds'. Secondly, during processing, impurities are introduced at the interface and in the oxide. Both mechanisms give rise to localized energy states known as 'traps'. These traps generate the RTS noise of figure 1.1 that dominates MOSFET LF noise behaviour.

In 1991, it was noted for the first time [6, 7] that MOSFET LF noise is reduced when the device is subjected to large signal excitation (LSE). In other words, turning it 'off' for some time before turning it 'on' reduces its noise when it is 'on'. This means that the LF noise of the device not only depends on the present bias state of the device but also on the bias history of the device. Soon afterwards this effect was associated with the emptying of traps that cause RTS noise [14]. In 1996, the effect was independently discovered by Hoogzaad and Gierkink at the University of Twente [19], leading to a demonstration of the LF noise reduction effect by LSE in a ring oscillator [20] and in a coupled sawtooth oscillator [36].

1.3 Scope of this thesis

In this thesis, we cover the issues that need to be addressed before LF noise reduction by LSE can be applied to analog circuit design. Detailed physical modelling of trap behaviour under LSE is given in [38].

For a long time, Moore's self-fulfilling prophecy [52] has made MOSFET dimensions go down according to a fairly well described exponential curve. This is not only very important to device physicists who use this curve to predict what process hurdles will need to be overcome and when this will need to be done, but it is also very interesting for people designing analog circuitry.

In chapter two, an introduction is given to the mathematics and physics of LF noise in MOSFETs.

Existing LF noise models are combined with what is known about CMOS process downscaling, which allows predictions to be made as to what will be the dominant problems in future process generations. From this, we show that LF noise, in contrast to many other problems, is not something Moore's 'law' will automatically solve, thereby justifying the work in this thesis.

Before LF noise under large signal excitation can be understood, data needs to be gathered. LF noise shows considerable spread: different nominally identical devices may have LF noise powers that vary by several orders of magnitude. For small-area devices, the spread is worse. This makes LF noise characterization challenging: it is not possible to characterize a single device and generalize its behaviour; large numbers of devices are required before general conclusions can be drawn. LF noise behaviour under LSE is also very variable: not only the magnitude of the noise change varies, but the direction as well: sometimes subjecting a device to LSE will increase its LF noise, while other devices show an LF noise decrease under the same conditions.

p-Channel devices are sometimes favoured by designers because they exhibit less LF noise than n-channel devices. Also, it is often noted that p-channel devices have LF noise that can best be explained using the $\Delta\mu$ model, whereas n-channel devices are often seen to behave in accordance with the ΔN model. In the light of these discrepancies, it is by no means certain that LF noise under LSE will behave in the same manner in both types of devices.

In chapter three, LF noise measurements under LSE for both device types

are compared. Another issue that is investigated is the relation between the change in LF noise under LSE and how the LSE is applied to the device: which device terminals allow the LF noise to be influenced? Finally, large numbers of nominally identical devices are measured and the spread of their LF noise and their LF noise under LSE is characterized.

Having characterized the effect, it needs to be removed from the realm of 'just another interesting effect' to something that can be related to established device parameters. Initially, not a lot was known about the physical origins of the LF noise measured, but it rapidly became apparent in measurements that the noise observed was RTS-dominated. Both in the time domain where the characteristic two-level signal of an RTS was observed, and in the frequency domain, where the shape of the LF noise power spectral density was seen to vary between f^0 and f^{-2}.

In chapter four, the experimental observations are related to the vast amount of known data on the subject of interface traps. This allows us to model why the LF noise usually decreases under LSE: LF noise under LSE probes traps nearer the center of the bandgap, and the trap density there is lower than near the bandgap edge. With our model, based on Shockley-Read-Hall theory applied under non-steady-state conditions, LF noise behaviour under LSE follows logically from the characteristics of the device, and no additional fitting parameters are required. An LF noise simulator based on the same model is also presented.

Regarding application of LF noise reduction by LSE, the important question to be answered is of course how it compares to established techniques. Unfortunately, on the whole, the established noise reduction techniques are quite effective at reducing noise, so competing with them is no trivial task. In addition, while LF noise reduction by LSE showed great promise in older processes, the noise reduction in newer CMOS processes is much lower. In correlated double sampling circuits, LSE may be applied but only in such a way that the bias history for both sample moments is identical. When this is not the case LSE is shown to be more likely to increase than to decrease the noise. Only in RF circuits, where other techniques are not applicable does LF noise reduction by LSE show some promise, especially since the frequency of the LSE is not important for the LF noise performance.

In chapter five, the different ways in which LF noise reduction by LSE can be applied to analog circuit design are explored. Three classes of circuits are discussed and the applicability of LF noise reduction by LSE in each class is treated. Though LF noise reduction by LSE can provide a benefit in specific cases, the accuracy of steady state LF noise models still needs to

improve a lot before the limited benefits of the LSE technique can become useful.

Chapter 2

Background

2.1 Introduction

In this chapter, the mathematical and physical background to the remainder of this work will be presented. We start with a brief summary of some important concepts in time and frequency domain analysis. Having done that, some important concepts in stochastics are reviewed. Next, the classical MOSFET square law approximation that forms the basis for a lot of the classical LF noise modelling work is presented. Finally, the future of LF noise in MOSFETs is treated. The framework presented in this chapter forms the basis for the modelling work in chapter 4.

2.2 Time and frequency domain analysis

A brief review of some important concepts in frequency domain analysis is presented. More exhaustive treatment of the subject matter can be found in [12] or other textbooks.

Units of the Fourier Transform.

The Fourier transform is the basis of frequency domain analysis. Using the Fourier transform we can analyse signals in the frequency and in the time domain, however some care is required in the use of units. The Fourier

transform is defined by [12]:

$$x(f) = \int_{-\infty}^{+\infty} x(t) \mathrm{e}^{-j2\pi ft} dt \qquad [\mathrm{Hz}^{-1}] \qquad \text{FT}$$
$$x(t) = \int_{-\infty}^{+\infty} x(f) \mathrm{e}^{j2\pi ft} df \qquad \text{IFT} \qquad (2.1)$$

It gives units for $x(f)$ of Hz^{-1} if the input signal $x(t)$ is dimensionless. It is common practice (eg. done by programs such as Maple) to substitute $\omega = 2\pi f$ in this definition, but this does not change the units of $x(\omega)$ which remain Hz^{-1}. A more logical approach, if rad/s is used as the unit of frequency, is to divide the FT by 2π; this gives units of $(\mathrm{rad/s})^{-1}$ in the frequency domain. This requires that the IFT is multiplied by 2π to ensure consistency. This definition is used by Machlup [48].

The curious but common practice of expressing frequency in rad/s while at the same time using a frequency domain description in which the units are Hz^{-1} is a frequent source of confusion.

Use of the DDFT to approximate the CCFT

A digital computer is used to calculate the Fourier transform of time-domain signals. This means the continuous-time to continuous-frequency Fourier transform (CCFT) cannot be used, but rather we have to use the discrete-time to discrete-frequency Fourier transform (DDFT) to approximate the continuous-time case. The definition of the DDFT of a time signal (as used by Matlab) is:

$$x(k) = \sum_{n=1}^{N} x(n) \mathrm{e}^{\frac{-j2\pi(k-1)(n-1)}{N}} \qquad 1 \leq k \leq N \qquad (2.2)$$

The inverse transform is given by:

$$x(n) = \frac{1}{N} \sum_{k=1}^{N} x(k) \mathrm{e}^{\frac{j2\pi(k-1)(n-1)}{N}} \qquad 1 \leq n \leq N \qquad (2.3)$$

In these definitions, $x(n)$ is the sampled time-domain signal with n the sample number between 1 and N, and $x(k)$ is the sampled frequency domain signal with k going from 1 to N, representing the sample frequency f_s. N is the number of points so that the total time $T = Nt_\mathrm{s}$ where t_s is the sample period. The relation between the CCFT and the DDFT is given by [12]:

$$x(f) = x(k)t_\mathrm{s} \qquad [\mathrm{Hz}^{-1}] \qquad (2.4)$$

The DDFT will cover the frequency range of 0 to $f_s/2$ in $N/2$ steps. The frequency resolution (distance between adjacent samples in the frequency domain) is $1/T$ [Hz]. This is also the lowest frequency measurable.

One-and two sided frequency domain description of signals

A common cause of confusion is the indiscriminate mix of one- and two-sided frequency domain descriptions. The Fourier transform, which in essence defines the frequency domain, gives rise to a frequency domain description of the signal that goes from $f = -\infty \ldots +\infty$, and this is mathematically the most complete way of describing a signal in the frequency domain. If $x(t)$ is real, however, $x(f)$ will have a negative half that is the complex conjugate of the positive half. In these cases, giving the complete $x(f)$ is redundant, and it is sufficient to give only the positive half of $x(f)$. To preserve the total energy of the signal, $x(f)$ is then multiplied by $\sqrt{2}$. The resultant $x(f)$ is known as 'one-sided'. For real signals, a one-sided frequency-domain description is convenient and adequate. However, when transforming back to the time domain, it is important that the appropriate integration limits be used: $-\infty$ and $+\infty$ for a two-sided $x(f)$ and 0 and $+\infty$ for a one-sided $x(f)$.

2.3 Stochastic signals

In describing noise, it is useful to briefly review some important concepts in stochastics. A stochastic signal is random; i.e. the value of the signal at a particular time cannot be predicted. The signal can, however, be described in terms of its stochastic parameters. Some useful stochastic parameters are defined in this section. Random signals are then examined in the frequency domain, and finally some important random signals (RTS noise and $1/f$ noise) are discussed in terms of the definitions given.

Describing stochastic signals.

The expected value of a random signal $x(t)$ is its time average. It is given by

$$\mathrm{E}[x(t)] = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)dt \tag{2.5}$$

The n^{th} moment of a signal is the expected value of $x(t)^n$. It is given by:

$$\mathrm{E}[x(t)^n] = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)^n dt \tag{2.6}$$

Some common moments are:

- The 1^{st} moment of $x(t)$ is its average, often denoted by m_x.
- The 2^{nd} moment of $x(t)$ is the total power of the signal.
- The 2^{nd} moment of $[x(t) - m_x]$ is its variance, $\mathrm{var}(x)$. This is the AC power of the signal.

Higher order moments may be used for mathematical and formal completeness, but they have limited practical significance.

The autocorrelation function of a signal is defined by:

$$R_{xx}(t, t+\tau) = \mathrm{E}\left[x(t)\, x(t+\tau)\right] \tag{2.7}$$

$$\begin{aligned} R_{xx}(t,t) =& \mathrm{E}\left[x(t)\, x(t)\right] \\ =& \mathrm{E}\left[x(t)^2\right] \\ =& 2^{\text{nd}} \text{ moment of } x(t) \\ =& \text{total power of the signal} \end{aligned} \tag{2.8}$$

In a similar way, the autocovariance is given by:

$$C_{xx}(t, t+\tau) = \mathrm{E}\left[(x(t) - m_x)\, (x(t+\tau) - m_x)\right] \tag{2.9}$$

$$\begin{aligned} C_{xx}(t,t) =& \mathrm{E}\left[(x(t) - m_x)\, (x(t) - m_x)\right] \\ =& \mathrm{E}\left[(x(t) - m_x)^2\right] \\ =& \mathrm{var}(x) \end{aligned} \tag{2.10}$$

The autocorrelation function and the autocovariance function differ only in that the autocorrelation function encompasses DC whereas the autocovariance function does not. The autocorrelation function is therefore more complete, and the autocovariance function should only be used in cases where the random signal in question has a non-zero mean and the intent is specifically to disregard this mean.

Ergodicity, stationarity and cyclostationarity

A random signal is said to be stationary if the moments of the random signal are not a function of time. A signal is 'strictly stationary' if this is satisfied

for all its moments. A signal is 'wide-sense stationary' if this is satisfied for all moments up to and including the second moment.

If a signal is stationary, the autocorrelation function given in eq. 2.7 will no longer be a function of the absolute time t but only of the time difference τ:

$$R_{xx}(t, t+\tau) = R_{xx}(\tau) \tag{2.11}$$

A random signal is said to be ergodic if the ensemble average (average over a number of realizations) is equal to the time average of one realization of the signal. A signal is wide-sense ergodic if the first and second moments are ergodic, and ergodic in the strict sense if all its moments are ergodic. If a random signal is ergodic, it must also be stationary. A signal whose moments are periodic in T is said to be cyclostationary in T.

Frequency domain analysis of stochastic signals: the PSD

The PSD and the Wiener-Khinchin theorem

The Power Spectral Density (PSD) of a signal is a plot of the power per unit of bandwidth against frequency. The units of a PSD are W Hz^{-1}. The units $\mathrm{V}^2\mathrm{Hz}^{-1}$ or $\mathrm{A}^2\mathrm{Hz}^{-1}$ also find common usage, in those cases a constant impedance is assumed. The Wiener-Khinchin theorem allows us to go from the time domain to the frequency domain. It states that the PSD of a signal is the Fourier transform of the autocorrelation function.

$$S_x(f) = \mathrm{FT}\left(R_{xx}(t, t+\tau)\right) \tag{2.12}$$

The power of the signal in the frequency domain and in the time domain is equal, as shown by equation 2.13. Care must be taken to use the correct units for $S_x(f)$. If integrating over f, the units for $S_x(f)$ must be in terms of power *per Hertz*.

$$\int_{-\infty}^{+\infty} S_x(f)\,df = P_{\text{signal}} = \mathrm{E}\left[x(t)^2\right] \tag{2.13}$$

The PSD can also be calculated without first calculating the autocorrelation function. This is sometimes referred to as the 'direct' way of calculating the PSD [12]:

$$S_x(f) = \lim_{T\to\infty} \frac{|x(f)|^2}{T} \tag{2.14}$$

The PSD contains less information than the time-domain signal. The first way of calculating the PSD (eq. 2.12) discards information when the autocorrelation function R_{xx} is computed; different time domain signals may

have the same autocorrelation function. The second way of calculating the PSD (eq. 2.14) discards phase information by taking the magnitude of $x(f)$. Though the PSD of a random signal is very useful, a time-domain description is formally more complete.

The PSD of a cyclostationary signal

A signal which is cyclostationary in T will have moments that are periodic in T. The autocorrelation function of such a cyclostationary signal will also be periodic in T; i.e.

$$R_{xx}(t,t+\tau) = R_{xx}(t+T,t+T+\tau) \tag{2.15}$$

Additionally, the autocorrelation function at any particular time t will have periodic components. Such an autocorrelation function can in theory be used to calculate the time-variant PSD of the cyclostationary signal by rigourous application of the Wiener-Khinchin theorem, but this not always necessary. A more productive approach is often to first 'stationarize' the signal by averaging the moments of the signal over one period, and then Fourier transforming to find the 'stationarized' PSD [16]. Stationarizing the signal may equivalently be done by modelling the time reference as a random variable that is uniformly distributed over one cycle. The observer, in both cases, has no knowledge anymore of the phase of the signal. This approach is therefore valid whenever the 'measuring' system is not synchronized with the cyclostationary noise source, for example when a spectrum analyser is used to measure cyclostationary noise.

Random Telegraph Signals

A Random Telegraph Signal (RTS) is shown in fig. 2.1. It is a time-continuous, amplitude-discrete signal. The conditional probability of a transition from one state to another (given that it is in that particular state) is proportional to dt; this makes the time spent in each state ('1' and '0' in the figure) exponentially distributed.

This can be understood as follows: if the *conditional* probability of a transition from a state, given that the signal is in that state, per unit time, is constant, then the *absolute* probability of a transition from that state is pro-

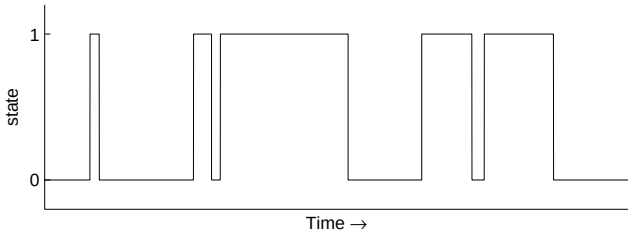


Figure 2.1: A Random Telegraph Signal (RTS)

portional to the probability of being in the state. Mathematically:

$$\begin{aligned} \mathrm{P}(\text{transition}) &= \mathrm{P}(\text{transition}|\text{state})\ \mathrm{P}(\text{state}) \\ \mathrm{P}(\text{transition}|\text{state}) &\propto dt \\ \mathrm{P}(\text{transition}) &\propto \mathrm{P}(\text{state})dt \\ d\mathrm{P}(\text{state}) &\propto \mathrm{P}(\text{state})dt \\ \frac{d\mathrm{P}(\text{state})}{dt} &\propto \mathrm{P}(\text{state}) \end{aligned} \tag{2.16}$$

Equation 2.16 is a first-order differential equation that can be solved to give an exponential $\mathrm{P}(\text{state})(t)$.

The RTS is completely characterized by three parameters: The amplitude, the mean 'high' time and the mean 'low' time. The autocorrelation function and PSD of an RTS will be derived. The derivation follows Machlup [48]. For mathematical convenience, the amplitude is chosen as 1, and the two states of the RTS are named '0' and '1'. The autocorrelation function of the RTS is given by:

$$\begin{aligned} R_{RTS} &= \mathrm{E}\left[RTS(t)\ RTS(t+\tau)\right] \\ &= \sum_{ij} \left[x_i\, x_j\ \mathrm{P}(x(t)=x_i)\ \mathrm{P}(x(t+\tau)=x_j|_{x(t)=x_i})\right] \end{aligned} \tag{2.17}$$

Since one of the states is called '0', three of the four terms in this sum vanish, and we have:

$$\begin{aligned} R_{RTS} &= 1\times 1\times \mathrm{P}(x(t)=1)\times \mathrm{P}(n_{\text{trans}} \text{ is even, given a start in state } 1) \\ &= \mathrm{P}(x(t)=1)\ P_{11}(\tau) \end{aligned} \tag{2.18}$$

If the mean 'high' time is denoted by τ_1 and the mean 'low' time by τ_0 then the probability of being in state '1' at any particular time is given by

$$\mathrm{P}(x(t)=1)=\frac{\tau_1}{\tau_0+\tau_1} \tag{2.19}$$

and the autocorrelation function is found as

$$R_{RTS}=\frac{\tau_1}{\tau_0+\tau_1}P_{11}(\tau) \tag{2.20}$$

P_{11} can be found by first realizing

$$P_{11}(\tau)+P_{10}(\tau)=1 \tag{2.21}$$

We can then write:

$$P_{11}(\tau+d\tau)=P_{10}(\tau)\frac{d\tau}{\tau_0}+P_{11}(\tau)\left(1-\frac{d\tau}{\tau_1}\right) \tag{2.22}$$

In words: An even number of transitions in time $\tau+d\tau$, given that a start is made in state '1', is possible in two ways: *either* an odd number of transitions in time τ followed by a transition from '0' to '1' in time $d\tau$, *or* an even number of transitions in time τ followed by no transition from state '1' to '0' in time $d\tau$. If the time $d\tau$ is made small enough, the probability of more than one transition in time $d\tau$ is negligible. Substituting eq. 2.21, the limit can then be taken for $d\tau\to 0$ to arrive at the differential equation for P_{11}:

$$\frac{dP_{11}}{d\tau}+\left(\frac{1}{\tau_1}+\frac{1}{\tau_0}\right)P_{11}=\frac{1}{\tau_0} \tag{2.23}$$

This differential equation can be solved using the initial condition $P_{11}(0)=1$ to arrive at:

$$P_{11}(\tau)=\frac{1}{\tau_1+\tau_0}\left[\tau_0 \mathrm{e}^{-\left(\frac{1}{\tau_0}+\frac{1}{\tau_1}\right)\tau}+\tau_1\right] \tag{2.24}$$

This leads to:

$$R_{RTS}(\tau)=\frac{\tau_1}{(\tau_1+\tau_0)^2}\left[\tau_0 \mathrm{e}^{-\left(\frac{1}{\tau_0}+\frac{1}{\tau_1}\right)\tau}+\tau_1\right] \tag{2.25}$$

Fourier transforming the autocorrelation function, the spectrum of the RTS can be found:

$$\begin{aligned} S_{RTS}(\omega)&=\int_{-\infty}^{+\infty}R_{RTS}\mathrm{e}^{-j\omega\tau}d\tau \quad [\mathrm{Hz}^{-1}] \\ &=2\frac{\tau_0\tau_1}{(\tau_0+\tau_1)^2}\frac{\left(\frac{1}{\tau_0}+\frac{1}{\tau_1}\right)}{\omega^2+\left(\frac{1}{\tau_0}+\frac{1}{\tau_1}\right)^2}+2\pi\left(\frac{\tau_1}{\tau_0+\tau_1}\right)^2\delta(\omega) \end{aligned} \tag{2.26}$$

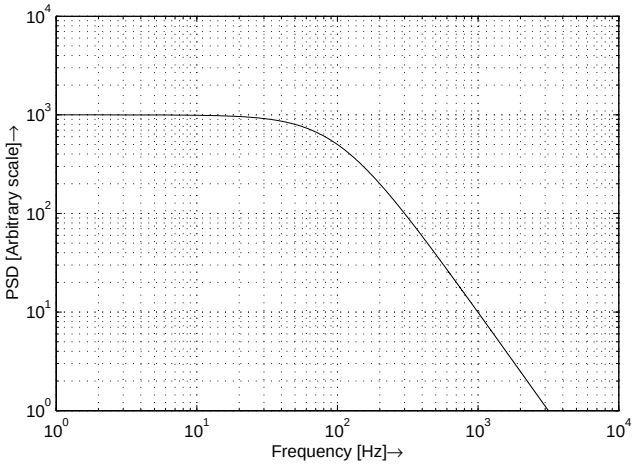


Figure 2.2: PSD of an RTS; $f_{0\mathrm{RTS}} = 100$ Hz

This is a Lorentzian spectrum. Ignoring the DC term and making two substitutions, the result is:

$$\begin{aligned} \beta &= \frac{\tau_0}{\tau_1} \\ \omega_{0\mathrm{RTS}} &= \frac{1}{\tau_0} + \frac{1}{\tau_1} && [\mathrm{rad/s}] \\ S_{RTS}(\omega) &= 2\frac{\beta}{(1+\beta)^2}\frac{1}{\omega_{0\mathrm{RTS}}}\frac{1}{1+\frac{\omega^2}{\omega_{0\mathrm{RTS}}^2}} && [\mathrm{Hz}^{-1}] \end{aligned} \tag{2.27}$$

A PSD of an RTS is given in fig. 2.2. The LF power of the RTS is inversely proportional to the RTS corner frequency, $\omega_{0\mathrm{RTS}}$. The final term defines the frequency dependence of the RTS: flat at low frequencies and decaying with $1/\omega^2$ above the RTS corner frequency. Note that the spectrum of the RTS is symmetrical with respect to τ_1 and τ_0; only the DC term is sensitive to the difference. The power of the PSD depends on the 'asymmetry factor' β. This is shown in fig. 2.3. If $\beta = 1$; i.e. $\tau_0 = \tau_1$, the PSD of the RTS will be at a maximum.

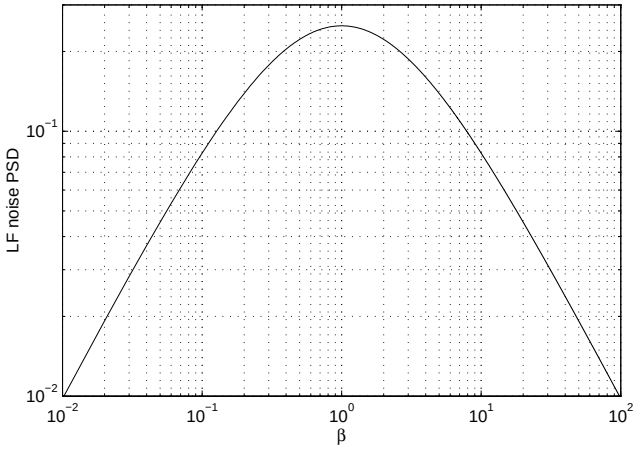


Figure 2.3: RTS noise power as function of the asymmetry of the RTS. $\omega_{0\mathrm{RTS}} = 1$, amplitude $= 1$. Maximum power for $\beta = 1$ (symmetrical RTS).

$1/f$ noise as a stochastic signal

Pure $1/f$ noise is described by a PSD as follows:

$$S_{1/\mathrm{f}} = \frac{K}{f} \tag{2.28}$$

In this equation, K is an arbitrary constant. This is a strictly mathematical construction, as pure $1/f$ noise cannot be observed in practice, since an infinite observation time is required to ascertain that the spectrum indeed has a $1/f$ character. Even though 'pure' $1/f$ noise looks very simple in equation form, and it is therefore tempting to use it to model the physical phenomena that *can* be observed, it is problematic in that the integral of the PSD is not convergent and therefore 'pure' $1/f$ noise implies an infinite signal power (eq. 2.13). This is not physically plausible. If a mathematical model of a physical process that has a $1/f$ character in the observable range of frequencies is desired, a more complex model is required. A rather pragmatic solution to the problem is given by Hooge and Bobbert [25]. Since the integral of pure $1/f$ noise does not converge, some lower limit frequency below which the noise PSD gets an f^0 shape must be assumed, as must some upper limit frequency above which the noise PSD gets an f^{-2} shape. Neither of these areas are observable, the lower limit due to the limits on observation time (see, for example [50]), and the upper limit due to thermal

noise, so these rather practical assumptions suffice. The PSD may now be integrated in three parts to find the power of the signal:

$$\begin{aligned} P_{\text{signal}} &= 2[\int_0^{f_\text{l}} \frac{K}{f_\text{l}} df \quad + \int_{f_\text{l}}^{f_\text{h}} \frac{K}{f} df \quad + \int_{f_\text{h}}^{+\infty} \frac{K f_\text{h}}{f^2} df \quad] \\ &= 2[\, K \quad + K \ln(\frac{f_\text{h}}{f_\text{l}}) \quad + K \quad] \end{aligned} \tag{2.29}$$

Choosing f_l suitably low and f_h suitably high, we have a description of observable $1/f$ noise. A suitable choice for f_l and f_h is still required but within rather wide limits, this is not critical. Moreover, if they are far enough apart, the power contribution of the LF and HF part of the integral is not significant, and this justifies the convenient choice of the shape of the spectrum in those parts.

The problem that the integral of the PSD of pure $1/f$ noise does not converge is tackled in another way [32, 49] by modelling $1/f$ noise as a non-stationary random process. This makes the variance and the PSD time dependent. Observing such a process for a limited time always gives a finite variance.

2.4 MOSFETs

Square law approximation

The current-voltage behaviour of a MOSFET can be described by the square law model, also known as the Sah-model [55]. Though this view of the MOSFET from 1964 does not take into account short-channel effects, it is nevertheless useful because it gives insight into the bias and geometry dependencies of many of the MOSFET noise models.

In strong inversion,

$$I_\text{D} = \frac{W}{L} \mu C_\text{ox}^\square [V_\text{GT} V_\text{DS} - \frac{1}{2} V_\text{DS}^2] \tag{2.30}$$

V_GT is the gate-threshold overdrive voltage, V_GT =V_GS-V_T. For low V_DS, expression 2.30 reduces to

$$I_\text{D} = \frac{W}{L} \mu C_\text{ox}^\square [V_\text{GT} V_\text{DS}] \tag{2.31}$$

In saturation, the voltage across the channel is V_{GT} and the expression becomes:

$$I_{\mathrm{D}} = \frac{W}{L}\mu C_{\mathrm{ox}}^{\square}[\frac{1}{2}V_{\mathrm{GT}}^2] \tag{2.32}$$

Transconductance is given by:

$$\begin{aligned} g_{\mathrm{m}} &= \frac{W}{L}\mu C_{\mathrm{ox}}^{\square} V_{\mathrm{DS}} \qquad \text{(linear region)} \\ g_{\mathrm{m}} &= \frac{W}{L}\mu C_{\mathrm{ox}}^{\square} V_{\mathrm{GT}} \qquad \text{(saturated region)} \end{aligned} \tag{2.33}$$

An estimate for the number of free carriers in the device can be made by taking the total inversion charge and dividing by the elementary charge q.

$$\begin{aligned} N_{\mathrm{lin}} &= \frac{C_{\mathrm{ox}}^{\square} V_{\mathrm{GT}} WL}{q} \\ N_{\mathrm{sat}} &= \frac{2}{3} N_{\mathrm{lin}} \end{aligned} \tag{2.34}$$

The factor $2/3$ in saturation comes from the integration of the channel charge from source to pinch-off point [45, 63].

Thermal noise in a MOSFET

Before going into $1/f$ noise it is relevant to briefly examine the equations governing the thermal noise of a MOSFET. In a resistor, thermal agitation of the electrons gives rise to thermal noise. It can be shown [12] that the (double sided) PSD of thermal noise is given by:

$$S_{\mathrm{th}} = 2R\frac{h|f|}{\mathrm{e}^{h|f|/kT} - 1} \qquad [\mathrm{V}^2\mathrm{Hz}^{-1}] \tag{2.35}$$

For $|f| << kT/h$, this simplifies to the well known

$$\begin{aligned} S_{\mathrm{th}} &= 2RkT \qquad [\mathrm{V}^2\mathrm{Hz}^{-1}] \qquad \text{(double sided)} \\ &= 4RkT \qquad [\mathrm{V}^2\mathrm{Hz}^{-1}] \qquad \text{(single sided)} \end{aligned} \tag{2.36}$$

It is common practice to use single sided spectra and it is the convention followed in this thesis. The channel of a MOSFET behaves as a resistor, and it exhibits corresponding thermal noise.

$$S_{I_{\mathrm{D}}} = 4kTg_{\mathrm{channel}}\xi \qquad [\mathrm{A}^2\mathrm{Hz}^{-1}] \tag{2.37}$$

g_{channel} is the conductance of the channel, given by I_D/V_{DS} in the linear regime, and to a first order approximation by I_D/V_{GT} in saturation. The factor ξ is 1 for low V_{DS}, and more than 1 when the channel shows non-linearity, as it does for higher V_{DS} and in saturation. [1] On theoretical grounds, ξ can be shown to be $4/3$ in saturation.

In the linear regime with $\xi = 1$, this leads to:

$$\begin{aligned} S_{I_D} &= \frac{4kTI_D}{V_{DS}} \\ S_{V_G} &= \frac{4kTI_D}{g_m^2 V_{DS}} \\ &= \frac{4kT}{\mu^2 C_{ox}^{\square 2}} \frac{L^2}{W^2} \frac{I_D}{V_{DS}^3} \\ &= \frac{4kT}{\mu C_{ox}^{\square}} \frac{L}{W} \frac{V_{GT}}{V_{DS}^2} \end{aligned} \tag{2.38}$$

In saturation:

$$S_{I_D} = \frac{4kTI_D}{V_{GT}} \xi \tag{2.39}$$

Filling in the expression for I_D and g_m leads to the oft-seen[1]

$$\begin{aligned} S_{I_D} &= 4kTg_m \frac{\xi}{2} \\ S_{V_G} &= \frac{4kT\frac{\xi}{2}}{g_m} \end{aligned} \tag{2.40}$$

which is a reasonable approximation but which gives little insight into the physics of MOSFET noise because it insinuates that: (a) g_m is somehow very important to the thermal noise of the MOSFET, and (b) the MOSFET exhibits 'less than thermal' noise, whereas in reality, (a) thermal noise originates in the channel and (b) the non-linear V/I characteristic of the conducting channel makes the noise worse, not better. The familiar form of eq. 2.40 is just an unfortunate coincidence (though useful for circuit design as g_m is often constrained). g_m Can be filled in:

$$S_{V_G} = 4kT \frac{\xi}{2} \frac{L}{W\mu C_{ox}^{\square} V_{GT}} \tag{2.41}$$

From which it can be seen that gate-referred thermal noise in a MOSFET is proportional to the absolute temperature T, L/W, and inversely proportional

[1] In literature, $\xi/2$ is often denoted as γ

to $C_{ox}^{\square}$. In the linear region, it is proportional to V_{GT}/V_{DS}^2, and in saturation, it is inversely proportional to V_{GT}.

2.5 LF noise modelling

MOSFETs not only exhibit thermal noise but also significant LF noise. There are different physical models for MOSFET LF noise, which will be described in this section. In MOSFETs with a small number of free carriers, single-electron trapping-detrapping events generate RTS noise.

In order to better understand the different LF noise models, we will first derive the dependence of the relative conductivity fluctuation on the free carrier concentration. This enables us to identify the different physical mechanisms that cause conductivity fluctuations. Current in the channel of a MOSFET is carried by mobile charge carriers.

$$\sigma = nq\mu \qquad [\mathrm{CV^{-1}s^{-1}m^{-1}}] \tag{2.42}$$

If the conductivity σ fluctuates, this is due to either μ fluctuating, n fluctuating, or it could be predominantly due to fluctuations in μ that result from fluctuations in n.

Externally, only the conductivity can be observed, but we can ascertain which parameter is fluctuating by examining the dependence of the relative conductivity fluctuation on the carrier concentration n: If a fluctuation in the conductivity is attributed to a fluctuation in mobility, we may write:

$$\frac{S_{\Delta\sigma}}{\sigma^2} = \frac{S_{\Delta\mu}}{\mu^2} \tag{2.43}$$

which is not dependent on n. If, on the other hand, a fluctuation in the conductivity is attributed to a fluctuation in the concentration of mobile carriers n,

$$\frac{S_{\Delta\sigma}}{\sigma^2} = \frac{S_{\Delta n}}{n^2} \tag{2.44}$$

which can be rewritten as:

$$\frac{S_{\Delta\sigma}}{\sigma^2} = \frac{S_{\Delta n}}{n}\,\frac{1}{n} \tag{2.45}$$

$S_{\Delta n}/n$ is independent of n if individual Δn events are considered to be uncorrelated, so $S_{\Delta\sigma}/\sigma^2$ is proportional to $1/n$. The third possibility is that

mobility fluctuations *caused by* number fluctuations are the dominant source of conductivity fluctuations. In this case,

$$\begin{aligned}\Delta\sigma &= qn\frac{d\mu}{dn}\Delta n \\ S_{\Delta\sigma} &= q^2n^2(\frac{d\mu}{dn})^2 S_{\Delta n} \\ \frac{S_{\Delta\sigma}}{\sigma^2} &= \frac{(\frac{d\mu}{dn})^2 S_{\Delta n}}{\mu^2} \\ &= \frac{(\frac{d\mu}{dn})^2 \frac{S_{\Delta n}}{n}}{\mu^2}\, n \end{aligned} \tag{2.46}$$

which means the relative conductivity fluctuations are now proportional to n. Using equation 2.34 the different dependencies of the noise on n can be related to V_{GT} and $C_{\mathrm{ox}}^{\square}$.

In the $\Delta\mu$ model, LF noise in MOSFETs is related to fluctuations in μ. The ΔN model assumes fluctuations in n are the dominant source of LF noise. The Hung model [29] postulates the presence of all three LF noise mechanisms. The implications of the models will be reviewed below.

