UCL
Université catholique de Louvain

École polytechnique de Louvain (EPL)

AC Compact Modelling of Low-Frequency Noise for $28\mathrm{nm}$ CMOS Technology and Beyond

Mémoire présenté par
Léopold Van Brandt

en vue de l'obtention du grade de Master
ingénieur civil électricien

Promoteur(s)
Denis Flandre, Jean-Pierre Raskin

Lecteur(s)
Valeria Kilchytska, Bertrand Parvais

Année académique 2016-2017

Acknowledgements

Ces mots seront les seuls écrits en français de tout le manuscrit.

Ma plus grande gratitude revient au Prof. Denis Flandre qui, en plus d'avoir éveillé mon intérêt pour l'électronique et la physique des dispositifs dès la troisième année de bachelier, a été d'excellent conseil tout au long de ma formation et bien entendu également comme promoteur principal de ce mémoire. Bien que j'aie côtoyé de nombreux grands chercheurs, tant à l'UCL qu'à imec, ce sont incontestablement les pistes et apports de Denis qui ont laissé le plus de traces sur ce travail. Son expérience, sa connaissance, mais aussi son pragmatisme vis-à-vis des applications circuits sont énormes.

J'ai réellement rencontré il y a un peu plus d'un an dans le cadre du cours "Transistors Avancés", et je suis très heureux qu'il ait fait partie de l'équipe encadrante de ce mémoire. Si l'orientation du sujet s'est vraisemblablement éloignée de ses domaines de prédilections, ses conseils, questions et remarques ponctuels ne m'auront pas moins aidé à aller dans la bonne direction.

Un tout grand merci également à Valeria avec qui je suis régulièrement en contact depuis un an. Sa connaissance de la technologie CMOS et sa maîtrise des outils de caractérisation expérimentale ont fait d'elle une précieuse alliée, tant dans le cadre de ce mémoire que dans tous mes travaux touchant à l'électronique.

I wish to thank all the great guys that I meet at imec and who has supported this work as much as they could. Many thanks to Bertrand for all his inputs and the bridge between UCL and imec that he build for me. I pay particular attention to Marko for all his patience and answers to my questions regarding the Smart array. Thanks also to Pieter for his very accurate answers during the meetings. I will not forget neither Ben, who originally launched this work by throwing the idea of getting interested in power spectral density, after some discussion and evaluation of my expertise.

Great thank to Kashif from UCL for the flicker noise simulation and some interesting discussion.

Je remercie également tous mes proches dont la plupart sont assez éloignés de mes activités. Merci à ma famille pour tous les encouragements tacites et explicites, et pour la confiance que vous m'avez toujours accordée.

Je terminerai par un merci à Alice. Merci de m'avoir soutenu et d'avoir patiente en cette période chargée par un travail chronophage mais ô combien passionnant.

Abstract

Electrical noise is defined as a randomly time-fluctuating electrical signal. In large-area MOSFETs, the stochastic component of the drain current is dominated by the flicker ($1/f$) noise in low frequency, and by the white Gaussian thermal noise in high frequency. These two noise sources are nowadays well incorporated into circuit simulators (i.e. SPICE), which allows analog designers to perform noise analysis both in transient and AC. In 28nm technology and beyond, a granular non-Gaussian low-frequency noise arises from the presence of a few traps within the MOSFET gate oxide: the *random telegraph noise* (RTN). Its amplitude is found to increase as the inverse of the gate area. Study and modelling of this noise get thus substantially motivated by the use of advanced CMOS technologies in VLSI design.

This work aims to derive an AC compact modelling of the RTN, and to discuss the cointegration of such a model with existing models for $1/f$ noise. The analytical expression of the RTN power spectral density is exactly known from stochastic process considerations. One goal of this work is to validate this expression trough spectral analysis based on experimental measurement in the time domain. The use of the analytical expression is only possible if accurate estimations of the RTN parameters are available. Thus, we propose a comprehensive methodology for both time-domain and spectral analysis of the RTN. A robust software tool has been implemented in Matlab®. Recommendations for a suitable choice of the measurement parameters such as the sampling frequency and the time window are made, in a large-scale characterization perspective.

Chapter 1

Introduction

1.1 Context

The challenges related to CMOS scaling go beyond electrostatic consideration, for now more than one decade. Variability of as-fabricated (i.e. time-zero) parameters of modern VLSI devices has been considered in circuit design tools for some time [1]. This time-zero variability is also known as statistical variability [2–5] and corresponds to what analog designers call Pelgrom's mismatch [6]. The impact of static variability is now well captured by precise numerical simulation of the MOSFET [2, 3]. Underlying compact or SPICE models allow the circuit designer to deal with worst case corners, that he takes into account to ensure enough margin for a good behaviour of the circuit.

Deeply-scaled CMOS devices suffer from an additional, time-dependant variability. It is due to the presence of a few stochastically-behaving defects, or traps, in the FET gate oxide [1, 8]. The gate area $A_G = W \cdot L$ can be so small that only *one* trap is active. The noise associated with capture/emission of carrier from/to the channel by the *single* defect (or trap) [9–15] is

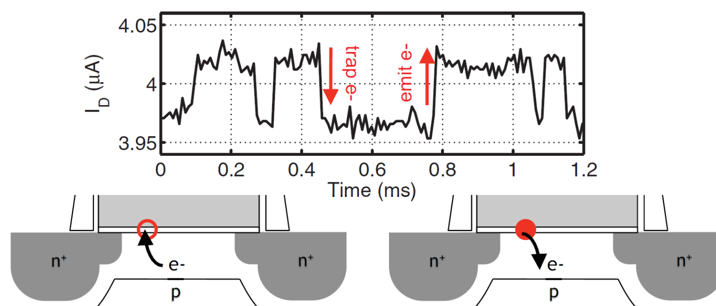


Figure 1.1: *Two-level* RTN trace (top), and illustration of carrier trapping process (bottom). Reproduced from [7].

called *random telegraph* (or *burst* or *popcorn*) *noise* (RTN). A typical RTN realization in the time domain is given in Figure 1.1. We typically observe two well defined current levels, depending whether the trap is charged or not. We note the presence of some superimposed "background" noise, that could be $1/f$ noise or thermal noise [16, 17].

As soon as inherent noise of MOSFET becomes an issue in the design of CMOS Integrated Circuits (ICs) by limiting the margins [18], accurate physics-based models are necessary [19,20]. Current fluctuations due to low-frequency noise (LFN) was found to be more significant than the Random Dopant Fluctuation (RDF) for MOSFETs scaled beyond 22nm [21, 22]. Thermal noise and flicker 1/f noise are well embedded in SPICE simulator. Equations describing their behaviour both in transient (.NOISETRAN) or AC (.NOISE) are available in industrial tools. Random telegraph noise which is essentially a granular version of 1/f noise [9, 12, 16, 23], is an additional Low-Frequency Noise (LFN) source that emerges strongly as quickly as the gate area is scaled down. Compact modelling of such a stochastic process in transient and AC is motivated by the design of analog integrated circuits, whose cointegration with digital blocks in SoC generally implies the use of deeply scaled technology.

1.2 Contribution

This work aims to propose a comprehensive methodology to process a large amount of RTN measurement in a AC compact modelling perspective. The power spectral density of the RTN is a *Lorentzian*:

$$\boxed{S_{\mathcal{L}}(f) = (\Delta I^*)^2 \, \frac{4\bar{\tau}^2}{\bar{\tau}_e + \bar{\tau}_c} \, \frac{1}{1 + (2\pi f \bar{\tau})^2}} \tag{1.1}$$

Three parameters are involved:

- the mean time constants $\bar{\tau}_e$ and $\bar{\tau}_c$;
- the trap electrical impact ΔI^*.

Hence, accurate estimations of these parameters is requires.

The state-of-the-art measurement technique to extract the time constants is the TDDS (Time-Dependent Defect Spectroscopy), which relies on multiple stress and relaxation phases [1, 11, 24]. This technique require significant data post-processing, and present to disadvantage of being far from circuit consideration.

Rather, this work propose a complete analysis of simple traces measured in the time-domain. A stochastic formalism is developed in Chapter 1 and used to describe the noise. The major part of this work is about power spectral density and its estimation given some samples, as outlined in Chapter 3.

This work can be qualified of methodologic in the sense that our target is to suggest an optimized set of measurement parameters dedicated to the RTN characterization. This is the purpose of Chapter 4. Finally, the largest part is Chapter 5 that treats in details a RTN study case from imec Smart array.

Chapter 2

Stochastic study of the RTN

Electrical noise can be defined as a randomly time-fluctuation electrical signal. The total MOSFET drain current $I_D(t)$ is the sum of both the deterministic DC component $I_{D,ref}$, resulting from the applied bias (V_{GS}, V_{DS}), and the randomly time-fluctuating quantity $\Delta I_D(t)$, which also depends on the bias. Mathematically, we have

$$I_D(V_{GS}, V_{DS}, t) = I_{D,ref}(V_{GS}, V_{DS}) + \Delta I_D(V_{GS}, V_{DS}, t).$$

or, simply

$$I_D(t) = I_{D,ref} + \Delta I_D(t).$$

In the following, the drain current fluctuation $\Delta I_D(t)$, or the corresponding threshold voltage fluctuation $\Delta V_{th}(t)$ is generally denoted by $X(t)$.

The theory of stochastic processes provides a framework for the study of noise, e.g. the RTN. Some theoretical background is given in the first section of this chapter. Emphasis is placed on the properties of the different noise sources in MOSFETs. Afterwards, a section is dedicated to the stochastic modelling of the RTN. The distribution function of the time constants is derived. Finally, RTN is then reinterpreted as a Markov chain, which allows a software implementation of a RTN generator. Finally, the full calculation of the correlation function of the RTN is given.

2.1 Stochastic process

We say that a process is *stochastic* if a knowledge of its values up to and including t allows us to *probabilistically* predict its value at any infinitesimally later time $t + \mathrm{d}t$ [25]. More precisely, the values $X(t')$ for $t' \leq t$ determine the *probability* that $X(t + \mathrm{d}t)$ will equal any particular value $x(t + \mathrm{d}t)$ for any given positive infinitesimal $\mathrm{d}t$.

A stochastic process $X(t)$ is a rule for assigning to every outcome of an experiment a time function $x(t)$ [26]. Thus a stochastic process is a family

(or *ensemble*) of time functions

$$\{x_1(t), x_2(t), \dots\}.$$

$x_1(t), x_2(t), \dots$ are fully deterministic single time functions and are called *realizations* or *samples* of the stochastic process $X(t)$.

2.1.1 Statistics of stochastic processes

The stochastic process $X(t)$ can be manifested in the form of an infinite number of realizations in the time domain. The tools to predict how a typical realization $x(t)$ looks like, or what several realizations share in common, are the *statistics* of the stochastic process.

This text is only concerned about *real* stochastic process.

First-order statistical properties

If the time t is fixed, then $X(t)$ is a *random variable*. Its *first-order* properties are the following:

- The *first-order probability density function* (pdf) is denoted by $f_{X(t)}(x(t))$. Thereafter, we will use the compact notation $f_X(x(t))$. This fundamental deterministic function can be interpreted as follows. For any fixed t, the product

  $$f_X(x(t))\,\mathrm{d}x$$

  represents the probability of $X(t)$ falling within a $\mathrm{d}x$-range of values around $x(t)$.

- The *mean* $\eta_X(t)$ of $X(t)$ is, for any fixed t, the expected value of the random variable $X(t)$:

  $$\begin{aligned}\eta_X(t) &\equiv \mathbb{E}\{X(t)\} \\ &= \int_{-\infty}^{+\infty} x\, f_X(x(t))\,\mathrm{d}x.\end{aligned} \tag{2.1}$$

- The *variance* is

  $$\begin{aligned}\sigma_X^2(t) &\equiv \mathbb{E}\left\{\left|X(t)-\eta_X(t)\right|^2\right\} \\ &= \int_{-\infty}^{+\infty} (x-\eta_X(t))^2\, f_X(x(t))\,\mathrm{d}x.\end{aligned} \tag{2.2}$$

Examples

- Thermal noise and flicker noise are usually assumed to be Gaussian [1]. This property characterize their pdf:

$$f_X(x(t)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right). \tag{2.3}$$

  $\eta_X = \mu = 0$ for a zero-mean noise.

- The sample space of a pure RTN contains only two values, the high and the low levels x_0 and x_1. Thus, its pdf is of the form

$$f_X(x(t)) = (1-d) \cdot \delta(x - x_0) + d \cdot \delta(x - x_1) \tag{2.4}$$

  where d is the average *duty cycle* of the RTN.

These two examples allow to appreciate the non-Gaussianity of the RTN which is a very different from the other well known noise sources in MOSFETs.

Stationarity We observe that the pdf's (2.3) and (2.4) does not depend on t:

$$f_X(x(t)) = f_X(x). \tag{2.5}$$

As a consequence, both the mean and the variance are also independent of t:

$$\begin{aligned} \eta_X(t) &= \eta_X = \text{cst} \\ \sigma_X^2(t) &= \sigma_X^2 = \text{cst}. \end{aligned} \tag{2.6}$$

These properties partially define what is known as *wide-sense stationarity*; the definition will be completed below once the second-order statistics have been introduced.

Second-order statistical properties

Unlike one single random variable X, the stochastic process $X(t)$ is defined either on a continuous range of t or at discrete but multiple times t_n. We way then wonder how values taken by the realizations at two different times t_1 and t_2 are related. The rigorous concept is the *correlation*.

- The *second-order probability density function* is denoted by $f_X(x(t_1), x(t_2))$. It is interpreted as follows. For any fixed t_1 and t_2, the product

$$f_X(x(t_1), x(t_2))\, \mathrm{d}x_1 \mathrm{d}x_2$$

[1]Theoretical support for this assumption is the *central limit theorem*.

represents the probability

$$P\left(\begin{cases} x\left(t_1\right)-\dfrac{\mathrm{d}x_1}{2} \leq X\left(t_1\right) \leq x\left(t_1\right)+\dfrac{\mathrm{d}x_1}{2} \\ x\left(t_2\right)-\dfrac{\mathrm{d}x_2}{2} \leq X\left(t_2\right) \leq x\left(t_2\right)+\dfrac{\mathrm{d}x_2}{2} \end{cases}\right)$$

- The *correlation function* $R_X\left(t_1,t_2\right)$ of $X\left(t\right)$ is the expected value of the product $X\left(t_1\right)X\left(t_2\right)$:

$$\begin{aligned} R_X\left(t_1,t_2\right) &\equiv \mathbb{E}\left\{X\left(t_1\right)X\left(t_2\right)\right\} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1x_2\, f_X\left(x\left(t_1\right),x\left(t_2\right)\right)\,\mathrm{d}x_1\mathrm{d}x_2. \end{aligned} \tag{2.7}$$

Stationarity A stochastic process is called *wide-sense stationary* (WSS) when the mean is constant and the correlation function depends only on $\tau \equiv t_1 - t_2$:

$$\begin{aligned} &\eta_X\left(t\right) = \eta_X = \text{cst} \\ &R_X\left(t_1,t_2\right) = R_X\left(\tau = t_1-t_2\right) = \mathbb{E}\left\{X\left(t+\tau\right)X\left(t\right)\right\}. \end{aligned} \tag{2.8}$$

Definition (2.6) is a particular case of (2.8). We note that if $X\left(t\right)$ is a WSS stochastic process with zero-mean, we have

$$\sigma_X^2\left(t\right) = \sigma_X^2 = \mathbb{E}\left\{\left|X\left(t\right)\right|^2\right\} = R_X\left(t,t\right) = R_X\left(0\right). \tag{2.9}$$

Variance is a quantity of utmost importance in electrical engineering and will be further discussed and interpreted in term of *power* in the next chapter.

Examples

- Thermal noise is *white noise.* The amplitude of the process at some instant t_1 is fully *independent* and *decorrelated* from the value at a different time t_2, even if t_1 and t_2 are very closed (but distinct):

$$R_W\left(\tau = t_1-t_2\right) = \sigma_W^2\delta\left(t_1-t_2\right) = \sigma_W^2\delta\left(\tau\right). \tag{2.10}$$

- $1/f$ noise and RTN are *coloured noise*, their samples at different times are correlated. This will be interpreted when relating the power spectral density to the correlation function in the next chapter.

  The correlation function of a RTN process will be computed further once the stochastic model has been established.

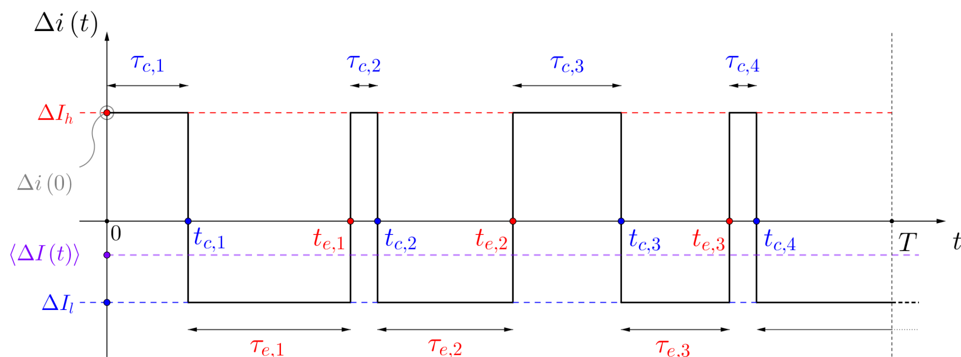


Figure 2.1: Typical RTN realization, with the useful definitions for the stochastic model.

2.2 Stochastic model for the random telegraph noise

The random telegraph noise physically originates from the carrier trapping process. Given its importance, we study in detail the capture/emission phenomenon. The experimental traces, such as the one of Figure 1.1, suggest a *two-state model*. Afterwards, the parameters are discussed from a physical point of view, using the simplest existing theories.

2.2.1 Poisson point process

A typical realization $\Delta i\,(t)$ of a pure two-level RTN could be the one depicted in Figure 2.1. The simple fact that the transitions from one state to another occurs only from time to time leads with very few light assumptions, almost miraculously, to the distribution of the time constants and the definition of the transition rates.

Description of the model parameters

The conventional (and the most simple) defect models assume that the defect (or trap) can exist in two states [11]:

- one *charged* (or *occupied* or *filled*).

  Some time before, an electron from the channel has been *captured* (or *trapped*) by the trap, and the drain current is low.

- The other *neutral* (that is *uncharged*, or *unoccupied* or *empty*).

  If the trap previously contained one electron, this electron has been *emitted* (or *released*) to the channel, and the drain current is high.

A *telegraph signal*, limited to a finite time period T, is *a priori* characterized by several parameters:

- the high- and low-current levels, respectively ΔI_h and ΔI_l. It is important to remember that these two current levels are *deterministic* parameters, for a given trap and a fixed bias.

- Its initial state, characterized by a random variable $\Delta I(0)$, taking ΔI_h and ΔI_l values with equal probability 1/2. For the realization of Figure 2.1, we find $\Delta i(0) = \Delta I_h$.

- The *capture* and *emission* events (or points), respectively denoted by $t_{c,1}, t_{c,2}, \ldots$ and $t_{e,1}, t_{e,2}, \ldots$ on the time axis. These are realizations of random points $T_{c,1}, T_{c,2}, \ldots$ and $T_{e,1}, T_{e,2}, \ldots$. Indeed, every new RTS realization will have its state transitions located at different instants. The numbers of respective events are contained is two random variables N_c and N_e that differ at most by one unit. These point on the time axis are called *random Poisson points* [26].

We are allowed to define random variables associated with the different up and down times, that is the capture and emission times

$$\begin{aligned}
\mathcal{T}_{c,1} &= T_{c,1} & \mathcal{T}_{e,1} &= T_{e,1} - T_{c,1} \\
\mathcal{T}_{c,2} &= T_{c,2} - T_{e,1} & \mathcal{T}_{e,2} &= T_{e,2} - T_{c,2} \\
\mathcal{T}_{c,3} &= T_{c,3} - T_{e,2} & \mathcal{T}_{e,3} &= T_{e,3} - T_{c,3} \\
&\ldots
\end{aligned}$$

whose realizations are $\tau_{c,1}, \tau_{c,2}, \ldots$ and $\tau_{e,1}, \tau_{e,2}, \ldots$ (also shown in Figure 2.1). We consider these as the realizations of only two random variables, $\mathcal{T}_c$ and $\mathcal{T}_e$. For each RTS realization, we get a particular number n_c of outcomes for $\mathcal{T}_c$, corresponding to $\tau_{c,1}, \tau_{c,2}, \ldots, \tau_{c,n_c}$

Time constant distributions

We are especially interested in knowing the distribution of $\mathcal{T}_c$, and similarly for $\mathcal{T}_e$. The capture and emission events are actually independent, in the sense that they occur alternatively and at their own rate. We may therefore virtually split the t-axis into two independent axis, t_c and t_e.

The study of the $\mathcal{T}_c$-distribution starts as follow. The N_c capture points are randomly placed in the interval $[0, T_c]$, where T_c is the total up time, that is

$$T_c = \mathcal{T}_{c,1} + \mathcal{T}_{c,2} + \ldots + \mathcal{T}_{c,N_c} \tag{2.11}$$

which is *a priori* unknown. Note that we obviously have $T_c + T_e = T$. The *average density of capture points* $\bar{\lambda}_c$ is computed as

$$\bar{\lambda}_c \equiv \mathbb{E}\{\Lambda_c\} = \mathbb{E}\left\{\frac{N_c}{T_c}\right\} = \mathbb{E}\left\{\frac{N_c}{\mathcal{T}_{c,1} + \mathcal{T}_{c,2} + \ldots + \mathcal{T}_{c,N_c}}\right\}. \tag{2.12}$$

We wish to evaluate the probability that k capture points lie in an interval of length Δt_c, chosen such that $\Delta t_c \ll T_c$. We make the assumption that T is large enough so that $N_c \gg 1$. We may thus invoke *Poisson theorem* [26] to evaluate the desired probability

$$P\left(k \text{ points in } \Delta t_c\right) = \exp\left(-\bar{\lambda}_c\,\Delta t_c\right)\frac{\left(\bar{\lambda}_c\,\Delta t_c\right)^k}{k\,!}. \tag{2.13}$$

In particular, the probability that one single point lies in Δt_c is

$$\begin{aligned} P\left(\text{one point in } \Delta t_c\right) &= \exp\left(-\bar{\lambda}_c\,\Delta t_c\right)\left(\bar{\lambda}_c\,\Delta t_c\right) \\ &\approx \bar{\lambda}_c\,\Delta t_c \end{aligned} \tag{2.14}$$

as Δt_c is small and $\bar{\lambda}_c$ is "moderate". Finally, the probability that there is no point in Δt_c is

$$P\left(\text{no point in } \Delta t_c\right) = \exp\left(-\bar{\lambda}_c\,\Delta t_c\right) \tag{2.15}$$

The *cumulative distribution function* of $\mathcal{T}_c$ is now easily computed:

$$\begin{aligned} F_{\mathcal{T}_c}\left(\tau_c\right) &\equiv P\left(\mathcal{T}_c \leq \tau_c\right) \\ &= P\left\{\text{at least one point in } \Delta t_c = \tau_c\right\} \\ &= 1 - P\left\{\text{no point in } \Delta t_c = \tau_c\right\} \\ &= 1 - \exp\left(-\bar{\lambda}_c\,\tau_c\right) \end{aligned}. \tag{2.16}$$

Its derivative gives the *probability density function* of $\mathcal{T}_c$

$$\boxed{\begin{aligned} f_{\mathcal{T}_c}\left(\tau_c\right) \equiv \frac{\mathrm{d}F_{\mathcal{T}_c}}{\mathrm{d}\tau_c}\left(\tau_c\right) &= \bar{\lambda}_c \exp\left(-\bar{\lambda}_c\,\tau_c\right) \\ &\equiv \frac{1}{\bar{\tau}_c}\exp\left(-\frac{\tau_c}{\bar{\tau}_c}\right) \end{aligned}} \tag{2.17}$$

where we defined

$$\boxed{\bar{\tau}_c \triangleq 1/\bar{\lambda}_c.} \tag{2.18}$$

An exactly similar development can be done for the emission time to obtain

$$\boxed{\begin{aligned} f_{\mathcal{T}_e}\left(\tau_e\right) &= \bar{\lambda}_e \exp\left(-\bar{\lambda}_e\,\tau_e\right) \\ &\equiv \frac{1}{\bar{\tau}_e}\exp\left(-\frac{\tau_e}{\bar{\tau}_e}\right) \end{aligned}} \tag{2.19}$$

with

$$\boxed{\bar{\tau}_e \triangleq 1/\bar{\lambda}_e.} \tag{2.20}$$

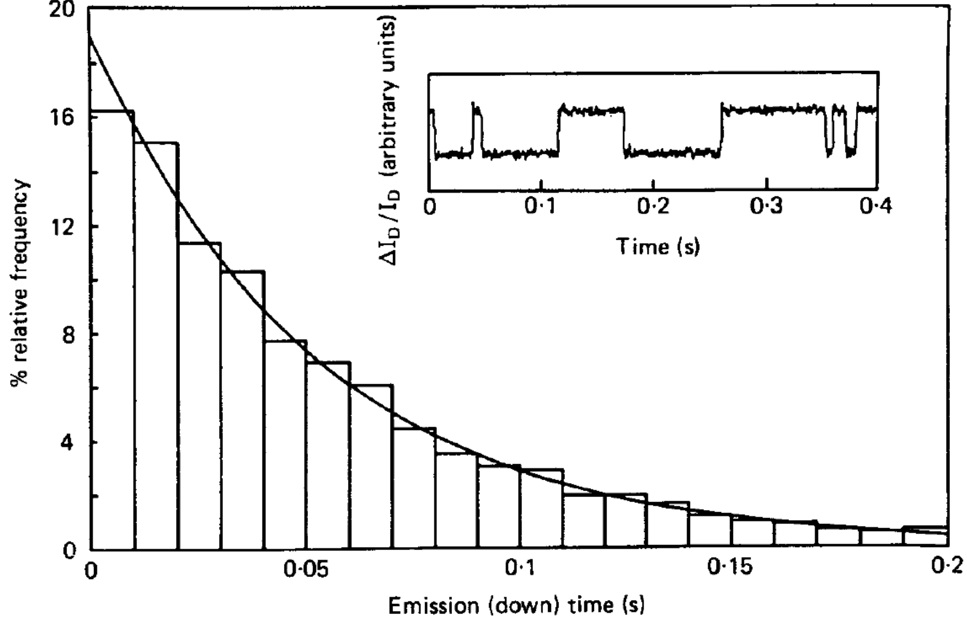


Figure 2.2: Distribution of 4425 emission times, showing that $\mathcal{T}_e$ is distributed exponentially, with mean $\bar{\tau}_e = 0.0528 \approx 0.05$s. The inset shows a portion of the corresponding RTN trace $\Delta I_D(t)/I_{D,ref}$. Reproduced from [10].

$\bar{\lambda}_c$ and $\bar{\lambda}_e$ are called *capture* and *emission rate*, respectively. $\bar{\tau}_c$ and $\bar{\tau}_e$ are the *mean time constants.* The average duty cycle, already introduced in (2.4) simply reads

$$d = \frac{\bar{\tau}_c}{\bar{\tau}_c + \bar{\tau}_e}. \tag{2.21}$$

We conclude that *the capture and emission times are exponentially distributed with means $\bar{\tau}_c$ and $\bar{\tau}_e$.* Figure 2.2 illustrates the excellent agreement with an experimental measurement of the emission times. This agreement was expected, as our single assumption to derive (2.17) was to consider a two-state model for the defect. All the literature admits such a distribution of $\bar{\tau}_c$ and $\bar{\tau}_e$ [1, 10–12, 15].

