THIAGO HANNA BOTH

Autocorrelation Analysis in Frequency Domain as a Tool for MOSFET Low-Frequency Noise Characterization

Thesis presented in partial fulfillment of the requirements for the degree of PhD in Microelectronics

Advisor: Prof. Dr. Gilson Inácio Wirth

Porto Alegre
December 2017
Both, Thiago Hanna


121 f.: il.


“One person’s data is another person’s noise.”

— K.C. Cole
ACKNOWLEDGEMENT

I would like to thank my thesis advisor, Prof. Dr. Gilson I. Wirth, for directing my scientific research over the last 9 years, starting at the undergraduate level, through the master’s degree and the current thesis.

I would also like to thank NXP Semiconductors, particularly the MCSL-Eindhoven group, for the invaluable opportunity of working as an invited researcher. The staff at NXP Semiconductors in Eindhoven, the Netherlands, was extremely hospitable, providing an ideal environment for the work to flourish. Particularly noteworthy are the contributions of Dr. Andries J. Scholten and Dr. Jeroen A. Croon, who not only suggested the correlation technique presented in this work, but also knowledgeably contributed to the analyses; Adrie Zegers-van Duijnhoven, who was responsible for the low-frequency noise measurements; and Dr. Hans P. Tuinhout, whose contribution to this work far exceeds that of a colleague or a supervisor.

The assistance of my colleagues at Laprot/UFRGS also deserves mention. Dr. Fábio F. Vidor provided valuable discussions regarding the results and also revised the final draft of the thesis; Dr. Maurício Banaszeski da Silva was instrumental for the derivation of the mathematical expressions; M.Sc. Alan Carlos J. Rossetto, M.Sc. Gabriela F. Furtado, M.Sc. Pablo I. Vaz and M.Sc. Thiago Copetti also contributed with ideas and insights.

Finally, I am exceedingly thankful to my parents, Carlos Estevan and Maria Lúcia, that, through their unconditional support, allowed me to improve as a scholar and as a person; to my brother and friend, Lucas, with whom I (unwillingly) learned about sharing; and to my beloved Bruna, who has been a safe harbour and my confidant.
ABSTRACT

Low-frequency noise (LFN) is a performance limiter for analog, digital and RF circuits, introducing phase noise in oscillators and reducing the stability of SRAM cells, for example. Metal-oxide-semiconductor field-effect-transistors (MOSFETs) are known for their particularly high $1/f$ and random telegraph noise levels, whose power may be orders of magnitude larger than thermal noise for frequencies up to dozens of kHz. With the technology scaling, the corner frequency — i.e. the frequency at which the contributions of thermal and shot noises to noise power overshadow that of the $1/f$ noise — is increased, making $1/f$ and random telegraph signal (RTS) the dominant noise mechanism in CMOS technologies for frequencies up to several MHz. Additionally, the LFN levels from device-to-device can vary several orders of magnitude in deeply-scaled devices, making LFN variability a major concern in advanced MOS technologies. Therefore, to assure proper circuit design in this scenario, it is necessary to identify the fundamental mechanisms responsible for MOSFET LFN, in order to provide accurate LFN models that account not only for the average noise power, but also for its variability and dependences on geometry, bias and temperature. In this work, a new variability-based LFN analysis technique is introduced, employing the autocorrelation of multiple LFN spectra in terms of parameters such as frequency, bias and temperature. This technique reveals information about the mechanisms responsible for the $1/f$ noise that is difficult to obtain otherwise. The correlation analyses performed on three different commercial mixed-signal CMOS technologies (140-nm, 65-nm and 40-nm) provide strong evidence that the LFN of both n- and p-type MOS transistors is primarily composed of the superposition of thermally activated random telegraph signals (RTS).

Keywords: $1/f$ noise. CMOS. low-frequency noise. MOSFET. Random Telegraph Noise.
RESUMO

O ruído de baixa frequência é um limitador de desempenho em circuitos analógicos, digitais e de radiofrequência, introduzindo ruído de fase em osciladores e reduzindo a estabilidade de células SRAM, por exemplo. Transistores de efeito de campo de metal-óxido-semicondutor (MOSFETs) são conhecidos pelos elevados níveis de ruído $1/f$ e telegráfico, cuja potência pode ser ordens de magnitude maior do que a observada para ruído térmico para frequências de até dezenas de kHz. Além disso, com o avanço da tecnologia, a frequência de corner — isto é, a frequência na qual as contribuições dos ruídos térmico e shot superam a contribuição do ruído $1/f$ — aumenta, tornando os ruídos $1/f$ e telegráfico os mecanismos dominantes de ruído na tecnologia CMOS para frequências de até centenas de MHz. Mais ainda, o ruído de baixa frequência em transistores nanométricos pode variar significativamente de dispositivo para dispositivo, o que torna a variabilidade de ruído um aspecto importante para tecnologias MOS modernas. Para assegurar o projeto adequado de circuitos do ponto de vista de ruído, é necessário, portanto, identificar os mecanismos fundamentais responsáveis pelo ruído de baixa frequência em MOSFETs e desenvolver modelos capazes de considerar as dependências do ruído com geometria, polarização e temperatura. Neste trabalho é proposta uma técnica para análise de ruído de baixa frequência baseada na autocorrelação dos espectros de ruído em função de parâmetros como frequência, polarização e temperatura. A metodologia apresentada revela informações importantes sobre os mecanismos responsáveis pelo ruído $1/f$ que são difíceis de obter de outras formas. As análises de correlação realizadas em três tecnologias CMOS comerciais (140 nm, 65 nm e 45 nm) fornecem evidências contundentes de que o ruído de baixa frequência em transistores MOS tipo-n e tipo-p é composto por um somatório de sinais telegráficos termicamente ativados.

Palavras-chave: CMOS, MOSFET, Ruído $1/f$, Ruído de Baixa Frequência, Ruído Telegráfico.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>Autocorrelation Function</td>
</tr>
<tr>
<td>ACM</td>
<td>Advanced Compact MOSFET model</td>
</tr>
<tr>
<td>BSIM</td>
<td>Berkeley Short-Channel IGFET model</td>
</tr>
<tr>
<td>BTI</td>
<td>Bias Temperature Instability</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary Metal-Oxide-Semiconductor</td>
</tr>
<tr>
<td>DUT</td>
<td>Device Under Test</td>
</tr>
<tr>
<td>EKV</td>
<td>Enz-Krummenacher-Vittoz MOSFET model</td>
</tr>
<tr>
<td>FinFET</td>
<td>Fin Field-Effect Transistor</td>
</tr>
<tr>
<td>GR</td>
<td>Generation-Recombination</td>
</tr>
<tr>
<td>HEMT</td>
<td>High-Electron-Mobility Transistor</td>
</tr>
<tr>
<td>IGFET</td>
<td>Insulated-Gate Field-Effect Transistor</td>
</tr>
<tr>
<td>LDD</td>
<td>Light Doped Drain</td>
</tr>
<tr>
<td>LFN</td>
<td>Low-Frequency Noise</td>
</tr>
<tr>
<td>MOSFET</td>
<td>Metal-Oxide-Semiconductor Field-Effect Transistor</td>
</tr>
<tr>
<td>NFET</td>
<td>n-type Metal-Oxide-Semiconductor Field-Effect Transistor</td>
</tr>
<tr>
<td>PFET</td>
<td>p-type Metal-Oxide-Semiconductor Field-Effect Transistor</td>
</tr>
<tr>
<td>PSD</td>
<td>Power Spectral Density</td>
</tr>
<tr>
<td>PSP</td>
<td>Penn-State Phillip MOSFET model</td>
</tr>
<tr>
<td>RTN</td>
<td>Random Telegraph Noise</td>
</tr>
<tr>
<td>RTS</td>
<td>Random Telegraph Signal</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SRAM</td>
<td>Static Random Access Memory</td>
</tr>
<tr>
<td>WKB</td>
<td>Wentzel-Kramers-Brillouin approximation</td>
</tr>
</tbody>
</table>
LIST OF SYMBOLS

\( C_{ox} \)  Gate oxide capacitance per unit area

\( E \)  Energy

\( E_B \)  Activation Energy

\( E_c \)  Conduction band energy level

\( E_f \)  Fermi Energy

\( E_{Fn} \)  Electron quasi Fermi energy

\( E_i \)  Intrinsic Fermi energy

\( E_{ox} \)  Electric field across the oxide

\( E_{sat} \)  Electric field at which carrier velocity saturates

\( E_T \)  Trap energy level

\( E_{T0} \)  Trap energy in flat band conditions

\( E_v \)  Valence band energy level

\( E_x \)  Electric field along the \( x \)-axis

\( f \)  Frequency in Hz

\( f_H \)  Upper limit of the bandwidth

\( f_L \)  Lower limit of the bandwidth

\( f_{\text{max}} \)  Upper limit of the observation window

\( f_{\text{min}} \)  Lower limit of the observation window

\( f_{\text{ref}} \)  Reference frequency for autocorrelation analysis

\( f_{\text{rolloff}} \)  Measurement system roll-off frequency

\( f_t \)  Fermi-Dirac distribution

\( g_{ds} \)  Drain-to-source transconductance

\( g_m \)  Gate-to-source transconductance

\( g_{mbs} \)  Substrate-to-source transconductance
$k$ Boltzmann constant in eV·K$^{-1}$

$k_B$ Boltzmann constant in J·K$^{-1}$

$K$ Normalized variance

$h$ Planck constant in J·s

$I$ Electric current

$I_D$ Drain current

$I_0$ Dark saturation current

$L$ Transistor Length

$L_{eff}$ Effective channel length

$N$ Surface carrier concentration

Total number of traps in a device

$N_1$ Estimated surface carrier concentration at the drain side

$N_t$ Surface trap density of occupied traps

$N_{tr}$ Surface trap density in area and energy

$N_{tr}^*$ Equivalent trap density in area and energy incorporating mobility fluctuations

$N_{tr}/\gamma$ Surface trap density in energy, area, and log-domain of time constants

$N_{t, empty}$ Average number of empty traps

$N_{t, full}$ Average number of occupied

$N_0$ Estimated surface carrier concentration at the source side

$q$ Elementary charge

$Q_{inv}$ Inversion layer charge

$R$ Resistance

Coupling coefficient between carrier number fluctuation and trap occupancy

$R_{DS}$ Drain-to-source resistance

$R_X$ Autocorrelation function of $X(t)$

$S_{ItD}$ Power spectral density of the drain current noise
$S_{ID,\text{inv}}$ 1/f Noise power spectral density estimated in strong inversion
$S_{ID,\text{subV}\text{t}}$ 1/f Noise power spectral density estimated in subthreshold region
$S_N$ Power spectral density of the fluctuations in the charge carriers concentration
$S_{VG}$ Gate referred noise power spectral density
$S_X$ Power spectral density of $X(t)$
$T$ Temperature
$T_{\text{ox}}$ Gate oxide thickness
$T_{\text{ref}}$ Reference Temperature for autocorrelation analysis
$V_D$ Potential applied across a semiconductor junction
$V_{DS}$ Drain-to-source voltage
$V_{DS,\text{sat}}$ Saturation drain-to-source voltage
$V_{dseff}$ Effective drain-to-source voltage
$V_{FB}$ Flat band voltage
$V_{GS}$ Gate-to-substrate voltage
$V_{GS}$ Gate-to-source voltage
$V_{gsteff}$ Effective gate-to-source voltage
$V_{GS,\text{ref}}$ Reference gate voltage for autocorrelation analysis
$V_T$ Estimated threshold voltage of a device population
$V_{th0}$ Threshold voltage of the long-channel device at zero substrate bias
$W$ Transistor Width
$W_p$ Noise power integrated over a bandwidth
$x_0$ Location parameter of the Cauchy-Lorentz distribution
$x_{tr}$ Trap position along the channel
$X_{tr}$ Uniformly distributed trap position along the channel
$z_{tr}$ Trap depth within the oxide
$\alpha$ Scattering coefficient
$\alpha_H$  Hooge parameter
$\gamma$  Wave function attenuation factor
$\gamma'$  Log-uniform distribution parameter
$\Delta f$  Difference between $f_1$ and $f_2$
$\Delta f'$  Ratio between $f_1$ and $f_2$
$\Delta I$  Difference between high and low current states
$\Delta I_D$  Drain current difference between occupied and empty trap states
$\Delta I_D'$  Normalized drain current deviation
$\epsilon$  MOSFET thermal noise bias dependent parameter
$\lambda$  Frequency exponent of the low-frequency noise power spectral density
$\mu$  Charge carrier mobility
$\mu_{eff}$  Effective charge carrier mobility
$\rho$  Correlation coefficient
$\sigma$  Scale parameter of the log-normal distribution
$\tau$  Characteristic time constant
$\tau_c$  Average capture time constant
$\tau_e$  Average emission time constant
$\tau_{max}$  Maximum characteristic time constant in the defined bandwidth
$\tau_{min}$  Minimum characteristic time constant in the defined bandwidth
$\tau_o$  Lowest possible characteristic time constant
$\tau_0$  Time spent at low current state
$\tau_1$  Time spent at high current state
$\phi_s$  Surface potential
$\psi$  Scale parameter of the Cauchy-Lorentz distribution
$\omega$  Angular frequency
LIST OF FIGURES

Figure 3.1 Schematics for the BSIM4 channel thermal noise modelling. ..........36
Figure 3.2 Basic processes involved in the generation and recombination of carriers through traps. .................................................................39
Figure 3.3 Example of a Random Telegraph Signal. ..............................................40
Figure 3.4 Measured LFN PSD of a large-area NFET (140-nm technology)........42
Figure 4.1 Sum of log-uniformly spaced Lorentzians. .........................................47
Figure 4.2 Behaviour of the $f_t(1 - f_t)$ term for different quasi Fermi levels. ......49
Figure 4.3 Behaviour of the $f_t(1 - f_t)$ as a function of energy and temperature. 49
Figure 5.1 Noise PSD and measured frequency exponent as a function of temperature for a PFET.................................................................61
Figure 5.2 Frequency exponent as a function of temperature for NFETs with different insulators. .................................................................62
Figure 5.3 Noise PSD for both N- and PFETs at different temperatures in linear region. .................................................................63
Figure 5.4 Noise PSD of an NFET at 1 Hz and frequency exponent as a function of temperature. .................................................................63
Figure 5.5 $1/f$ noise spectra of N- and PFETs at different temperatures. ...............64
Figure 5.6 Noise PSD at 10 Hz for two FINFETs in linear region. ............................65
Figure 5.7 Frequency exponent as a function of temperature at 10 Hz for two FINFETs. .................................................................65
Figure 6.1 EDGE™ low-frequency noise measurement system............................69
Figure 6.2 Schematic of low-frequency noise test setup.........................................70
Figure 6.3 Measured and smoothed LFN spectra for small-area and large-area NFETs.................................................................70
Figure 6.4 Quantile plot of the LFN PSD of N- and PFET populations (140-nm technology). .................................................................72
Figure 6.5 Quantile plot of the LFN PSD of N- and PFET populations (40-nm technology). .................................................................72
Figure 6.6 Quantile plot of the LFN PSD of NFET populations (65-nm technology). .................................................................72
Figure 6.7 Quantile plots of the LFN PSD at 20 Hz for two 160-DUT NFET population (40-nm node). .................................................................73
Figure 6.8 Simulated and measured quantile plots of the LFN PSD of a small-area NFET population.................................................................74
Figure 6.9 Expected value of the LFN PSD as a function of bias for two NFET populations (140-nm technology). .................................................................74
Figure 6.10 Expected value of the LFN PSD as a function of bias for two NFET populations (140-nm technology). .................................................................75
Figure 6.11 Expected value of the LFN PSD as a function of bias for two PFET populations (140-nm technology). .................................................................76
Figure 6.12 Expected value of the LFN PSD as a function of bias for two PFET populations (140-nm technology). .................................................................76
Figure 6.13 Standard deviation of $\ln(S_{f_t}/I_0^2)$ as a function of device geometry. .........76
Figure 6.14 K parameter as a function of area for 10 NFET populations.................77
| Figure 6.15 | Standard deviation of the noise PSD for two NFET populations (140-nm technology). | 77 |
| Figure 6.16 | Expected value of the LFN PSD for five NFET populations (65-nm technology). | 78 |
| Figure 6.17 | K parameter as a function of normalized drain current for four NFET populations (65-nm technology). | 78 |
| Figure 6.18 | Expected value of the LFN PSD for two NFET populations (40-nm technology). | 79 |
| Figure 6.19 | Expected value of the LFN PSD for two NFET populations (40-nm technology). | 79 |
| Figure 6.20 | Expected value of the LFN PSD as a function of drain current for two NFET populations (40-nm technology). | 80 |
| Figure 6.21 | Standard deviation of the LFN PSD for five NFET and PFET populations. | 80 |
| Figure 6.22 | Scatter plot of the LFN PSD versus drain current for two NFET populations (40-nm technology). | 81 |
| Figure 6.23 | Average value of $f \cdot S_{I_D}/I_{D}^2$ for different frequencies and bias for a 282-DUT NFET population (65-nm technology). | 82 |
| Figure 6.24 | Variance of $f \cdot S_{I_D}/I_{D}^2$ for different frequencies and bias for a 282-DUT NFET population (65-nm technology). | 83 |
| Figure 6.25 | Correlation coefficient of $S_{I_D}/I_{D}^2$ between experiments 1 and 2 as a function of frequency. | 83 |
| Figure 6.26 | Expected value of the integrated noise PSD over different bandwidths (40-nm technology). | 84 |
| Figure 6.27 | Expected value of the integrated noise PSD over different bandwidths (65-nm technology). | 85 |
| Figure 6.28 | Expected value of the integrated noise PSD over different bandwidths (140-nm technology). | 85 |
| Figure 6.29 | LFN spectra of two NFET devices at the same bias conditions and at two different temperatures. | 86 |
| Figure 6.30 | LFN spectra of an NFET device at the same bias conditions and at two different temperatures. | 87 |
| Figure 6.31 | LFN spectra of two NFET devices at the same bias conditions and at two different temperatures. | 87 |
| Figure 6.32 | LFN spectra of a device at the same bias conditions at two different temperatures. No visible Lorentzians. | 88 |
| Figure 6.33 | LFN spectra of a device at the same bias conditions at two different temperatures. No visible Lorentzians. | 88 |
| Figure 6.34 | Cumulative probability plot of $\log_{10} \tau$. | 88 |
| Figure 6.35 | Scatter plot of $\log_{10} \tau$ at two different temperatures. | 89 |
| Figure 6.36 | Cumulative probability plot of the activation energy extracted from (6.9). | 89 |
| Figure 7.1 | Correlation coefficient as a function of $\Delta f'$. | 94 |
| Figure 7.2 | LFN spectra of 320 small-area devices. | 95 |
| Figure 7.3 | Correlation coefficient of the LFN PSD as a function of frequency calculated theoretically and extracted from a 320-DUT NFET population (40-nm technology). | 96 |
| Figure 7.4 | LFN spectra and frequency correlation coefficient for a small-area NFET population. | 96 |
LIST OF TABLES

Table 4.1 BSIM4 Noise Simulation Parameters........................................................54
1 INTRODUCTION

The word *noise* is used to describe a plethora of phenomena that are (usually) unwanted. Commonly, *noise* refers to acoustic noise, which is characterized by disturbing, often loud, sounds, caused by e.g. traffic, aircraft, machines and bad musicians. Physically, noise and sound are indistinguishable; both are propagating mechanical waves of pressure. There is, therefore, an innate subjectivity in acoustic noise, i.e. what is agreeable and what is not to human hearing. The 1980 track by Australian band AC/DC, entitled *Rock and Roll ain’t noise pollution*, plays on this idea by implying that their music, though loud and disturbing to some, is not *noise*.

Even though acoustic noise is certainly the most recognizable form of *noise*, several other fields — such as electronics, physics, biology and economy — are familiarized with the concept. And while the word *noise* usually carries the stigma of unwantedness, it is certainly not always the case: in computer graphics, for example, gradient and value noises (e.g. Perlin Noise) are widely employed for procedural generation of textures; in physics, Albert Einstein demonstrated in one of his works during the *Annus Mirabilis* of 1905 that the Avogadro number could be derived from a stochastic quantity (Brownian motion) (COHEN, 2005).