$1/f$ Noise in a MOSFET, $\Delta\mu$ model

Measurements on homogenous semiconductor samples [26] indicate that in those samples, the observed $1/f$ noise obeys the relation

$$\frac{S_I}{I^2} = \frac{\alpha_{\mathrm{H}}}{fN} \tag{2.47}$$

This is known as the Hooge equation. It states that the spectral density of current noise, normalized to the square of the DC current, is inversely proportional to N, the total number of free carriers in the system, and f, the frequency. Regarding the inverse dependence on N, this indicates that in those samples, the observed $1/f$ noise was a bulk rather than a surface phenomenon.

Applying Hooge's equation to a MOSFET and assuming that individual noise sources in the channel of the device are uncorrelated, it can be shown that [34, 56]:

$$S_{V_{\mathrm{G}}} = \frac{\alpha_{\mathrm{H}}}{f}\frac{1}{WL}\frac{q}{C_{\mathrm{ox}}^{\square}}(V_{\mathrm{GT}} - \frac{1}{2}V_{\mathrm{DS}}) \tag{2.48}$$

In the linear region with low V_{DS} this reduces to:

$$S_{V_G} = \frac{\alpha_H}{f}\frac{1}{WL}\frac{qV_{GT}}{C_{ox}^{\square}} \tag{2.49}$$

and in saturation (substituting V_{GT} in the place of V_{DS} as the voltage over the channel) it becomes:

$$S_{V_G} = \frac{\alpha_H}{f}\frac{1}{WL}\frac{qV_{GT}}{2C_{ox}^{\square}} \tag{2.50}$$

The $\Delta\mu$ model predicts that S_{V_G} should scale with the inverse of the device area, should be linearly dependent on V_{GT} and inversely dependent on $C_{ox}^{\square}$. p-Channel devices are often reported to exhibit $1/f$ noise behaviour in accordance with the $\Delta\mu$ model [10, 56].

Substituting

$$V_{GT} = \sqrt{\frac{2I_D L}{W\mu C_{ox}^{\square}}} \tag{2.51}$$

it may be concluded that in saturation, with I_D kept constant, the gate-referred noise is proportional to $W^{-3/2}L^{-1/2}$.

Given a measured S_{V_G} at a particular frequency, it is always possible to calculate α_H if the parameters of the device are known. This does not have to imply a $\Delta\mu$ origin of the LF noise (although α_H is often primarily associated with the $\Delta\mu$ model), but can be useful simply as an empirical measure of the LF noise normalized to the drain current, the frequency and the number of free carriers in the device.

$1/f$ Noise in a MOSFET, ΔN model

Number fluctuations in a MOSFET can also cause low frequency noise. Number fluctuations are caused by trapping and detrapping of mobile carriers in traps at the interface or in the oxide. The area, bias and $C_{ox}^{\square}$ dependency predicted by this so-called ΔN model will be reviewed. In the linear region,

$$\begin{aligned} S_{V_G} &= \frac{S_{Q_{SS}}}{C_{ox}^{\square 2}W^2L^2} \\ &= \frac{q^2 S_{\Delta N}}{C_{ox}^{\square 2}W^2L^2} \end{aligned} \tag{2.52}$$

$S_{Q_{SS}}$ Is the spectrum of surface state charge fluctuations. Assuming that the noise sources in the device are uncorrelated, $S_{\Delta N} = S^{\square}_{\Delta N} WL$ may be substituted to derive:

$$S_{V_G} = \frac{q^2}{C^{\square 2}_{ox} WL} S^{\square}_{\Delta N} \tag{2.53}$$

The spectrum of number fluctuations per unit area ($S^{\square}_{\Delta N}$) is independent of the size of the device if the device is uniform. This means, that like the $\Delta\mu$ model, the ΔN model predicts that the gate referred PSD of noise should scale with the inverse of the device area. To a reasonable approximation, $S^{\square}_{\Delta N}$ is not bias dependent. This is because the trapping process is rate-limited by the number of available traps, not the number of available free carriers. Hence, the ΔN model predicts that the gate-referred noise PSD should be independent of V_{GT} and proportional to $C^{\square -2}_{ox}$. n-Channel MOSFETs are often seen to exhibit $1/f$ noise behaviour in accordance with the ΔN model.

Equation 2.53 is equally valid in the linear region and in saturation [18]. Since S_{V_G} is not dependent on V_{GT}, eq. 2.53 gives the correct scaling regardless of whether V_{GT} or I_D is kept constant.

McWhorter's ΔN model

So far, the area and bias dependence of ΔN fluctuations has been addressed without discussing what the shape of the spectrum described by $S^{\square}_{\Delta N}$ is. McWhorter was the first to show that a trapping/detrapping process can lead to a $1/f$ type spectrum [51]. His derivation will be briefly summarized here.

A single trap gives rise to a Lorentzian PSD, as shown in section 2.3 above. If a MOSFET contains a 'large' number of such traps, and these traps do not interact, their PSD's may be added. If the traps all generate an RTS with the same amplitude, and their time constants are exponentially distributed, the summation of their PSD's will have a $1/f$ shape. This is shown in fig. 2.4. An exponential distribution of time constants may result from a uniform distribution of the distance from the channel to the traps in the oxide, if tunnelling is the mechanism by which charge carriers from the channel interact with the traps [23].

There are some issues regarding the McWhorter model that deserve mention:

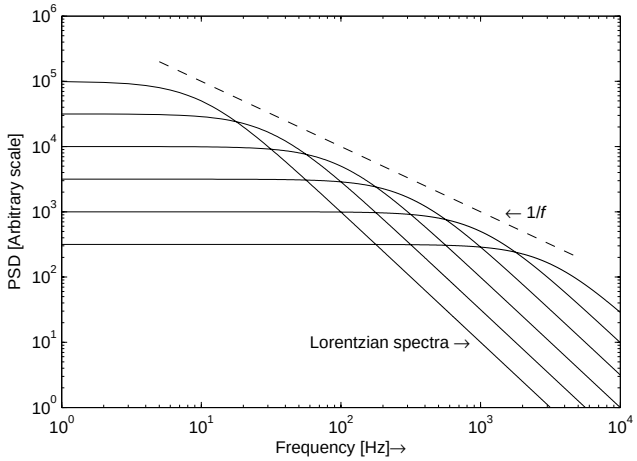


Figure 2.4: Addition of Lorentzian spectra results in a $1/f$ spectrum

- The McWhorter model describes *sufficient* conditions for the emergence of a $1/f$ spectrum, not *necessary* conditions. For example, in [53], far less stringent conditions for the emergence of a $1/f$ spectrum from RTS noise are derived.
- If the RTS time constants are solely due to the distance of the trap in the oxide, one would expect a low frequency cutoff for the $1/f$ noise in devices with very thin oxides. Such a cutoff has not been observed.
- The McWhorter model was derived in a time when devices were 'large', and the postulate of a 'large' number of traps seemed reasonable. In very small area devices, (eg Brederlow [9] measures 30..300 traps per μm^2 in a 0.25 μm process) averaging does not occur and the McWhorter model is not directly applicable.
- Measurements of RTS noise on small devices indicate that traps may in fact interact with one another. [33].

Hung $1/f$ noise model

The Hung $1/f$ noise model [28, 29] is often used in circuit simulators. It is based on a model of number fluctuations and correlated mobility fluctuations. With suitable parameters, it is able to achieve excellent correlation with measured results [71]. The basic form of the Hung $1/f$ model for

small V_{DS} is given in [29]:

$$S_{V_{\mathrm{G}}} \propto \frac{kTq^2}{\gamma f WLC_{\mathrm{ox}}^{\square 2}}(1+\alpha\mu N^{\square})^2 \tag{2.54}$$

This equation expresses that the capture of a mobile charge carrier causes a correlated fluctuation in mobility due to Coulomb scattering at the interface. Depending on the strength of the coupling coefficient α, it predicts the existence of three regions[2].

For $\alpha\mu N^{\square} << 1$, the model reduces to the simple ΔN model, predicting $S_{V_{\mathrm{G}}}$ to be independent on V_{GT}, and $C_{\mathrm{ox}}^{\square -2}$.

For $\alpha\mu N^{\square} \approx 1$, the model states

$$S_{V_{\mathrm{G}}} \propto \frac{kTq^2N^{\square}}{WLC_{ox}^{\square 2}} = \frac{kTqV_{\mathrm{GT}}}{WLC_{\mathrm{ox}}^{\square}} \tag{2.55}$$

Which is equivalent to the $\Delta\mu$ model, predicting a dependency of $S_{V_{\mathrm{G}}}$ on $V_{\mathrm{GT}}\, C_{ox}^{\square -1}$.

The third area of the Hung model is for $\alpha\mu N^{\square} >> 1$, this is when mobility fluctuations *resulting from* number fluctuations dominate. In this area, $S_{V_{\mathrm{G}}}$ is proportional to V_{GT}^2.

RTS noise in MOSFETs

In MOSFETs with a small number of free carriers, RTS noise rather than $1/f$ noise is observed. A useful review of the history of RTS observations in MOSFETs is given in [33]. Assuming that:

- $1/f$ Noise is caused by mobility fluctuations with a given α_{H},
- The RTS noise is pure ΔN, i.e. there is no mobility fluctuation as a result of the Coulomb scattering from trapping,
- The RTS noise is visible if the amplitude of the RTS is larger than the RMS value of the $1/f$ noise in the observation bandwidth,

a simple condition can be derived [35] for the visibility of the RTS noise:

$$N \leq \frac{1}{\alpha_{\mathrm{H}} \ln(f_{\mathrm{m}}(\tau_{\mathrm{c}}+\tau_{\mathrm{e}}))} \tag{2.56}$$

α_{H} is the Hooge empirical $1/f$ noise factor, f_{m} is the bandwidth of the measurement setup, and τ_{e} and τ_{c} are the parameters of the RTS.

[2] Note that this α is not the same as α_{H} used in the context of the $\Delta\mu$ LF noise model.

Noise source	Scaling		
	W and L	$C_{ox}^{\square}$ ($\propto 1/t_{ox}$)	V_{GT}
Thermal	$W^{-1}L$	$C_{ox}^{\square -1}$	V_{GT} (lin) V_{GT}^{-1}(sat)
$\Delta\mu$	$W^{-1}L^{-1}$	$C_{ox}^{\square -1}$	V_{GT}
ΔN	$W^{-1}L^{-1}$	$C_{ox}^{\square -2}$	V_{GT}^{0}
Correlated ΔN and $\Delta\mu$	$W^{-1}L^{-1}$	$C_{ox}^{\square 0}$	V_{GT}^{2}

Table 2.1: Overview of LF noise models

The visibility of the RTS improves with smaller α_H, smaller measurement bandwidth and higher RTS corner frequency.

Overview

The important scaling properties of the different models are given in table 2.1. W and L scaling is given, as is the $C_{ox}^{\square}$ and V_{GT} dependency. Scaling rules for the $1/f$ corner frequency f_c can be derived if so desired. Note that whereas thermal noise is always present in the channel of a MOSFET, and $S_{I_D,th}$ is non-zero for any V_{DS}, this is not the case for LF noise. As predicted by the different noise models, $S_{I_D} = 0$ when $I_D = 0$. Regarding the $\Delta\mu$ model, this is immediately obvious from equation 2.47. Regarding the ΔN model, S_{V_G} does not go to 0 for V_{DS} =0 (equation 2.53) but g_m does, thereby reducing S_{I_D} to 0 for V_{DS} =0.

2.6 The future of LF noise in MOSFETs

CMOS downscaling is expected to continue for a number of process generations [30]. A change in gate dielectric from SiO_2 to a so-called 'High-κ' dielectric is expected by 2006 or 2007 [30]. This is expected to aggravate trapping and associated LF noise problems, as high-κ gate dielectrics are reported to have interface state densities one or two orders of magnitude higher than SiO_2 [69].

To investigate the effect of downsizing MOSFETs on their LF noise, a next process generation is modelled by a shrink factor s. We discuss how this

Current process generation	Next process generation
t_{ox}	t_{ox}/s
$C_{ox}^{\square}$	$s\,C_{ox}^{\square}$
V_{DD}	V_{DD}/s
V_{GT}	V_{GT}

Table 2.2: Simple model for CMOS scaling

affects $1/f$ noise. The upcoming change to high-κ gate dielectrics should be considered in isolation.

The main difference between one process generation and the next (table 2.2) is the decreasing t_{ox}, leading to an increase in $C_{ox}^{\square}$. To prevent oxide breakdown, this mandates an equivalent decrease of V_{DD}. V_T scales down from one process generation to the next, but analog CMOS design over the past few process generations has shown that V_{GT} in typical analog circuits does not change much from one generation to the next [2], so V_{GT} is modelled as constant. To evaluate how the $1/f$ noise changes from one process generation to the next, boundary conditions have to be selected. Several can be chosen, and some are mutually exclusive. Some possibilities are:

- **Keep σ_{V_T}, the spread in V_T, constant.** This is realistic in the light of the observation that V_{GT} does not change much from one process generation to the next. Using the relation for the spread in V_T:

$$\sigma_{V_T} = \frac{A_{V_T}}{\sqrt{WL}} \tag{2.57}$$

  where A_{V_T} is a constant proportional to t_{ox}, this means WL will have to decrease by a factor s^2 from one process generation to the next.
- **Keep A_{tot}, the total area, constant.** This means WL will remain the same from one process generation to the next. Keeping the total area constant is equivalent to keeping the relative spread in V_T; σ_{V_T}/V_{DD} constant.
- **Keep SNR_{th}, the signal-to-thermal noise-ratio, constant.** At constant V_T, keeping SNR_{th} constant means keeping $S_{V_{th}}$ constant. From equation 2.40, this means L/W should increase by s from one process generation to the next.
- **Keep P_{DD}, the supply power, constant.** Since V_{DD} is dropping by a factor s, this means increasing I_D by a factor s. Making use of

		P_{DD} constant	SNR_{th} constant
A_{tot} constant	W	1	$s^{-1/2}$
	L	1	$s^{1/2}$
	WL	1	1
	S_{V_G} ($C_{ox}^0/C_{ox}^{-1}/C_{ox}^{-2}$)	$1/s^{-1}/s^{-2}$	$1/s^{-1}/s^{-2}$
	f_c ($C_{ox}^0/C_{ox}^{-1}/C_{ox}^{-2}$)	$s/1/s^{-1}$	$1/s^{-1}/s^{-2}$
σ_{V_T} constant	W	s^{-1}	$s^{-3/2}$
	L	s^{-1}	$s^{-1/2}$
	WL	s^{-2}	s^{-2}
	S_{V_G} ($C_{ox}^0/C_{ox}^{-1}/C_{ox}^{-2}$)	$s^2/s/1$	$s^2/s/1$
	f_c ($C_{ox}^0/C_{ox}^{-1}/C_{ox}^{-2}$)	$s^3/s^2/s$	$s^2/s/1$

Table 2.3: The future of $1/f$ noise. Downscaling by factor s under given constraints. $C_{ox}^0/C_{ox}^{-1}/C_{ox}^{-2}$ refers to 'correlated ΔN and $\Delta\mu$', $\Delta\mu$ and ΔN $1/f$ noise models respectively

equation 2.32, this means keeping W/L constant from one process generation to the next.

Either σ_{V_T} or A_{tot} can be kept constant, in combination with either P_{DD} or SNR_{th}. This leads to the scaling rules for W and L as given in table 2.3. It is up to the designer to decide which of the four possibilities is most applicable to the situation in question. S_{V_G} Is probably the fairest indicator of whether the $1/f$ noise is becoming more or less in absolute terms while f_c, the $1/f$ noise corner frequency, indicates the relative importance of $1/f$ noise compared to thermal noise.

The most realistic approach to scaling is to keep (a) SNR_{th} and (b) σ_{V_T} constant (the lower right hand corner in table 2.3). In this way, we are (a) comparing the LF noise to the thermal noise, and (b) keeping the matching properties of the circuit the same. Under these conditions, f_c will increase for the correlated ΔN and $\Delta\mu$ model and the $\Delta\mu$ model, and remain unchanged for the ΔN model. Other boundary conditions may be chosen to make this evaluation (eg. in [71] a simulation-based rather than an analytical approach is used), but similar conclusions result.

2.7 Conclusion

In this chapter, we have reviewed frequency and time domain analysis of stochastic signals, and common LF noise models for MOSFETs and their implications are discussed.

The dependence of the relative conductivity fluctuation ($S_{\Delta\sigma}/\sigma^2$) on the free carrier concentration n is derived. This is the basis on which the different LF noise models for MOSFETs, namely the $\Delta\mu$, ΔN and the correlated $\Delta\mu$ and ΔN models are distinguished.

It is expected that LF noise in small-area devices will continue to be RTS-dominated. For future CMOS process generations, the relative importance of LF noise compared to thermal noise is not expected to decrease. (table 2.3). The expected change in future CMOS processes from SiO_2 to a high-κ gate dielectric is expected to increase LF noise problems as trap densities in high-κ gate dielectrics are much higher than in SiO_2.

Chapter 3

Measurement of LF noise under large signal excitation

3.1 Introduction

In our study of LF noise under Large Signal Excitation (LSE) numerous measurements of MOSFET LF noise behaviour under LSE have been performed. In this chapter, the most important experiments and their results will be presented. First, LF noise measures will be briefly discussed so that these can be used to describe the results of the experiments. We review the published data on the subject. Next, the most important measurement results are presented. Factors that influence LF noise are presented; most importantly: what is the waveform and amplitude of the large signal that the device is subjected to and to what terminal of the device is it applied. Device factors that influence noise and noise under LSE will be described; such as process factors (in particular: t_{ox}, device size, and device polarity (n or p-channel). There is a significant difference between measurements on single selected 'golden samples', which exhibit spectacular behaviour when subjected to LSE, and 'real-world' performance of randomly selected sets of devices. Though both types of measurement are valuable, it is important to be aware of this distinction. To investigate spread of LF noise and LF noise under LSE, measurements of LF noise under LSE on large sets of unselected devices are presented. Finally, we will summarize.

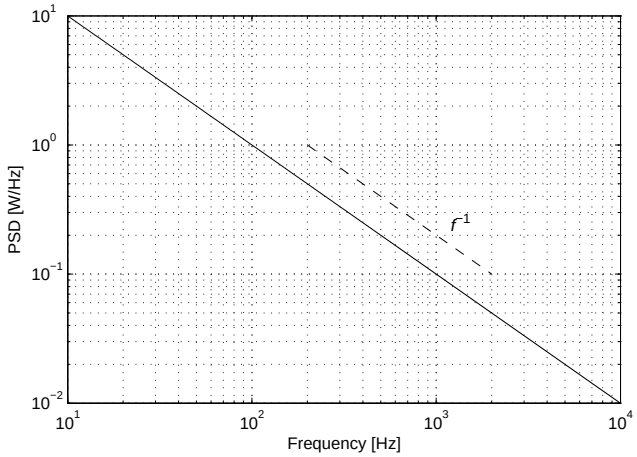


Figure 3.1: $1/f$ Spectrum

3.2 Measures for LF noise

LF noise may be examined in the frequency domain and in the time domain. The two, though fundamentally related, give quite different insights in noise behaviour. Both are useful.

LF noise in the frequency domain

Noise can be characterized by a plot of its Power Spectral Density (PSD). The PSD gives the noise power per unit of bandwidth as a function of frequency. The units of the PSD are W/Hz. This is often plotted on a logarithmic scale by converting to decibels, using a suitable reference power. In many practical experiments, the noise voltage or the noise current rather than the noise power is measured. A constant load impedance is assumed, and the relative noise power is simply expressed as V^2/Hz, or as A^2/Hz. Again, decibels are commonly used, leading to the oft-seen dBV^2/Hz (where the 0 dB reference is $1V^2/Hz$) or dBA^2/Hz (where the 0 dB reference is 1 A^2/Hz).

$1/f$ Noise derives its name from the $1/f$ shape of its PSD as in fig. 3.1.

$$S_{1/f} = \frac{K_f}{f^\gamma} \tag{3.1}$$

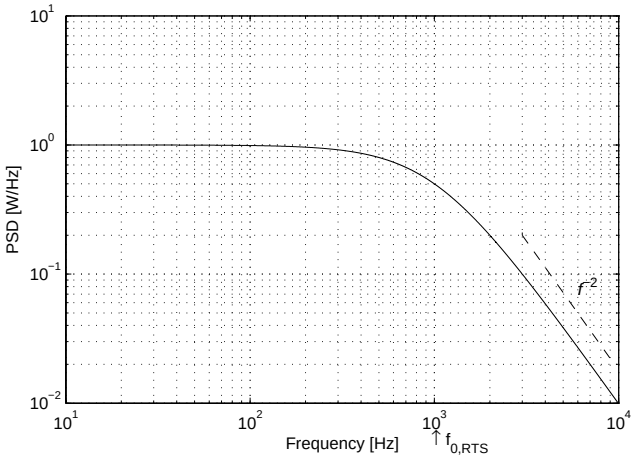


Figure 3.2: Spectrum of a Lorentzian signal

where K_f is a constant and γ is close to 1. A common limit for noise to be classified as $1/f$ noise is $0.9 < \gamma < 1.1$ [27]. When examining LF noise in small MOSFETs, a more complex spectrum with γ frequency dependent and varying between 0 and 2 is often observed. This type of spectrum is usually considered to consist of a summation of various so-called 'Lorentzian' spectra. A single Lorentzian has the following PSD:

$$S_{\text{Lorentzian}} = \frac{K_l}{1+(f/f_0)^2} \tag{3.2}$$

In which K_l is a constant and f_0 is the corner frequency of the Lorentzian. Below f_0, the PSD of the Lorentzian is flat; above f_0 it it has a f^{-2} characteristic. A Lorentzian PSD is given in fig. 3.2.

When studying LF noise, it is often desired to compare the LF noise in the steady state to the LF noise under LSE. Whereas the steady state spectrum will generally have a reasonably regular form similar to that in fig. 3.1 or fig. 3.2, the measured noise spectrum under large signal excitation is more complex: Not only will it contain the LF noise of interest, but also harmonics of the excitation signal and associated modulated LF spectra. In many cases, these modulated spectra are not of primary interest, as it is only desired to quantify whether the large signal excitation has changed the LF noise, and if so, in which direction and by how much.

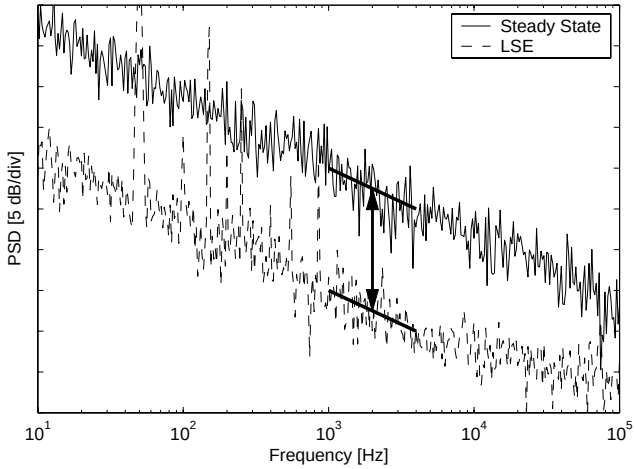


Figure 3.3: Spot noise measurement is suitable if PSD's to be compared have the same slope

Spot noise measurement

The difference between two spectra may be quantified in terms of the distance between them at a particular frequency. The 'spot' frequency at which the noise measurement is made should be given, but the spot noise measurement is only really suitable if the PSDs have the same slope for a significant frequency range as, for example, in figure 3.3, so this is often not very critical. A spot frequency should be selected that is free from spurs (undesired harmonic components in the PSD). The most obvious spur frequencies in the lab that should be avoided are multiples of 50 Hz; several are visible in fig. 3.3. The spot noise measurement is simple and insightful and is useful for rapid noise estimates in the lab. However, being a measurement at a single frequency, the spot noise measurement gives only limited information, and it gives no information on the shape of the spectrum at all.

Integrated noise measurement

When the PSD's to be compared are not parallel, for example in fig. 3.4 in which a significant Lorentzian component is seen in the steady state measurement, the spot noise measurement will give very variable results depending on the frequency at which it is applied, and it is no longer a

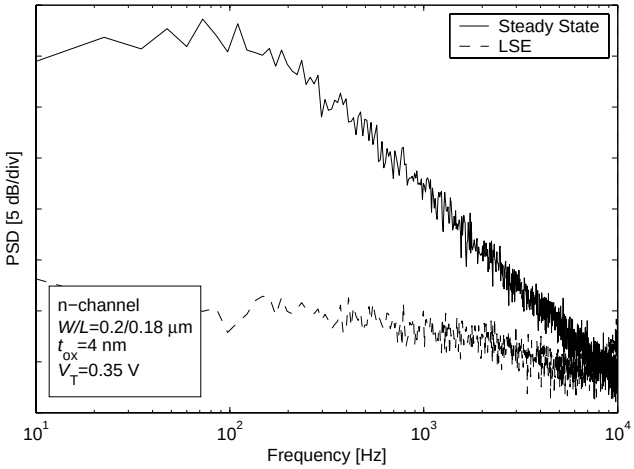


Figure 3.4: Steady state and LSE noise

representative estimate of the difference between the two PSD's. In such cases, a better approach is to integrate the noise over some bandwidth and compare the integrated noise of the PSD's. Care should again be taken to avoid spurs in the integration interval, and to give the best measure of noise, the integration interval should be as wide as reasonable.

Systematic effects and 'net noise reduction'

In some experiments, when comparing LF noise measurements in the steady state to those under LSE, a correction needs to be made for the fact that the device is 'off' for some of the time when it is subjected to LSE. One such example will be treated here. In this experiment [65], the device is periodically turned 'off' and 'on' in an abrupt manner and with a 50% duty cycle. The two signals of which the PSD's will be compared are given in fig. 3.5.

If the steady-state noise is described by

$$n(t) \tag{3.3}$$

the modulated signal can be described by

$$C(t) = n(t)B(t) \tag{3.4}$$

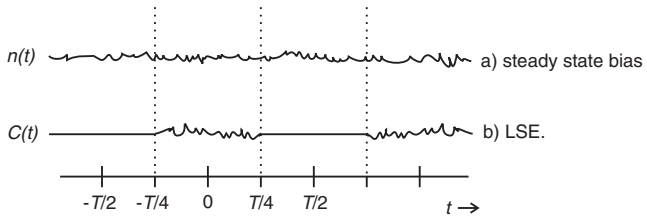


Figure 3.5: Systematic noise decrease under LSE: comparing equation 3.3 and 3.4

Where $B(t)$ is a square wave varying between 0 and 1, and $C(t)$ is cyclostationary. Its PSD can be calculated by first calculating its cyclostationary autocorrelation function $R_{CC}(t,t+\tau)$, 'stationarizing' it by averaging the autocorrelation function over one period [16], (this discards phase information which is not important for the PSD), and finally Fourier transforming it to obtain the PSD.

$$R_{CC}(t,t+\tau) = \mathrm{E}[n(t)B(t) \times n(t+\tau)B(t+\tau)] \tag{3.5}$$

Since $n(t)$ and $B(t)$ are uncorrelated, this can be rewritten as:

$$\begin{aligned} R_{CC}(t,t+\tau) =& \mathrm{E}[n(t)n(t+\tau) \times B(t)B(t+\tau)] \\ =& \mathrm{E}[n(t)n(t+\tau)] \times \mathrm{E}[B(t)B(t+\tau)] \\ =& R_{nn}(t,t+\tau) \times R_{BB}(t,t+\tau) \end{aligned} \tag{3.6}$$

$R_{nn}(t,t+\tau)$ is the autocorrelation function of the noise. If the noise is stationary, $R_{nn}(t,t+\tau) = R_{nn}(\tau)$. $R_{BB}(t,t+\tau)$ is not stationary (it is a function of both t and τ) and needs to be stationarized first. Referring to fig. 3.5, we first write:

$$\begin{aligned} R_{BB}(t,t+\tau) =& \underline{1}(t+\tau+T/4) \times \underline{1}(-(t+\tau)+T/4) \quad \text{for } \mid t \mid < T/4 \\ =& 0 \quad \text{for} \quad T/4 < \mid t \mid < T/2 \end{aligned} \tag{3.7}$$

In this equation, t is the time at which the autocorrelation function is calculated, and τ is the time offset that describes the autocorrelation function at that time. To stationarize this signal, the autocorrelation function is averaged over one cycle. The stationarized autocorrelation function is denoted

by R_{BBs}.

$$\begin{aligned} R_{BBs}(\tau) &= \frac{1}{T}\int_{-T/2}^{T/2} R_{BB}(t,t+\tau)dt \\ &= \frac{1}{2}\underline{1}(\frac{T}{2}+\tau)+\frac{\tau}{T}\underline{1}(\frac{T}{2}+\tau)-\frac{2\tau}{T}\underline{1}(\tau) \\ &\quad -\frac{1}{2}\underline{1}(-\frac{T}{2}+\tau)+\frac{\tau}{T}\underline{1}(-\frac{T}{2}+\tau) \end{aligned} \tag{3.8}$$

This is a triangular function of τ with height 0.5 and width $(-T/2..T/2)$. The Fourier transform can now be performed. The autocorrelation function is a multiplication of

$$R_{CCs}(\tau) = R_{nn}(\tau)\, R_{BBs}(\tau) \tag{3.9}$$

This will give a convolution of the PSD of the noise and the PSD of the square wave in the frequency domain. The relative amplitude of the baseband alias and the first alias is most interesting. The baseband alias will have a power of $1/4$ relative to the steady state case and the n^{th} alias will have an power of $2/(n\pi)^2$ relative to the steady state case. In the PSD, this is -6.0 dB for the baseband alias and -6.9 dB for the first alias. The 3$^{\text{rd}}$ alias will be at -16.5 dB. This is illustrated in fig. 3.6. Assuming that the baseband alias of the PSD can be validly interpreted if the first alias is some distance below the baseband alias of the PSD, the maximum frequency up to which the baseband alias can be correctly interpreted can be derived if we assume a $1/f$ shape of the PSD:

$$\frac{f_{\text{ex}}}{f_{\text{max}}} = 1+\frac{4}{1}\frac{2}{\pi^2}10^{\frac{d}{10}} \tag{3.10}$$

In this equation f_{max} is the highest frequency for which the baseband alias can be validly interpreted, f_{ex} is the frequency of excitation, and d is the distance in dB that is required between the baseband alias and the first alias. If a 3 dB distance between the baseband and the first alias is required, the baseband alias is free from aliasing up to a maximum frequency of $0.38\,f_{\text{ex}}$, as can be seen in fig. 3.6.

LF noise in the time domain

In the time domain, the most obvious way to characterize a stochastic signal is by its autocorrelation function. If the signal is cyclostationary, it can

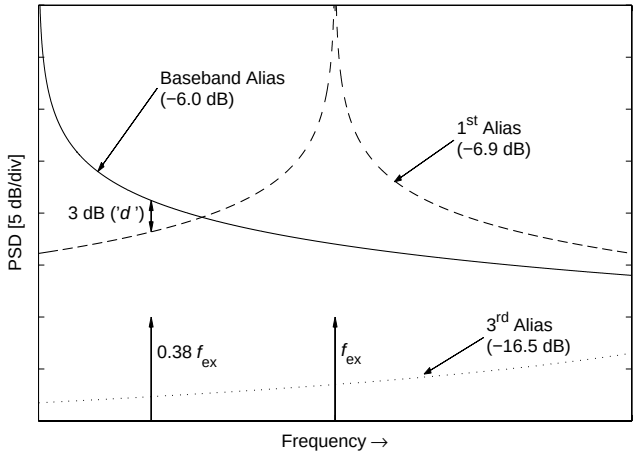


Figure 3.6: 1st Alias limits measurement of baseband alias

be stationarized first and the PSD can then be calculated using the stationarized autocorrelation function. However, stationarizing the autocorrelation function (as is done when a PSD is examined) is not always desirable. For example, in switched capacitor circuits, a noisy transistor that is subjected to a large (on-off) bias transient is used to take a sample of a signal. The cyclostationarity of the noise is relevant because the noise of the transistor *at a particular time in the period* is relevant, and stationarizing the noise by averaging the statistical parameters over the whole period is not appropriate [42]. For these sorts of situations a measure of noise at a particular time following a bias transient will be defined. This will be called the time-dependent noise.

Time-dependent noise

If the period of large-signal excitation is T, the time-dependent noise is defined as the series of samples taken at time instants $nT + t_{\text{offset}}$. n is the number of the period ($n = 1..n_{\max}$), and t_{offset} is the time offset in the period relative to the start of the period (fig. 3.7). The noise being subjected to LSE is cyclostationary so the time-dependent noise series is stationary, but its statistical parameters (including its mean and variance) may be dependent on t_{offset}. Extracting the time-dependent noise is equivalent to subsampling the original noise with a sampling frequency of f_{ex}, so the variance of the

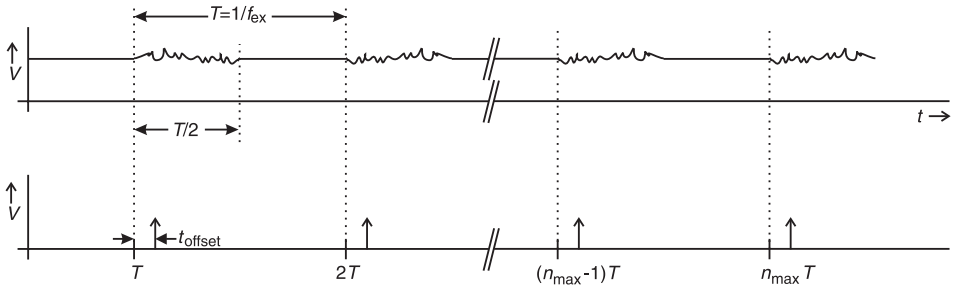


Figure 3.7: Time-dependent noise

time-dependent noise is equal to the integrated noise PSD from $1/n_{\max}T$ to the bandwidth of the noise at a particular point in the period. Referring to fig. 3.7, if $t_{\text{offset}} > T/2$ the time-dependent noise will be 0.

The time-dependent noise is an appropriate noise measure in systems where the use of the noisy device is synchronous with the large signal that the device is subjected to. Switched-capacitor circuits are an obvious example. In these systems, the variance of the time-dependent noise is of primary importance. The average value of the time dependent noise may also be a function of t_{offset}. This is important if the device is used at different times following a bias transient. A good example of this is the correlated double sampling circuit discussed in section 5.3

3.3 Review of published measurement results

In this section, previously published LF noise measurements under LSE are presented. The first observation of the fact that LF device behaviour depends on the bias history as well as on the bias state of the MOSFET dates from around 1985, however this was not published until much later [70]. First published measurement results are from 1991 [7]; later we presented more extensive measurement results on large-area HEF 4007 MOSFETs [65].

In [70], Eric Swanson discusses the design of the Crystal Semiconductor CS5016, a 50 kHz, 16 bit AD converter, which took place around 1985. A slow tail in a comparator offset voltage was observed following a comparator overdrive. Measurements were done on a 3 μm process from Orbit Semiconductor. It was also found ('5x worse') in 3N169 discrete MOSFETs.