Three-parameter model

Furthermore, we are now convinced that the *the RTN process is fully determined by a set of four parameters* $\{\Delta I_h, \Delta I_l, \bar{\tau}_c, \bar{\tau}_e\}$, as the capture and emission events can be drawn from distributions (2.17) and (2.19). The *mean* of the RTN (also represented in Figure 2.1) is calculated as

$$\langle \Delta I(t) \rangle = \frac{\bar{\tau}_c \cdot \Delta I_h + \bar{\tau}_e \cdot \Delta I_l}{\bar{\tau}_c + \bar{\tau}_e}. \tag{2.22}$$

The literature [1,10–12,15] always characterizes the RTN with the reduced set of three parameters

$$\{\Delta I_D^*, \bar{\tau_c}, \bar{\tau_e}\} \text{ with } \Delta I_D^* \equiv \Delta I_h - \Delta I_l,$$

or, equivalently, by converting a shift in the drain current to shift in the threshold voltage V_{th},

$$\{\Delta V_{th}^*, \bar{\tau_c}, \bar{\tau_e}\}.$$

ΔI_D^* and ΔV_{th}^* are called the *defect impact on the MOSFET* [1,11]. There is obviously an implicit assumption that allows us to move from a four-parameter representation to a three-parameter set. Our hypothesis is that all the groups treat the RTN as a *zero-mean* signal:

$$\langle \Delta I(t) \rangle = \frac{\bar{\tau_c} \cdot \Delta I_h + \bar{\tau_e} \cdot \Delta I_l}{\bar{\tau_c} + \bar{\tau_e}} = 0, \tag{2.23}$$

potentially by shifting the two current levels ΔI_h and ΔI_l from their physical values, so that (2.23) is verified. The physical mean of the RTN is probably included in *time-zero* (or *static*) *variability* [11]. Equation (2.23) then clearly shows that one of the four parameters depends on three others, yielding a minimal set of only three parameters. This is the *defect-centric* approach [1, 12], each trap is fully characterized by its mean capture and emission times $\bar{\tau_c}$ and $\bar{\tau_e}e$, and its electrical impact ΔI^* (or ΔV_{th}^*).

2.2.2 Markov process

Definitions

A process is said *memoryless* of the value of $X(t)$ *alone* uniquely determines $X(t + \mathrm{d}t)$ [25]. The value $X(t + \mathrm{d}t)$ is obtained from the value $X(t)$ A *memoryless stochastic process* is called a *Markov process* [25,27]. The value of $X(t)$ *alone* uniquely determines the probabilities assigned to the possible values of $X(t + \mathrm{d}t)$, so the process can advance in time without having to recall its past values, which are forgotten [25]. Mathematically,

$$f_X\Big(x(t + \mathrm{d}t)\,|x(t_0), x(t_1), \ldots, x(t_n), x(t)\Big) = f_X\Big(x(t + \mathrm{d}t)\,|x(t)\Big) \tag{2.24}$$

$\forall\, t_0 < t_1 < \ldots < t_n < t < t + \mathrm{d}t$, where we use the compact notation $f_X\Big(x(t + \mathrm{d}t)\,|x(t)\Big)$ for the *conditional probability density function*

$$f_{X(t+\mathrm{d}t)|X(t)}\Big(x(t + \mathrm{d}t)\,|x(t)\Big).$$

Thus, a Markov process $X(t)$ is the state function of some system whose state value at time $t + \mathrm{d}t$ can be probabilistically predicted from its state value at time t, but in a way that cannot be enhanced by a knowledge of any values of the process prior to time t [25].

RTN software generator

The RTN is a Markov process. The next capture or emission events depends only on the current charge state of the trap, given the transition rates. For the sake of convenience, we set $X(t) \equiv \Delta I(t)$ and we assume $\Delta I_l = 0$ and $\Delta I_h = \Delta I^*$. Kinetic equations of the process are directly derived from the stochastic defect modelling of previous section. Using (2.14), we obtain

$$P\left(X(t + \mathrm{d}t) = \Delta I^* \,|X(t) = 0\right) = \lambda_e \cdot \mathrm{d}t \tag{2.25}$$

and

$$P\left(X(t + \mathrm{d}t) = 0 \,|X(t) = \Delta I^*\right) = \lambda_c \cdot \mathrm{d}t \tag{2.26}$$

```
% Ns = number of samples
% dt = time step
for n = 1:Ns-1

    % 1. Draw a uniformly distributed random number in [0,1]
    u = rand;

    % Transition rate associated with the current state,
    lambda = lambdae*x(n,:) + lambdac*not(x(n,:));
    % P{exactly one point in dt}
    P = exp(-lambda*dt).*(lambda*dt);
    % P = lambda*dt; % approximation for lambda*dt << 1

    % 2. If u > P{exactly one point in dt}, the state changes.
    x(n+1) = (u < P)*not(x(n)) + (u >= P)*x(n);

    % Update time index n -> n+1.
    % Return to 1.

end
```

Code 2.1: Implementation of the RTS generator.

These two very simple equations allows us to implement a RTN generator in discrete time. A discrete-time Markov process that takes a finite number of possible values is called a *Markov chain*. Algorithm is presented in the form of a Matlab® code in Code 2.1. A few lines are sufficient to generate a RTN trace. As an illustration, a few telegraph signals output by our generator are depicted in Fig. 2.3. The user is obviously free to chose all the parameters: time window T, number of samples N_s, time step $\mathrm{d}t$, the RTN properties $\bar{\tau}_c$, $\bar{\tau}_e$ and ΔI^*. This is a powerful tool to validate analysis methods, for instance spectral estimation as described in the next chapter.

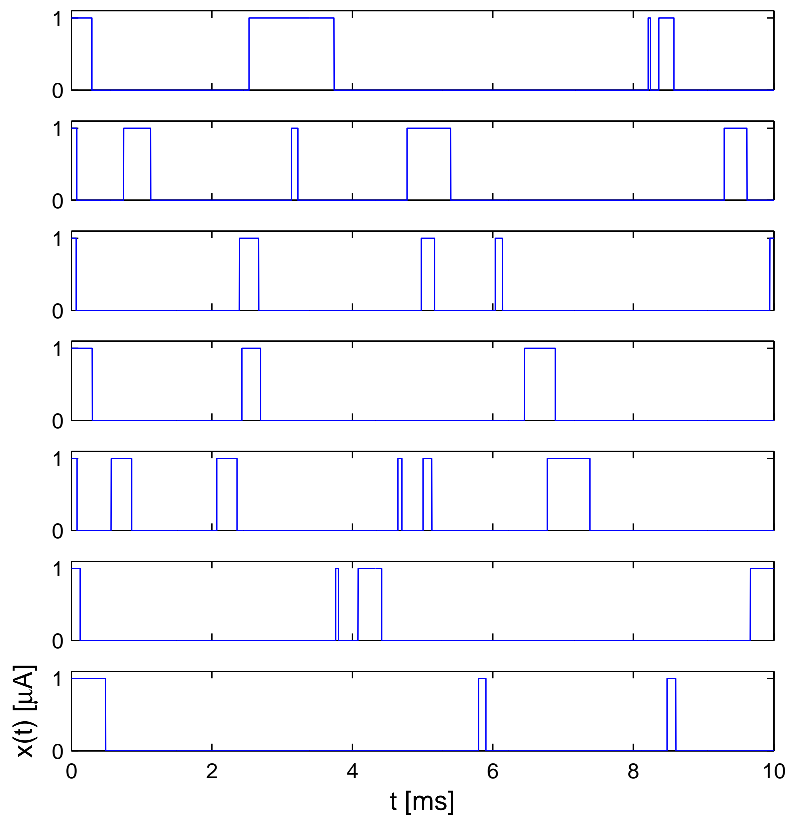


Figure 2.3: Seven RTN traces output by our RTS generator. Parameters are $T = 10\text{ms}$, $N_s = 10\text{k}$, $\bar{\tau}_c = 3\text{ms}$, $\bar{\tau}_e = 0.3\text{ms}$ and $\Delta I^* = 1\mu\text{A}$.

2.2.3 Correlation function of the RTN

We end up this chapter with the computation of the RTN correlation function as defined in (2.7), which will allows us later to obtain the analytical expression of the RTN power spectral density, i.e. the Lorentzian. Our demonstration is inspired by two pioneer papers that are [10, 13], although our derivation contains more details.

We apply the definition (2.7) for a discrete-state WSS process process:

$$\begin{aligned} R_X(t) &\equiv \mathbb{E}\left\{X(s+t)\,X(s)\right\} \\ &= \sum_i \sum_j x_i x_j \, P\left(X(s+t)=x_i,\, X(s)=x_j\right) \\ &= \left(\Delta I^*\right)^2 \cdot P\left(X(s+t)=\Delta I^*,\, X(s)=\Delta I^*\right). \\ &= \left(\Delta I^*\right)^2 \cdot P\Big(X(s+t)=\Delta I^* \,|\, X(s)=\Delta I^*\Big) \cdot P\left(X(s)=\Delta I^*\right). \end{aligned} \tag{2.27}$$

Two probabilities are involved:

- $$P\left(X(s)=\Delta I^*\right) = d = \frac{\bar{\tau}_c}{\bar{\tau}_c+\bar{\tau}_e} = \frac{\bar{\lambda}_e}{\bar{\lambda}_c+\bar{\lambda}_e} \tag{2.28}$$

  derived from the definition of the mean time constants.

- We may observe that

  $$P\Big(X(s+t)=\Delta I^* \,|\, X(s)=\Delta I^*\Big)$$

  is also

  $$\begin{aligned} &P\Big(X(t)=\Delta I^* \,|\, X(0)=\Delta I^*\Big) \\ &= P\left(\text{even number of transitions in } t, \text{ starting in state } \Delta I^*\right) \\ &\triangleq P_{even}(t). \end{aligned} \tag{2.29}$$

  Defining

  $$P_{odd}(t) \triangleq P\left(\text{odd number of transitions in } t, \text{ starting in state } \Delta I^*\right), \tag{2.30}$$

  we have

  $$P_{even}(t) + P_{odd}(t) = 1. \tag{2.31}$$

  The time evolution of $P_{even}(t)$ and $P_{odd}(t)$ is based on the kinetic

equations (2.26) and (2.25):

$$\begin{aligned}
& P_{even}(t+\mathrm{d}t) \\
&= P\left(X(t)=\Delta I^* \,|\, X(0)=\Delta I^*\right)\cdot P\left(\text{no transition } \Delta I^* \to 0 \text{ during d}t\right) \\
&\quad + P\left(X(t)=0 \,|\, X(0)=\Delta I^*\right)\cdot P\left(\text{transition } \Delta I^* \to 0 \text{ during d}t\right) \\
&= P_{even}(t)\cdot\left(1-\bar{\lambda}_c\cdot\mathrm{d}t\right)+P_{odd}(t)\cdot\bar{\lambda}_e\cdot\mathrm{d}t \\
&= P_{even}(t)\cdot\left(1-\bar{\lambda}_c\cdot\mathrm{d}t\right)+\left(1-P_{even}(t)\right)\cdot\bar{\lambda}_e\cdot\mathrm{d}t \\
&= P_{even}(t)-P_{even}(t)\cdot\left(\bar{\lambda}_c+\bar{\lambda}_e\right)\cdot\mathrm{d}t+\bar{\lambda}_e\cdot\mathrm{d}t.
\end{aligned} \tag{2.32}$$

Rearranging and taking $\mathrm{d}t \to 0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}P_{even}(t) = -\left(\bar{\lambda}_c+\bar{\lambda}_e\right)\cdot P_{even}(t)+\bar{\lambda}_e, \tag{2.33}$$

which can be solved, given the initial condition $P_{even}(0)=1$ (there is zero transition at time 0), to yield

$$P_{even}(t) = \frac{\bar{\lambda}_e}{\bar{\lambda}_c+\bar{\lambda}_e}+\frac{\bar{\lambda}_c}{\bar{\lambda}_c+\bar{\lambda}_e}\exp\left(-\left(\bar{\lambda}_c+\bar{\lambda}_e\right)\cdot t\right) \tag{2.34}$$

Finally, injecting (2.28) and (2.34) in (2.27), the correlation function reads

$$\boxed{\begin{aligned}
R_X(t) &= \left(\Delta I^*\right)^2\cdot\left(\frac{\bar{\lambda}_e}{\bar{\lambda}_c+\bar{\lambda}_e}\right)^2 \\
&\quad + \left(\Delta I^*\right)^2\cdot\frac{\bar{\lambda}_c\,\bar{\lambda}_e}{\left(\bar{\lambda}_c+\bar{\lambda}_e\right)^2}\cdot\exp\left(-\left(\bar{\lambda}_c+\bar{\lambda}_e\right)\cdot t\right).
\end{aligned}} \tag{2.35}$$

2.3 Conclusion

This chapter has set the basis for the future. The essential notions about stochastic processes have been reviewed in the specific context of noise in MOSFET. One remaining fundamental notion has been repeatedly announced, the power spectral density which is the topic of the next chapter.

Capture and emission times were found to be exponentially distributed around their means. This has to been confirmed from the processing of experimental data and will be done in Chapter 5. A RTN generator has been successfully implemented in Matlab® and constitutes a useful tool to test some processing methods on the RTN. Finally, the correlation function of the RTN has been calculated, anticipating development of the next chapter.

Chapter 3

Spectral representation of the noise

This chapter develop another representation of the second-order statistics of a stochastic process, the power spectral density. Integration of random telegraph noise in SPICE simulators is only possible if accurate noise power spectral densities has been extracted in a given technology. This work uses a simple spectral estimation method based on measured traces. The implementation is validated using our RTN generator. Finally, a decisive criterion to select the experimental time window is suggested.

3.1 Power spectral density

The transient simulation of the noise in a circuit is a computational intensive task. We might rather be interested in the average electrical impact of the noise. In this regard, the *average noise power* delivered by a source noise is an important quantity. We address step by step the all the notions of power of a stochastic process.

3.1.1 Constructive definition

We construct the definition of the *power spectral density* (PSD) entirely from power consideration by exploiting Parseval's theorem.

Notions of power of a stochastic process

Definitions Let the noise be $X(t)$, a real wide-sense stationary (WSS) stochastic process with zero mean

$$\mathbb{E}\left\{X(t)\right\} = 0. \tag{3.1}$$

It is further assumed that the noise is a causal signal:

$$X(t \leq 0) = 0. \tag{3.2}$$

- The *instantaneous noise power*

$$|X(t)|^2 \tag{3.3}$$

is also a stochastic process.

- The *variance*

$$\begin{aligned}\sigma_X^2 &\equiv \mathbb{E}\left\{|X(t)|^2\right\} \\ &= \int_{-\infty}^{+\infty} x^2 f_X(x)\, \mathrm{d}x.\end{aligned} \tag{3.4}$$

is nothing but the *expected instantaneous power*. It is deterministic and independent of t due to the stationarity.

- The *time-averaged noise power*, or the *mean-square variation* is by definition

$$\left\langle |X(t)|^2 \right\rangle \equiv \lim_{T\to\infty} \frac{1}{T} \int_0^T |X(t)|^2 \, \mathrm{d}t. \tag{3.5}$$

In general, it is a random variable because (3.5) gives different values for different realizations of $X(t)$.

Ergodic process If and only we assume that $X(t)$ is *variance-ergodic*, the time-averaged noise power (3.5) is a deterministic quantity coinciding with the variance [26] (3.4):

$$\sigma_X^2 \equiv \mathbb{E}\left\{|X(t)|^2\right\} = \left\langle |X(t)|^2 \right\rangle \equiv \lim_{T\to\infty} \frac{1}{T} \int_0^T |X(t)|^2 \, \mathrm{d}t \tag{3.6}$$

Ensemble average, or statistical average, carried out by the operator $\mathbb{E}\{\cdot\}$, can be replaced by the time average $\langle \cdot \rangle$ for the ergodic process.

Nonergodic process In the most general case, we note that the *expected time-averaged noise power*

$$\mathbb{E}\left\{\left\langle |X(t)|^2 \right\rangle\right\} \tag{3.7}$$

is always deterministic and coincides with the variance (3.4) provided that expectation operator $\mathbb{E}\left\{\,\cdot\,\right\}$ commutes with the time-average operator $\langle\,\cdot\,\rangle$:

$$\begin{aligned}\mathbb{E}\left\{\left\langle\left|X\left(t\right)\right|^{2}\right\rangle\right\} &= \mathbb{E}\left\{\lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\left|X\left(t\right)\right|^{2}\,\mathrm{d}t\right\}\\ &= \lim_{T\to\infty}\mathbb{E}\left\{\frac{1}{T}\int_{0}^{T}\left|X\left(t\right)\right|^{2}\,\mathrm{d}t\right\}\\ &= \lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\mathbb{E}\left\{\left|X\left(t\right)\right|^{2}\right\}\,\mathrm{d}t\\ &= \lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\sigma_X^2\,\mathrm{d}t\\ &= \sigma_X^2.\end{aligned} \tag{3.8}$$

More compactly written:

$$\begin{aligned}\mathbb{E}\left\{\left\langle\left|X\left(t\right)\right|^{2}\right\rangle\right\} &= \left\langle\mathbb{E}\left\{\left|X\left(t\right)\right|^{2}\right\}\right\rangle\\ &= \left\langle\sigma_X^2\right\rangle\\ &= \sigma_X^2.\end{aligned}$$

Summary: variance is noise power It has been shown that the variance may always be interpreted as a power, for both ergotic and nonergodic processes. Table 3.1 summaries the relations. Despite the fact that the average noise power is computed for an infinitely long time, an expectation is required in the nonergotic case.

Process	Relation variance $\longleftrightarrow$ power
Ergodic	Average noise power $\sigma_X^2 = \left\langle\left|X\left(t\right)\right|^{2}\right\rangle \equiv \lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\left|X\left(t\right)\right|^{2}\,\mathrm{d}t$
Nonergodic	Expected average noise power $\sigma_X^2 = \mathbb{E}\left\{\left\langle\left|X\left(t\right)\right|^{2}\right\rangle\right\}$

Table 3.1: Power interpretation of variance for ergodic and nonergodic stochastic processes.

Parseval's theorem and power spectral density

Fourier transform and Parseval's theorem The Fourier transform $\mathcal{F}$ of a stochastic process $F(t)$

$$F(f) = \mathcal{F}\{F(t)\} = \int_{-\infty}^{+\infty} F(t) \exp(-j2\pi f t)\, \mathrm{d}t \tag{3.9}$$

is also a stochastic process. The Fourier transform is understood as an operator $\mathcal{F}\{\cdot\}$ converting every sample of $F(t)$ to its deterministic Fourier transform.

Parseval's theorem states the *total energy* of a signal can be calculated both in the time domain and the frequency domain:

$$\int_{-\infty}^{+\infty} |F(t)|^2\, \mathrm{d}t = \int_{-\infty}^{+\infty} |F(f)|^2\, \mathrm{d}f. \tag{3.10}$$

The modulus $|F(f)|$ is even when $F(t)$ is real. We use this parity to rewrite (3.10) as

$$\int_{-\infty}^{+\infty} |F(t)|^2\, \mathrm{d}t = 2 \cdot \int_{0}^{+\infty} |F(f)|^2\, \mathrm{d}f. \tag{3.11}$$

Definition of the PSD We define the *power spectral density* (PSD) $S_X(f)$ of the real WSS stochastic process $X(t)$ as the *unique* and *deterministic function of the frequency* f describing how the expected average noise power (or simply the average noise power if ergodicity is assumed) is distributed in frequency:

$$\sigma_X^2 \equiv \int_{0}^{+\infty} S_X(f)\, \mathrm{d}f \quad \text{(definition of the PSD)}. \tag{3.12}$$

Word "power" comes from the power interpretation of the variance. "Spectral Density" stands for the fact that the total average noise power is computed trough a integration of $S_X(f)$ over all frequencies.

Quite obviously, power calculation in the time domain may be related to some calculation in the frequency domain thanks to Parseval's theorem that will be our starting point. Let us consider $X_T(t)$, a truncated version of $X(t)$ by a rectangular window of duration T:

$$X_T(t) = X(t) \cdot \mathrm{rect}_{[0,T]}(t) = \begin{cases} X(t) & \text{if } 0 \leq t \leq T \\ 0 & \text{else.} \end{cases} \tag{3.13}$$

By the way, our approach is constructive as electrical noise is in practice only measured during some limited period T and not an infinitely long time. Parseval's theorem (3.11) for $F(t) = X_T(t)$ reads

$$\int_{0}^{T} |X_T(t)|^2\, \mathrm{d}t = \int_{0}^{+\infty} 2 \cdot |X_T(f)|^2\, \mathrm{d}f. \tag{3.14}$$

Dividing through by T:

$$\frac{1}{T}\int_0^T |X_T(t)|^2 \, \mathrm{d}t = \int_0^{+\infty} \frac{2}{T}|X_T(f)|^2 \, \mathrm{d}f. \tag{3.15}$$

Given a window of time T, the one-sided *periodogram* of $X(t)$ is by definition [26, 28, 29]

$$\boxed{\begin{aligned} \widehat{S}_T(f) &\triangleq \frac{2}{T}|X_T(f)|^2 \\ &= \frac{2}{T}\left|\int_{-\infty}^{+\infty} X_T(t)\exp(-j2\pi f t)\, \mathrm{d}t\right|^2 \\ &= \frac{2}{T}\left|\int_0^T X(t)\exp(-j2\pi f t)\, \mathrm{d}t\right|^2. \end{aligned}} \tag{3.16}$$

Equation (3.15) is shortly rewritten as

$$\frac{1}{T}\int_0^T |X_T(t)|^2 \, \mathrm{d}t = \int_0^{+\infty} \widehat{S}_T(f) \, \mathrm{d}f, \tag{3.17}$$

So far, quantities of (3.17) are random, and depends on T.

- Application of the operator $\lim_{T\to\infty}\{\cdot\}$ to both members of (3.17) yields

$$\begin{aligned} \lim_{T\to\infty}\frac{1}{T}\int_0^T |X_T(t)|^2 \, \mathrm{d}t &= \lim_{T\to\infty}\int_0^{+\infty} \widehat{S}_T(f)\, \mathrm{d}f \\ \lim_{T\to\infty}\frac{1}{T}\int_0^T |X(t)|^2 \, \mathrm{d}t &= \int_0^{+\infty} \lim_{T\to\infty}\widehat{S}_T(f)\, \mathrm{d}f. \end{aligned} \tag{3.18}$$

  - If and only if ergodicity is assumed, the left-hand side may be identified with the variance:

$$\sigma_X^2 = \lim_{T\to\infty}\frac{1}{T}\int_0^T |X_T(t)|^2 \, \mathrm{d}t = \int_0^{+\infty} \lim_{T\to\infty}\widehat{S}_T(f)\, \mathrm{d}f. \tag{3.19}$$

    And the power spectral density defined in (3.12) reads

$$S_X(f) = \lim_{T\to\infty}\widehat{S}_T(f). \tag{3.20}$$

    Ergodicity makes unnecessary the application of the expectation operator $\mathbb{E}\{\cdot\}$ to recover the variance.

  - In the general case of a nonergodic process, both members of (3.18) are random quantities so that a unique deterministic power spectral density cannot be identified.

- Application of the operator $\mathbb{E}\{\cdot\}$ to both members of (3.17) yields

$$\begin{aligned}\mathbb{E}\left\{\frac{1}{T}\int_0^T |X_T(t)|^2 \,\mathrm{d}t\right\} &= \mathbb{E}\left\{\int_0^{+\infty} \widehat{S}_T(f)\,\mathrm{d}f\right\} \\ \sigma_X^2 &= \int_0^{+\infty} \mathbb{E}\left\{\widehat{S}_T(f)\right\}\mathrm{d}f.\end{aligned} \tag{3.21}$$

  We were able to identify the variance. However, the deterministic quantity

$$\mathbb{E}\left\{\widehat{S}_T(f)\right\} \tag{3.22}$$

  is a function of f with a parametric dependence on T and is thus not unique. Uniqueness is recovered by taking $\lim_{T\to\infty}\{\cdot\}$.

As a summary, the unique and deterministic function of frequency f frequency verifying (3.12) is the one-sided power spectral density defined as

$$\boxed{\begin{aligned}S_X(f) &= \lim_{T\to\infty}\mathbb{E}\left\{\widehat{S}_T(f)\right\} \\ &= \lim_{T\to\infty}\mathbb{E}\left\{\frac{2}{T}|X_T(f)|^2\right\} \\ &= \lim_{T\to\infty}\mathbb{E}\left\{\frac{2}{T}\left|\int_0^T X(t)\exp(-j2\pi f t)\,\mathrm{d}t\right|^2\right\}.\end{aligned}} \tag{3.23}$$

3.1.2 Link with the statistics

Wiener-Khinchin theorem

The *Wiener-Khinchin theorem*, given here without proof, states that the power spectral density is the Fourier transform of the correlation function defined in (2.7) [26, 30]:

$$\boxed{\begin{aligned}S_X(f) &= 2\cdot\mathcal{F}\{R(t)\} \\ &= 2\cdot\int_{-\infty}^{+\infty} R_X(t)\exp(-j2\pi f t)\,\mathrm{d}t \\ &= 4\cdot\int_0^{+\infty} R_X(t)\cos(2\pi f t)\,\mathrm{d}t\end{aligned}} \tag{3.24}$$

The factor 2 stands because we have defined a *one-sided* PSD. Definition (3.24) is proved to be equivalent is (3.23).

As of it, the power spectral density is a description of the second-order statistics of a stochastic process in the frequency domain. It does not tell us anything about the first-order probability function of the stochastic process. Noise sources in MOSFETs provide us great example to illustrate this fact.

Examples: noise sources in MOSFETs

- Thermal noise and flicker noise have a Gaussian distribution (2.3). Though, they have very different PSDs:
  - Thermal noise is a stochastic process with decorrelated values, i.e. a white noise. From (2.10), the Fourier transform pair reads

    $$R_W(t) = \sigma_W^2 \delta(t) \overset{\mathcal{F}}{\longleftrightarrow} S_W(f) = 2 \cdot \sigma_W^2. \tag{3.25}$$

    The *color of noise* refers to the spectral profile. The thermal noise is *white* because its power spectral density is constant.
  - Flicker noise is also called *pink* noise, because its power spectral density decreases in frequency as $1/f^a$, with a a parameter close to one.
- Theorem (3.24) is the tool that allows us to calculate the power spectral density of the RTN. Plugging $R(t)$ of (2.35) into the integral yields after some calculation:

  $$\begin{aligned} S_X(f) = (\Delta I^*)^2 \frac{4\bar{\tau}^2}{\bar{\tau}_e + \bar{\tau}_c} \frac{1}{1 + (2\pi f \bar{\tau})^2} \\ + (\Delta I^*)^2 \left(\frac{\bar{\tau}_c^2}{\bar{\tau}_e + \bar{\tau}_c} \right)^2 \delta(f) \end{aligned} \tag{3.26}$$

  with

  $$\bar{\tau} \triangleq \left(\frac{1}{\bar{\tau}_e} + \frac{1}{\bar{\tau}_c} \right)^{-1}. \tag{3.27}$$

  The second term of (3.28) is a pure DC term that will be omitted most of the time, assuming the RTN has a zero mean. The *Lorentzian* is the RTN PSD:

  $$\boxed{S_{\mathcal{L}}(f) = (\Delta I^*)^2 \frac{4\bar{\tau}^2}{\bar{\tau}_e + \bar{\tau}_c} \frac{1}{1 + (2\pi f \bar{\tau})^2}} \tag{3.28}$$

  Again, we emphasis that the RTN process is not fully defined by its power spectral density (3.28). A first argument is that the Lorentzian expression in perfectly symmetric in $\bar{\tau}_e$ and $\bar{\tau}_c$, while these two parameters obviously determine the shape of the two-state traces in the time domain, and consequently the amplitude of the two Dirac distributions of the pdf (2.4).

Spectral factorization and innovation To go further in this fact, we synthesize realizations of a Gaussian stochastic process but whose PSD is also the Lorentzian (3.28). Let us consider a system whose transfer function is $L(f)$. The input is $W(t)$, while the output is $X(t)$. The effect of the

system on the power spectral density can be derived from Wiener-Khinchin theorem (3.24) [26, 29]:

$$S_X(f) = |L(f)|^2 \cdot S_W(f) \tag{3.29}$$

If we desire the shape a stochastic process with some power spectral density $S_X(f)$, (3.29) tells us that such a PSD is obtained by filtering unit variance white noise with the a transfer function $L(f)$ such that

$$S_X(f) = |L(f)|^2 = L(f)\, L(f)^* . \tag{3.30}$$

This is known as *spectral factorization*, and $L(f)$ is called an *innovation filter* [26, 29].

The synthesized realizations of $X(t)$ are plotted in Fig. 3.1. Correlation between samples is clearly observed, the process is not white. However, it is normally distributed and is definitely not a RTN, while its PSD is exactly the Lorentzian (3.28), as demonstrated in Fig. 3.2.

3.2 Spectral estimation

The spectral estimation problem can be stated as follow. From a single finite-duration realization $x(t)$ of a real WSS stochastic process $X(t)$, we wish to determine an estimate $\widehat{S}_T(f)$ of its power spectral density $S_X(f)$ [26, 30, 31].

3.2.1 Periodogram method

Description

Nonparametric methods of spectral estimation rely directly and entirely on the definition (3.23) of the PSD [30]. The definition includes two operators:

- $\lim_{T\to\infty}$, stating that the power spectral density must theoretically be computed from a infinitely long trace.
- $\mathbb{E}\{\,\cdot\,\}$, that requires either that the realization be perfectly representative of the stochastic process overall behaviour, or that the PSD is computed as a ensemble average.