In electronics, however, *noise* is understood as undesirable fluctuations in measurable quantities, such as current and voltage. These fluctuations carry no meaningful information, and may obscure a signal. Using the acoustic analogy, imagine trying to convey a message to a friend on the other side of a bar. If the bar is empty and silent, it is easy to convey the message as the speech is easily audible. On the other hand, if the bar is crowded and loud, transmitting the information becomes difficult, meaning the transmitter may need to shout. In electronics the same is valid: transmitting a signal through a noisy medium requires a larger signal, i.e. more signal power.

Interestingly, the randomness and unpredictability are fundamental properties of both information and noise (HAYKIN; MOHER, 2007). If one could predict information, there would be no need to communicate it, as the receiver would be able to foresee it. Similarly, were the receiver able to predict noise, then the receiver would be able to remove the noise, negating its effects. The goal is, thus, to separate noise from information to the maximum extent possible (HAYKIN; MOHER, 2007).

Regarding its origins, noise in electronics can rise from external and internal sources. External sources cause deterministic disturbances due to electromagnetic and
electrostatic effects. One common example of this kind of noise is the crosstalk between telephone lines; the coupling between the lines causes the signal from one line to disturb another. This disturbance is not truly random, but may obscure the information on the line nonetheless. Internal sources, on the other hand, are stochastic processes, i.e. the value of the disturbance is random at any given time. The most fundamental example of internal noise source in electronics is the thermal noise (also known as Johnson or Nyquist noise), caused by the random-walk of electrons due to thermal agitation. Chapter 3 of this work provides a deeper discussion on the fundamental noise sources in electronics.

Due to its stochastic (i.e. random) nature, it is convenient to describe noise from internal sources using statistical quantities. While observing the probability density function and the autocorrelation function of noise signals are possible methods, the primary metric for noise analysis used in this work is the power spectral density (PSD), which provides the amount of noise power per unit frequency. It is possible to categorize noise based on its PSD behaviour; this classification is referred to as colours of noise. Typically, noise can be white, pink, blue, brown and violet. Noise sources that have a constant power over a bandwidth (that is, the power is independent of the frequency) are called white; those whose power decreases with $1/f$ and $1/f^2$ factors over the bandwidth are called pink and brown, respectively; those whose power increases with $f$ and $f^2$ factors over the bandwidth are called blue and violet, respectively. This work focus mainly on pink noise, also known as $1/f$ noise or flicker noise.

Fluctuations with a $1/f$ power spectrum have been reported from voltages and currents in diodes and vacuum tubes; to frequency of quartz oscillators; average seasonal temperature; annual amount of rainfall; rate of traffic flow; rate of insulin take by diabetics; economic data; loudness and pitch of music; flow of the river Nile; and luminosity of stars (KESHNER, 1982; BAK; TANG; WIESENFELD, 1987). The observation of $1/f$ noise in a multitude of systems has led Bak, Tang and Wiesenfeld (1987) to propose a model for ubiquitous $1/f$ noise based on self-organized criticality, though Jensen, Christensen and Fogedby (1989) claimed that such model produces a $1/f^2$ spectrum.

Due to the $1/f$ decay of pink noise, lower frequencies contribute more significantly to noise power than higher frequencies. Thus, $1/f$ noise is one of the major components of low-frequency noise (LFN) in integrated circuits. Other important components of LFN are the random telegraph signal (RTS) and generation-recombination noise. These components will be further discussed in Chapters 3 and 4.

From a designers perspective, low-frequency noise is a performance limiter for
analog, digital and RF circuits. Providing accurate LFN models that account not only for the average noise power but also for variability and its dependences on geometry, bias and temperature is therefore imperative for proper circuit design.

Historically, two main schools of thought have emerged to explain the $1/f$ behaviour of the LFN power-spectral density in metal-oxide-semiconductor field-effect transistors (MOSFETs). The first assumes the origin of the LFN is related to the capture and emission of charge carriers by defects (traps) (MCWHORTER, 1957), and it is commonly referred to as number fluctuation theory. The second attributes the LFN to bulk mobility fluctuations (HOOG, 1969), being referred to as mobility fluctuation theory.

While trapping/de-trapping has emerged as the prevalent model in recent years, the trapping mechanism itself is not thoroughly understood. While studies on random telegraph signals and $1/f$ noise (SURYA; HSIANG, 1988; KIRTON; UREN, 1989) suggest a thermally activated process, direct tunnelling (MCWHORTER, 1957), which is essentially temperature independent, is still the basis of the LFN in compact models, such as BSIM and PSP (HUNG et al., 1990b; HUNG et al., 1990c).

In this work, a novel technique, based on the analysis of the autocorrelation of numerous LFN power-spectra density measurements, is introduced. The analyses reveal the statistical relationship between the LFN levels at a given frequency and temperature compared to their levels under different conditions. It is shown that the correlation coefficient is extremely sensitive to the physical mechanisms assumed for interpreting and modelling the LFN PSD. Consequently, it proves possible to draw valuable conclusions about the frequency dependence of the fundamental noise sources causing the $1/f$ noise, as well as the origin of the trapping mechanism (tunnelling/thermal activation).

The analyses presented in this work were performed on measured LFN data from n- and p-type MOS transistors in 140-nm, 65-nm and 40-nm mixed-signal CMOS technologies. The results provide strong evidence for a unified description of N- and PFET LFN in terms of thermally activated random telegraph signals, providing a statistical approach to the classical work of Dutta and Horn (1981).

This work is organized as follows: the introductory chapters 2 and 3 provide a mathematical background for noise analysis, and a brief discussion on the fundamental noise sources in electronics, respectively. An in-depth review of MOSFET LFN modelling is presented in chapter 4, followed by a discussion on the temperature dependence of MOSFET LFN in chapter 5. In chapter 6, the noise measurement set-up as well as important noise quantities (such as expected value and variance) from commercial mixed-
signal CMOS technologies are displayed. The autocorrelation technique proposed for LFN analysis is discussed in chapter 7 and applied for several case studies. The closing remarks of this work are presented in chapter 8.
2 MATHEMATICAL REPRESENTATION OF NOISE

The study of noise in electronics consists in the analysis of random, unwanted, electric signals in devices, circuits and systems, aiming at providing an useful description of these processes for circuit designers. However, due to the stochastic nature of most noise processes (e.g. thermal noise), detailed knowledge about the time variation of a random signal is unavailable, meaning that it is not possible to predict exactly what waveform will be observed in the future.

For this reason, to deal with noise one must speak in terms of probabilities and statistical properties in order to convey a useful description of the process. The purpose of this chapter is to provide a brief introduction on the mathematical description of noise, in particular the definition of the power spectral density, which is the primary noise quantity analysed in this work.

The discussion in this chapter will consider only real-valued random processes.

2.1 Random Processes

A random process or stochastic process, \( X(t) \), is a time-varying function which, due to randomness, can have many outcomes. Each outcome is a time function known as sample function, \( x(t) \). The collection of all possible outcomes of an experiment is called the sample space or ensemble (CARLSON; CRILLY, 2010; LATHI; DING, 2009).

When the statistical characteristics of the random process are independent of time, the process is described as being stationary. It is said to be strictly stationary if the joint distribution of any set of random variables obtained by observing the random process \( X(t) \) is independent of the location of the origin \( t = 0 \) (HAYKIN; MOHER, 2007), and wide-sense stationary if the average is independent of time and the autocorrelation function, \( R_X(t,t+s) \), depends only on the time difference \( s \) (CARLSON; CRILLY, 2010), that is

\[
E[X(t)] = \text{constant} \tag{2.1}
\]

and

\[
R_X(t,t+s) = R(s) \,, \tag{2.2}
\]

where \( E \) is the expected value operator.
2.2 Autocorrelation Function of a Random Process

Even though random processes are unpredictable by definition, similarities between samples of the process at different times may be observed, meaning they are correlated (HAYKIN; MOHER, 2007). The autocorrelation function (ACF) is the mathematical description of the similarity between observations of a random process with a certain lag in time. In other words, it describes the rapidity of the amplitude change of a signal in time (LATHI; DING, 2009).

For a generic wide-sense stationary random process, $X(t)$, the autocorrelation function between times $t$ and $s$, $R_X(t, t+s)$, is given by

$$ R_X = E[X(t) \cdot X(t+s)] \ . \quad (2.3) $$

If the ensemble average of the random process $X(t)$, $E[X(t)]$ is equal to the time average of any sample function, $\langle x(t) \rangle$; and the autocorrelation function of the process, $R_X(t, t+s)$, is equal to the time autocorrelation function of any sample function, so that

$$ E[X(t)] = \langle x(t) \rangle \ , \quad (2.4) $$

$$ E[X(t) \cdot X(t+s)] = \langle x(t) \cdot x(t+s) \rangle \ , \quad (2.5) $$
the process is defined as *ergodic*.

For an ergodic wide-sense stationary random process

$$ R_X(s) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot x(t+s)dt \ . \quad (2.6) $$

It is worth noting that the ACF of a wide-sense stationary random process is maximum at the origins, i.e. $R_X(0) \geq |R_X(s)|$, and that its value at the origin is equal to the mean-square value of the random process, i.e. $R_X(0) = E[X(t)^2]$ (HAYKIN; MOHER, 2007).
2.3 Power Spectral Density

The power spectral density describes the amount of power per unit frequency (Hz) of a given signal or random process. In order to determine the PSD of a random process from the sample function \( x(t) \), it is necessary to define its Fourier transform. However, since power signals have infinite energy, the sample function may not be Fourier transformable (HAYKIN; MOHER, 2007). To overcome this limitation, the following steps are taken: consider the finite-duration (truncated) version of the sample function of an ergodic stationary process, given by

\[
x_T(t) = \begin{cases} 
  x(t), & \text{if } |t| < T/2 \\
  0, & \text{if } |t| > T/2 .
\end{cases}
\] (2.7)

As long as \( T \) is finite, the truncated sample function, \( x_T(t) \) has finite energy, which enables the determination of its Fourier transform \( X_T(f) \). The energy spectral density of \( x_T(t) \) is given by \( |X_T(f)|^2 \) and the average power, \( P_X \), from Rayleigh’s energy theorem (Parseval’s Theorem), in the form of

\[
\frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt = \int x_T(t) dt = \int |X_T(f)|^2 df,
\] (2.8)

is given by

\[
P_X = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} E[x_T(t)^2] \, dt = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} E[|X_T(f)|^2] \, df .
\] (2.9)

Notice that, as \( T \) increases, the energy in \( x_T(t) \) also increases. Correspondingly, \( |X_T(f)|^2 \) increases with \( T \) and, as \( T \to \infty \), so will \( |X_T(f)|^2 \to \infty \). However, since that average power has to be finite, \( |X_T(f)|^2 \) must approach infinity at the same rate as \( T \). This convergence allows the interchange between the limiting operation and the integration in the right-hand side of (2.9) (CARLSON; CRILLY, 2010; HAYKIN; MOHER, 2007; LATHI; DING, 2009), that is

\[
P_X = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} E[|X_T(f)|^2] \, df = \int \lim_{T \to \infty} \frac{1}{T} E[|X_T(f)|^2] \, df .
\] (2.10)
The power spectral density is then defined as

\[ S_X = \lim_{T \to \infty} \frac{1}{T} E \left[ |X_T(f)|^2 \right] , \quad (2.11) \]

and, consequently,

\[ P_X = \int_{-\infty}^{\infty} S_X(f) df . \quad (2.12) \]

Following to the Wiener-Khinchine theorem (HAYKIN; MOHER, 2007; LATHI; DING, 2009; CARLSON; CRILLY, 2010), the PSD of a random process, \( S_X(f) \), is related to the autocorrelation function of the process, \( R_X(\tau) \), by the Fourier transform

\[ S_X(f) = \int_{-\infty}^{\infty} R_X(s) e^{-j2\pi f \tau} ds , \quad (2.13) \]

\[ R_X(s) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f \tau} df . \quad (2.14) \]

which may be simplified for even, real-valued, functions in the form of

\[ S_X(f) = \int_{-\infty}^{\infty} R_X(s) \cdot \cos(2\pi fs) ds , \quad (2.15) \]

\[ R_X(s) = \int_{-\infty}^{\infty} S_X(f) \cdot \cos(2\pi fs) df . \quad (2.16) \]

Even though the PSD and the ACF contain, essentially, the same information about the noise source, the former is more commonly used to describe its behaviour. The reasons for that include the convenience of working in frequency-domain and the ease of determining the power within a bandwidth, required for signal-to-noise ratio (SNR) estimation.

### 2.3.1 Power Spectral Density Estimation

The problem of PSD estimation consists of obtaining, from a finite record of stationary data sequence, how the total power is distributed over frequency. There are two broad approaches to spectral analysis: the first, called non-parametric, is based entirely
on the definition of the PSD (2.11), which, for discrete-time data can be rewritten as

\[ S_X = \lim_{N \to \infty} \frac{1}{N} E[X[k]^2]; \tag{2.17} \]

the second, postulates a model for the data, providing a means to parametrize the spectrum, thereby reducing the spectral estimation problem to that of estimating the parameters in the assumed model.

The more common non-parametric estimators are the periodogram and the correlogram. The periodogram relies on the definition (2.17), neglecting the expected value and limit operations, such that

\[ \tilde{S}_X = \frac{1}{N} \sum_{n=1}^{N} x[n] \cdot e^{-j2\pi fn}, \tag{2.18} \]

where \( x[n] \) and \( \tilde{S}_X \) are the sampled signal, and the PSD estimate, respectively.

The correlogram, on the other hand, uses the definition (2.13), so that

\[ \tilde{S}_X = \sum_{k=-(N-1)}^{N-1} R[k] \cdot e^{-j2\pi k}, \tag{2.19} \]

where \( R[k] \) and \( \tilde{S}_X \) are covariance lag obtained from the sampled function (equivalent to the autocorrelation function), and the PSD estimate.

Notice, however, that both the periodogram and the correlogram are poor estimates of the PSD (STOICA; MOSES, 1997). To circumvent the natural limitations of the periodogram, several periodogram-based methodologies have been proposed. Examples include the Bartlett (1950) method — which consists of splitting the available sample of \( N \) observations into \( L = N/M \) subsamples of \( M \) observations each, and then averaging the periodograms obtained from the subsamples for each value of \( f \) —, and the Welch (1967) method — which is similar to the Bartlett method, but the data segments are allowed to overlap.
3 FUNDAMENTAL NOISE SOURCES IN ELECTRONICS

In order to properly address and model noise in integrated circuits it is necessary to identify the individual noise sources in electronic devices. Primarily, disturbances can originate from external sources — due to electrostatic and electromagnetic interaction — and from internal sources, i.e. random fluctuations in the physical mechanisms that govern electron transport. In this chapter, and henceforth in this work, the term noise will refer to the latter (internal sources).

The following sections provide a background on the fundamental noise sources present in electronics, which are: thermal noise, shot noise, generation-recombination noise (GR noise), random telegraph signal and 1/f noise.

3.1 Thermal Noise

The thermal noise is a broadband white noise — i.e. it has a constant power spectral density — that arises from the random motion of electron and is present in any passive resistor above absolute zero temperature, even without bias applied. Experimentally reported by Johnson (1928) and theoretically explained by Nyquist (1928), the thermal noise is also known as Johnson or Nyquist noise.

Each time an electron is scattered, its velocity is randomized; therefore, instantly, there could be a number of electrons flowing in a given direction, resulting in a small net current. Due to its random characteristic, this current fluctuates in amplitude and direction, but its expected value is always zero (HAARTMAN; ÖSTLING, 2007). The PSD of the thermal noise of a piece of material with resistance $R$ and at a temperature $T$ can be written as

$$S_I = \frac{4k_BT}{R},$$

(3.1)

where $k_B$ is the Boltzmann constant in Joules per Kelvin.

Notice, from (3.1), that the thermal noise in an ideal resistor is white, i.e. its intensity is constant throughout the frequency spectra. This approximation, however, is only valid if $k_BT \gg hf$ (where $h$ and $f$ are the Planck constant in J-s and $f$ in Hz, respectively). At low temperatures and high frequencies, quantum effects become relevant...
and the more general expression

\[ S_I = 4 \cdot \frac{h f}{\exp \left( \frac{h f}{k_B T} \right) - 1} \cdot \frac{1}{R}, \]  

(3.2)

proposed by Nyquist (1928), must be used. Note that if \( k_B T \gg h f \), the quantum correction factor reduces to \( k_B T \), resulting in equation (3.1).

For long-channel MOS transistors, the thermal noise can be estimated using

\[ S_{I_D} = 4k_B T \varepsilon g_m \]  

(3.3)

where \( g_m \) is the small-signal transconductance and at the bias point and \( \varepsilon \) is a bias dependent parameter (\( \varepsilon = 1/2 \) for weak inversion and \( \varepsilon = 2/3 \) for strong inversion) (TEDJA; SPIEGEL; WILLIAMS, 1994).

From a transistor modelling perspective, the BSIM4 equation model provides two methods for channel thermal noise calculation: a charge-based model and an holistic model, whose schematics are shown in figure 3.1.

Figure 3.1: Schematics for the BSIM4 channel thermal noise modelling for the a) charge-based model and b) holistic model.

![Schematics for the BSIM4 channel thermal noise modelling](image)

(a) \( \text{tnoiMod} = 0 \)  
(b) \( \text{tnoiMod} = 1 \)


The charge-based model estimates the noise current using

\[ \overline{i_d^2} = \frac{4k_B T \Delta f}{R_{DS}(V) + \frac{i_{eff}^2}{\mu_{eff} Q_{inv}}} \cdot NTNOI, \]  

(3.4)

where \( \Delta f \) is the system bandwidth, \( R_{DS}(V) \) is the bias-dependent light doped drain (LDD) source-drain resistance, \( L_{eff} \) is the effective channel length, \( \mu_{eff} \) is the effective carrier mobility, \( NTNOI \) is a fitting parameter for short-channel devices, and \( Q_{inv} \) is the inversion layer charge.

Conversely, the holistic model estimates the thermal noise incorporating short-channel and velocity saturation effects, the thermal noise amplification through gate and substrate transconductances (\( g_m \) and \( g_{mbs} \), respectively), as well as the induced-gate noise.
with partial correlation to the channel thermal noise. The noise voltage partitioned to the source is given by

\[ \overline{v_d^2} = 4k_B T \theta_{tnoi}^2 \frac{V_{dseff} \Delta f}{I_{ds}} \]  

(3.5)

whereas the noise current source in the channel region is given by

\[ \overline{i_d^2} = 4k_B T \frac{V_{dseff} \Delta f}{I_{ds}} \cdot [g_{ds} + \beta_{tnoi} \cdot (g_m + g_{mbs})]^2 - \overline{v_d^2} \cdot (g_m + g_{ds} + g_{mbs})^2 \]  

(3.6)

where

\[ \theta_{tnoi} = RNOIB \cdot \left[ 1 + TNOIB \cdot L_{eff} \cdot \left( \frac{V_{gsteff}}{E_{sat} L_{eff}} \right)^2 \right] \]  

(3.7)

\[ \beta_{tnoi} = RNOIA \cdot \left[ 1 + TNOIA \cdot L_{eff} \cdot \left( \frac{V_{gsteff}}{E_{sat} L_{eff}} \right)^2 \right] \]  

(3.8)

and \( V_{dseff} \) is the effective voltage between drain and source; \( I_{ds} \) is the drain current; \( g_{ds} \) is the drain transconductance; \( V_{gsteff} \) is the the effective gate voltage subtracted of the threshold voltage; \( E_{sat} \) is the critical electric field at which the carrier velocity becomes saturated; \( TNOIA \) and \( TNOIB \) are the channel-length dependence of the total channel thermal noise and the channel thermal noise partitioning, respectively; and \( RNOIA \) and \( RNOIB \) are model parameters (DUNGA et al., 2006).

From a designer’s perspective, thermal noise from resistive elements is unavoidable, but it is possible to minimize it through clever circuit design. Reactive elements, for example, do not generate thermal noise. Therefore, input matching techniques using reactive elements can be used to reduce thermal noise. Also, narrowing the system bandwidth reduces the thermal noise power. Finally, noise cancelling techniques have also been proposed to mitigate the thermal noise (BRUCCOLERI; KLUMPERINK; NAUTA, 2004).

