Universidade de São Paulo
Instituto de Física

Modelo estocástico de flutuações turbulentas de plasma aplicado ao tokamak TCABR

Martim Zurita

Orientador: Prof. Dr. Zwinglio de Oliveira Guimarães Filho

Dissertação de mestrado apresentada ao Instituto de Física da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Ciências.

Banca examinadora:
Prof. Dr. Zwinglio de Oliveira Guimarães Filho – Orientador (IFUSP)
Prof. Dr. Dennis Lozano Toufen (IFSP)
Prof. Dr. Ricardo Luiz Viana (UFPR)

São Paulo
2022

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Zurita, Martim

Modelo estocástico de flutuações turbulentas de plasma aplicado ao tokamak TCABR / Stochastic modeling of turbulent plasma fluctuations applied to the TCABR tokamak. São Paulo, 2022.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Depto. de Física Aplicada

Orientador(a): Prof. Dr. Zwinglio de Oliveira Guimarães Filho

Área de Concentração: Física de Plasmas

Unitermos: 1. Física de Plasmas; 2. Tokamaks; 3. Turbulência eletrostática; 4. Análise de séries temporais; 5. Modelos para processos estocásticos.

USP/IF/SBI-002/2022

University of São Paulo
Physics Institute

Stochastic modeling of turbulent plasma fluctuations applied to the TCABR tokamak

Martim Zurita

Supervisor: Prof. Zwinglio de Oliveira Guimarães Filho

Dissertation submitted to the Physics Institute of the University of São Paulo in partial fulfillment of the requirements for the degree of Master of Science.

Examining Committee:
Prof. Zwinglio de Oliveira Guimarães Filho — Supervisor (IFUSP)
Prof. Dennis Lozano Toufen (IFSP)
Prof. Ricardo Luiz Viana (UFPR)

São Paulo
2022

To my parents

Acknowledgements

First of all, this dissertation is dedicated to my parents, for their support in my career and aspirations. My girlfriend also deserves a special place here for listening so much about my research, helping to polish my explanations.

Secondly, I am grateful for Prof. Zwinglio O. Guimarães-Filho and his mentorship. Zwinglio is an excellent advisor. He worries about his students and, at the same time, gives us intellectual freedom. Also, his enthusiasm for science is a great source of inspiration for everyone around him. I have learned a lot with Zwinglio and feel very lucky to have him as my master's degree supervisor.

Additionally, other scientists who assisted directly or indirectly in the process of this dissertation should be acknowledged. Early in my undergraduate studies, Victor S. Fernandes suggested to me an undergraduate project (a.k.a. scientific initiation) with Zwinglio, and that was the first seed of this work. Victor is a very encouraging and caring person. Referring directly to this thesis, I highlight Prof. Iberê L. Caldas, who is the chief of the thematic project under which this research was financed. Iberê is very supportive and remarkable for connecting scientists. Also, the TCABR team was responsible for the experiments used here. In special, Gustavo G. Grenfell coordinated the experiment 34132, which was fundamental for the dissertation. Additionally, I thank Wanderley P. de Sá for helping me with the TCABR computer cluster. Finally, I would like to acknowledge Prof. Gustavo P. Canal. He is a very inspiring and dedicated scientist, who is always willing to aid students to pursue the best for their careers.

Furthermore, I thank the coauthors of my first article (apart from Zwinglio): W. Andres Hernandez, Caíke Crepaldi and Felipe A. C. Pereira. They have shared their work with me and each one of them helped me to learn important concepts of plasma physics or data analysis.

The importance of CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) and FAPESP (*Fundação de Amparo à Pesquisa do Estado de São Paulo*) should also be recognized. CNPq provided financial support to this research from March to September of 2020, through grant 131254/2020-0, while FAPESP funded it from October of 2020 to February of 2022, with grant 2020/03912-4.[1] It is vital for the progress of science that state-owned funding agencies are valued.

[1]Following clause 7.2 of the FAPESP grant term, I highlight that the hypotheses and conclusions of this dissertation are of my responsibility, not reflecting FAPESP's point of view.

Abstract

In toroidal devices known as tokamaks, high-temperature plasmas are confined by intense magnetic fields. Nevertheless, this confinement is deteriorated by turbulence at the edge of the devices. This turbulence has an intermittent behavior with the presence of high-amplitude bursts. To describe local measurements of density with bursts, a stochastic pulse train model (SPTM) has been developed since the last decade. For such a model, different categories of background signals have been considered in the literature, namely, backgrounds with Gaussian noises (correlated and uncorrelated) or with small-amplitude pulses. However, until now these models with different background signals weren't simultaneously compared to an experiment. Moreover, there isn't a fitting method for the SPTM that can evaluate all its parameters in a unified and objective way. The present dissertation aims to fulfill these two gaps. Having created the SPTM fit, we applied it to the TCABR tokamak. For this analysis, we utilized measurements of ion saturation current, a signal proportional to the local plasma density. In addition, we introduced to the context of the SPTM two non-linear tools: the complexity-entropy diagram and the determinism from recurrence quantification analysis. With them and the frequency spectrum, we concluded that, for the analyzed experiment, the model with a pulse background described the structure of plasma density fluctuations better than the models with Gaussian noise.

Keywords: plasma physics. Tokamaks. Electrostatic turbulence. Ion saturation current. Stochastic modeling.

Resumo

Em dispositivos toroidais conhecidos como tokamaks, plasmas com temperaturas solares são confinados por intensos campos magnéticos. Esse confinamento, todavia, é deteriorado pela turbulência na borda dos dispositivos. Essa turbulência possui um comportamento intermitente marcado pela presença de rajadas de alta amplitude, denominadas como *bursts*. Para descrever medidas locais de densidade com *bursts*, um modelo estocástico de trem de pulsos (METP) vem sendo desenvolvido desde a última década. Para tal modelo, foram considerados na literatura diferentes categorias de sinais de fundo, a saber, fundos com diferentes ruídos gaussianos (descorrelacionados e correlacionados) ou com pulsos de pequena amplitude. Entretanto, até o momento esses sinais de fundo não foram comparados simultaneamente a um sinal experimental. Além disso, ainda não há um método de ajuste para o METP que obtenha todos os seus parâmetros de forma unificada e objetiva. A presente dissertação almeja preencher essas duas lacunas. Criado o ajuste do METP, o aplicamos para o tokamak TCABR. Para tal análise, utilizamos medidas de corrente de saturação iônica, um sinal proporcional a densidade local do plasma. Adicionalmente, introduzimos para o contexto do METP duas ferramentas não lineares: o diagrama de complexidade-entropia e o determinismo da análise de recorrência. Com elas e com o espectro de potência, concluímos que, para o experimento analisado, o modelo com fundo de pulsos descreveu a estrutura de flutuações de densidade melhor do que os modelos com ruído gaussiano.