According to our previous discussion, we might let open the question of the presence of $\mathbb{E}\{\,\cdot\,\}$ in the case where $X(t)$ is variance-ergodic. As proving the ergodicity of a given process is a difficult task [26], we remain fully general.

The simplest estimate is obtained by "forgetting" or at least loosening the constraints imposed by these two operators:

$$\begin{aligned} S_X(f) &\approx \widehat{S}_T(f) \\ &= \frac{2}{T}|X_T(f)|^2 \\ &= \frac{2}{T}\left|\int_0^T x_T(t)\exp(-j2\pi f t)\;\mathrm{d}t\right|^2 . \end{aligned} \tag{3.31}$$

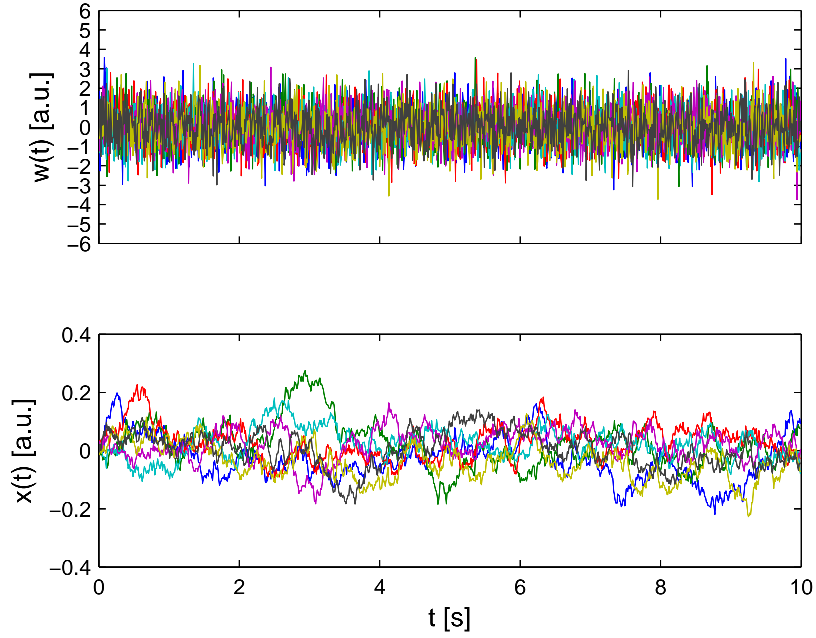


Figure 3.1: Seven $x(t)$ traces (bottom) synthesized from white Gaussian noise (top) with an innovation filter. Parameters are $\bar{\tau}_c = 1\text{s}$, $\bar{\tau}_e = 1\text{s}$ and $\Delta I^* = 1a.u.$.

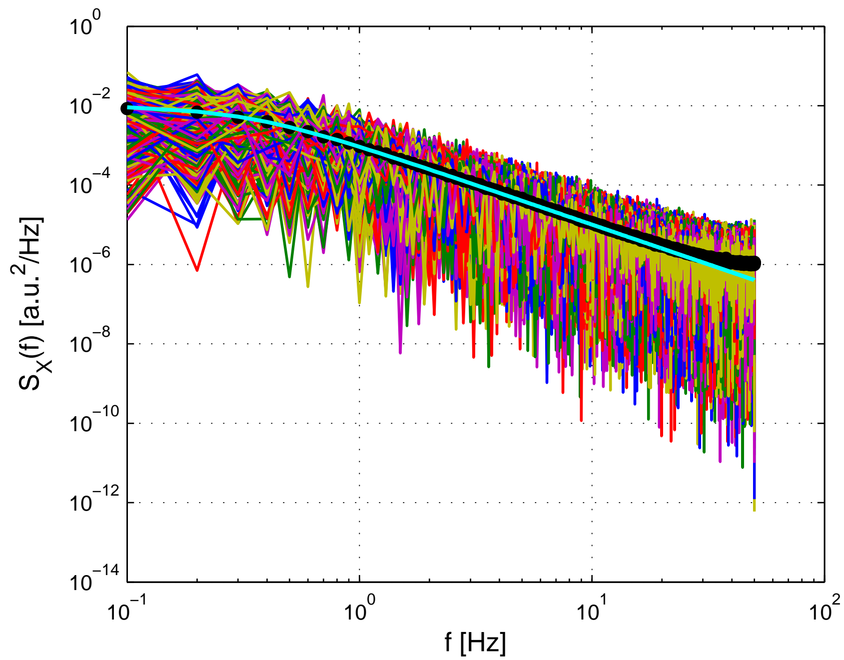


Figure 3.2: Corresponding power spectral density, which is well a Lorentzian, while the synthesized process is not a random telegraph noise.

Hopefully, several realizations

$$\{x_{T,1}(t), x_{T,2}(t), \dots\}$$

are available. We form the periodograms $\widehat{S}_{T,i}(f)$ for each of them and perform an ensemble average:

$$\boxed{\bar{\widehat{S}}_T(f) = \frac{1}{N_{traces}} \sum_{i=1}^{N_{traces}} \widehat{S}_{T,i}(f).} \tag{3.32}$$

Properties of the periodogram method

A general and quantitative discussion of the statistical properties of the spectral estimate is difficult. This advanced topic is widely detailed in dedicated books [30, 31], more summarily in [26]. Without reproducing all the very technical proofs, we outline the accuracy and the limitations of the periodogram method.

Bias As the time window T is finite, the true spectrum of the infinite duration signal $X(f)$ is polluted by the convolution with the spectrum of the rectangular window:

$$\begin{aligned} X_T(f) &= \mathcal{F}\{x(t)\} \\ &= \mathcal{F}\left\{x(t) \cdot \mathrm{rect}_{[0,T]}(t)\right\} \\ &= \mathcal{F}\{x(t)\} * \mathcal{F}\left\{\mathrm{rect}_{[0,T]}(t)\right\} \\ &= X(f) * \frac{\sin(\pi f T)}{\pi f}. \end{aligned} \tag{3.33}$$

Equation (3.33) makes clear that the exact spectrum is not obtained when T is finite. This inaccurately estimated spectrum causes the periodogram (3.31) to be a poor estimate of the PSD. The bias of the estimate (3.31) depends thus on the width of the window. Main results from the analysis developed in [26, 30] are

- If T is sufficiently large, so that $S_X(f)$ is nearly constant in a frequency interval of the order of $1/T$, $\widehat{S}_T(f)$ is asymptotically an unbiased estimator of $S_X(f)$:

$$\mathbb{E}\left\{\widehat{S}_T(f)\right\} = S_X(f). \tag{3.34}$$

- Otherwise, when T is short, the periodogram is a biased estimate of $S_X(f)$. This is an important point for the practical cases: a too short measurement period can lead to a an inaccurate spectral estimation, even if the number of available traces is large. The reason being that the expected value of each individual periodogram is not the sought $S_X(f)$.

Variance The variance of the periodogram can be shown to be [26, 30]

$$\mathrm{Var}\left\{\widehat{S}_T(f)\right\} = S_X^2(f) \text{ for } f \gg 1/T. \tag{3.35}$$

As a consequence, the larger the PSD of interested, the larger the variance. This was expected, and it does not tell anything about the influence of T. However, if we consider the ensemble average of several periodogram as stated in (3.32), we find

$$\mathrm{Var}\left\{\bar{\widehat{S}}_T(f)\right\} = \frac{1}{N_{traces}} S_X^2(f) \ . \tag{3.36}$$

This simple result is fundamental: it ensures us that increasing the number of realization allows to reduce the variance. However, we point out the fact that the *relatie standard error of the mean*

$$\varepsilon_{\bar{\widehat{S}}} \triangleq \frac{\sqrt{\mathrm{Var}\left\{\bar{\widehat{S}}_T\right\}}}{S_X} \tag{3.37}$$

decreases only as $\sim 1/\sqrt{N_{traces}}$, which means that reducing the error by a factor of two is expensive: the number of realizations must be raised by a factor four.

3.2.2 Implementation

From a practical point of view, the periodogram (3.31) is numerically evaluated for discrete frequencies

$$\begin{aligned} f_k &= k \cdot \mathrm{d}f \\ &= k \cdot \frac{1}{T} \quad k = 0, 1, \ldots, \frac{N_{NFFT}}{2} \end{aligned} \tag{3.38}$$

with a *discrete Fourier transform* (DFT), whose an efficient implementation is the fast Fourier transform (FFT). The spectrum of one realizations is computed as

$$\begin{aligned} X_T(f) \longrightarrow X_T(f_k) &= \int_0^T x_T(t) \exp(-j2\pi f_k t) \, \mathrm{d}t \\ &\approx \underbrace{\sum_{n=0}^{N_s-1} x_T(n \cdot \mathrm{d}t) \exp(-j2\pi f_k n \mathrm{d}t) \cdot \mathrm{d}t}_{X_{FFT}[k]}. \end{aligned} \tag{3.39}$$

The selection of appropriate measurement parameters is discussed in the next chapter that deals with sampling and digitizing.

```
F_two_sided = (1/NFFT)*(0:NFFT-1)';
kmax = floor(NFFT/2)+1;
indices_one_sided = (1:kmax)';
F = F_two_sided(indices_one_sided);
f = F*fs;

X = fft(x,NFFT)*dt;
X = X(indices_one_sided,:);
S = 2*(abs(X).^2)/T;
```

Code 3.1: Implementation of the DFT-based periodogram method.

As of it, the periodogram method requires only a few lines of Matlab® code, using the `fft` build-in function. Code 3.1 is a piece of program whose importance is huge for this work, oriented toward spectral analysis.

3.2.3 Validation of the software tool

The efficiency and accuracy of the developed spectral estimation method is tested based for two importance processes: a white Gaussian noise, as well as RTN directly generated by our software tool as described in Code 2.1 of previous chapter.

First test: White noise PSD

The most elementary test to validate Code 3.1 is to be sure that it properly and accurately estimate the constant power spectral density of a white noise.

Fig. 3.3 shows that our peridogram method smoothly converges to the flat PSD. Equation (3.36) is numerically confirmed by Fig. 3.4 that plots the relative mean-square error. About 100 traces are necessary to get as estimate as precise as 1%.

Second test: RTN PSD

As we did in previous chapter, we use the RTN generator to output a lot of traces such as the ones we showed in Fig. 2.3. Then, spectral estimation routine of Code 3.1 is applied to this traces.

The impact of the time window is illustrated by Fig. 3.5 and Fig. 3.6. When T is only 3ms, even the averaged periodogram is a biased, i.e. inaccurate, estimator of the analytical Lorentzian. This bias is observed as a vertical shift. The estimation is very satisfying for $T = 20$ms. The conclusion of this analysis is that the measurement period for experimental RTN traces should be much larger than the RTN time constants for accurate spectral estimation. Typically:

$$\boxed{T \sim 5 - 10 \cdot \max\left(\bar{\tau}_c, \bar{\tau}_e\right).} \tag{3.40}$$

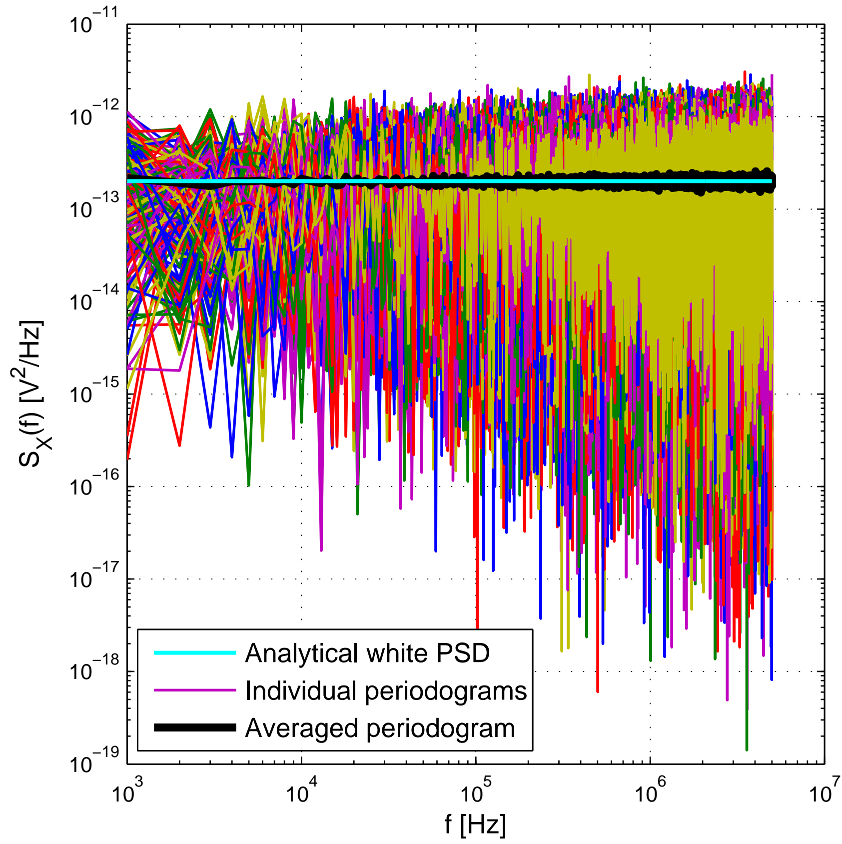


Figure 3.3: Estimation of white noise PSD using up to 250 traces.

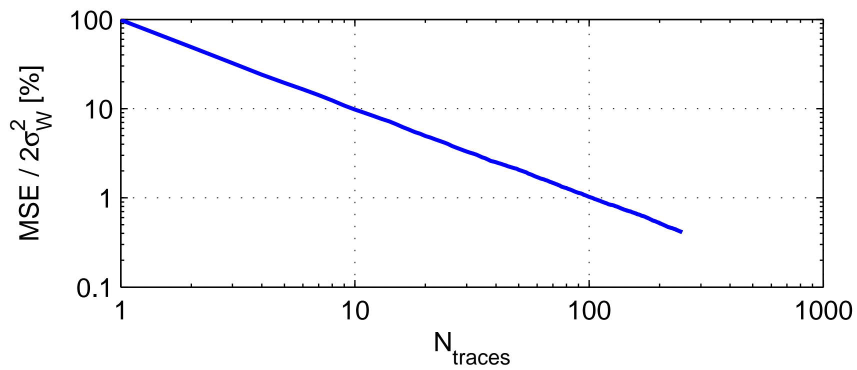


Figure 3.4: Computed mean-square error (MSE) scales as $1/N_{traces}$ as predicted.

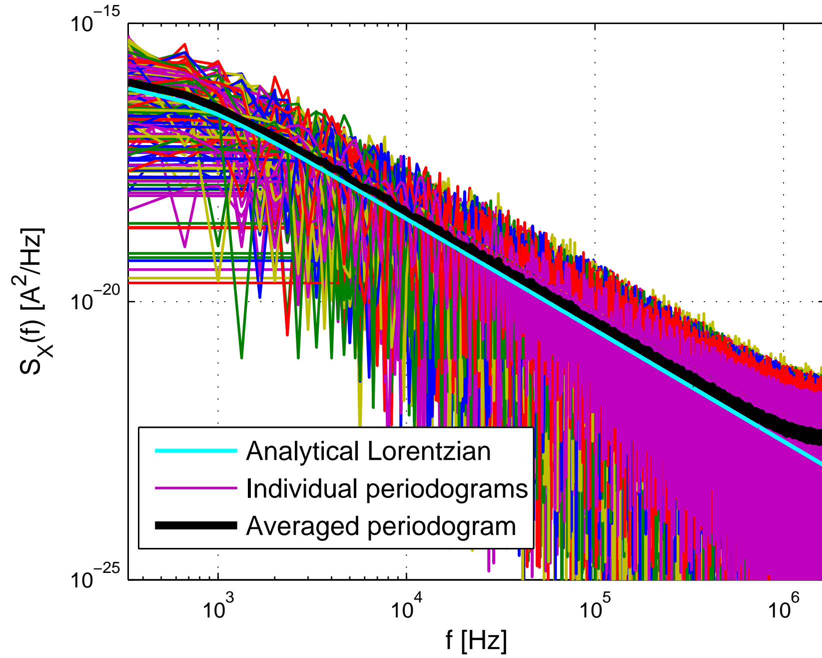


Figure 3.5: Estimation of the RTN PSD using 250 traces of duration $T = 3$ms. RTN parameters are $\bar{\tau}_c = 3$ms, $\bar{\tau}_e = 0.3$ms and $\Delta I^* = 1\mu$A. T is short regarding the RTN speed and the estimate is biased.

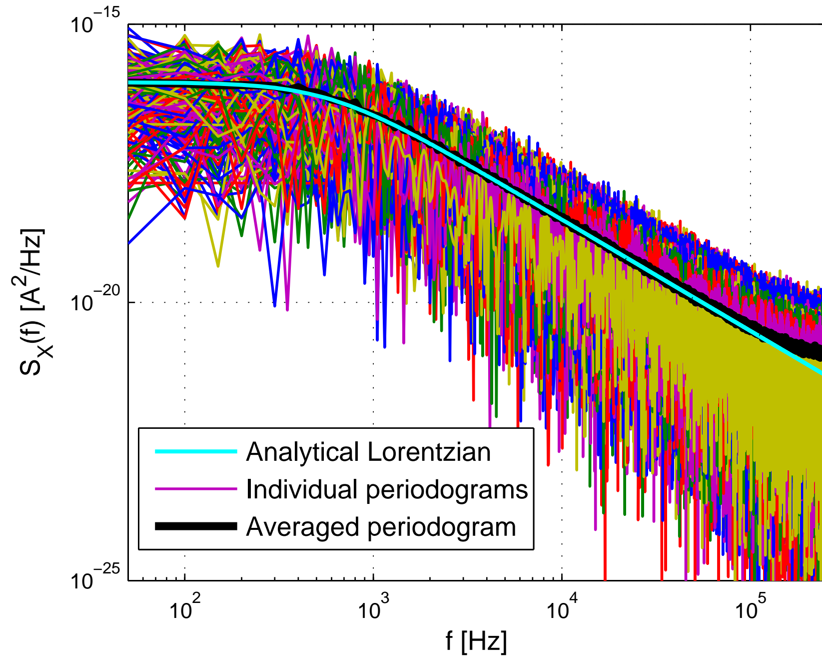


Figure 3.6: $T = 20$ms, the estimator gets unbiased thank to the large time window.

The number of traces is selected based on (3.36). A hundred traces is commonly satisfying.

3.3 Conclusion

Power spectral density is a fundamental quantity that describes how the average noise power is spread along the frequency axis on one hand, how the samples of a realization are correlated on the other hand.

Spectral estimation aims to estimate such a PSD based on experimental traces in the time-domain. This work focuses on the intuitive periodogram method. A quantitative criterion to select the time window has been established, combining theoretical prediction and results from numerical computation.

Chapter 4

Measurement parameter selection

The experimental characterization of RTN is a challenge. The may reason is that the values of $\bar{\tau}_c$ and $\bar{\tau}_e$ drastically vary from defect to defect, from tens of ns to years, based on energy level and spatial properties [1, 11]. When we want to perform direct time-domain measurement, equipment must be set to carefully selected time window T and sampling frequency f_s. A tradeoff between these two parameters arises from the memory limitation of the oscilloscope. Our choice will thus depend on the properties of the studied RTN.

We summarize the Figure of Merits (FoMs) of measurement equipment, in view of both time- and frequency-domain analysis of the RTN. It will be explained that the suitable choice of experimental parameters is strongly dependent on the RTN time constants and is not obvious at all without any prior knowledge. Theoretical study of the convergence and the error of important estimators, $\widehat{\bar{\tau}}$ and $\widehat{S}_T(f)$, provides a rough estimation of the required number of traces N_{traces} to be measured . Results of this chapter will be widely exploited in the next chapter to discuss performance of RTN characterization systems published in the literature on one hand, to review the results from our analysis of practical cases on the other hand.

4.1 From analog world to digital measurement

Noise is like any physical signal a pure analog signal, i.e. a continuous-time signal with a continuous range of values. From the experimental point of view, all the signals are discrete-time signals. Indeed, the measurement equipment, recording for instance the MOSFET drain current trace $I_D(t)$ on a given period T, captures only *samples* of the physical signal. The equipment is set to a given *sampling frequency*, denoted by f_s. The *sampling period*, or *time step* which corresponds to the time between two successive measurements, is

thus simply given by

$$\mathrm{d}t = \frac{1}{f_s}. \tag{4.1}$$

Consequently, all the time domain traces are only available for discrete-time values

$$t_0 = 0 < t_1 = 1 \cdot \mathrm{d}t < \ldots < t_n = n \cdot \mathrm{d}t < \ldots < t_{N_s} = T = N_s \cdot \mathrm{d}t$$

where N_s denotes the number of samples. The time origin t_0 has been set to 0 without loss of generality.

Given this experimental constraint, we move from a time-function representation to a time-sequence representation, which is convenient for numerical data processing. As a summary, taking the MOSFET drain current as an example, we have

$$I_D(t)\ \forall t \in \mathbb{R} \xrightarrow{\text{Sampling at } f_s=1/\mathrm{d}t} \{I_D(0), I_D(\mathrm{d}t), \ldots, I_D(n \cdot \mathrm{d}t), \ldots, I_D(T)\}.$$

Note that we voluntarily make a double usage of the notation $\mathrm{d}t$, both for a positive infinitesimal and for the time difference between two samples. The sampling period $\mathrm{d}t$ is assumed to be "small enough" to be treated as an infinitesimal. It means that no "major" change of the state of the system is missed when moving from the continuous-time physical signal to a measured time-sequence. "Small enough" actually means that $\mathrm{d}t$ is much smaller that the *characteristic time constants* of the studied stochastic process, as explained further.

Beyond the sampling in the time-domain, the amplitude of a signals is also sampled. While the analog signal takes its values over a continuous range, the digital signal is known with a finite accuary, determined by the the I_{LSB} or V_{LSB} of the equipment, e.g. the oscilloscope.

4.2 Time-domain analysis requirements

The capture and emission times of one given trap within one given device set suitable values for f_s and T ensuring a statistically accurate time-domain analysis, that is essentially time-constant extraction.

Relations are derived from the exponential distribution of the time constants around their means $\bar{\tau}_c$ and $\bar{\tau}_e$, (2.17) and (2.19) respectively:

$$f_{\mathcal{T}}(\tau) = \frac{1}{\bar{\tau}} \exp\left(-\frac{\tau}{\bar{\tau}}\right). \tag{4.2}$$

4.2.1 Sampling frequency and measurement period

The measurement process of any I or V traces is performed during a period T at a sampling frequency f_s, which yields a number of samples

$$N_s = f_s \cdot T \tag{4.3}$$

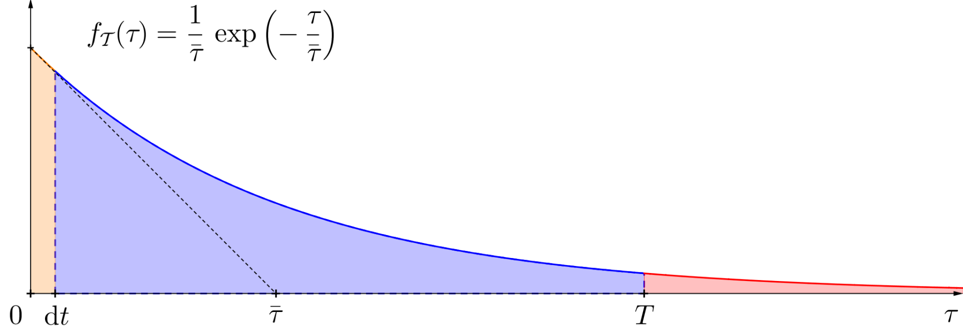


Figure 4.1: Impact of nonzero time step dt and finite T on the statistical relevance of the measured time constants.

Ideally, we wish to record a trace during an infinitely long time at an infinite sampling rate (i.e. zero time step dt), in order to recover the physical analog signal. The limiting figure for RTN characterization is generally the memory depth, which is a key parameter of the oscilloscopes. As a result of which we face a tough tradeoff between f_s and T, according to (4.3).

Capturing most of the events

- Sampling frequency should be high enough so that even very short times can be captured. If dt is the sampling period, it is clear that time constants such that

  $$\tau < \mathrm{d}t$$

  are missed. Our idea is to set a lower limit to dt so that less than 1% of short time-constants are lost:

  $$\begin{aligned} F_{\mathcal{T}}(\mathrm{d}t) &= P\left(\mathcal{T} \leq \mathrm{d}t\right) \\ &= \int_0^{\mathrm{d}t} f_{\mathcal{T}}(\tau)\, \mathrm{d}\tau \\ &= 1 - \exp\left(-\frac{\mathrm{d}t}{\bar{\tau}}\right) \\ &\approx \frac{\mathrm{d}t}{\bar{\tau}} \leq 1\%. \end{aligned} \tag{4.4}$$

  This probability, which is a miss rate, correponds to the orange surface below the pdf up to dt in Fig. 4.1. This calculation holds for both $\mathcal{T}_c$ and $\mathcal{T}_e$. We request

  $$\boxed{f_s = \frac{1}{\mathrm{d}t} \geq 100 \cdot 1/\min\left(\bar{\tau}_c, \bar{\tau}_e\right).} \tag{4.5}$$

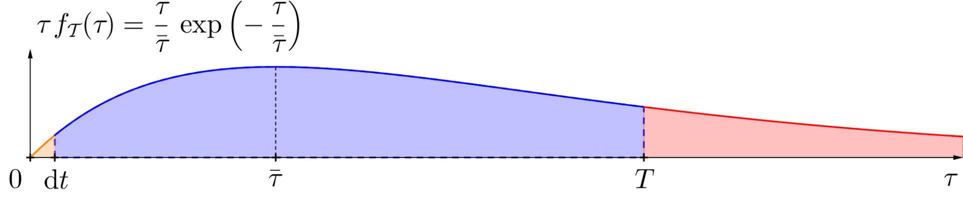


Figure 4.2: Impact of nonzero time step dt and finite T on estimation of the mean time constants.

- The experimental extraction of the tail of the distribution (5.39) will be accurate as long as the rare long times are well recorded. If T is the measurement period, it is clear that time constants such that

$$\tau > T$$

are missed. A lower limit for T is found so that less 1% than of long time-constants are lost:

$$\begin{aligned}\bar{F}_{\mathcal{T}}(T) &= P(\mathcal{T} \geq T) \\ &= 1 - P(\mathcal{T} \leq T) \\ &= 1 - \int_0^T f_{\mathcal{T}}(\tau)\,\mathrm{d}\tau \\ &= \exp\left(-\frac{T}{\bar{\tau}}\right) \leq 1\%,\end{aligned} \tag{4.6}$$

which can be solved for T:

$$T \geq \log(100)\,\bar{\tau} \approx 5 \cdot \bar{\tau}. \tag{4.7}$$

The miss rate for long time constants is the red surface below the pdf from T in Fig. 4.1. Again, calculation holds for both $\mathcal{T}_c$ and $\mathcal{T}_e$ and the specification reads

$$\boxed{T \geq 5 \cdot \max(\bar{\tau}_c, \bar{\tau}_e).} \tag{4.8}$$

Accurate estimation of the mean time constants

Recording most of the time constants is a major concern as we want obtain representative traces of the studied RTN. We observe however small and larger time constants does not equally contribute to the mean time constant. This section discussion the combined effect of dt and T on the mean estimation.

Effect of a too large time step Missing one rare long time constant has a stronger deleterious impact on the mean estimation than the loss of a few ultra short time constants. This is better understood by inspecting the shape of the function

$$\tau\, f_{\mathcal{T}}\left(\tau\right)$$

that we integrated to calculate the mean. As shown in Fig. 4.2, the effect of the time step is negligible as soon as $\mathrm{d}t$ is one order of magnitude smaller than $\bar{\tau}$. Miss rate condition (4.5) is thus strong and also ensures accurate estimation of $\bar{\tau}$ regarding $\mathrm{d}t$.