### 3.2 Shot Noise

Due to the discrete nature of charge carriers (electrons), which have a quantized charge equal to the elementary charge, the current flow through a potential barrier, like a \( pn \) junction is not continuous; thus, shot noise current is generated when electrons flow independently and at random through the barrier (HAARTMAN; ÖSTLING, 2007). Simply put, if one could "observe" the flow of carriers through a barrier for an ensemble of
time intervals, the average number of carriers flowing in each interval would represent the DC current whereas the deviation from this average for each time interval would represent the shot noise current.

This effect was first described in the context of vacuum tubes by Schottky (1918). Since the shot noise is related to the corpuscular nature of the current flow through a potential barrier, a current is imperative for the existence of shot noise. The PSD of the shot noise can be expressed as

\[ S_I = 2qI \]  

where \( I \) is the DC current across the barrier and \( q \) is the elementary electron charge. From equation (3.9), the shot noise PSD is constant with frequency, resulting in a white noise spectrum.

For an ideal \( pn \) junction, the current flowing through the potential barrier is given by

\[ I = I_0 \left[ \exp \left( \frac{qV_D}{k_B T} \right) - 1 \right] \]  

where \( I_0 \) is the dark saturation current and \( V_D \) is the potential applied across the junction. From equation (3.10), when no bias is applied, the forward and backward components of the current cancel each other, resulting in zero net current. The shot noise, however, is not zero; the total shot noise is given by the contributions to the shot noise of the forward and backward currents separately, as

\[ S_I = 2qI_0 \cdot \exp \left( \frac{qV_D}{k_B T} \right) + 2qI_0, \]  

hence, for \( V_D = 0 \)

\[ S_I = 4qI_0. \]  

### 3.3 Generation-Recombination Noise

The generation-recombination (GR) noise in semiconductors originates from random transitions of the charge carriers between electronic states, which causes the number of free carriers available for current transport to fluctuate. These transitions occur not only between the valence and conduction bands, but also between the bands and localized states (bulk traps) within the semiconductor forbidden gap (SAH; NOYCE; SHOCKLEY, 1957).
These energy states exist due to various defects, such as the presence of impurities in the semiconductor and imperfections in its crystalline structure, and may be neutral or charged in its empty state. From the work of Sah, Noyce and Shockley (1957), there are four basic processes involved in the carrier generation and recombination. If an energy state is occupied by a hole, it may either emit the hole to the valence band or capture an electron from the conduction band to recombine with the hole; if it is occupied by an electron, it may either emit the electron to the conduction band or capture a hole from the valence band to recombine with the electron. These processes are illustrated in figure 3.2, where \( E_v \), \( E_c \) and \( E_T \) are the top of the valence band, the bottom of the conduction band and the trap energy level, respectively.

Figure 3.2: Basic processes involved in the generation and recombination of carriers through traps. From left to right: a) electron capture; b) electron emission; c) hole capture; d) hole emission. Processes a) and c) result in recombination of a charge carrier; processes b) and d) result in generation of a charge carrier.

\[
S_N(f) = 4\Delta N^2 \cdot \frac{\tau}{1 + (2\pi f \tau)^2} \tag{3.13}
\]

where \( \tau \) is the time constant of the transitions and \( \Delta N^2 \) is the variance, which can be
expressed as
\[
\frac{1}{\Delta N^2} = \frac{1}{N} + \frac{1}{N_{t,full}} + \frac{1}{N_{t,empty}}
\]  
(3.14)

where \( N_{t,full} \) and \( N_{t,empty} \) are the average number of full and empty traps, respectively (HAARTMAN; ÖSTLING, 2007). The spectrum given by (3.13) is called a Lorentzian due to its similarity with the probability density function of the Lorentz(ian) distribution (also known as Cauchy and Cauchy-Lorentz distribution), with parameters \( x_0 \) and \( \psi \), shown in equation (3.15) for the sake of completion. The Lorentzian has a constant PSD for frequencies below the characteristic time constant, \( 1/\tau \), and decreases with \( 1/f^2 \) for frequencies above \( 1/\tau \).

\[
f(x; x_0, \psi) = \frac{1}{\pi \psi \sqrt{1 + \left(\frac{x - x_0}{\psi}\right)^2}}
\]  
(3.15)

### 3.4 Random Telegraph Signal

The random telegraph signal, also known as random telegraph noise (RTN), is a special case of GR noise (HAARTMAN; ÖSTLING, 2007). In MOS transistors, the telegraphic noise is characterized by discrete fluctuations in the channel conduction, producing a random telegraph signal (RTS). These fluctuations occur due to capture and emission of charge carriers by defects (traps) located in the oxide and at the oxide-semiconductor interface of these devices (KIRTON; UREN, 1989; HUNG et al., 1990a). Figure 3.3 illustrates the effect of RTS in the drain current of a MOS transistor under constant bias.

Figure 3.3: Random telegraph signal observed in the drain current of a transistor due to capture and emission of carriers by a single trap.

According to Yamamoto (2004), it can be demonstrated that the times spent at the high and low current states (\( \tau_1 \) and \( \tau_0 \), respectively) are exponentially distributed,
resulting in a Poisson process. For an ergodic process, the average time the signal spends at the high current state, $\langle \tau_1 \rangle$, is given by

$$\int_0^\infty t p_1(t) dt = \langle \tau_1 \rangle = \overline{\tau}_1,$$

where $\overline{\tau}_1$ is the ensemble average of $\tau_1$ and $p_1(t)$ is the probability that state 1 will not transition between times 0 and $t$; the standard deviation is given by

$$\left[ \int_0^\infty t^2 p_1(t) dt - \overline{\tau}_1^2 \right]^{\frac{1}{2}} = \langle \tau_1 \rangle = \overline{\tau}_1.$$

In essence, despite the inherent similarities between GR noise and RTS — as it will be discussed in the following chapter, RTS can also be modelled using (3.13) —, the first originates in the semiconductor bulk, whereas the latter originates from trapping in the oxide and at the oxide-semiconductor interface.

Random telegraph signals are usually observed in small-area devices for two main reasons: i) the reduced number of traps in these devices makes it easier to identify the two-level signal; and ii) the trap impact on the drain current is increased in small-area devices. Under certain circumstances (such as high drain bias), however, it is also possible to observe RTS in larger devices.

### 3.5 1/f Noise

The 1/f noise, also known as flicker or pink noise, is characterized by fluctuations whose PSD have a $1/f^\lambda$ dependence, being $\lambda \approx 1$. Figure 3.4 shows noise PSD of a large-area NFET under uniformly charged channel conditions, which approximately follows the 1/f behaviour.

Given its importance, several MOSFET 1/f noise models have emerged over the last decades, assuming different physical mechanisms to explain its bias, area and temperature behaviour. These models, as well as the physical mechanisms underlying the 1/f noise, are discussed in detail in chapter 4.

To put the importance of the 1/f noise into perspective, the predicted thermal and shot noises are in the order of $10^{-23} \text{ A}^2/\text{Hz}$ for the device shown in Figure 3.4. Thus, the corner frequency (the frequency at which the contributions of the shot and thermal noises
Figure 3.4: Measured LFN PSD of a large-area NFET (140-nm technology). The frequency exponent ($\lambda$), measured between 10 Hz and 30 Hz, is approximately 1.176. Bias conditions are $V_{DS} = 0.1$ V and $V_{GS} = 1.8$ V.

become more relevant than that of the $1/f$ noise) is expected to be between 10 kHz and 100 kHz. With the technology scaling, the corner frequency is further increased, making the $1/f$ noise the dominant noise mechanism, together with the RTS, in deeply-scaled CMOS technologies.

Throughout most of this work, a normalized noise PSD quantity is used, i.e. $S_{ID}/I_{D}^2$ in Hz$^{-1}$. While this quantity is useful for the analysis proposed, from a designer’s perspective, however, the input referred noise (or gate referred noise) provides better information. The conversion from drain current noise to gate referred noise is straightforward through

$$S_{V_G} = \frac{S_{ID}}{g_m^2},$$

(3.18)

where $g_m$ is the transistor transconductance.
4 LOW-FREQUENCY NOISE MODEL IN MOSFETS

The modelling of LFN in MOSFET requires understanding of fundamental physical mechanisms responsible for the current fluctuations. The first section of this chapter presents the mathematical description of the RTS, critical for the derivation of the RTS-based $1/f$ noise model presented in section 4.2.4, which states that $1/f$ noise and RTS originate from the same physical mechanism: capture and emission of charges by traps (MCWHORTER, 1957; KIRTON; UREN, 1989; BANASZESKI DA SILVA et al., 2014). On the opposite hand, bulk mobility fluctuation models, more prominently the model by Hooge (1969), state that the $1/f$ noise and RTS have different origins.

In this chapter, the conventionally used RTS model as well as popular $1/f$ noise models are presented and discussed.

4.1 Random Telegraph Noise Model

The PSD of a telegraphic signal can be calculated based on its capture and emission time constants, as presented by Machlup (1954). From the Wiener-Khintchine theorem, the PSD of an ergodic, wide-sense stationary, random process, $S(f)$, can be obtained through the Fourier transform of the autocorrelation function of the process, $R(s)$, according to

$$S(f) = 4 \int_0^\infty R(s) \cdot \cos(2\pi fs) ds.$$  \hspace{1cm} (4.1)

From Machlup (1954), the autocorrelation function of a random telegraph process can be described by

$$R(s) = \sum_i \sum_j x_i \cdot x_j \cdot \{ \text{prob. that } x(t) = x_i \} \cdot \{ \text{prob. that } x(t+s) = x_j, \text{ given that } x(t) = x_i \},$$  \hspace{1cm} (4.2)

where $x_i$ and $x_j$ are the possible values for the telegraphic signal; $x(t)$ and $x(t+s)$ are the values of the telegraphic signal at times $t$ and $t+s$, respectively; and $i$ and $j$ are the indexes of the RTS states (0 being the low state and 1 being the high state). The probability that, at any given time, the RTS in state 1 is given by

$$p_{x=1} = \frac{\tau_1}{\tau_1 + \tau_0},$$  \hspace{1cm} (4.3)
whereas the probability of being in state 0 is given by

\[ p_{x=0} = \frac{\tau_0}{\tau_1 + \tau_0}, \quad (4.4) \]

being \( \tau_1 \) and \( \tau_0 \) the average times spent on states 1 and 0, respectively.

By assuming that \( x_0 = 0 \) and \( x_1 = \Delta I \), where \( \Delta I \) is the difference between the current between states 1 (high) and 0 (low), three out of the four summation terms in (4.2) are cancelled, reducing it to

\[ R(s) = \Delta I^2 \cdot \frac{\tau_1}{\tau_1 + \tau_0} \cdot P_{11}(s), \quad (4.5) \]

where \( P_{11}(s) \) is the probability that the signal remains on state 1 after time \( s \), that is, an even number of transitions occur during time \( s \), starting at state 1. Taking \( P_{01}(s) \) as the probability of an odd number of transitions occurring in a given time \( s \), starting at state 1, from the Law of Total Probability

\[ P_{11}(s) + P_{01}(s) = 1. \quad (4.6) \]

The probability of an even number of transitions, \( P_{11} \), at a given time \( s + ds \) — being \( ds \) a sufficiently small time step so that the probability of more than one transition occurring is negligible — is comprised of two mutually exclusive probabilities: \( i \) the probability of an odd number of transitions during time \( t \), \( P_{01}(s) \), followed by one transition during \( ds \); \( ii \) the probability of an even number transitions during time \( t \), \( P_{11}(s) \), and no transitions during \( ds \). This probability can be written in differential form as

\[ P_{11}(s + ds) = P_{10}(s) \cdot \frac{ds}{\tau_0} + P_{11}(s) \cdot \left( 1 - \frac{ds}{\tau_1} \right). \quad (4.7) \]

Substituting (4.6) in (4.7) and applying the limit \( ds \to 0 \), results in

\[ \frac{d}{ds} [P_{11}(s + ds)] + \left( \frac{1}{\tau_0} + \frac{1}{\tau_1} \right) \cdot P_{11}(s) = \frac{1}{\tau_0} \quad (4.8) \]

which can be solved for the initial condition \( P_{11}(0) = 0 \), since there can be no transitions at time 0, yielding

\[ P_{11}(s) = \frac{\tau_1}{\tau_0 + \tau_1} + \frac{\tau_1}{\tau_0 + \tau_1} \cdot \exp \left[ - \left( \frac{1}{\tau_0} + \frac{1}{\tau_1} \right) \cdot s \right]. \quad (4.9) \]
Replacing (4.9) in (4.5) gives the autocorrelation function of the process as

\[ R(s) = \Delta I^2 \cdot \frac{\tau_1}{\tau_0 + \tau_1} \cdot \left\{ \frac{\tau_1}{\tau_0 + \tau_1} + \frac{\tau_0}{\tau_0 + \tau_1} \cdot \exp \left[ -\left( \frac{1}{\tau_0} + \frac{1}{\tau_1} \right) \cdot s \right] \right\}. \quad (4.10) \]

Then solving (4.1) using (4.10) gives

\[ S(f) = 4 \int_0^{\infty} R(s) \cdot \cos(2\pi fs) ds = \frac{4\Delta I^2}{(\frac{1}{\tau_0} + \frac{1}{\tau_1}) \cdot \left[ \left( \frac{1}{\tau_0} + \frac{1}{\tau_1} \right)^2 + (2\pi f)^2 \right]} \]

which gives the PSD of a telegraphic signal as a function of its average times spent on high and low states. In the context of trapping and de-trapping of charge carriers by defects, the average times spent on low and high states can be understood as the average time required to emit (\( \tau_e \)) and capture (\( \tau_c \)) a charge carrier, respectively.

Equation (4.11) can be rewritten in a more useful form by substituting

\[ \beta = \frac{\tau_e}{\tau_c} \]

and

\[ \frac{1}{\tau} = \frac{1}{\tau_c} + \frac{1}{\tau_e}, \]

resulting in

\[ S(f) = 4\Delta I^2 \cdot \frac{\beta}{(1 + \beta)^2} \cdot \frac{\tau}{1 + (2\pi f \tau)^2}, \]

where \( \beta \) and \( \tau \) are the ratio between emission and capture, and the characteristic time constant, respectively. From detailed balance (KIRTON; UREN, 1989),

\[ \beta = \frac{\tau_e}{\tau_c} = g \cdot \exp \left( \frac{E_F - E_T}{kT} \right), \]

where \( g, k, E_F \) and \( E_T \) are the degeneracy (approximately equal to one), Boltzmann constant in eV·K\(^{-1}\), Fermi energy and trap energy, respectively.

For the particular case in which \( \tau_0 = \tau_1 = \tau_c \), (4.14) can be simplified into

\[ S(f) = 2\Delta I^2 \cdot \frac{\tau_c}{4 + (2\pi f \tau_c)^2}, \]

resulting in maximum total noise power. For this reason, experimentally identifiable telegraphic signals usually display \( \tau_c \approx \tau_e \), which may be wrongly interpreted as a correlation between the two parameters that are, essentially, uncorrelated (GRASSER et al., 2009).
4.2 1/f Noise Models

Historically, two main schools of thought have emerged to explain and model the 1/f behaviour of the power-spectral density. The first assumes the origin is related to the capture and emission of charge carriers by surface defects (traps) (MCWHORTER, 1957), and is commonly referred to as number fluctuation theory. The second attributes the 1/f to bulk mobility fluctuations (HOOG, 1969; HOOG; VANDAMME, 1978), and is commonly referred to as mobility fluctuation theory. A unified theory has been proposed (HUNG et al., 1990b; HUNG et al., 1990c; GHIBAUDO et al., 1991), incorporating mobility fluctuations due to Coulomb scattering to the number fluctuation model. This unified model is not, however, a combination of both models (SCHOLTEN et al., 2003), as the mobility fluctuations share the same origin as number fluctuations and are, thus, correlated.

While bulk mobility fluctuation models have been used to describe the 1/f noise bias dependence of PFETs (CHANG; ABIDI; VISWANATHAN, 1994; VANDAMME; HOOG, 2008; HAARTMAN; ÖSTLING, 2007), number fluctuation models have been recently shown to explain the 1/f bias dependence of both P- and NFETs if a non-uniform distribution of traps in energy (VAN DER WEL et al., 2005; WIRTH; SILVA; KACZER, 2011) is assumed for p-type devices (BANASZESKI DA SILVA et al., 2014; BANASZESKI DA SILVA et al., 2016). Even in the context of trapping of charges, the trapping mechanism itself is not thoroughly understood. While studies on random telegraph signals and 1/f noise (DUTTA; HORN, 1981; SURYA; HSIANG, 1988; KIRTON; UREN, 1989) suggest a thermally activated process, compact 1/f noise models models, such as BSIM and PSP (HUNG et al., 1990b; HUNG et al., 1990c), still employ direct tunnelling (MCWHORTER, 1957), which is temperature independent.

In this section, number and mobility fluctuation models are presented. The correlated mobility model is also presented, as well as its compact version.

4.2.1 Number Fluctuation Model - McWhorter Model

Number fluctuation models assume that the primary mechanism responsible for the 1/f noise are the capture and emission of carriers by traps, which cause fluctuations to the surface potential of the device. As discussed in section 4.1, the activity of a single trap produces a Lorentzian power-spectrum. The work of Surdin (1939) proved that a 1/f
A physical model was proposed by McWhorter (1957) for a semiconductor with an oxide on the surface. This model, widely regarded as the McWhorter model, assumes that the uniformly distributed $\tau$ arise from direct tunnelling of charge carriers to traps located within the oxide, causing the current to fluctuate. Essentially, variations in trap occupancy result in carrier number fluctuations in the channel. The normalized drain current fluctuation can be written as

$$\frac{\partial I_D}{I_D} = \left[ \frac{1}{\Delta N} \frac{\partial \Delta N}{\partial \Delta N_t} \right] \partial \Delta N_t,$$

where $\Delta N = N W \Delta x$ and $\Delta N_t = N_i W \Delta x$, being $N$ and $N_i$ the number of channel carriers and and occupied traps per unit area, respectively; $W$ the transistor width and $\Delta x$ a distance along the transistor length. The ratio $\partial \Delta N / \partial \Delta N_t = R$ is the coupling coefficient between the carrier number fluctuation in the channel and the trap occupancy. This ratio, $R$, is approximately 1 for strong inversion, but may assume smaller values for other regions of operation, being expressed by

$$R = \frac{C_i}{C_{ox} + C_i + C_d + C_{it}},$$

where $C_{ox}$, $C_i$, $C_d$ and $C_{it}$ are the oxide, inversion layer, depletion, and interface trap capacitances per unit area. Substituting

$$C_i \approx \frac{q^2 N}{k_BT},$$

spectrum could be obtained by the superposition of Lorentzian spectra with $\tau$ homogeneously distributed in log-scale, as shown in Figure 4.1.

Figure 4.1: Sum of uniformly-spaced (in log-scale) Lorentzians, resulting in an approximately $1/f$ spectrum.
equation (4.18) can be rewritten as

$$ R = - \frac{N}{N + N^*}, $$

(4.20)

where

$$ N^* = \frac{k_B T}{q^2} \cdot (C_{ox} + C_d + C_{it}), $$

(4.21)

with typical $N^*$ values being of the order of $10^{-10}$ cm$^{-2}$ (HUNG et al., 1990b).

Consequently, the power spectral density of the local drain current fluctuation is given by

$$ S_{\Delta I_D}(x, f) = \left( \frac{I_D}{W \Delta x} \right)^2 \cdot \left( \frac{R}{N} \right)^2 \cdot S_{\Delta N_t}(x, f), $$

(4.22)

where $S_{\Delta N_t}(x, f)$ is the PSD of the mean-square fluctuations in the number of occupied traps in the area $W \Delta x$, and it is given by

$$ S_{\Delta N_t}(x, f) = \int_{E_c}^{E_v} \int_0^W \int_0^{T_{tx}} 4N_{tr}(E, x, y, z) \Delta x \cdot f_t(1 - f_t) \cdot \frac{\tau(E, x, y, z)}{1 + \omega^2 \tau(E, x, y, z)^2} dz dy dE, $$

(4.23)

where $N_{tr}(E, x, y, z)$ is the distribution of the traps in the oxide and over the energy, length, width, and depth; $\tau(E, x, y, z)$ is the trapping time constants; $\omega = 2\pi f$ is the angular frequency; $T_{tx}$ is the oxide thickness; $E_c - E_v$ is the silicon band gap; and $f_t$ is the trap occupancy function. This function, $f_t$, is given by

$$ f_t = \left[ 1 + \exp \left( \frac{E - E_{Fn}}{kT} \right) \right]^{-1}, $$

(4.24)

where $E_{Fn}$ is the electron quasi Fermi level.