Palavras-chave: Física de Plasmas. Tokamaks. Turbulência eletrostática. Corrente de saturação iônica. Modelagem estocástica.

Contents

List of Tables

List of abbreviations

1. CAB: conditionally averaged burst
2. CDF: cumulative distribution function
3. CF: characteristic function
4. CH: complexity-entropy (H comes from the Shannon entropy H_S)
5. CN: colored noise
6. COMPASS: tokamak name derived from COMPact ASSembly
7. DF: degrees of freedom
8. DFT: discrete Fourier transform
9. ECF: empirical characteristic function
10. ELM: edge-localized mode
11. EPFL: *École Polytechnique Fédérale de Lausanne*
12. fBm: fractional Brownian motion
13. FFT: fast Fourier transform
14. FPP: filtered Poisson process
15. GPI: gas puff imaging
16. JET: Joint European Torus
17. KSTAR: Korea Superconducting Tokamak Advanced Research
18. LCFS: last closed flux surface (synonym of LCMS)
19. LCMS: last closed magnetic surface (synonym of LCFS)
20. Li-BES: lithium-beam emission spectroscopy
21. MHD: magneto-hydro-dynamics
22. NSTX: National Spherical Torus Experiment
23. OU: Ornstein-Uhlenbeck

24. OUN: Ornstein-Uhlenbeck noise
25. RP: recurrence plot
26. RQA: recurrence quantification analysis
27. RR: recurrence rate (set at 10% for this dissertation)
28. PB: pulse background
29. PDF: probability density function
30. PSD: power spectral density
31. SPTM: stochastic pulse train model
32. SPTM-N: stochastic pulse train model with noise
33. SPTM-PB: stochastic pulse train model with pulse background
34. SOL: scrape-off layer
35. TCABR: *Tokamak Chauffage Alfvén Brésilien*
36. TCV: *Tokamak à Configuration Variable*

List of the most used symbols

Here is the list of the symbols most used in this dissertation, in alphabetical order. Some letters are repeated, but their meaning can be understood in their context. The greek alphabet is:

$$\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta, \iota, \kappa, \lambda, \mu, \nu, \xi, o, \pi, \rho, \sigma, \tau, \upsilon, \phi, \chi, \psi, \omega$$

1. a : minor radius of a tokamak with a limiter ($a = 18.0$ cm in TCABR)
2. $\langle A \rangle$: average amplitude of the main pulses
3. $\left\langle A^{(b)} \right\rangle$: average amplitude of the background pulses
4. C_{JS} : Jensen-Shannon complexity
5. $C_{\Phi}(u)$: characteristic function of a times series or random variable $\Phi(t)$
6. DET: determinism quantifier of the recurrence plot
7. d_{CH} : embedding dimension of the CH diagram
8. d_{RQA} : embedding dimension of the RQA recurrence plot
9. f : frequency
10. f_j : j-th point of a reference function for the chi-square analysis
11. H_S : Shannon normalized entropy
12. i : imaginary unit, $i^2 = -1$
13. i : natural index, $i \in \mathbb{N}$
14. I_b : pulse background of the saturation current
15. I_B : main pulses of the saturation current (related to the bursts)
16. I_p : plasma current
17. $I_{\mathcal{N}}$: Gaussian noise background of the saturation current
18. I_{sat} : ion saturation current
19. $\tilde{I}_{sat}$: normalized I_{sat} (*i.e.*, $\tilde{I}_{sat} = (I_{sat} - \langle I_{sat} \rangle)/\sigma_I$)
20. j : natural index, $j \in \mathbb{N}$

21. K_Φ : kurtosis of a time series or random variable $\Phi(t)$
22. n : natural index, $n \in \mathbb{N}$
23. n : plasma density
24. n_e : electron density
25. N : number of points of a data set
26. N_P : total number of burst-related pulses
27. $N_P^{(b)}$: total number of background pulses
28. N_{RQA} : length of the recurrence plots of the RQA
29. $P_\Phi(\Phi)$: probability density function of a random variable Φ (PDF)
30. r : radial coordinate with respect to the center of the plasma column
31. R_0 : major radius of the tokamak (equal to $R_0 = 61.5$ cm in TCABR)
32. $R_\Phi(\tau)$: autocorrelation function of a random variable $\Phi(t)$ with lag τ
33. RR : recurrence rate of the RQA
34. $s_n(\lambda, \zeta_r, \zeta_f)$: stretch coefficient of order $n \in \mathbb{N}$
35. S_Φ : skewness of a random variable or random variable $\Phi(t)$
36. t : time
37. T : total duration of the I_{sat} analysis
38. $\mathcal{T}$: threshold for bursts detection (fixated at $\mathcal{T} = 2.5$)
39. u : argument of the characteristic function
40. y_j : j-th point of a experimental function in the chi-square analysis
41. γ : intermittency parameter, equal to $\tau_d/\tau_w = \tau_d N_P/T$
42. Γ: the Gamma function, $\Gamma(z) \equiv \int_0^\infty x^{z-1}e^{-x}dx, \quad \mathrm{Re}(z) > 0.$
43. Δt : time step of the I_{sat} measurements ($\Delta t = 0.5$ μs in TCABR)
44. Δt_B : minimum waiting time between bursts (fixated at $\Delta t_B = 15$ μs)
45. ϵ : noise parameter, equal to $\sigma_\mathcal{N}^2/\sigma_B^2$
46. ζ : stretching exponent of the stretched exponential *(e.g.,* $\exp(-x^\zeta)$)
47. η : inverse frequency power (or exponent) of the colored noise ($\Omega_\mathcal{N} \propto 1/f^\eta$)