Effect of a too short time window The effect of a finite-time window is to truncate the exponential pdf of the random variable $\mathcal{T}$ at $\tau = T$, as time constants longer than T are never recorded. The deleterious consequence of cutting thr tail off is that the expected value of one time-constant $\mathcal{T}$

$$\mathbb{E}_T\left\{\mathcal{T}\right\} \equiv \int_{\mathrm{d}t}^{T} \tau\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau \tag{4.9}$$

may be substantially smaller than the real mean time constant

$$\bar{\tau} = \mathbb{E}\left\{\mathcal{T}\right\} \equiv \int_{0}^{\infty} \tau\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau. \tag{4.10}$$

We quantify the impact of T by computing

$$\begin{aligned}
\mathbb{E}_T\left\{\mathcal{T}\right\} &\equiv \int_{\mathrm{d}t}^{T} \tau\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau \\
&= \int_{0}^{\infty} \tau\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau - \int_{0}^{\mathrm{d}t} \tau\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau - \int_{T}^{\infty} \tau\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau \\
&\approx \bar{\tau}\left(\left(1 + \frac{1}{2}\left(\frac{\mathrm{d}t}{\bar{\tau}}\right)^2\right) - \int_{T}^{\infty} \frac{\tau}{\bar{\tau}}\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau\right) \\
&\approx \bar{\tau}\left(1 - \int_{T}^{\infty} \frac{\tau}{\bar{\tau}}\, f_{\mathcal{T}}\left(\tau\right)\, \mathrm{d}\tau\right) \\
&= \bar{\tau}\left(1 - \int_{T}^{\infty} \frac{\tau}{\bar{\tau}}\, \exp\left(-\frac{\tau}{\bar{\tau}}\right)\, \mathrm{d}\left(\frac{\tau}{\bar{\tau}}\right)\right) \\
&= \bar{\tau}\left(1 - \int_{T/\bar{\tau}}^{\infty} x\, e^{-x}\, \mathrm{d}x\right) \\
&= \bar{\tau}\left(1 - \left[-\left(x+1\right)\, e^{-x}\right]_{T/\bar{\tau}}^{\infty}\right) \\
&= \bar{\tau}\left(1 - \left(\frac{T}{\bar{\tau}} + 1\right)\, \exp\left(\frac{T}{\bar{\tau}}\right)\right).
\end{aligned} \tag{4.11}$$

Provided with (4.11), we find:

- for $T = 5 \cdot \bar{\tau}$,

$$\mathbb{E}_T \left\{ \mathcal{T} \right\} \approx 0.960 \cdot \bar{\tau},$$

which confirms *a posteriori* that this value of T is satisfying if we accept a relative finite-T induced error of $\sim 4\%$. Accuracy of $\sim 1\%$ is reached for $T \geq 6.64 \cdot \bar{\tau}$.

- for $T = 1 \cdot \bar{\tau}$,

$$\mathbb{E}_T \left\{ \mathcal{T} \right\} \approx 0.264 \cdot \bar{\tau},$$

If the measurement duration is similar to the largest mean time constant, either $\bar{\tau}_c$ or $\bar{\tau}_e$, the latter is underestimated by a factor of 4. This not that much intuitive result is fundamental.

4.2.2 Number of traces

The number of capture and emission events within *one single* trace never differ by more than one unit. The mean number of capture/emission events within one trace of duration T is estimated as

$$N_\tau \approx T / \max \left(\bar{\tau}_c, \bar{\tau}_e \right) \tag{4.12}$$

If $T = 5 \cdot \max \left(\bar{\tau}_c, \bar{\tau}_e \right)$ is used, we obtain $N_\tau \approx 5$, which is not enough to get a statistically accurate estimation of $\bar{\tau}_c$ and $\bar{\tau}_e$. This intuitive statement will be proved below.

The maximum likelihood estimator of $\bar{\tau}$ is the random variable

$$\hat{\bar{\tau}} = \frac{1}{N_\tau} \left(\mathcal{T}_1 + \mathcal{T}_2 + \ldots + \mathcal{T}_{N_\tau} \right) \tag{4.13}$$

corresponding to the *sample mean*. We work under the assumption that one sample $\mathcal{T}$ is rigorously exponentially distributed as (5.39). This is valid as long as T is sufficiently large, so that the tail is not cut off. The value $T \approx 5 \cdot \max \left(\bar{\tau}_c, \bar{\tau}_e \right)$ fulfills such a condition, as we saw. Estimator (4.13) is trivially shown to be unbiased:

$$\begin{aligned} \mathbb{E} \left\{ \hat{\bar{\tau}} \right\} &= \mathbb{E} \left\{ \frac{1}{N_\tau} \left(\mathcal{T}_1 + \mathcal{T}_2 + \ldots + \mathcal{T}_{N_\tau} \right) \right\} \\ &= \frac{1}{N_\tau} \left(\mathbb{E} \left\{ \mathcal{T}_1 \right\} + \mathbb{E} \left\{ \mathcal{T}_2 \right\} + \ldots + \mathbb{E} \left\{ \mathcal{T}_{N_\tau} \right\} \right) \\ &= \frac{N_\tau \cdot \bar{\tau}}{N_\tau} \\ &= \bar{\tau}. \end{aligned} \tag{4.14}$$

More useful then is the expression of the variance of the estimator:

$$\begin{aligned}
\mathbb{E}\left\{\left(\widehat{\bar{\tau}}-\bar{\tau}\right)^2\right\} \equiv \mathrm{Var}\left\{\widehat{\bar{\tau}}\right\} &= \mathrm{Var}\left\{\frac{1}{N_\tau}\left(\mathcal{T}_1+\mathcal{T}_2+\ldots+\mathcal{T}_{N_\tau}\right)\right\} \\
&= \frac{1}{N_\tau^2}\left(\mathrm{Var}\left\{\mathcal{T}_1\right\}+\mathrm{Var}\left\{\mathcal{T}_2\right\}+\ldots+\mathrm{Var}\left\{\mathcal{T}_{N_\tau}\right\}\right) \\
&= \frac{N_\tau\cdot\bar{\tau}^2}{N_\tau^2} \\
&= \frac{\bar{\tau}^2}{N_\tau},
\end{aligned} \tag{4.15}$$

where we have used the variance of an exponentially distributed random variable $\mathcal{T}$:

$$\mathrm{Var}\left\{\mathcal{T}\right\} = \bar{\tau}^2. \tag{4.16}$$

Equation (4.15) yields an expression for the *relative standard error of the mean*:

$$\begin{aligned}
\varepsilon_{\bar{\tau}} \triangleq \frac{\sqrt{\mathrm{Var}\left\{\widehat{\bar{\tau}}\right\}}}{\bar{\tau}} &= \sqrt{\mathbb{E}\left\{\frac{\left(\widehat{\bar{\tau}}-\bar{\tau}\right)^2}{\bar{\tau}^2}\right\}} \\
&= \frac{1}{\sqrt{N_\tau}}.
\end{aligned} \tag{4.17}$$

A reliable statistical data set, ensuring accurate estimation of $\bar{\tau}_c$ and $\bar{\tau}_e$, should be composed of at least a number N_τ of realizations such that

$$\varepsilon_{\bar{\tau}} = \frac{1}{\sqrt{N_\tau}} \leq 1\%,$$

which gives us

$$N_\tau \geq 10000. \tag{4.18}$$

Still assuming $T \approx 5\cdot\max\left(\bar{\tau}_c,\bar{\tau}_e\right)$), we also obtain the minimum number of traces

$$N_{traces} = N_\tau/5 \geq 2000, \tag{4.19}$$

which is huge and may not always be applicable in practice. We pay dearly, in term of N_τ and N_{traces}, the reduction of $\varepsilon_{\bar{\tau}}$ as only $1/\sqrt{N_\tau}$. If the specification is released to an error of of $\sim 5\%$, the number of traces falls to 80, which is fairly acceptable.

We emphasis that the measurement of *one single* very long time window does not solve this issue, as T_{long} in this case should be such that

$$N_\tau \approx \frac{T_{long}}{\bar{\tau}} \geq 10000,$$

hence

$$T_{long} \geq 10000\cdot\bar{\tau},$$

which, combined with

$$f_s \geq 100 \cdot 1/\bar{\tau},$$

would require

$$N_s = f_s \cdot T_{long} \geq 10^6$$

samples, which is above the memory depth of most oscilloscopes. Again, loosening to constraint to 5% seems more realistic.

4.3 Spectral analysis specifications

First of all, it is worth remembering that a comprehensive spectral analysis analysis must imperatively be preceded by the time-domain one. The main reasons are:

- The purpose of spectral analysis is to estimate the noise Power Spectral Density (PSD). This figure is related by a Fourier transform to the correlation between noise samples. It is determined by the *second order* probability density function, while it is fully independent of the first order pdf that can only be studied through histograms in the time-domain.
- This work propose a spectral estimation method based on the DFT of analog traces. Measured signals are always a combination of both RTN and some background noise, especially $1/f$ noise. It is of the greatest to perform PSD estimation on the neatly separated signal, which is one outcome of the time-domain analysis.
- We wish to validate the spectral estimation of the RTN PSD with the analytical Lorentzian. Matching between both is only possible if accurate estimations of $\bar{\tau}_c$ and $\bar{\tau}_e$ are plugged into the expression.

As a consequence, it is not possible to select experimental parameters f_s and T uniquely based on spectral considerations. The guidelines that we developed for the time-domain extraction should be applied.

4.3.1 DFT parameters

In addition, we keep in mind that f_s and T respectively set the maximum frequency and the frequency resolution of the discrete spectrum computed by DFT.

As announced in the previous chapter, the periodogram (3.16) is numerically evaluated for discrete frequencies

$$f_k = k \cdot \mathrm{d}f = k \cdot \frac{1}{T}, \tag{4.20}$$

with a *discrete Fourier transform* (DFT), whose an efficient implementation is the fast Fourier transform (FFT):

$$\begin{aligned} S_X(f_k) \approx \widehat{S}_T(f_k) &= \frac{2}{T}\left|X_T(f_k)\right|^2 \\ &= \frac{2}{T} X_{FFT}[k] \cdot \mathrm{d}t, \end{aligned} \tag{4.21}$$

$$k = 0, 1, \ldots, \frac{N_{NFFT}}{2}$$

indexes the discrete frequencies. The difference between two successive frequencies is called the *spectral resolution*, denoted $\mathrm{d}f$ and is related to the measurement period T:

$$\mathrm{d}f = 1/T. \tag{4.22}$$

It is also the first non-zero discrete frequency, and the first point that we see in a logarithmic representation where 0Hz can not rigorously appear. If N_s is the number of samples forming the discrete sequence

$$\left\{x[n] = x(n \cdot \mathrm{d}t)\right\}_{n=0,1,\ldots,N_s-1}, \tag{4.23}$$

we always set $N_{FFT} = N_s$ so that there is no zero padding whose interpretation is sometimes troublesome. Using $N_{FFT} = 2^m$ is usually recommended [28, 30] yet, to fully enjoy the efficiency of the FFT algorithm.

As the traces are actually discrete-time sequences, Nyquist-Shannon sampling theorem states that

$$f_{max} = f_s/2. \tag{4.24}$$

The spectral characterization of the RTN itself is not expected to be limited by (4.24), as f_s was forced to be much larger than the corner frequency of the Lorentzian

$$f_c = \frac{1}{2\pi\bar{\tau}} \tag{4.25}$$

for the RTN extraction. Nevertheless, high-frequency estimation is motivated by the presence of $1/f$ noise that always exceed the RTN starting from some frequency depending on the RTN parameters and $1/f$-noise level. The higher f_c, the higher f_{max} (and thus f_s) should be, which is fully consistent with the time-domain constraint. As explained further, potential frequency limitation can be overcome by fitting and extrapolation provided that the lower-frequency points are accurate and smooth enough.

Our position is that neither the spectral resolution $\mathrm{d}f$ nor the maximum frequency f_{max} are expected to dominantly constraint T and f_s, which are, so far, already constrained by (4.8) and (4.5).

4.3.2 Accuracy of the periodogram method

$\widehat{S}_T(f)$ is not a single number such as $\hat{\bar{\tau}}$, it is a PSD and thus a function of f, Fourier transforms is involved. Obtaining inequality base an accuracy specification is more difficult than for $\hat{\bar{\tau}}$. Invoked quantities often depend on f, the values of the sought PSD,...

The previous chapter has quantitatively established some criterion's that guarantee the convergence of the spectral estimation resorting to the periodogram method.

- The periodogram suffer from bias when T is finite and not large enough, resulting from the convolution of the true spectrum with the parasitic rectangular window. We have came to the conclusion that the measurement period for experimental RTN traces should be much larger than the RTN time constants for accurate spectral estimation. Typically:

$$T \sim 5 - 10 \cdot \max\left(\bar{\tau}_c, \bar{\tau}_e\right), \tag{4.26}$$

  which is fortunately similar to (4.8).

- The averaged peridogram was found to have a relative standard error expressed as

$$\varepsilon_{\hat{\bar{S}}} \triangleq \frac{\sqrt{\mathrm{Var}\left\{\hat{\bar{S}}_T\right\}}}{S_X} \sim 1/\sqrt{N_{traces}} \tag{4.27}$$

  from which we have estimated at about a hundred the number of traces required to maintain the error below $\sim 1\%$. This was confirmed by our numerical validation tests. $N_{traces} \approx 100$ is, again, very consistent with the specified valued for the time-domain analysis.

4.4 Summary and discussion

We have pointed out the strong influence of the RTN parameters on the choice of suitable sampling frequency and time window. Given the tradeoff between T and f_s, neither of the two should be excessive. We recommend to select them based on the following rules of thumb that were proved mathematically and numerically to be relevant:

$$\boxed{\begin{array}{l} T \approx 5 \cdot \max\left(\bar{\tau}_c, \bar{\tau}_e\right) \\ f_s \approx \dfrac{100}{\min\left(\bar{\tau}_c, \bar{\tau}_e\right)} \end{array}} \tag{4.28}$$

Parameter		RTN speed	
		Slow	Fast
	$\bar{\tau}$	$\sim 0.1 - 1\text{s}$	$0.1 - 1\text{ms}$
Sampling frequency	$f_s = \dfrac{1}{\mathrm{d}t}$	$\sim 1\text{kHz}$	$\sim 1\text{MHz}$
Measurement period	T	$\sim 1 - 10\text{s}$	$\sim 1 - 10\text{ms}$

Table 4.1: Specifications of the parameters of the measurement equipment for both an accurate extraction of a symmetric RTN, depending on its speed defined by the time constants $\bar{\tau}_c$ and $\bar{\tau}_e$.

The product of the two yields the number of samples

$$N_s = f_s \cdot T = 500 \frac{\max(\bar{\tau}_c, \bar{\tau}_e)}{\min(\bar{\tau}_c, \bar{\tau}_e)}. \tag{4.29}$$

Equation (4.29) is very interesting, because it does tell us how more or less pronounced asymmetry of the RTN leads to a very high number of samples. The problem is even more complex if we consider the multi-trap RTN.

4.4.1 Impact of the RTN symmetry

Symmetric RTN

Considering first the case of a more or less symmetric RTN

$$\bar{\tau}_c \sim \bar{\tau}_e,$$

it is instructive to roughly classify the RTNs in two groups, depending on their speed. Table 4.1 focuses on typical RTN speeds that we met during this work. This case is very few restrictive and even $T = 10 \cdot \bar{\tau}$ can be selected while keeping the number of samples at 1k.

Asymmetric RTN

The next chapter deals with a single-trap but asymmetric RTN:

$$\bar{\tau}_c > 10 \cdot \bar{\tau}_e,$$

Plugging $\bar{\tau}_c = 10 \cdot \bar{\tau}_e$ in (4.29) gives $N_s = 5\text{k}$. If the oscilloscope has a memory depth of 10k, the constraints can be satisfied. Our mathematical analysis was developed several mounts after the measurement phase of the Smart array, and it appears that f_s was actually chosen to high to the detriment of T which will be proved to be to short.

In case of still more severe asymmetry in the RTN behaviour, for instance a factor 100, it becomes more and more difficult to simultaneously have the suitable f_s and T.

Remark It could happen that we are not able to fulfill all the requirements and specifications, for instance of our oscilloscope has a weak memory depth. Something has to be relaxed and we have to be critic about the results and very smart during the phase of post-processing of the collected data. We give two examples:

- If the time window T was found to be too small to record a statistically sufficient number of long capture or emission times occurring regularly, the estimation of the corresponding mean time constant, based a direct ensemble average, gets strongly inaccurate. However, the exploitation of *censored data* and some statistical post-processing make an accurate estimation possible.

- If the spectral analysis has a lower limit of $\mathrm{d}f = 1/T$, the extracted PSDs can be extended to lower frequency by fitting or extrapolation, as the shapes of $1/f$ or RTN PSD are well knowns. This is also true for the upper frequency limit f_{max}: the PSD can be extrapolated to higher frequency provided that the shape of the PSD obtained below f_{max} is consistent and the values sufficiently accurate.

Chapter 5

Study case: single-trap RTN

Previous chapter set all the theoretical and experimental background regarding noise characterization and analysis. This chapter aims to describe in details a comprehensive methodology to deal with measured RTN traces, and extracting relevant informations in a AC compact modelling perspective. Both time-domain and frequency-domain analysis are presented.

The study case is a golden device from imec Smart array, fabricated in 28nm bulk CMOS technology. It exhibits a dominant two-level RTN behaviour.

State-of-the-art regarding on-chip characterization is firstly reviewed. The main part of this chapter is the noise analysis using the Matlab® software tool developed in the frame of this work. The consistency of the different methods is frequently questioned and verified. The apotheosis of this chapter is probably the validation trough spectral estimation of the Lorentzian as the RTN power spectral density. Additional noise was found to be attributed to the flicker noise.

5.1 On-chip measurement of RTN

In order to accurately capture the statistics of variability and reliability mechanisms such as bias temperature instability (BTI) and random telegraph noise (RTN) in a given technology, large scale transistor arrays are necessary. The development of on-chip characterization systems for 45nm and 28nm CMOS technology have been addressed by a few groups.

We focus on systems whose output are full analog current of voltage traces which serve our purposes of time-domain analysis and DFT-based spectral estimation. Based on the developments of previous chapter, some recently published realizations are briefly discussed. Emphasis is placed on their respective limitations through quantitative Figures of Merit (FoMs) of measurement equipment, in view of both time- and frequency-domain analysis of the RTN. Particular attention is paid to imec Smart array described in [32],

paper in press that will be presented very soon during the 2017 Symposia on VLSI Technology and Circuits.

5.1.1 Recent on-chip characterization systems

Tang *et al.* (2013) [33] worked with a ring oscillator array in 65nm process. The architecture presents the advantage of giving an insight of the RTN impact on a fundamental analog block that are ring oscillators, by measuring the RTN induced frequency shit. Sampling time step goes as low as $\sim 1\mu s$, and the multiple traps behaviour was covered. Unfortunately, the output of the circuit are only digitized Δf-RTN traces, and not the full analog ones. These are of great interest as we wish to study the combined effect of RTN and $1/f$ noise in a compact modelling perspective. In the following, we present only state-of-the-art chips enabling collection of a large number of analog traces.

Realov and Sheapard (2013) [7], from Columbia University, implemented a system in 45nm CMOS process for time-domain measurement of RTN:

- The DUT array contains only 312 nMOSFETs (and as many pMOSFETs) for each of the six available dimensions W/L. Regarding device-to-device noise variability, tails can thus not be systematically and confidently observed. However, the accurate knowledge of the tails of the distributions is a major concern for the circuit applications whose design is mostly based on the worst case.

- The on-chip measurement unit consists of both current and voltage-mode analog-to-digital converters (ADCs), with $I_{LSB} = 2.44$nA and $V_{LSB} = 0.976$mV. Such values are not expected to be limiting, except for the capture of traps with very small electrical impact ΔI_D or ΔV_{th}.

- We may mainly blame the realization for its limited maximum sampling frequency of only 75kHz. Such a low value prevents from measuring fast RTN events whose typical time constants are smaller than $\sim 100\mu s - 1ms$, while these short-time RTNs may be deleterious for high-speed circuit applications. In addition to that, the trace-based spectral estimation (which was, obviously, not the authors' orientation) would be unavoidably limited to 37.5kHz.

- Finally, from the experimental point of view, their chip needs four channels for the DUT bias. We do not have such a number at UCL [1].

Beyond a few limitations of the implemented chip, their data post-processing is very rich and clearly described in the paper, and inspired some of our own analysis and software development.

[1] according to a private discussion with Valeria Kilchytska.

In 2015, KUL and imec published together about a variability and reliability characterization array fabricated in 28nm High-κ/Metal Gate (HK/MG) CMOS [34, 35]:

- The DUT array contains 4536 MOSFETs for each of the twelve available dimensions W/L.
- Full analog I_D-traces are extracted, but measurement is external using source measure unit (SMU) and thus not on-chip. Measurement process is less automate. Moreover, parasitic capacitances related to the external measurement environment are expected to be higher than typical on-chip capacitances.
- No information is available about I_{LSB}, nor the maximum sampling frequency. The bottleneck is expected to be the instrumentation rather than the chip itself.

This chip is more reliability (BTI) oriented and was not fully intended for RTN characterization, unlike its successor "Smart array" described in details thereafter.

Whitcombe *et al.* (2016) [36], from University of California (Berkeley), fabricated an improved version of ADC-based measurement platform, also in 28nm HK/MG bulk process:

- The DUT array contains 41472 nMOSFETs (and as many pMOSFETs), evenly split between two lengths L and three different V_{th}. We compute $\sim$ 7000 transistors for each case.
- 96×18 array of ADCs enables parallel measurement of hundreds devices to quickly obtain meaningful statistical data for a wide range of RTN time constants [36].
- No mention of I_{LSB} value is made in the paper.
- Measurement up to at least $f_s = 400$kHz has been reported.

5.1.2 imec Smart array

During the 2017 Symposia on VLSI Technology and Circuits, held in June 2017 in Kyoto, paper "A fully-integrated method for RTN parameter extraction" [32] will be presented by imec reliability group. Array as well as biasing, measurement and processing blocks are implemented in 28nm HK/MG bulk CMOS from TSMC®.

Beyond on-chip measurement of RTN traces, a fully-integrated analysis of these traces is performed: mean capture and emission time constants, $\bar{\tau}_c$ and $\bar{\tau}_e$ respectively, are extracted.

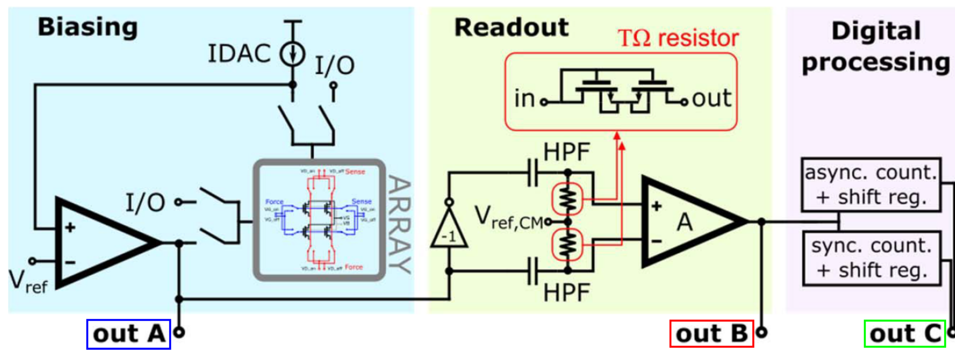


Figure 5.1: Circuit implementation of imec Smart array. From [32]. *out A*: analog ΔV_{th} traces. *out B*: digital ΔV_{th} traces. *out C*: $\bar{\tau}_c$ and $\bar{\tau}_e$. Adapted from [32].

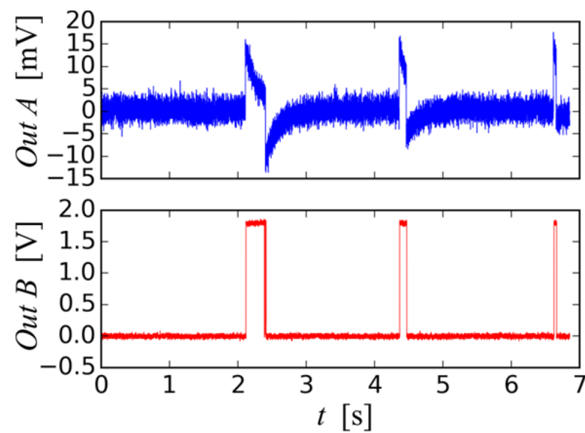


Figure 5.2: Full analog traces and digitized RTN traces. Mean time constants $\bar{\tau}_c$ and $\bar{\tau}_e$ are extracted from the latter. From [32].

Description of the measurement infrastructure

The full circuit implementation is illustrated in Fig. 5.1.

- The array contains 18144 nMOSFETs in four different geometries, hence 4536 for each, similarly to [32].
- Biasing and measurement of the DUT works as follows. Drain current I_D^* is fixed by the current source IDAC. Drain voltage V_D^* is maintained at a constant chosen value V_{ref} by the operational amplifier in feedback loop. Gate voltage V_g is left floating at the output of the op amp. The RTN-induced fluctuation is recorded by an AC measurement. This voltage shift is $\Delta V_{th}(t)$ and corresponds to *out A* of the chip (see Fig. 5.2). The DC value V_G (for fixed I_D^* and V_D^*) is also output.
- The readout block, based on a high-pass filter (HPF), yields in *out B* a rail-to-rail digital trace of the measured RTN (Fig. 5.2).
- The digital processing block consists of
  - one asynchronous counter for the number of capture/emission events;
  - the other synchronized with the clock signal that counts the number of clock periods the trap was charged or neutral.

  Knowing T_{clk}^*, mean time constants can directly be computed.

The whole circuit works properly up to at least 10MHz. The bottleneck is rather the instrumentation. *out A* and *out B* are recorded with a Tektronix ® TDS3054 oscilloscope. This has a very limited memory depth of 10k points per channel [2]. As a result of which we face the usual tradeoff up between the sampling frequency and the measurement period, according to

$$N_s = f_s \cdot T \leq 10\text{k}.$$

Fig. 5.2 shows a slow RTN, with $\bar{\tau}_e \sim 0.1 - 1$s. In that case, parasitic exponential decay of the analog signal is observed. This was found by the group to be due to the high-pass filter of the oscilloscope. For faster RTNs requiring high sampling rate, *out A* and *out B* were not measured synchronously. Consequently, the matching between the two and the accuracy of the on-chip extraction of time constants can only be verified statistically by computing $\bar{\tau}_e$ and $\bar{\tau}_c$ with a large number of traces. This time-consuming work has been done by Simicic *et al.* in [32] to validate the digital part of the circuit consisting in the readout and the digital processing blocks and their respective outputs *out B* and *out C*. Our further coming analysis also exploits digital traces as a validation tool.

[2] "TDS3000 Series User Manual" p.3-11 about Normal Acquisition.

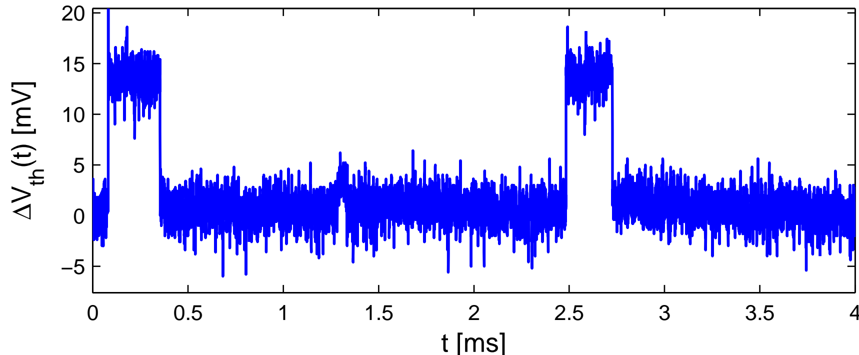


Figure 5.3: One noise trace among all the 999 measured on a 28nm bulk nMOSFET from TSMC®, with $W = 100$nm and $L = 30$nm. Bias is $V_{GS} = 470$mV and $V_{DS} = 50$mV. A golden two-level, i.e. single-trap, Random Telegraph Noise (RTN) is observed, with some background noise on top of it.

Our review of the literature leads us to think of the Smart array as the most advanced large-scale characterization platform of RTN. The RTN data exploited in depth in this chapter come from this state-of-the-art chip.

Our case study: single-trap RTN

Among the thousands transistors of the array, Marko Simicic identified at least one *golden device* in the sense that it exhibits a single active trap RTN behaviour on one hand, and that the RTN is fast enough to be essentially insensible to the exponential decay. One typical trace is depicted in Fig. 5.3.

Dimensions are $W = 100$nm and $L = 30$nm. Source-referenced applied voltages are $V_{GS} = 470$mV and $V_{DS} = 50$mV. I_D-V_{GS} curve seen in Fig. 5.4a and allows to identify the operation region of the transistor. We find $I_D \approx 2\mu$A. Device seems slightly above threshold and at the boundary of saturation and triode regions, despite low V_{DS}. These observations are made more rigorous by extracting the threshold voltage V_{th} that not provided to us.