In order to evaluate the integral in (4.23), the following suppositions are made: a) the distribution of the traps is uniform close to the interface, thus $N_{tr}(E, x, y, z) = N_{tr}(E)$; b) the probability of an electron penetrating into the oxide decreases exponentially with the distance from the interface; as a result, the trapping time constant is given by

$$ \tau = \tau_0(E) \cdot \exp (\gamma z), $$

(4.25)

where $\tau_0(E)$ is the time constant at the interface and $\gamma$ is the wave function attenuation factor given by the WKB theory for carrier tunnelling in the Si-SiO$_2$ system (HUNG et al., 1990b). Notice that, while the McWhorter model traditionally employs direct tunnelling,
thermal activation, given by

$$\tau = \tau_0 \cdot \exp \left( \frac{E_B}{kT} \right), \quad (4.26)$$

where $E_B$ is an activation energy, has been suggested by Surya and Hsiang (1988). The temperature dependence of the characteristic time constant is further explored in chapter 5 of this work.

Figure 4.2 illustrates the $f_t(1 - f_t)$ term of equation (4.23), which behaves like a delta function around the quasi Fermi level. Thus, the main contribution for the integral is for traps close to the quasi Fermi level, hence one can approximate $N_{tr}(E) = N_{tr}(E_{Fn})$, removing $N_{tr}(E_{Fn})$ from the integral. The spread of the curve around the quasi Fermi level is related to the temperature of the system, as shown in figure 4.3. The area below the curve can be shown to increase with $kT$, meaning a larger number of traps are active at higher temperatures.

Figure 4.2: Behaviour of the $f_t(1 - f_t)$ term for different electron quasi Fermi levels, $E_{Fn}$.

Figure 4.3: Behaviour of the $f_t(1 - f_t)$ term as a function of trap energy for different temperatures and $E_F = 0.6$ eV.

Thus, replacing $f_t(1 - f_t) = -kT \frac{df_t}{dE}$ in equation (4.23) and solving the integration, results in

$$S_{\Delta N_t}(x, f) = N_{tr}(E_{Fn}) \cdot \frac{kTW \Delta x}{\gamma f}. \quad (4.27)$$
Hence, the total drain current noise power spectral density can be calculated using

\[ S_{I_D} = \frac{1}{L^2} \int_0^L S_{\Delta I_D}(x, f) \cdot \Delta x \, dx, \quad (4.28) \]

where \( L \) is the channel length.

Finally, replacing (4.27) in (4.22) and then (4.22) in (4.28), results in

\[ S_{I_D} = \frac{kT I_D^2}{\gamma fWL^2} \int_0^L N_{tr}(E_{Fn}) \left( \frac{R}{N(x)} \right)^2 \, dx, \quad (4.29) \]

which provides a description of the \( 1/f \) noise in terms of bias, temperature and geometry.

### 4.2.2 Mobility Fluctuation Model - Hooge Model

The work of Hooge (1969) demonstrated, based on \( 1/f \) noise measurements of homogeneous samples, that surfaces effects were not primarily responsible for the \( 1/f \) noise, indicating that the noise was essentially a bulk effect. This observation naturally contradicts the explanation given by McWhorter (1957), that attributes the \( 1/f \) to surface effects (traps).

The bulk mobility fluctuation model proposed by Hooge (1969), and Hooge and Vandamme (1978) is given by the empirical formula

\[ S_{I_D}(f) = \frac{q\alpha_H I_D^2}{fWL^2} \int_0^L \frac{dx}{Q_{inv}(x)} \]

where \( \alpha_H \) and \( Q_{inv}(x) \) are the Hooge parameter and inversion layer charge, respectively. These bulk mobility fluctuations were attributed to phonon scattering (HOOG; VANDAMME, 1978). Since mobility in the channel depends not only on technology but also on the electric field, the Hooge parameter depends both on technology and bias conditions.

The works of Handel (1980) and Ngai (1980) attempted to explain the bulk mobility fluctuations through quantum beats between elastically scattered and weakly inelastically scattered carriers, respectively, being the distinct inelastic scattering mechanism the difference between both theories. These hypothesis, however, require unrealistic assumptions, as stated by Weissman (1988). The work of Jindal and van der Ziel (1981) proposed an explanation based on the fluctuation in the mean free path of the carriers; this
explanation is also insufficient according to Weissman (1988), as the characteristic times would be too short to produce low-frequency fluctuations.

For deeply-scaled devices (sub-1 \( \mu \text{m}^2 \) area), the Lorentzian-like profiles of single traps become more prominent than the typical \( 1/f \) spectra, which indicates that trapping is, at least for small-area devices, the dominant noise mechanism. According to Vandamme and Hooge (2008), this is not evidence that the \( 1/f \) spectra of large-area devices is caused by the superposition of these Lorentzians, reinforcing the importance of bulk mobility fluctuations.

4.2.3 Correlated Fluctuations Model - Hung Model

The \( 1/f \) noise model implemented in the BSIM3, BSIM4 and PSP models was developed by Hung et al. (1990b) and Hung et al. (1990c). This model is considered an 'unified' model due to the fact that it accounts for both carrier number fluctuations and mobility fluctuations in the channel of the device. Notice that this unified model is not a combination of both models, as pointed out by Scholten et al. (2003). The mobility fluctuations are a consequence of the Coulombic scattering due to trap occupation. Therefore, these mobilities fluctuations are not a bulk effect, rather sharing the same surface origin as number fluctuations. Hence, these fluctuations are correlated.

The derivation of the model, is similar to that of the McWhorter model. Variations in trap occupancy, however, result not only in fluctuations in carrier number, but also in correlated fluctuations in carrier mobility in the channel (HUNG et al., 1990b). The normalized drain current fluctuation is similar to that of (4.17), but has an added mobility factor, and is given by

\[
\frac{\partial I_D}{I_D} = - \left[ \frac{1}{\Delta N} \frac{\partial \Delta N}{\partial \Delta N_t} \pm \frac{1}{\mu} \frac{\partial \mu}{\partial \Delta N_t} \right] \partial \Delta N_t. \tag{4.31}
\]

The signal ahead of the mobility term in equation (4.31) is chosen according to the trap charge (if the trap is neutral or charged when occupied). The ratio \( \partial \mu / \partial \Delta N_t \) is the coupling coefficient between the carrier mobility fluctuation and the trap occupancy. According to Hung et al. (1990b), it can be demonstrated, through Mathiessen rule, that

\[
\frac{\partial \mu}{\partial \Delta N_t} = \frac{\alpha \mu^2}{W \Delta x}. \tag{4.32}
\]
where $\alpha$ is the scattering coefficient.

Given that the relation $\partial \Delta N / \partial \Delta N_t = R$ from the McWhorter model is still valid, then (4.31) can be rewritten as

$$\frac{\partial I_D}{I_D} = - \left( \frac{R}{N} \pm \alpha \mu \right) \cdot \frac{\partial \Delta N_t}{W \Delta x}. \quad (4.33)$$

Consequently, the power spectral density of the local drain current fluctuation is given by

$$S_{\Delta I_D}(x,f) = \left( \frac{I_D}{W \Delta x} \right)^2 \cdot \left( \frac{R}{N} \pm \alpha \mu \right)^2 \cdot S_{\Delta N_t}(x,f) \quad (4.34)$$

where $S_{\Delta N_t}(x,f)$ is the PSD of the mean-square fluctuations in the number of occupied traps in the area $W \Delta x$, given by (4.23).

Following the same steps of the McWhorter model derivation — including the assumption that direct tunnelling is responsible for the distribution of characteristic time constants —, the $1/f$ noise PSD is found to be

$$S_{I_D} = \frac{kTI_D^2}{fWL^2} \int_0^L N_{tr}(E_{Fn}) \cdot \left[ \frac{R}{N(x)} \pm \alpha \mu \right]^2 \cdot R \cdot dV. \quad (4.35)$$

From (4.35), the works of Hung et al. (1990b), Hung et al. (1990c) derived a compact model that has been widely employed in MOSFET models, such as BSIM and PSP. Substituting $dx = -dV/E$ and obtaining $E$ from the drain current equation

$$I_D = W \mu q N E_x \quad (4.36)$$

results in

$$S_{I_D} = \frac{kTqI_D \mu}{fWL^2} \int_0^{V_D} N_{tr}(E_{Fn}) \cdot \left( 1 \pm \alpha \mu NR^{-1} \right)^2 \cdot R^2 \cdot dV. \quad (4.37)$$

In order to solve (4.37) the bias dependence of parameters $\alpha$, $\mu$ and $N_t(E_{Fn})$ is required. Since $\alpha$ and $\mu$ are a function of the carrier concentration, $N$, which, in turn, is a function of the carrier quasi Fermi level, (4.37) can be rewritten as

$$S_{I_D} = \frac{kTqI_D \mu}{\lambda fL^2} \int_0^{V_D} N_{tr}^*(E_{Fn}) \cdot \frac{R^2}{N} \cdot dV. \quad (4.38)$$

where

$$N_{tr}^* = N_t(E_{Fn}) \cdot \left( 1 \pm \alpha \mu NR^{-1} \right)^2. \quad (4.39)$$

In this form, $N_{tr}^*(E_{Fn})$ is the apparent oxide trap density that produces the same
noise power as if there were no contribution from the mobility fluctuation (HUNG et al., 1990b). The model proposed by Hung et al. (1990b) provides an empirical parametric function which is sufficiently simple to provide an analytical expression for $S_{ID}$, but robust enough to be applicable for MOSFETs fabricated by different technological processes. The proposed model, equation (4.40), has three parameters — NOIA, NOIB and NOIC — and is a function of the carrier concentration, $N$.

$$N_{rr}^* = NOIA + NOIB \cdot N + NOIC \cdot N^2$$  \hspace{1cm} (4.40)

Parameters NOIA, NOIB and NOIC are technology-dependent parameters. Assuming that the trap density, $N_t$, is independent of energy, then

$$NOIA = N_t$$  \hspace{1cm} (4.41)

$$NOIB = \pm 2 \alpha \mu_{eff} R^{-1} N_t$$  \hspace{1cm} (4.42)

$$NOIC = (\alpha \mu_{eff} R^{-1})^2 N_t.$$  \hspace{1cm} (4.43)

Substituting (4.40) in (4.38) an analytical expression can be derived for the $1/f$ noise PSD for different device operating regions, as presented by Hung et al. (1990b). In strong inversion, the $1/f$ noise PSD is given by

$$S_{ID,inv}(f) = \frac{kT q^2 \mu_D}{a \lambda f L^2 C_{ox}} [NOIA \cdot \log \left( \frac{N_0 + N^*}{N_l + N^*} \right) + NOIB \cdot (N_0 - N_l) + \frac{NOIC}{2} (N_0^2 - N_l^2)]$$

$$+ \Delta L \frac{kT_D^2}{\lambda f W L^2} \left[ \frac{NOIA + NOIB \cdot N_l + NOIC \cdot N_l^2}{(N_l + N^*)^2} \right],$$

(4.44)

where

$$N_0 = N(0) = \frac{C_{ox}}{q} \cdot (V_{GS} - V_{th0}) ,$$

(4.45)

$$N_l = N(L) = \frac{C_{ox}}{q} \cdot (V_{GS} - V_{th0} - aV_{DS}),$$

(4.46)

being $a$, $V_{th0}$, $V_{DS}$ and $V_{GS}$ a bulk charge effect modelling parameter, threshold voltage of the long channel device at zero substrate bias, gate-to-source voltage, and drain-to-source voltage (in the case of saturation, the saturation voltage, $V_{DS_{sat}}$), respectively (DUNGA et
al., 2006). For subthreshold region, $S_{ID}(f)$ is given by

$$S_{ID,subVt}(f) = \frac{NOIA \cdot kTI_D^2}{WL\lambda fN^*}.$$ (4.47)

The parameters of equations (4.44) and (4.47) are presented in Table 4.1.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>Boltzmann Constant</td>
</tr>
<tr>
<td>$T$</td>
<td>Temperature</td>
</tr>
<tr>
<td>$q$</td>
<td>Elementary Charge</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Effective Carrier Mobility</td>
</tr>
<tr>
<td>$I_D$</td>
<td>Drain Current</td>
</tr>
<tr>
<td>$C_{ox}$</td>
<td>Oxide Capacitance</td>
</tr>
<tr>
<td>$L$</td>
<td>Effective Channel Length</td>
</tr>
<tr>
<td>$W$</td>
<td>Effective Channel Width</td>
</tr>
<tr>
<td>$a$</td>
<td>Bulk charge effect modelling parameter</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Attenuation coefficient given by WKB theory</td>
</tr>
<tr>
<td>NOIA</td>
<td>Noise modelling parameter A</td>
</tr>
<tr>
<td>NOIB</td>
<td>Noise modelling parameter B</td>
</tr>
<tr>
<td>NOIC</td>
<td>Noise modelling parameter C</td>
</tr>
<tr>
<td>$N_0$</td>
<td>Carrier density at source terminal</td>
</tr>
<tr>
<td>$N_I$</td>
<td>Carrier density at drain terminal</td>
</tr>
<tr>
<td>$\Delta L$</td>
<td>Channel length modulation parameter</td>
</tr>
</tbody>
</table>

The 1/f noise PSD at any operating point is then given by

$$S_{ID}(f) = \frac{S_{ID,inv}(f) \cdot S_{ID,subVt}(f)}{S_{ID,inv}(f) + S_{ID,subVt}(f)}.$$ (4.48)

Despite its popularity, the correlated fluctuations model has been criticized by Vandamme and Hooge (2008), arguing that the correlated mobility parameters is but a mere curve fitting parameter, with unrealistic values.

### 4.2.4 RTS-Based Statistical Model

RTS-based models are fundamentally similar to number fluctuation models, as capture and emission of carriers by traps is also assumed, and the LFN PSD is a result
of the summation of Lorentzians. In opposition to the McWhorter model, however, these models attribute statistical distribution to specific noise parameters — such as number of traps in a device and individual trap impact —, which allow the estimation of the noise variability, for instance. Among the models that propose a RTS-based LFN are the works of Wirth et al. (2005); Wirth, da Silva and Brederlow (2007); da Silva, Wirth and Brusamarello (2008); Banaszeski da Silva et al. (2014); Banaszeski da Silva et al. (2016).

The model proposed by Banaszeski da Silva et al. (2014) provides an analytical expression for the noise variance. The following assumptions are taken for the derivation of the model:

(i) the total number of traps in a device (N) is a Poisson distributed random variable;
(ii) \( \tau \) is a log-uniformly distributed random variable;
(iii) \( \beta/(1+\beta)^2 \) and \( \beta^2/(1+\beta)^4 \) are approximated by delta functions multiplied by \( kT \) and \( kT/6 \), respectively;
(iv) \( \Delta I_D, N \) and \( \tau \) are independent.

### 4.2.4.1 Expected Value

The work of Banaszeski da Silva et al. (2014) demonstrated that, from the RTS framework, the expected value of the LFN in MOS devices is given by

\[
E[S_{I_D}(f)] = \frac{kT I_D^2}{2WL^2 f} \int_0^L E\left[ \Delta \tilde{I}_D^2 | X_{tr} = x \right] \cdot \frac{N_{tr}(E_{Fn})}{\gamma'} dx, \tag{4.49}
\]

where the trap position along the channel, \( X_{tr} \), is assumed to be a uniformly distributed random variable; \( \Delta \tilde{I}_D \) is the normalized current deviation, given by

\[
\Delta \tilde{I}_D = \frac{WL\Delta I_D}{I_D}; \tag{4.50}
\]

and \( N_{tr}(E_{Fn})/\gamma \) gives the trap density in the area, energy and in the log-domain of time constants \((\text{cm}^{-2}\text{eV}^{-1} \text{per neper})\). It is important to notice that \( \gamma' \) in the model is the parameter of the log-uniform distribution of time constants.

For the particular case in which the channel is uniformly charged, if the drain current fluctuations are assumed to be exponentially distributed, (4.49) can be simplified to

\[
E[S_{I_D}(f)] = 2\frac{kT I_D^2}{WL f} \cdot \frac{N_{tr}(E_{Fn})}{\gamma'} \cdot E\left[ \Delta \tilde{I}_D \right]^2. \tag{4.51}
\]
Incorporating number fluctuation and correlated mobility fluctuations, the average drain current fluctuation, \( E[\Delta \tilde{I}_D] \), is given by

\[
E[\Delta \tilde{I}_D] = \left( \frac{R}{N} + \alpha \mu \right),
\]

and \( R \) is given by (4.20).

**4.2.4.2 Variance**

Similarly to the expected value, the variance of the LFN PSD can be derived from the RTS-based framework, as shown by Banaszeski da Silva et al. (2014). The variance is given by

\[
\text{Var}[S_{I_D}(f)] = \frac{kT I_D^4}{3 \pi^2 f^2 W^3 L^4} \int_0^L E[\Delta \tilde{I}_D^4 | X_T = x] \cdot \frac{N_{tr}(E_{Fn})}{\gamma'} dx.
\]

For the particular case in which the channel is uniformly charged, if the drain current fluctuations are assumed to be exponentially distributed, and carrier number fluctuations and correlated mobility fluctuations are considered, (4.53) can be simplified to

\[
\text{Var}[S_{I_D}(f)] = 24 \frac{kT I_D^4}{3 \pi^2 f^2 W^3 L^3} \cdot \frac{N_{tr}(E_{Fn})}{\gamma'} \left( \frac{R}{N} + \alpha \mu \right)^4.
\]

The aforementioned work has also shown that it is possible to relate the expected value and variance of the LFN PSD to the standard deviation of \( \ln(S_{I_D}) \) using

\[
\sigma[\ln(S_{I_D})] = \sqrt{\ln \left(1 + \frac{\text{Var}[S_{I_D}]}{E[S_{I_D}^2]}\right)},
\]

which can be rewritten as

\[
\sigma[\ln(S_{I_D})] = \sqrt{\ln \left(1 + \frac{K}{WL}\right)},
\]

for \( K = WL \cdot \text{Var}[S_{I_D}]/E[S_{I_D}^2] \). For the uniformly inverted channel condition, the \( K \) parameter is easily obtained analytically through

\[
K = \frac{2}{\pi^2 kT} \cdot \frac{\gamma'}{N_{tr}(E_{Fn})}.
\]
Notice that $K$ is geometry independent and, thus, is a technology parameter, suitable for comparing noise variability in different technology nodes.

### 4.2.4.3 Correlation Coefficient

For upcoming analyses presented in this work, the correlation coefficient between two LFN PSD populations is pivotal. Thus, it is necessary estimate the correlation coefficient using the same RTS-based framework employed for the expected value and variance expressions. The Pearson product-moment correlation coefficient, $R$, of two random variables, $X$ and $Y$, is given by the covariance of the two variables divided by the square root of the product of their variance, so that

$$R[X, Y] = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X] \cdot \text{Var}[Y]}},$$

where $\text{Cov}$ is the covariance operator. In the context of this work, $X$ and $Y$ are the PSDs of the LFN and are treated as stochastic processes in respect to a given system variable – such as frequency, temperature and biasing. Since the same stochastic process is been compared to itself under different conditions, these analyses are called autocorrelation analyses.

From the RTS-based framework presented in section 4.1, the LFN PSD of a MOS-FET is given by the sum of the contribution of individual traps, thus

$$E[S_{d}] = E \left[ \sum_{i=1}^{N} L_{x, i} \right],$$

in which $N$ and $L_{x,i}(f)$ are the number of traps in the device (which, itself is a random variable) and the Lorentzian PSD of the individual trap $i$ at a given condition $x$, respectively.