The problem was solved by periodically pulling the source of the comparator transistors up to V_{DD}. A few ns of such 'flush time' was found to be sufficient to cure the effect, which was attributed to slow trapping and detrapping in the oxide. This is the first mention of the use of LSE to combat low-frequency effects in MOSFETs. No mention is made of improved LF noise characteristics of the transistors involved, though it is entirely possible that this happened as a side effect.

In [6, 7], Bloom and Nemirovsky present an experimental setup for and measurement results of LF noise in Si-MOSFETs subjected to LSE at the gate. Large MOSFETs were used ($100/8$ and $80/8$ μm from a CMOS process from 'Technion' in Israel), biased in the linear region and in saturation. The switching frequency was 600 Hz, and noise measurements were done between 1 and 100 Hz. A noise power reduction by a factor 3 was observed at 1 Hz. The effect was found to be independent of duty cycle for duty cycles between 25% and 88%. Various other transistors were measured, and a noise power reduction by a factor between 1.2 and 3.7 was observed at 1 Hz. The observed noise power reduction showed little sensitivity to the position of the sampling pulse within the 'on' period.

In [14], more noise measurements under LSE are presented by Dierickx and Simoen. A noise reduction was found 'Systematically on large-area devices'. It was observed that when subjected to LSE, 'RTS noise disappears', and the suggestion is made that subjecting the device to LSE could be 'a possibility to separate the contributions of different sources of $1/f$ noise in MOSFETs.' Experimental results are presented on $3.5/3.5$ μm MOSFETs from a 3 μm process from IMEC in Belgium. $t_{ox} = 60$ nm, And the devices have no LDD. For four n-channel devices, an average noise decrease by a factor 8.4 at 1 Hz was found. For four p-channel devices, this was a factor 3.9. The devices were in saturation. RTS noise was measured on four $0.4/0.8$ μm MOSFETs with a total of 8 traps. This measurement was done in the linear region. In steady state, τ_e was found to be independent of V_{GS}, while τ_c decreased with increasing V_{GS}. Under LSE, the same behaviour was observed, except for very low V_{GS}, where τ_e was seen to decrease with decreasing V_{GS}. This is associated with accumulation.

In [20], Gierkink shows that the phase noise of a discrete ring oscillator equipped with HEF 4007 MOSFETs depends on the amplitude of oscillation. Varying the amplitude of oscillation varies almost everything in the oscillator, so measurements of upconversion were made using a pilot-tone to calibrate the measurement setup. The remaining LF noise change after calibration is attributed to the LF noise change of the devices. An 8 dB

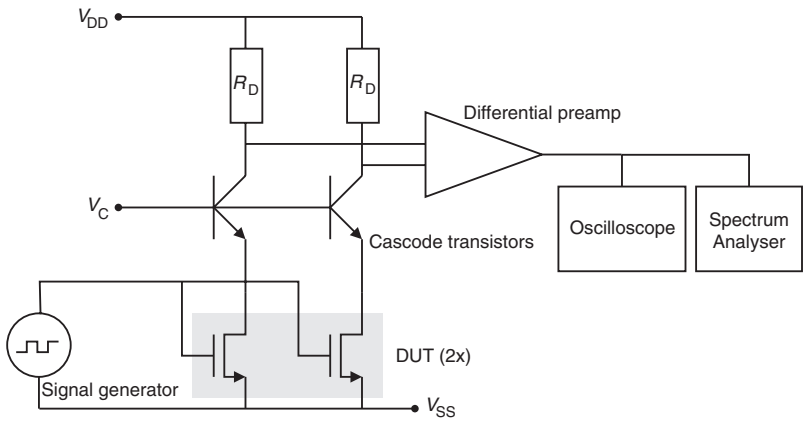


Figure 3.8: Noise measurement setup to measure LF noise in steady state and under LSE

reduction in LF noise is noted. This is reported for one sample, though measurements on devices from different manufacturers were performed.

In [65], we present measurements of HEF 4007 MOSFETs. A noise reduction due to LSE of 6...8 dB was systematically found in these devices. The devices are not specified for analog applications, and the process used is not specified either. Hence, exact device size and oxide thickness is not known. It is clear however that these devices are large-area (input capacitance of an n-channel device is typically 5.5 pF) and that they have a thick gate oxide ($V_T = 1.9$ V, and the gate oxide breakdown voltage is > 18 V).

Figure 3.8 shows the measurement setup used to measure LF noise under LSE as used in [65]. The cascode transistors serve to keep V_{DS} constant during the measurement, and the differential nature of the circuit (the two DUTs are driven in-phase) helps make this setup immune to external noise sources such as supply noise and noise from the signal generator. One of the two resistors R_D is variable to adjust for mismatch between the two DUTs. The LSE applied to the gates of the device is typically a 50% square wave with the high level of the square wave identical to the bias level of the steady state measurement; in this way the steady state measurement may be directly compared to the LSE measurement; bias dependency of LF noise does not influence the measurement.

A typical measurement is shown in fig. 3.9. At low frequencies, 6 dB between the LSE and the steady state curve is expected on mathematical

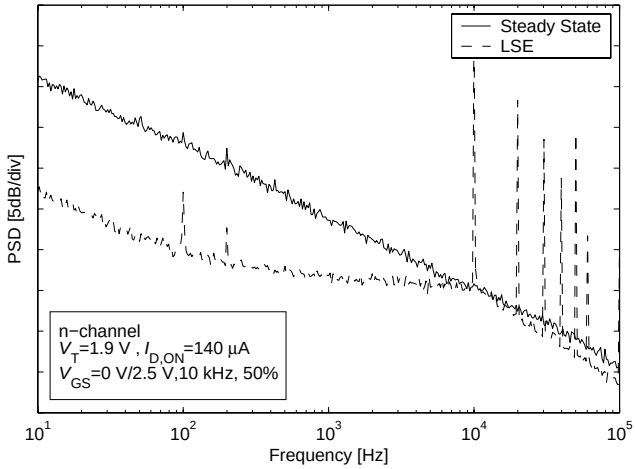


Figure 3.9: LF noise measurement on HEF 4007 device

grounds; the observed difference is 14 dB of which 8 dB is attributed to the decrease of LF noise due to LSE. Several devices were measured and an LF noise decrease of this magnitude was consistently observed.

A measurement was also performed in which the frequency of excitation was varied (fig. 3.10). The measurement results show that the frequency of excitation is not important for the LF noise as long as the frequency of excitation is high compared to the frequency of LF noise.

Finally, the 'off' voltage of the switching waveform was varied. It was observed that the devices needed to be cycled well below threshold before a significant noise reduction was seen. These results are plotted in fig. 3.11.

3.4 Large signal bias factors and their influence on LF noise

Introduction

In this section experimental results are presented that show how MOSFET LF noise under large signal excitation is influenced by different parameters of the large signal waveform. We will review the dependence of LF noise under LSE on:

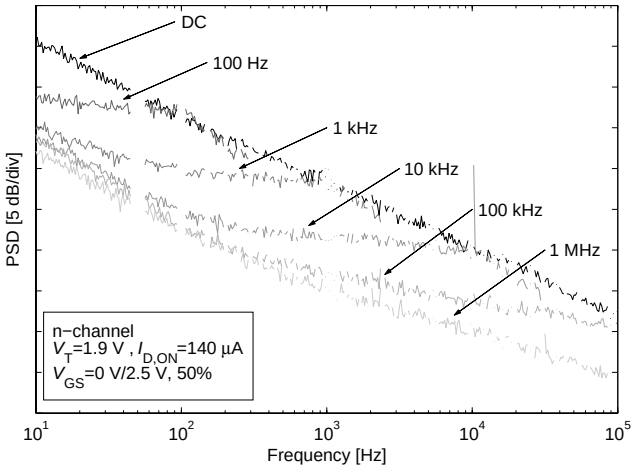


Figure 3.10: LF noise under LSE as a function of the frequency of excitation

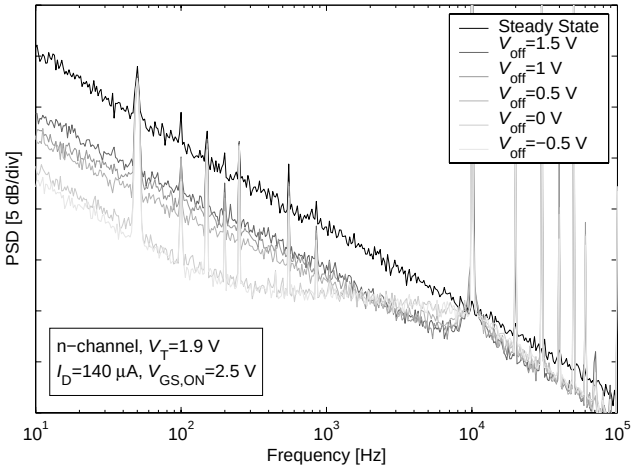


Figure 3.11: Device needs to be switched to well below threshold for noise reduction

- The circuit configuration,
- The duty cycle,
- How far the device is turned 'off'.

Circuit configuration

In a circuit, a MOSFET may encounter LSE in several ways. Depending on the design of the circuit, either V_G or V_S may be periodically time variant. V_D is not critical to the behaviour of the MOSFET as long as it is high enough to keep the device in saturation, and V_B is not always available as a separate terminal, or it may be shared between several devices. It is therefore relevant to examine the effect of periodic LSE at the gate and at the source. In particular, we want to find out whether there is a significant difference between these two modes of excitation as far as the LF noise performance of the MOSFET is concerned.

Comparative measurement method

To compare the effect of LSE at the gate and LSE at the source, the device is turned off periodically by either:

- varying V_G periodically between an 'on' value and an 'off' value while keeping V_S at a constant voltage, or
- varying V_S periodically between a low 'on' value and a high 'off' value, while keeping V_G at a constant voltage.

The LF noise in each of these situations is compared to the LF noise in steady state. The three states to be compared are shown in fig. 3.12. To make the steady state measurements directly comparable to the LSE measurements, the 'on' bias state in each of the three situations is made identical: the current is kept at 20 μA exactly. Depending on the W/L of the device, this results in a device that is more or less strongly inverted.

In each of the three situations, the noise at the source terminal of the device is measured using the measurement setup of fig. 3.13. Details of each operating condition are given in table 3.1. The first half of the setup is fully differential; i.e. all parts including the DUTs are duplicated and the preamplifier is differential. The large periodic signal is common-mode, and is therefore attenuated by the common mode rejection ratio of the differential difference amplifier. This reduces the dynamic range requirements of the remainder of the measurement setup. Working differentially also helps to

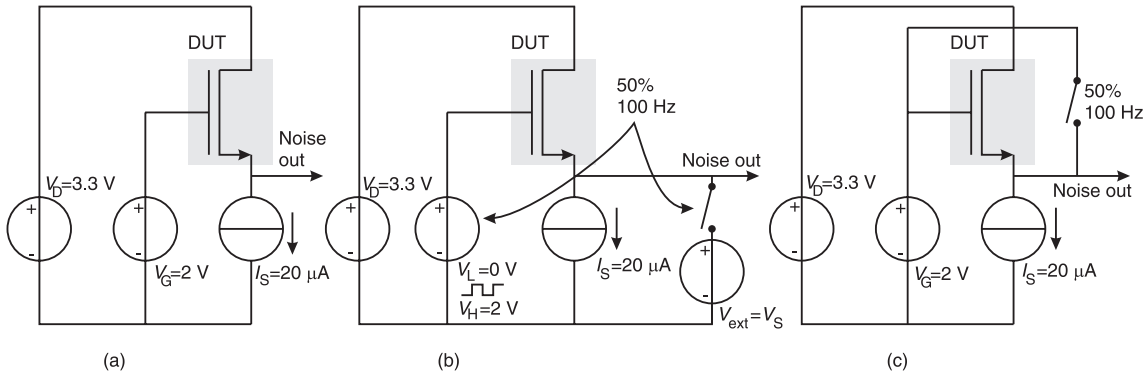


Figure 3.12: Comparison of three different bias conditions: (a) Steady state, (b) LSE at gate, (c) LSE at source.

suppress interference from the measurement setup and the different voltage sources such as V_{ext}. The device noise of the two DUTs is uncorrelated and their noise powers therefore add at the output.

The operating conditions of the different sources and switches is given in table 3.1. Note that switches B and C are also operated for the constant-bias measurement to make the operating conditions of the amplifiers as close as possible to the switched gate or switched source case. In this way, the possibility that the measurement results are caused by some transient effect in the amplifier can be precluded since all the measurements drive the amplifier in the same way. V_{ext} is an external voltage approximately equal to the common-mode input voltage of the differential amplifier, that prevents the amplifier from saturating when the device is disconnected from its input (when switch B is open). It is also used to prevent V_S from changing when V_G is varied.

Since I_D is determined by the gate-source voltage V_{GS}, the gate-referred voltage noise (which is of interest) is equal to the source-referred voltage noise (which is measured). The steady-state measurement is repeated twice for each device to check whether the measurement is consistent and reproducible. To further preclude possible effects of drift and temperature differences, the steady state and the switched measurements on each device are done in an alternate fashion.

For the LSE measurements, the device is turned on and off with a frequency of 100 Hz. The noise is sampled for 5 μs immediately following the turn-on

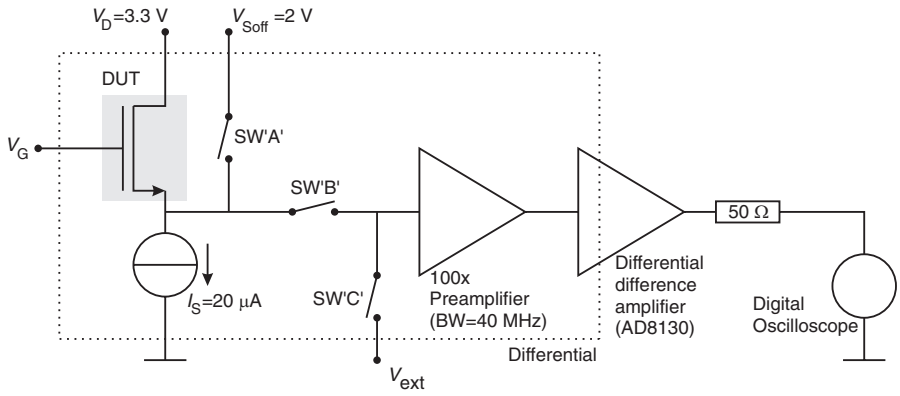


Figure 3.13: Measurement setup for LSE at gate and LSE at source

	Steady state	LSE at gate	LSE at source
V_G	2 V DC	0 V/2 V 50% 100 Hz	2 V DC
SW 'A'	OPEN	OPEN	50% 100 Hz
SW 'B'	50 % 100 Hz	CLOSED	Out of phase with SW 'A'
SW 'C'	Out of phase with SW 'B'	CLOSED when $V_G = 0$ V	In phase with SW 'A'

Table 3.1: Operating conditions for the measurement setup for steady state, LSE at gate and LSE at source

	Steady state	LSE at gate	LSE at source
V_G	2.0 V (DC)	Square wave, 0 V / 2 V, 50%, 100 Hz	2.0 V (DC)
V_S	Determined by I_D	V_{ext} / determined by I_D	Square wave, 2.0 V / determined by I_D, 50%, 100 Hz
V_D	3.3 V (DC)	3.3 V (DC)	3.3 V (DC)
I_D	20 μA per device	0 / 20 μA per device	0 / 20 μA per device

Table 3.2: Bias conditions for Steady state bias conditions, LSE at the gate and LSE at the source.

transient. The sampling frequency is 200 MHz, giving 1000 data points for each turn-on transient. After storage of the data (this takes approximately 1 second), the experiment is repeated. The experiment is repeated a total of 500 times for each bias condition (steady state, LSE at gate, LSE at source) for each sample. We want to determine the time-dependent noise (see page 38). To this end, the time-domain records are first passed through a lowpass filter with a cutoff frequency of 4 MHz to improve the LF noise to broadband noise ratio. After that, a sample is taken of the time domain data 0.5 μs after the turn-on transient. The variance of these 500 samples is an empirical measure for the LF noise in a frequency range up to about 0.5 Hz in all three bias conditions.

Measurement results

Measurements in all three bias conditions were carried out on a set of 42 n-channel devices from one wafer. The devices are from an industrial 0.35 μm process. Several different device sizes were available, with a W/L ratio of $0.8/0.35$, $1.025/0.475$, $1.125/0.525$, $1.25/0.575$, $0.5/0.35$ and $5/0.35$ μm. To assign equal weight to the different size devices, the variance of the time dependent noise is multiplied by the device area to normalize for the theoretical dependence of the LF noise on the area at constant I_D (see section 2.5). The bias conditions are given in table 3.2 below.

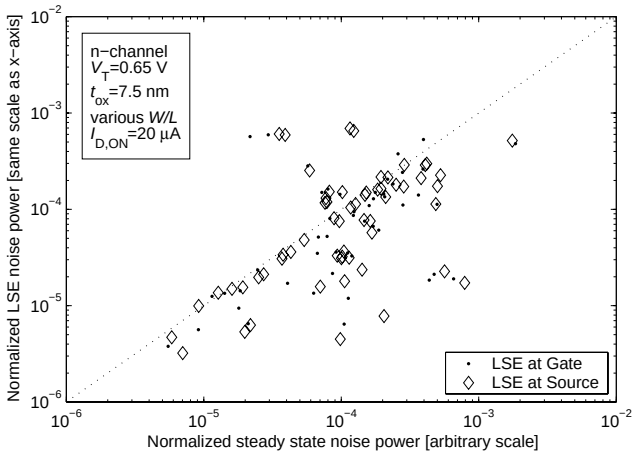


Figure 3.14: Noise under LSE at gate and LSE at source

An overview of the 'LSE at gate' and the 'LSE at source' measurements is given in figure 3.14 below. Along the *x*-axis, the LF noise power in the steady state is given, and along the *y*-axis, the LF noise power for LSE is given. Different symbols differentiate between the measurements with LSE at the gate and LSE at the source.

It is immediately apparent that the noise is very variable: the sample-to-sample spread is over 2 orders of magnitude. The noise under LSE shows similar spread. For reference, a dotted line signifying 'LSE noise=steady state noise' has been included in the figure. It can be seen that the majority of points lie below the line; indicating that on average, LF noise decreases when a device is subjected to LSE.

To investigate more closely the relation between LSE at the gate and LSE at the source, figure 3.15 was plotted. Along the *x*-axis the noise change for LSE at the source has been plotted. This is the LF noise under LSE at the source divided by the steady state LF noise for the same device. A value larger than 1 indicates an increase in noise for the LSE case; conversely, a value smaller than 1 indicates a decrease in noise for the LSE case. Along the *y*-axis, the same quantity is plotted for the 'LSE at gate' experiment. For reference, a line $y = x$ is also included in the figure. It is immediately apparent the noise change for LSE at the source correlates closely with the noise change for LSE at the gate. There is one data point that is far-removed from the 45^o line. On further investigation, it was found that this device had

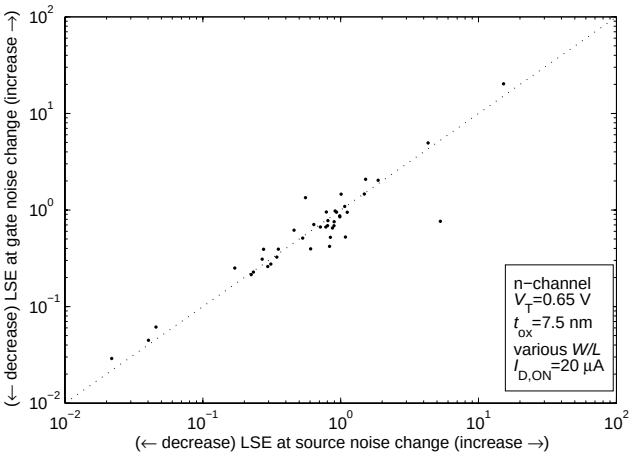


Figure 3.15: Switched Gate compared to Switched Source

a very slow RTS, with time constants of the order of the measurement-time (approx. 500 s), which influenced the measurement. A longer measurement on this device would give better results.

Conclusion

Subjecting the device to LSE at its gate and at its source is largely equivalent in terms of LF device noise.

Duty cycle

When subjecting a device to LSE, it is interesting to examine the effect of duty cycle on the LF noise. The duty cycle is the fraction of the period that the device is 'on' and generating noise. In terms of LF noise under LSE, it is important to know whether the device needs to be turned 'off' for a large percentage of the time to influence its LF noise, or whether subjecting the device to a brief 'off' pulse is sufficient to influence the LF noise of the device. Obviously turning a device 'off' for a large percentage of time reduces its average output noise, but the LF output noise is examined after correcting for such systematic effects.

The first report of an LF effect under LSE [70] notes that accumulating the surface of the device 'for a few ns' had significant effect, and removed the LF transient present in the device. Later, LSE experiments were carried out [7] where the duty cycle was found to be unimportant for duty cycles between 25 and 88%.

LF noise measurements were carried out [17] on p-channel devices with 10 and 20 nm t_{ox}. W/L was $10/0.3$ and $10/1$ μm. 16 samples were measured in total; out of these, 5 showed an LF noise decrease as the duty cycle was decreased; the other 11 showed the duty cycle to be largely unimportant for the LF noise change that was observed when subjecting the device to LSE.

Next, measurements were carried out on a number of RTS-dominated devices [24]. These devices were n-channel MOSFETs from a 0.18 μm process[1]. Three devices were selected for their very visible RTS behaviour. The LF noise of these devices (integrated from 1 to 20 Hz and corrected for systematic effects) is plotted as a function of the duty cycle in fig. 3.16. In this figure, one can clearly see that the behaviour of the RTS is strongly dependent on the duty cycle. Depending on the characteristics of the RTS, the behaviour is very different. RTS 1 has the lowest LF noise, and it decreases as the duty cycle decreases. RTS 2 has a maximum LF noise at a duty cycle of around 40..60%, and RTS 3 shows a maximum in LF noise for a duty cycle of 90%. Time domain measurements confirm that this behaviour is in correspondence with the 'effective τ's' of the RTS (see section 4.3). These are selected devices and the results may not be typical of all MOSFETs.

Conclusion

We have shown that whether or not the duty cycle of the switching waveform is important depends on the parameters of the RTS in the device in question. The LF noise power behaves in accordance with the RTS parameters in the 'on' and the 'off' state, more detail of which is given in chapter 4. When the RTS parameters are not influenced by LSE, the duty cycle of the switching waveform is not important; in other cases, turning the MOSFET 'off' for a brief period of time can be sufficient to reduce the LF noise of the device, and turning it off for longer does not produce additional LF noise benefits [7, 70].

[1] This is nominally a 0.18 μm process, but the use of phase-shift masks allows minimum device lengths of 0.13 μm

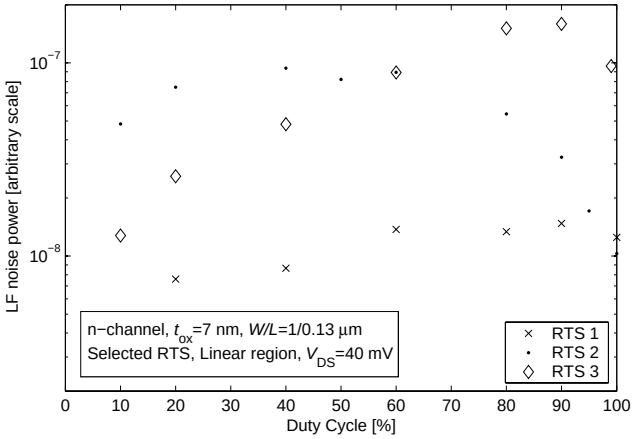


Figure 3.16: LF noise as a function of duty cycle of the LSE

Waveform

Several measurements indicate that in order to influence the LF noise, the LSE the device is subjected to has to cycle the gate-source voltage to well below threshold, eg. as observed in [65]. In this section, experimental data will be presented.

Measurements were performed on p-channel devices. A number of devices with different oxide thickness were measured. The devices were biased in saturation and LSE was applied at the gate of the device as a 10 kHz, 50% square wave. The measurement results are given in fig. 3.17 below. Clearly, cycling the device to well below threshold is beneficial in reducing the LF noise of the device.

Further research was done on RTS-dominated devices. Measurements in the time domain [24] indicate that the RTS time constants change significantly if the device is cycled to a voltage well below threshold. Measurements of the LF PSD (figure 3.18) of these devices are in agreement with this result.

Conclusion

In order to influence the LF noise of a MOSFET by subjecting it to LSE, the gate-source voltage needs to be cycled to ‘well below’ threshold.

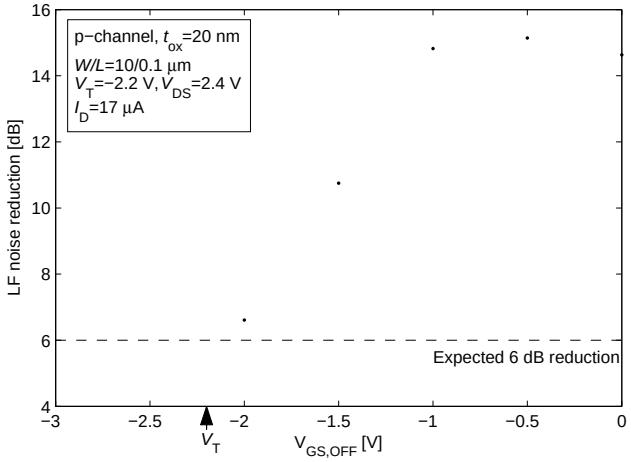


Figure 3.17: Device needs to be cycled to well below threshold for noise reduction

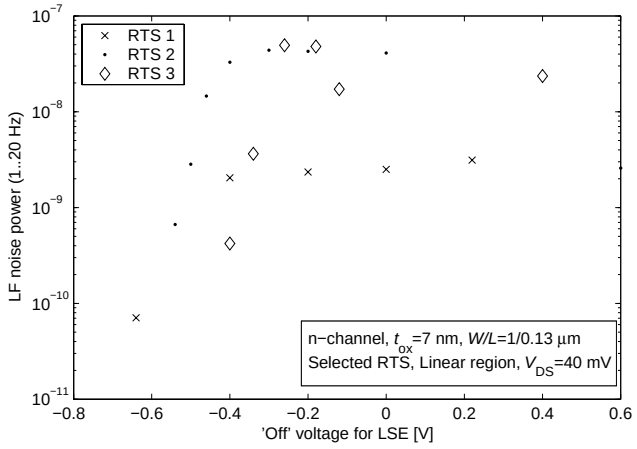


Figure 3.18: Device needs to be switched to well below threshold for noise reduction

The exact bias conditions under which a significant LF noise change is observed depends on the device parameters. In early literature [7, 70] it was noted that the device needs to be 'cycled to accumulation' for any significant change in the LF noise of the device to occur. For the n-channel MOSFETs used, this was done by making the gate-bulk voltage negative [7] or by pulling the source up to V_{DD} [70]. Later measurements [65] showed that cycling the device to accumulation was not always necessary: cycling V_{GS} to 0 V was sufficient. This is supported by later measurements, where turning the device 'off' by pulling V_S up to V_G was also found to influence the LF noise of the device.

3.5 Device factors and their influence on LF noise

Introduction

In this section, measurements are presented of LF noise under large signal excitation on a variety of MOSFETs. It will be shown that p-channel devices behave in the same way that n-channel devices do. Measurement results are given that show how LF noise scales with device area and oxide thickness.

Device type

The difference in LF noise behaviour in steady state and under LSE was measured for p-channel devices and n-channel devices. Fig. 3.19 shows the results for a large number of p-channel devices, and fig. 3.20 shows the same for a number of n-channel devices. We conclude that the p-channel devices show the same trends in noise behaviour as the n-channel devices.

Device size

To investigate the dependence of LF noise on device size, measurements were carried out on a set of devices with varying W and L. The devices are from a 0.18 μm process, and the measurement is carried out at constant I_D. The result is plotted in fig. 3.21; LF noise power is seen to scale inversely with the area of the device, which is in accordance with the ΔN model for LF noise (see section 2.5).

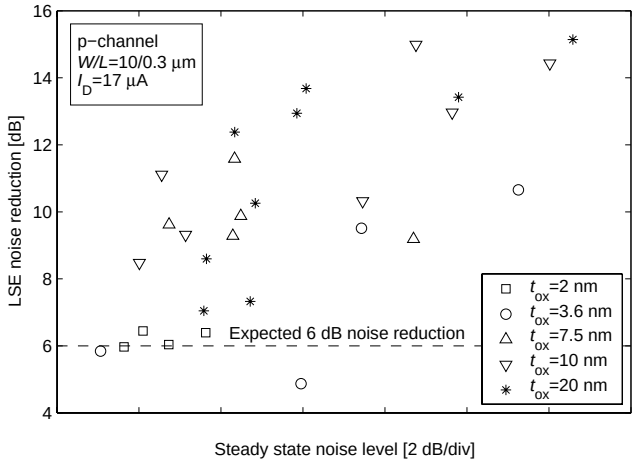


Figure 3.19: p-channel LF noise behaviour in steady state and under LSE

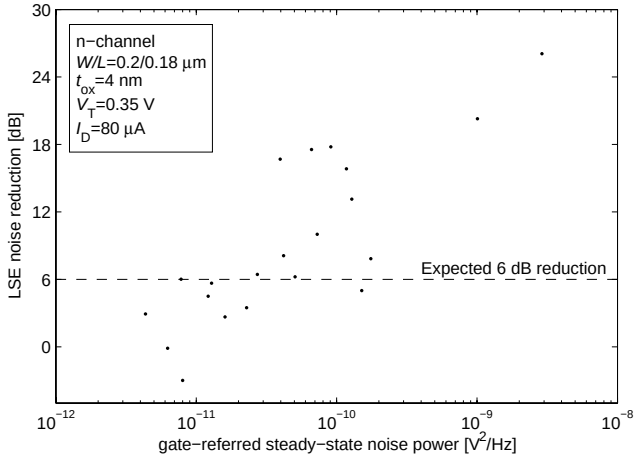


Figure 3.20: n-channel LF noise behaviour in steady state and under LSE

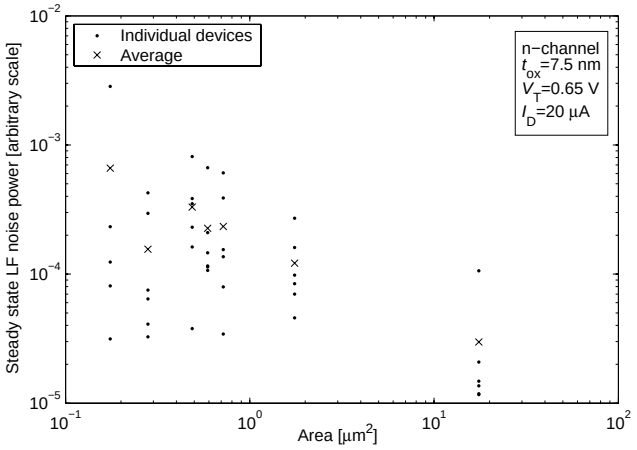


Figure 3.21: Noise power scales inversely with area

Technological factors

Measurements on a large number of different processes were carried out. An LF noise change under LSE was observed in almost all cases. Measurements were performed on: HEF 4007 devices (large devices from what is probably a 10 µm CMOS process), 0.35 µm CMOS processes, 0.25 µm CMOS processes and 0.18 µm CMOS processes. Additionally, measurements were performed on a special wafer set (nominal 0.18 µm process) that was processed with varying t_{ox}. In this way, the effect of the oxide thickness can be investigated in isolation.

The results in fig. 3.22 show that the LF noise power at the drain increases linearly with oxide thickness. The measurement is done at constant I_D. This is in accordance with the ΔN model (section 2.5).

Measurement results of LF noise reduction under LSE are given in fig. 3.23. The LF noise reduction is seen to increase as t_{ox} increases.

3.6 Spread in LF noise under large signal excitation

It was found that the LF noise performance of small MOSFETs (both in steady state and under LSE) shows great spread. As far as the steady state LF noise measurements are concerned, this has been reported before in li-

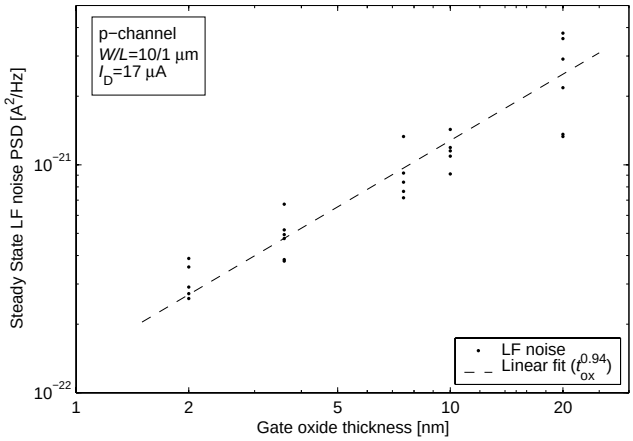


Figure 3.22: LF noise power increases with increasing t_{ox}. Measurements from special wafer set with varying oxide thickness; all devices same nominal size.

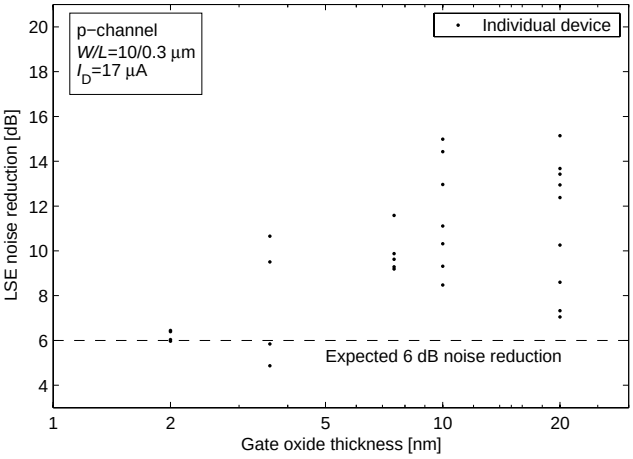


Figure 3.23: LF noise decrease under LSE as a function of t_{ox}. Measurements from special wafer set with varying t_{ox}; all devices same nominal size.

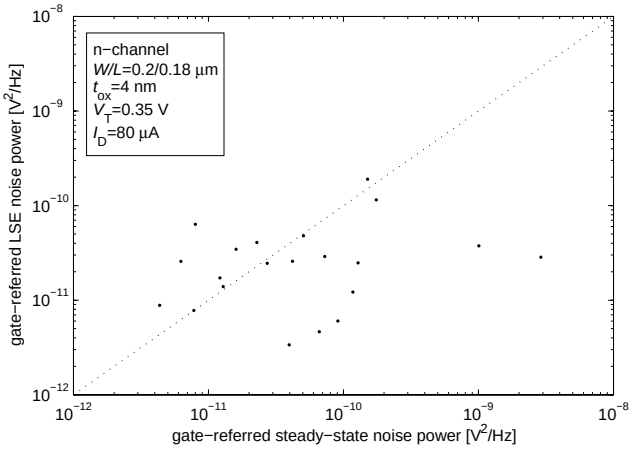


Figure 3.24: Results of 21 n-channel devices

terature [9]. As far as the LSE noise measurements are concerned, this data was published for the first time in [66]. If LF noise reduction by LSE is to be applied in circuit design, statistical behaviour of the effect is of paramount importance.