V_{th} extraction The I_D-V_{GS} data points of Fig. 5.4a were found to be to few for a numerically accurate g_m/I_D extraction method described in [37,38]. We opted for a dg_m/dV_{GS}-based extraction. I_D-V_{GS} was fitted by a polynomial as illustrated in Fig. 5.4a. Exact derivative of this polynomial yields the g_m-V_{GS} curve in 5.4b. $g_m \approx 23\mu$A/V is read. A second differentiation give the dg_m/dV_{GS} curve in Fig 5.5. The maximum is found at

$$V_{th} \approx 440\text{mV}, \qquad (5.1)$$

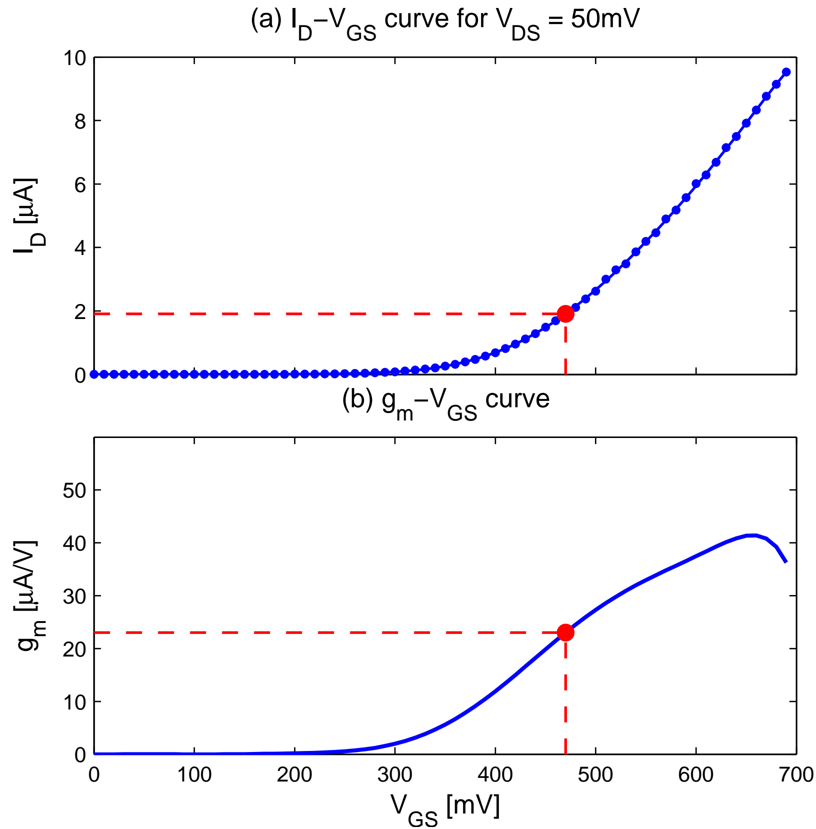


Figure 5.4: (a) I_D-V_{GS} and (b) g_m-V_{GS} curves with the DC measurement point emphasized. Device operates in moderate inversion and in saturation.

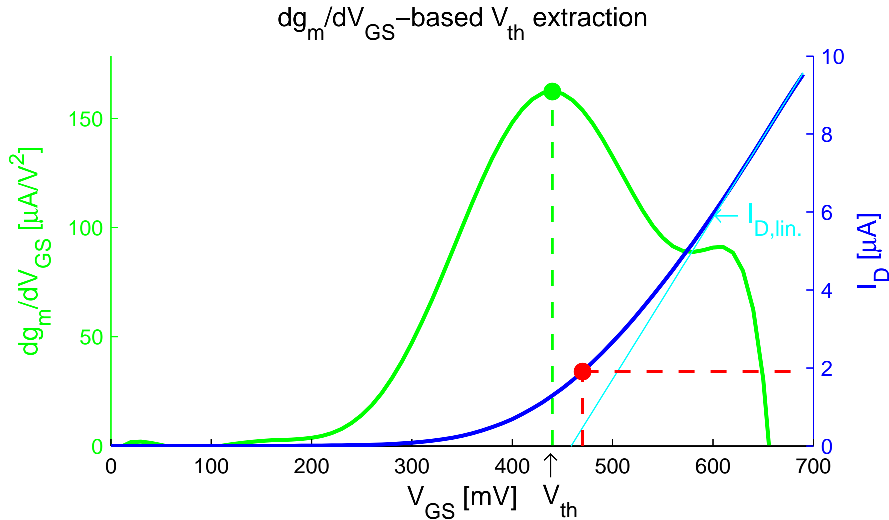


Figure 5.5: Threshold-voltage extraction based on the $\mathrm{d}g_m/\mathrm{d}V_{GS}$ methodology. Estimated $V_{th} \approx 440$mV is consistent with the value obtained by linear extrapolation of I_D, $V_{th,lin} \approx 459$mV.

Parameter	Symbol	Value
Number of traces	N_{traces}	996
Measurement period	T	4ms
Sampling frequency	f_s	2.5MHz
Time step	$\mathrm{d}t$	0.4µs
Number of samples	N_s	10k
Frequency resolution	$\mathrm{d}f$	250Hz
Maximal frequency	f_{max}	1.25MHz

Table 5.1: Values of the measurement parameters. T and f_s also defined the $\mathrm{d}f$ and f_{max} of our spectral analysis.

which is consistent with $V_{th,lin} \approx 459\text{mV}$ obtained by a simple linear extrapolation. No attention should be paid to the little hump around 600mV. When such humps are physical, they are usually attributed to second channel formation in multiple gate devices [39, 40]. In our case, it is a numerical consequence of the polynomial regression, and this non physical section of the curve should be ignored.

Operation region Computing

$$\frac{g_m}{I_D} \approx 12\text{V}^{-1},$$

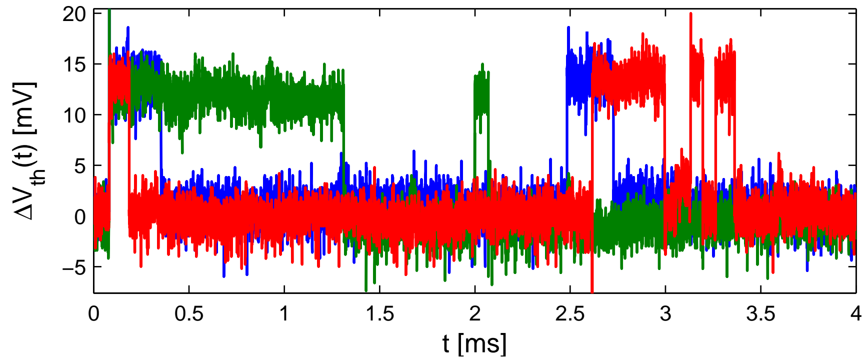
we convinced ourself that the MOSFET is in *moderate* inversion. Comparing gate overdrive

$$V_{OV} = V_{GS} - V_{th} \approx 30\text{mV}, \tag{5.2}$$

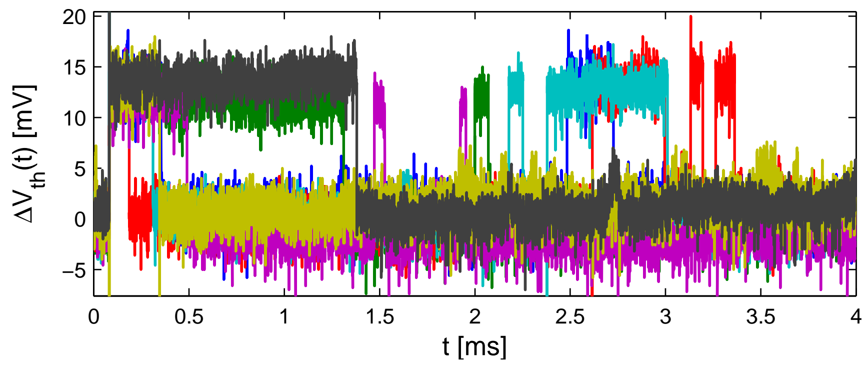
with $V_{DS} = 50\text{mV}$ also confirms that the device operates at the boundary between saturation and linear regions.

5.2 Time-domain analysis

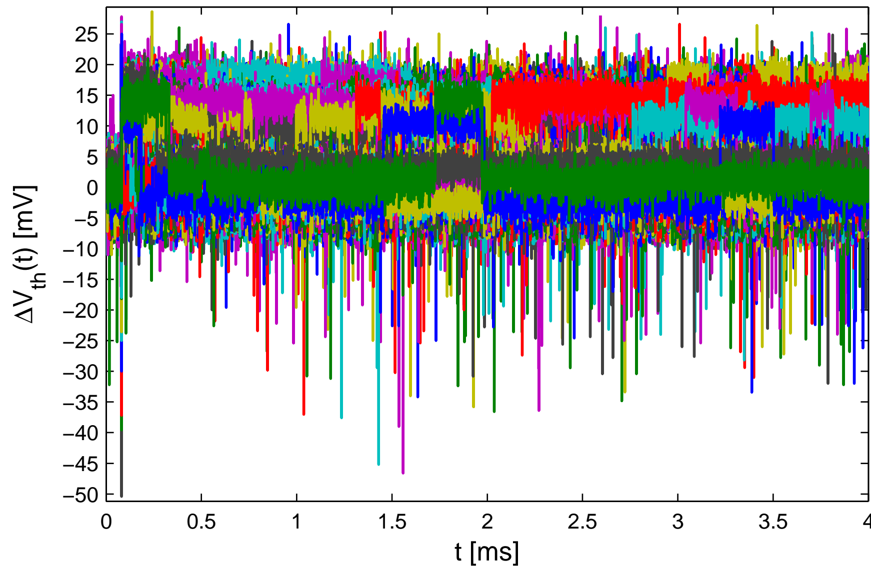
Measurement parameters were selected by Marko Simicic and are summarized in Table 5.1. Their relevance regarding the particular RTN observed in that golden device will be progressively discussed. Originally, 999 traces were recorded, but we eliminated three of them that have an erroneous DC level. 3, 7 and all the 996 are successively depicted in Fig. 5.6. For the purposes of illustration, our data processing will frequently be illustrated with a few representative traces, e.g. 3 or 7, and we will keep using the same colors for the first seven traces for clarity.



(a) 3 traces.



(b) 7 traces.



(c) 996 traces.

Figure 5.6: Measured analog traces ®.

Standard deviation σ of the background is computed and the two Gaussians are identified. The normality of the background noise is tested and approved. Different methods to extract a pure RTN from the analog traces are proposed. The analysis ends up with the estimations of the mean capture and emission time constants. The obtained values are compared with the one obtained from *out C* of Smart array.

5.2.1 Experimental pdf: histograms

A sum of Gaussians

Construction from traces The first-order probability density function (pdf) of the stochastic process $\Delta V_{th}(t)$

$$f_{\Delta V_{th}(t)}\left(\Delta V_{th}(t)\right)$$

does not depend on the time t if the noise is assumed to be *stationary* [3]:

$$f_{\Delta V_{th}}\left(\Delta V_{th}\right).$$

Consequently, an estimation of the pdf is obtained from the experimental measurement by constructing, for every trace, one histogram counting the number of occurrences of $\Delta V_{th}(t)$ for every *bin*. Almost miraculously, the hypothesis of stationarity eliminates any notion of time. Every data points

$$\left(t_i, \Delta V_{th}(t_i)\right)$$

equally contributes to the experimental pdf: its own time t_i is forgotten, and the counter of the bin in which $\Delta V_{th}(t_i)$ lies is incremented. Such construction is illustrated in Fig. 5.7 for the three first traces. Generally speaking, there is not precise rule for selecting the width of the intervals [41]. Here, the very large number of samples allows us to use the smallest relevant bin width corresponding to the oscilloscope resolution:

$$\mathrm{d}V_{th} = V_{LSB} = 200\mu\mathrm{V}.$$

All the individual histograms, each associated with one trace, can be collected together into one single and global histogram, which is, hopefully, an accurate estimation of the noise pdf. As seen in Fig. 5.8 the latter consists of two Gaussians centered around two values that will be related to the two levels of the RTN.

[3] Here, the hypothesis of stationarity cannot be verified from the experimental traces: the time origin is distorted by the trigger of the first capture event at $t \approx 80\mu\mathrm{s}$ (see Fig. 5.6).

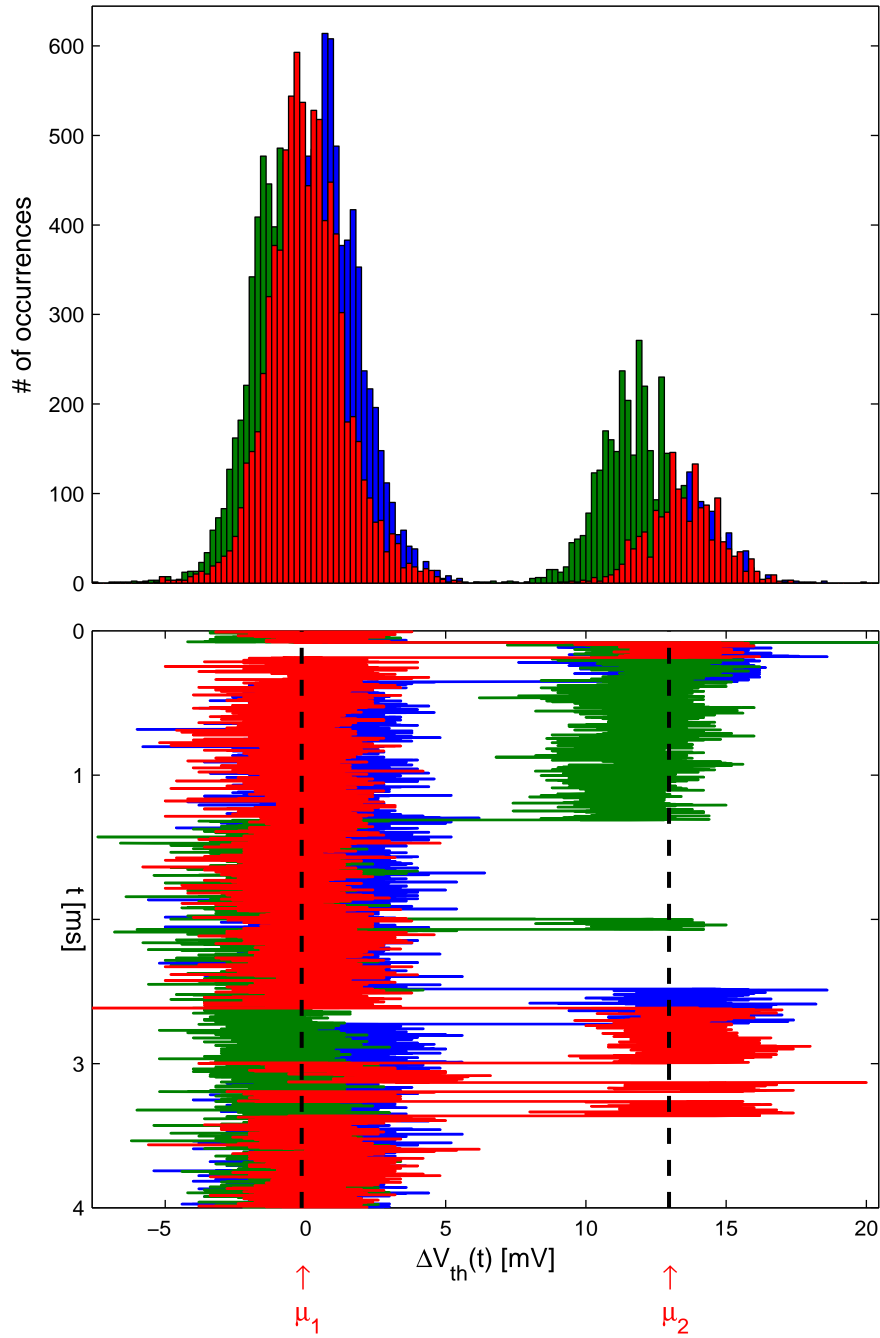


Figure 5.7: The three traces from Fig. 5.6a and their associated histograms. The two RTN levels μ_1 and μ_2, extracted further, are shown for clarity.

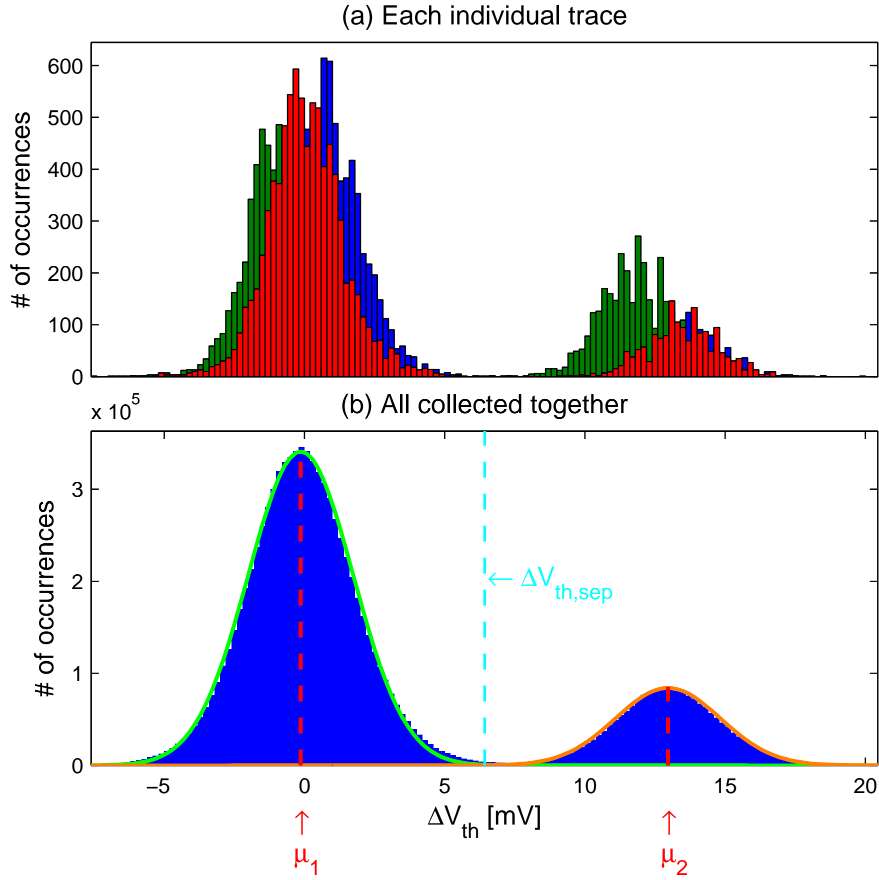


Figure 5.8: Experimental pdf $f_{\Delta V_{th}}(\Delta V_{th})$ of ΔV_{th}: histogram of the number of occurrences. Bin width is $V_{LSB} = 200\mu V$. (a) Three individual histograms from Fig. 5.7. (b) All the 996 traces collected together. The obtained pdf consists of two Gaussians (green and orange lines) centered around the two RTN levels μ_1 and μ_2. For this nice study case, Gaussians are well separate by a boundary $\Delta V_{th,sep}$.

Theoretical support Such an observed pdf is explained by decomposing the total noise trace into one RTN with some background noise on top of it:

$$\Delta V_{th}\left(t\right) = \Delta V_{th,RTN}\left(t\right) + \Delta V_{th,noise}\left(t\right). \tag{5.3}$$

The RTN pdf is composed of two Dirac delta whose amplitude depends on $\bar{\tau}_c$ and $\bar{\tau}_c$:

$$\begin{aligned} f_{\Delta V_{th,RTN}}\left(\Delta V_{th,RTN}\right) =& \frac{\bar{\tau}_c}{\bar{\tau}_c + \bar{\tau}_e}\,\delta\left(\Delta V_{th,RTN} - \mu_1\right) \\ &+ \frac{\bar{\tau}_e}{\bar{\tau}_c + \bar{\tau}_e}\,\delta\left(\Delta V_{th,RTN} - \mu_2\right). \end{aligned} \tag{5.4}$$

The background noise will be further proved to be normally distributed:

$$f_{\Delta V_{th,noise}}\left(\Delta V_{th,noise}\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{\Delta V_{th,noise}}{\sigma}\right)^2\right). \tag{5.5}$$

Equation (5.3) states that the total ΔV_{th} is the sum of two stochastic processes that we assume independent. Its pdf is then obtained as the convolution of the RTN pdf (5.4) with the background noise pdf (5.5):

$$\begin{aligned} f_{\Delta V_{th}}\left(\Delta V_{th}\right) &= \left(f_{\Delta V_{th,RTN}} * f_{\Delta V_{th,noise}}\right)\left(\Delta V_{th}\right) \\ &= \int_{-\infty}^{+\infty} f_{\Delta V_{th,RTN}}\left(V\right) f_{\Delta V_{th,noise}}\left(\Delta V_{th} - V\right) \mathrm{d}V \\ &= \frac{\bar{\tau}_c}{\bar{\tau}_c + \bar{\tau}_e}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{\Delta V_{th} - \mu_1}{\sigma}\right)^2\right) \\ &\quad + \frac{\bar{\tau}_e}{\bar{\tau}_c + \bar{\tau}_e}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{\Delta V_{th} - \mu_2}{\sigma}\right)^2\right), \end{aligned} \tag{5.6}$$

where the colors refer to Fig. 5.8. We emphasis that μ_1, μ_2 and σ are not known at this point, and that the Gaussianity still has to be tested.

Remark: link between histogram and pdf A histogram is essentially an experimental estimation of the exact pdf of the random process. All our histograms are expressed in number of occurrences rather than in normalized frequency. If $N_{\Delta V_{th}}$ denotes the total number of occurrences and N_i the counter of bin i centered around $\Delta V_{th,i}$, we have

$$\begin{aligned} \frac{N_i}{N_{\Delta V_{th}}} &\approx P\left(\Delta V_{th} \in \left[\Delta V_{th,i} - \mathrm{d}V/2, \Delta V_{th,i} + \mathrm{d}V/2\right]\right) \\ &= f_{\Delta V_{th}}\left(\Delta V_{th,i}\right)\mathrm{d}V, \end{aligned} \tag{5.7}$$

so that the true continuous pdf is related to the histogram with the relation

$$N_i \approx N_{\Delta V_{th}} \cdot f_{\Delta V_{th}} \left(\Delta V_{th,i}\right) \cdot \mathrm{d}V, \tag{5.8}$$

which ensures the normalization of the pdf:

$$\begin{aligned} \int_{-\infty}^{+\infty} f_{\Delta V_{th}} \left(\Delta V_{th}\right) \mathrm{d}V &\approx \sum_{i=-\infty}^{\infty} f_{\Delta V_{th}} \left(\Delta V_{th,i}\right) \mathrm{d}V \\ &\approx \sum_{i=-\infty}^{\infty} \frac{N_i}{N_{\Delta V_{th}}} = 1. \end{aligned} \tag{5.9}$$

Extraction of the RTN levels and σ

Based on the expected pdf given in (5.6), the two levels of the RTN μ_1 and μ_2 must be extracted as the positions of the two global maxima. Different methods were investigated.

Polynomial regression As the histogram of Fig. 5.8 is a discrete data set, these maxima cannot be found by differential calculus. We attempted a polynomial regression to obtain one continuous curve that might be differentiated analytically. We emphasis that the polynomial does not aim to smoothly fit the histogram everywhere: the hope is to reproduce the two large peaks and to find their locations by differentiation. A very high-degree polynomial must be used, e.g. degree 50. This leads to bad conditioning due to very closed $\Delta V_{th,i}$. The roots of the derivative are so numerous that they require additional processing to identify the two corresponding to the maxima. So that few has been gained with respect to the original histogram.

Threshold-based peak detection algorithm μ_1 and μ_2 can be extracted with great accuracy with few computational intensive method. We have implemented a *peak detection algorithm* enabling to identify the positions of the two large maxima of Fig. 5.8. The main difficulty is to properly define these two main peaks. From a mathematical point of view, the global maximum is unique, it is the highest point of the whole histogram which is located in position μ_1. We say that there is a *local peak* at index i if

$$N_{i-1} \leq N_i \geq N_{i+1}. \tag{5.10}$$

Build-in Matlab® function `findpeaks` searches for these peaks. The second maximum of interest, located in μ_2 is formally only one local peaks among others. What differentiates it, visually, from all the others? Firstly, the fact that it the highest point on a large neighbourhood. Some ridiculous local peak located in the "no man's land" between the two Gaussians might well fulfills such a condition. We therefore complete our definition with a notion

of *threshold* that a peak candidate must exceed. Threshold was here set to 100 occurrences.

```
nPeaks = 2;
threshold = 100;
d = 1;
% all the local peaks that are above threshold
[height_peaks, indices_peaks] = findpeaks(f_dVth,...
        'minPeakHeight',threshold);
% small local peaks are progressively eliminated
while ((length(indices_peaks) > nPeaks))
    d = d+1; % extending the tested neighbourhood
    if ((indices_peaks(1)-d) < 1)
        indices_peaks = indices_peaks(2:end);
        height_peaks = height_peaks(2:end);
    end
    if ((indices_peaks(end)+d) > nBins)
        indices_peaks = indices_peaks(1:end-1);
        height_peaks = height_peaks(1:end-1);
    end
    % remaining candidates
    isBigPeak = (height_peaks >= f_dVth(indices_peaks-d))...
              & (height_peaks >= f_dVth(indices_peaks+d));
    indices_peaks = indices_peaks(isBigPeak);
    height_peaks = height_peaks(isBigPeak);
    % two close identical peaks
    indices_same = find(~diff(val_peaks));
    % deleting the second one
    indices_peaks(indices_same+1) = [];
    height_peaks(indices_same+1) = [];
end
max1 = height_peaks(1);
max2 = height_peaks(2);
index_max1 = indices_peaks(1);
index_max2 = indices_peaks(2);
```

Code 5.1: Implementation of the peak detection algorithm.

The Matlab® implementation of the algorithm is given in Code 5.1. $\mu_1 = -0.100$mV et $\mu_2 = 12.964$mV are output.

This study case is considered as nice in the sense that the *random telegraph signal-to-noise ratio*, defined as

$$\frac{\Delta\mu}{\sigma} = \frac{\mu_2 - \mu_1}{\sigma} \tag{5.11}$$

is large [12]. σ is the RMS background noise level, i.e. the square root of the variance

$$\begin{aligned}\sigma^2 &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \left|\Delta V_{th,noise}\left(t\right)\right|^2 \mathrm{d}t \\ &\approx \frac{1}{T} \int_0^T \left|\Delta V_{th,noise}\left(t\right)\right|^2 \mathrm{d}t \\ &\approx \frac{1}{N_s} \sum_{n=1}^{N_s} \left|\Delta V_{th,noise}\left(t_n\right)\right|^2 .\end{aligned} \tag{5.12}$$

It is exactly related to the standard deviation σ of the Gaussian distribution (5.5). This is a consequence of the *variance-ergodicity* stating that the variance, defined as a ensemble average

$$\begin{aligned}\sigma^2 &= \mathbb{E}\left\{\left|\Delta V_{th,noise}\left(t\right)\right|^2\right\} \\ &= \int_{-\infty}^{+\infty} \left|\Delta V_{th,noise}\left(t\right)\right|^2 f_{\Delta V_{th,noise}}\left(\Delta V_{th,noise}\right) \mathrm{d}\left(\Delta V_{th,noise}\right)\end{aligned} \tag{5.13}$$

can be computed as the time average (5.12) [26].

We have extracted the pdf from the experimental realizations. Our two Gaussians are well separate. A clear boundary located in

$$\Delta V_{th,sep} \approx \frac{\mu_1 + \mu_2}{2} \tag{5.14}$$

split the histogram in two, as shown in Fig. (5.8). It has been checked *a posteriori* that

$$\mu_1 - 3\sigma \leq \Delta V_{th,sep} \leq \mu_2 + 3\sigma.$$

Hence, we can estimate σ based on a expectation, i.e. ensemble average, as (5.13) for each of two Gaussians:

$$\begin{aligned}\sigma_1^2 &\triangleq \int_{-\infty}^{\Delta V_{th,sep}} \left|\Delta V_{th} - \mu_1\right|^2 f_{\Delta V_{th}}\left(\Delta V_{th}\right) \mathrm{d}\left(\Delta V_{th}\right) \\ &\approx \sum_{i=-\infty}^{index_{sep}} \left|\Delta V_{th,i} - \mu_1\right|^2 f_{\Delta V_{th}}\left(\Delta V_{th,i}\right) \cdot \mathrm{d}V \\ &= \sum_{i=-\infty}^{index_{sep}} \left|\Delta V_{th,i} - \mu_1\right|^2 \frac{N_i}{N_{\Delta V_{th}}},\end{aligned} \tag{5.15}$$

$$\begin{aligned}\sigma_2^2 &\triangleq \int_{\Delta V_{th,sep}}^{+\infty} \left|\Delta V_{th} - \mu_2\right|^2 f_{\Delta V_{th}}\left(\Delta V_{th}\right) \mathrm{d}\left(\Delta V_{th}\right) \\ &\approx \sum_{i=index_{sep}}^{+\infty} \left|\Delta V_{th,i} - \mu_2\right|^2 f_{\Delta V_{th}}\left(\Delta V_{th,i}\right) \cdot \mathrm{d}V \\ &= \sum_{i=index_{sep}}^{+\infty} \left|\Delta V_{th,i} - \mu_2\right|^2 \frac{N_i}{N_{\Delta V_{th}}}.\end{aligned} \tag{5.16}$$

μ_1	μ_2	σ
-0.100mV	12.964mV	1.967mV

Table 5.2: Parameter values obtained from the peak detection algorithm.