48. θ : normalized time ($\theta = t/\tau_d$)
49. θ_t : normalized time step, equal to $\Delta t/\tau_d$
50. $\vec{\theta}$: vector of parameters for the chi-square map
51. ι : relative mean of the noise, equal to $\langle I_{\mathcal{N}} \rangle \,/\, \langle I_B \rangle$
52. λ : pulse asymmetry parameter, $\lambda = \tau_r/\tau_d$
53. σ_j : uncertainty associated with the j-th point of the chi-square analysis
54. σ_Φ : standard deviation of a time series or random variable $\Phi(t)$
55. τ : lag of the autocorrelation function
56. τ_d : characteristic duration of the pulses, $\tau_d = \tau_r + \tau_f$
57. τ_{ef} : effective pulse duration for stretched pulses, $\tau_{ef} = s_1\tau_d$
58. τ_f : characteristic duration of the pulses fall, $\tau_f = (1-\lambda)\tau_d$
59. τ_r : characteristic duration of the pulses rise, $\tau_r = \lambda\tau_d$
60. $\tau_{\mathcal{N}}$: correlation time of the Ornstein-Uhlenbeck noise
61. τ_{RQA} : delay of the embedding vectors of the RQA
62. τ_w : average waiting time between pulses, $\tau_w = T/N_p$
63. ϕ : potential applied in a Langmuir probe (only used in Section 3.1)
64. ϕ : waveform of the pulses
65. ϕ_B : waveform of the conditionally averaged burst
66. $\Phi(t)$: a time series and/or a random variable
67. $\langle \Phi \rangle$: average of a time series or random variable $\Phi(t)$
68. χ^2: the chi-square, defined as the sum of squared residuals
69. ω : angular velocity ($\omega = 2\pi f$)
70. Ω_Φ : power spectral density (PSD) of a time series or random variable $\Phi(t)$

Chapter 1

Introduction

Chapter 1 begins by explaining the main motivation to study magnetically confined plasmas—fusion energy. Later, it briefly describes what plasmas are and how they are confined in tokamaks. The TCABR tokamak is then presented. Later, the concept of plasma bursts, central for this work, is introduced and, subsequently, the stochastic model used for the dissertation is rapidly contextualized. At the end, the dissertation chapters and objectives are summarized.

1.1 Fusion Energy

Today, around 80% of the world's energy matrix comes from fossil fuels [1]. In 2019, natural gas, coal and oil were responsible respectively for 22%, 26% and 31% of the global energy consumption, as seen in Fig. 1.1. Nevertheless, a recent prediction [2] estimates that, if consumed with the annual production levels of 2015, coal reserves won't last more than 120 years, while the reserves of natural gas and oil won't last more than 60 years.

For this reason, alternative energy sources urge to be sought, and fusion is one promising option [3]. Nuclear fusion is how stars produce energy. On Earth, researchers aim to replicate this procedure with controlled thermonuclear fusion. It promises to be clean and safe, with resources that could last more than a thousand years [4, 5].

Fusion nuclear reactions occur by merging lighter nuclei into heavier ones, releasing energy within the process. On reactors, the least difficult way to achieve this is by converting deuterium and tritium into helium (Chapter 4 of Ref. 6). Deuterium (D or H^2) and tritium (T or H^3) are both hydrogen isotopes. The first contains one proton and one neutron in its nucleus, while the second, one proton and two neutrons. The reaction can be represented as below.

$$D + T \rightarrow \alpha + n + 17.6 \text{ MeV}$$

An α particle is a helium nucleus (it has two protons and two neutrons), while n stands for neutron. A portion of the reaction mass is converted to kinetic energy as predicted by Einstein's mass-energy equivalence, $E = mc^2$ (where c is the light-speed). This energy sums up to about 17.6 MeV and is mostly driven to the neutron (about 14.1 MeV). In envisioned fusion reactors, it is mainly the neutrons' energy that must be transformed into electrical energy to power cities (Section 2 of Ref. 7). In its turn, the principal role of the helium nuclei would be to heat the plasma (Section 1.5 of Ref. 7). A scheme of the fusion reaction

is depicted in Fig. 1.2.

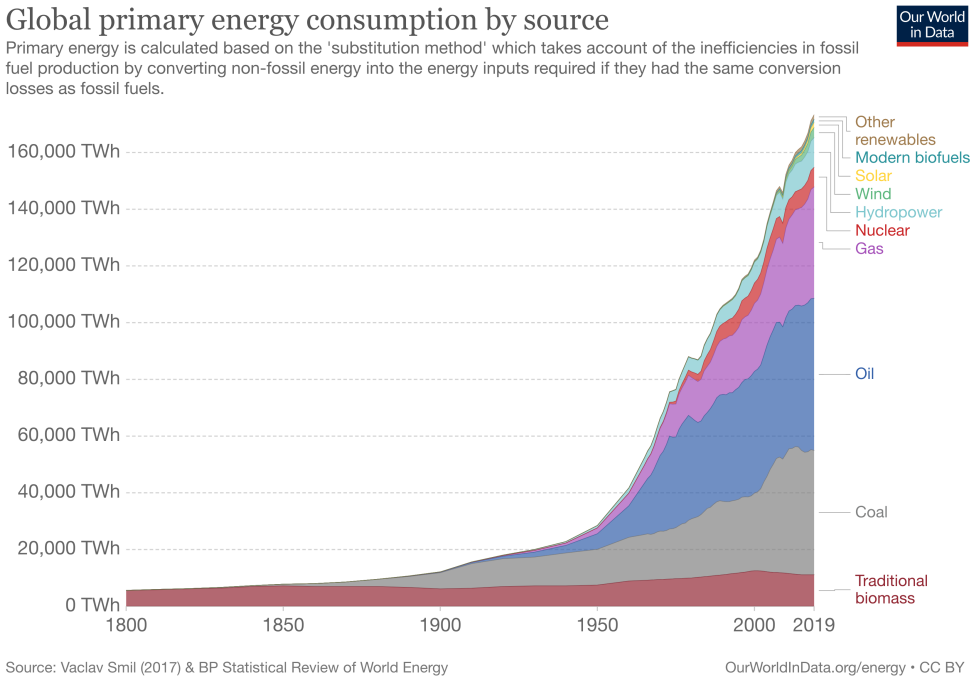


Figure 1.1: Global energy consumption by time, retrieved from Our World in Data [1]. See the original interative graph.