The first measurement on a large number of nominally identical devices is given in fig. 3.24. These are results of 21 n-channel devices from a 0.18 µm process. Noise is measured in a bandwidth of 10 Hz-150 Hz. LSE is a square wave of 50% duty cycle, 10 kHz, applied at the gate.

In fig. 3.25 results for 41 n-channel devices from a 0.35 µm process are given. These devices were of varying sizes (W/L from $0.5/0.35$ µm to $5/0.35$ µm). The LSE is a 50% square wave of 100 Hz, applied at the gate. The LF noise quantity measured is the variance of the time dependent noise measured 0.5 µs after turn-on of the device.

In fig 3.26 results for 24 n-channel devices are given. These devices are from a 0.25 µm process. The LSE is a 10 kHz square wave with a 50% duty cycle and $V_{\text{off}} = -0.6$ V. The LF noise is integrated between 10 Hz and 1 kHz for this measurement. Devices were in saturation. This measurement was repeated on 56 devices from the same process, but with V_{GS} cycled to 0 V rather than -0.6 V. The results are plotted in fig. 3.27. A group of 30 p-channel devices was also measured. The results are plotted in fig. 3.28. I_{D} is constant for this measurement, and t_{ox} varies between 2 and 20 nm.

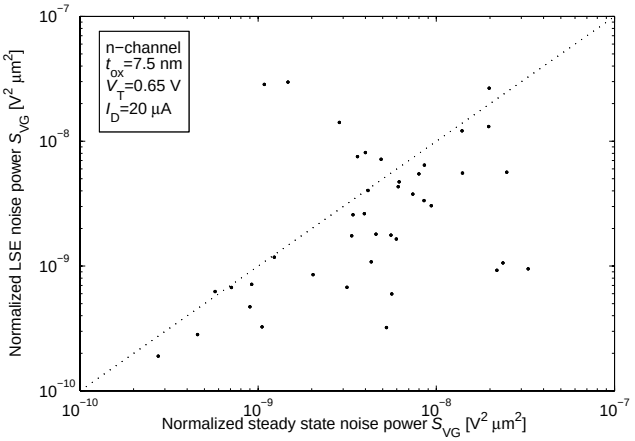


Figure 3.25: Results of 41 n-channel devices

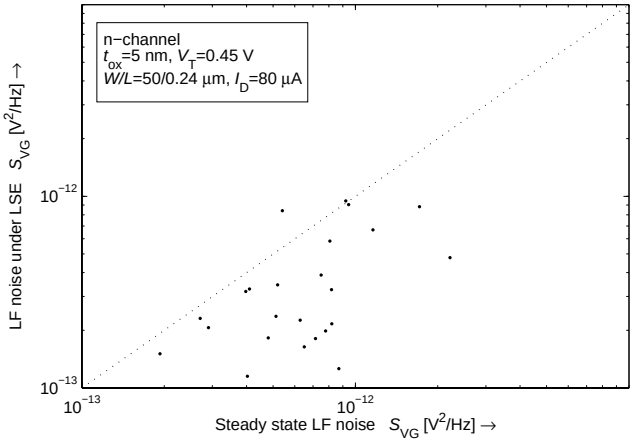


Figure 3.26: Results of devices from a 0.25 µm process cycled to $V_{GS} = -0.6$ V

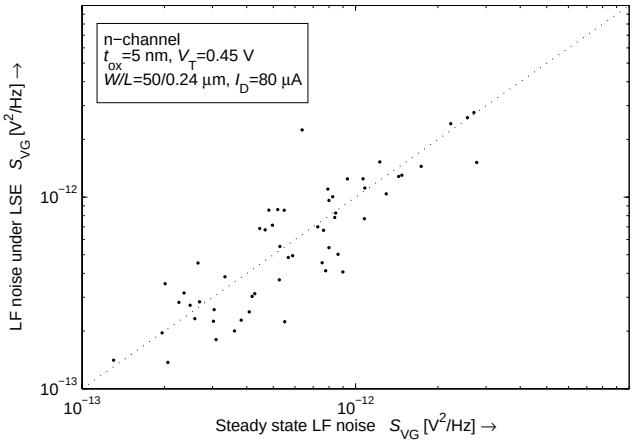


Figure 3.27: Results of devices from a 0.25 µm process cycled to V_{GS} =0 V

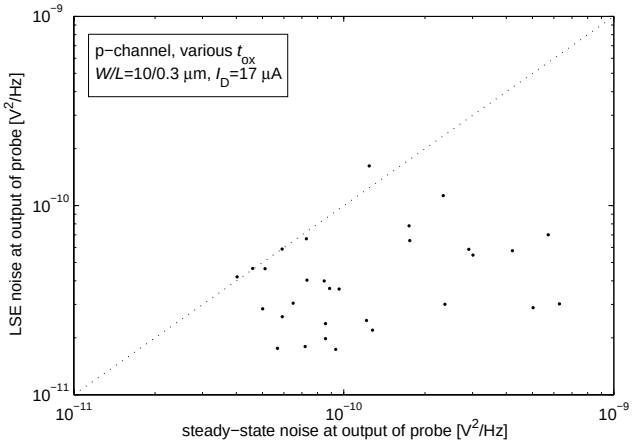


Figure 3.28: Results of $>$ 30 p-channel devices

Measurement from	Average noise power			Spread in noise power (σ)		
	Steady State	LSE	LSE / S.State	Steady State	LSE	LSE / S.State
fig. 3.24 (0.18 μm n)	2.4×10^{-10}	3.6×10^{-11}	0.15	6.5×10^{-10}	4.3×10^{-11}	0.07
fig. 3.25 (0.35 μm n)	7.3×10^{-9}	5.3×10^{-9}	0.73	7.8×10^{-9}	7.4×10^{-9}	0.95
fig. 3.26 (0.25 μm n)	7.4×10^{-13}	3.8×10^{-13}	0.51	4.5×10^{-13}	2.7×10^{-13}	0.6
fig. 3.27 (0.25 μm n)	7.7×10^{-13}	7.6×10^{-13}	0.99	6.2×10^{-13}	6.2×10^{-13}	1
fig. 3.28 (0.18 μm p)	1.7×10^{-10}	4.6×10^{-11}	0.27	1.6×10^{-10}	3×10^{-11}	0.19

Table 3.3: Measurements of spread

Table 3.3 gives a summary of the measurement results as far as spread is concerned. It is seen that in almost every case LSE makes the average LF noise go down, and in some cases this is accompanied by a significant reduction in the LF noise spread. For the 0.25 μm process (fig. 3.26 and 3.27), the devices need to be turned 'off' to -0.6 V before any significant effect is noticed. The larger devices (0.25 μm n in the table; $A = 12$ μm^2) show less spread than the smaller devices. The spread in LF noise mandates correct use of statistics when designing close to the LF noise limit. Care should be taken since the distribution of the LF noise is not well known; σ is not much smaller than the average, so it is clearly not Gaussian.

3.7 Conclusion

In this chapter, measurements of LF noise under LSE have been presented. The majority of devices measured showed RTS-dominated LF noise. In all measurements, subjecting the devices to LSE decreased the average LF noise of the devices.

As far as bias factors are concerned, turning the device 'off' by varying V_S changes the LF noise in the same way as by varying V_G. However, in all cases, V_{GS} needs to be cycled well below V_T before any significant effect on

the LF noise is observed. The voltage below which this happens depends on the parameters of the process in question.

A reduction in LF noise when the device is subjected to LSE is observed in all processes that were measured, ranging from very old processes (HEF 4007) to new (0.18 μm) processes, from experimental to mature industrial processes. In all cases, subjecting the MOSFET to LSE was seen to influence its LF noise. This was observed for both n-channel and p-channel devices. Measurements on similar devices with varying oxide thicknesses (a special wafer set with varying t_{ox} but with otherwise identical devices was used) were performed; no significant deviation from existing LF noise models (ΔN) was observed.

Also, for the first time, measurements of spread of LF noise under LSE are presented. LF noise of the devices is not normally distributed, so care should be taken when using such devices in noise-limited applications, as the distribution of the noise is not, in general, known. While it is as yet impossible to predict the LF noise behaviour of a single device when subjected to LSE, the average LF noise of a group of devices as well as the spread of the LF noise will go down when the devices are subjected to LSE.

Chapter 4

Modelling

4.1 Introduction

The low frequency noise observed in n-channel MOSFETs, especially in small-area devices is caused predominantly by Random Telegraph Signals (RTS). RTS noise (see section 2.3) is always observed in small devices with a low number of free carriers as shown in section 2.5. It may also be observed in larger devices.

An RTS in a MOSFET is caused by a so-called 'trap' which is a localized energy state. Trapping behaviour in a MOSFET is governed by Fermi-Dirac statistics, as described in the classical Shockley-Read-Hall (SRH) model for trapping and detrapping [60]. The SRH model is commonly applied under steady state conditions to derive the effective occupancy of a trap. As is well noted in the original paper by Shockley and Read and in [54], it is equally valid under transient conditions, which is the case treated here.

In this chapter, a model is presented that explains how the RTS noise changes in a MOSFET subjected to large signal excitation (LSE). It is based on the Shockley-Read-Hall model for capture and emission of electrons by traps. Earlier models to predict the behaviour of RTS noise under LSE [19, 62] cannot account for the experimental observations that in some cases, the LF noise of the device increases when subjected to LSE, and that not all of the LF noise of the device disappears when the device is subjected to LSE.

This chapter is organized as follows: First, the SRH model for capture and

emission will be briefly reviewed. Next, the predictions made by SRH model under transient conditions will be treated analytically. This leads to closed-form expressions for some limiting cases, which explain why an LF noise decrease is often observed when a MOSFET is subjected to LSE. Next, a simulator is presented [68] for more complex cases where the analytical approach is unpractical. The simulator is seen to produce results in excellent accordance with measurement results. Finally, the model is discussed. In this chapter we focus on the macroscopic properties of the model; a more detailed description of the physics involved is given in [40] and [38]. Though the model works well in an experimental setting, the main shortcoming is that it cannot predict in advance what the parameters of the traps in MOSFETs are.

4.2 Model

Shockley-Read-Hall statistics

A trap is a localized energy state in the bandgap that has an energy level between the conduction band energy level (E_c) and the valence band energy level (E_v). Traps occur at the Si-SiO_2 interface or in the oxide and are attributed to dangling bonds or impurities. Traps that are close in energy to the conductance band edge interact with the conductance band, and traps that are close to the valence band edge will interact with the valence band. In the following section, the behaviour of traps near the conductance band edge is discussed; these are the traps that dominate the LF noise behaviour of n-channel MOSFETs. For p-channel devices, an analogous derivation can be made for traps near the valence band edge.

A trap (fig. 4.1) near the conductance band edge can interact with the conductance band by capturing or by releasing an electron. Capture or emission of an electron from the conduction band causes RTS noise (observed in the drain current of the device) in two ways: First of all, the electron that is captured from the channel does not take further part in the conduction process in the device; this is known as the ΔN effect. Secondly, capture of an electron in a trap changes the charge of the trap. Capture of an electron in a trap will make the trap more negatively charged, regardless of its initial charge state. This has the effect of modulating the position of the channel; and is known as 'Coulomb scattering' or as a $\Delta\mu$ effect caused by and correlated to the capture of an electron (ΔN). This second effect [57] often causes a much larger drain current fluctuation than can be explained on the basis of

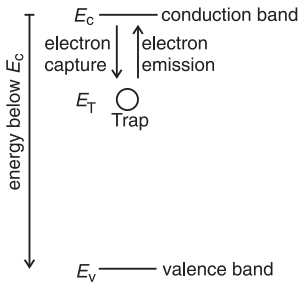


Figure 4.1: In n-channel MOSFETs, traps near the conduction band edge cause RTS noise

the simple ΔN effect.

The parameters of the RTS noise are: (1): Its amplitude, (2): The mean time before capture of an electron (τ_c), and (3): The mean time before emission of an electron (τ_e). This is discussed in detail in section 2.3.

On the basis of Fermi-Dirac statistics, it can be shown that:

$$\begin{aligned} \tau_c &= [\sigma(E,x) v_{th} n]^{-1} \\ \tau_e &= \left[\sigma(E,x) v_{th} N_c e^{-(E)/kT}\right]^{-1} \end{aligned} \tag{4.1}$$

$\sigma(E,x)$ Is the capture cross section of the trap, which is dependent on the energy level of the trap below the conduction band edge ($E = E_c - E_T$) and its depth in the oxide away from the channel (x). v_{th} Is the thermal velocity of the electrons, n is the electron density in the conduction band, and N_c is the effective density of states in the conduction band.

The rates of capture and emission are given by:

$$\begin{aligned} r_c(E,t) &= \frac{1 - f(E,t)}{\tau_c} \qquad [s^{-1}] \\ r_e(E,t) &= \frac{f(E,t)}{\tau_e} \qquad [s^{-1}] \end{aligned} \tag{4.2}$$

In these equations, $f(E,t)$ is the occupancy of a trap: it is the probability of finding a trap with energy level E filled at time t. The net rate of capture is given by:

$$\frac{df(E,t)}{dt} = r_c(E,t) - r_e(E,t) \tag{4.3}$$

This equation is central in the study of traps, since it is valid in the steady state as well as during a transient. It is a first order differential equation in $f(E,t)$ and can be solved [54] to give:

$$f(E,t) = \frac{\tau_e}{\tau_e + \tau_c} + K e^{-(\frac{1}{\tau_e}+\frac{1}{\tau_c})t} \tag{4.4}$$

K is an integration constant that will depend on the initial condition. In equilibrium ($t \to \infty$), equation 4.4 reduces to the well known form:

$$f(E) = \frac{\tau_e}{\tau_c + \tau_e} \tag{4.5}$$

Bias dependence of τ_e and τ_c

According to basic SRH-theory (eq. 4.1), τ_c is bias dependent via the bias dependency of n. τ_e should not be dependent on bias, only on the energy level E of the trap. If the trap is situated some distance in the oxide, however, E may depend on V_{GS}. As a result of this, τ_e may also be bias dependent.

This is consistent with measurements of RTS time constants done under steady state conditions. Measurements of the bias dependency of RTS time constants in n-channel MOSFETs are given in [24, 33, 41, 47, 59] and [38]. In all cases, it is found that as V_{GS} is decreased τ_c increases and τ_e decreases, (the probability of a trap being full decreases and the probability of a trap being empty increases). The change in τ is commonly up to two orders of magnitude, though even more is observed in selected devices [24].

$$\begin{aligned} \tau_c &\quad \text{increases with decreasing } V_{GS} \\ \tau_e &\quad \text{decreases with decreasing } V_{GS} \end{aligned} \tag{4.6}$$

RTS noise under LSE is modelled by making the following assumptions:

- The LF noise of the MOSFET is dominated by a single RTS with a particular energy level E and location in the oxide x.
- The behaviour of the RTS is governed by Shockley-Read-Hall statistics. This is true in steady state and also under transient conditions.
- τ_e And τ_c exhibit bias dependence in accordance with 4.6.
- τ_e And τ_c are instantaneous functions of the bias of the device.

4.3 Analytical

The occupancy of a trap with a given energy level as a function of time, $f(t)$, is derived for the simple case where the bias voltage alternates abruptly and

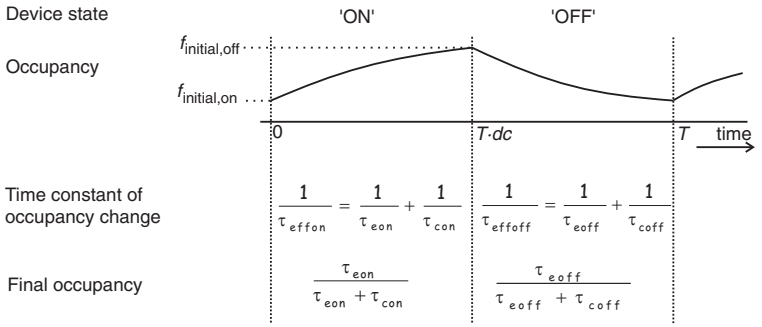


Figure 4.2: Occupancy of trap

periodically between two states. This is analytically the simplest case. For insight, the hypothetical case where the RTS amplitude does not depend on the bias state of the device is treated, and only the RTS time constants are time-variant. Under the condition that the excitation frequency is high compared to the RTS corner frequency, it is seen that a cyclostationary RTS can be modelled by an equivalent stationary RTS; the time constants of this equivalent stationary RTS will be derived. The fact that in the measurement, the RTS is invisible during a fraction of the period can be considered in isolation and will be accounted for separately.

In deriving the occupancy $f(t)$, reference is made to fig. 4.2.

dc Is the duty cycle of the cyclostationary RTS, which varies between 0 and 1 and expresses the fraction of the period in which the device is in the 'on' state. During the first part of the period, from $t = 0 \ldots Tdc$, the device is 'on' During this time, RTS behaviour is governed by $\tau_{e,on}$ and $\tau_{c,on}$. During the second part of the period, from $t = Tdc \ldots T$, the device is 'off', and RTS behaviour is governed by $\tau_{e,off}$ and $\tau_{c,off}$. The occupancy is given by eq. 4.4 at all times, with the addition that in eq. 4.4 there were only two time constants, τ_e and τ_c, and now there are four: $\tau_{e,on}, \tau_{e,off}, \tau_{c,on}$ and $\tau_{c,off}$.

Making the substitutions

$$\begin{aligned} \tau_{\text{eff,on}} &= \left[\frac{1}{\tau_{\text{e,on}}} + \frac{1}{\tau_{\text{c,on}}} \right]^{-1} \\ \tau_{\text{eff,off}} &= \left[\frac{1}{\tau_{\text{e,off}}} + \frac{1}{\tau_{\text{c,off}}} \right]^{-1} \end{aligned} \tag{4.7}$$

we may write, for the first half of the period,

$$f(t) = \frac{\tau_{\text{e,on}}}{\tau_{\text{e,on}} + \tau_{\text{c,on}}} - \left[\frac{\tau_{\text{e,on}}}{\tau_{\text{e,on}} + \tau_{\text{c,on}}} - f_{\text{initial,on}} \right] \text{e}^{-t/\tau_{\text{eff,on}}} \quad \text{for } t = 0..Tdc \tag{4.8}$$

In the same way, during the second half of the period,

$$f(t) = \frac{\tau_{\text{e,off}}}{\tau_{\text{e,off}} + \tau_{\text{c,off}}} - \left[\frac{\tau_{\text{e,off}}}{\tau_{\text{e,off}} + \tau_{\text{c,off}}} - f_{\text{initial,off}} \right] \text{e}^{-(t-Tdc)/\tau_{\text{eff,off}}} \quad \text{for } t = Tdc..T \tag{4.9}$$

Since the occupancy of the trap is continuous in time, and the behaviour is cyclostationary, $f(Tdc)$ as given by the first equation must equal $f_{\text{intital,off}}$, and $f(T)$ as given by the second equation must equal $f_{\text{initial,on}}$. In this way, we can solve for $f_{\text{initial,on}}$ and $f_{\text{initial,off}}$:

$$\begin{aligned} f_{\text{initial,on}} &= f(T) \\ &= \frac{\frac{\tau_{\text{e,off}}}{\tau_{\text{e,off}}+\tau_{\text{c,off}}} \left(1 - \text{e}^{\frac{-T(1-dc)}{\tau_{\text{eff,off}}}} \right) + \frac{\tau_{\text{e,on}}}{\tau_{\text{e,on}}+\tau_{\text{c,on}}} \text{e}^{\frac{-T(1-dc)}{\tau_{\text{eff,off}}}} \left(1 - \text{e}^{\frac{-Tdc}{\tau_{\text{eff,on}}}} \right)}{1 - \text{e}^{\frac{-Tdc}{\tau_{\text{eff,on}}}} \text{e}^{\frac{-T(1-dc)}{\tau_{\text{eff,off}}}}} \end{aligned} \tag{4.10}$$

$f_{\text{initial,off}}$ can be found in a similar way.

It is useful to examine what happens if the switching frequency is made very high compared to the RTS corner frequency. To this end, the limit is taken for $T \to 0$. In this limit, the frequency can be considered so high that any analysis interval *dt* will contain an integer number of periods of the switching signal, and since the signal is *cyclostationary*, it becomes stationary in the limit for $T \to 0$. If it is stationary, the probability of a transition per unit time is constant, i.e. it is an RTS. This equivalent stationary RTS will be described in terms of the parameters of the cyclostationary RTS. The amplitude of the RTS is the same as that of the original RTS. We will call the equivalent capture and emission time constants for this equivalent stationary RTS $\tau_{\text{c,eff}}$ and $\tau_{\text{e,eff}}$. They are derived below:

The 'effective occupancy' of the equivalent stationary RTS is found by examining $f_{\text{initial,on}}$ and $f_{\text{initial,off}}$, which, in the limit for $T \to 0$ converge to the

same value:

$$f_{\text{effective}} = \frac{1/\left[\frac{dc}{\tau_{\text{e,on}}} + \frac{1-dc}{\tau_{\text{e,off}}}\right]}{1/\left[\frac{dc}{\tau_{\text{e,on}}} + \frac{1-dc}{\tau_{\text{e,off}}}\right] + 1/\left[\frac{dc}{\tau_{\text{c,on}}} + \frac{1-dc}{\tau_{\text{c,off}}}\right]} \tag{4.11}$$

Since this describes the effective occupancy of the equivalent stationary RTS, it can be seen from eq. 4.5 that this is equal to:

$$f_{\text{effective}} = \frac{\tau_{\text{e,eff}}}{\tau_{\text{e,eff}} + \tau_{\text{c,eff}}} \tag{4.12}$$

It follows that:

$$\tau_{\text{e,eff}} = 1/\left[\frac{dc}{\tau_{\text{e,on}}} + \frac{1-dc}{\tau_{\text{e,off}}}\right] \times \frac{\tau_{\text{c,eff}}}{1/\left[\frac{dc}{\tau_{\text{c,on}}} + \frac{1-dc}{\tau_{\text{c,off}}}\right]} \tag{4.13}$$

In the limit for $dc \to 1$, $\tau_{\text{e,eff}}$ must obviously equal $\tau_{\text{e,on}}$, which is only satisfied in eq. 4.13 if

$$\frac{1}{\tau_{\text{c,eff}}} = \frac{dc}{\tau_{\text{c,on}}} + \frac{1-dc}{\tau_{\text{c,off}}} \tag{4.14}$$

$\tau_{\text{e,eff}}$ is therefore given by

$$\frac{1}{\tau_{\text{e,eff}}} = \frac{dc}{\tau_{\text{e,on}}} + \frac{1-dc}{\tau_{\text{e,off}}} \tag{4.15}$$

Conclusion

A cyclostationary RTS with a constant amplitude and two states, an ‘on’ state from $t = 0 \ldots Tdc$, in which behaviour is governed by $\tau_{\text{e,on}}$ and $\tau_{\text{c,on}}$, and an ‘off’ state from $t = Tdc \ldots T$ in which behaviour is governed by $\tau_{\text{e,off}}$ and $\tau_{\text{c,off}}$, can, in the limit for $T \to 0$ be described by an equivalent stationary RTS with parameters $\tau_{\text{e,eff}}$ and $\tau_{\text{c,eff}}$. The parameters $\tau_{\text{e,eff}}$ and $\tau_{\text{c,eff}}$ are given by equations 4.14 and 4.15.

Modulating a cyclostationary RTS

In actual measurements, the cyclostationary RTS is not visible when the device is ‘off’. This can be modelled mathematically by a modulation with a square wave of period T and width Tdc.

In the limit for $T \to 0$, the occupancy of the RTS is constant, and the cyclostationary RTS is described by an equivalent stationary RTS. This equivalent stationary RTS is modulated by a square wave with a period T and a duty cycle dc, so it will have an LF PSD corresponding to that of the equivalent stationary RTS, but a factor $1/dc^2$ lower due to the modulation of the RTS by the square wave.

Since the equivalent RTS is stationary, it has no phase relation with the excitation signal, and therefore it does not matter what the phase of the modulating signal is. In other words, it does not matter in which part of the period the RTS is visible. This can be understood because the number of RTS transitions in one period of the excitation waveform is 0 in the limit for $T \to 0$.

Noise power of an RTS

The PSD of an RTS is given by eq. 2.27. τ_0 Is replaced by τ_c, τ_1 by τ_e and the amplitude of 1 is replaced by an amplitude ΔI to come to:

$$\beta = \frac{\tau_c}{\tau_e} \tag{4.16}$$

$$\omega_{0RTS} = \frac{1}{\tau_e} + \frac{1}{\tau_c} \qquad [\mathrm{rad/s}] \tag{4.17}$$

$$S_{RTS}(\omega) = 2(\Delta I)^2 \frac{\beta}{(1+\beta)^2} \frac{1}{\omega_{0RTS}} \frac{1}{1+\frac{\omega^2}{\omega_{0RTS}^2}} \qquad [\mathrm{A^2 Hz^{-1}}] \tag{4.18}$$

This will also hold for the stationary 'effective RTS' that results from the limit for $T \to 0$ of a cyclostationary RTS that is periodically and abruptly switched between two states. The PSD behaviour for that case can now be investigated. For simplicity, we will consider the case for $dc = 0.5$.

In fig. 4.3, the RTS parameters β_{on} and $\omega_{0,on}$ are plotted as a function of $\tau_{e,on}$ and $\tau_{c,on}$. The logarithm of $\tau_{e,on}$ and $\tau_{c,on}$ are plotted along the x and y axes respectively, and contour lines for each factor 10 difference in β_{on} and $\omega_{0,on}$ are plotted. In this way, we can easily visualize the RTS parameters of any RTS given $\tau_{e,on}$ and $\tau_{c,on}$. Since τ's have an exponential dependence on the energy E of the trap and the distance x in the oxide, a plot in the $\log \tau$ domain makes sense. A log axis for τ is also convenient for a multiplicative change in τ since it becomes a shift along a $\log \tau$ axis. Making use of eq. 4.1, it is noted that a decade change in τ_e at constant τ_c corresponds to a change in trap energy of approximately 60 meV.

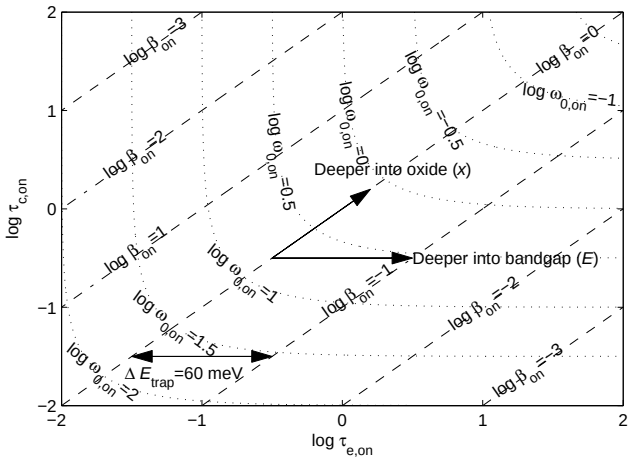


Figure 4.3: Parameters of RTS as a function of $\tau_{\mathrm{e,on}}$ and $\tau_{\mathrm{c,on}}$

To illustrate which RTS are the dominant contributors to the LF noise at the output, the PSD of the RTS at a particular frequency (in this case at $\omega = 1$ rad/s) can be plotted as a function of $\tau_{\mathrm{e,on}}$ and $\tau_{\mathrm{c,on}}$. This is done in fig. 4.4. The contour lines in the plot denote the τ's of traps with a relative noise power of 80, 60, 40 and 20% compared to the dominant RTS which is at the heart of the contours.

It is immediately obvious that the RTS'es with β_{on} and $\omega_{0,\mathrm{on}}$ close to 1 give the dominant noise contribution to the PSD at $\omega = 1$ rad/s. The traps with large β_{on} are mostly empty and do not contribute significantly to the noise. The traps with small β_{on} are mostly full and do not contribute significantly either. If the RTS corner frequency is too low or too high, the contribution of the RTS is insignificant as well. In fact, fig. 4.4 is nothing but a graphical representation of eq. 4.18.

We are now in a position to examine the LSE RTS behaviour. To this end, a simple but effective and insightful model for the bias dependency of τ_{e} and τ_{c} is used:

$$\begin{aligned} \tau_{\mathrm{e,off}} &= \tau_{\mathrm{e,on}} / m_{\mathrm{e}} \\ \tau_{\mathrm{c,off}} &= \tau_{\mathrm{c,on}} \times m_{\mathrm{c}} \end{aligned} \tag{4.19}$$

It might be physically more realistic to model $\tau_{\mathrm{c,off}}$ as ∞, since capture of an electron in a trap is very unlikely when the device is 'off', but for the results

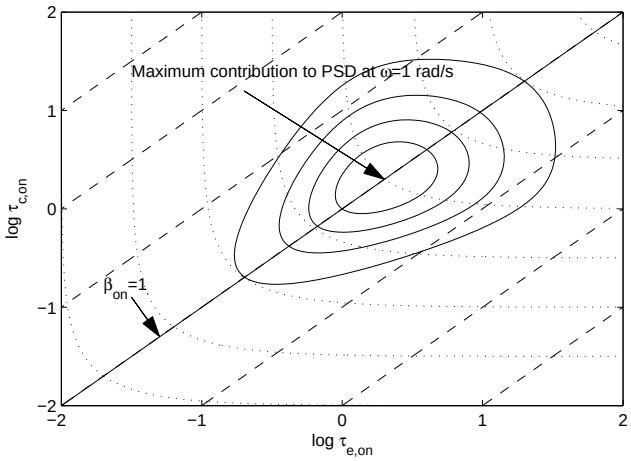


Figure 4.4: Contribution to PSD at $\omega = 1$ rad/s for different RTS'es, steady state

of the model, the difference is not important as the significant effect of subjecting the device to LSE lies in the change in τ_e, not τ_c. From equation 4.19, 4.15 and 4.14, it follows that:

$$\begin{aligned} \tau_{e,eff} &= \frac{2}{m_e + 1}\tau_{e,on} \\ \tau_{c,eff} &= \frac{2m_c}{m_c + 1}\tau_{c,on} \end{aligned} \tag{4.20}$$

This immediately explains why a change in τ_c is not as important as a change in τ_e: even in the limit for $m_c \to \infty$, $\tau_{c,eff}$ only changes by a factor of 2 compared to $\tau_{c,on}$. $\tau_{e,eff}$, on the other hand, can change by several orders of magnitude compared to $\tau_{e,on}$.

The parameters (β, ω_0 and the PSD at 1 rad/s) of the effective RTS for the LSE case can now be plotted. Since the relation between the on-state τ's and the effective τ's (given in eq. 4.20) is known, the plot of the effective RTS parameters can be made as a function of $\tau_{e,on}$ and $\tau_{c,on}$. This is done in fig. 4.5 for $m_e = m_c = 10$.

This figure shows that the main contribution to the PSD is now from different RTS'es. Traps that were mostly full in the steady state case now have an occupancy of 50% and a β of 1, and they dominate the output noise con-

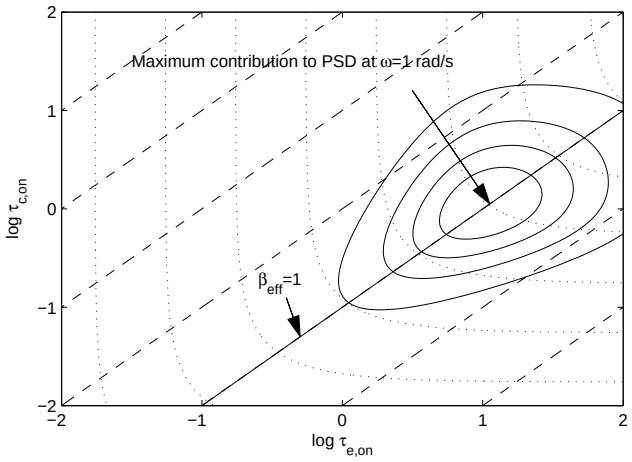


Figure 4.5: Contribution to PSD at PSD at $\omega = 1$ rad/s for different RTS'es, LSE

tribution. In fact, the whole figure has simply been shifted down by a factor $2m_c/(m_c+1)$ and has been shifted to the right by a factor $(m_e+1)/2$.

The question of whether subjecting a device to LSE leads to a decrease in the LF noise PSD as observed in measurements can now be addressed. To this end, the distribution of τ's is examined. This can be the distribution of τ's in a 'large' device with very 'many' traps, but it can also be the distribution of τ's over an ensemble of small devices, each with only a limited number of traps. In the latter case, an individual device will show behaviour depending on the traps it happens to contain, but the ensemble average of the noise performance will be the same as for a single large device with many traps, under the assumption that individual traps make uncorrelated contributions to the output noise power.

Since the shape of the noise contribution curve in fig. 4.4 and 4.5 is the same, and it has only been shifted in the $\log t$ domain, it may be concluded that the noise PSD will not change if the distribution of τ's is uniform in $\log t$ (exponential in t).

A uniform distribution in $\log t$ results if two conditions are satisfied. First of all, the distribution of trap depth in the oxide, x, should be uniform. This is the basis of McWhorter's model [51] and responsible for the emergence

of a $1/f$ spectrum[1]. If, additionally, the trap energy level E is uniformly distributed, a uniform distribution of traps in $\log \tau_e$ and $\log \tau_c$ results.

Since in measurements it is observed that subjecting the device to LSE reduces its LF noise, we conclude that the distribution of time constants is not uniform in the $\log t$ domain.

One explanation for the non-uniformity of the distribution of traps in the $\log t$ domain is that the trap density near the center of the bandgap is lower. This is often reported in literature, eg. [72]. Comparing $\tau_{e,eff}$ to $\tau_{e,on}$, a decrease is noted, and comparing $\tau_{c,eff}$ to $\tau_{c,on}$, an increase is observed (eq. 4.20). This means that the traps nearer the center of the bandgap are the ones contributing LF noise under LSE, and if there are fewer traps, this explains why LSE is observed to lower the LF noise. Like this, LF noise measurements under LSE can be used as a trap density characterization tool.

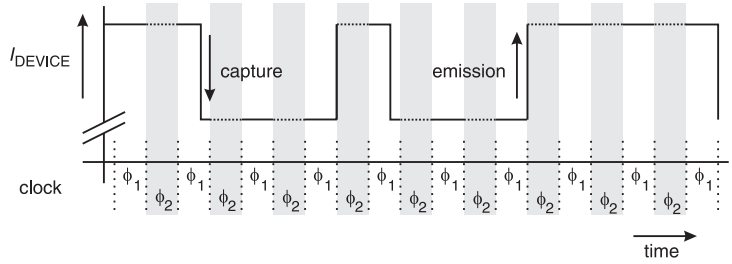
4.4 Cyclostationary RTS noise Simulator

For those cases where the analytical approach does not suffice, a numerical simulator to simulate cyclostationary RTS noise was made. It can be used to simulate cases that are too complex to for the analytical approach and it can also be used to validate the limiting case described by the analytical model.