$\sigma_1 = 1.965$mV and $\sigma_2 = 1.969$mV are found. The great matching between the two is encouraging, as the two Gaussians come from the same physical process that is background noise. RTN only influences mean location and peak amplitude.

Values of the means μ_1 and μ_2 may be refined following a similar approach. As a summary, the peak detection algorithm approach yields values of Table 5.2. These will be cross-checked with a second method described below. In addition to that, we will be able to evaluate the time average (5.12) as soon as the background noise has been separated from the RTN, which constitutes another way to estimate σ.

Nonlinear fitting A very computational intensive approach is directly based on (5.6) that we want to fit to the histogram in the least square sense (taking into account the normalization factors in (5.7) linking the histogram to the pdf). The optimization problem is written as

$$\min_{(\bar{\tau}_c, \bar{\tau}_e, \mu_1, \mu_2, \sigma)} SSE, \tag{5.17}$$

where SSE is the *sum of squared errors of predictions*, that is the sum of the squares of residuals:

$$\begin{aligned} SSE &= \sum_i \left(\frac{N_i}{N_{\Delta V_{th}} \cdot \mathrm{d}V} - f_{\Delta V_{th}} \left(\Delta V_{th,i} \right) \right)^2 \\ &= \sum_i \left(\frac{N_i}{N_{\Delta V_{th}} \cdot \mathrm{d}V} - \frac{\bar{\tau}_c}{\bar{\tau}_c + \bar{\tau}_e} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{\Delta V_{th,i} - \mu_1}{\sigma} \right)^2 \right) \right. \\ &\qquad \left. - \frac{\bar{\tau}_e}{\bar{\tau}_c + \bar{\tau}_e} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{\Delta V_{th,i} - \mu_2}{\sigma} \right)^2 \right) \right)^2 . \end{aligned} \tag{5.18}$$

Optimization parameters have been coloured to highlight the strongly non-linear character of the problem. Minimization is performed in Matlab® using the `fminsearch` build-in function. Optimal values of variables $(\bar{\tau}_c, \bar{\tau}_e, \mu_1, \mu_2, \sigma)$ are given in Table 5.3. The high quality of the fit is appreciated in Fig. 5.9.

$\bar{\tau}_c$	$\bar{\tau}_e$	μ_1	μ_2	σ
0.585ms	0.144ms	−0.130mV	12.963mV	1.869mV

Table 5.3: Optimal parameters of the nonlinear regression defined by (5.17) and (5.18).

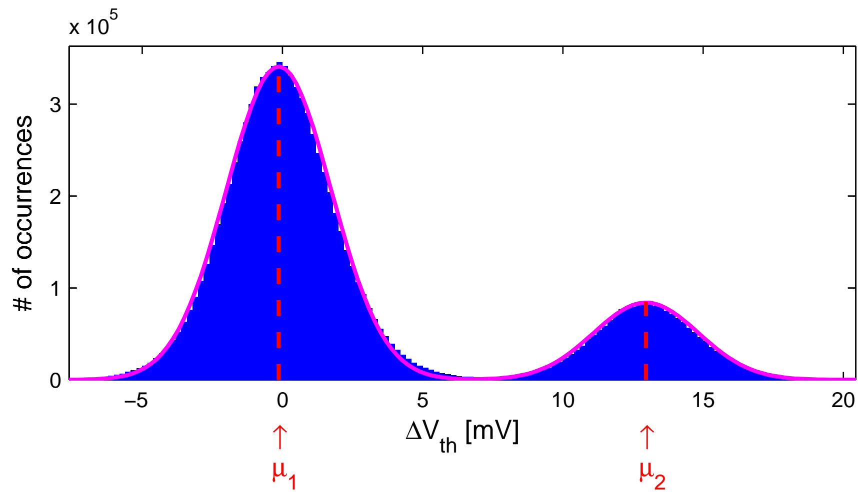


Figure 5.9: Nonlinear fitting, in the least squares sense, of (5.6) to the whole histogram.

This method provides an estimation of σ without any additional computation. Inspection of Fig.5.6a or Fig. 5.6b led us to think that the values for $\bar{\tau}_c$ and $\bar{\tau}_e$ are underestimated. This does not question the quality of the fit. It rather raises the question of whether the measured traces are statistically representative or not. We saw in the previous chapter that underestimated mean time constants mostly results from the effect of a too short time window. This important question will be addressed in details later.

We compare the values for μ_1, μ_2 and σ of the two approaches, Table 5.2 for the peak detection algorithm based extraction, Table 5.3 for the nonlinear multi-Gaussian fit. The means are almost perfectly similar. Difference of $\sim 5\%$ is observed for the standard deviation. Time average (5.12) will allow us to decide for one or the other.

Normality test

There are many tests for normality. Two excellent references are [42, 43]. The most elementary and qualitative one is a graphical representation of the experimental distribution, such as the histograms of Fig. 5.8 or Fig. 5.17. The distribution of some experimentally-derived data set must also be assessed quantitatively. The most advanced tests are beyond the scope of this work and we content ourselves with a *Q-Q plot*, where Q stands for *quantile*, as well as the closely related 68-95-99.7 rule.

Q-Q plot If a random variable X is normally distributed with mean μ and standard deviation σ, the cumulative distribution function (CDF) is

$$\begin{aligned} F_X(x) = P(X \leq x) &= \int_{-\infty}^{x} f_X(x)\,\mathrm{d}x \\ &= \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right). \end{aligned} \tag{5.19}$$

Theoretical p-quantile is then obtained by inverting relation (5.19):

$$z_p \triangleq \frac{x_p - \mu}{\sigma} = \sqrt{2} \cdot \operatorname{erf}^{-1}(2p-1)\,. \tag{5.20}$$

For instance, $p = 2.5\%$ gives $z_{2.5\%} \approx -1.96$, or $x = \mu - 1.96 \cdot \sigma$.

From the histogram of Fig. 5.8, we compute two experimental CDFs (one for each Gaussian) as

$$\begin{aligned} CDF_1\left(\Delta V_{th,j}\right) &= \sum_{i=-\infty}^{j} \frac{N_i}{N_{\Delta V_{th}}} \\ CDF_2\left(\Delta V_{th,j}\right) &= \sum_{i=index_{sep}}^{j} \frac{N_i}{N_{\Delta V_{th}}} \end{aligned} \tag{5.21}$$

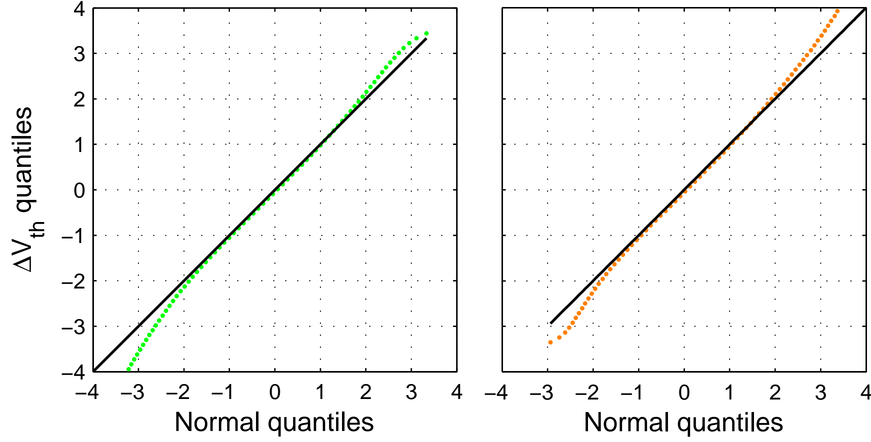


Figure 5.10: Q-Q plot to assess normality of background noise. Colors green and orange refers to the two Gaussians of Fig. 5.8.

For comparison purposes, theoretical quantiles (5.20) are evaluated for probabilities

$$p_j = CDF\left(\Delta V_{th,j}\right), \tag{5.22}$$

and the corresponding experimental quantiles are

$$q_{p_j} = \frac{\Delta V_{th,j} - \mu}{\sigma}. \tag{5.23}$$

The Q-Q plot is the parametric curve

$$p_j \in [0,1] \to \left(z_{p_j}, q_{p_j}\right). \tag{5.24}$$

Q-Q plots for the two Gaussian-like data sets are depicted in Fig. 5.10. Parameters μ_1, μ_2 and σ from Table 5.3 were used. The normality is excellent in the $[-2\sigma, +2\sigma]$. Slight deviations from the straight line

$$q_{p_j} = z_{p_j} \tag{5.25}$$

are observed at $\pm 3\sigma$ for both Gaussians. The tails are distorted because of some outliers (see, for instance, a few glitches in the traces in Fig. 5.7). Moreover, CDF_1 gets inaccurate starting from $\mu_1 + 3 \cdot \sigma$ because of the interference with the second Gaussian around $\Delta V_{th,sep}$, and conversely.

Verification of the 68-95-99.7 rule While the Q-Q plots are one-tailed tests, the symmetry of the Gaussian-like distribution tends to encourage a two-tailed test such as the $\mu \pm N \cdot \sigma$ rule derived from

$$\begin{aligned} P\left(\mu - N \cdot \sigma \leq X \leq \mu + N \cdot \sigma\right) &= F_X\left(\mu + N \cdot \sigma\right) - F_X\left(\mu - N \cdot \sigma\right) \\ &= \operatorname{erf}\left(\frac{N}{\sqrt{2}}\right). \end{aligned} \tag{5.26}$$

Range	Normal	Gaussian 1	Gaussian 2
$\mu \pm \sigma$	68.3%	68.5%	68.5%
$\mu \pm 2\sigma$	95.4%	95.0%	93.6%
$\mu \pm 3\sigma$	99.7%	99.3%	99.2%

Table 5.4: Coverage probability test.

If our experimental pdf is well the sum of two Gaussians, their respective coverage probability must be found to matches the very famous coverage probability 68% for σ, 95% for 2σ and 99.7% for 3σ. Test is summarized in Table 5.4. All values are satisfying, excepted the 93.6% of Gaussian 2 for 2σ. This may be explained by outliers that inflate the deep tails beyond 3σ (see in Fig. 5.10 the deviation of the Q-Q plot between 3σ and 4σ).

Conclusion about the background noise We reach the expected conclusion that the background noise $\Delta V_{th,noise}$ is a Gaussian stochastic process. This property characterizes its *first-order* probability density function (pdf). It does *not* tell us anything about the *correlation function* (2.7) or, equivalently, the *power spectral density* (PSD) (3.24)

$$R_{\Delta V_{th,noise}}(t) \overset{\mathcal{F}}{\longleftrightarrow} S_{\Delta V_{th,noise}}(f) . \tag{5.27}$$

related to the *second-order* probability density function. This provides an essential motivation for spectral analysis beyond the study of the RTN itself, to confidently relate the background noise to the well known flicker ($1/f$) noise.

5.2.2 Pure RTN extraction

We get back to traces of Fig. 5.6. We wish to separate the pure two-level RTN and the Gaussian noise in order to study them independently. Different methods are investigated through a few representative traces, including pathological ones. Further, the extraction is statistically verified by discussing the obtained mean time constants.

Boundary-based RTN extraction

A direct and naive way to extract the RTN from the analog traces is to code the following equation:

$$\Delta V_{th,RTN}(t) = \begin{cases} \mu_1 & \text{if } \Delta V_{th}(t) < \Delta V_{th,sep} \\ \mu_2 & \text{if } \Delta V_{th}(t) \geq \Delta V_{th,sep} \end{cases} \tag{5.28}$$

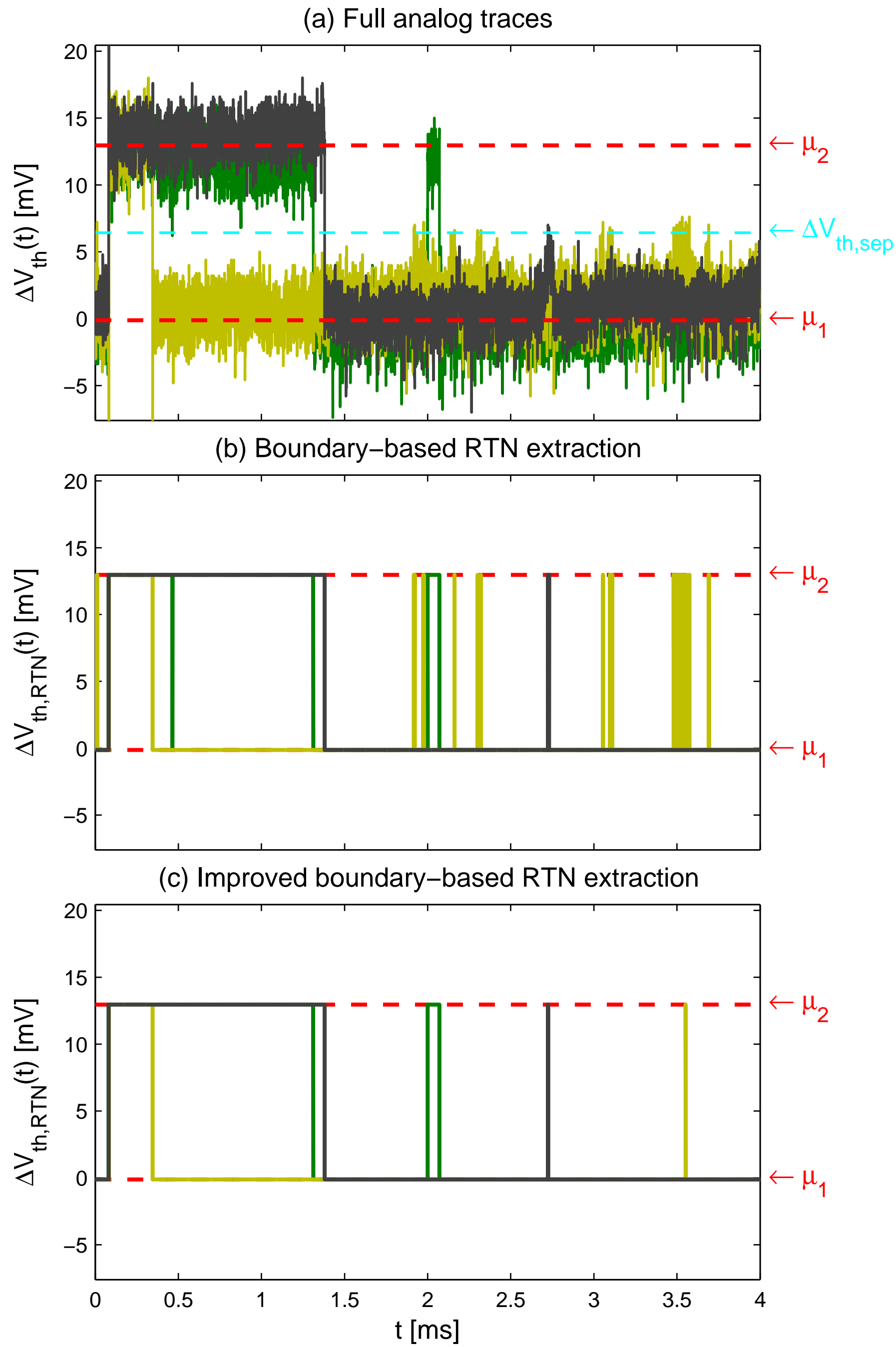


Figure 5.11: Boundary $\Delta V_{th,sep}$-based RTN extraction (b) for three representative analog traces (a). This extraction method is much to sensitive to the background noise level. (c)

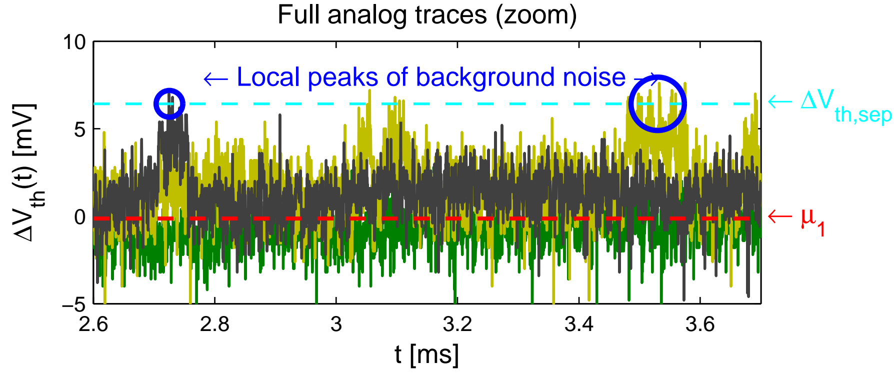


Figure 5.12: Illustration of the limitation of basic boundary-based RTN extraction. Local peaks of background noise makes $\Delta V_{th}(t)$ exceed the splitting level $\Delta V_{th,sep}$ and cause erroneous transitions from low to high state.

This method is illustrated in Fig. 5.11b. A quick glance convinces us of its terrible lack of robustness. Loads of erroneous edges are detected, due to the background nice that may be locally high. Zoom performed in Fig. 5.12 makes it even more clear.

Charge trapping and detrapping in the background noise By the way, one may questions the existence of one very small second trap with $\Delta\mu \approx 5$mV that would be charged from $t \approx 2.71$ms to $t \approx 2.75$ms for the black trace, or from $t \approx 3.48$ms to $t \approx 3.58$ms for the yellow one. Advanced methods as the hidden Markov model would allow to decide [12, 44]. In this work, negligible-impact traps are treated as $1/f$ noise. Thus, the presence of this trapping-like behaviour in the background noise is not a problem *a priori*.

Reducing the number of erroneous transitions An improvement of this basic method can be made by refusing transitions arising from noise peak whose duration is only one time step. This is acceptable is

$$\mathrm{d}t = 0.4\mu\mathrm{s} \ll \min(\bar{\tau}_c, \bar{\tau}_e) = \bar{\tau}_e,$$

with $\bar{\tau}_e \geq 100\mu$s at least, so that no true edge are rejected with this correction. Fig. 5.11c shows that the extracted RTNs are still unsatisfying as parasitic edges remain for the black and the yellow traces. We might want to reject also transitions induced by noise peaks whose duration is longer than dt. This approach is too arbitrary since the risk of wrongly ignoring very short time constants increases.

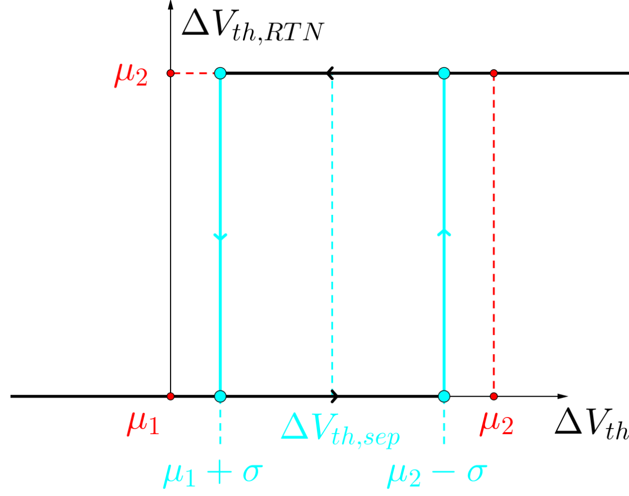


Figure 5.13: Schmitt trigger operation. $\Delta V_{th}(t)$ is the analog input signal. Output is the two-state $\Delta V_{th,RTN}(t)$.

Schmitt trigger based RTN extraction

A robust algorithm to extract the RTN is based on a *Schmitt trigger*, which has been integrated in our Matlab® software tool. The operation principle is better understood by thinking of the Schmitt trigger as a circuit whose input is the full analog trace $\Delta V_{th}(t)$ and output is the sought $\Delta V_{th,RTN}(t)$. The transfer function is depicted in Fig. 5.13. It is naturally nonlinear: there is an hysteresis, the output state depends on the history of the input signal; this was not the case of the boundary-based extraction dictated by (5.28).

It is not sufficient that the trace crosses $\Delta V_{th,sep}$ to assume a capture or emission event. That single boundary $\Delta V_{th,sep}$ is replaced by two thresholds placed at $\mu_1 + \sigma$ and $\mu_2 - \sigma$. For a low-to-high transition to be validated, i.e. for $\Delta V_{th,RTN}(t)$ to switch from μ_1 to μ_2, it is necessary that the trace crosses the $\mu_2 - \sigma$. If the transition is confirmed by the crossing of the $\mu_2 - \sigma$, the sharp edge of the RTN may actually be placed at the time when $\mu_1 + \sigma$ was crossed. Then, RTN will remain in high-state unless the trace crosses $\mu_1 + \sigma$ over again.

The happy consequence is that the probability that the background takes a high enough value to be responsible for a parasitic transition gets now almost zero for this study case. This can be proved using the fact that μ_1 and μ_2 levels are as far as more than 6σ, and then plugging $N = 5$ into (5.26).

Fig. 5.14 shows that the RTN is properly extracted even for pathological traces presenting high peaks of background noise such as the black and the yellow. The extraction for some additional traces is seen in Fig. 5.15, and we check again the consistency of the Schmitt trigger method. Checking the

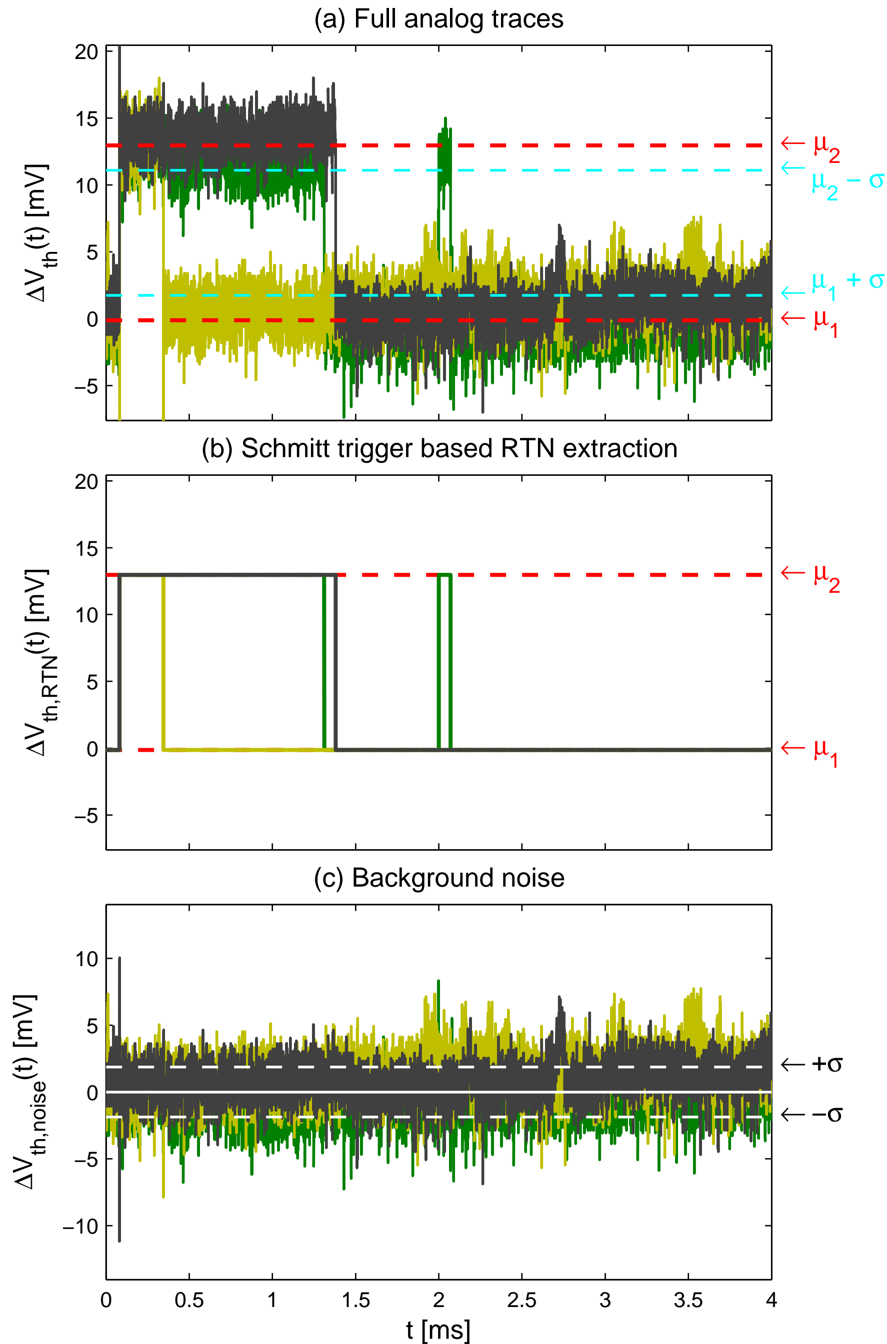


Figure 5.14: Schmitt trigger based RTN extraction for three representative traces: separation of **(a)** the measured noise traces into **(b)** a dominant pure RTN and, **(c)** the additional Gaussian background noise.

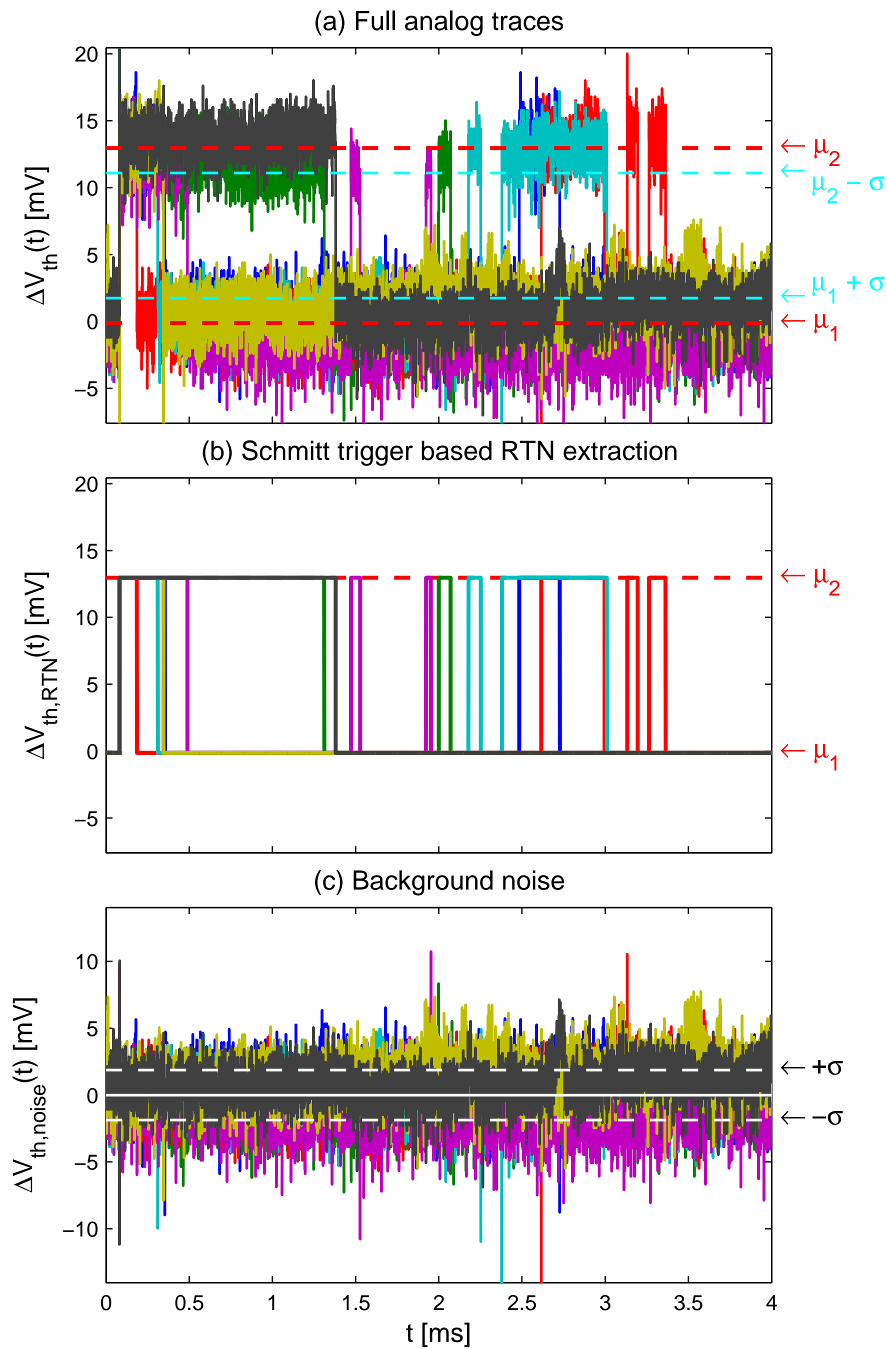


Figure 5.15: Schmitt trigger based RTN extraction for seven traces.

traces "with the eye" quickly gets a tedious task. The next section goes one step further in the validation by exploiting statistically all the 996 available traces, that we do not wish the inspect individually with the eye.