Using the Law of Total Variance gives

$$\text{Var} \left[ \sum_{i=1}^{N} L_{x, i} \right] = E \left[ \text{Var} \left[ \sum_{i=1}^{N} L_{x, i} \mid N \right] \right] + \text{Var} \left[ E \left[ \sum_{i=1}^{N} L_{x, i} \mid N \right] \right].$$

For $N$ and $L_{x,i}$ independent yields

$$E \left[ \text{Var} \left[ \sum_{i=1}^{N} L_{x, i} \mid N \right] \right] = E[N] \cdot \text{Var}[L_{x}],$$
and

$$\text{Var} \left[ \mathbb{E} \left[ \sum_{i=1}^{N} L_{xi} \mid N \right] \right] = \text{Var}[N] \cdot \mathbb{E}[L_x]^2 . \quad (4.62)$$

Assuming that $N$ is Poisson distributed and, thus

$$\mathbb{E}[N] = \text{Var}[N] , \quad (4.63)$$

and writing the variance as

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 , \quad (4.64)$$

it is possible to simplify (4.60) to

$$\text{Var} \left[ \sum_{i=1}^{N} L_{xi} \right] = \mathbb{E}[N] \cdot \mathbb{E}[L_x^2] . \quad (4.65)$$

Similarly, using the Law of Total Covariance

$$\text{Cov} \left[ \sum_{i=1}^{N} L_{xi}, \sum_{i=1}^{N} L_{yi} \right]$$

$$= \mathbb{E} \left[ \text{Cov} \left[ \sum_{i=1}^{N} L_{xi}, \sum_{i=1}^{N} L_{yi} \mid N \right] \right] + \text{Cov} \left[ \mathbb{E} \left[ \sum_{i=1}^{N} L_{xi} \mid N \right], \mathbb{E} \left[ \sum_{i=1}^{N} L_{yi} \mid N \right] \right] . \quad (4.66)$$

For $N$ independent of $L_{i,x}$ and $L_{i,y}$, (4.66) reduces to

$$\text{Cov} \left[ \sum_{i=1}^{N} L_{xi}, \sum_{i=1}^{N} L_{yi} \right] = \mathbb{E}[N] \cdot \text{Cov}[L_x, L_y] + \mathbb{E}[N, N] \cdot \mathbb{E}[L_x] \cdot \mathbb{E}[L_y] . \quad (4.67)$$

Naturally, Cov$[N, N] = \text{Var}[N]$. If $N$ is again assumed to be Poisson distributed, then (4.67) can be further simplified to

$$\text{Cov} \left[ \sum_{i=1}^{N} L_{xi}, \sum_{i=1}^{N} L_{yi} \right] = \mathbb{E}[N] \cdot (\text{Cov}[L_x, L_y] + \mathbb{E}[L_x] \cdot \mathbb{E}[L_y]) . \quad (4.68)$$

Finally, using the definition of covariance

$$\text{Cov}[X, Y] = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y] \quad (4.69)$$
simplifies (4.68) to

\[
\text{Cov} \left[ \sum_{i=1}^{N} L_{x,i}, \sum_{i=1}^{N} L_{y,i} \right] = \mathbf{E}[N] \cdot \mathbf{E}[L_x \cdot L_y]. \tag{4.70}
\]

Substituting (4.65) and (4.70) into (4.58) results in

\[
R[S_{D,x}, S_{D,y}] = \frac{\mathbf{E}[L_x \cdot L_y]}{\sqrt{\mathbf{E}[L_x^2] \cdot \mathbf{E}[L_y^2]}}. \tag{4.71}
\]

Equation (4.71) will be used in Chapter 7 of this work to analytically and numerically calculate the correlation coefficient between two LFN PSD populations under different conditions using the RTS framework. These predictions can be compared to measured data in order to verify whether the models properly describe the behaviour of the fundamental noise sources.

4.3 Summary and Discussion

Random telegraph noise and 1/f noise are the primary contributors to CMOS LFN power. While bulk-mobility fluctuation models (more prominently the Hooge model) support that RTS and 1/f noise have distinguished origins, number fluctuation models (including the correlated mobility fluctuations) not only attribute both RTS and 1/f noise to the trapping and de-trapping of charge carrier, but also support the view that they are, essentially, the same process.

Vandamme and Hooge (2008) questioned this unified view, stating that “the typical bumpy spectra in small MOSTs do not represent a conclusive proof that the 1/f noise in large devices result from the superposition of many elementary RTS fluctuations.” Using the methodology proposed in this work, it is demonstrated in section 7.2 that there is strong evidence to support the claim that 1/f noise in large-area MOS transistors is indeed a product of RTS fluctuations, even if no bumps are visible in the spectrum.

In the context of charge trapping, the assumption that the characteristic time constant of traps is governed by direct tunnelling has been employed not only in noise models (MCWHORTER, 1957; HUNG et al., 1990b; HUNG et al., 1990c), but also in works where the position of the trap within the oxide is extracted, such as Lukyanchikova et al. (2009) and Ding, Misra and Srinivasan (2016), despite observations of a thermally activated process being reported in several works, such as Surya and Hsiang (1988), Kirton
and Uren (1989), Xiong et al. (2002) and Wei, Xiong and Zhou (2009).

Notice that in section 7.3, the methodology proposed in this work is applied to experimental data and the results support the view direct tunnelling is not sufficient to describe the behaviour of the LFN with the temperature; a thermally activated model, in the likes of that proposed by Dutta and Horn (1981); Surya and Hsiang (1988); and Kirton and Uren (1989), is required to fully explain the observed LFN behaviour. Consequently, it should be accounted for in the aforementioned LFN models.

The following chapter addresses this important discussion on the temperature dependence of MOSFET LFN.
5 LOW-FREQUENCY NOISE TEMPERATURE DEPENDENCE

The temperature dependence of the $1/f$ noise has been an important aspect for the modelling of $1/f$ noise and the understanding of its dynamics. Among the first studies of the $1/f$ noise in semiconductor devices is the work by McWhorter (1957), which observed a weak dependence of the $1/f$ noise magnitude on temperature. This behaviour led to the proposition of a noise model based on electron tunnelling, which is inherently a temperature independent process, stating that the time constant distribution is related to the spatial distribution of traps depth-wise. Recent works, however, have shown modern oxides are simply too thin to support the elastic tunnelling model (CAMPBELL et al., 2009) and capture and emission times are uncorrelated to the trap distance from the interface (NAGUMO et al., 2010).

In contradiction to the model proposed by McWhorter (1957), the work of Surya and Hsiang (1988) observed that the $1/f$ noise of commercial p-channel MOSFETs varied significantly, not only in magnitude, but also in form, when operated from 60 K up to 290 K. Based on these experimental results, Surya and Hsiang (1988) proposed a different approach to the calculation of the times constants, based on the thermal activated model for $1/f$ noise in semiconductors by Dutta and Horn (1981)\footnote{If the reader is not familiarised with the Dutta and Horn identity, please refer to section 5.1 of this work.}.

Figure 5.1a displays experimental data for the noise PSD at 200 Hz as a function of the absolute temperature. The result indicates that the magnitude of the $1/f$ noise PSD is, in fact, temperature dependent. Additionally, Figure 5.1b displays extracted frequency exponent, $\gamma$, of the $1/f$ noise spectrum at 200 Hz as a function of temperature, indicating that the shape of noise spectrum is strongly dependent on the temperature and that this relation is non-linear.

Figure 5.1: a) Noise PSD and b) measured frequency exponent as a function of temperature for a p-channel device in linear region.

Source: Adapted from Surya and Hsiang (1988)
According to Surya and Hsiang (1988), the simplistic spatial distribution of traps is insufficient to fully account for the temperature dependence of the noise magnitude, thus questioning the validity of the elastic tunnelling model. Conversely, the work claims that using the thermal activation model it is possible to account for both the magnitude and the functional form of the $1/f$ noise PSD.

The work of Wong and Cheng (1990) delves further into the temperature dependence of the $1/f$ noise by analysing n-channel MOS devices with different gate insulators. Figure 5.2 displays the frequency exponent as a function of the temperature for different devices. According to their work, there is a correlation of the peak locations in the plots and the major trapping centres in the Si-SiO$_2$ interface.

Figure 5.2: Frequency exponent as a function of absolute temperature for n-channel devices with different insulators.

A comparison between commercially available CMOS transistors from subthreshold to strong inversion was performed by Chang, Abidi and Viswanathan (1994). For NFETs in strong inversion, no significant magnitude variation was observed in the input-referred noise spectrum at different temperatures, and only a slope change was verified, as shown in Figure 5.3a. These results, associated to the weak bias dependence observed, gave credence to the tunnelling model proposed by McWhorter (1957).
The PFETs, on the other hand, showed strong bias dependence on the input-referred noise spectra at all temperatures. Figure 5.3b displays the input referred noise for the p-channel device in strong inversion for different temperatures; a significant $1/f$ noise spectra variation with temperature can be observed, in contrast to that of the NFET. According to Chang, Abidi and Viswanathan (1994), the behaviour of the p-channel transistor can be modelled using mobility fluctuations. Therefore, it concludes that noise in the n-channel device is more conveniently modelled using a number fluctuation model, whereas the p-channel device is more conveniently modelled using a bulk mobility fluctuation model.

Figure 5.3: Noise PSD for a) p-channel transistor and b) n-channel transistor at different temperatures in the linear region of operation.

A study regarding the impact of total ionizing dose (TID) on the $1/f$ noise spectra of NFETs at different temperatures was performed by Xiong et al. (2002). The noise magnitude variation with temperature of an NFET device before and after irradiation, and after annealing, are shown in figure 5.4a. Their findings reinforce the idea that the $1/f$ noise in n-channel devices is a thermally activated process, in accordance to Surya and Hsiang (1988), agreeing that the relation presented by Dutta and Horn (1981) provides a very reasonable quantitative description of the noise, as shown in figure 5.4b. The success of the Dutta and Horn relation in describing the correlated temperatures and frequency dependence of the noise allows the energy distribution of the defect responsible for the noise to be estimated from noise measurements as a function of temperature (XIONG et al., 2002).
Figure 5.4: a) Noise PSD of an NFET at 1 Hz as a function of temperature prior and after irradiation. b) Measured and calculated frequency exponent for the pre-irradiated condition.

Source: Adapted from Xiong et al. (2002)

Lee et al. (2006) studied the $1/f$ noise in sub-100 nm MOSFETs. Figure 5.5 shows the $1/f$ noise PSD for NFETs and PFETs at different temperatures. For the p-channel device, a small noise magnitude reduction from 27 °C to 70 °C is observed, but little variation is observed from 70 °C to 150 °C. For the n-channel device, the result is similar to the one presented in figure 5.3, as there is little noise power variation with temperature. Here, however, the Lorentzian-like lumps are clearly visible in the n-channel $1/f$ noise spectra due to the small device area (W=150 nm and L=50 nm). Interestingly, the lumps of the Lorentzian curves seem to shift with temperature, giving a complex relation with temperature. Also, due to these shifts, the frequency exponent varies significantly with temperature.

Figure 5.5: $1/f$ noise spectra of N- and PFETs at different temperatures.

Source: Lee et al. (2006)

Regarding the temperature dependence of the $1/f$ noise in FINFETs, the work of Wei, Xiong and Zhou (2009) provided experimental data on the issue, concluding that the noise characteristics of FINFETs are comparable to their bulk-CMOS counterparts.
Figure 5.6 displays the temperature dependence of the $1/f$ noise of two FINFETs, device 1 and device 2, in the linear region of operation.

Figure 5.6: Measured drain-voltage noise spectral density at 10 Hz versus temperature for two FINFETs in linear region of operation.

Their results indicate that for the thermal activation model is more convincing for FINFETs than the elastic tunnelling model. Figure 5.7 displays the measured and calculated frequency exponent using the Dutta and Horn relation, presented in equation (5.10) of the following section.

Figure 5.7: Frequency exponent as a function of temperature at 10 Hz for a) device 1 and b) device 2.

In the context of AlGaN/GaN HEMTs (high-electron-mobility transistors), Wang et al. (2017) demonstrated that, while bulk-mobility fluctuations are commonly applied to analyse $1/f$ noise in such devices, they do not properly describe the temperature, bias and frequency dependence of the noise. In addition, it was also reported that there is close agreement between experimental data and the Dutta-Horn model. The findings reinforce that the noise is well-described by number fluctuations and that the capture and emission process is thermally activated.
5.1 The Dutta-Horn Identity

The work of Dutta and Horn (1981) provides an important relation between the slope of the PSD spectrum (derivative of the noise spectrum in terms of frequency) and the derivative of the noise spectrum in terms of temperature. If the Dutta-Horn identity is satisfied, it is a strong indication that underlying mechanisms of trapping and de-trapping of carriers is thermally activated (DUTTA; HORN, 1981; SURYA; HSIANG, 1988; WEI; XIONG; ZHOU, 2009).

From Chapter 4, equation (4.23), if number fluctuations are assumed as the primary responsible for the $1/f$ noise in MOS devices, then the noise PSD follows

$$ S_{I_d}(\omega, T) \propto \int_{-\infty}^{\infty} \frac{\tau(E,x,y,z)}{1 + \omega^2 \tau(E,x,y,z)^2} \cdot f_{\tau} d\tau \ . \quad (5.1) $$

If thermal activation and position independence is assumed for $\tau$, then (5.1) can be rewritten as

$$ S_{I_d}(\omega, T) \propto \int_{-\infty}^{\infty} \frac{\tau_o \cdot e^{E_B/kT}}{1 + (\omega \tau_o \cdot e^{E_B/kT})^2} \cdot D(E_B) dE_B \ , \quad (5.2) $$

where $D(E_B)$ is the probability density function of the activation energies distribution. If $D(E_B)$ is constant within the window of interest, i.e. the activation energies are uniformly distributed, then $S(\omega, T)$ follows a $1/\omega$ behaviour. If $D(E_B)$ is not constant, however, the solution is not simple, though a solution is possible after expanding $D(E_B)$ in a Taylor series (DUTTA; HORN, 1981). When $D(E_B)$ varies slowly over the range of $\Delta E_B \approx kT$ then (5.2) reduces to

$$ S_{I_d}(\omega, T) \propto \frac{kT}{\omega} \cdot D(\overline{E_B}) \ , \quad (5.3) $$

where $\overline{E_B}$ is the activation energy at which the function

$$ h(\omega, T) = \frac{\tau_o \cdot e^{E_B/kT}}{1 + (\omega \tau_o \cdot e^{E_B/kT})^2} $$

peaks. This energy can be obtained by calculating the maximum of (5.4) through

$$ \frac{\partial h}{\partial E_B} = 0 \ , \quad (5.5) $$

resulting in

$$ \overline{E_B} = -kT \cdot \ln(\omega \tau_o) \ . \quad (5.6) $$
Rewriting (5.3) as

\[ D(E_B) \propto \frac{\omega}{kT} S_{I_D}(\omega, T), \tag{5.7} \]

it becomes clear that if \( \omega \) and \( T \) are varied in order to keep \( E_B \) unchanged in (5.6), then \((\omega/kT) \cdot S_{I_D}\) must also remain unchanged, since \(D(E_B)\) is constant. The conclusion is that there must be a ratio \( \Delta \omega / \Delta T \) that satisfy

\[ \left( \Delta \omega \cdot \frac{\partial}{\partial \omega} + \Delta T \frac{\partial}{\partial T} \right) \cdot \left[ -kT \cdot \ln(\omega \tau_o) \right] = 0 \tag{5.8} \]

and

\[ \left( \Delta \omega \cdot \frac{\partial}{\partial \omega} + \Delta T \frac{\partial}{\partial T} \right) \cdot \left[ \frac{\omega}{kT} \cdot S_{I_D}(\omega, T) \right] = 0. \tag{5.9} \]

Equations (5.8) and (5.9) essentially translate the previous ideas into their mathematical form. There should be a combination of \( \Delta \omega \) and \( \Delta T \) values that cause \( E_B \) to remain constant and, consequently, also \( D(E_B) \). Solving these equations results in the Dutta and Horn identity, given by

\[ \frac{\partial \ln [S_{I_D}(\omega, T)]}{\partial \ln (\omega)} = \frac{1}{\ln(\omega \tau_o)} \cdot \left[ \frac{\partial S_{I_D}(\omega, T)}{\partial \ln (T)} - 1 \right]. \tag{5.10} \]

Though the relation may seem surprising at first, relating the slope of the noise PSD (i.e. the derivative of the noise in terms of frequency) to the derivative of the noise PSD in terms of temperature, one should keep in mind that the characteristic time constant of the traps is exponentially dependent on temperature when thermal activation is assumed. Therefore, temperature increments cause the characteristic time constants to decrease, shifting the noise PSD towards higher frequencies. It is important to notice that the sum of finite individual traps do not produce perfect 1/f noise and, thus, small bumps are always present in the noise PSD. This means that increasing the temperature, these bumps “slide” towards higher frequencies.

The Dutta-Horn identity thus assumes that, at different temperatures, the shape of the PSD is preserved but shifted in frequency, so that the same slope of the PSD is found at different frequencies for different temperatures. This idea will be further explored in upcoming chapters of this work.

This relation has been used in a multitude of works (SURYA; HSIANG, 1988; XIONG et al., 2002; WEI; XIONG; ZHOU, 2009) to provide strong evidence that the 1/f noise originates from a thermally activated process. The drawback of this method is that it requires several noise measurements of a single transistor at different temperatures.
5.2 Summary and Discussion

While several works in the literature exist on the temperature dependence of the LFN, they do not necessarily converge to a single theory. While the Dutta-Horn identity provides a strong case for a thermally activated process, conventionally used compact models (such as BSIM, PSP, ACM and EKV) still employ elastic tunnelling to account for the distribution of time constants (HUNG et al., 1990b, 1990b; ARNAUD; GALUP-MONTORO, 2003; ARNAUD; GALUP-MONTORO, 2004; MAVREDAKIS et al., 2016). Moreover, statistical studies on the temperature dependence of the $1/f$ are lacking, as most results shown in this chapter are for individual devices.

Identifying these gaps, this work provides a methodology to study, statistically, the temperature dependence of the LFN. Based on the work of Dutta and Horn (1981) it is clear that a thorough statistical analysis of the frequency dependence of the fundamental LFN sources is also necessary, given its intimate relation to the temperature.
6 NOISE MEASUREMENTS

The DC and LFN measurements were performed on large populations (from 43 up to 320 devices) using the EDGE™ system from Cascade Microtech, depicted in Figure 6.1, at NXP Semiconductors, Eindhoven, The Netherlands. The measurements were supervised and conducted by Dr. Hans Tuinhout and technician Adrie Zegers-van Duijhoven. Three different CMOS technologies from different foundries — 140-nm, 65-nm and 40-nm — were analysed under several bias conditions. All devices were measured over a frequency range from 10 Hz to 100 kHz. Averaging was set to 20 times for the lowest band (10-100 Hz) and 80 times for the highest bands (TUINHOUT; DUIJNHOVEN, 2013). Unless stated otherwise, the measurements were conducted at 25 °C.

Figure 6.1: EDGE™ low-frequency noise measurement system.

![Source: Tuinhout and Duijnhoven (2013)](image)

The schematic for the low-frequency noise measurement is shown in Figure 6.2. The source and load resistors, \( R_s \) and \( R_{\text{load}} \), respectively, set the impedance conditions for the device-under test (DUT) during noise measurements and can be adjusted to accommodate different devices. The low noise amplifier (LNA) amplifies the noise signals from the DUT without significantly adding noise of its own. Its performance is determinant for the system noise floor. The filter units (low-pass filters) block the noise generated from the SMUs. The dynamic signal analyser (DSA) conditions and converts the signal to a digital format so that it can be processed by a computer, where the Fast Fourier Transform (FFT) is determined (HANSEN, 2009).

The frequency at which the frequency response of the system dominates the be-

---

1 Undisclosed for contractual reasons.
The behaviour of the noise PSD is called roll-off frequency. It can be calculated as

$$f_{rolloff} = \frac{1}{2\pi RC},$$

(6.1)

where $R$ is the Thevenin equivalent of the resistive loads and $C$ is the capacitance of the measuring system and cables. The load resistor can be optimized in order to increase the roll-off frequency.