Principle

The generation of RTS noise in a MOSFET that is operated under cyclostationary bias conditions is simulated. The basic assumption is again that τ_e and τ_c vary instantaneously with V_{GS}. The simulator can easily be used for a variety of non-trivial situations that are not easily tackled analytically. For example, it is not only able to simulate discrete, two-state cyclostationary RTS'es, but can handle any time-dependent RTS, and it can cope with any bias dependency of τ_e and τ_c. Also, it can be used for transient simulation, where the observation interval immediately follows a single bias transient. In this section, a comparatively simple case will be treated in-depth to show that the simulation results are in direct agreement with measurement results: we will examine measurements carried out on $0.2/0.18$ µm

[1] The mechanism by which the electrons from the channel interact with traps is assumed to be direct tunnelling, making the capture cross section a negative exponential function of x.



Clock Phase	ϕ_1	ϕ_2
Gate voltage	V_{GS1}	V_{GS2}
Capture time constant	τ_{con}	$\tau_{con}^{*}m_c$
Emission time constant	τ_{eon}	τ_{eon}/m_e

Figure 4.6: Cyclostationary RTS has periodically varying τ_c and τ_e that directly depend on V_{GS}. This is the case in MOSFETs where the gate bias is periodically cycled between two values V_{GS1} and V_{GS2}.

n-channel MOSFETs [66] that are dominated by RTS noise. In these measurements, the gate bias V_{GS} is periodically switched between two states (an 'on' state ϕ_1 and an 'off' state ϕ_2) with a 50% duty cycle. Both the drain-source voltage V_{DS} and the bulk-source voltage V_{BS} are kept constant. The two bias states that are periodically alternated between correspond to two different gate voltages and hence, different τ_c and τ_e for the RTS noise generating traps in the device. This is shown in fig. 4.6; V_{GS1} corresponds to the 'on' state of the device, and V_{GS2} corresponds to the 'off' state of the device. The steady state RTS time constants are $\tau_{c,on}$ and $\tau_{e,on}$ and the 'off' state time constants are modelled as $\tau_{c,off} = \tau_{c,on} \times m_c$ and $\tau_{e,off} = \tau_{e,on}/m_e$ where m_c and m_e are parameters that are to be determined. When m_c and m_e are unequal to 1, the parameters of the RTS vary periodically, and the RTS is therefore cyclostationary.

Obviously, when cycling the gate bias, the drain current I_D of the device also changes, and with it, the visibility of the trap at the terminals of the device: When the device is 'on', the effect of the trap is visible as a fluctuation in I_D and when the device is 'off', I_D is negligible and the behaviour of the trap is not visible in the drain current. This is shown in fig. 4.7a: The cyclostationary RTS is modulated by a square wave. The drain current of the device, I_D, (fig. 4.7a), can be considered as a superposition of two sig-

nals: a large square wave with an amplitude of I_L (deterministic; fig. 4.7b), and a small modulated cyclostationary RTS with an amplitude $I_H - I_L$. (stochastic; fig. 4.7c). The crucial point is that the RTS continues to 'exist' even at times when it cannot be directly observed! The deterministic current component of fig. 4.7b contributes a series of δ-functions in the output power spectrum. (At 0 Hz, $\pm f_{\text{switch}}$, $\pm 3 f_{\text{switch}}$, $\pm 5 f_{\text{switch}}$, etc...) In measurements, this current component is suppressed as much as possible, and it is therefore not simulated. The stochastic current component of fig. 4.7c, however, is interesting: it is the modulated, cyclostationary RTS noise that is responsible for the LF noise that is observed. In the frequency domain, it contributes a series of aliases around harmonics (again: around 0 Hz, $\pm f_{\text{switch}}$, $\pm 3 f_{\text{switch}}$, $\pm 5 f_{\text{switch}}$, etc...) of the modulating frequency. In measurements, the baseband alias of this signal is measured using a spectrum analyser.

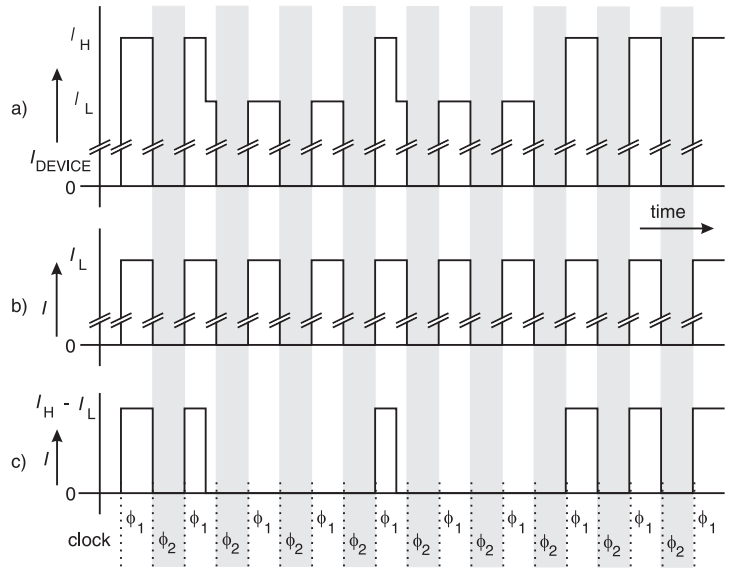
The simulator produces the signal of fig. 4.7c. This is done by first generating a cyclostationary RTS. This models the internal stochastic process; using a single factor $m = m_c = m_e$ to model the change in τ_c and τ_e produces excellent agreement with measurement results. The cyclostationary RTS is then modulated by a square wave. This models the effect of switching on the visibility of the RTS. The modulated DC drain current (fig. 4.7b) is not simulated as it is uninteresting.

This method of simulation is more subtle than the method suggested by Tian and El Gamal [62]. Their calculation is based on the assumption that when the MOSFET is turned 'off', $\tau_c \to \infty$ and $\tau_e \to 0$; or in terms of our model: $m \to \infty$. They correctly note that their method underestimates the noise coming from the device in actual measurements. In contrast, our model can account for the very variable noise reduction that is observed in different small-area devices, and it can explain that the LF noise of some devices increases when they are subjected to large signal excitation.

Implementation

The simulator is implemented in MATLAB. In general, the transition probabilities per unit time are defined as [48]:

$$\begin{aligned} P_{\text{capture}}(V_{GS}(t)) &= \frac{dt}{\tau_c(V_{GS}(t))} \\ P_{\text{emission}}(V_{GS}(t)) &= \frac{dt}{\tau_e(V_{GS}(t))} \end{aligned} \tag{4.21}$$



Clock Phase	ϕ_1	ϕ_2
Gate voltage	V_{GS1}	V_{GS2}
Capture time constant	τ_{con}	$\tau_{con} * m_c$
Emission time constant	τ_{eon}	τ_{eon} / m_e
Transfer function to output node	1	0

Figure 4.7: Cyclostationary, modulated RTS. (a) Shows the total drain current of the device that is cycled between V_{GS1} and V_{GS2}. (b) And (c) isolate the deterministic and stochastic components of this current, respectively.

Since the simulator is time-discrete, (with a constant sample time t_{sample}), t_{sample} needs to be substituted in place of dt:

$$P_{\text{capture}}(V_{\text{GS}}(t)) = \frac{t_{\text{sample}}}{\tau_{\text{c}}(V_{\text{GS}}(t))} \qquad [\text{sample}^{-1}]$$
$$P_{\text{emission}}(V_{\text{GS}}(t)) = \frac{t_{\text{sample}}}{\tau_{\text{e}}(V_{\text{GS}}(t))} \qquad [\text{sample}^{-1}] \tag{4.22}$$

Due care is taken that the time-discrete nature of the simulation does not introduce significant errors: P_{capture} and $P_{\text{emission}} \ll 1$ [sample^{-1}]. For each simulation run a large number of transitions is generated to produce statistically significant results. For consistency with the measurements, the lower frequency is chosen to be 10 Hz, and the switching frequency is 10 kHz. The sample rate of the simulation is an order of magnitude larger than the highest frequency of interest. To produce a smoother plot, the simulation is carried out a number of times with different seeds for the random generator and the results are averaged.

Validation

The effect of various parameters on the spectrum of the cyclostationary RTS can now be explored. Some simulation results are plotted in fig. 4.8: 'A' is the PSD of a stationary RTS for $\beta = 1$. 'B' shows the effect of *only modulating* this RTS ($m = m_{\text{c}} = m_{\text{e}} = 1$). As predicted by basic modulation theory [61], the PSD below the switching frequency has decreased by a factor 4 (6 dB). Apart from the clearly visible LF spectrum, an alias of this spectrum is visible around the switching frequency and its odd multiples, as expected. The modulated RTS is now made cyclostationary: m_{c} and $m_{\text{e}} \neq 1$. Curve 'C' shows the PSD of the cyclostationary RTS with $m_{\text{c}} = m_{\text{e}} = 10$. Curve 'D' shows the PSD for $m_{\text{c}} = m_{\text{e}} = 50$ and curve 'E' shows the PSD when m_{c} and $m_{\text{e}} \to \infty$. This is the limiting case, where in the 'off' state $\tau_{\text{c}} \to 0$ and $\tau_{\text{e}} \to \infty$. Curve 'E' has the lowest LF PSD. It corresponds to the method of [19] and [62]. For comparison, the solid lines in the figure are the results from the analytical expression which are seen to be in excellent agreement with the simulator output. Direct comparison is possible because this is a simple case. The analytical expression for curve 'E' is not plotted in the figure, as the requirement that the switching frequency be 'very high' is not adequately satisfied for this case.

Our model and our simulator are able to explain the variable noise reduction that is seen in measurements. They can also account for increased LF noise

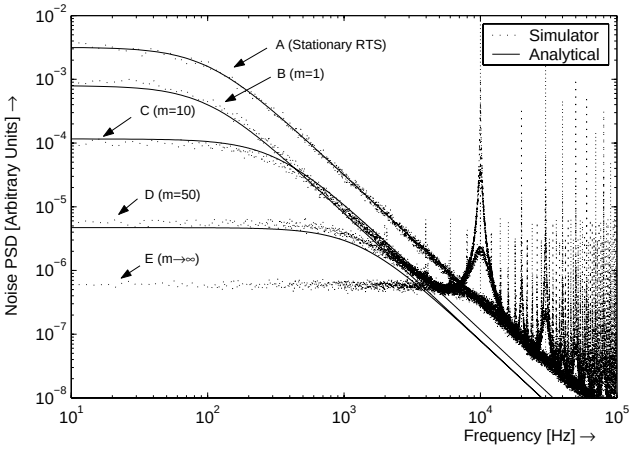


Figure 4.8: Simulation highlights the difference between simply modulating an RTS and making it cyclostationary. Modulating the RTS gives a 6 dB decrease in LF noise PSD; making it cyclostationary gives a far larger (and variable) decrease.

in a MOSFET under large signal excitation. To illustrate this, a simulation is performed with an asymmetric RTS; $\beta = 0.02$. The results are shown in fig. 4.9. The stationary RTS looks much like any other stationary RTS, and the modulated RTS ($m_c = m_e = 1$) brings no surprises either. The interesting curve here is the one where $m_c = m_e = 10$, for which the LF PSD is seen to rise *above* the PSD of the stationary RTS. For much larger values of m_c and m_e, the LF PSD drops once more.

In the next section, we will show that the simulator can reproduce the measurement results by fitting β, m_c and m_e to the measurement results. Before that, however, the measurement results will be examined in more detail.

Comparison between simulator and measurement results

Our cyclostationary RTS generation model can reproduce a number of measurement results. First of all, a measurement with a varying β is carried out by varying the V_{GS} of a single device, and secondly a number of randomly selected devices from the same wafer, exhibiting widely varying noise performance, are measured. Noise measurements are carried out on minimal-

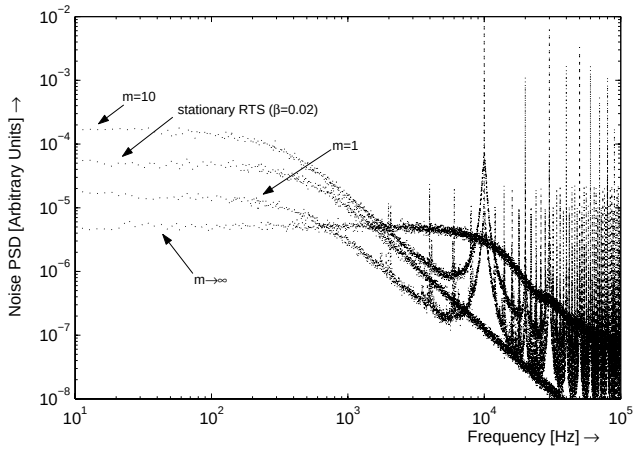


Figure 4.9: Generation of cyclostationary RTS starting with asymmetric RTS. An asymmetric RTS can give an LF noise increase when the RTS is made cyclostationary.

size n-channel MOSFETs from a 0.18 μm process. Device size (W/L) was $0.2/0.18$ μm.

- Steady state noise measurements are performed by applying a constant bias voltage to the gate.
- Large signal noise measurements are carried out by subjecting the device to a 50% duty cycle square wave; i.e. 50% of the time 'on', (V_{GS} same as in the steady state case),and 50% of the time 'off' ($V_{GS} = 0$ V). The switching frequency is 10 kHz.

For details of the measurement, the reader is referred to chapter 3 or references [65] and [66].

Measurement with variable V_{GS}

First of all, a measurement was carried out on a single device that was selected for having a very visible RTS. V_{GS} Of this device was stepped through three different values:

- One where the trap was observed to be approximately half-filled ($\tau_c \approx \tau_e$, figure 4.10 b),
- one where it was observed to be mostly full ($\beta \approx 0.3$, figure 4.10 a),

- and one where it was observed to be mostly empty ($\beta \approx 10$, figure 4.10 c).

For each V_{GS}, a steady state bias and a large signal measurement was carried out. For all measurements, the RTS amplitude at the drain was observed to remain constant in the steady state case as well as in the large signal case. Hence, the difference in the amplitude of the measured spectra is attributed to differences in the time domain behaviour of the RTS.

The simulator can correctly model this measurement. First of all, the steady state simulations are fitted to the steady state measurement results. Correspondence between simulation and measurement is achieved when, in the simulator, f_{0RTS} is set to 1.5 kHz and β is chosen as 0.3, 1 and 10 for measurements (a), (b) and (c) respectively. The simulated steady state results are shown as '$\circ$' in the figure. To fit the model to the LSE results, $m_c = m_e = 12$ was chosen. The simulated LSE results are shown as '$\triangle$' in the figure. As can be seen, the steady state spectra are not as sensitive to variation in β as the LSE spectra. Thus, after fitting the steady state results with varying β's, a single assumption ($m = 12$) is adequate to coarsely model all three LSE results. The only difference between the simulation results for fig. 4.10 a, b, and c is in the β of the RTS that is being modelled. In the measurements, the spectrum is observed to rise again at low frequencies. This is due to additional slow traps that are not modelled in the simulator.

Measurement of 21 different devices

Next, 21 nominally identical devices, selected randomly from the whole wafer were measured. These devices were found to exhibit widely varying RTS noise, in terms of amplitude, f_{0RTS} and β.

The measurement results are plotted in fig. 4.11, which shows steady state noise along the *x*-axis versus the LSE noise *reduction* along the *y*-axis. The steady state noise (*x*-axis) is expressed as average gate-referred noise $[\mathrm{V\ Hz^{-1/2}}]$ in a measurement band from 10 to 150 Hz. The net LSE noise reduction (*y*-axis) is measured by subjecting the device to large signal conditions. In this state, the average LF noise power is again measured from 10 to 150 Hz. The difference between the steady state noise measurement and the LSE noise measurement, corrected for the modulation effect (6 dB), is the noise reduction[2] that is plotted along the *y*-axis.

[2] this is comparable to the LF difference between curve 'B' and curve 'C', 'D' or 'E' in fig. 4.8.

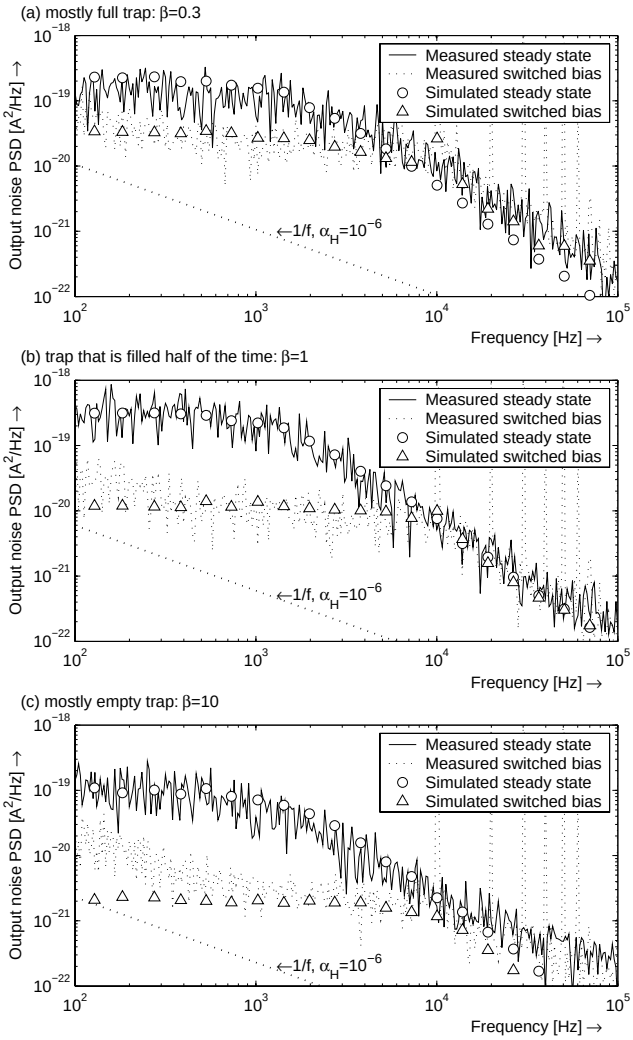


Figure 4.10: Different types of RTS noise in a single MOSFET can also be explained very well using the cyclostationary RTS noise model.

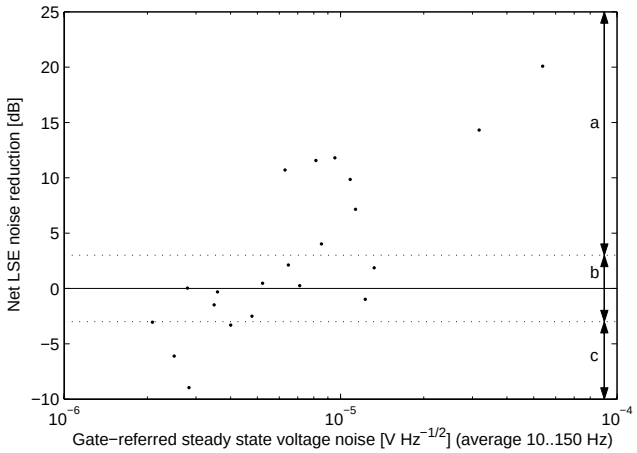


Figure 4.11: Net LSE noise reduction vs. steady state noise power for randomly selected devices: Noise reduction varies and in some cases noise increases.

In fig. 4.11, three different classes of devices, ‘a’, ‘b’ and ‘c’ can be identified. The LF noise properties of these different devices are summarized in table 4.1. In figures 4.12-4.14, the noise spectrum of one device from each category is examined, and in each case, a stationary RTS noise simulation and a corresponding cyclostationary RTS noise simulation is shown to come into close qualitative agreement with the measured data for the steady state noise measurement and the large signal noise measurement respectively. The parameters of the RTS noise simulation for each case are given in table 4.1.

In each of the figures 4.12-4.14, a calculated line for the bulk silicon $1/f$ noise is shown, with α_{H} chosen as 10^{-6}. This is a realistic value for bulk Si [11], and represents the level of bulk $1/f$ noise expected in the measurement. It can be seen that the RTS noise level is much higher than the bulk $1/f$ noise. In each figure, the output-referred noise PSD is plotted. This can be referred to the gate by dividing by the g_{m}^2 of the device ($g_{\mathrm{m}} = 130\ \mu\mathrm{A/V}$). In fig. 4.13 and 4.14, the measured LSE spectrum exhibits a rise at low frequencies. As in fig. 4.10, this is probably due to additional slow traps in the device or bulk $1/f$ noise and is not modelled. In fig. 4.14, both the LSE measurement result and the simulation show a rise in the PSD close to the switching frequency. This is the first alias of the

Device Category	'a'	'b'	'c'
LF noise under LSE	Decrease >3 dB	No change	Increase >3 dB
Location in fig. 4.11	Upper right	Middle	Lower left
RTS simulation: f_o [Hz]	150	5 k	1.5 k
RTS simulation: β	1	0.3	0.02
RTS simulation: m_c $(= m_e)$	50	10	50
Measurement and simulation shown in fig.	4.12	4.13	4.14

Table 4.1: Devices of figure 4.11 can be grouped into three categories

noise spectrum around the switching frequency, which is clearly visible in fig. 4.8 as well.

Simulating transient noise

The simulator can also be used to simulate noise under transient bias conditions. By this the noise briefly after turn-on of a device is meant. For example, in a sampled system, a device may be off for a relatively long time, and then be turned on and immediately used to take a sample of the input signal. If the time between turn-on and the sample moment is short relative to the time constants of the device noise, then a steady state approach to noise is obviously not appropriate. The cyclostationary limit for high frequencies treated analytically in section 4.3 is also not applicable in a case like this. The simulator, however, can aid in describing such noise behaviour.

As an example, consider the noise measurement in fig. 4.15 (a) and (b). This measurement was carried out on a $0.5/0.35$ μm n-channel MOSFET from a 0.35 μm process.

Fig. 4.15 (a) shows the steady state noise measurement. Time is plotted along the *x*-axis, and the trial# is plotted along the *y*-axis. The device noise

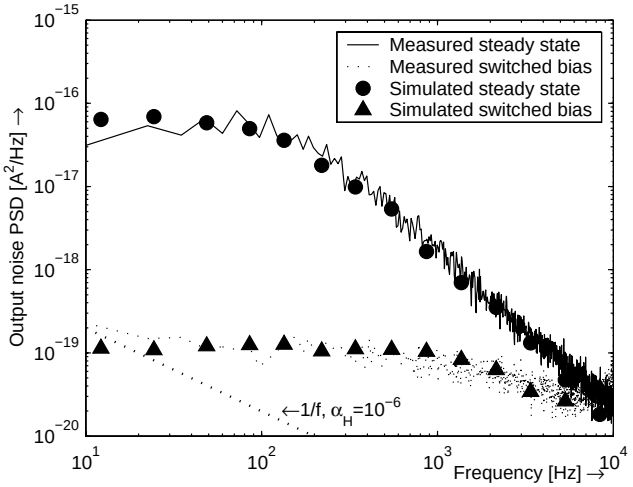


Figure 4.12: Measurement and simulation result showing strong decrease in LF noise under LSE

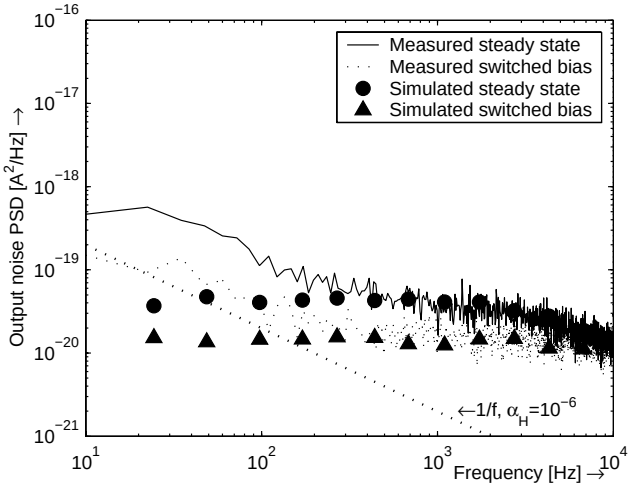


Figure 4.13: Measurement and simulation result showing no change in LF noise PSD when the device is subjected to LSE

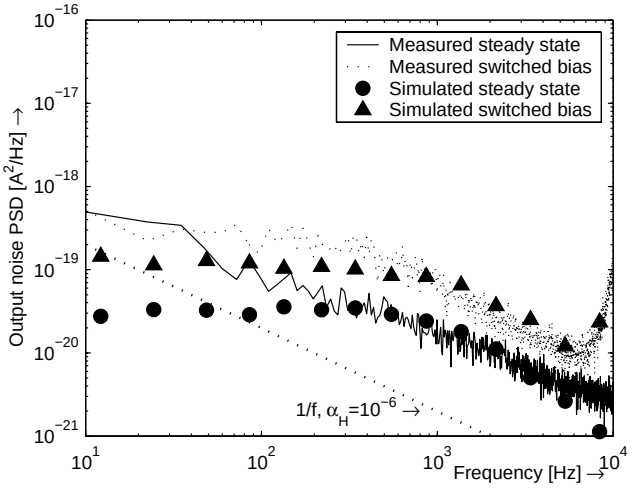


Figure 4.14: Measurement and simulation result showing an increase in LF noise when the device is subjected to LSE

is dominated by an RTS; the figure is colour coded so that one state of the RTS is white and the other state of the RTS is black. The RTS in question is clearly strongly asymmetric.

Fig. 4.15 (b) shows the noise of the same device, but now immediately after turn-on of the device: immediately preceding the measurement interval of 4 μs the device was 'off' for 5 ms. It is clear that the trap starts in the 'black' state immediately after turn-on, and after some time relaxes into the 'white' state, after which steady state behaviour is observed. Since the 'black' state is the one the trap is found in immediately after turn-on, it may be associated with an empty trap, and the 'white' state with a full trap.

The simulator is able to reproduce this steady state and transient behaviour by suitable selection of $\tau_{\mathrm{c,on}}$ (6×10^{-7} s) and $\tau_{\mathrm{e,on}}$ (4×10^{-5} s). $\tau_{\mathrm{e,off}}$ is made very small to ensure that the trap is mostly empty at the instant of turn-on, just as in the measurement. The simulation results are plotted in fig. 4.15 (c) and (d).

It is clear that the simulation results are very similar to the measurement results, but further validation is desired. To this end, the average of the noise voltage over all trials is plotted as a function of time, as is the variance of the noise voltage over all trials as a function of time. This is done in fig. 4.16.

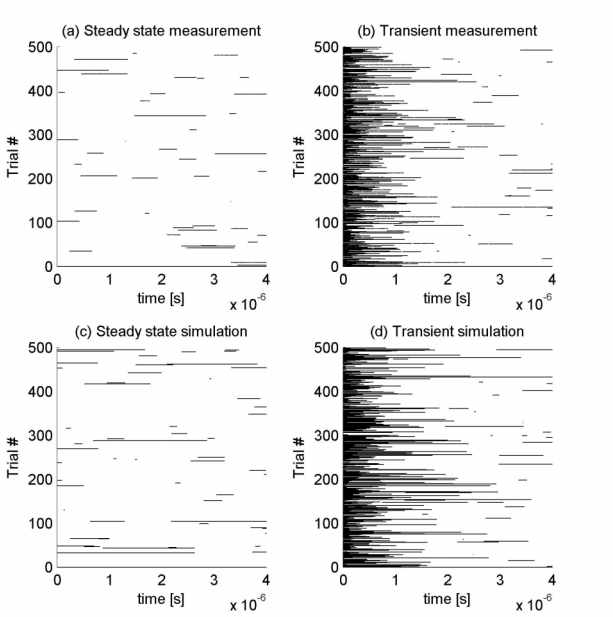


Figure 4.15: Comparison between measurement (a and b) and simulation (c and d) of an RTS under steady state (a and c) and transient (b and d) conditions.

Once again, it is clear that the simulation results are in close agreement with the measurement results.

In fig. 4.16 (a), it is seen that for a steady state RTS, the average noise voltage is constant. The steady state variance (c) of the noise is also constant, though the strongly asymmetric nature of this particular RTS means that the variance is low, and the curve looks noisy. In fig. 4.16 (b), the average noise voltage immediately following turn-on of the device is shown, and in (d), the variance of the noise immediately following the bias transient is shown. The variance exhibits a very characteristic 'hump'; at $t = 0$ the variance is low because the trap is mostly empty, at $t = 4$ μs the variance is low because the trap is mostly full, and in between, around $t = 0.5$ μs after turn-on, the variance is at a maximum due to the large number of transitions occurring at that time.

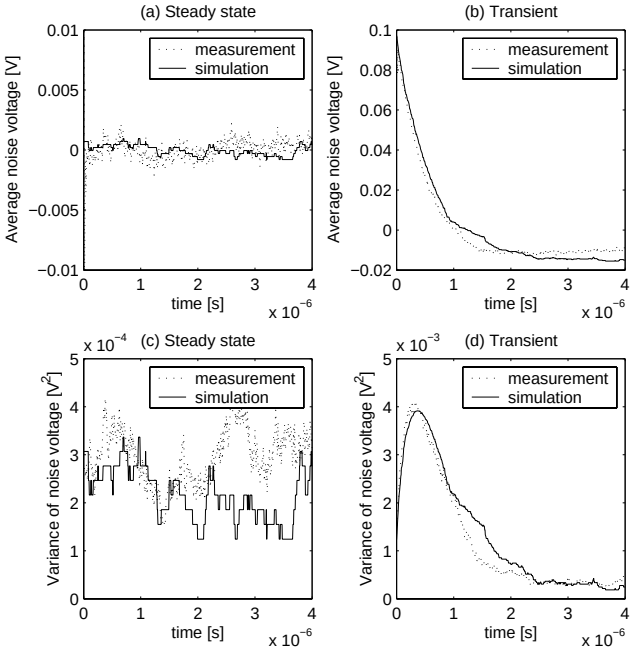


Figure 4.16: Comparison between measurement and simulation of average noise voltage and variance of noise voltage in steady state and after biasing transient

Summary

When a large signal bias voltage is applied to a transistor in which noise behaviour is dominated by the effect of a single trap, several different types of behaviour are seen.

In one experiment, different β's are generated in a single MOSFET by varying V_{GS}. The variation in β has an influence on the steady state noise spectrum, and a strong effect on the LSE noise spectrum. In the simulation, the same β as observed in the experiments is used, after which only the appropriate m needs to be chosen to fit the model to the experimental results.

In another experiment, different β's are found in different nominally identical devices on the same wafer. In the same way as in the first experiment,

cyclostationary RTS noise simulation is able to model the measurement results. Depending on the parameter β of the trap in question, the LF noise of the device is seen to go down, go up or not change significantly. The two extreme possibilities are:

- If β is close to 1 in the steady state, (upper righthand corner of fig. 4.11) then subjecting the device to LSE will strongly reduce the noise coming from the device.
- If β is much smaller than 1 in the steady state, (lower left-hand corner of fig. 4.11) large signal excitation can cause an increase in the noise coming from the device. Such a device contains a trap that is mostly filled in the steady state, and applying LSE to the device will cause it to be periodically emptied, thereby increasing its contribution to the LF noise PSD. The traps with β deviating far from 1 are those with a rather low steady state LF noise PSD. The simulation corresponds to the measurements in this respect.

The strength of the simulation presented here is that a wide range of observations can be explained using a single, physically realistic assumption, namely that the τ_c and τ_e of a trap in a MOSFET are bias dependent. In modelling the experimental results, in each case it was assumed that $m_c = m_e$.

For demonstration of the model, the reasonable assumption was made that $m_c = m_e$ which works well but for which there is no strong physical basis. The power of the model, however, is that it can handle *any* bias dependency of τ_e and τ_c. For example, in [40] and [38], a more detailed characterization of τ_e and τ_c as a function of V_{GS} is given. Such characterization effort can easily be incorporated.

In summary:

- The model explains why the LF noise of a device can go up, remain the same or decrease.
- If the additional assumption of a lower density of traps in the center of the bandgap is made, the model explains why the LF noise of large devices containing 'many' traps decreases when they are subjected to LSE.
- When a distribution of traps is mentioned, this can equivalently refer to a distribution of traps in a large device or to a distribution of traps over many small devices each containing only a few traps.
- In the limit for high enough switching frequencies, the switching frequency is not important for the LF noise of the device. $\tau_{e,eff}$ And $\tau_{c,eff}$ are not a function of frequency.

4.5 Conclusion

In this chapter, Shockley-Read-Hall theory in conjunction with a bias dependent trap energy level is used to model RTS noise in MOSFETs under nonstationary conditions. In this way, the various experimental results were modelled.

The simple case where the device is periodically and abruptly switched between two states, and where the switching frequency is made very high relative to the RTS corner frequency is treated analytically. Under the assumptions of this analysis, a cyclostationary RTS can be modelled by an equivalent stationary RTS. It is shown that under large signal conditions device noise is dominated by different traps compared to the steady state situation. A non-exponential distribution of trap time constants (either in a single large device or over an ensemble of small devices) can lead to a reduction in LF noise when the device is subjected to large signal excitation. For example, if the trap density near the center of the bandgap is lower than near the conduction band edge, subjecting the device to LSE will lower its LF noise.

To analyse more complex situations, a numerical simulator is presented that generates cyclostationary random telegraph signals. This simulator is useful in more complex situations, for example where the bias of the device is varied in a continuous manner or where we are interested in device noise briefly following a single bias transient. The simulator produces results that are in excellent agreement with a wide variety of measurement results.

The predominant limitation of the modelling work is that we do not, in general, know what the distribution of traps in a device is. This holds for both the distribution of the energy level of the trap and for the distribution of the distance of the trap from the channel in the oxide. Alternatively, LF noise measurement under LSE may be used as a characterization method for the trap distribution in energy. This can provide independent validation of other trap characterization methods.

Chapter 5

Perspective on application

5.1 Introduction

In this chapter LF noise is treated from a designer's point of view. A review of literature is presented, as is a discussion in what ways scaling can be used to improve the LF noise performance of circuits. Next, we will discuss how LF noise reduction by large signal excitation (LSE) can be applied to analog CMOS circuits. Rather than discussing every possible application of LSE in detail, three classes of circuits will be examined. In the first, time continuous circuits, a transistor is periodically switched out of the circuit, and replaced by another. This technique can be applied to single devices or to whole (sub) circuits. The second class of circuits are sampled data circuits. The time-discrete nature of these circuits allows, in principle, the use of LSE to reduce LF device noise when the devices are not used. An example of this type of circuit is an imager pixel in a CMOS image sensor, which will be covered in some detail. The third class of circuits are circuits where large signal excitation is present by virtue of the operation of the circuit. Examples of this type of circuit include VCO's and mixers. In these circuits, it is sometimes possible to incorporate LSE against LF noise by ensuring that devices in the circuit experience bias excursions to well below threshold. Next, some novel uses of LF noise under LSE will be briefly discussed. Finally, we will review LF noise modelling work and how this impacts the application of LF noise under LSE.