Background noise Once the pure RTN has been extracted, the additional noise is isolated by immediate subtraction from the whole trace:

$$\Delta V_{th,noise}(t) = \Delta V_{th}(t) - \Delta V_{th,RTN}(t). \tag{5.29}$$

The realizations of this noise plotted in Fig. 5.14c or Fig. 5.15c looks like our usual perception of a zero-mean Gaussian noise.

- Assuming *mean-ergodicity*, we verify the zero-mean property with a time average over all the traces:

$$\begin{aligned}\mathbb{E}\left\{\Delta V_{th,noise}(t)\right\} &= \eta_{noise} \\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \Delta V_{th,noise}(t)\,\mathrm{d}t \\ &\approx \frac{1}{N_{traces}}\sum_{n=1}^{N_{traces}}\left(\frac{1}{N_s}\sum_{i=1}^{N_s}\Delta V_{th,noise}^{(n)}(t_i)\right) = 131\mu\mathrm{V}.\end{aligned} \tag{5.30}$$

  Given the resolution $V_{LSB} = 200\mu\mathrm{V}$ of the oscilloscope, the estimated mean is small enough to be considered as zero.

- We are now able to evaluate the *average noise power* with the time integration (5.12) for every trace. Carrying out an ensemble average over all traces, we compute

$$\sigma^2 \approx \frac{1}{N_{traces}}\sum_{n=1}^{N_{traces}}\left(\frac{1}{N_s}\sum_{i=1}^{N_s}\left|\Delta V_{th,noise}^{(n)}(t_i)\right|^2\right) = 3.936\mu\mathrm{V}^2, \tag{5.31}$$

  from which we derived the noise RMS value:

$$\sigma = \sqrt{\sigma^2} = 1.984\mathrm{mV}.$$

  This value is very consistent with the previous σ estimations of Table 5.2 (ensemble average after the peak detection algorithm) and Table 5.3 (multi-Gaussian fit), and rather tends to favour the first of two approaches.

5.2.3 Analysis of time constants

Mean time-constant estimation

A large number of time-constant realizations

$$\tau_{e,1}, \tau_{e,2}, \ldots, \tau_{e,N_{\tau_e}}$$

$$\tau_{c,1}, \tau_{c,2}, \ldots, \tau_{c,N_{\tau_c}}$$

Data	This work		
	out A		*out B*
Method	Boundary-based	Schmitt trigger	Digitized RTN traces
$\bar{\tau}_e$ [ms]	0.193	0.326	0.343
$\bar{\tau}_c$ [ms]	0.547	1.067	1.205

	Simicic *et al.* [32]	
	out C	*out A*
		Censored data
$\bar{\tau}_e$ [ms]	0.314	idem
$\bar{\tau}_c$ [ms]	3.754	idem

Table 5.5: Estimations of the mean time constants. Results of our RTN extraction methods are compared with the values provided by *out C* of the Smart array, whose accuracy has been validated by Simicic *et al.* with an advanced post-processing of the analog traces of *out A*.

are obtained from the pure RTN traces of Fig. 5.15b, by counting the number of time steps between two successive edges. Using all the 996 traces, we have a total of $N_{\tau_e} = 2233$ emission time constants and $N_{\tau_c} = 1364$ capture time constants.

The mean time constants are naturally estimated as sample means:

$$\begin{aligned} \hat{\bar{\tau}}_e &= \frac{1}{N_{\tau_e}} \left(\tau_{e,1} + \tau_{e,2} + \ldots + \tau_{e,N_{\tau_e}} \right) \\ \hat{\bar{\tau}}_c &= \frac{1}{N_{\tau_c}} \left(\tau_{c,1} + \tau_{c,2} + \ldots + \tau_{c,N_{\tau_c}} \right) . \end{aligned} \tag{5.32}$$

These empirical means allows us to discuss our methods for RTN extraction. The values of $\bar{\tau}_e$ and $\bar{\tau}_c$ available from *out C* of the Smart array constitute a very useful reference. For now, we assume that they are exact. Table 5.5 compares all the found values.

- The boundary-based RTN extraction led to an estimation of very poor quality due to the erroneous edges. This method is thus rejected for good.

- We are very pleased to note that our $\hat{\bar{\tau}}_e$ is consistent with the value from the Smart array. Let us quantify the relative error:

$$\varepsilon_{\bar{\tau}_e} = \frac{\bar{\tau}_e - \hat{\bar{\tau}}_e}{\bar{\tau}_e} = 3.8\%. \tag{5.33}$$

Relation for the relative standard error of the mean from previous chapter

$$\varepsilon_{\bar{\tau}} = \frac{1}{\sqrt{N_\tau}}, \tag{5.34}$$

yields 2.1% with $N_\tau = 2233$. The theoretical expression seems to predict very well the order of magnitude of the error for a given number of realizations. Equation (5.34) is also there to recall that a reduction of the estimation error by a factor of ten would one hundred times more time constants, hence roughly one hundred times more traces.

- However, it appears that $\bar{\tau}_c$ is underestimated by a factor of almost four:

$$\frac{\widehat{\bar{\tau}}_c}{\bar{\tau}_c} = 0.284. \tag{5.35}$$

This is a dramatic consequence of a too short time window. We have indeed

$$\frac{T}{\bar{\tau}_c} = \frac{4\text{ms}}{3.754\text{ms}} = 1.07. \tag{5.36}$$

The effect of T on the mean time constant has been quantified in the previous chapter:

$$\begin{aligned} \widehat{\bar{\tau}} \approx \mathbb{E}_T\left\{\mathcal{T}\right\} &\equiv \int_0^T \tau\, f_{\mathcal{T}}\left(\tau\right)\,\mathrm{d}\tau \\ &= \bar{\tau}\left(1 - \left(\frac{T}{\bar{\tau}} + 1\right)\exp\left(\frac{T}{\bar{\tau}}\right)\right). \end{aligned} \tag{5.37}$$

Substituting value (5.36) in (5.37) gives

$$\widehat{\bar{\tau}} \approx 0.288 \cdot \bar{\tau}. \tag{5.38}$$

The great matching between (5.35) and (5.38) reinforces the power of (5.37) to predict the impact of the measurement period on the mean time constant estimator.

Beyond the equations, the inaccurate estimation of $\bar{\tau}_c$ is more simply explained by inspecting the traces, e.g. the one shown in Fig. 5.16. One realization τ_e is obviously recorded. Then, the RTN remains is the low-state for a long time. A so long time that no new capture event occurs before the end of the time window. The value τ_c is never recorded in this case, and all the information from $t \approx 1.38$ms is thrown in the trash.

We can convinced ourself that it was not possible to do better with the estimator $\widehat{\tau}_c$ of (5.32) without changing the measurement period by exploiting the digital RTN traces directly provided by *out B* of the Smart array. Values reported in Table 5.5 validate, at least statistically, the extraction of the

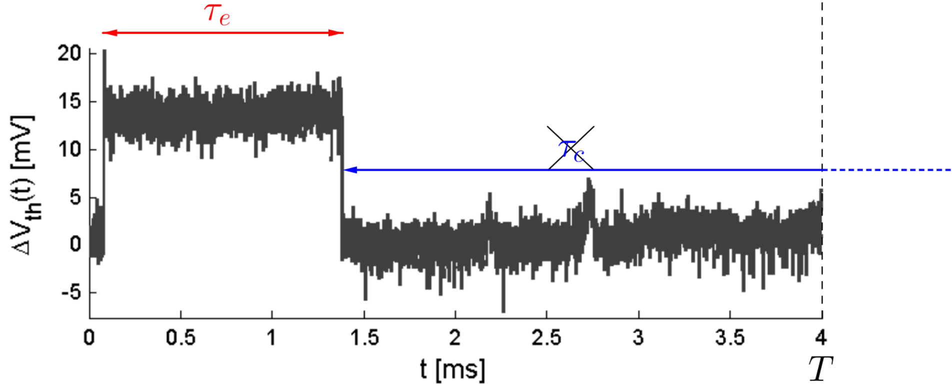


Figure 5.16: The measurement period T is too short with respect to $\bar{\tau}_c$ so that loads of long capture times τ_c are not recorded.

RTN with the Schmitt trigger method. The only culprit for the bad $\bar{\tau}_c$ is the short time window.

Would it be possible to exploit the censored observations of τ_c, for instance the low-state starting at $t \approx 1.38$ms in Fig. 5.16? The answer is yes, using Cox's *likelihood function for censored data* [45–48]. Simicic *et al.* applied this method together with a Kaplan-Meier estimator on analog traces of *out A* to demonstrate the accuracy of the mean time constants output by the SoC at *out C*. Such an approach is undoubtedly a statistical wonder but is beyond our skills and the scope of this text.

Time-constant pdf

Using the time-constant realizations, histograms are constructed as depicted in Fig. 5.17. These are the experimental pdfs of $\mathcal{T}_e$ and $\mathcal{T}_c$. Both representations, in linear and logarithm scale, allow to recognize the theoretical exponential distribution:

$$f_{\mathcal{T}}(\tau) = \frac{1}{\bar{\tau}} \exp\left(-\frac{\tau}{\bar{\tau}}\right). \tag{5.39}$$

- Regarding $\mathcal{T}_e$, the histogram is really smooth. Almost no deviation from (5.39), into which parameter τ_e from Table 5.5 is plugged [4], is observed.

- Histogram of $\mathcal{T}_c$ is less smooth, and especially its tail is cut off slightly before $\tau_c = T = 4$ms, related to our previous discussion about the time window. The impact on the histogram is interesting. The relative frequency of recorded short time constants is too large. With more

[4]Either our value $\widehat{\bar{\tau}}_e$ or the one output from the Smart array can be used, as they are very closed.

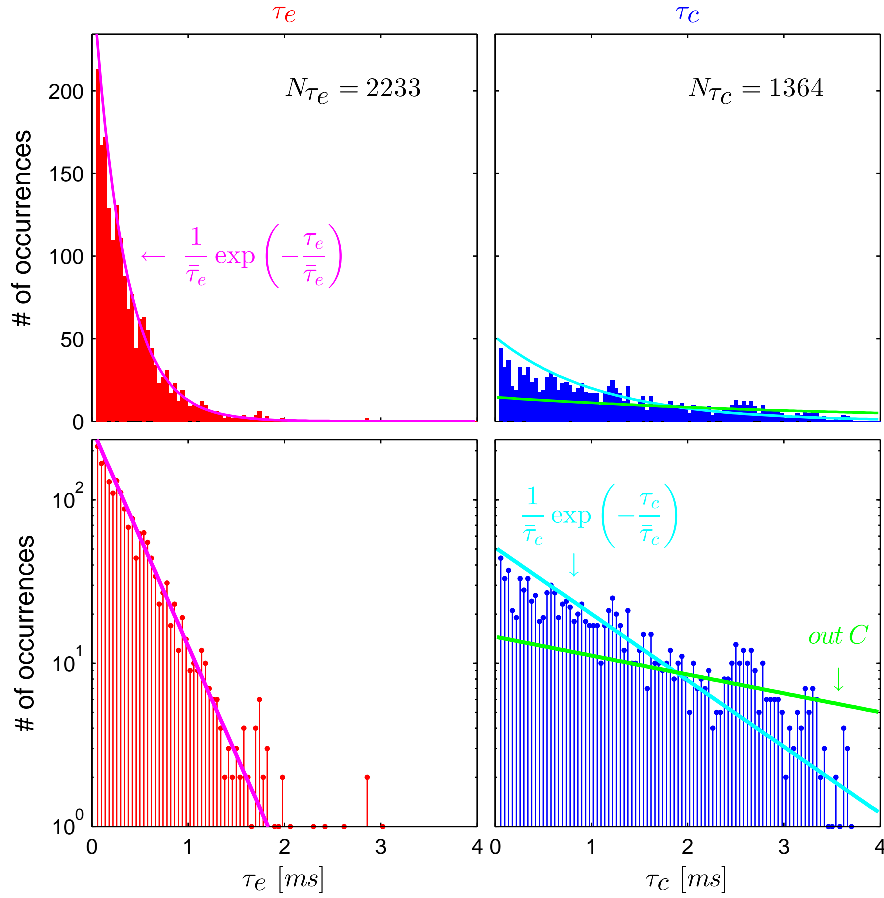


Figure 5.17: Experimental distributions of $\mathcal{T}_e$ and $\mathcal{T}_c$ based on 996 traces, in both linear (top) and logarithm scale (bottom). Bin width is $100 \cdot dt = 40\mu s$. Theoretical pdf (5.39) with our estimations of $\bar{\tau}_e$ and $\bar{\tau}_c$ (magenta and cyan lines) are superimposed. The right exponential distribution is recovered for $\mathcal{T}_e$. Experimental pdf of $\mathcal{T}_c$ is also exponential-like, but the tail is cut off according to the discussion related to Figure 5.16. Green curve is (5.39) but with $\bar{\tau}_c$ from the Smart array.

than 1000 samples, a lot of time constants as long as $\bar{\tau}_e = 3.754$ should occur, as

$$\begin{aligned} \bar{F}_{\mathcal{T}}(\bar{\tau}) &= P(\mathcal{T} \geq \bar{\tau}) \\ &= 1 - P(\mathcal{T} \leq \bar{\tau}) \\ &= 1 - \int_0^{\bar{\tau}} f_{\mathcal{T}}(\tau)\, \mathrm{d}\tau \\ &= e^{-1} \approx 37\%, \end{aligned} \tag{5.40}$$

but this calculation is annihilated by the limited measurement period that cuts the tail and inflate the pdf for low τ values. If the window was infinite, we would recover the exact pdf (5.39) into with $\bar{\tau}_e = 3.754$ plugged into it, i.e. the green line in Fig. 5.17 (right-bottom).

The time-domain analysis is complete. RTN and Gaussian background noise are now available separately. All the parameters, μ_1, μ_2, σ, $\bar{\tau}_e$ and $\bar{\tau}_c$ are accurately known.

5.3 Spectral analysis

This section aims to estimate the power spectral density of the stochastic process $\Delta(t)$. We are also interested in the PSDs of the pure two-level RTN $\Delta V_{th,RTN}(t)$ and of the background noise $\Delta V_{th,noise}(t)$. Questions guiding our spectral analysis are the following:

- For this golden device exhibiting a high RTS-to-noise ratio, we expect that the total PSD will be dominated by the RTN, at least from a few Hz up to a certain high frequency. Does our analysis confirm this intuition?
- We know the exact analytical expression of the RTN PSD which is a Lorentzian:

$$S_{\mathcal{L}}(f) = S_{\Delta V_{th,RTN}}(f) = (\Delta\mu)^2 \frac{4\bar{\tau}^2}{\bar{\tau}_e + \bar{\tau}_c} \frac{1}{1 + (2\pi f \bar{\tau})^2} \tag{5.41}$$

  with

$$\bar{\tau} \triangleq \left(\frac{1}{\bar{\tau}_e} + \frac{1}{\bar{\tau}_c} \right)^{-1}. \tag{5.42}$$

  Thus, without any additional effort, we way plug all the parameters extracted from the time-domain analysis into 5.41 and we have got the PSD the studied RTN. Does the DFT-based PSD estimation match this expression?
- Can we identify the background noise with flicker ($1/f$) noise and/or thermal noise? A spectral analysis addresses this question by firstly

looking at the frequency dependence of the estimated PSD. Then comes into play device physics: beyond a qualitative trend, the presence of one of these famous noise sources can be validated or invalidated through simulations and theoretical predictions.

5.3.1 Application of the periodogram method

Spectral estimation is performed using the periodogram method developed in Chapter 5. To each realization $X^n(t)$ of one noise source that may be ΔV_{th}, $\Delta V_{th,RTN}$ or $\Delta V_{th,noise}$ is associated one *periodogram*

$$\widehat{S}_T^{(n)}(f) = \frac{2}{T}\left|X_T^{(n)}(f)\right|^2. \tag{5.43}$$

The spectrum $X_T^{(n)}(f)$ is computed with a DFT:

$$X_T^{(n)}(f) \longrightarrow X_T^{(n)}(f_k), \tag{5.44}$$

with $f_k = 0, \mathrm{d}f, 2\cdot\mathrm{d}f, \ldots, f_s/2$. The spectral resolution limit is 250Hz as indicated in Table 5.1.

It will be shown that individual estimations (5.43) are highly inaccurate. Forming the peridograms $\widehat{S}_{T,n}(f)$ of several traces, we obtain a smoother estimation of the desired PSD:

$$\widehat{S}_T(f) \equiv \bar{S}_T(f) = \frac{1}{N_{traces}} \sum_{n=1}^{N_{traces}} \widehat{S}_{T,n}(f). \tag{5.45}$$

As T is fixed, we will omit the subscript for the sake of convenience. We simply write:

$$\widehat{S}(f) = \frac{1}{N_{traces}} \sum_{n=1}^{N_{traces}} \widehat{S}_n(f) \tag{5.46}$$

We recall that the variance of estimator $\widehat{S}(f)$ decreases as $1/N_{traces}$:

$$\mathrm{Var}\left\{\widehat{S}(f)\right\} \sim \frac{1}{N_{traces}}, \tag{5.47}$$

which theoretically ensures that the spectral estimation converges as N_{traces} increases. This result is valid provided that T is large [26, 30, 31].

The method is now applied to three full ΔV_{th} traces in Fig. 5.18. According to 5.46, the estimated PSD is a ensemble average of the three computed peridograms:

$$\widehat{S}(f) = \frac{1}{3}\left(\widehat{S}_1(f) + \widehat{S}_2(f) + \widehat{S}_1(f)\right) \tag{5.48}$$

This is more easily understood in Fig. 5.18a where the power spectral density is represented in a linear scale, which is quite usual but useful to have an insight of method. The conventional representation in a log-log scale plot is provided in Fig. 5.18b. Several important explanations and comments are made:

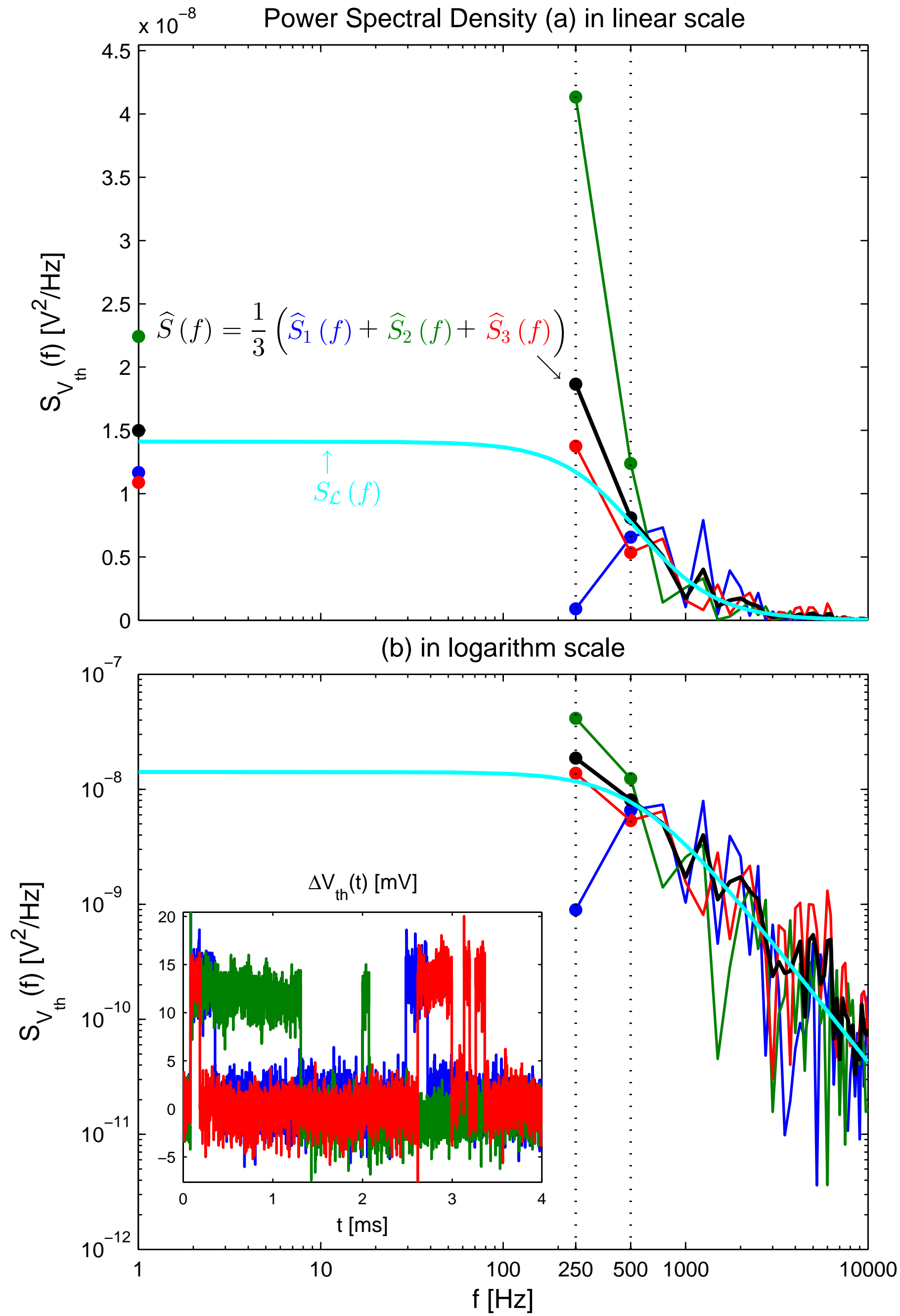


Figure 5.18: Peridogram method applied to the three full traces reproduced in the inset. $S_{\mathcal{L}}(f)$ is the analytical Lorentzian (5.41). Values read at 1Hz are actually the DC values.

- Its appears that the RTN contribution to the PSD is dominant.
- Fig. 5.18 shows that individual periodograms (in blue, green and red) are generally poor estimation of the analytical Lorentzian (in cyan) (5.41). Their ensemble average (in black) already follows the good trend. We obviously expect to improve it by increasing the number of traces.
- We may notice that actually the periodogram $\widehat{S}_1(f)$ of the third trace is luckily much more accurate that the two others. This illustrates to random behaviour of individual peridograms, depending on a particular trace shape, whether this trace is representative of the studied RTN. This further supports the need for an ensemble average.
- We will not pay to much of attention to the "DC" component. As the frequency is represented in a logarithm scale, 0Hz should even not appear. For illustration purpose, it has been arbitrary shifted to 1Hz.

5.3.2 Validation of the Lorentzian as the RTN PSD

We can confirm the domination of the RTN by comparing the estimated PSD for the full analog traces on one hand, and the extracted pure RTN on the other done. These are respectively represented in Fig. 5.19a and Fig. 5.19b. Only the fine eye can observe slight difference between the two, mainly in higher frequency starting from $\sim$ 100kHz which might roughly be the frequency at which the background noise start to become non negligible.

All the 996 are exploited in Fig. 5.20. The peridogram-based PSD estimation of the RTN perfectly matches the analytical Lorentzian given in (5.41), assumed that accurately extracted $\bar{\tau}_e$ and $\bar{\tau}_c$ are plugged into its expression. Above the corner frequency, the slope of $-40\text{dB}/dec$ is precisely recovered.

Also depicted is the obtained PSD for the Gaussian background noise, applying peridogram method to $\Delta V_{th,noise}(t)$. It is well below the Lorentzian up to $\sim 50-100$kHz. The spectral estimation of the noise gets oscillating and inaccurate above $\sim$ 30kHz. However, the region of confidence is large enough to reveal the presence of $1/f$ noise. Slope seems somewhat smaller than $-20\text{dB}/dec$. The study and the validation of the pink color of the background noise is addressed in detailed in the next section.

5.3.3 Background noise is flicker noise

Parameter extractions

Presumed form of a flicker noise is

$$S_{1/f}(f) = \frac{S_{1/f}(1\text{Hz})}{f^a}. \tag{5.49}$$

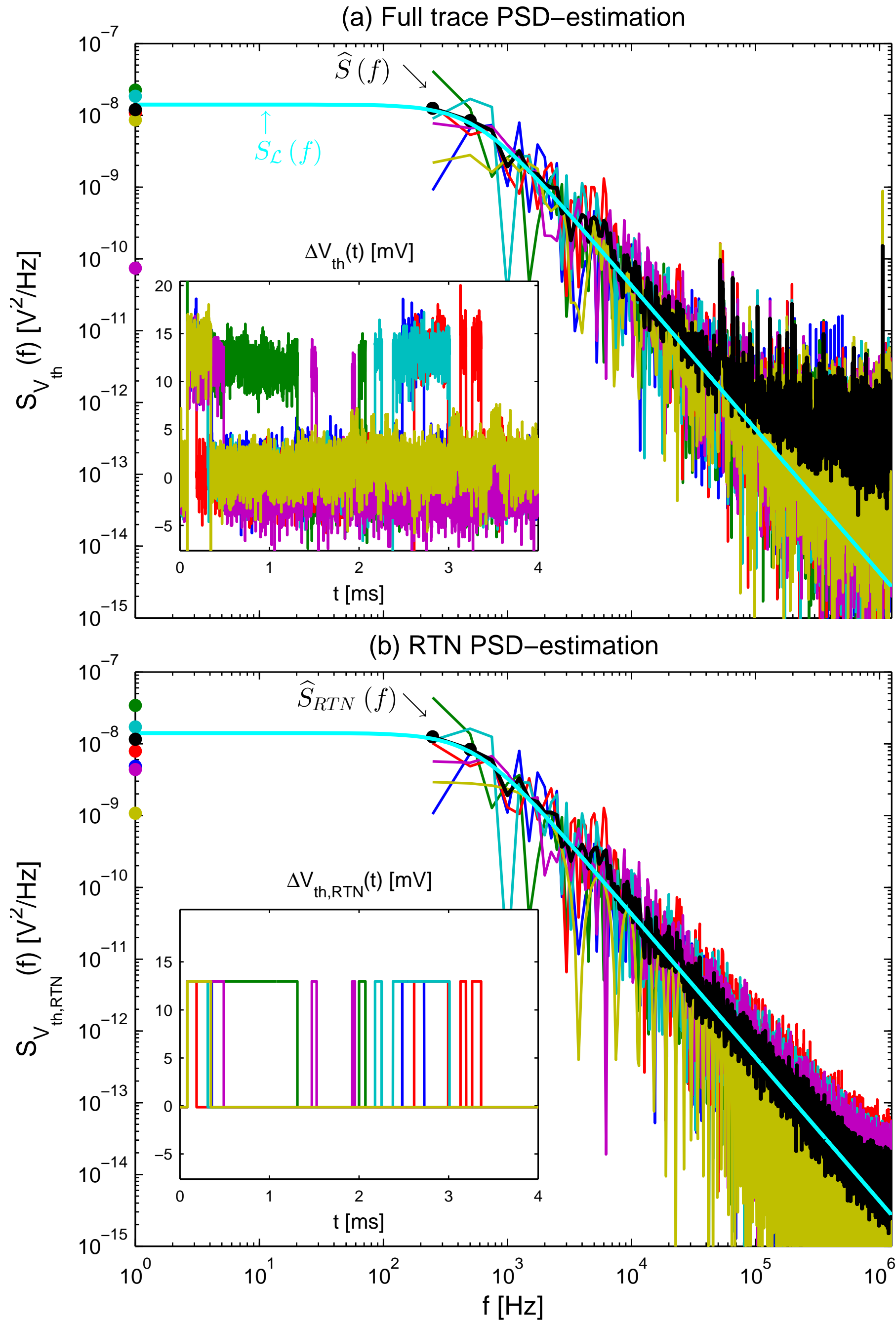


Figure 5.19: Peridogram method applied to (a) six full $\Delta V_{th}(t)$ traces, and (b) the corresponding $\Delta V_{th,RTN}(t)$ traces. The similarity between the two indicates that the RTN dominates the PSD.