Notice that the estimation of the PSD is performed using a non-parametric method and, therefore, is inherently noisy, particularly for low frequencies, where the averaging is lower (20 times instead of 80 times). Therefore, additional smoothing of the spectra was performed via MATLAB. Figure 6.3 illustrate the data before and after the smoothing.

For all measurements presented in this work, the statistical uncertainty is displayed through error bars obtained using bootstrapping, which is a technique to estimate the sample distribution that relies on random sampling with replacement. In this work, the size of the bootstrapping samples are the same as that of the original sample, as proposed by Efron (1979). Unless stated otherwise, the number of bootstrapping samples is equal to 5,000.
This chapter presents the measured data from a statistical standpoint. First, the
distribution of the LFN PSD at a given frequency is analysed, followed by its expected
value and variance. Then, the reproducibility of the measurements are verified, in order to
ascertain that the LFN PSD is unchanged with time (i.e. the noise process is wide-sense
stationary). The average power integrated over a bandwidth is also discussed. Finally, an
analysis of the spectra of small-area devices as a function of temperature is presented.

6.1 Distribution

Based on the works of Yu et al. (2012) and Banaszeski da Silva et al. (2014), as
well as experimental data obtained from several populations from three different tech-
nology nodes (140-nm, 65-nm and 40-nm), the LFN at a given frequency is found to be
approximately log-normally distributed. By definition, a log-normal distribution is nor-
mally distributed in log-scale, i.e. if $X$ is a log-normally distributed random variable, then
$Y = \ln(X)$ is normally distributed.

By establishing that LFN PSD at a given frequency is, approximately, log-normally
distributed, it is possible to estimate the expected value and variance of the distribution
using the maximum likelihood estimators of the log-normal distribution, as performed by
Banaszeski da Silva et al. (2014) and Banaszeski da Silva et al. (2016). These parameters
are

$$\mu = E[\ln(X)], \quad (6.2)$$

$$\sigma = \text{SD}[\ln(X)], \quad (6.3)$$

where $\text{SD}$ is the standard deviation operator. Using these parameters, the expected value
and variance of a log-normal distributed population can be estimated from the sampled
data using

$$E[X] = e^{\mu + \frac{1}{2}\sigma^2}, \quad (6.4)$$

$$\text{Var}[X] = \left(e^{\sigma^2} - 1\right)e^{2\mu + \sigma^2}. \quad (6.5)$$

Figures 6.4 to 6.6 display the quantile-quantile plot (QQ-Plot) of several LFN
measurements from different populations. The natural logarithm of the noise quantity
(data quantiles) is displayed on the y-axis. Using standard normal quantiles on the x-axis,
if the data falls on a straight line, it indicates that the data is normally distributed. Though small deviations are observed, it is clear that the measured data follow a straight line, indicating a normal behaviour, meaning the $S_{ID}$ of both N- and PFETs are well-described by a log-normal distribution.

Figure 6.4: Quantile plot of the LFN PSD at 20 Hz for six 43-DUT a) NFET and b) PFET populations (140-nm). Bias conditions are $V_{GS} = 1.8$ V and $V_{DS} = 0.1$ V.

Figure 6.5: Quantile plot of the LFN PSD at 20 Hz for six 60-DUT a) NFET and b) PFET populations (40-nm node). Bias conditions are $V_{GS} = 0.9$ V and $V_{DS} = 50$ mV.

Figure 6.6: Quantile plot of the LFN PSD at 20 Hz for three 282-DUT NFET populations (65-nm node). Bias conditions are $V_{GS} = 1.2$ V and $V_{DS} = 0.5$ V.
While measurements under low drain bias (uniformly charged channel conditions) usually display approximately log-normal behaviour, as shown in figures 6.4 to 6.6, deviations may appear under high drain bias, particularly for long-channel devices. Figure 6.7a illustrates the deviation observed in a 160-DUT 1 μm long NFET population. Bias conditions are $V_{GS} = 0.7$ V and $V_{DS} = 0.5$ V. From the quantile plot, it is clear that the left-side tail is lighter than the normal distribution, whereas the right-side tail is heavier. To put the length dependence into perspective, figure 6.7b displays the quantile plot of a 0.04 μm long NFET population under the same bias conditions. While the left-side tail is also lighter than the normal distribution, only a slight deviation on the right-side tail is observed. It should be noted that these deviations on the right-side tail are particularly important when considering transistor yield, as outliers (devices with large noise) will appear more often than expected if normality is assumed.

Figure 6.7: Quantile plots of the LFN PSD at 20 Hz for two 160-DUT NFET population (40-nm node): a) long-channel devices and b) short-channel device. Bias conditions are $V_{GS} = 0.7$ V and $V_{DS} = 0.5$ V.

An explanation for this phenomenon was given by Banaszeski da Silva et al. (2016). Under high drain bias (non-uniformly charged channel condition), the trap impact depends heavily on the trap position along the channel. For long-channel devices in saturation regime, traps located near the pocket implants, particularly at the drain side, have their impact amplified. This trap-impact enhancement due to trap position along the channel is important because it may increase the LFN variability of long-channel devices. Accurately modelling this effect is, therefore, crucial for proper noise simulation in electronic circuits.

In order to verify whether the RTS-based model discussed in Section 4.2.4 produce log-normal distribution, Monte Carlo simulations were performed. These simulations, which employed the assumptions discussed in Section 4.2.4, reinforce that the
LFN distribution is indeed approximately log-normal for most devices under uniformly inverted channel condition. The predicted behaviour for small-area devices (less than 1 µm²), however, deviates from log-normality. For these devices, the Monte Carlo simulation predicts a different distribution, as shown in figure 6.8a. Measured LFN PSD for smallest area devices available (0.3 × 0.04 µm²), however, do follow the approximately log-normal behaviour, as seen in figure 6.8b. Notice that the Monte Carlo simulation was performed using the same transistor size and the parameters extracted from figure 6.8b using equations (4.51), (4.54) and (4.57).

While no conclusive explanation for this distinct behaviour between simulation and measurement has been reached, the log-normal behaviour is approached when the average number of traps per device increases. If that is the case, then it would implicate that the exponential distribution of trap impacts is incorrectly assumed. There are several works that attempt to estimate or simulate the distribution of trap impacts (ASENOV et al., 2003; TEGA et al., 2009; TAKEUCHI et al., 2009), being the exponential and log-normal distributions the most prominent ones. Notice that for the Monte Carlo simulations performed, using the exponential distribution of trap impacts resulted the average number of traps to be sufficiently small so that approximately 1% of the simulated devices had no traps whatsoever.
6.2 Expected Value and Variability

In this section the bias dependence of the expected value and variance of the measured data is presented. For low drain bias (i.e. uniformly charged channel condition) the extracted values of $K$, described in eq. (4.57), are also presented. The experimental results are in agreement with typical $1/f$ noise behaviour of MOSFETs found in other works (SIMOEN; CLAEYS, 1999; MAVREDAKIS; ANTONOPOULOS; BUCHER, 2010).

6.2.1 140-nm Technology

The 140-nm technology devices measured consist of 24 populations (12 NFET and 12 PFET) composed of 43 devices each. Three drain voltages were applied (0.1 V, 0.5 V and 1.8 V) for several different gate biases, from weak inversion to strong inversion. The threshold voltage of the devices is around 0.4 V.

Figure 6.9: Expected value of the normalized LFN PSD at 20 Hz as a function of gate and drain bias for two 43-DUT NFET populations (140-nm technology). Both device populations have the same width (8 μm), but different lengths (a) 0.336 μm and b) 8 μm).

Figure 6.10: Expected value of the normalized LFN PSD at 20 Hz as a function of gate and drain bias for two 43-device NFET populations. Both device populations have the same length (0.16 μm), but different widths (a) 0.232 μm and b) 8 μm).
Figure 6.11: Expected value of the normalized LFN PSD at 20 Hz as a function of gate and drain bias for two 43-device PFET populations. Both device populations have the same width (8 µm), but different lengths (a) 0.336 µm and b) 8 µm).

Figure 6.12: Expected value of the normalized LFN PSD at 20 Hz as a function of gate and drain bias for two 43-device PFET populations. Both device populations have the same length (0.16 µm), but different widths (a) 0.232 µm and b) 8 µm).

The expected value of the LFN PSD as a function of gate voltage for four different geometries is presented in Figures 6.9 to 6.12. From moderate to strong inversion, the expected value of \( S_{ID}/I^2_D \) shows a stronger bias dependence for NFET than for PFETs.

The standard deviation of \( \ln\left(\frac{S_{ID}}{I^2_D}\right) \) as a function of area, Figure 6.13, did not follow a \( 1/\sqrt{WL} \) dependence, but rather one that follows equation (4.56) (BANASZESKI DA SILVA et al., 2014).

Figure 6.13: Standard deviation of \( \ln(S_{ID}/I^2_D) \) as a function of a) \( 1/\sqrt{WL} \) and b) \( WL \) for ten 43-device NFET populations. Bias conditions are \( V_{DS} = 0.1 \) V and \( V_{GS} = 1.4 \) V.
Under uniformly charged channel conditions, a constant $K$ parameter — i.e. the normalized variance multiplied by the area, discussed in section 4.2.4.2 — is able to describe the LFN variability for all geometries, following (4.56) (BANASZESKI DA SILVA et al., 2014). The $K$ parameter extracted from ten 43-device NFET populations, each with a different geometry, is shown in Figure 6.14.

Figure 6.14: Extracted and fitted $K$ parameter as a function of area $WL$ for ten 43-device NFET populations. Bias conditions are $V_{DS} = 0.1$ V and $V_{GS} = 1.4$ V.

Under high drain bias, however, the enhancement of trap impacts near the drain region causes long channel devices to exhibit larger-than-predicted variability. Figure 6.15 displays the standard deviation of $\ln[\frac{S_{ID}/I_{D}^{2}}{\sigma}]$ as a function of area under high drain bias, in which it is visible that the variability in long-channel devices deviates from the behaviour predicted under uniformly inverted channel condition.

Figure 6.15: Standard deviation of $\ln[\frac{S_{ID}/I_{D}^{2}}{\sigma}]$ as a function of area for two 43-device NFET populations. Bias conditions are a) $V_{DS} = 0.5$ V and $V_{GS} = 1.4$ V and b) $V_{DS} = 1.8$ V and $V_{GS} = 1.4$ V. Deviations from the predicted behaviour are observed for long channel devices when $V_{DS} > V_{GS} - V_{T}$.
6.2.2 65-nm Technology

For the 65-nm technology, devices from two different foundries (A and B) were measured. From foundry A, two 68-device NFET populations were measured at 4 different bias conditions. From foundry B, three 282-device NFET populations were measured, also at 4 different bias. The expected value of the normalized LFN PSD as a function of the $I_D \cdot L/W$ is shown in Figure 6.16. Note that for comparing two different technology processes, using the normalized drain current instead of the gate voltage is more convenient.

Notice that the noise levels are similar for both technologies (devices from foundry B show slightly lower noise levels). Additionally, both technology processes display similar $K$ parameter (normalized variability), as shown in Figure 6.17, indicating that $N_{tr}/\gamma'$ is similar for both technology processes. The increased variability observed for the $10 \times 10 \mu m^2$ at lower gate biases is attributed to the increased impact of the halo implants in this region of operation (BANASZESKI DA SILVA et al., 2016).

Figure 6.16: Expected value of the normalized LFN PSD at 20 Hz as a function of the normalized drain current ($I_D \cdot L/W$) for two 68-device NFET populations from foundry A and three 282-device NFET populations from foundry B ($V_T \approx 0.4$ V).

Figure 6.17: $K$ parameter at 20 Hz as a function of the normalized drain current ($I_D \cdot L/W$) for two 68-device NFET populations from foundry A and two 282-device NFET populations from foundry B (65-nm node).
6.2.3 40-nm Technology

The 40-nm technology devices measured consist of 10 populations (5 NFET and 5 PFET) composed of 60 devices each. These populations were measured at three drain voltages (50 mV, 0.55 V and 1.1 V) for several different gate biases, from weak to strong inversion. Additionally, two 180-device NFET populations were also measured under $V_{DS} = 0.5$ V. The threshold voltage of these devices is approximately 0.4 V.

The expected value of the LFN PSD as a function of gate voltage for two different geometries is presented in Figures 6.18 and 6.19. As previously observed for the 140-nm technology, from moderate to strong inversion, the expected value of $S_{I_D}/I_D^2$ shows a stronger bias dependence for NFET than for PFETs.

Figure 6.18: Expected value of the normalized LFN PSD at 20 Hz as a function of gate and drain bias for two 60-DUT NFET populations. Both device populations have the same width (1 $\mu$m), but different lengths a) 1 $\mu$m and b) 0.2 $\mu$m).

Figure 6.19: Expected value of the normalized LFN PSD at 20 Hz as a function of gate and drain bias for two 60-DUT PFET populations. Both device populations have the same width (1 $\mu$m), but different lengths a) 1 $\mu$m and b) 0.2 $\mu$m).

In figure 6.20, the expected value of the normalized LFN PSD at 20 Hz as a function of drain current and temperature is shown. The advantage of plotting versus drain
current in this case is the fact that the threshold voltage shifts with temperature and, thus, the bias condition \( V_{GS} - V_T \) is modified. Notice that the expected value of the LFN is weakly temperature dependent, specially in strong inversion.

Figure 6.20: Expected value of the normalized LFN PSD at 20 Hz as a function of drain current and temperature for two 80-DUT NFET populations.

Following the behaviour observed for the 140-nm technology, the standard deviation of \( \ln \left( \frac{S_{ID}}{I_D^2} \right) \) as a function of area, Figure 6.21, did not follow a \( 1 / \sqrt{WL} \) dependence, but rather one that predicted by the log-normal model. Notice that the normalized variability parameter \( (K) \) extracted for PFET devices is smaller than that of NFET devices. From equation 4.57, reproduced here for simplicity, it is possible to conclude that the trap density in the area, energy and in the log-domain of time constants (cm\(^{-2}\)eV\(^{-1}\) per neper), \( N_{tr}(E_{Fn})/\gamma \), is larger in PFET devices than in NFET devices for this particular technology at the given bias condition.

\[
K = \frac{2}{\pi^2 kT} \frac{\gamma'}{N_{tr}(E_{Fn})}. \tag{6.6}
\]

Figure 6.21: Standard deviation of the natural logarithm of \( S_{ID}/I_D^2 \) for five NFET and PFET devices populations as a function of \( 1 / \sqrt{WL} \) for \( V_{DS} = 50 \text{ mV} \) and \( V_{GS} = 1.1 \text{ V} \).
It is important to observe that the normalized LFN PSD \( \left( \frac{S_{I_D}}{I_D^2} \right) \) is not correlated to the drain current, as shown in Figure 6.22 for two 180-DUT NFET populations. It is also possible to confirm that drain current variability is not the primary source of LFN variability, as the drain current variability is several orders magnitude lower than that of LFN.

6.3 Reproducibility

For the analyses presented in chapter 7 of this work, it is critical that the measured LFN PSD of each device remains the same if the experiment is reproduced after a few hours or days (under the same bias conditions). Particularly for the correlation coefficient analyses presented in Chapter 7 of this work, it is critical to ascertain that defects are not created or annealed during device measurement or wafer storage, i.e. the PSD of the LFN remains roughly unchanged over time.

To verify the reproducibility of the PSD measurements within the desired bandwidth, a population of 282 65-nm technology devices \((W = 1 \ \mu m \text{ and } L = 0.06 \ \mu m)\) was measured twice with seven weeks of interval between experiments. The average and variance of \( f \cdot \frac{S_{I_D}}{I_D^2} \) at different bias conditions is shown in Figures 6.23 and 6.24. The multiplication of \( \frac{S_{I_D}}{I_D^2} \) by the frequency is for compactness of the figures. The correlation coefficient as a function of frequency are shown in Figure 6.25. Notice that drain bias is kept low in order to avoid defect creation due to hot carrier injection during measurement.

The near-identical behaviour of the expected value and variance, as well as the high value of correlation coefficient (above 0.95) between experimental results indicate
that the measurement set-up is stable and that the LFN PSD of the population at a given frequency can be considered a stationary random process, i.e. the average and variance do not change over time.

Figure 6.23: Average value of $f \cdot S_{I_D}/I_D^2$ for different frequencies and biases. Blue and red lines indicate the first and second measurements, respectively. Despite the time interval between experiments, the average value falls within the statistical uncertainty obtained using bootstrapping. Drain voltage is 0.5 V.
Figure 6.24: Variance of $f \cdot S_{ID}/I_D^2$ for different frequencies and biases. Blue and red lines indicate the first and second measurements, respectively. Despite the time interval between experiments, the average value falls within the statistical uncertainty obtained using bootstrapping. Drain voltage is 0.5 V.

Figure 6.25: Correlation coefficient of $S_{ID}/I_D^2$ between experiments 1 and 2 as a function of frequency for different bias. Drain voltage is 0.5 V.
6.4 Noise Power Integrated over a Bandwidth

The noise power of a given device integrated over a bandwidth \([f_L, f_H]\) is given by

\[
W_p = \int_{f_L}^{f_H} S_{ID}(f) df. \tag{6.7}
\]

The work of Silva, Wirth and Brederlow (2006) demonstrated that for any frequency range region the mean noise as well as the relative error are invariant under a scale transformation \(f_H \rightarrow c f_H\) and \(f_L \rightarrow c f_L\), where \(c\) is a constant. Following the assumptions from Section 4.2.4, under the premise that \(f_L > 1/2\pi \tau_{max}\) and \(f_H < 1/2\pi \tau_{min}\) (i.e. the observation window is within the log-uniform distribution of time constants), it can be demonstrated that

\[
E[W_p] = 2E[\Delta I_D]^2 \cdot kT W L \cdot \frac{N_{tr}(E_F)}{\gamma} \cdot \ln \left(\frac{f_H}{f_L}\right), \tag{6.8}
\]

where \(N_{tr}(E_F) / \gamma'\) is the trap density in the area, energy and in the log-domain of time constants (\(\text{cm}^{-2}\text{eV}^{-1}\) per neper).

Experimental results from two different technologies (40-nm and 65-nm) shown in Figures 6.26 and 6.27 corroborate this finding, given that the quantity \(E[W_p] / \ln(f_H / f_L)\) is approximately constant, regardless of the bandwidth. Notice that, in Figure 6.26, the same population was measured twice, at different drain bias conditions, whereas two different populations (i.e. different geometries) were measured at the same bias conditions in Figure 6.27.

Figure 6.26: Expected value of the integrated noise normalized by \(I_D^2\) and \(\log(f_H / f_L)\) over different bandwidths for a 78-DUT 40-nm technology NFET population \((W = 1 \mu m\) and \(L = 1 \mu m)\) under two different drain biases. Dashed red line is a guide for the eye.
Figure 6.27: Expected value of the integrated noise normalized by $I_D^2$ and $\log(f_H/f_L)$ over different bandwidths for two 282-DUT 65-nm technology NFET population under $V_{GS} = 1.2$ V and $V_{DS} = 1.2$ V. Dashed red line is a guide for the eye.

Notice, however, that if the noise is not exactly $1/f$ (i.e. when the distribution of time constants is not exactly uniform in log-scale), the relation is not valid. One example is shown in Figure 6.28 for a 140-nm technology PFET population, in which the frequency exponent is larger than 1.

Figure 6.28: Expected value of the integrated noise normalized by $I_D^2$ and $\log(f_H/f_L)$ over different bandwidths for two 43-DUT 140-nm technology PFET population under $V_{GS} = 1.8$ V and $V_{DS} = 1.8$ V. Dashed red line in a) represents an ideal $1/f$.

6.5 Analysis of Individual Lorentzians

The identification and analysis of individual Lorentzian spectra enables the verification of basic assumptions, such as the log-uniform distribution of time constants. Each clearly visible Lorentzian spectrum is attributed to the RTS of a single trap, enabling
the extraction of the characteristic time constant of such trap. By multiplying \( S_{I_D}/I_{D}^2 \) by the frequency, the cut off frequency of the trap \( (f_c) \) is found at the peaks of \( f \cdot S_{I_D}/I_{D}^2 \). The characteristic time constant can then be estimated through \( \tau = 1/(2\pi f) \). There are drawbacks to such method though — e.g. Lorentzian are only clearly identifiable in small-area devices and traps that have similar characteristic time constant can be difficult to distinguish —, which are further discussed in this section. For this reason, in chapter 7 a methodology is proposed to extract important information from the LFN spectra without the need of identifying individual Lorentzians; the methodology not only enables analysing large-area devices, but also for small-area devices without visible Lorentzians.