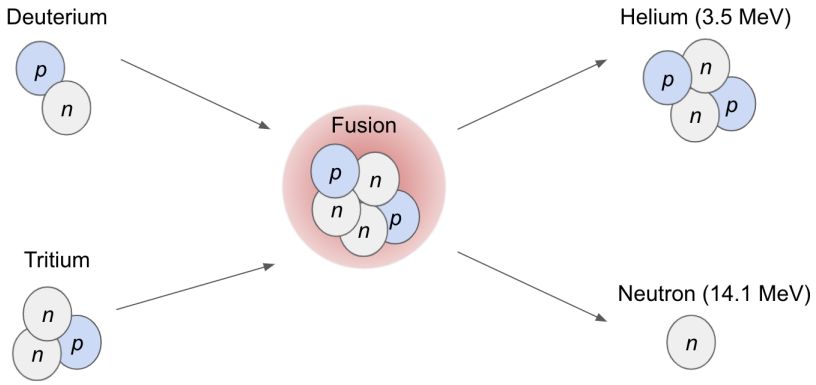


Figure 1.2: Scheme of a fusion reaction with deuterium and tritium resulting in a helium nucleus (with kinetic energy of 3.5 MeV) and a neutron (with kinetic energy of 14.1 MeV).

Achieving controlled thermonuclear fusion is among the greatest challenges of the century (introduction of Part I of Ref. 6). For this, a collaboration between the USA, Europe, Russia, China, India, South Korea and Japan is being made to develop the International Thermonuclear Experimental Reactor, or ITER, which is being built in the south of France [4]. ITER intends to be the first reactor with sufficient temperature to produce more energy than it

consumes. Also, in the second half of the 21st century, this partnership expects to build DEMO, which promises to be the first prototype fusion power plant to fuel an electrical network. There are also companies and public-private partnerships aiming to achieve controlled fusion. Some examples are Tri Alpha Energy, General Fusion and SPARC [8, 9].

1.2 Magnetically confined plasmas

This dissertation is inserted into the field of magnetically confined plasmas. Obtaining controlled thermonuclear fusion is the main motivation of this research area, as fusion reactions occur inside of plasmas. In this section, it will be presented the concept of plasmas and how they can be confined.

1.2.1 Plasmas

If a gas is heated enough, its electrons will have the energy to escape the electric attraction from their nuclei, ionizing the gas and transforming it into a plasma. However, not every ionized gas can be called a plasma, since in general any gas has at least a small degree of ionization.

To formalize the concept, a plasma is defined as *quasi-neutral* gas of ionized and neutral particles which exhibits electromagnetic collective behavior (Chapter 1 of Ref. 10). The meaning of "collective behavior" is what follows. In a gas of neutral particles, the main interactions are due to collisions, as external electromagnetic fields do not exert influence. In plasmas, on the other hand, free ions and electrons create electromagnetic fields that can influence other ionized particles, inducing a collective behavior in the system.

However, this influence decays with a characteristic distance, called the Debye length, λ_D. In a plasma, the potential of a particle with charge q can be written as (Eq. (2.18) of Chapter 11 of Ref. 11),

$$V(r) = \frac{1}{4\pi\epsilon_0}\frac{q}{r}\exp\left(-\frac{\sqrt{2}r}{\lambda_D}\right) \tag{1.1}$$

where the Debye length is

$$\lambda_D = \sqrt{\frac{\epsilon_0 k_B T}{n_e e^2}} \tag{1.2}$$

ϵ_0 is the vacuum permittivity, k_B is the Boltzmann constant, T is the plasma temperature (supposed approximately equal to electrons and ions), n_e is the electron density and e is its charge. From Eq. (1.1), one notes that the electric potential of a charge in a plasma decays faster than the Coulomb potential, $V_C(r) = q/(4\pi\epsilon_0 r)$. For $r \gg \lambda_D$, the potential is negligible. Consequently, plasmas act to shield charge fluctuations for $r > \lambda_D$, and particles interact mainly at a distance of $r < \lambda_D$. This is what is meant by *quasi-neutral* (*quasi* comes from Latin and means "as if"): for scales bigger than λ_D, the plasma acts as a neutral gas.

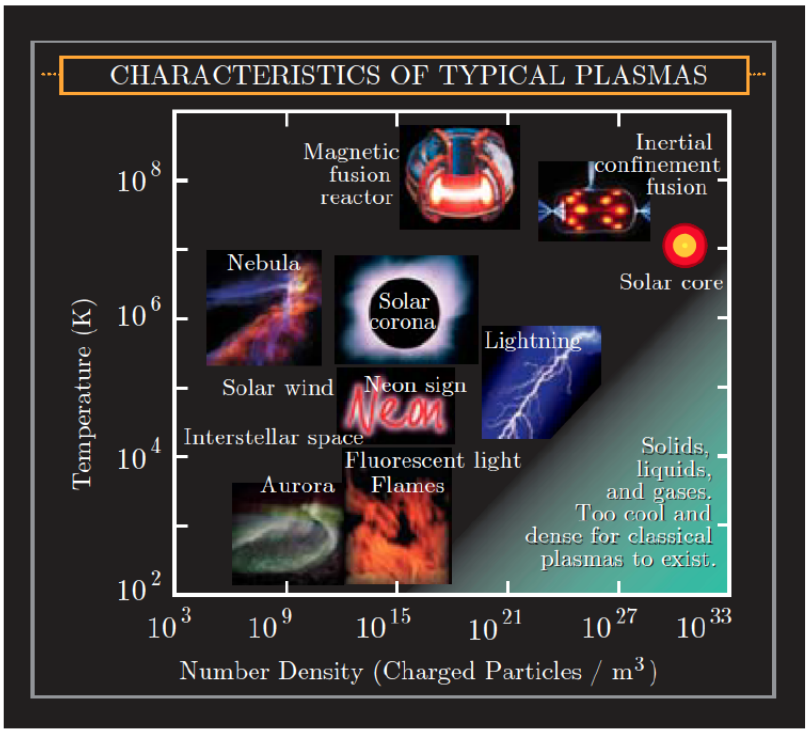


Figure 1.3: Temperature and density for various plasmas. Source: CPEP physics.