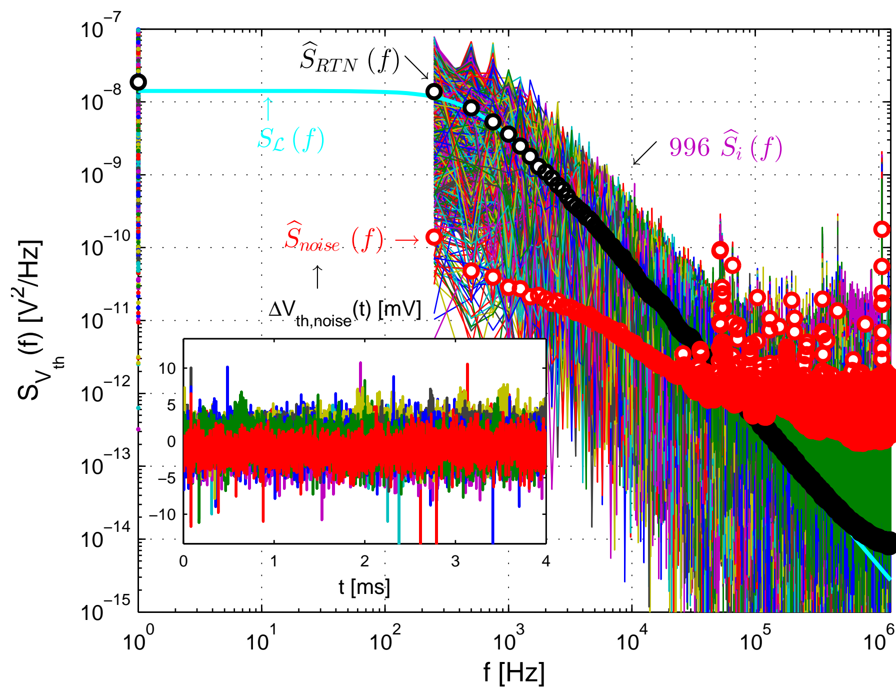


Figure 5.20: Peridogram method applied to all the 996 traces. Black circles are $\widehat{S}(f)$ that perfectly matches the analytical Lorentzian $S_{\mathcal{L}}(f)$. $\widehat{S}_{noise}(f)$ is computed from the background noise $\Delta V_{th,noise}(t)$ traces reproduced in the inset.

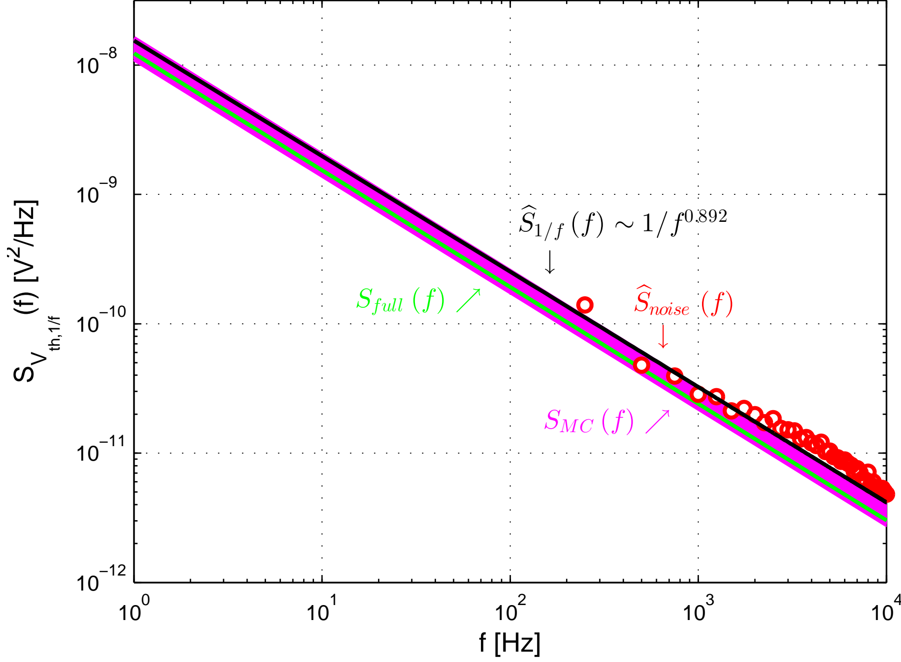


Figure 5.21: Spectral study of the background noise that appear to be a flicker noise. Linear regression yield an exponent $a = 0.892 \approx 0.9$. $1/f$ noise is physically validated through simulations, both a top-level one and a Monte-Carlo one.

This expression is exactly similar to the model used by early circuit simulator, e.g. SPICE, [19, 49]:

$$S_{\Delta I_D,1/f}(f) = \frac{KF \cdot I_D^{AF}}{(C'_{ox})^{\alpha} WL} \cdot \frac{1}{f^a} \tag{5.50}$$

where KF, AF, α and a are empirical factors. Knowing C'_{ox}, W and L, and the DC bias I_D and g_m, our quantity $S_{1/f}(1\text{Hz})$ might well be related to KF if requested. If the noise is assumed be to treated as a small signal [5], we easily convert $S_{\Delta V_{th}}(f)$ to $S_{\Delta I_D}(f)$, and conversely, using

$$S_{\Delta I_D}(f) = g_m^2 \cdot S_{\Delta V_{th}}(f)\,. \tag{5.51}$$

Taking $10\log_{10}(\,\cdot\,)$ of both member of (5.49):

$$10\log_{10}\left(S_{1/f}(f)\right) = 10\log_{10}\left(S_{1/f}(1\text{Hz})\right) - a \cdot 10\log_{10}(f)\,, \tag{5.52}$$

[5] By the way, this is far from obvious for a discontinuous jump process such as the RTN, whose relative current amplitude larger than 50% were reported [2,3]. Such a debate would be interesting although it goes beyond the scope of this work

Device	Param	Noise Contribution	% Of Total
/M0	fn	4.58881e-07	90.44
/M2	id	6.02629e-09	1.19
/M0	id	5.82742e-09	1.15
/I67/I48/M2b	id	4.98091e-09	0.98
/M17	fn	4.47211e-09	0.88
/I67/I48/M1b	id	2.6041e-09	0.51
/I67/I48/M1a	id	2.59722e-09	0.51
/M2	fn	2.25341e-09	0.44
/I67/I48/M2a	id	1.80059e-09	0.35
/I67/I48/M4	id	9.4583e-10	0.19
/I180/I_BG/M3	fn	8.79726e-10	0.17
I67.I48.R1.r5	thermal_noise	7.2754e-10	0.14
I67.I48.R1.r4	thermal_noise	7.27531e-10	0.14
I67.I48.R1.r3	thermal_noise	7.27522e-10	0.14
I67.I48.R1.r2	thermal_noise	7.27513e-10	0.14
/I67/I48/M3q	id	7.1101e-10	0.14
/M1	id	6.20568e-10	0.12
/I67/I48/M2b	fn	5.3669e-10	0.11
/I180/I_BG/Mbin	fn	4.89289e-10	0.10
/I67/I48/M5a	id	4.72493e-10	0.09

Integrated Noise Summary (in V^2) Sorted By Noise Contributors
Total Summarized Noise = 5.07376e-07
No input referred noise available
The above noise summary info is for noise data

Figure 5.22: Report of the noise contributions from the top-level simulation of the analog block of Fig. 5.1.

$S_{1/f}$ (1Hz) and a parameters are extracted by linear fit of the first points

$$\left(f_k, \widehat{S}_{noise}\left(f_k\right)\right),$$

the red circles in Fig. 5.21. $S_{1/f}$ (1Hz) $= -78.1$dB and $a = 0.892$ are found.

Quantitative validation through simulations

The $1/f$ like shape of the PSD is thus quantitatively validated. Are we sure that this noise is well the drain current flicker of the DUT? We address this question. All the simulation were performed by Marko Simicic.

A top-level simulation of the whole measurement circuit of 5.1, assumed in the typical corner, allows to ensure that the noise is well dominated by the DUT. Report presented in Fig. 5.22 fortunately confirms it. In addition, this top-level simulation yields a reference PSD for the flicker noise,

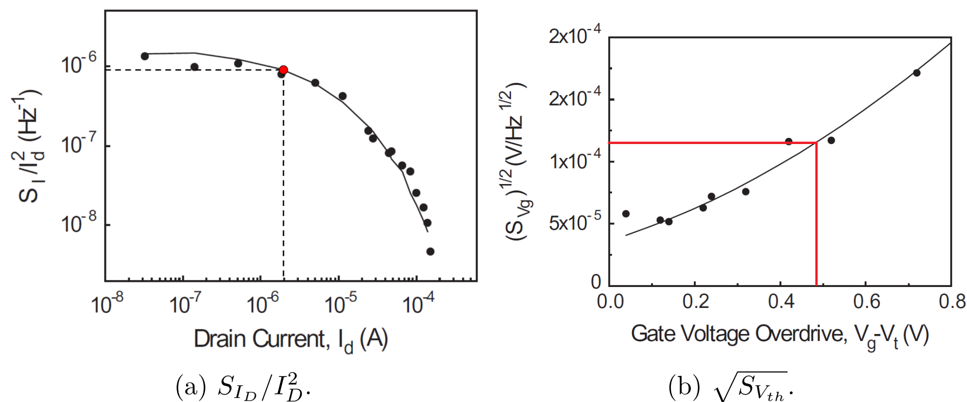
$$S_{full}\left(f\right),$$

also depicted in Fig. 5.21.

Once the domination of the DUT has been checked, top-level simulation gets necessarily computational intensive. Traditional noise AC simulations of a single device may be used to validate the $1/f$ noise. As we have no idea about the corner in which the golden device, a Monte-Carlo simulation accounting for device-to-device PSD variability (mismatch only) was performed. All the obtained curves

$$S_{MC}\left(f\right),$$

are added to Fig. 5.21. The matching between all the simulations and our extracted background noise is excellent.



(a) S_{I_D}/I_D^2. (b) $\sqrt{S_{V_{th}}}$.

Figure 5.23: Power spectral density measured at $V_{DS} = 50$mV and $f = 1$Hz. Data from [50].

Quantity	Units	Spectral analysis	Simulation	Measurement
$S_{\Delta V_{th}}$	[dB]	-78.1	-78.8	-78.5
$S_{\Delta I_D}$	[dB]	-170.9	-170.9	-174.8

Table 5.6: Flicker noise power spectral density evaluated at 1Hz. Comparison between our spectral analysis based extraction, Marko Simicic's simulation, and measurement from [50].

Comparison with experimental data from the literature

Paper [50] results from a collaboration between the IMEP-LAHC (LETI, Grenoble), the Aristotle University of Thessaloniki (Greece), and ST Microelectronics®. Spectral measurement of 28nm bulk devices in linear region are proposed. This offers us a comparison with experimental data. The specific bias point $I_D - V_{GS}$ of our device is marked in Fig. 5.23. Table 5.6 arises from these values, our extracted $1/f$ noise (5.49). Equation(5.51) has also been used.

Thermal noise

So far, we have point out, through spectral analysis, that the noise is dominated by the pure two-level RTN. $1/f$ noise was demonstrated to be present. It is only significant with respected to to RTN below 1Hz or above 50kHz. Fig. 5.24 is a nice summarizing picture. We have separated RTN and $1/f$ noise by data post-processing. Similar results have already been reported in the literature, for instance in [17]. However, they used very switched bias conditions while our analysis requires only constant-bias measurement.

We may wonder whether the really not smooth plateau between 100kHz

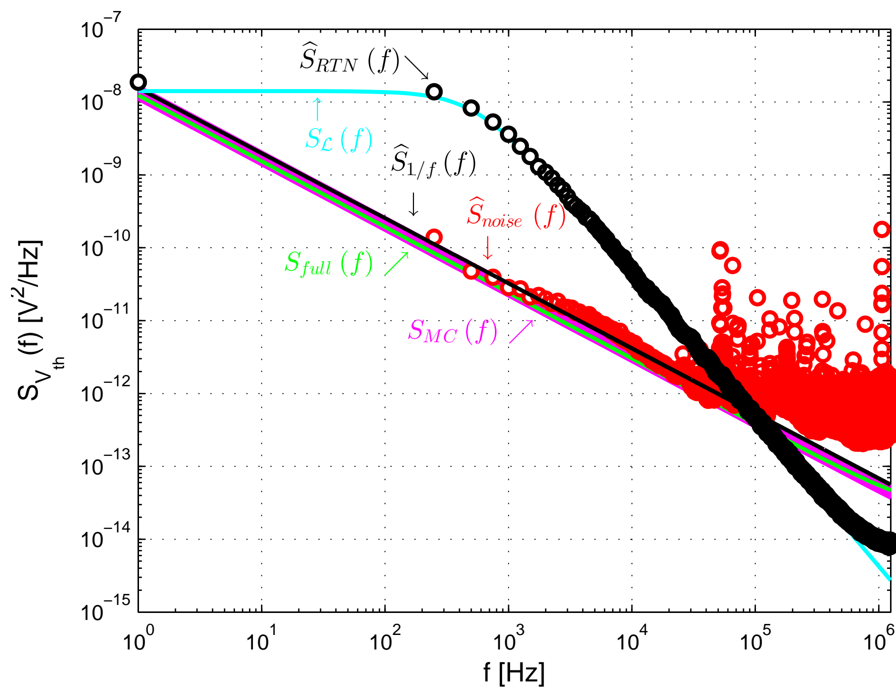


Figure 5.24: Summary of the spectral analysis. Both RTN and $1/f$ were extracted and validated.

and 1MHz can be attributed to the thermal noise. The observed "flattening" of the PSD could rather be attributable to some bad performance of the DFT-based periodogram method close to the maximum frequency $f_{max} = f_s/2 \approx 1\text{MHz}$.

At the DC bias point, i.e. moderate inversion and more or less linear region, white thermal noise PSD can be estimated as [51, 52]

$$S_{\Delta V_{th},thermal}\left(f\right) \approx 4k_BT\left(\frac{1}{g_m}\right). \tag{5.53}$$

We find

$$S_{\Delta V_{th,thermal}} \approx 4\cdot 10^{-16}\text{V}^2/\text{Hz} \tag{5.54}$$

or, expressed in dB,

$$S_{\Delta V_{th,thermal}} \approx -154\text{dB} \tag{5.55}$$

which is much below the values observed even at 1MHz in Fig. 5.24. As conclusion, at our sampling frequency, i.e. the bandwidth of our system, the background is dominated by the flicker noise.

One remaining question: RMS value of the background noise computed from the $1/f$ PSD Time-domain analysis has provided us different estimation of σ, the RMS value of the Gaussian noise. All were consistent with each others, see for instance Table 5.3 or (5.12):

$$\sigma \approx 2\text{mV}.$$

We should, in principle, be able to calculate this value by integration the $1/f$ noise PSD over the bandwidth:

$$\begin{aligned}\sigma &= \int_{1/T}^{f_s/2} S_{1/f}\left(f\right)\,\mathrm{d}f\\ &\approx \sum_{k=1}^{N_{FFT}/2} S_{1/f}\left(k\cdot \mathrm{d}f\right)\cdot \mathrm{d}f.\end{aligned} \tag{5.56}$$

Only 0.757mV is found. Explanation for such a low value would deserve a deeper investigation in a future release.

5.4 Conclusion of this study case

This chapter is undoubtedly the most important of this work. Based of a large amount of data from a state-of-the-art on-chip characterization system, the so called Smart array, a comprehensive analysis of a pure single-trap, two-level RTN was presented. Most of the results were output of the Matlab® software tool that we developed.

RTN levels and background noise standard deviation were extracted using different methods that are consistent with each others. All are quite easily extend to the multi-trap case. Time-domain analysis ended up with the Schmitt trigger based extraction of the RTN, as well as the distributions of the time constants. Spectral analysis confirms that the analytical Lorentzian is the exact expression to be incorporated in a AC compact model, provided that the right parameters are plugged into it. Finally, the background noise was proved to be Gaussian and nothing but that the drain current flicker noise of the DUT.

Chapter 6

Conclusion and perspectives

This work was finally very oriented toward applied mathematics, numerical simulation and stochastic processes. However, all this theoretical background helps us in understanding the noise and yield the tools to model it.

In the perspective of AC compact modelling of the RTN, this work has proposed a comprehensive methodology for both time-domain and spectral analysis. The analytical Lorentzian for the RTN power spectral density has been validated, while the background has fortunately been related to the usual flicker noise.

We can regret the fact that only one practical study case has been addressed in details. We wish to apply our methodology and Matlab® software tool to others devices, in different bias and temperature conditions. Large-scale characterization is indeed necessary to validate a serious compact model.

The development of the software has required a lot of effort and time. However, it should be made more systematic, more robust and more use-friendly. We would like to extend it to the general multi-trap case, and potentially make of it a industrial quality tool.

Bibliography

[1] B. Kaczer, J. Franco, P. Weckx, P. Roussel, M. Simicic, V. Putcha, E. Bury, M. Cho, R. Degraeve, D. Linten, G. Groeseneken, P. Debacker, B. Parvais, P. Raghavan, F. Catthoor, G. Rzepa, M. Waltl, W. Goes, and T. Grasser, “The defect-centric perspective of device and circuit reliability—from gate oxide defects to circuits,” *Solid-State Electronics*, vol. 125, pp. 52–62, nov 2016. [Online]. Available: http://dx.doi.org/10.1016/j.sse.2016.07.010

[2] A. Asenov, R. Balasubramaniam, A. Brown, and J. Davies, “RTS amplitudes in decananometer MOSFETs: 3-d simulation study,” *IEEE Transactions on Electron Devices*, vol. 50, no. 3, pp. 839–845, mar 2003. [Online]. Available: http://dx.doi.org/10.1109/TED.2003.811418

[3] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, “Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs,” *IEEE Transactions on Electron Devices*, vol. 50, no. 9, pp. 1837–1852, Sept 2003.

[4] T. Skotnicki, C. Fenouillet-Beranger, C. Gallon, F. Boeuf, S. Monfray, F. Payet, A. Pouydebasque, M. Szczap, A. Farcy, F. Arnaud, S. Clerc, M. Sellier, A. Cathignol, J. P. Schoellkopf, E. Perea, R. Ferrant, and H. Mingam, “Innovative Materials, Devices, and CMOS Technologies for Low-Power Mobile Multimedia,” *IEEE Transactions on Electron Devices*, vol. 55, no. 1, pp. 96–130, Jan 2008.

[5] K. J. Kuhn, M. D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. T. Ma, A. Maheshwari, and S. Mudanai, “Process Technology Variation,” *IEEE Transactions on Electron Devices*, vol. 58, no. 8, pp. 2197–2208, Aug 2011.

[6] M. Pelgrom, A. Duinmaijer, and A. Welbers, “Matching properties of MOS transistors,” *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, oct 1989. [Online]. Available: http://dx.doi.org/10.1109/JSSC.1989.572629

[7] S. Realov and K. L. Shepard, "Analysis of random telegraph noise in 45-nm CMOS using on-chip characterization system," *IEEE Transactions on Electron Devices*, vol. 60, no. 5, pp. 1716–1722, 2013.

[8] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction-Diffusion to Switching Oxide Traps," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, pp. 3652–3666, nov 2011. [Online]. Available: http://dx.doi.org/10.1109/TED.2011.2164543

[9] A. L. McWhorter *et al.*, "$1/f$ noise and related surface effects in germanium." 1955.

[10] M. Kirton and M. Uren, "Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency ($1/f$) noise," *Advances in Physics*, vol. 38, no. 4, pp. 367–468, jan 1989. [Online]. Available: http://dx.doi.org/10.1080/00018738900101122

[11] T. Grasser, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectronics Reliability*, vol. 52, no. 1, pp. 39–70, jan 2012. [Online]. Available: http://dx.doi.org/10.1016/j.microrel.2011.09.002

[12] T. Grasser, Ed., *Bias Temperature Instability for Devices and Circuits*. Springer Nature, 2014. [Online]. Available: http://dx.doi.org/10.1007/978-1-4614-7909-3

[13] S. Machlup, "Noise in semiconductors: Spectrum of a two-parameter random signal," *Journal of Applied Physics*, vol. 25, no. 3, p. 341, 1954. [Online]. Available: http://dx.doi.org/10.1063/1.1721637

[14] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, "Random telegraph noise of deep-submicrometer MOSFETs," *IEEE electron device letters*, vol. 11, no. 2, pp. 90–92, 1990.

[15] G. Ghibaudo and T. Boutchacha, "Electrical noise and RTS fluctuations in advanced CMOS devices," *Microelectronics Reliability*, vol. 42, no. 4-5, pp. 573–582, apr 2002. [Online]. Available: http://dx.doi.org/10.1016/S0026-2714(02)00025-2

[16] L. K. J. Vandamme and F. N. Hooge, "What do we certainly know about $1/f$ noise in MOSTs?" *IEEE Transactions on Electron Devices*, vol. 55, no. 11, pp. 3070–3085, nov 2008. [Online]. Available: http://dx.doi.org/10.1109/TED.2008.2005167

[17] J. Kolhatkar, L. Vandamme, C. Salm, and H. Wallinga, "Separation of random telegraph signals from 1/f noise in mosfets under constant and switched bias conditions," in *European Solid-State Device Research, 2003. ESSDERC'03. 33rd Conference on.* IEEE, 2003, pp. 549–552.

[18] M.-L. Fan, V. P.-H. Hu, Y.-N. Chen, P. Su, and C.-T. Chuang, "Analysis of single-trap-induced random telegraph noise on FinFET devices, 6t SRAM cell, and logic circuits," *IEEE Transactions on Electron Devices*, vol. 59, no. 8, pp. 2227–2234, aug 2012. [Online]. Available: http://dx.doi.org/10.1109/TED.2012.2200686

[19] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, "A physics-based mosfet noise model for circuit simulators," *IEEE Transactions on Electron Devices*, vol. 37, no. 5, pp. 1323–1333, 1990.

[20] ——, "A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 37, no. 3, pp. 654–665, 1990.

[21] N. Tega, H. Miki, F. Pagette, D. Frank, A. Ray, M. Rooks, W. Haensch, and K. Torii, "Increasing threshold voltage variation due to random telegraph noise in fets as gate lengths scale to 20 nm," in *2009 Symposium on VLSI Technology.* IEEE, 2009, pp. 50–51.

[22] H. Luo, Y. Wang, Y. Cao, Y. Xie, Y. Ma, and H. Yang, "Temporal performance degradation under RTN: Evaluation and mitigation for nanoscale circuits," in *2012 IEEE Computer Society Annual Symposium on VLSI.* Institute of Electrical and Electronics Engineers (IEEE), aug 2012. [Online]. Available: http://dx.doi.org/10.1109/ISVLSI.2012.35

[23] F. Hooge, T. Kleinpenning, and L. Vandamme, "Experimental studies on 1/f noise," *Reports on progress in Physics*, vol. 44, no. 5, p. 479, 1981.

[24] T. Grasser, H. Reisinger, P. J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, "The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability," in *2010 IEEE International Reliability Physics Symposium*, May 2010, pp. 16–25.

[25] D. T. Gillespie, *Markov processes: an introduction for physical scientists.* Elsevier, 1991.

[26] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, ser. McGraw-Hill Series in Electrical Engineering. McGraw-Hill, 1991.

[27] D. T. Gillespie, "The mathematics of Brownian motion and Johnson noise," *American Journal of Physics*, vol. 64, no. 3, p. 225, 1996. [Online]. Available: http://dx.doi.org/10.1119/1.18210

[28] W. C. Chu, *Speech coding algorithms: foundation and evolution of standardized coders.* John Wiley & Sons, 2004.

[29] R. Brown and P. Hwang, *Introduction to Random Signals and Applied Kalman Filtering with MATLAB Exercises, 4th Edition.* Wiley, 2012.

[30] P. Stoica, R. L. Moses *et al.*, *Spectral analysis of signals.* Pearson Prentice Hall Upper Saddle River, NJ, 2005, vol. 452.

[31] P. Stoica and R. L. Moses, *Introduction to spectral analysis.* Prentice hall Upper Saddle River, 1997, vol. 1.

[32] M. Simicic, S. Morrison, B. Parvais, P. Weckx, B. Kaczer, K. Sewada, H. Ammo, S. Yamakawa, K. Nomoto, M. Ono, D. Linten, D. Verkest, P. Wambacq, G. Groeseneken, and G. Gielen, "A fully-integrated method for RTN parameters extraction," in *2017 Symposium on VLSI Technology.* IEEE, in press.

[33] Q. Tang, X. Wang, J. Keane, and C. H. Kim, "RTN induced frequency shift measurements using a ring oscillator based circuit," in *2013 Symposium on VLSI Technology*, June 2013, pp. T188–T189.

[34] M. Simicic, V. Putcha, B. Parvais, P. Weckx, B. Kaczer, G. Groeseneken, G. Gielen, D. Linten, and A. Thean, "Advanced MOSFET variability and reliability characterization array," in *2015 IEEE International Integrated Reliability Workshop (IIRW)*, Oct 2015, pp. 73–76.

[35] V. Putcha, M. Simicic, P. Weckx, B. Parvais, J. Franco, B. Kaczer, D. Linten, D. Verkest, A. Thean, and G. Groeseneken, "Smart-array for pipelined BTI characterization," in *2015 IEEE International Integrated Reliability Workshop (IIRW)*, Oct 2015, pp. 95–98.

[36] A. Whitcombe, S. Taylor, M. Denham, V. Milovanovic, and B. Nikolic, "On-chip i-v variability and random telegraph noise characterization in 28 nm CMOS," in *2016 46th European Solid-State Device Research Conference (ESSDERC)*, Sept 2016, pp. 248–251.

[37] D. Flandre, V. Kilchytska, and T. Rudenko, "g_m/i_d method for threshold voltage extraction applicable in advanced MOSFETs with nonlinear behavior above threshold," *IEEE Electron Device Letters*, vol. 31, no. 9, pp. 930–932, Sept 2010.

[38] T. Rudenko, V. Kilchytska, M. K. M. Arshad, J. P. Raskin, A. Nazarov, and D. Flandre, "On the MOSFET threshold voltage extraction by transconductance and transconductance-to-current ratio change methods: Part i ;effect of gate-voltage-dependent mobility," *IEEE Transactions on Electron Devices*, vol. 58, no. 12, pp. 4172–4179, Dec 2011.

[39] C.-W. Lee, S.-R.-N. Yun, C.-G. Yu, J.-T. Park, and J.-P. Colinge, "Device design guidelines for nano-scale MuGFETs," *Solid-State Electronics*, vol. 51, no. 3, pp. 505–510, mar 2007. [Online]. Available: http://dx.doi.org/10.1016/j.sse.2006.11.013

[40] J.-P. Colinge, *FinFETs and Other Multi-Gate Transistors*, 1st ed. Springer Publishing Company, Incorporated, 2007.

[41] D. Wackerly, W. Mendenhall, and R. Scheaffer, *Mathematical statistics with applications.* Nelson Education, 2007.

[42] H. C. Thode, *Testing for normality.* CRC press, 2002, vol. 164.

[43] A. R. Henderson, "Testing experimental data for univariate normality," *Clinica Chimica Acta*, vol. 366, no. 1, pp. 112–129, 2006.

[44] H. Miki, N. Tega, M. Yamaoka, D. J. Frank, A. Bansal, M. Kobayashi, K. Cheng, C. P. D'Emic, Z. Ren, S. Wu, J. B. Yau, Y. Zhu, M. A. Guillorn, D. G. Park, W. Haensch, E. Leobandung, and K. Torii, "Statistical measurement of random telegraph noise and its impact in scaled-down high-k/metal-gate mosfets," in *2012 International Electron Devices Meeting*, Dec 2012, pp. 19.1.1–19.1.4.

[45] D. R. Cox and D. V. Hinkley, *Theoretical statistics.* CRC Press, 1979.

[46] B. Efron, "The efficiency of cox's likelihood function for censored data," *Journal of the American statistical Association*, vol. 72, no. 359, pp. 557–565, 1977.

[47] J. P. Klein and M. L. Moeschberger, *Survival analysis: techniques for censored and truncated data.* Springer Science & Business Media, 2005.

[48] R. Peto, "Experimental survival curves for interval-censored data," *Applied Statistics*, pp. 86–91, 1973.

[49] R. J. Baker, *CMOS Circuit Design, Layout, and Simulation*, 3rd ed. Wiley-IEEE Press, 2010.

[50] E. Ioannidis, C. Dimitriadis, S. Haendler, R. Bianchi, J. Jomaah, and G. Ghibaudo, "Improved analysis and modeling of low-frequency noise in nanoscale MOSFETs," *Solid-State Electronics*, vol. 76, pp. 54–59, 2012.

[51] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, ser. The Oxford Series in Electrical and Computer Engineering Series. Oxford University Press, 2011.

[52] C.-H. Chen and M. J. Deen, "Channel noise modeling of deep submicron MOSFETs," *IEEE Transactions on Electron Devices*, vol. 49, no. 8, pp. 1484–1487, 2002.

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve www.uclouvain.be/epl