From a 320-DUT NFET population \((W = 0.3\mu m \text{ and } L = 0.04\mu m)\), several bumpy, Lorentzian-like, spectra were identified from the LFN PSD at two different temperatures \((25 \, ^\circ C \text{ and } 50 \, ^\circ C)\). The window of observation was set between 20 Hz and 20 kHz. This is due to the fact that at lower frequencies the number of samples is small and, thus, uncertainty is large; at higher frequencies, on the other hand, the noise level approaches the noise floor of the measurement set-up. In Figure 6.29, the LFN PSD of devices with pronounced Lorentzians at both temperatures are shown. Notice that the cut off frequencies of the Lorentzian is increased with increasing temperature, meaning that the characteristic time constant decreased, following a thermally activated process.

**Figure 6.29:** LFN spectra of two NFET devices at the same bias conditions and at two different temperatures.

On the other hand, while a Lorentzian spectrum is visible at both temperatures in Figure 6.30, there is little observable shift in the cut-off. This temperature independent behaviour is in line with that expected for direct tunnelling. Overall, however, the majority of the identified Lorentzian spectra seemed to follow a thermally activated behaviour, which is explored further in this section.
Figure 6.30: LFN spectra of an NFET device at the same bias conditions and at two different temperatures.

Analysing the behaviour of these Lorentzians individually from an ensemble, however, is not simple. Even in small-area devices, these Lorentzians are not always easily identifiable. Their spectra may be less prominent due to, for example, the presence of other Lorentzians with $\tau$ near or within the observation window. Figure 6.31 illustrates two devices with less pronounced, yet identifiable, profiles.

Figure 6.31: LFN spectra of two NFET devices at the same bias conditions and at two different temperatures.

A significant share of devices exhibited no Lorentzians at all at either temperatures. Such devices displayed a approximately $1/f$ noise spectrum, as illustrated in Figure 6.32. Finally, in order to illustrate the difficulty of identifying these individual Lorentzians, Figure 6.33 displays the LFN spectra of a device at 25 ºC and 50 ºC in which the lumps that are not visible to the naked eye may have been incorrectly identified by the algorithm.

From the 320-DUT population, 187 and 193 individual Lorentzians were identified at 25 ºC and 50 ºC, respectively. Figure 6.34 displays the cumulative probability plot of $\log_{10}\tau$ extracted from the measured PSD at 25 ºC and 50 ºC. The linear behaviour indicates the characteristic time constant is approximately uniformly distributed in log scale within the window of observation. From the frequency range defined (20 Hz to 20 kHz), the uniform distribution of $\log_{10}\tau$ spans from $\tau_{\text{min}} = 1/(2\pi f_{\text{max}})$ to $\tau_{\text{max}} = 1/(2\pi f_{\text{min}})$. 
Figure 6.32: LFN spectra of a device at the same bias conditions at two different temperatures. No visible Lorentzians.

Figure 6.33: LFN spectra of a device at the same bias conditions at two different temperatures. No visible Lorentzians.

The observed log-uniform distribution of $\tau$ reinforce the relation between $1/f$ noise and trap activity, since such distribution produces a $1/f$ spectrum when multiple Lorentzians are present within the observation window (SURDIN, 1939; MCWHORTER, 1957). It is important to observe that, by identifying prominent Lorentzians, only information about traps that contribute significantly to noise power is obtained; thus, it is not possible to ascertain, from this analysis, that the $1/f$ noise in devices in which no Lorentzian profile was identified is also produced by trap activity.

Figure 6.34: Cumulative probability plot compared to the expected uniform distribution of $\log_{10} \tau$. The observation window is defined between 20 Hz and 20 kHz.
Figure 6.35: Scatter plot of $\log_{10}\tau$ at 25 ºC and 50 ºC for the condition when $\tau(25^\circ C) > \tau(50^\circ C)$, meaning that the Lorentzian is shifted towards higher frequencies with increasing temperature (indicating thermal activation). The number of Lorentzians that satisfy the restriction is 131.

Figure 6.35 displays the scatter plot of $\tau(25^\circ C)$ versus $\tau(50^\circ C)$, aiming to identify whether the characteristic time constants extracted at 25 ºC and 50 ºC are correlated. Assuming thermal activation, then $\tau(25^\circ C) > \tau(50^\circ C)$; thus, only Lorentzians that satisfy this constraint were included in the analysis (a total of 131 spectra). While the correlation is strong ($\rho = 0.94$), the fact that only Lorentzians that satisfied the constraint were included in the analysis shuns any attempts to generalize the behaviour observed for the LFN as a whole. Nevertheless, there is strong indication that thermal activation governs the activity of the observed traps.

Conversely, if $\tau$ is assumed to originate from a thermally activated process, then the activation energy can be estimated using

$$E_B = \ln \left( \frac{\tau(T_1)}{\tau(T_2)} \right) \cdot \frac{kT_1T_2}{(T_2 - T_1)}.$$  \hspace{1cm} (6.9)

If equation (6.9) is applied to data from 6.35, the resulting distribution is not uniform (as shown in figure 6.36). This result indicates that simple thermal activation may not be enough to account for the temperature dependence of $\tau$.

Figure 6.36: Cumulative probability plot of the activation energy extracted from (6.9). The distribution is clearly not uniform.
6.6 Summary and Discussion

This chapter provided a discussion on the measured LFN spectra from three different technology nodes (40-nm, 65-nm and 140-nm). The noise distribution at a given frequency, expected value and variance were evaluated, as well the integrated power over a bandwidth. The results were consistent with other works in the literature. Additionally, noise measurements were shown to remain stable when measured after the span of a few weeks.

The analysis of individual Lorentzians from the LFN spectra measured reinforces the view that thermal activation governs the capture and emission process. However, it is insufficient for determining whether trapping/de-trapping is the dominant LFN mechanism; whether thermal activation is primarily responsible for the log-uniform distribution of time constants; or whether the conclusions hold true when no Lorentzians are visible ($1/f$ spectrum). In order to provide a more robust analysis, a statistical methodology is developed and presented in the following chapter of this work.
7 CORRELATION COEFFICIENT ANALYSES

In this chapter, a novel noise analysis technique which is extremely sensitive to the physical mechanisms assumed for the interpretation and modelling of the LFN in MOSFETs is proposed. This technique is based on the analysis of the autocorrelation of the LFN PSD, meaning the analysis is performed entirely in frequency domain. Using this methodology it is possible to draw valuable conclusions about the frequency, temperature and bias dependence of the fundamental noise sources underlying the $1/f$ noise in MOSFETs, as well as the origin of the trapping mechanism.

Notice that the term autocorrelation here refers to the fact that the technique estimates the correlation between the noise power density of a given population at a given condition (bias, temperature, frequency), with the same population at a different condition — e.g. $S_{I_D}(f)$ and $S_{I_D}(f + \Delta f)$. While, in principle, such autocorrelation could be defined as autocorrelation functions (in frequency domain), the author refrains from using the term in order to avoid misinterpretation given the classic usage of the term for time functions 1.

7.1 Correlation Coefficient Extraction

As discussed in section 6.1, the LFN PSD follows a log-normal distribution. Therefore, the extraction of the correlation coefficient from experimental data was performed using the maximum likelihood estimators from the log-normal distribution. From (4.58), rewritten for convenience, the correlation coefficient between two random variables, $X$ and $Y$, is given by

$$R[X,Y] = \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X] \cdot \text{Var}[Y]}}. \quad (7.1)$$

Using the definition of Covariance, (4.69), and the properties of the log-normal distribution

$$E[X] = e^{\mu_x + \frac{1}{2} \sigma^2}, \quad (7.2)$$

$$\text{Var}[X] = \left(e^{\sigma^2_x} - 1\right)e^{2\mu_x + \sigma^2_x}, \quad (7.3)$$

1 See section 2.2.
where $\mu_x$ and $\mu_y$ are the expected value of the logarithm of the random variables $X$ and $Y$, respectively (notice that $\ln(X)$ and $\ln(Y)$ are normally distributed); $\sigma_x$ and $\sigma_y$ are the standard deviations of $\ln(X)$ and $\ln(Y)$, respectively; and $\rho$ is the correlation coefficient between $\ln(X)$ and $\ln(Y)$, (7.1) can be rewritten as

$$R[X, Y] = \frac{e^{\rho \sigma_x \sigma_y} - 1}{\sqrt{(e^{\sigma_x^2} - 1) \cdot (e^{\sigma_y^2} - 1)}}$$

for log-normal distributions. This expression will be used throughout this chapter to extract the correlation coefficient between measured LFN PSD populations.

### 7.2 Frequency Autocorrelation

The frequency autocorrelation analysis provide valuable information on the frequency dependence of the fundamental noise sources underlying the $1/f$ noise. In order to verify whether the assumption that the $1/f$ spectrum observed in large-area devices is indeed composed of a superposition of Lorentzians, an analytical expression for the correlation coefficient in terms of $\Delta f$ is derived from the RTS framework and compared to the data.

#### 7.2.1 Model Derivation

The correlation coefficient of the LFN PSD in terms of $\Delta f$ when the LFN is assumed to result from multiple switching traps — i.e. multiple superimposed Lorentzian spectra, $L(f)$ — can be calculated theoretically using (4.71) as

$$R \left[ S_{Id} (f_{ref}), S_{Id} (f_{ref} + \Delta f) \right] = \frac{E[L(f_{ref}) \cdot L(f_{ref} + \Delta f)]}{\sqrt{E[L(f_{ref})^2] \cdot E[L(f_{ref} + \Delta f)^2]}}$$

where $L(f)$ is given in the context of RTS by (4.14), rewritten here for convenience

$$L(f) = 4\Delta I_D^2 \cdot \frac{\beta}{(1 + \beta)^2} \cdot \frac{2\pi f}{1 + (2\pi f \tau)^2}.$$
where $\Delta I_D$, $\beta$, and $\tau$ are the current fluctuation caused by a trap, the ratio between the emission and capture time constants ($\tau_e / \tau_c$) of a trap and the characteristic time constant of a trap $(1 / \tau = 1 / \tau_c + 1 / \tau_e)$, respectively. The ratio between emission and capture times is given by $\beta = \exp [(E_F - E_T) / kT]$, where $E_T$, $E_F$, $k$, and $T$ are the trap energy, Fermi energy, Boltzmann’s constant, and temperature, respectively. For simplicity, $f_{\text{ref}}$ and $f_{\text{ref}} + \Delta f$ are substituted for $f_1$ and $f_2$, respectively.

The problem is thus reduced to the solution of $E[L(f_1) \cdot L(f_2)]$ and $E[L(f_1)^2]$. Using

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X dx , \quad (7.8)$$

yields

$$E[L(f_1)^2] = 16 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta I_D^4 \cdot \frac{\beta^2}{(1 + \beta)^2} \cdot \frac{\tau^2}{[1 + (2\pi f_1 \tau)^2]^2} \cdot f_{\Delta I_D} f_{\beta} f_{\tau} d\Delta I_D d\beta d\tau , \quad (7.9)$$

and

$$E[L(f_1) \cdot L(f_2)] = 4 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta I_D^4 \cdot \frac{\beta^2}{(1 + \beta)^2} \cdot \frac{\tau^2}{[1 + (2\pi f_1 \tau)^2]^2} \cdot [1 + (2\pi f_2 \tau)^2] \cdot f_{\Delta I_D} f_{\beta} f_{\tau} d\Delta I_D d\beta d\tau . \quad (7.10)$$

If the random variables $\Delta I_D$, $\beta$ and $\tau$ of any given trap are assumed to be frequency independent (meaning they remain constant at all frequencies), $\tau$ to be log-uniformly distributed between $\tau_{\text{min}}$ and $\tau_{\text{max}}$, and the channel to be uniformly inverted (no position dependence of the traps), (7.6) simplifies to

$$R[S(f_1), S(f_2)] = \frac{\int_{-\infty}^{\infty} \frac{\tau^2}{[1 + (2\pi f_1 \tau)^2]^2} f_{\tau} d\tau}{\sqrt{\int_{-\infty}^{\infty} \frac{\tau^2}{[1 + (2\pi f_1 \tau)^2]^2} f_{\tau} d\tau \cdot \int_{-\infty}^{\infty} \frac{\tau^2}{[1 + (2\pi f_2 \tau)^2]^2} f_{\tau} d\tau}} , \quad (7.11)$$

which can be solved to

$$R[S(f_1), S(f_2)] = 2 \cdot \frac{\ln \left( \frac{f_1}{f_2} \right)}{(f_1^2 - f_2^2)} f_1 f_2 . \quad (7.12)$$

This result is particularly noteworthy because it is free of parameters; the shape of
R only depends on the spectra of the fundamental noise sources and their superposition. If, for instance, the observed $1/f$ noise would be the sum of microscopic $1/f$ noise sources with a perfect slope of -1, R would be equal to 1 for all $\Delta f$. Therefore, this correlation coefficient is a signature of the frequency dependence of the fundamental noise sources constituting the $1/f$ noise.

Equation (7.12) can be written in terms of $\Delta f'$, which is the ratio between $f_1$ and $f_2$, i.e. $\Delta f' = f_1/f_2$. In this case, (7.12) simplifies to

$$R[S(f_1), S(f_2)] = 2 \cdot \frac{\ln(\Delta f')}{(1 - \Delta f'^2)} \cdot \Delta f'$$

(7.13)

highlighting the fact that the shape of the correlation coefficient curve is independent of the reference frequency, $f_{ref}$ when analysed in log-scale. Figure 7.1 depicts the correlation coefficient as a function of $\Delta f'$.

Figure 7.1: Correlation coefficient as a function of $\Delta f'$.

The correlation coefficient decreases as $\Delta f$ increases. This means that for small frequency increments (small $\Delta f$), devices that have the higher noise levels at $f_{ref}$ are more likely to have the higher noise levels at $f_{ref} + \Delta f$. Conversely, for large frequency increments (large $\Delta f$), as the correlation coefficient approaches zero, it is not possible to predict which devices have the higher noise levels at $f_{ref} + \Delta f$ based on their levels at $f_{ref}$.

The bell-like shape of the correlation coefficient curve is directly related to the frequency dependence of the fundamental noise sources underlying the LFN.
7.2.2 Experimental Results

Using equation (7.5), the correlation coefficient as a function of frequency was extracted for several N- and PFET populations from different technology nodes. In this section, the extracted correlation coefficient is compared to that obtained analytically, (7.12), and the results are discussed.

For small-area devices, the impact of individual traps is more significant, resulting in prominent lumps in the spectrum due to the Lorentzian-like PSD of a single trap. For these devices, however, telegraphic signals are usually visible in time domain measurements; therefore, there is usually consensus that RTS is the dominant noise mechanism in small-area devices. Figure 7.2 displays the LFN PSD of several small-area NFETs, highlighting Lorentzian-like spectra. For large-area devices as well as for PFETs, on the other hand, the primary mechanism underlying the LFN is debated (HAARTMAN; ÖSTLING, 2007; VANDAMME; HOOGE, 2008). While Lorentzian-like spectra may not be clearly visible, in the previous section it was shown that, if RTS is the primary mechanism underlying the LFN in MOSFETs, then experimental data should satisfy (7.12). The findings for 40-nm, 65-nm and 140-nm technologies are reported.

Following (7.12), the correlation coefficient is independent of the trap impact ($\Delta I_D$). One should note, however, that the LFN distribution deviates from the log-normal distribution under high drain bias, as shown in section 6.1. Thus, in order to avoid distortions due to these deviations, comparisons are made, preferably, under low drain bias (uniformly charge channel condition).

Figure 7.2: LFN spectra of 320 small-area devices and highlighted Lorentzian-like spectra.
7.2.2.1 40-nm Technology

The population presented in Figure 7.2 has prominent Lorentzian-like spectra, indicating that RTS is indeed the primary cause of LFN. The comparison between extracted and calculated correlation coefficient is shown in Figure 7.3. As expected, (7.12) provides an accurate prediction of the dependence of the correlation coefficient with frequency, indicating, from a statistical perspective, that the RTS is the dominating noise mechanism in this population. For a similar small-area NFET population (albeit fewer samples), shown in Figure 7.4, the theoretically calculated correlation coefficient also gives a reasonable prediction of the measured correlation coefficient.

Figure 7.4: a) LFN spectra of a 60-DUT small-area NFET population (40-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are W=1 µm and L=0.04 µm; bias conditions are $V_{GS} = 0.9$ V and $V_{DS} = 0.05$ V.
Interestingly, when a PFET population is analysed (Figure 7.5), Lorentzian-like spectra are also visible. Despite the reduced number of prominent lumps, the measured correlation coefficient is also well predicted by the model. This is a strong indication that the primary LFN mechanism in PFETs is the capture and emission of charge carriers.

For large-area devices (Figures 7.6 and 7.7), Lorentzian-like spectra are hardly visible for both N- and PFET populations. The correlation coefficient, however, follows the RTS-predicted behaviour. This is a strong indication that the $1/f$ noise in large-area devices is composed by a superposition of Lorentzians, similarly to small-area devices. Notice that the de-correlation with increasing $\Delta f$ means that the spectra do not have a constant slope; they intertwine at a certain rate which is well described by the RTS model.
Figure 7.7: a) LFN spectra of a 60-DUT large-area PFET population (40-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are $W=16 \, \mu m$ and $L=0.2 \, \mu m$; bias conditions are $V_{GS} = 1.1 \, V$ and $V_{DS} = 0.05 \, V$.

7.2.2.2 65-nm Technology

Since the correlation coefficient is independent of transistors parameters and technology process, different technology processes should yield similar results. For the small-area NFET population shown in Figure 7.8, Lorentzian-like spectra are clearly visible. As expected, the correlation coefficient is accurately predicted by the RTS-based model, despite the larger drain voltage when compared to that used for the 40-nm technology.

Figure 7.8: a) LFN spectra of a 282-DUT small-area NFET population (65-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are $W=1 \, \mu m$ and $L=0.06 \, \mu m$; bias conditions are $V_{GS} = 1.2 \, V$ and $V_{DS} = 0.5 \, V$.

For a large-area NFET population (Figure 7.9), the extracted correlation coefficient also indicates that a superposition of Lorentzians is the primary source of LFN, despite the small number of visible lumps. Notice that, for this population, measurements
above 10 kHz are unreliable due to the noise level being close or below the measurement noise floor.

Figure 7.9: a) LFN spectra of a 282-DUT large-area NFET population (65-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are W=10 µm and L=10 µm; bias conditions are \( V_{GS} = 1.2 \text{ V} \) and \( V_{DS} = 0.5 \text{ V} \).

**7.2.2.3 140-nm Technology**

Similarly to the results obtained for the previous technologies, both small-area N- and PFET populations display Lorentzian-like spectra (Figures 7.10 and 7.11), indicating RTS-based LFN.

Figure 7.10: a) LFN spectra of a 43-DUT small-area NFET population (140-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are W=0.232 µm and L=0.336 µm; bias conditions are \( V_{GS} = 1.0 \text{ V} \) and \( V_{DS} = 0.1 \text{ V} \).

For the large-area devices (7.12 and 7.13), while the spectra are approximately \( 1/f \), i.e. lumps are not visible in the spectra, the correlation coefficient follows the be-
Figure 7.11: a) LFN spectra of a 43-DUT small-area PFET population (140-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are $W=0.232 \ \mu m$ and $L=0.336 \ \mu m$; bias conditions are $V_{GS} = 1.8 \ \text{V}$ and $V_{DS} = 0.1 \ \text{V}$.

Figure 7.12: a) LFN spectra of a 43-DUT large-area NFET population (140-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are $W=8 \ \mu m$ and $L=1 \ \mu m$; bias conditions are $V_{GS} = 1.0 \ \text{V}$ and $V_{DS} = 0.1 \ \text{V}$.