Formally, a plasma is defined by three conditions (Chapter 1 of Refs. 10 and 11):

1. The characteristic scale of the plasma system, L, has to be much larger than the Debye length, $L \gg \lambda_D$, so that the plasma can shield local electromagnetic fluctuations.

2. Since the shielding effect results from the collective behavior of various particles, the number of particles in a Debye sphere, $N_D = \frac{4\pi}{3}\lambda_D^3 n_e$, has to be much bigger than one, $N_D \gg 1$.

3. The characteristic interval of collisions between electrons and neutral particles, τ_{en}, needs to be much larger than the characteristic oscillation time τ_{pe} that the electrons take to respond to electromagnetic oscillations: $\tau_{en} \gg \tau_{pe}$. Otherwise, collisions could ruin the process of shielding.

The three above conditions for the existence of a plasma can be summarized as $L \gg \lambda_D$, $N_D \gg 1$ and $\tau_{en} \gg \tau_{pe}$. These parameters are functions of temperature and density. So plasmas can be characterized by T and n, as seen in Fig. 1.3.

1.2.2 Tokamaks and the TCABR

To achieve fusion on Earth, plasmas must be heated at temperatures of the order of 10^8 K, hotter than the solar core, $T_{sun} \sim 10^7$ K (as seen in Fig. 1.3). It is a Herculean task to confine such hot fluids, and the most used solution is by powerful magnetic fields (e.g., pp.

191–195 of Refs. 6). Due to the Lorentz magnetic force $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, ionized particles in the plasma follow a helical path along the magnetic field lines. This is illustrated in Fig. 1.4(a). However, this wouldn't prevent particles from reaching a wall at the end of a linear device. The key, then, is to connect the end of the machine to its beginning, constructing a toroid. A picture of a toroid can be seen in Fig. 1.4(b).

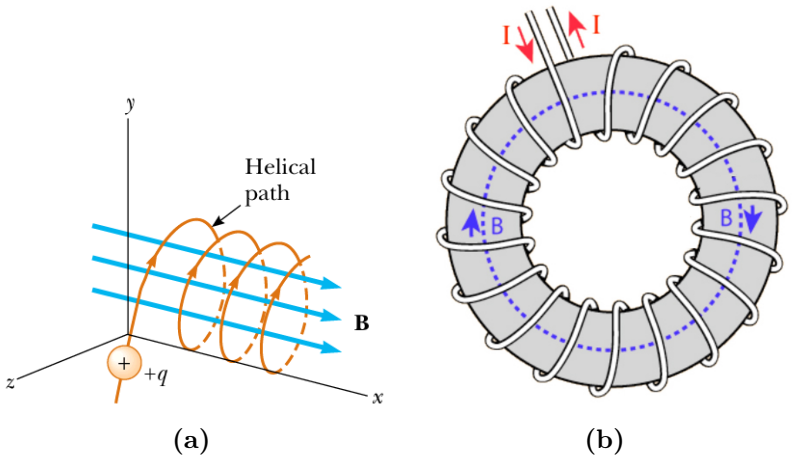


Figure 1.4: **(a)** A positive charge q following a helical path along the magnetic field $\mathbf{B}$. Source: Physics Exams. **(b)** A simple toroid for plasma confinement. A current I in the *poloidal* direction of the toroid creates a magnetic field in the *toroidal* direction, around which the charged particles circulate. Source: Chegg Study.

With this toroidal configuration, used in tokamaks, plasma particles are trapped in helicoidal motions around the magnetic field line, making loops in the toroidal direction.[1] Tokamak is a transliteration of the Russian word "токамак", and can be understood as a toroidal chamber with magnetic coils. These devices are widely used to confine plasmas (see for example cap. 4–6 of Ref. 6) and are seen as the leading candidates for a practical fusion reactor [4, 9].

The Institute of Physics of the University of São Paulo (USP) has a tokamak called TCABR [12, 13]. It was originally designed and built at the *École Polytechnique Fédérale de Lausanne* (EPFL), in Switzerland, and operated there from 1980 until 1992, under the name of TCA—a French acronym for *Tokamak Chauffage Alfvén*. The main objective of TCA was to investigate and enhance plasma heating with Alfvén waves [14]. A few years later, the device was transferred to USP, passing through an upgrade and adding *Brésilien* to its name. Its operation began in 1999. Fig. 1.5 shows a photograph of TCABR.

[1] Due to the existence of drifts and instabilities, corrections on the magnetic field must be made to maximize the confinement time in tokamaks. This includes using the own plasma current and poloidal coils to create desirable poloidal magnetic fields. Further information can be found for example in pp. 195–201 of Ref. 6, and a detailed description of tokamaks is given in Chapters 4–6 of the same reference.

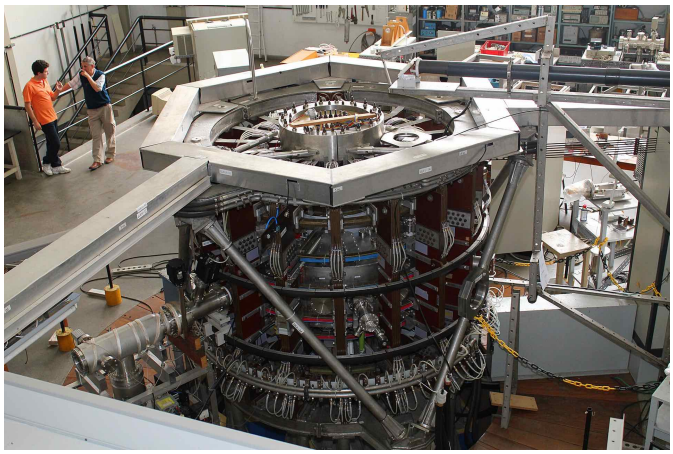


Figure 1.5: Photograph of the TCABR. Source: imagens.usp.br/?p=10222.