Review of literature

In literature, some applications of LF noise reduction by LSE have appeared. Unfortunately, much of the work is lacking in important ways. Many authors mistakenly believe that the LF noise reduction through LSE is modelled in current circuit simulators, and therefore mistake device-physical effects of LSE with circuit effects of LSE. Not many authors present measurement results of LF noise under LSE, and none present any measurements of spread of the effect on a group of devices.

- Several authors mistakenly believe the change in LF noise by LSE can be simulated, which is, to the best of my knowledge, incorrect at this time [3, 8].
- The distinction between a change in upconversion factor and a change in LF noise level is not made [3, 8].
- It is not always clear that LF noise limits the performance of the application in question [43].
- Noise measurements are not always given [8, 43].
- A comparison between simulation and measurement results is not always made [43].
- Measurements are given but only from a single device [73].

Scaling to combat $1/f$ noise

Faced with a circuit that does not meet the required performance due to LF noise, a designer can scale the circuit up to improve performance. The scaling rules given here can always be applied to the whole circuit, but the conclusions are equally valid if only part of the circuit is scaled up. The resources that a designer can expend to achieve his goals are power (P) and area (A), and the goals he can achieve are a higher SNR_{th} (directly proportional to P) and a higher $SNR_{1/\text{f}}$ (directly proportional to A). Scaling both A and P up at the same time is commonly known as W-scaling or conductance-scaling [37].

To give an overview of scaling possibilities, the simple quadratic MOSFET model is applied, i.e. short-channel effects are ignored. Keeping V_{GT} constant during scaling is a reasonable approach since almost all parameters of the circuit depend on V_{GT} in a non-linear fashion. It makes sense, therefore, to first optimize the performance of the circuit, and once all performance criteria are met, the designer can scale the circuit to meet noise requirements without changing V_{GT}. Specifying a specific $1/f$ noise model is not

Original	Increase only A	Increase only P	Increase A and P.
W	$n^{1/2}W$	$n^{1/2}W$	nW
L	$n^{1/2}L$	$n^{-1/2}L$	L
$WL(\propto A, C_{\text{in}})$	nWL	WL	nWL
$W/L(\propto P, I_{\text{signal}}, g_{\text{m}})$	W/L	nW/L	nW/L
V_{GT}	V_{GT}		
P_{signal}	P_{signal}	n^2P_{signal}	n^2P_{signal}
$S_{V_{\text{G}}\,\text{signal}}$	$S_{V_{\text{G}}\,\text{signal}}$		
$S_{V_{\text{G}}\,\text{th}}$	$S_{V_{\text{G}}\,\text{th}}$	$n^{-1}S_{V_{\text{G}}\,\text{th}}$	$n^{-1}S_{V_{\text{G}}\,\text{th}}$
$S_{V_{\text{G}}\,1/\text{f}}$	$n^{-1}S_{V_{\text{G}}\,1/\text{f}}$	$S_{V_{\text{G}}\,1/\text{f}}$	$n^{-1}S_{V_{\text{G}}\,1/\text{f}}$
SNR_{th}	SNR_{th}	$n\,SNR_{\text{th}}$	$n\,SNR_{\text{th}}$
$SNR_{1/\text{f}}$	$n\,SNR_{1/\text{f}}$	$SNR_{1/\text{f}}$	$n\,SNR_{1/\text{f}}$
BW $(\propto g_{\text{m}}/C_{\text{in}})$	n^{-1}BW	nBW	BW

Table 5.1: Scaling by a factor n to reduce thermal and LF noise

required since all $1/f$ noise models have the same geometry dependence. The scaling overview is given in table 5.1

The bandwidth of the circuit is assumed to be proportional to $g_{\text{m}}/C_{\text{in}}$. Since g_{m} scales with P and C_{in} scales with A, the bandwidth of the circuit scales with P/A. Hence, the required bandwidth fixes the ratio P/A. The only remaining liberty the circuit designer has is increasing A and P proportionally until the noise targets are met.

5.2 Time-continuous circuits

Introduction

One rather simple way of combating LF noise is to simply duplicate an entire circuit a number of times. If switches and a multi phase clock source are added, the circuit of fig. 5.1 results.

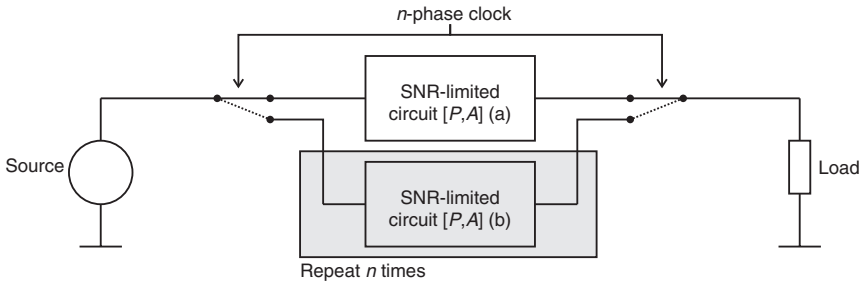


Figure 5.1: Multipath scaling to combat noise

The principle of operation of the circuit in fig. 5.1 is that at any instant in time, one of the available circuits is active, and the others are turned off. The inactive circuits do not draw any power when they are off and noise-limiting devices in the inactive circuits are (or can easily be) subjected to a large signal (eg. V_{GS} is made '0'), in order to reduce their LF noise in the subsequent phase, when they are active. This schematic is representative of a large class of circuits, and represents either a complete circuit or a noise-limited part of a larger circuit. In the analysis, it is assumed that the addition of switches before and after the circuit in question is feasible and without significant penalty in terms of area and power spent. This is not always the case, of course, but is nevertheless a useful assumption that enables a baseline comparison of different noise reduction strategies.

Comparison between multipath and area scaling

Since the multipath topology spends die area and not power, the most appropriate comparison is to area scaling as described in table 5.1. For this, the systematic effect on the noise performance of the multipath topology needs to be considered first. The comparison will be made under the conditions that the original circuit is simply duplicated twice.

Thermal noise

In terms of thermal noise, the performance of the circuit has not changed in any significant way. During the first half of the period, circuit (a) is contributing thermal noise to the output, and during the second half of the period, circuit (b) is contributing thermal noise to the output. The PSD of

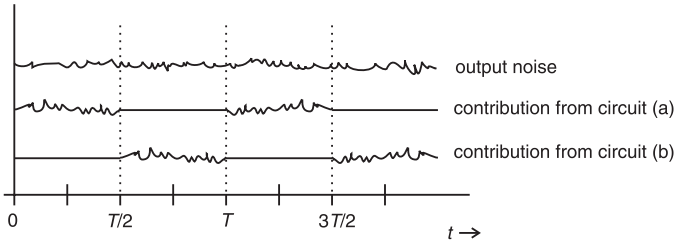


Figure 5.2: Noise performance of two-path system

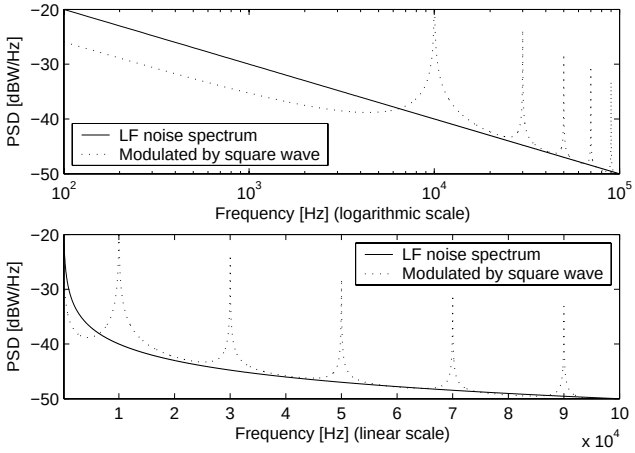


Figure 5.3: Noise PSD of two-path system

thermal noise at the output is at all times equal to the PSD of thermal noise of one of the two circuits. Since circuits (a) and (b) are nominally identical, the thermal noise performance of the multipath topology is the same as that of a single circuit.

LF noise

The output noise of the multipath topology is given in fig. 5.2. It can be seen as the addition of two LF noise sources, that of circuit (a), modulated by a square wave, and that of circuit (b), modulated by a phase shifted version of the same square wave.

In the frequency domain (fig. 5.3), the noise contribution of circuit (a) consists of aliases of the original LF noise spectrum around DC and odd multiples of the switching frequency. Following the derivation of section 3.2, the relative power of the aliases is $1/4$ for the baseband alias (-6 dB), and $2/(n\pi)^2$ for the higher order odd aliases (-6.9, -16.5 and -20.9 dB for the first, third and fifth harmonic respectively). Summing the power of the baseband and all odd aliases, exactly half the power of of the original LF noise spectrum (-3 dB) results, as is immediately obvious from the time-domain view of fig. 5.2 where it is seen that one circuit is active for half of the total time.

Since the output noise contribution of circuit (b) is not correlated to the output noise contribution of circuit (a), the PSD's of both noise contributions may simply be added. This means that at the output, we get a relative amplitude of the aliases of the LF noise of -3.0 dB around DC, and -3.9 dB, -13.5 dB, and -17.9 dB around the first, third and fifth harmonic, respectively. Summing the power of the baseband and all odd aliases, the result is twice the noise power contribution of a single circuit, i.e. exactly the same total noise power as the original LF noise spectrum.

The performance of the multi-path topology can now be compared to the area-scaled version. This is done in table 5.2 where increasing the area of a circuit n times is compared to duplicating the circuit n times and periodically switching between them. A distinction is made between simply switching between the different circuits ('Multipath scaling, unfiltered'), and including a lowpass filter at the output. ('Multipath scaling, lowpass filtered').

Considering only the LF part of the output noise PSD, the multipath topology works as well as area scaling. Area scaling has the disadvantage that the bandwidth of the circuit is reduced, whereas the multipath topology has the disadvantage that the higher aliases of the LF noise need to be removed by filtering at the output. This effectively limits the bandwidth of the multipath topology as well.

There is one other point in the comparison between area scaling and multipath scaling, and that is that if the two circuits in the multipath topology are purposely made different, then the multipath topology opens the possibility of tuning the parameters of the circuit by varying the duty cycle of the switching waveform. An example of this approach is treated below.

In conclusion, area scaling and multipath scaling will attain the same noise performance to a first order approximation, and area scaling will do so wit-

Original	Area scaling	Multipath scaling	
		unfiltered	lowpass filtered
A	nA	nA, switches	nA, switches, filter
P	P		
$S_{V_{\mathrm{G}}\,\mathrm{sig}}$	$S_{V_{\mathrm{G}}\,\mathrm{sig}}$		
$S_{V_{\mathrm{G}}\,\mathrm{th}}$	$S_{V_{\mathrm{G}}\,\mathrm{th}}$		
$S_{V_{\mathrm{G}}\,1/\mathrm{f}}$	$n^{-1}S_{V_{\mathrm{G}}\,1/\mathrm{f}}$	$S_{V_{\mathrm{G}}\,1/\mathrm{f}}$	$n^{-1}S_{V_{\mathrm{G}}\,1/\mathrm{f}}$
SNR_{th}	SNR_{th}		
$SNR_{1/\mathrm{f}}$	$nSNR_{1/\mathrm{f}}$	$SNR_{1/\mathrm{f}}$	$nSNR_{1/\mathrm{f}}$

Table 5.2: Comparison between area scaling and multipath scaling

hout the added complexity of a clock source, switches and a lowpass filter. On the other hand, multipath scaling allows the LF noise of the circuit to be reduced by subjecting the devices to LSE, and enables tuning of circuit parameters if the two circuits are not identical. How the benefits of LF noise reduction by LSE, available in the multipath topology, weigh against the implementation problems (switches, a filter and a clock source) will depend to a large degree on the LF noise reduction by LSE available in a particular technology as given in table 3.3.

Comparison between multipath scaling and a chopper structure

In comparison to area scaling, the multipath topology may seem to be a competitive approach, the only significant disadvantages being the added switches, clock source and lowpass filter. However, in situations where such additions are feasible there is often a much better approach to combat LF noise, namely making a chopper structure.

In order to implement a chopper structure, the circuit first needs to be made differential.

According to table 5.1, each half of the differential circuit can be made using half the area and half the power of the original circuit, so that the complete differential circuit occupies the same area and draws the same power as the original circuit. Each half will have the same bandwidth, but

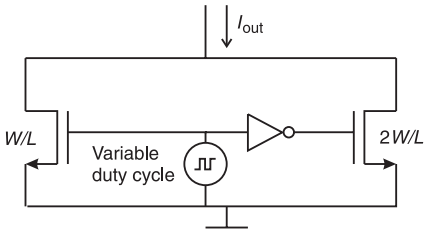


Figure 5.4: Tunable current source with LF noise reduction due to LSE

half the SNR of the original circuit. Since the output is defined as the difference between the outputs of the two signal paths, and the noise power of the two halves is uncorrelated, the SNR of the differential structure is equal to that of the original circuit. So, to a first order approximation, making a circuit differential is without penalty in terms of area and power spent. Next, a clock source, switches and a lowpass filter are required, just as for the multipath topology.

Detailed analysis of the chopper structure [15] shows that a chopper structure removes the $1/f$ noise completely, without increasing the thermal noise level. This is often a much better approach, since all the desired benefits are obtained while the only added area is for a clock source, switches and a filter.

Example of a multipath circuit: current source

One example of a multipath topology that combines LF noise reduction by LSE and tunability by unequal design of the different paths is the tunable current source in fig. 5.4. The two MOSFETs are driven in antiphase. Because they are driven with the same V_{GS}, and they differ by a factor 2 in their W/L ratio, the output current varies periodically in time between I and $2I$. The average output current is continuously variable between I and $2I$ by changing the duty cycle.

A prototype of this circuit was realized on a breadboard using HEF 4007 n-channel MOSFETs. V_T is approximately 1.9 V. To make a $2W/L$ device, two devices were placed in parallel. V_{GS} was 2 V when the device was 'on' and 0 V when the device was 'off'. The average output current was measured to vary linearly between 140 μA and 280 μA depending on the duty cycle of the driving signal. These devices were observed to exhibit an

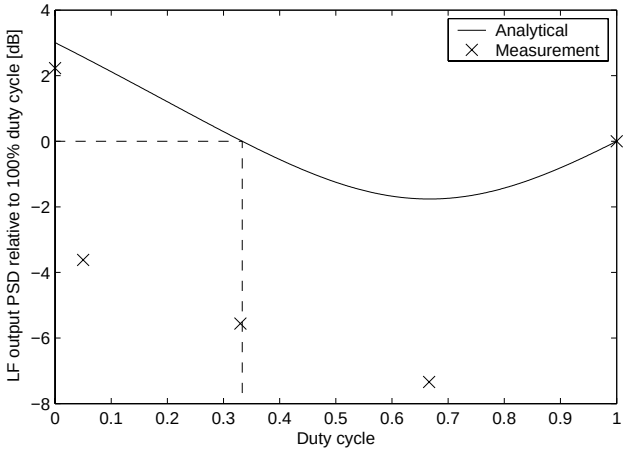


Figure 5.5: Systematic LF noise behaviour of tunable current source

LF noise reduction of 6 to 8 dB when subjected to LSE (see section 3.3 for details), so the same reduction in this circuit application is expected.

Apart from the noise reduction due to LSE, there is also a systematic effect that needs to be considered. $dc = 0$ (0% Duty cycle) is the case where the $2W/L$ device is 'on', and $dc = 1$ is the case where the W/L device is 'on'. Looking at the output LF noise PSD of both devices:

$$\begin{aligned} PSD_{\text{W/L}} &\propto (dc)^2 \\ PSD_{\text{2W/L}} &\propto 2(1-dc)^2 \\ PSD_{\text{total}} &= PSD_{\text{W/L}} + PSD_{\text{2W/L}} \propto dc^2 + 2(1-dc)^2 \end{aligned} \tag{5.1}$$

Taking the output noise for $dc = 1$ (where the W/L device is 'on') as the reference, the expected output noise power as a function of the duty cycle can be plotted. This is the solid line in fig. 5.5. In the absence of any device-physical effects, the same LF output noise PSD for a duty cycle of 0.33 as for a duty cycle of 1 is expected. Any difference observed between the LF output noise for those two duty cycles is attributed to the reduction of LF device noise due to LSE. The measured data is also plotted in the figure for several different duty cycles; it is noted that the LF noise for $dc = 0.33$ is approximately 6 dB lower than for $dc = 1$.

Conclusion

Whereas the multipath technique to combat $1/f$ noise is feasible, and allows use of the LF-LSE technique to reduce LF device noise, it compares unfavourably in terms of power and area spent to the 'classical' chopper structure, which, unlike the multipath technique, not only reduces in-band LF noise but eliminates it altogether. Like the multipath technique, a chopper structure allows parameter tuning by designing both 'halves' of the circuit differently. In situations where addition of a clock, switches and a lowpass filter is not possible, area scaling offers a viable alternative to the multipath topology. Though area scaling reduces the bandwidth of the circuit, the multipath technique requires lowpass filtering to achieve a systematic advantage in the first place, as shown in table 5.2.

In this light, use of the multipath technique to combat $1/f$ noise by the LSE effect cannot be recommended in general.

5.3 Time-discrete circuits

Introduction

An important function of digital computers is signal processing involving inputs from analog sensors and outputs to analog actuators of many sorts. Computers perform signal processing in a time-discrete fashion, and this mandates conversion from the time-continuous to the time-discrete domain and back again. This is done by a sampling system in the ADC at the input, and a reconstruction filter in the DAC at the output.

A generic system is shown in fig. 5.6. It consists of an analog sensor, followed by some analog signal processing/preamplification circuitry, an ADC, a processor, and an output via a suitable DAC to some actuator.

Before the ADC, filtering is required to limit the bandwidth of the input signal to prevent aliasing, and amplification is required to alleviate the noise and dynamic range requirements of the ADC. Noise plays an important role in such a preamp as signal levels may be very low. MOSFET $1/f$ noise is also important in such a preamp since it is often desired to integrate the preamp with the digital CMOS core for single-chip solutions to signal-processing tasks.

The time-discrete nature of the digital signal processor means that there is some design freedom in the analog front-end. In particular, since the

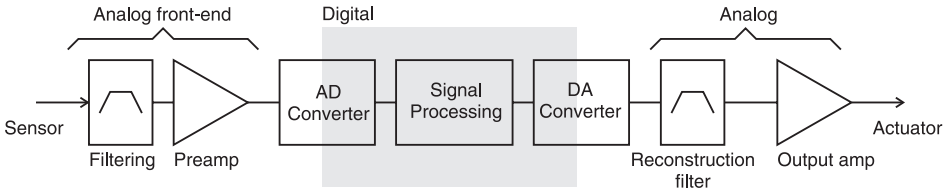


Figure 5.6: Generic signal processing system with digital core

ADC does not require the input signal to be valid at all times, but only around its sample instants, the transfer function of the preamp does not need to be continuous in time. Designing an analog front-end that operates discontinuously in time in synchronism with the system clock will be shown to bring benefits in terms of noise performance.

In simple sampled circuits, LF noise of the active devices may limit the noise performance of the circuit. Since the circuit does not have to operate in a time-continuous fashion, the noise limiting devices in the circuit can be subjected to LSE to reduce their $1/f$ noise. This does not require expenditure of significant area or power. In fact, power consumption may even go down because devices in the preamp do not draw power when they are turned 'off'.

Correlated double sampling

One well-known technique that can be applied in time-discrete systems before sampling to reduce the LF noise of the preamp before the sampler is 'Correlated Double Sampling' (CDS). Despite the positive effect it has on LF input noise, the LF noise may still play a significant role at the output of the CDS circuit. In this section, the conditions under which this is the case will be derived. A generic CDS system is shown in fig. 5.7.

We assume a signal source, followed by a preamp of bandwidth ψf_s that suffers from additive white and $1/f$ noise. The circuit samples the signal at a frequency f_s. The two-phase sampling clock allows us, in phase ϕ_2, to make the input '0' and take a sample of the noise of the preamp before taking a sample of the input signal in phase ϕ_1.

The sample of the noise is then subtracted from the sample of the signal+noise. In figure 5.7 this is done in the digital domain, though it is also possible (and advantageous with respect to quantization noise) to im-

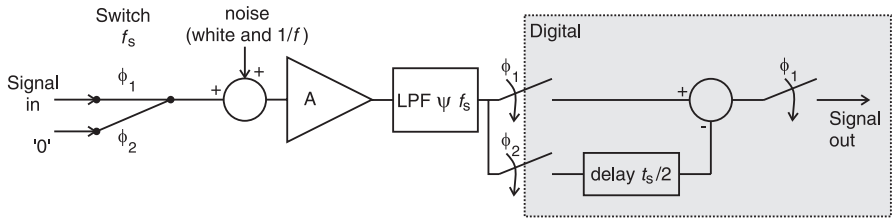


Figure 5.7: Correlated double sampling

plement the subtraction in the analog domain. It is important to note that the CDS can only be applied 'around' an amplifier if its input can be accessed; noise originating from before the first switch is in no way influenced.

Intuitively, the operation of the CDS is obvious: If the noise is strongly correlated in time (i.e. it has a significant component that is LF compared to f_s), the two samples of the noise from phase ϕ_1 and ϕ_2 will be almost identical and the noise will be largely cancelled by the CDS operation. If, on the other hand, the two noise samples are uncorrelated in time, subtracting them is equivalent to adding their powers. Because of this, the CDS operation will double the output noise power if the noise is white.

Despite the attenuation of LF noise by CDS, LF input noise may still be dominant at the output. In appendix A, the condition for which $1/f$ noise is the dominant noise source despite the CDS is derived.

For the realistic case $\psi = 5$; i.e. an amplifier bandwidth $5f_s$,

$$f_c \geq 1.5 f_s \tag{5.2}$$

is found as the condition for which $1/f$ noise dominates despite the CDS. In this equation, f_c is the $1/f$ 'corner frequency' where $S_{1/f}$ equals S_{white}, and f_s is the CDS sampling frequency, i.e. the frequency at which the CDS circuit generates output samples. Numeric solutions for other amplifier bandwidths are given in table 5.3.

If $1/f$ noise is the dominant noise source at the output of the CDS circuit, LF components of the $1/f$ noise are adequately suppressed, but the HF part of the $1/f$ noise contributes noise at the output. From fig. A.5 it is noted that the dominant $1/f$ noise contribution comes from a frequency around $0.74\ f_s$. One example of a circuit where $1/f$ noise is typically the dominant noise source despite the use of CDS is an imager pixel. In the following section, we will discuss how subjecting the devices in an imager pixel to

Relative BW (ψ)	$1/f$ noise dominates if $f_c >$
1	$0.61\,f_s$
2	$0.82\,f_s$
5	$1.50\,f_s$
10	$2.48\,f_s$

Table 5.3: Smaller BW of amplifier in CDS system increases sensitivity to $1/f$ noise

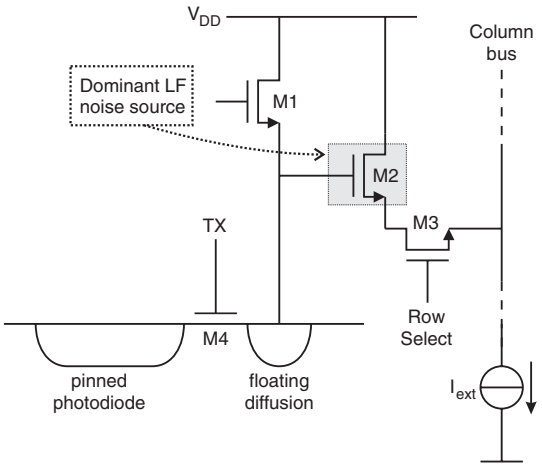


Figure 5.8: CMOS imager pixel

LSE influences their LF noise.

Example of a CDS circuit: Imager

One well known application of CDS is in a CMOS imager pixel. In fig. 5.8, an imager pixel is shown. After a certain integration period during which light is incident on the photodiode, transistor M1 is switched on and resets the floating diffusion to a high potential. This reset voltage is read out by asserting the Row Select line, which connects transistor M2 to the current source I_{ext} that is external to the pixel. After the read-out of this

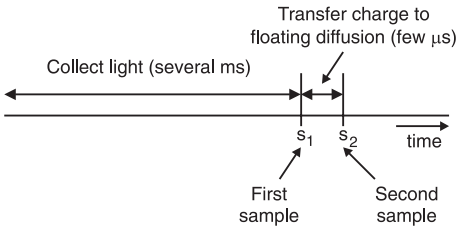


Figure 5.9: Sample positions for CDS in imager

reset voltage, the photo-charge is transferred to the floating diffusion via transfer gate M4, and read-out in the same way. By subtracting both read-out signals a correlated double sampling operation is performed, allowing removal of offset and kT/C noise of the photodiode. M1 to M4 have to be very small to maximize the photosensitive area in the pixel. M1, M3 and M4 are switches and do not exhibit much LF noise, leaving the noise of M2 as the dominant LF noise source in the front-end, despite the correlated double sampling (see table 5.3 above).

LSE can be applied to M2 to influence its LF noise. Applying LSE at the gate of M2 is not possible as parasitic charge injection (if LSE is applied between the two sampling moments of the CDS) would corrupt the photo charge at its gate. Also, this would mean addition of circuitry to every pixel, thereby increasing the pixel size or decreasing the size of the photodiode and thereby the sensitivity. LSE can however be applied at the source of M2 via the column bus. This has the additional advantage that the circuitry to pull 'up' the column bus only needs to be replicated once per column, not once per pixel.

To investigate whether this is beneficial, measurements were performed on a number of devices that were subject to the same bias conditions as M2 would be in a real circuit. In the noise measurements, CDS operation of the actual circuit is replicated so that precisely the noise that is interesting to us is observed. This is illustrated in fig. 5.9. The simplest way to apply LSE is to first keep the device 'off' for a relatively long time, and 0.5 μs after its turn-on transient, a first sample of the noise can be taken (s_1). 3 μs later, a second sample of the noise is taken (s_2). The mean square difference between these two samples is calculated: $\overline{(s_2-s_1)^2}$. This is the output noise power of the CDS, and by measuring it directly and comparing it to $\overline{(s_2-s_1)^2}$ in the steady state, a judgement can be made on whether subjecting M2 to LSE in this manner is useful.

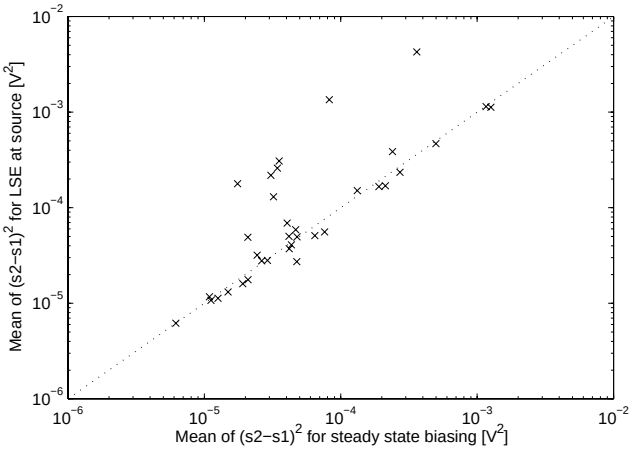


Figure 5.10: Noise measurement when applying CDS

In fig. 5.10, measurement results obtained in this way for 35 devices with areas between 0.175 μm^2 and 1.75 μm^2 are given. In the figure, the noise under steady state bias conditions is plotted along the *x*-axis, and the noise when the device is turned 'off' briefly before the sampling instants by pulling the source terminal up to V_{DD} is plotted along the *y*-axis. The 45^o line is the boundary between those devices showing more, and those showing less noise when subjected to LSE. The majority of devices in this plot lie above the 45^o line: on average, $\overline{(s_2 - s_1)^2}$ rises from 1.5×10^{-4} to 3.2×10^{-4} when the device is subjected to LSE. Spread in the results is considerable; over two orders of magnitude for both the steady state and the LSE measurement. This is a rather disappointing result for the LSE technique which was seen to reduce the LF noise of the same devices in chapter 3.

Explanation of the measurement results

These measurement results can be explained by assuming that the input noise of the CDS circuit is predominantly RTS noise. We will show that the measurement results are in line with the predictions made by the non-steady-state SRH model of chapter 4.

From the analysis of the operation of the CDS circuit, it was found that the dominant noise contribution at the output comes from around 0.74 f_s. To

facilitate analysis, it will be assumed that there is only a single RTS present at the input of the CDS, and that it has an $f_{0,\mathrm{RTS}}$ of 0.74 f_s. RTS noise with a much lower corner frequency is suppressed by the CDS action, and RTS noise with a much higher corner frequency is suppressed by the limited bandwidth of the amplifier. It will further be assumed that the trap causing this RTS is emptied by turning the device 'off' prior to the CDS operation. The time between the two CDS samples is $t_s/2$, which in this application is 3 μs, so t_s =6 μs, f_s =167 kHz and 0.74 f_s is 123 kHz.

When the device is turned 'on', the RTS exhibits transient behaviour. From eq. 4.4, the occupancy of the trap as a function of time can be calculated:

$$f(t) = f_{\text{steady state}}\left[1 - \mathrm{e}^{-t/\tau_{\text{eff}}}\right] \tag{5.3}$$

where $\tau_{\text{eff}} = 1/\omega_{0\text{RTS}}$. This is a simple exponential function starting from 0 and ending at $f_{\text{steady state}}$ for $t \to \infty$. From eq. 4.18, it is known that the steady state variance of a steady state RTS is given by

$$\begin{aligned}\text{var}_{\text{steady state}} &= \frac{\beta}{(1+\beta)^2} \\ &= f_{\text{steady state}}(1 - f_{\text{steady state}})\end{aligned} \tag{5.4}$$

This not only holds for the steady state but also during a transient:

$$\text{var}(t) = f(t)\,(1 - f(t)) \tag{5.5}$$

In fig. 5.11, the transient variance of an RTS is plotted as a function of t/τ_{eff} for three different RTS'es with a steady state β of 0.1, 1 and 10. The characteristic 'hump' in the variance for $\beta = 0.1$ that is here predicted by theory is the same that is observed in the transient measurement of section 4.4.

The CDS operation when the input is the transient occupancy and variance of fig. 5.11 can now be examined. The output noise of the CDS operation is $\overline{(s_2 - s_1)^2}$. A useful simplification can be made by assuming that s_1 and s_2 are not correlated. This simplification is justified by the fact that the RTS in question is so fast that its contribution to the output noise is significant, i.e. as far as CDS is concerned, s_1 and s_2 show little correlation. Under these conditions,

$$\begin{aligned}P_{\text{output CDS}} &= \overline{(s_2 - s_1)^2} \\ &= \text{var}(t_2) + \text{var}(t_1) + [f(t_2) - f(t_1)]^2\end{aligned} \tag{5.6}$$

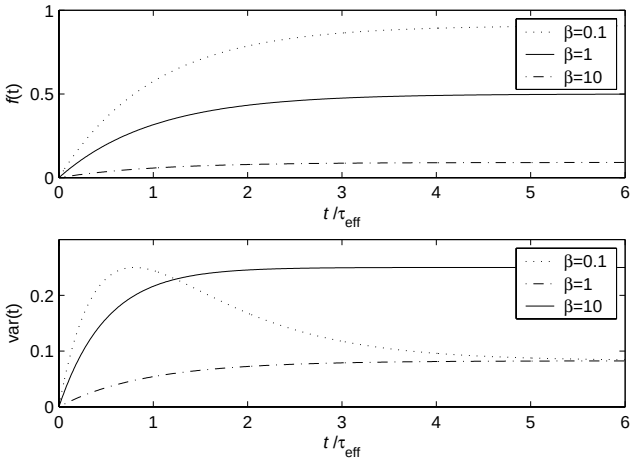


Figure 5.11: Transient occupancy and variance of RTS

In the steady state, $P_{\text{output CDS}}$ is simply given by $\text{var}(t_1)+\text{var}(t_2)$ since the occupancy $f(t)$ is constant. Under transient conditions, the first sample is taken almost instantly after turn-on, as fig. 5.11 shows that the variance is low at that time. In this way, it is hoped $\text{var}(t_1)$ will be minimal. This seems beneficial, however reducing $\text{var}(t_1)$ comes at the expense of a difference between $f(t_1)$ and $f(t_2)$. Taking the extreme case where the first sample is taken at the turn-on instant, and the second sample is taken in steady state, (This condition is not entirely satisfied as the second sample will be taken after approximately $2.3\tau_{\text{eff}}$, but it is a fair approximation for the insight we desire), the output noise of the CDS operation in steady state can be compared to the output noise for the LSE case[1].

In steady state:

$$\begin{aligned} P_{\text{output CDS}} &= \text{var}(t_1)+\text{var}(t_2) \\ &= 2\text{var}_{\text{steady state}} \\ &= 2f_{\text{steady state}}(1-f_{\text{steady state}}) \end{aligned} \tag{5.7}$$

[1] If it is indeed possible to take the first sample at the turn-on instant when the variance of the LF noise is 0, taking a second sample seems rather pointless since we already have a noise-free sample, however the CDS operation is still required to combat offset and kT/C noise of the photodiode.

Immediately after turn-on:

$$\begin{aligned} P_{\text{output CDS}} &= \text{var}(t_2) + \text{var}(t_1) + [f(t_2) - f(t_1)]^2 \\ &= \text{var}_{\text{steady state}} + 0 + (f_{\text{steady state}} - 0)^2 \\ &= f_{\text{steady state}}(1 - f_{\text{steady state}}) + f^2_{\text{steady state}} \\ &= f_{\text{steady state}} \end{aligned} \tag{5.8}$$

The CDS output noise for the steady state case and the LSE case as a function of $f_{\text{steady state}}$ is plotted in fig. 5.12. In the top figure, the CDS output noise is given as a function of the steady state occupancy for both the steady state case and the LSE case. The steady state curve shows a maximum output noise for an occupancy of 0.5. This is twice the variance of the RTS, which is the result expected since it was assumed that the two samples of the RTS were uncorrelated. The LSE curve is linear; as the variance of s_1 decreases, the average difference between s_1 and s_2 increases. Since the average difference between s_1 and s_2 is due to a shift in DC level of the noise, it may seem strange to group it under 'noise'. However, since there is no a priori knowledge of the trap, it is not known what the size of this DC component is and it must consequently be treated as undesired noise. In the lower figure, the ratio between the CDS output noise power for the LSE case and the CDS output noise power for the steady state case is plotted. This figure is central in understanding the measurement results: when subjecting a device to LSE, a small decrease in the output noise (a factor 2 at most) may result; this will happen for traps that have a very low steady state occupancy. For traps that have an occupancy of 0.5, the output noise power is the same for the LSE and the steady state case. For traps with an occupancy of more than 0.5, the ratio between the LSE output noise and the steady state output noise grows rapidly and is not limited.

This is in excellent correspondence with the measurement results of fig. 5.10, where some devices were seen to exhibit a slight decrease in CDS output noise under LSE, many were observed not to be influenced at all, and some showed a strong increase in CDS output noise when subjected to LSE. As further support for this conclusion, it is noted that the transient occupancy and variance as predicted by eq. 5.3 and 5.5 have been observed experimentally in the transient noise measurement presented in section 4.4.