These results reinforce the validity of LFN models based on the summation of individual telegraphic signals resulting in a $1/f$ spectra. If there is a fundamental $1/f$ noise source underlying the LFN — from bulk mobility fluctuations, for instance —, its effects are overshadowed by the Lorentzian-like PSD of the multiple RTS, as observed by the autocorrelation coefficient analysis.
Figure 7.13: a) LFN spectra of a 43-DUT large-area PFET population (140-nm technology); and b) extracted (blue markers) versus theoretically calculated (red line) correlation coefficients. Transistor dimensions are $W=8 \ \mu m$ and $L=0.16 \ \mu m$; bias conditions are $V_{GS} = 1.8 \ \text{V}$ and $V_{DS} = 0.1 \ \text{V}$.

7.3 Temperature Autocorrelation

The autocorrelation can also be analysed in terms of temperature shifts. Even though $1/f$ noise is known for being weakly temperature dependent, the sensitivity of the autocorrelation technique proves useful to draw conclusions regarding the temperature dependence of the characteristic time constant of the traps.

Three different trapping models were considered for comparison. The first is the direct-tunnelling model (MCWHORTER, 1957), in which

$$
\tau = \tau_o \cdot \exp(\gamma z),
$$

(7.14)

with $z$ and $\gamma$ being the trap depth in the oxide, and the wave function attenuation factor, respectively. In this model, $\tau$ is unaffected by temperature shifts, causing the PSD distribution to be strongly correlated to itself, regardless of $\Delta T$. This assumptions hold true if the Fermi energy is weakly affected by the temperature shifts, which is reasonable for strong inversion.

The second model introduces thermal activation (DUTTA; HORN, 1981; SURYA; HSIANG, 1988) using

$$
\tau = \tau_o \cdot \exp \left( \frac{E_B}{kT} \right),
$$

(7.15)

where $E_B$ is an activation energy. In this model, increasing temperature reduces $\tau$, shifting the entire LFN PSD towards higher frequencies, causing de-correlation at $f_{ref}$. Notice that for a wide spread of activation energies, this shift means the number of traps in a given
interval of time constants increases.

The third model is the Kirton and Uren model (KIRTON; UREN, 1989). It combines tunnelling and thermal activation, and is given by

\[
\tau = \frac{1}{n v_{th} \sigma_0} \cdot \exp \left( \gamma z + \frac{E_B}{kT} \right) \cdot \frac{\beta}{1 + \beta},
\]

(7.16)

where \( n \), \( v_{th} \), and \( \sigma_0 \) are the carrier concentration, thermal velocity and capture cross-section pre-factor, respectively. For all models, \( z \) and \( E_B \) were assumed uniformly distributed. As observed by (KIRTON; UREN, 1989; VANDAMME; HOOGE, 2008), (7.16) does not produce a log-uniformly distributed random variable for \( E_B \) and \( z \) uniform, though the resulting PSD is approximately \( 1/f \).

Figure 7.14 shows the measured correlation coefficient between different temperatures for two NFET populations. The theoretical correlation coefficient was calculated numerically using the three models introduced above.

Figure 7.14: Correlation coefficient as a function of temperature (\( T_{ref} = 25 \) °C) at 100 Hz for a) two independent 80-DUT NFET populations (1 × 0.04 µm²) and b) two independent 78-DUT NFET populations (1 × 1 µm²). The first population of each geometry was measured at -25 °C, 0 °C, and 25 °C; whereas the second was measured at 25 °C, 50 °C and, 75 °C. All populations were measured on the same wafer.

The decorrelation observed in Figure 7.14 with increasing \( \Delta T \) indicates that, at the reference frequency (100 Hz), different traps are responsible for the LFN PSD at different temperatures. Two explanations are possible for the observed behaviour, however. The first, previously discussed, is that the characteristic time constant of the traps is temperature dependent; in this case, with increasing temperature traps become faster, causing the entire noise PSD to shift towards higher frequencies. The second is the case in which different traps are active altogether when temperature is modified, causing the LFN spectrum...
to change its shape entirely with temperature.

While the analysis in Figure 7.14 is not sufficient to clarify on the matter, a combined temperature and frequency correlation analysis, presented in the following section, provides an answer to this question.

### 7.4 Temperature and Frequency Correlation

In the previous section it was established that temperature shifts cause the de-correlation of the LFN PSD at a given frequency, even though it was not possible to determine whether the temperature affects the occupation probability of the traps (i.e. the $\beta$ parameter) or the characteristic time constant of the traps ($\tau$). If the first assumption is true — that is the $\beta$ parameters is responsible for the observed de-correlation with $\Delta T$ — then different traps are responsible for the noise PSD at different temperatures, meaning the correlation would drop at all frequencies with increasing temperature. On the other hand, if the second assumption is true — $\tau$ is thermally activated — then the same traps would be responsible for the noise PSD at a different temperature; however, their influence would be shifted in frequency.

While the findings of Dutta and Horn (1981) as well as the analysis of individual Lorentzians presented in Section 6.5 point towards a thermally activated $\tau$, these techniques are inadequate for a statistical analysis. The first — the Dutta and Horn identity — requires several temperature measurements (thus, is suitable for individual devices), while the second — the analysis of individual Lorentzians — requires identification of PSD profiles that are not present in large area devices. Conveniently, if the correlation technique presented in this work is applied to the measured data, combining both temperature and frequency correlations. ($R[S_{l_0}(f_{ref}, T_{ref}), S_{l_0}(f_{ref} + \Delta f, T_{ref} + \Delta T)]$), it is possible to statistically demonstrate that the time constants of the traps are thermally activated, for both small-area (in which Lorentzian profiles are visible) and large-area devices (in which the LFN PSD is approximately $1/f$).

Figures 7.15 and 7.16 display the extracted correlation coefficient as a function of frequency and temperature for a reference frequency of 1 kHz and a reference temperature of 25 °C for two NFET populations from the same wafer (symbols). The prediction given by the theoretical model, calculated numerically using the Kirton and Uren model, is also presented (solid line). At $\Delta T = 0$ °C the analysis is reduced to the frequency autocorrelation. Notice, however, that as $\Delta T$ increases the peak of the correlation coefficient
Figure 7.15: Correlation coefficient extracted from measured data (symbols) from two 80-DUT NFET, 40-nm technology, populations (A and B) from the same wafer. Each population was measured at three temperatures as shown in the legend. Device geometry is $W = 1 \mu m$ and $L = 0.04 \mu m$; $T_{ref} = 25$ °C and $f_{ref} = 1$ Hz; bias conditions were a) $V_{GS} = 1.1$ V, $V_{DS} = 0.1$ V and b) $V_{GS} = 0.7$ V, $V_{DS} = 0.5$ V. Solid lines indicate theoretical prediction.

is shifted in frequency. This provides strong indication that the "shape" of the PSD is unaltered with temperature, but shifted in frequency; in other words, it is an indication that the same traps are active at different temperatures (i.e. their occupation probability, $\beta$ is primarily unaffected with temperature), but their characteristic time, $\tau$, is strongly temperature dependent. This is in line with the concept that $\tau$ decreases with increasing $T$. This is a well-known observation for RTS, but is now also revealed in the context of $1/f$ noise.

Despite the fact that the frequency-temperature autocorrelation follows the now-familiar bell shape, it does not fully reach 1 when $\Delta T > 0$, which indicates that the stochastic process underlying the LFN contains at least two independent stochastic components as is, e.g., the case for the Kirton and Uren model. Figure 7.17 illustrates the difference between the correlation coefficient estimation using thermal activation and using Kirton and Uren model.

Due to the complexity of measuring the same device several times (e.g. degradation of the pads), each population was split into two groups. The first, named group A,
was measured at -25 °C, 0 °C and 25 °C; while the second, group B, was measured at 25 °C, 50 °C and 75 °C.

Figure 7.16: Correlation coefficient extracted from measured data (symbols) from two 78-DUT NFET, 40-nm technology, populations (A and B) from the same wafer. Each population was measured at three temperatures as shown in the legend. Device geometry is \( W = 1 \mu m \) and \( L = 1 \mu m \); \( T_{ref} = 25 °C \) and \( f_{ref} = 1 \) Hz; bias conditions were a) \( V_{GS} = 1.1 \) V, \( V_{DS} = 0.1 \) V and b) \( V_{GS} = 0.7 \) V, \( V_{DS} = 0.5 \) V. Solid lines indicate theoretical prediction.

Figure 7.17: Estimated correlation coefficient using thermal activation (dashed lines) and the Kirton and Uren model (solid lines).
7.5 Gate Voltage Autocorrelation

The proposed technique can also be applied to study the gate voltage dependence of the LFN PSD. From the RTS and bias temperature instability (BTI) framework, the capture and emission time constants are expected to be strongly gate bias dependent. Therefore, the $\beta$ parameter (the ratio between capture and emission time constants), which describes how much a given trap contributes to noise power (maximum contribution when $\tau_c = \tau_e$), should also be gate bias dependent. Simply put, given that $\beta = \tau_e / \tau_c = \exp[(E_T - E_F)/kT]$, if the trap energy, $E_T$, is close to the Fermi energy, $E_F$, then $\tau_c \approx \tau_e$, causing the trap to switch often, in which case the trap contributes significantly to noise power. Thus, since $E_F$ varies with gate bias, different traps are expected to contribute to noise power at different gate biases. In order to demonstrate that shifts in gate bias indeed have a profound effect on the contribution of a single trap to noise power, as predicted by the RTS model, the frequency and gate bias autocorrelations are analysed in tandem ($R[S_{I_D}(f_{ref}, V_{GS,ref}), S_{I_D}(f_{ref} + \Delta f, V_{GS,ref} + \Delta V_{GS})]$).

In figures 7.18 and 7.19, the correlation coefficient as a function of frequency and gate bias was extracted for two NFET and two PFET 60-DUT, 40-nm technology, populations, respectively. The significant drop in the correlation with increasing $\Delta V_{GS}$ demonstrates that there is a strong dependence of $\beta$ with $V_{GS}$. It is important to notice that, contrary to the temperature and frequency autocorrelation, there is no clear shift in the peak of the correlation coefficient, indicating the $\beta$ parameter was more significantly affected than $\tau$ (i.e. traps did not become overall faster or slower with increasing $\Delta V_{GS}$; their overall contribution to noise power was modified with increasing $\Delta V_{GS}$).

Figure 7.18: Extracted correlation coefficient as a function of frequency and gate bias for two 60-DUT NFET populations (40-nm technology). The reference gate bias is 0.9 V and drain bias is 50 mV.
Notice, however, the observed de-correlation with $\Delta V_{GS}$ may be attributed to $E_T$ or $E_F$; that is, not only may $E_F$ depend on gate voltage, but also $E_T$. Typically, if traps are assumed to be located at the oxide/semiconductor interface, then $E_T$ is constant, with only $E_F$ bias dependent. If, on the other hand, the trap is assumed to be at a certain depth within the oxide, then $E_T$ and $E_F$ are both bias dependent. The dependence of $E_T$ on the electric field across the oxide is given by

$$E_T = E_{T0} + qz_{tr}E_{ox}$$  \hspace{1cm} (7.17)

where $E_{T0}$ and $z_{tr}$ are the trap energy in flat band conditions and trap depth, respectively; and the electric field, $E_{ox}$, is given by

$$E_{ox} = \frac{1}{T_{ox}} (V_{GB} - V_{FB} - \phi_s)$$  \hspace{1cm} (7.18)

where $V_{GB}$, $V_{FB}$, $\phi_s$ and $T_{ox}$ are the gate to substrate voltage, flat band voltage, surface potential and gate oxide thickness, respectively. The impact of trap depth on the bias dependence of the trap energy is illustrated in Figure 7.20, in which the transition from weak to strong inversion ($\Delta V_{GS,1}$) and the application of an additional voltage in strong inversion ($\Delta V_{GS,2}$) are shown.

In the first case, notice that both $E_F - E_c$ and $E_{ox}$ increase with $\Delta V_{GB,1}$, causing the trap to shift from well above $E_F$ to slightly below $E_F$. In the second case, however, further increases in $V_{GB}$ barely affect $E_F$, primarily increasing $E_{ox}$. If the trap were located at the semiconductor/oxide interface, then $\Delta E_T$ would be equal to zero. Due to the trap depth however, $\Delta V_{GB,2}$ causes the trap to shift from slightly below $E_F$ to deeply below $E_F$. 

Figure 7.19: Extracted correlation coefficient as a function of frequency and gate bias for two 60-DUT PFET populations (40-nm technology). The reference gate bias is 0.9 V and drain bias is 50 mV.
Figure 7.20: MOS capacitor energy band diagram for three different gate bias conditions ($V_{blue} < V_{red} < V_{green}$).

Consequently, if traps are located inside the oxide and not at the semiconductor/oxide interface, increasing $\Delta V_{GS}$ may cause significant de-correlation even if $E_F$ remains roughly constant.

Figure 7.21 displays the estimated correlation coefficient from a generic device ($V_T \approx 0.65V$), using the Monte Carlo simulation tool, assuming traps are located at the interface and 0.5 nm inside the gate oxide. The comparison between measured and simulated results indicate that, on average, traps are not located at the interface, but rather inside the oxide, so that $E_T$ is also bias dependent. One should be aware, however, that...

Figure 7.21: Estimated correlation coefficient versus frequency for different gate voltages applied using Monte Carlo simulation. Solid black line is the frequency correlation at the reference gate voltage (0.9 V). Solid lines (blue and red) are the estimated correlation coefficient assuming traps are located at the interface for 0.7 V and 1.1 V, respectively; dashed lines (blue and red) are assuming traps are located approximately 0.5 nm inside the gate oxide for 0.7 and 1.1 V, respectively.
the estimation of the Fermi energy through basic semiconductor equations may not be adequate for sub-micrometer devices. Therefore, in order to properly estimate the average trap depth using the proposed technique, additional analyses, supported by TCAD simulations, are required for future studies.

7.6 Drain and Source Swap

Another interesting analysis is presented in Figure 7.22, in which the LFN PSD of a population of thick gate oxide PFETs is measured twice, swapping the drain and source terminals. Due to the symmetry of the MOSFET, the average value and the variance of the noise are similar regardless of which terminal is taken as drain or source. The correlation coefficient, however, is strongly dependent on the region of operation of the device. This is a consequence of the trap impact dependence on the position along the channel. Following number fluctuation solely,

$$
E \left[ \Delta I_D(x_{tr}) \right] = \frac{1}{N(x_{tr}) + \left( k_B T / q^2 \right) (C_{ox} + C_D + C_{it})},
$$

where $N(x_t)$ is the surface carrier density at position $x_{tr}$.

Figure 7.22: Correlation coefficient of the LFN PSD at 20 Hz as a function of gate voltage for a 68-DUT PFET population measured swapping the drain and source terminals (solid blue dots). Black dashed line obtained through Monte Carlo simulation of 5,000 devices. The threshold voltage of the devices is approximately 0.5 V.

In weak inversion, the channel is not yet formed; due to the low surface carrier density, the trap impact is dominated by the oxide capacitance, which is position independent. Therefore, correlation coefficient is high, given that the impact of a trap is independent of its position along the channel.

As gate voltage increases, the channel is formed; however, given that $V_{DS} > V_{GS}$ —
$V_T$ the device operates in saturation region. In this region, the surface carrier concentration varies along the channel, thus the trap impact becomes position dependent. Traps near the pinch-off region become more prominent due to the lower surface carrier density, an effect which is enhanced by the pocket (halo) implants, availing to the observed drop in correlation.

Further increasing the gate voltage causes the device to operate in linear region ($V_{DS} < V_{GS} - V_T$), causing the surface carrier concentration to become approximately constant along the channel for sufficiently high gate voltages. The trap impact becomes again position independent, and, therefore the correlation coefficient increases.

For the Monte Carlo simulation presented in Figure 7.22, the surface carrier density, required in (7.19), was obtained through numerical solution of the 1-d Poisson equation of the MOSFET. The trap impact as a function of position for three different gate voltages are presented in Figure 7.23. As expected, in subthreshold and linear region, $V_{GS} = 0.45$ V and $V_{GS} = 2.4$ V, respectively, $E\left[\Delta I_D(x_{tr})\right]$ is approximately constant, and can be approximated by $q^2/(k_BT\varepsilon_{ox})$ and $1/N$, respectively. On the other hand, when $V_{GS} - V_T \approx V_{DS}$, $E\left[\Delta I_D(x_{tr})\right]$ is strongly position dependent, since the surface carrier concentration, $N$, is also position dependent. It is important to observe that the purpose of the Monte Carlo simulation is to provide the qualitative behaviour of the correlation coefficient across multiple gate voltages for a generic device population. This means that the simulation was not parametrized to represent the measured 65-nm technology (with the exception of the threshold voltage).

Figure 7.23: Average trap impact as a function of trap position for three different gate voltages.

The trap impact as a function of channel position for different device geometries and biases was more deeply explored through TCAD simulation by Banaszeski da Silva et al. (2016), accounting for non-uniform channel doping.
7.7 Summary and Discussions

The correlation coefficient methodology developed and presented in this chapter provided strong indication that the low-frequency noise in the analysed CMOS technologies is primarily composed by a superposition of Lorentzians, reinforcing trapping/detrapping noise models. No evidence was found to support bulk mobility fluctuations in these devices. Moreover, it was demonstrated that the characteristic time constant of these Lorentzians is temperature dependent, in line with a thermally activated process, as proposed by Dutta and Horn (1981), Surya and Hsiang (1988), and Kirton and Uren (1989).

Additionally, the results indicate that the gate bias primarily affects the occupation probability of the traps — i.e. the ratio between capture and emission time constants —, but not necessarily their characteristic time constants. Also, the observed de-correlation when gate bias is varied in strong inversion indicates that traps are located within the oxide rather than at the oxide/semiconductor interface.

Finally, by swapping the source and drain terminals, it was demonstrated that trap impact is strongly position dependent when the device is operated in saturation, which must be accounted for in LFN models. This is particularly important for modern CMOS technologies due to channel engineering. Pocket implants, for instance, may enhance trap impact, causing both the expected value and the variance of the LFN to significantly increase, as demonstrated by Banaszeski da Silva et al. (2016).
8 FINAL REMARKS

In this work, a new technique for low-frequency noise analysis has been proposed, in which autocorrelations between LFN spectra at different frequencies, biases, and temperatures are investigated. The strength of the method lies in the fact that different physical models yield significantly different predictions for these autocorrelations. Additionally, being a statistical method, it captures the dominant behaviour of the entire population, despite the subtleties of individual devices.

An analytical model for the correlation coefficient as a function of frequency was derived from the RTS-based framework and compared to extracted correlation coefficient from measured data from three different CMOS mixed-signal technologies. The comparison reinforced that the $1/f$ noise in MOSFETs is composed by a summation of Lorentzians.

The temperature dependence of the LFN was also evaluated using the autocorrelation technique; the extracted correlation coefficient as a function of frequency and temperature was compared to a RTS-based numerical model. The results provided a strong case for predominantly thermally activated time constants for the trapping process, in accordance with the temperature dependence observed by Surya and Hsiang (1988) and Kirton and Uren (1989), which should be accounted for in compact LFN models.

A similar analysis was conducted in terms of the gate bias, which indicated that, while gate bias has no noticeable impact on the characteristic time constant, it significantly affects trap occupancy, i.e. which traps contribute more significantly to noise at a given condition. Monte Carlo simulations were employed to evaluate the impact of trap depth within the oxide on the correlation coefficient, suggesting that traps might not be located at the interface.

Also, by swapping the drain and source contacts it was possible to observe that trap impact is significantly affected by trap position. Therefore, for accurate LFN modelling it is imperative to account for the non-uniform channel doping of modern MOSFET devices.

Overall, the measured data from three different commercial CMOS mixed-signal technologies provide strong support for an LFN model based on the sum of thermally activated random telegraph signals for all measured N-and PFETs, regardless of device area. No evidence supporting bulk mobility fluctuations was found.

The methodology presented in this work can be applied to study LFN in advanced CMOS technologies (such as FinFETs) and in other device structures (such as bipolar
junction transistors). These analysis avail the understanding and quantification of the physical mechanisms underlying LFN in new devices and materials.
REFERENCES


PUBLISHED WORK