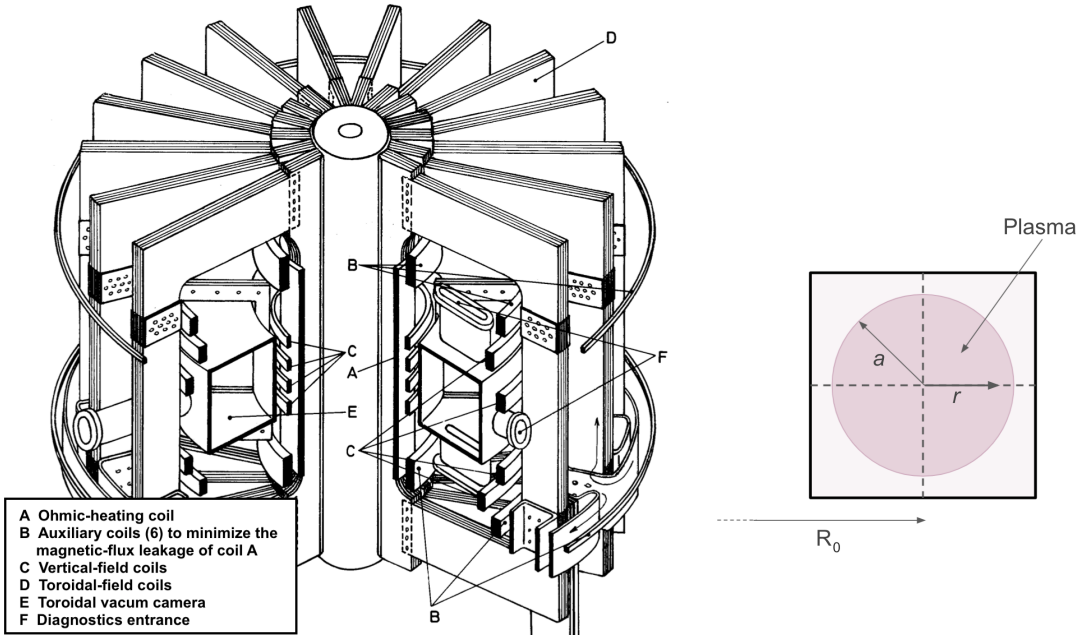


Figure 1.6: **(a)** Illustration of the TCABR tokamak. Coil A produces the toroidal electric field that ohmically heats the plasma. The D coils generate the toroidal magnetic field that confines the fluid. The plasma is kept inside the toroidal vacuum camera E. Source: adapted from Ref. 14. **(b)** Scheme of the TCABR poloidal cross-section, showing the major radius $R_0 = 61.5$ cm and the minor radius $a = 18.0$ cm.

The current purpose of the TCABR tokamak is not restricted to Alfvén heating. New objectives include (i) the characterization of magneto-hydro-dynamics (MHD) instabilities and (ii) the study of an improved confinement regime by edge bias polarization [12]. In the TCABR it is also possible to study the behavior of plasma turbulence at the edge and the scrape-off layer regions, which is the focus of this dissertation.

The TCABR plasma is made of hydrogen and has a circular poloidal section. It is in general ohmically heated, meaning that the plasma is heated by collisions due to the generalized Ohm's law (pp. 214–216 of Ref. 6, or Section 5.7 of Ref. 10). An external coil with varying current $i_{OHT}(t)$ induces an electric field in the toroidal direction, making a loop around the tokamak. This electric field then induces a plasma current I_p in the chamber and, due to the fluid collisional resistance, the current dissipates heat in the plasma. The plasma current in TCABR reaches up to $I_p \leq 100$ kA. Fig. 1.6 displays a scheme of the tokamak with its components, while Fig. 1.7 shows time series of the plasma current I_p, the loop voltage V_{loop} (associated with the toroidal electric field) and the current i_{OHT} in the coil for ohmic heating.

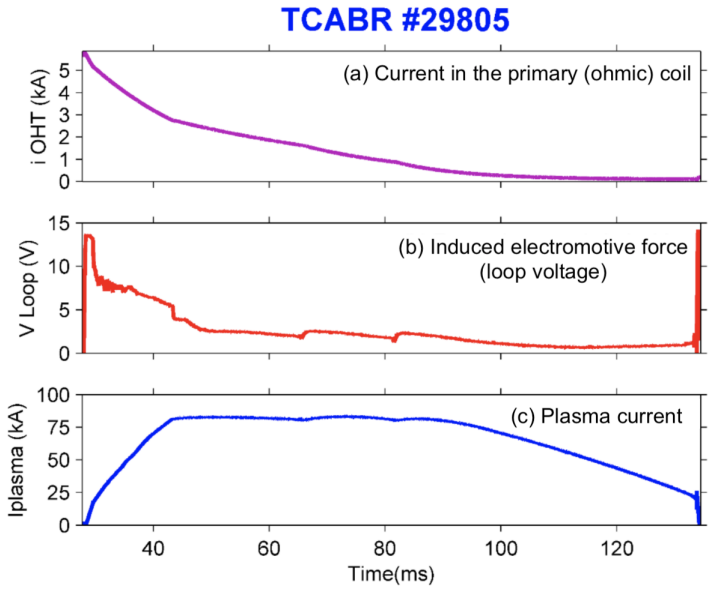


Figure 1.7: The first panel shows the current i_{OHT} of the TCABR ohmic-heating coil (A in Fig. 1.6). The temporal variation of i_{OHT} is responsible for the loop voltage arising around the tokamak (panel (b)), which in its turn causes the plasma current I_p shown in the panel (c). I_p then heats the plasma. Author of the figure: Z. O. Guimarães-Filho.