The important question that remains to be answered is whether an LSE biasing scheme can be devised that gives an LF noise benefit in combination with CDS. From the above analysis, one characteristic of such a biasing scheme is clear, namely that it should ensure that the bias history of both sample instants is the same. An on/off square wave with a duty cycle that

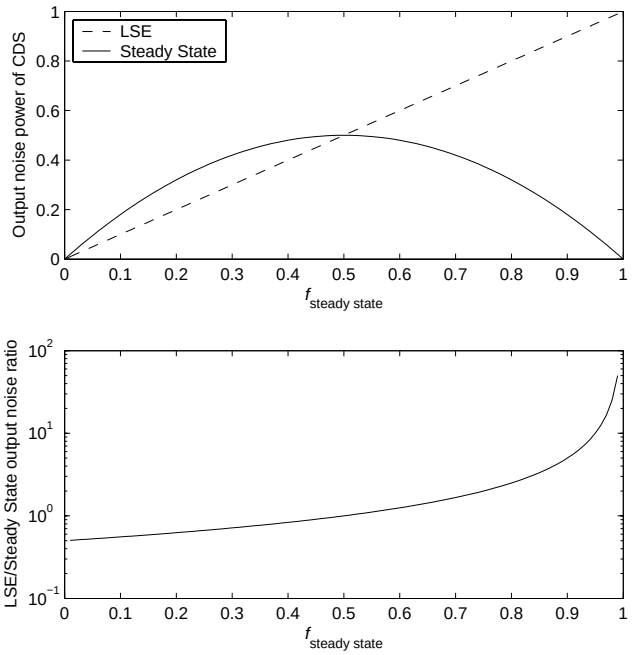


Figure 5.12: Output noise of CDS operation for steady state biasing and LSE

is as low as possible would satisfy this requirement, where the two samples for the CDS are taken in subsequent 'on' periods of the square wave, as long as possible after the start of the square wave.

Conclusion

Simple preamps for time-discrete systems suffer from LF noise, and a reduction in the LF noise is welcome and possible. With minimal area expenditure, it is often possible to turn noise-limiting devices in the preamp off, and this reduces their LF noise. If done in synchronism with the system clock, it can be entirely transparent from a signal processing point of view: the circuit behaves exactly as before with the exception that the LF noise performance of the circuit is improved. Power consumption of the circuit may also decrease.

In correlated double sampling systems, one can, despite the use of CDS, still have a dominant LF noise contribution at the output. Though the low frequency part of the LF noise spectrum is well filtered out by the CDS operation, the high-frequency part of the LF noise spectrum can still contribute significantly to the output noise.

The case where the device is turned off once prior to using it for CDS is analysed. Making the realistic assumptions that the device noise is predominantly RTS noise and that the traps causing the RTS noise are emptied by turning the device off, it was shown that the transient SRH model developed in chapter 4 explains the measurement results. In view of the predictions made by the model, which are in correspondence with the measurement results, subjecting the device to LSE in this manner in order to reduce its LF noise is not recommended in CDS circuits as the disadvantages outweigh the potential benefits. For CDS circuits it is important that the bias history of both sample moments is made as similar as possible.

5.4 Large-signal circuits

Introduction

Large-signal circuits are an important class of circuits where the devices are operated under large-signal conditions. In RF-CMOS design, the supply voltage drops for each next process generation, while at the same time high demands are placed on the required signal to noise ratio or the output power. This mandates large signal operation of the circuit: rail to rail signal swing is not uncommon. Examples of RF circuits where the LF device noise is important and the signal swing is large include VCO's, mixers and PLL's. The high frequency of operation of these circuits means that it is not practical to apply techniques such as correlated double sampling or chopping. Expending power to reduce the LF noise (see sec. 5.1) is not desirable as it increases the often already significant power consumption of the circuit, and a reduction in bandwidth by area-scaling is often not acceptable either. In these circuits, simple measures can sometimes be taken to ensure that the noise limiting devices are cycled to well below threshold, thereby reducing their LF noise.

Upconversion mechanisms

Before solutions to LF noise problems in RF circuits can be considered, we must consider how LF noise comes to play an important role in RF circuits in the first place.

If the circuit is a baseband circuit, the signal is in direct competition with the LF device noise. Whereas traditionally, LF device noise has been a phenomenon relevant at very low frequencies, it is not uncommon for typical devices used in CMOS circuit design to have an LF noise corner frequency, f_c, of up to several MHz. From table 2.3 it can be seen that the corner frequency is expected to rise further in future process generations. As a result of this, LF noise is a factor of importance in virtually all baseband circuits.

LF noise also plays an important role in RF circuits that operate far above the LF noise corner frequency. A brief review of how this happens in some common circuits is presented.

- VCO's
  In VCO's, LF device noise is upconverted to reappear as phase noise close to the carrier. Upconversion mechanisms are discussed in detail in [21]. The analysis is linear, time-variant; using an 'Impulse Sensitivity Function' (ISF) the sensitivity of the oscillator to disturbances such as LF noise is modelled. The model explains that white noise reappears as $1/\Delta f^2$ phase noise, and $1/f$ noise appears as $1/\Delta f^3$ phase noise. In theory, a perfectly symmetrical waveform gives a DC-free ISF which should prevent upconversion of LF noise to phase noise altogether. Upconversion mechanisms in ring oscillators are discussed in [22] and [19].
- Mixers
  In an active mixer as given in fig. 5.13, a simple analysis might lead to the false conclusion that there is no LF device noise at the output. After all, LF noise of the input transconductor M1 is upconverted by the mixer to appear around multiples of the LO frequency at the output, where it is filtered out by an IF lowpass filter. The switch transistors M2 and M3 might be assumed to switch in an instantaneous fashion, so the switch transistor is either off or in series with the input transconductor. If the output impedance of the input transconductor is high, no LF switch noise is expected at the output.
  A more detailed analysis of the active mixer [13] reveals that there are in fact several mechanisms by which LF device noise appears at the output in direct competition with the desired IF signal. The most

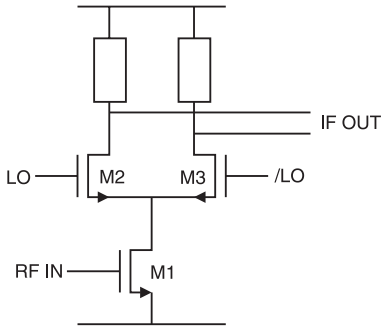


Figure 5.13: Active mixer

important contributions to LF noise at the output are: (1) Gate referred LF noise of the switches that modulates the switching instant; (2) Input transconductor LF noise that leaks to the output via offset of the switches, and (3) LF switch noise that charges and discharges the parasitic capacitor at the source of the switch pair. This current is rectified by the mixer to appear in baseband at the output.
Driving the mixer with an ideal rail-to-rail square wave at the LO port (this is not possible in practice due to limits on LO drive capability and power dissipation) could in theory eliminate output LF noise due to effect (1) and (2), however effect (3) can never completely be eliminated. The problem of LF noise at the output is obviously more acute if the mixer is used in a low or zero-IF topology as is common practice in RF-CMOS front-ends.
Passive mixers are sometimes favoured for their good linearity, the fact that they have no static power consumption and, since they are operated at V_{DS} =0, the promise that they will be free from LF noise. Detailed analysis [46], however, shows that the amplifiers following such mixers contribute significant LF noise to the output.

- Phase-Locked Loop (PLL)
  A typical function of a PLL (fig. 5.14) is to multiply a reference frequency by an integer factor to produce a HF output signal that is precisely locked in phase and frequency to the reference frequency. The reference frequency is typically generated by a quartz crystal, and is of the order of several tens of MHz at most. In its simplest form, the PLL consists of a Phase Frequency Detector (PFD), a charge pump and loop filter driving a VCO and a frequency divider.
  Jitter performance of the PLL is very important. At offset frequen-

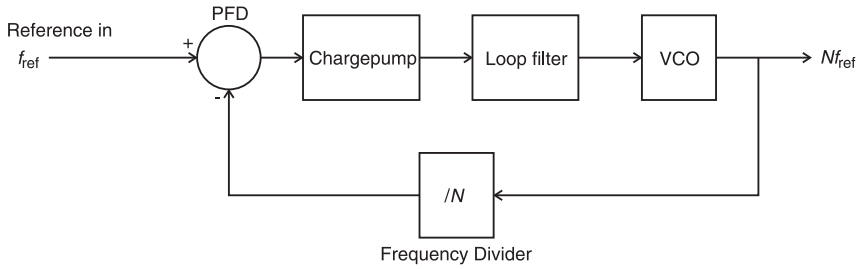


Figure 5.14: Phase Locked Loop

cies larger than the loop bandwidth of the PLL, jitter performance is dominated by the characteristics of the VCO. For smaller offset frequencies (the 'in-band' performance of the PLL), jitter is dominated by contributions from the charge pump, the PFD and the frequency divider. Detailed jitter analysis of PLL's is given in [44], [5] and [4], as well as [64].

Measurement of LF noise under LSE at high frequencies

To verify that subjecting the devices to LSE can usefully reduce their LF noise in applications as those mentioned above, LF device noise was measured while the devices were subjected to LSE at a high frequency [67]. The LF noise of the MOSFET is reduced by subjecting it to LSE at frequencies up to 3 GHz. The experimental method and results are given below.

Method

The easiest way to measure the LF noise of the MOSFET under LSE is to bias it with a square wave. In this way, the LF noise of the MOSFET under LSE can easily be compared to the LF noise of the MOSFET under steady-state bias conditions. A correction needs only to be made for the systematic effect of the device not being 'on' continuously (see section 3.2, page 35). If care is taken that the 'on' state of the square wave bias is identical to the steady state bias, then it is not necessary to measure or take into account a possible bias dependency of the LF noise. This method is useful for frequencies of excitation up to a few MHz. For higher frequencies a different approach is required. Generating a well-behaved square wave at

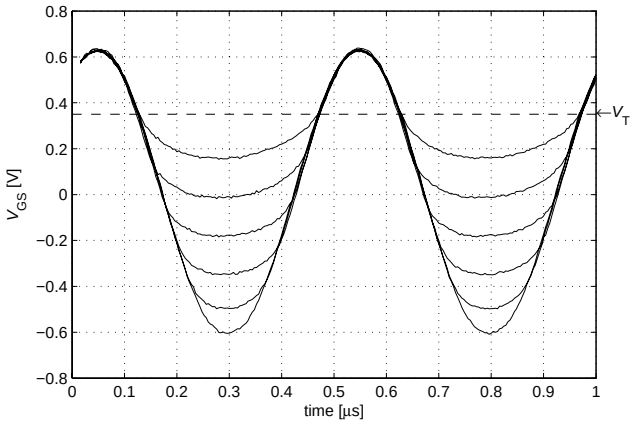


Figure 5.15: Excitation voltage waveform with a variable minimum voltage

RF is more difficult than generating a sine wave at the same frequency as fast square waves place extreme demands on the bandwidth of the circuitry. Moreover, typical fast square wave signal generators are not low-noise. (Equivalent noise resistance values as high as 1 MΩ are no exception for 50 Ω pulse-generators.) In contrast, low noise sine wave generators for RF frequencies are readily available, so this is a much more suitable choice for RF measurements.

While using a sine instead of a square wave is much more feasible at RF, it makes interpretation of the measurement results more difficult: The device passes through many biasing points each period and comparison to a DC-bias noise measurement is not straightforward. The solution is found by realizing that the MOS device only contributes significant noise during the time the gate-source voltage is above threshold (and $I_D \neq 0$). Also, it is known from other experiments (see section 3.4) that the device noise reduction is a function of the minimum off-voltage. Hence, the minimum off-voltage can be varied while keeping the waveform *above* threshold the same: Such a waveform is shown in fig. 5.15. The threshold voltage of the device under test is shown in the figure. As long as the waveform above threshold does not change, neither does the DC drain current of the device, and simulators using existing MOSFET noise models (see section 2.5) predict that the LF noise does not change either.

To produce the waveform of fig. 5.15, the circuit of fig. 5.16 is used. The

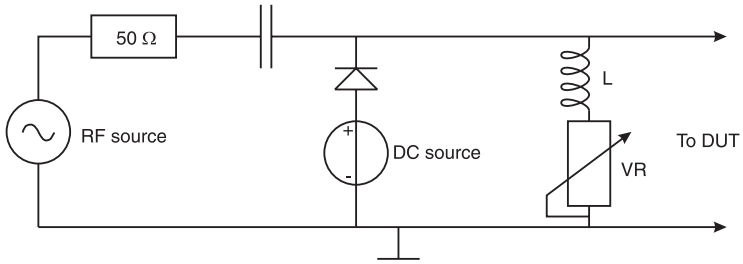


Figure 5.16: Circuit to produce the waveform of fig. 5.15

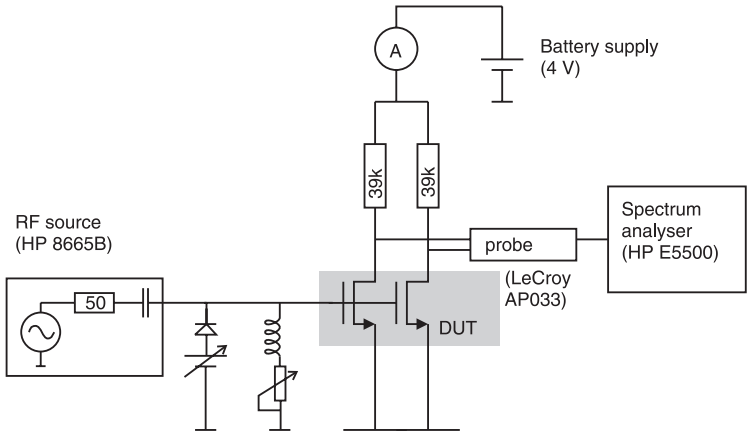


Figure 5.17: The complete measurement setup

diode in this circuit conducts during the negative half-period of the RF sine wave, limiting the negative excursion of the waveform. An RF Choke 'L' in combination with a variable resistor ensures that the positive half of the waveform does not shift as the minimum off-voltage is varied. The variable resistor needs to be adjusted only once. This waveform is applied to the gates of the devices under test (see fig. 5.17). Two identical DUTs are driven in-phase. This allows us to use the common mode rejection of a differential probe to filter out the RF driving signal and measure the LF noise of the devices. A low pass filter at the drains (not shown; cutoff frequency is 100 kHz) helps to keep the drain voltage of the devices constant during the measurement.

The diode in the circuit has to have a large current handling capacity; around

1 V from a 50 Ω source is driven into it; this is approximately 20 mA. At this current, a low forward voltage drop is desired. These demands are fulfilled by a relatively large, slow diode, which means the setup will not function at very high frequencies. Fortunately, this is not required and the measurement can be carried out in two steps: First, at 2 MHz, we show that varying the minimum 'off' voltage reduces the LF noise of the devices. Next, we show that the LF noise of the device does not change as the frequency of sine wave excitation is varied.

The first part of the measurement is carried out at 2 MHz. The variable resistor is adjusted so that the top part of the sine wave is not influenced by a variation in the minimum off-voltage, as witnessed by the fact that the DC drain current of the devices does not change as the minimum off-voltage is varied. For additional verification that varying the minimum off-voltage does not influence the effective transconductance of the devices, a small pilot tone (150 μV at 5 kHz) is injected into the gate of one of the devices under test and its amplitude at the drain is measured. By this method, the time-average transconductance of the devices was indeed found to be independent of the minimum off-voltage. The LF noise of the devices as a function of the minimum off-voltage of the driving waveform can now be measured. Having carried out this varying off-voltage measurement we can proceed to RF sine wave measurements. These are carried out with an unmodified sine wave driving the gates of the DUTs. (The full sine wave with an amplitude of 600 mV appears at the gate of the DUTs.) As the frequency of excitation is varied, the bias voltage trajectory of the devices does not change, and neither the average drain current nor the effective transconductance of the device changes. Again, injecting a pilot tone and measuring its amplitude at the drain verifies that this assumption is correct.

Summing up the two-step measurement method:

- At an excitation frequency of 2 MHz, it is demonstrated that the LF noise of the devices is reduced if the device is cycled to accumulation.
- It is then shown that the LF noise remains low for excitation frequencies of up to 3 GHz.

RF measurement results

Measurements were carried out on n-channel devices from a mature industrial 0.35 μm process: t_{ox} =7.5 nm, V_T =0.35 V and $W/L = 50/0.35$ μm. 6 pairs of devices from the same wafer were measured. Devices were contac-

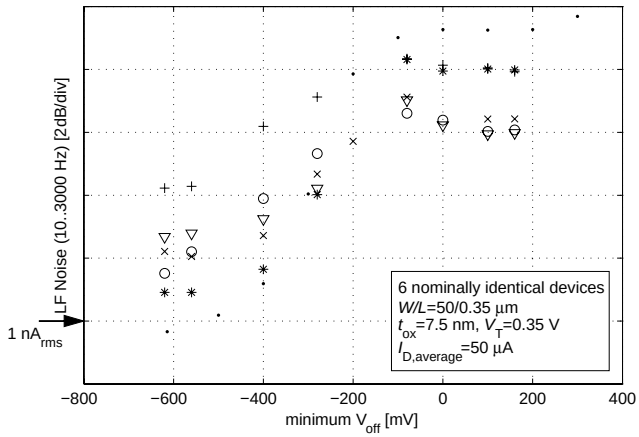


Figure 5.18: Measurement at 2 MHz showing that cycling the device well below threshold reduces its LF noise relative to cycling just below threshold

ted using a wafer prober and coaxial probes. The RF signal path is a 50 Ω line terminated at the DUTs. V_{DS} =2 V and the amplitude of the driving signal is 600 mV at the DUT gate. V_S =V_B =0 V. Average current through each device is 50 µA for all measurements.

For each measurement, the LF noise spectrum is measured using a spectrum analyser. The spectrum is integrated from 10 Hz to 3 kHz, and the integrated noise is plotted in fig. 5.18 and fig. 5.19. The results for six nominally identical device pairs are shown. For reference, 1 nA_{rms} drain current noise (in the same bandwidth) is shown in both figures. All devices show the same trend. Fig. 5.18 shows that as the minimum off-voltage is reduced, without changing the effective transconductance of the device, the LF noise of the device is reduced. The reduction (just as the absolute noise level) shows spread from device to device. It is in the range of 4-10 dB. This is in accordance with the results from section 3.6 where a large spread in the steady state noise level and the noise reduction from device to device was also observed.

Fig. 5.19 shows that, when driving the DUTs with an unmodified sine wave, the noise reduction observed at 2 MHz is retained as the frequency of excitation is increased. The accuracy of the measurements is estimated at better than +/- 0.5 dB at frequencies up to 500 MHz, and slightly less for

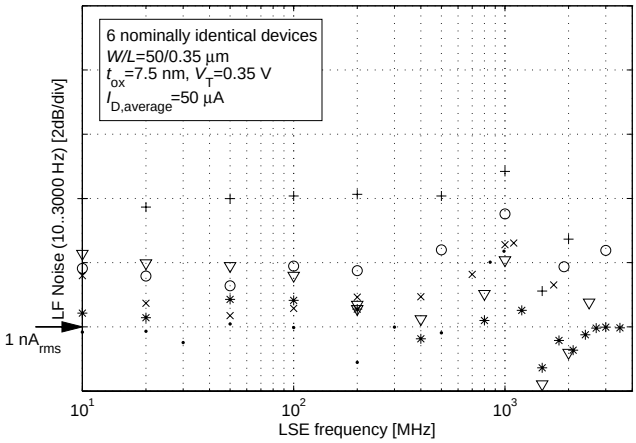


Figure 5.19: Measurement results at higher frequencies showing that LF noise is not a function of the frequency of excitation for frequencies up to 3 GHz

higher frequencies. In summary, it is shown using a two-step measurement method that cycling a MOSFET to well below threshold at 2 MHz reduces its LF noise by up to 10 dB, and that the noise reduction observed by this LSE is essentially independent of the frequency of excitation for excitation frequencies of up to at least 3 GHz.

Conclusion

The measurement results show that subjecting the devices to LSE reduces their LF noise. This effect can be usefully applied in RF circuits like the ones discussed in this section. In these circuits, three distinct phenomena need to be distinguished: First of all, upconversion of LF noise is a sensitive function of the signal waveform. Secondly, LF noise is bias dependent. Finally, subjecting the device to LSE reduces its LF noise. On the basis of experimental data, it is not always possible to make a clear distinction between these three effects, however taking measures to ensure that the devices are cycled to well below threshold to reduce their LF noise is a good idea.

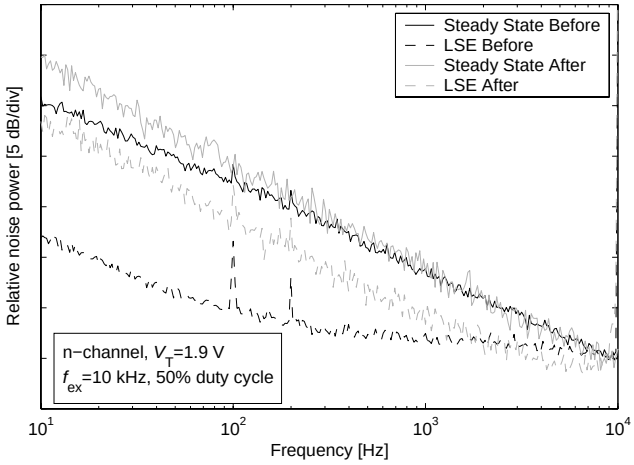


Figure 5.20: LF noise under LSE is a sensitive detector of radiation damage

5.5 Novel applications of LF noise under LSE

Apart from normal circuit design, in which subjecting a device to LSE can reduce its LF noise, there are some novel applications of LF noise under LSE in device characterization. These applications stem from the fact that under LSE, a different set of traps is contributing to the LF noise of the device compared to the steady state case.

One novel application is of LF noise under LSE as a sensitive detector of radiation damage to CMOS devices. To illustrate this, an LF noise measurement in steady state and under LSE was carried out. The device was then subjected to ionizing radiation. This causes defects in the device. Afterwards, the LF noise was measured again. Results are shown in fig. 5.20. Whereas there is no large difference in the steady state measurement before and after damaging the device, the LSE measurement has changed significantly. Integrating the noise from 10 to 500 Hz, the steady state noise power has increased by 2 dB. At the same time, the LSE noise power has increased by 8.5 dB. It is clear that the LSE noise measurement is far more sensitive to this kind of device damage than the DC noise measurement.

A very similar outcome is observed from an experiment in which devices are damaged by hot carrier degradation [39]. LF noise measurements in steady state and under LSE before and after stressing the devices are shown

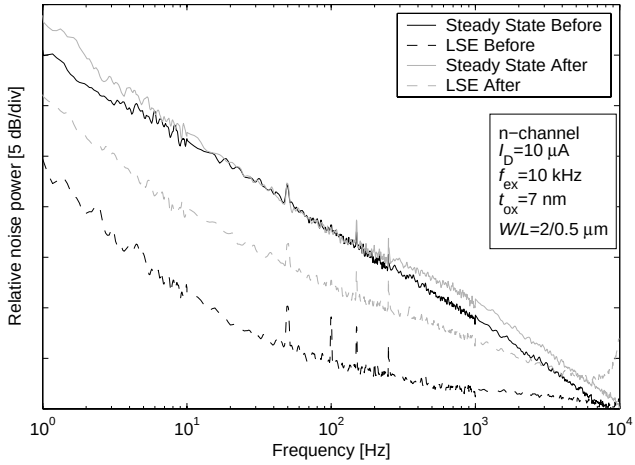


Figure 5.21: LF noise under LSE is a sensitive detector of hot carrier damage; average of 12 devices

in fig. 5.21. Considering the integrated noise power up to 500 Hz, it is noted that the steady state noise has gone up by 0.95 dB, whereas the LSE noise has increased by 7 dB in the same time. Here, too, LSE noise is a much more sensitive detector of device damage than steady-state LF noise.

Both results can be explained using the model of chapter 4. In steady state, LF noise is generated by traps near the conduction band edge and under LSE, LF noise is generated by traps deeper in the bandgap (see figures 4.3, 4.4 and 4.5). Before damage the trap density deep in the bandgap is apparently lower than near the conduction band edge; hence the difference between the steady state and the LSE noise measurement. After damaging the device, the trap distribution in energy is much more uniform which explains the fact that the LF noise reduction by LSE is almost absent after damaging the device.

5.6 Characterization of LF noise

In this chapter, we have discussed in what cases LF noise reduction by LSE can be exploited in circuit design. Measurement results were presented in chapter 3. In all measurements, it was found that the LF noise in nominally

identical devices exhibits considerable spread; in minimal-area devices of modern processes, the variation is over two orders of magnitude.

In the characterization of LF noise for compact modelling, a rather pragmatic approach is often followed: it is better to overestimate the LF noise rather than underestimate it, so parameters are set rather high to ensure that the model never underestimates the LF device noise [56]. While it is certainly true that overestimating is better than underestimating LF noise, it is even better to accurately predict the LF noise, since overestimation of LF noise leads to unnecessary expenditure of area as shown in section 5.1.

Though in small-area devices LF noise is dominated by RTS, and the LF noise spectrum may show significant deviation from a $1/f$ shape, this does not detract from the aim of accurate modelling: The model should accurately predict the noise level that the required percentage of devices fall below (eg. 99.7% for the commonly used 3σ limit). In a production environment, individual device behaviour is hardly ever relevant; more important is what the average and spread of a parameter is over a large number of devices.

If practical use is to be made of noise reduction by subjecting the devices to LSE, it is imperative first to accurately characterize the LF noise of the devices in steady state.

Only when correspondence between LF noise simulation and measurement is within a dB or so, as is presently the case for thermal noise modelling, will a reduction in LF noise by LSE offer a useful advance in performance. The difference between the steady state LF noise model and the actual LF noise performance of the devices should be less than the anticipated LF noise improvement from LSE, since otherwise, the benefits of the effect are not significant.

This means that when a process is characterized, a sufficient number of devices needs to be measured to accurately predict a noise level that the required percentage of devices will fall below (eg. 99.7%). Furthermore, wafer-to-wafer spread of LF noise needs to be considered, and the characterization should be repeated to ensure that the LF noise does not change significantly during production.

In terms of modelling, it is a good idea to express LF noise and bias dependence of LF noise in terms of $N_t(E,x)$: the trap density as a function of the energy level and the distance to the channel. In this way, the effects of LSE on LF noise can be incorporated in future models with a minimum of effort, as an accurate description of $N_t(E,x)$ is sufficient to model LF noise behaviour under LSE as shown in chapter 4.

5.7 Conclusion

In this chapter application of LF noise reduction by LSE was discussed. A comparison was made to traditional methods of tackling the LF noise problem, such as scaling, CDS and chopping.

In time-continuous circuits, LF noise reduction by LSE may be applied. The benefits are limited in comparison to simple area scaling. If bandwidth is not important and lowpass filtering at the output is viable, then LSE compares unfavourably to a differential structure with a chopper. In some specific cases (eg. bias current sources) making the circuit differential may not be possible and in those cases subjecting the devices to LSE may offer a benefit.

In time-discrete circuits, the transfer function of the circuit does not need to be continuous in time, and LF noise reduction by LSE may be conveniently applied. There is no significant penalty in terms of area spent, and power consumption of the circuit may even go down. A common technique to reduce the LF noise of analog front-ends to digital systems is correlated double sampling (CDS).

If LF noise reduction by LSE is to be applied in combination with CDS, care must be taken that the bias history of both sample moments is as similar as possible, since when this is not the case, LSE is more likely to increase than to decrease the CDS output noise.

In RF circuits, devices are often subjected to LSE by virtue of the operation of the circuit. In these circuits, it is sometimes easy to ensure that the noise-limiting devices are cycled well below threshold, thereby reducing their LF noise. We show that LF noise reduction by LSE works for frequencies of excitation of up to at least 3 GHz. This is in line with the model presented in chapter 4, which predicts that the LF noise under LSE is not a function of the frequency of excitation, as long as the frequency of excitation is high compared to the LF noise in question. When analysing such circuits, it is important to attribute observed changes in LF noise performance to the appropriate mechanism: changing the signal waveform in general changes three things: the bias dependence of the noise, the upconversion, and the LF noise due to LSE.

A novel application of LF noise under LSE is presented; namely that it is an extremely sensitive detector of device damage, eg. by ionizing radiation or by hot carrier stress. This is understood because LF noise under LSE is attributed to a different group of traps than the ones responsible for LF noise

in the steady state. Whilst this is useful for trap characterization, it is also important to realize that as a result of device damage, LF noise reduction due to LSE may disappear.

Finally, it is noted that while in certain cases, subjecting a device to LSE can usefully lower its LF noise, the attainable benefit is limited and the LF noise does not disappear completely when the device is subjected to LSE. For LF noise reduction under LSE to be useful in analog circuit design, it is first necessary that steady state LF noise models be made more accurate. Once steady state LF noise models can accurately predict LF noise, LSE can be applied as a technique to further reduce LF noise. LF noise models based on an accurate characterization of the distribution of traps in energy and position, $N_t(E,x)$ are beneficial since they not only allow prediction of steady state LF noise but also of LF noise under LSE.

Chapter 6

Conclusion

6.1 Summary of conclusions

A summary of the most important conclusions of this thesis is given below.

Regarding LF noise in MOSFETs, it is noted that the MOSFET is a surface channel device. Both n and p-channel devices exhibit similar low frequency (LF) noise behaviour that can be explained by a carrier number fluctuation model (section 3.5). LF noise in MOSFETs is predominantly caused by Random Telegraph Signals (RTS). This is the case for small devices, but also in larger devices with an area as large as 10 μm^2, LF noise is dominated by RTS.

Regarding the relevance of the topic, it is expected that LF noise in MOSFETs will remain an important issue in future process generations (table 2.3 in chapter 2). In the most realistic scenarios, the LF noise corner frequency will remain the same or increase as CMOS processes are downscaled which means that in many applications (both baseband and RF), LF noise will continue to be or will become a significant problem. Though an important question in this context is what kind of gate dielectric will be used for future CMOS generations, much of the work in this thesis is applicable to any gate dielectric as long as it contains traps. High-κ gate dielectrics are renowned for their high trap densities.

Regarding the influence of LSE on LF noise in MOSFETs, it is noted that biasing a device to an 'off' state prior to turning it on for use reduces its LF noise on average (table 3.3 in chapter 3). Also, the spread of the LF

noise is reduced. For this to happen, the device needs to be turned off to well below threshold. LF device noise is sensitive to V_{GS}, so the device can be turned off by either increasing V_S or decreasing V_G. LF noise in MOSFETs is a function of the time-average V_{GS}; the frequency of excitation is not important as long it is high relative to the frequency of LF noise (fig. 5.19).

The influence on LF noise by LSE is explained by a model based on the Shockley-Read-Hall (SRH) model in chapter 4. It explains LF noise behaviour in a device under LSE. Only one assumption (which is supported by experimental data) is made, namely that the emission time constant of a trap (τ_e) is dependent on V_{GS}. With this assumption, the SRH-model predicts that LF noise under LSE comes from a different group of traps compared to the steady state case. Making the additional assumption (which is again validated experimentally by several authors) that the distribution of traps in energy is non-uniform, it can be explained that the LF noise under LSE changes. In this way, we can account for all measurement results of LF noise under LSE.

LF noise reduction by LSE can be applied in circuit design, but it is by no means a cure-all solution to LF noise problems. Only in specific cases does it bring some benefit. In time-continuous circuits, LF noise reduction by LSE competes with established noise reduction methods such as area scaling or modulating techniques that separate the LF noise and the signal in frequency. The comparison in section 5.2 indicates that LF noise reduction by LSE is only competitive in very specific cases. Whether or not it is competitive depends on the process in question. In simple time discrete circuits, subjecting devices to LSE is without penalty so there is no reason not to do so. If LF noise reduction by LSE is to be successfully applied in conjunction with correlated double sampling, care must be taken that the bias history for both samples is identical, as otherwise LSE will degrade the noise performance of the circuit. In RF applications, LF noise reduction by LSE is most promising since other techniques are not applicable there.

Regarding recommendations for future work, the further development of noise models based on the distribution of traps in energy and position; $N_t(E,x)$; is to be recommended. If the characterization of trap densities in position and energy level is correct and complete, such models can explain LF noise in steady state and under LSE (chapter 4). To this end, it is beneficial if trap densities are characterized by some technique orthogonal to LF noise measurements. In creation of parameter sets, large numbers of devices need to be characterized to ensure that models accurately and reliably predict LF noise and spread in LF noise. Only in this way can designers

produce optimal designs.

6.2 Original contributions of this thesis

Chapter 2:

- The dependence of the relative conductivity fluctuation ($S_{\Delta\sigma}/\sigma^2$) on the free carrier concentration n is derived. This is the basis on which the $\Delta\mu$, ΔN and correlated ΔN and $\Delta\mu$ LF noise models are distinguished.

Chapter 3:

- It is shown that LF noise in MOSFETs is influenced by the history of V_{GS} and that V_{GS} may be changed by changing V_G and keeping V_S constant or equivalently by changing V_S and keeping V_G constant.
- LF noise in p- and n-channel devices is shown to behave in the same way when the device is subjected to LSE.
- LF noise reduction by LSE is shown to be a very variable effect. Subjecting a device to LSE is seen to make its LF noise go up or go down, but on average, the LF noise of a device is reduced if the device is subjected to LSE.

Chapter 4:

- LF noise behaviour under LSE is modelled using the Shockley-Read-Hall model applied under non-steady-state conditions.
- A cyclostationary RTS is described by an equivalent stationary RTS in the limit for $T \to 0$, and the time constants for this equivalent RTS are derived.
- Under LSE, the LF noise of a MOSFET is shown to be caused by a different group of traps compared to the steady state bias situation. If traps are not uniformly distributed in energy level, the LF noise of the device will change when the device is subjected to LSE.

Chapter 5:

- 'Multipath scaling', where multiple circuits are alternately used for signal processing is presented as a technique to reduce LF noise.
- LF noise reduction by LSE used in conjunction with correlated double sampling (CDS) is shown to be disadvantageous if the bias history of both sample instants is not identical.
- A measurement method is presented that allows measurement of LF noise under LSE for high excitation frequencies without making any assumptions about the bias dependency of LF noise. Using this measurement method, it is shown that LF noise under LSE is not dependent on the frequency of excitation, thereby validating the model of chapter 4.
- LF noise under LSE is seen to be an extremely sensitive detector of device defects, eg. due to radiation damage or hot electron damage.

Appendix A

Noise performance of CDS circuit.

Despite the attenuation of LF noise by Correlated Double Sampling (CDS), LF input noise may still be dominant at the output. The conditions for which this is the case will be derived below. For noise analysis, we assume a zero input signal, which reduces the circuit of fig. 5.7 to that of fig. A.1 (a). The analog equivalent of this circuit is given in fig. A.1 (b) [31]. For analysis purposes it is further assumed that $A = 1$.

The transfer function of the CDS circuit with regards to noise performance can be derived. The impulse response and transfer function of the circuit of

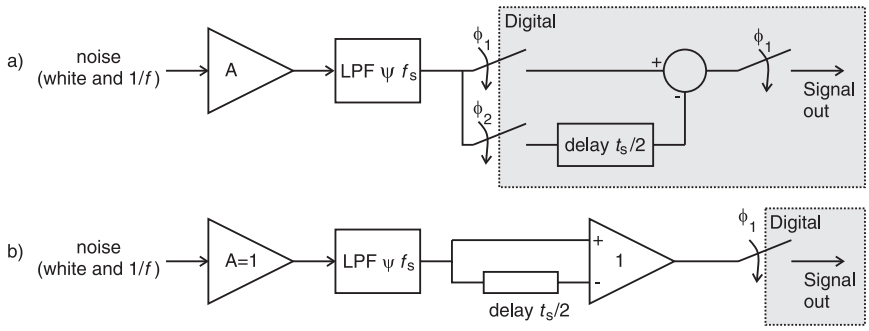


Figure A.1: Equivalent CDS circuit with regards to noise performance

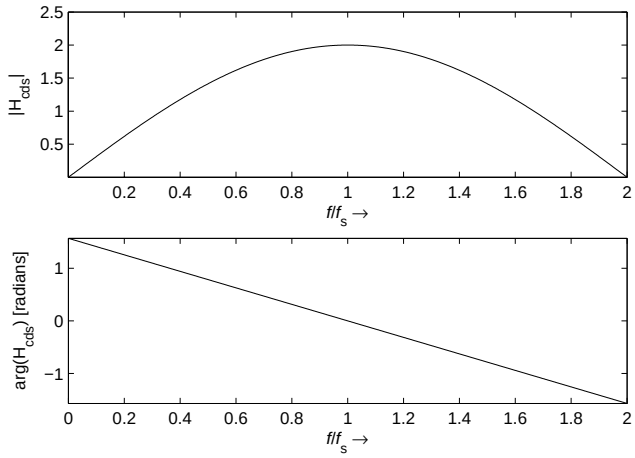


Figure A.2: Transfer function of CDS circuit

fig. A.1 (b) are given by:

$$\begin{aligned} h_{\text{cds}} &= \delta(t) - \delta(t - t_{\text{s}}/2) \\ H_{\text{cds}}(\omega) &= \text{FT}(h_{\text{cds}}) \\ &= 1 - \text{e}^{\frac{-j\omega t_{\text{s}}}{2}} \\ H_{\text{cds}}(f) &= 1 - \text{e}^{\frac{-j\pi f}{f_{\text{s}}}} \end{aligned} \tag{A.1}$$

In fig. A.2, the transfer function of the CDS circuit is given as a function of f/f_{s}. This transfer function is periodic in frequency; only the first period (from $f = 0$ to $f = 2f_{\text{s}}$) is shown. The output noise PSD is given by:

$$S_{\text{output}}(f) = S_{\text{input}}(f)\,|H_{\text{cds}}(f)|^2 \tag{A.2}$$

The integrated noise at the output of fig. A.1 (b) is given by:

$$\begin{aligned} P_{\text{output}} &= \int_0^{\psi f_{\text{s}}} S_{\text{output}}(f)\,df \\ &= \int_0^{\psi f_{\text{s}}} S_{\text{input}}(f)\,|H_{\text{cds}}(f)|^2\,df \end{aligned} \tag{A.3}$$

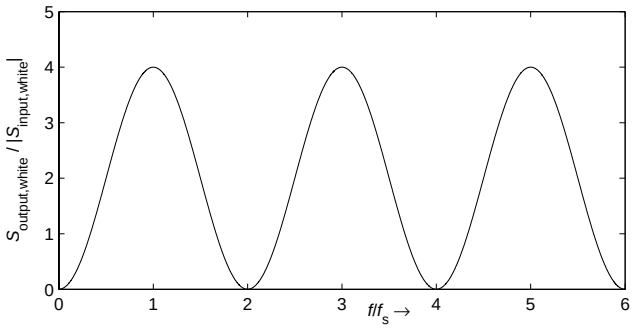


Figure A.3: Output PSD of CDS circuit when input noise is white.

For white noise, this is

$$\begin{aligned} P_{\text{output,white}} &= S_{\text{white}} \int_0^{\psi f_s} |H_{\text{cds}}(f)|^2 df \\ &= S_{\text{white}} \int_0^{\psi f_s} |1 - e^{\frac{-j\pi f}{f_s}}|^2 df \end{aligned} \tag{A.4}$$

The output noise PSD due to white noise is given in fig. A.3. Along the *x*-axis, the frequency, normalized to f_s is plotted, and along the *y*-axis, the output noise PSD, normalized to the input noise PSD is plotted.

In fig. A.4, the relative bandwidth of the amplifier, ψ, is plotted along the *x*-axis, and the ratio between the integrated output noise power and the integrated input noise power is plotted along the *y*-axis. It is noted that the power gain of the CDS for integrated noise is exactly 2 if the bandwidth of the white noise is nf_s, with n an integer ≥ 1. If the bandwidth of the white noise is much larger than $2f_s$, the power gain of the CDS approaches 2, as can be understood intuitively.

For $1/f$ noise, the same approach is followed: The $1/f$ input PSD is given by:

$$S_{\text{input,1/f}} = \frac{S_{\text{white}} f_c}{f} \tag{A.5}$$

in which f_c is the $1/f$ 'corner frequency', at which the $1/f$ noise PSD

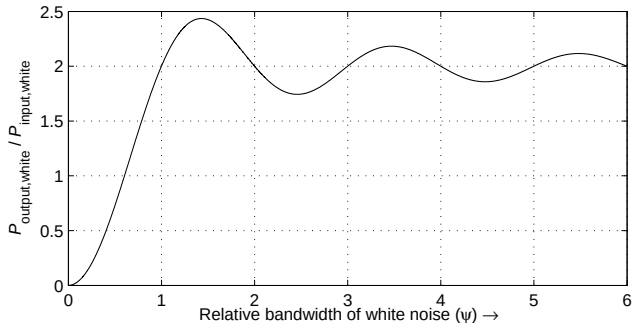


Figure A.4: Noise power gain of CDS circuit for band-limited white noise.

equals the white noise PSD.

$$\begin{aligned} P_{\text{output,1/f}} &= \int_0^{\psi f_s} S_{1/\text{f}}(f)\,|H_{\text{cds}}(f)|^2 df \\ &= \int_0^{\psi f_s} \frac{S_{\text{white}}\ f_c}{f}\,|H_{\text{cds}}(f)|^2 df \\ &= S_{\text{white}}\ f_c \int_0^{\psi f_s} \frac{|H_{\text{cds}}(f)|^2}{f} \quad df \end{aligned} \tag{A.6}$$

There are two things that should be noted. First, this integral is convergent despite the lower limit of $f = 0$ because $|H_{\text{cds}}(f)|^2$ goes to 0 faster than $1/f$ goes to infinity, and secondly, the integral depends on ψ and f_c but not on f_s. This is because the integration limit and H_{cds} both contain f_s terms that cancel.

The output PSD due to $1/f$ noise is given in fig. A.5. Along the *x*-axis, the frequency, normalized to the sample frequency is plotted, and along the *y*-axis, the output noise PSD normalized to $S_{\text{white}}\ f_c$ is plotted. This normalization is derived from the definition of $S_{1/\text{f}}$ by equation A.5.

The integrated output noise due to band-limited $1/f$ noise is given in fig. A.6. Along the *x*-axis, the relative bandwidth of the amplifier, ψ, is plotted. Along the *y*-axis, the output noise power relative to the input noise power is plotted.

To derive under what conditions the $1/f$ noise is dominant at the output, a comparison is made between $P_{\text{output,white}}$ and $P_{\text{output,1/f}}$, and a condition is

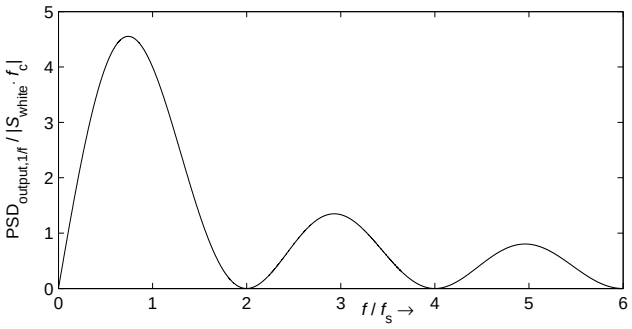


Figure A.5: Output noise PSD of CDS circuit due to $1/f$ noise

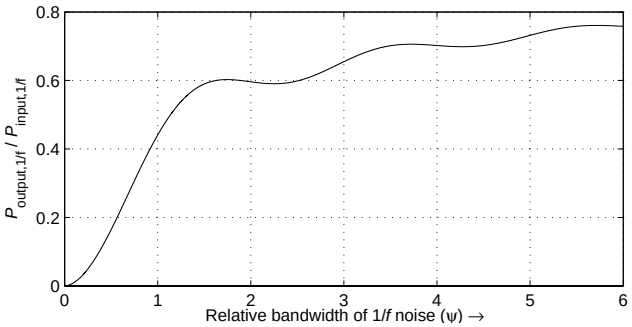


Figure A.6: Noise power gain of CDS circuit for band-limited $1/f$ noise

derived for f_c in terms of f_s.:

$$\begin{aligned} P_{\text{output,white}} &\leq P_{\text{output,1/f}} \\ S_{\text{white}} \int_0^{\psi f_s} |H_{\text{cds}}(f)|^2 df &\leq S_{\text{white}} f_c \int_0^{\psi f_s} \frac{|H_{\text{cds}}(f)|^2}{f} df \\ \int_0^{\psi f_s} |H_{\text{cds}}(f)|^2 df &\leq f_c \int_0^{\psi f_s} \frac{|H_{\text{cds}}(f)|^2}{f} df \\ 2 f_s \left[\psi - \frac{\sin \psi\pi}{\pi} \right] &\leq 2 f_c \left[\ln \pi + \ln \psi - \text{Ci}(\psi\pi) + \gamma \right] \\ \frac{f_c}{f_s} &\geq \frac{\psi - \frac{\sin \psi\pi}{\pi}}{\ln \pi + \ln \psi - \text{Ci}(\psi\pi) + \gamma} \end{aligned} \tag{A.7}$$

In this expression, Ci is the cosine integral, which can be computed numeri-

cally, and γ is Euler's constant, numerically equal to 0.577. For the realistic case $\psi = 5$; i.e. an amplifier bandwidth $5f_s$, it is found that

$$f_c \geq 1.5 f_s \tag{A.8}$$

is the condition for which $1/f$ noise dominates despite the CDS.

Bibliography

[1] A. J. Annema. Personal communication.

[2] A. J. Annema, B. Nauta, R. van Langevelde, and H. Tuinhout. Analog circuits in ultra-deep-submicron CMOS. *IEEE Journal of Solid-State Circuits*, 40(1):132–143, Jan. 2005.

[3] D. Baek, T. Song, E. Yoon, and S. Hong. 8-GHz CMOS quadrature VCO using transformer-based LC tank. *IEEE Microwave and Wireless Components Letters*, 13(10):446–448, Oct. 2003.

[4] R. C. H. van de Beek. *High-Speed Low-Jitter Frequency Multiplication in CMOS*. PhD thesis, University of Twente, 2004.

[5] R. C. H. van de Beek, E. A. M. Klumperink, C. S. Vaucher, and B. Nauta. Low-jitter clock multiplication: a comparison between PLLs and DLLs. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 49(8):555–566, Aug. 2002.

[6] I. Bloom and Y. Nemirovsky. $1/f$ noise reduction by interfering with the self correlation of the physical noisy process. *Proceedings of the National Convention of Electrical and Electronics Engineers in Israel*, 17:69–72, 1991.

[7] I. Bloom and Y. Nemirovsky. $1/f$ noise reduction of metal-oxide-semiconductor transistors by cycling from inversion to accumulation. *Applied Physics Letters*, 58(15):1664–1666, Apr. 1991.

[8] C. C. Boon, M. A. Do, K. S. Yeo, J. G. Ma, and X. L. Zhang. RF CMOS low-phase-noise LC oscillator through memory reduction tail transistor. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 51(2):85–90, Feb. 2004.

[9] R. Brederlow, W. Weber, D. Schmitt-Landsiedel, and R. Thewes. Fluctuations of the low frequency noise of MOS transistors and their modeling in analog and RF-circuits. In *IEDM Technical Digest*, pages 159–162, 1999.

[10] J. Chang, A. A. Abidi, and C. R. Viswanathan. Flicker noise in CMOS transistors from subthreshold to strong inversion at various temperatures. *IEEE Transactions on Electron Devices*, 41(11):1965–1971, Nov. 1994.

[11] R. H. M. Clevers. Volume and temperature dependency of the $1/f$ noise parameter α in Si. *Physica B*, 154(2):214–224, Apr. 1989.

[12] L. W. Couch. *Digital and Analog Communication Systems (6^{th} edition)*. Prentice Hall, 2001.

[13] H. Darabi and A. A. Abidi. Noise in RF-CMOS mixers: A simple physical model. *IEEE Journal of Solid-State Circuits*, 35(1):15–25, Jan. 2000.

[14] B. Dierickx and E. Simoen. The decrease of 'random telegraph signal' noise in metal-oxide-semicondutor field-effect transistors when cycled from inversion to accumulation. *Journal of Applied Physics*, 71(4):2028–2029, Feb. 1992.

[15] C. C. Enz, E. Vittoz, and F. Krummenacher. A CMOS chopper amplifier. *IEEE Journal of Solid-State Circuits*, 22(3):335–342, June 1987.

[16] W. A. Gardner and L. E. Franks. Characterization of cyclostationary random signal processes. *IEEE Transactions on Information Theory*, IT-21(1):4–14, Jan. 1975.

[17] J. Geerlings. *LF noise measurements under switched bias conditions with different duty cycles*. IOO report, University of Twente, 2003.

[18] G. Ghibaudo. On the theory of carrier number fluctuations in MOS devices. *Solid-State Electronics*, 32(7):563–565, 1989.

[19] S. L. J. Gierkink. *Control Linearity and Jitter of Relaxation Oscillators*. PhD thesis, University of Twente, 1999.

[20] S. L. J. Gierkink, E. A. M. Klumperink, A. P. van der Wel, G. Hoogzaad, A. J. M. 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*, 34(7):1022–1025, July 1999.

[21] A. Hajimiri and T. H. Lee. A general theory of phase noise in electrical oscillators. *IEEE Journal of Solid-State Circuits*, 33(2):179–194, Feb. 1998.

[22] A. Hajimiri, S. Limotyrakis, and T. H. Lee. Jitter and phase noise in ring oscillators. *IEEE Journal of Solid-State Circuits*, 34(6):790–804, June 1999.

[23] F. P. Heiman and G. Warfield. The effect of oxide traps on the MOS capacitance. *IEEE Transactions on Electron Devices*, ED-12:167–178, Apr. 1965.

[24] E. Hoekstra. Random telegraph noise in MOSFETs- time domain analysis under transient conditions. Master's thesis, University of Twente, Aug. 2004. Report No. 068.109.

[25] F. Hooge and P. Bobbert. On the correlation function of $1/f$ noise. *Physica B*, 239(3-4):223–230, 1997.

[26] F. N. Hooge. $1/f$ noise is no surface effect. *Physica*, 29A(3):139–140, Apr. 1969.

[27] F. N. Hooge. $1/f$ noise sources. *IEEE Transactions on Electron Devices*, 41(11):1926–1935, Nov. 1994.

[28] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng. A physics-based MOSFET noise model for circuit simulators. *IEEE Transactions on Electron Devices*, 37(5):1323–1333, May 1990.

[29] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng. A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors. *IEEE Transactions on Electron Devices*, 37(3):654–665, Mar. 1990.

[30] ITRS roadmap, 2003 edition, and 2004 update. International Technology Roadmap for Semiconductors, http://public.itrs.net, 2003,2004.

[31] R. J. Kansy. Response of a correlated double sampling circuit to $1/f$ noise. *IEEE Journal of Solid-State Circuits*, SC-15(3):373–375, June 1980.

[32] M. S. Keshner. $1/f$ noise. *Proceedings of the IEEE*, 70(3):212–218, Mar. 1982.

[33] M. J. Kirton and M. J. Uren. Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency ($1/f$) noise. *Advances In Physics*, 38(4):367–468, 1989.

[34] F. M. Klaassen. Characterization of low $1/f$ noise in MOS transistors. *IEEE Transactions on Electron Devices*, 18(10):887–891, Oct. 1971.

[35] T. G. M. Kleinpenning. On $1/f$ noise and random telegraph noise in very small electronic devices. *Physica B*, 164:331–334, Apr. 1990.

[36] E. A. M. Klumperink, S. L. J. Gierkink, A. P. van der Wel, and B. Nauta. Reducing MOSFET $1/f$ noise and power consumption by switched biasing. *IEEE Journal of Solid-State Circuits*, 35(7):994–1001, July 2000.

[37] E. A. M. Klumperink and B. Nauta. Systematic comparison of HF CMOS transconductors. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 50(10):728–741, Oct. 2003.

[38] J. S. Kolhatkar. *Steady-state and Cyclo-stationary RTS noise in MOSFETs*. PhD thesis, University of Twente, 2005.

[39] J. S. Kolhatkar, E. Hoekstra, A. J. Hof, C. Salm, J. Schmitz, and H. Wallinga. Investigating hot-carrier degradation in MOSFETs using constant and switched biased low-frequency noise measurements. In *Proceedings of the SAFE*, pages 700–703, Veldhoven, The Netherlands, Nov. 2004. STW Technology Foundation.

[40] J. S. Kolhatkar, E. Hoekstra, C. Salm, A. P. van der Wel, E. A. M. Klumperink, J. Schmitz, and H. Wallinga. Modelling of RTS noise in MOSFETs under steady-state and large-signal excitation. In *Proceedings of the IEEE International Electron Devices Meeting (IEDM)*, San Francisco, Dec. 2004.

[41] J. S. Kolhatkar, A. P. van der Wel, E. A. M. Klumperink, C. Salm, B. Nauta, and H. Wallinga. Measurement and extraction of RTS parameters under 'switched biased' conditions in MOSFETS. In J. Sikula, editor, *Proceedings of the 17th international Conference on Noise and Fluctuations (ICNF)*, pages 237–240, Prague, Czech Republic, Aug. 2003. CNRL s.r.o. ISBN: 80-239-1005-1.

[42] K. S. Kundert. Introduction to RF simulation and its application. *IEEE Journal of Solid-State Circuits*, 34(9):1298–1319, Sept. 1999.

[43] P. Larsson. A 2-1600 MHz CMOS clock recovery PLL with low-Vdd capability. *IEEE Journal of Solid-State Circuits*, 34(12):1951–1960, Dec. 1999.

[44] D. C. Lee. Analysis of jitter in phase-locked loops. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 49(11):704–711, Nov. 2002.

[45] T. H. Lee. *The Design of CMOS Radio Frequency Integrated Circuits*. Cambridge University Press, 1998.

[46] D. Leenaerts, J. van der Tang, and C. S. Vaucher. *Circuit Design for RF Transceivers*. Kluwer Academic Publishers, 2001. ISBN 0-7923-7551-3.

[47] N. B. Lukyanchikova, M. V. Petrichuk, and N. P. Garbar. Asymmetry of the RTSs capture and emission kinetics in nMOSFETS processed in a 0.35μm CMOS technology. In C. Claeys and E. Simoen, editors, *Proceedings of the 14th International Conference on Noise in Physical Systems and* $1/f$ *Fluctuations (ICNF)*, pages 232–235, Leuven, July 1997. World Scientific. ISBN 981-02-3141-5.

[48] S. Machlup. Noise in semiconductors: Spectrum of a two-parameter random signal. *Journal of Applied Physics*, 25(3):341–343, Mar. 1954.

[49] B. B. Mandelbrot. Some noises with $1/f$ spectrum, a bridge between direct current and white noise. *IEEE Transactions on Information Theory*, IT-13(2):289–298, 1967.

[50] J. A. McNeill. Jitter in ring oscillators. *IEEE Journal of Solid-State Circuits*, 32(6):870–879, June 1997.

[51] A. L. McWhorter. $1/f$ *noise and related surface effects in germanium*. PhD thesis, MIT, Cambridge, MA, 1955.

[52] G. E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38(8):114–117, Apr. 1965.

[53] A. Nemirovsky and A. Ron. A new approach to carrier trapping-detrapping $1/f$ noise. In *Proceedings of the 14th International Conference on Noise in Physical Systems and* $1/f$ *Fluctuations (ICNF)*, pages 85–88, 1997.

[54] R. G. M. Penning de Vries. *Charge Transfer Inefficiency in Surface Channel Charge-Coupled Devices*. PhD thesis, Technische Hogeschool Twente, 1984.

[55] C. T. Sah. Characteristics of the metal-oxide-semiconductor transistor. *IEEE Transactions on Electron Devices*, ED-11:324–345, July 1964.

[56] A. J. Scholten and D. B. M. Klaassen. New $1/f$ noise model in MOS model 9, level 903. Unclassified Report 816/98, Philips Research Nat. Lab., 1998.

[57] M. Schulz and H. H. Mueller. Single-electron trapping at semiconductor interfaces. *Advances in Solid State Physics*, 35:229–241, 1996.

[58] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.

[59] Z. Shi, J. Miéville, and M. Dutoit. Random telegraph signals in deep submicron n-MOSFETs. *IEEE Transactions on Electron Devices*, 41(7):1161–1168, July 1994.

[60] W. Shockley and W. T. Read, Jr. Statistics of the recombinations of holes and electrons. *Physical Review*, 87(5):835–842, Sept. 1952.

[61] H. Taub and D. L. Schilling. *Principles of communication systems*, chapter 1-3 and 1-6. McGraw-Hill Book Company, 1986.

[62] H. Tian and A. El Gamal. Analysis of $1/f$ noise in switched MOSFET circuits. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 48(2):151–157, Feb. 2001.

[63] L. K. J. Vandamme, X. Li, and D. Rigaud. $1/f$ noise in MOS devices, mobility or number fluctuations? *IEEE Transactions on Electron Devices*, 41(11):1936–1945, Nov. 1994.

[64] C. S. Vaucher. *Architectures for RF frequency synthesizers*. PhD thesis, University of Twente, 2001. Also available as a book in the "Kluwer International Series in Engineering and Computer Science", ISBN 0-306-47955-9.

[65] A. P. van der Wel, E. A. M. Klumperink, S. L. J. Gierkink, R. F. Wassenaar, and H. Wallinga. MOSFET $1/f$ noise measurement under switched bias conditions. *IEEE Electron Device Letters*, 21(1):43–46, Jan. 2000.

[66] A. P. van der Wel, E. A. M. Klumperink, and B. Nauta. Effect of switched biasing on $1/f$ noise and random telegraph signals in deep-submicron MOSFETs. *Electronics Letters*, 37(1):55–56, Jan. 2001.

[67] A. P. van der Wel, E. A. M. Klumperink, and B. Nauta. Measurement of MOSFET LF noise under large signal RF excitation. In E. G. G. Baccarani and M. Rudan, editors, *Proceedings of the 32nd European Solid-State Device Research Conference (ESSDERC)*, pages 91–95, Florence, Italy, Sept. 2002. University of Bologna.

[68] A. P. van der Wel, E. A. M. Klumperink, L. K. J. Vandamme, and B. Nauta. Modeling random telegraph noise under switched bias conditions using cyclostationary RTS noise. *IEEE Transactions on Electron Devices*, 50(5):1378–1384, May 2003.

[69] G. D. Wilk, R. M. Wallace, and J. M. Anthony. High-κ gate dielectrics: Current status and materials properties considerations. *Journal of Applied Physics*, 89(10):5243–5275, May 2001.

[70] J. Williams, editor. *The Art and Science of Analog Circuit Design*. The EDN Series for Design Engineers. Butterworth-Heinemann, 1998.

[71] P. H. Woerlee, M. J. Knitel, R. van Langevelde, D. B. M. Klaassen, L. F. Tiemeijer, A. J. Scholten, and A. T. A. Zegers-van Duijnhoven. RF-CMOS performance trends. *IEEE Transactions on Electron Devices*, 48(8):1776–1782, Aug. 2001.

[72] X. L. Xu, R. T. Kuhn, J. J. Wortman, and M. C. Öztürk. Rapid thermal chemical vapor deposition of thin silicon oxide films using silane and nitrous oxide. *Applied Physics Letters*, 60(24):6063–6065, June 1992.

[73] S. Ye, L. Jansson, and I. Galton. Techniques for in-band phase noise suppression in re-circulating DLL's. In *Proceedings of the 2003 Custom Integrated Circuits Conference*, pages 13.2.1–13.2.4. IEEE, 2003.

List of Publications

Journal papers

1. S. L. J. Gierkink, E. A. M. Klumperink, **A. P. van der Wel**, G. Hoogzaad, A. J. M. 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* 34(7):1022-1025, July 1999.

2. E. A. M. Klumperink, S. L. J. Gierkink, **A. P. van der Wel** and B. Nauta. Reducing MOSFET $1/f$ noise and power consumption by switched biasing. *IEEE Journal of Solid-State circuits* 35(7):994-1001, July 2000.

3. **A. P. van der Wel**, E. A. M. Klumperink, S. L. J. Gierkink, R. F. Wassenaar and H. Wallinga. MOSFET $1/f$ noise measurement under switched bias conditions *IEEE Electron Device Letters* 21(1):43-46, Jan. 2000.

4. **A. P. van der Wel**, E. A. M. Klumperink and B. Nauta. Effect of switched biasing on $1/f$ noise and random telegraph signals in deep-submicron MOSFETs. *Electronics Letters* 37(1):55-56, Jan. 2001.

5. **A. P. van der Wel**, E. A. M. Klumperink, L. K. J. Vandamme and B. Nauta. Modelling random telegraph noise under switched bias conditions using cyclostationary RTS noise. *IEEE Transactions on Electron Devices* 50(5):1378-1384, May 2003.

6. **A. P. van der Wel**, S. L. J. Gierkink, R. C. Frye, V. Boccuzzi, and B. Nauta. A Robust 43-GHz VCO in CMOS for OC-768 SONET applications. *IEEE Journal of Solid-State Circuits* 39(7):1159-1163, July 2004.

Conference papers

1. S. L. J. Gierkink, **A. P. van der Wel**, G. Hoogzaad, E. A. M. Klumperink and A. J. M. van Tuijl. Reduction of the $1/f$ noise induced phase noise in a CMOS ring oscillator by increasing the amplitude of oscillation. *IEEE International Symposium on Circuits and Systems (ISCAS)* Monterey, California, USA, May 31 - June 3, 1998.

2. **A. P. van der Wel**, E. A. M. Klumperink and B. Nauta. Switched Biasing' reduces $1/f$ noise in n-channel MOSFETs. *Proceedings of the ProRISC*, Veldhoven, The Netherlands, 2000.

3. **A. P. van der Wel**, E. A. M. Klumperink, B. Nauta, L. K. J. Vandamme and S. L. J. Gierkink. Simulation and measurement of RTS noise in N-channel MOSFETs under switched bias conditions. In *Proceedings of the 16th International Conference on Noise in Physical Systems and* $1/f$ *Fluctuations (ICNF)*, Gainesville, Florida, USA, pp. 391-394, Oct. 2001.

4. **A. P. Van der Wel**, E. A. M. Klumperink and B. Nauta. Simulation and measurement of n-channel MOSFET RTS noise under switched bias conditions. *Proceedings of the ProRISC*, Veldhoven, The Netherlands, 2001.

5. **A. P. van der Wel**, E. A. M. Klumperink and B. Nauta. Measurement of MOSFET LF noise under large signal RF excitation. In *Proceedings of the 32nd European Solid-State Device Research Conference (ESSDERC)*, Florence, Italy, pp. 91-95, Sept. 2002.

6. **A. P. van der Wel**, E. A. M. Klumperink and B. Nauta. Measurement of the Low Frequency Noise of MOSFETs Under Large Signal RF Excitation. *Proceedings of the ProRISC*, Veldhoven, The Netherlands, 2002.

7. J. S. Kolhatkar, **A. P. van der Wel**, E. A. M. Klumperink, C. Salm, B. Nauta and H. Wallinga. Measurement and extraction of RTS parameters under 'switched biased' conditions in MOSFETs. In *Proceedings of the 17th International Conference on Noise and Fluctuations (ICNF)*, Prague, Czech Republic, 2003.

8. **A. P. van der Wel**, S. L. J. Gierkink, R. C. Frye, V. Boccuzzi and B. Nauta. A robust 43 GHz VCO in standard CMOS for OC-768

SONET applications. In *Proceedings of the 29th European Solid-State Circuits Conference (ESSCIRC)*, Estoril, Portugal, pp. 345-348, Sept. 2003.

9. **A. P. van der Wel**, S. L. J. Gierkink, R. C. Frye, V. Boccuzzi and B. Nauta. A 43 GHz 1-V VCO in standard CMOS with high output drive capability. *Proceedings of the ProRISC*, Veldhoven, The Netherlands, 2003.

10. J. S. Kolhatkar, E. Hoekstra, C. Salm, **A. P. van der Wel**, E. A. M. Klumperink, J. Schmitz and H. Wallinga. Modelling of RTS noise in MOSFETs under steady-state and large-signal excitation. In *Proceedings of the International Electron Devices Meeting (IEDM)*, 2004.

11. **A. P. van der Wel**, E. A. M. Klumperink, J. S. Kolhatkar, E. Hoekstra and B. Nauta. Visualization techniques for random telegraph signals in MOSFETs. *Proceedings of the ProRISC*, Veldhoven, The Netherlands, 2004.

Dankwoord

Promoveren is een erg leuke bezigheid. Dat komt voor een groot deel omdat je zo veel capabele, enthousiaste en interessante mensen tegenkomt. Ik wil een aantal van hen hier bedanken.

In de eerste plaats mijn dagelijks begeleider Eric. Eric, je energie, drive, professionalisme en enthousiasme is wonderbaarlijk en heel inspirerend. Je weet als geen ander waar je het over hebt, waar we zijn en waar we naar toe willen. Je vraagt altijd de juiste dingen op de juiste momenten. Als ik dan weer in het weekend of om 00:30 een mailtje van je krijg met de mededeling 'Ik heb er nog eens over nagedacht, en ik denk dat. . . ' weet ik het weer: als ik weer zou promoveren zou ik je zo weer als begeleider kiezen.

Iets meer op de achtergrond maar zeker niet minder belangrijk is mijn promotor Bram. Bram, je bent een inspirerend figuur en ik vind het heel gaaf dat ik een aantal jaar samen met je heb mogen werken. Je bent niet alleen zeer betrokken bij de technisch-inhoudelijke kant van het werk, maar ook het managen van de vakgroep doe je met veel enthousiasme en overgave. Je bent een fijne 'baas' om voor te werken en ik hoop dat we de samenwerking in een of andere vorm nog lang voort zullen zetten.

Het werk in dit proefschrift heb ik niet alleen gedaan. Meer dan een beetje betrokken waren Jay, Cora en Hans. Als er zinnige dingen in dit proefschrift staan, zijn jullie daar medeverantwoordelijk voor.

De gebruikerscommissie kwam twee keer per jaar bijeen. Die bijeenkomsten heb ik als heel constructief en nuttig ervaren, en dat is primair te wijten aan de betrokkenheid en kunde van de leden. Jan-Wim Eikenbroek, Peter Kamp, Frank Karelse, Eduard Stikvoort, Lode Vandamme, Reinout Woltjer en Paul van Zeijl: bedankt!

Sander Gierkink en Mihai Banu wil ik bedanken voor een geweldige stage bij Agere in New Jersey. Het was niet alleen erg leerzaam maar ook nog

eens heel erg leuk!

Martijn Snoeij heeft de IC's gemaakt waar veel van de metingen in dit proefschrift op uitgevoerd zijn. Martijn, ik denk dat je een heel significante meerwaarde aan het geheel hebt gegeven door langs te komen met de vraag 'Kunnen we wat met LSE in een imager?'.

Ondersteuning bij het onderzoek kwam van vele kanten. Verschillende studenten hebben bijdragen geleverd die indirect of direct in dit proefschrift terug te vinden zijn. Bij het meetwerk was het altijd handig om Henk in de buurt te hebben, en computerproblemen werden altijd tijdig en kundig (niet altijd met een glimlach van mijn kant, vrees ik...) opgelost door Frederik en Cor. Organisatorische hulp was altijd ruim voorhanden van de vele mensen die door de jaren heen het secretariaat van de vakgroep geleid hebben: Altijd kreeg ik goede antwoorden op mijn vele vragen: 'Waar is...?', 'Hoe doen we...?', 'Hebben we...?, 'Kunnen jullie regelen dat...?' Margie, Marie-Christine, Miranda, Joke, Annemiek en Gerdien: Bedankt!

Wat promoveren op 'vloer 3' meerwaarde geeft zijn de enthousiaste collega's waar je dat mee mag doen. In de eerste plaats natuurlijk mijn kamergenoten: Svetoslav, David en Saša, en recenter: Rameswor en Eric. De andere AIO's wil ik ook bedanken voor de vele constructieve en creatieve uurtjes die we voor het whiteboard in de koffiekamer hebben doorgebracht. Mijn paranimfen Vincent en Simon: bedankt voor jullie hulp! Op het grensvlak van ADC's, software-defined radio en motorfietsen kunnen we vast nog veel leuke dingen doen! Anne-Johan, ik vond het heel plezierig samenwerken met je; ik hoop in de toekomst net zulke enthousiaste en toegewijde collega's als jij te ontmoeten. Ook de eerstejaars moet ik bedanken voor hun toch altijd weer inspirerende aanwezigheid in de collegezaal; jullie zorgden voor de broodnodige afwisseling naast het wetenschappelijke werk. Ik hoop dat jullie het ook leuk vonden...

Ook naast het werk zijn er veel mensen die veel betekend hebben voor me. Kampvuren stoken, prachtige reizen maken in en buiten Europa en heel veel dom ouwehoeren over techniek en mooie bochtjes doe ik met mijn vrienden van de Motorsportgroep. Arthur, Frans, Jelle, Wijnand: Ik hoop nog vele mooie reizen met jullie te maken!

Tot slot wil ik nog drie 'direct betrokkenen' bedanken. Aat en Gabi, jullie waren niet altijd direct in de buurt, maar ik geloof dat ons contact er kwalitatief nooit onder geleden heeft. Bedankt voor alles! Als allerlaatste en allerbelangrijkste wil ik mijn lieve vriendin Irma bedanken. Irma, je hebt geen idee hoe veel je voor me betekent.

About the Author

Arnoud van der Wel was born on the 31st of May, 1974. In 1992, he obtained his International Baccalaureate diploma from the International School Moshi in Moshi, Tanzania. He went on to study Electrical Engineering at the University of Twente in Enschede, The Netherlands, where he graduated with honours in 1997. In 2000, he graduated from the teacher's training college at the same university, making him a qualified physics teacher. In the past five years, he has been working towards the PhD degree in the IC-Design group of the University of Twente. The results of his research are contained in this thesis.