The TCABR central electron temperature is around $k_B T_{e0} \sim 0.5$ keV (i.e., $T_{e0} \sim 6{\times}10^6$ K) and its mean electron density is $0.9 \leq \bar{n}_{e0} \leq 3$, in units of 10^{19} m^{-3} [12]. Using Eq. (1.2), this electron temperature and the density $n_e = 2 \times 10^{19}$ m^{-3}, the Debye length results to be $\lambda_D \simeq 40$ μm. The TCABR minor radius for plasma confinement is $a = 18.0$ cm, meaning that $a \gg \lambda_D$, as required to an ionized gas to be a plasma (subsection 1.2.1). The major radius of TCABR is $R_0 = 61.5$ cm, giving an aspect ratio of $q = R_0/a = 3.4$.

Other parameters of TCABR include the toroidal magnetic field, $B_0 \sim 1.1$ T, the hydrogen filling pressure, $P_H \simeq 3 \times 10^{-4}$ Pa, a discharge duration of $T_D \simeq 100$ ms, and a steady regime duration up to $T \leq 60$ ms. All parameters are summarized in Table 1.1.

Table 1.1: TCABR parameters. Sources: Refs. 12, 15 and 16.

Parameter	Symbol	Value
Plasma format		Circular
Plasma composition	H	Hydrogen
Major radius	R_0	61.5 cm
Minor radius	a	18 cm
Aspect ratio (R_0/a)	q	3.4
Plasma current	I_p	$\leq$ 100 kA
Central (or toroidal) magnetic field	B_0	$\sim$ 1.1 T
Mean electronic density	$\overline{n}_{e0}$	$0.9 \sim 3 \times 10^{19}\mathrm{m}^{-3}$
Central electronic temperature	$k_B T_{e0}$	$\sim$ 0.5 keV
Hydrogen filling pressure	P_H	3×10^{-4} Pa
Discharge duration	T_D	$\simeq$ 100 ms
Steady regime duration	T	$\leq$ 60 ms

1.3 Turbulent transport and bursts

The first theoretical attempts to describe the transport of heat and particles in magnetized plasmas used cylindrical geometry. In these so-called classical models, it was assumed that diffusion (i.e. movement of single particles due to collisions) would be the principal cause of transport (see for ex. Section 4 of Ref. 17). However, the transport rate obtained in experiments was much greater than expected. Further models then considered the toroidal geometry of tokamaks, obtaining higher diffusion coefficients, and the corresponding description become known as the neoclassical theory for plasma transport [17]. Nevertheless, the neoclassical diffusion was still much smaller than the transport observed in experiments.

This gap was then credited to *anomalous* transport [17, 18]. Although a lot remains to be understood, further research found evidence that the anomalous transport is mainly convective (i.e., associated with coherent structures) and caused by turbulence—that is, due to non-linear micro and mesoscale fluctuations in particle density, temperature and electromagnetic fields [19–21].[2] Thus, another name for this anomalous phenomenon is "turbulent transport".

The confinement of plasmas in tokamaks is strongly affected by turbulent events that occur at the edge of the confinement region (Preface and Section 7.3 of Ref. 19, and also Refs. 22–25). The transport of ions resulting from this turbulence can also damage and erode the main chamber walls of fusion devices [19, 26]. Several machines share similar characteristics regarding these extreme events, indicating a universality of this type of phenomenon in magnetically confined plasmas [24, 25, 27].

Among the turbulent phenomena, bursts stand out. They are also called blobs, filaments and avaloids, and this name multiplicity will be addressed in a paragraph below. Bursts are

[2] For Ref. 19, see the Chapter 7, in special its abstract, Section 7.3 and Subsection 7.3.2.

density peaks with amplitude higher than the average of the signal in at various standard deviations.[3] Simulations and experiments also evidence that the high burst density is accompanied by high temperature [27–29]. Bursts are among the principal agents responsible for the deterioration of plasma confinement [24, 25]. Their shape is of filamentary coherent structures extended along magnetic field lines in the devices, as seen in Fig. 1.10. In tokamaks, these structures often propagate from the scrape off-layer (SOL) to the wall (e.g., Fig. 1.9), generating a considerable loss of particles that compromise the plasma confinement and its life-span. They have been found in various magnetically confined plasmas devices such as linear machines [24,25,27], tokamaks [24,25,27–38,40–44] (including TCABR [15,29,45,46]), stellarators [27,47] and in the Helimak [27,29,48,49].

Bursts [27–41,44,47] are also commonly called blobs [27–44,47], filaments [27,31–37,39–44, 47], or avaloids (as an avalanche reference) [24,25,27]. It is possible to identify a dimensional difference between the first three names. The density spikes in local measurements (Fig. 1.8) resemble explosions and, in this context, the term burst is fairly adequate. In its turn, we see a blob format in 2D graphs of these structures (Fig. 1.9). Lastly, 3D analysis of these events shows that they have a filamentary structure (Fig. 1.10). Thus, bursts, blobs and filaments refer to the same coherent structures, but each name is more appropriate in a different dimensional context. These nuances are used, for example, in [42, 43, 50]. In this dissertation, the term burst will be preferred, as here the analyzed signal is locally-measured. However, this paragraph highlights that all three names refer to the same structure.

Recently, a stochastic model was developed [50–55], aiming to describe statistical features of the edge plasma turbulence measured in a specific position. Therefore, in this model bursts are the main actors. This stochastic description has been tested and validated in various tokamaks around the globe, such as TCV [30,31], KSTAR [32], Alcator C-Mod [33–39], JET [40], COMPASS [41], and also in the Texas Helimak [48,49]. Three different density signals were analyzed: ion saturation current [30–32,34,38–40,48,49], gas puff imaging [33,35–37,39] and lithium-beam emission spectroscopy (Li-BES) [41]. In the present dissertation, this stochastic model will be adopted to describe turbulent fluctuations in the TCABR tokamak.

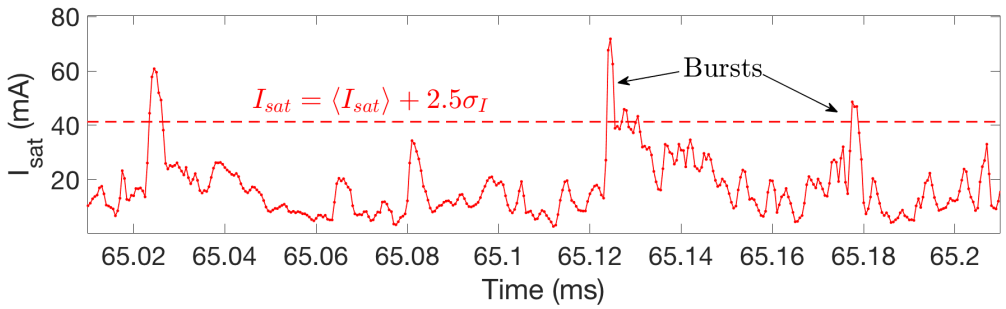


Figure 1.8: Excerpt of a saturation current (I_{sat}) measurement made in TCABR, showing three bursts (peaks bigger than the I_{sat} average in more than 2.5 standard deviations).

[3] The most common threshold is 2.5 standard deviations [30–35], but some works use 2 [48] or 3 [25].

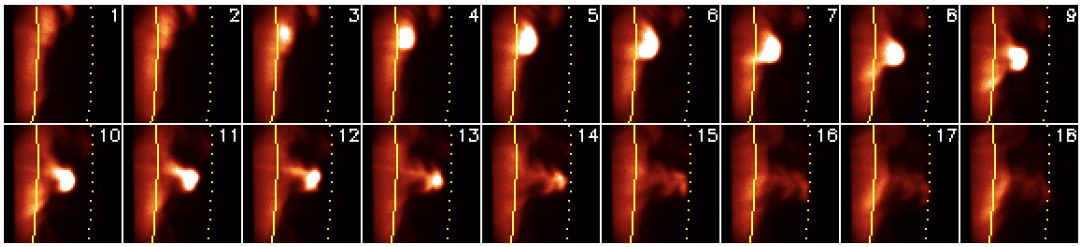
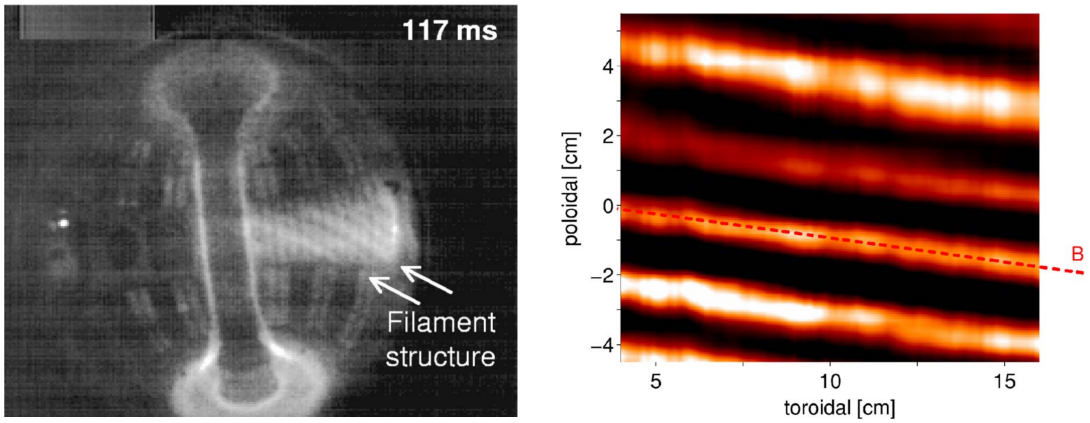


Figure 1.9: Gas puff imaging of a blob propagation in the NSXT late H-mode (inter-frame time of 7.0 μs). The solid line indicates the separatrix [59, 62].



(a) **(b)**

Figure 1.10: **(a)** Image of filaments in the NSTX tokamak edge, obtained with a 20 μs exposure photo [60]. **(b)** Poloidal-toroidal cross-section of filaments seen with D_α emissions in the Alcator C-Mod tokamak. A magnetic field line B is superimposed along the position of maximum D_α intensity [61, 62].

Eventually, it has been observed that high confinement modes (H-modes) in fusion plasmas decrease the amplitude of bursts [59]. Nevertheless, in late H-modes, blobs still play an important role and have not been fully understood [59] and filaments are also generated by edge-localized modes (ELMs) that can occur during H-modes [27]. Besides, recent works [36, 37] were successful in applying the stochastic model of Refs. 50–55 to discharges with high confinement regimes. This indicates that such stochastic model provides a powerful framework to describe turbulent density fluctuations in various scenarios of magnetized plasmas.

Although the stochastic description presented in Refs. 50–55 is phenomenological, its framework influenced further theoretical research [56–58]. Such works were able to explain some plasma turbulence properties with a first-principles model based on filament dynamics. The properties analyzed include the link between filament dynamics and density profiles in the scrape-off layer of magnetized plasmas.

The onset of blob-filaments still is a topic of active research. Ref. 63 showed in theory that filaments can be formed as a consequence of the nonlinear evolution of drift waves. A piece

of evidence in favor of this hypothesis was obtained in TORPEX, a low-temperature toroidal plasma device [64]. More specifically, blobs were observed detaching from interchange wave crests, which were sheared off by the motion of particles due to $\mathbf{E} \times \mathbf{B}$ drifts (Fig. 2a–c of Ref. 64).[4]

The mechanism of blob propagation is also understood and is supported by more experimental evidence (see for ex. Refs. 64 and 65 and Subsection 7.3.2 of Ref. 19). It is mainly ascribed to an inertial force effect, that is, the curvature of the toroidal magnetic field, which in its turn causes plasma polarization and a corresponding $\mathbf{E} \times \mathbf{B}$ drift. This last one then drives particles out of the plasma edge.

1.4 Objectives and chapter summary

This dissertation aims to apply the stochastic pulse train model (SPTM), deduced in Refs. 50–55, to the tokamak TCABR and compare different background choices for the SPTM, using a new fit developed here. Recent articles [48, 54] introduced the use of correlated Gaussian noise and pulse background to the SPTM. Here we will present an objective fit procedure to adjust the SPTM with noise, which can be used in other magnetic confinement machines. We will then apply the noise model to the TCABR tokamak and show evidence that the density fluctuations in this device are better described by a background with correlated noise than with white noise. We will also compare the model with noise to the one with a pulse background. Some evidence will indicate that the latter is better to describe the structure of density fluctuations. In particular, we will apply two diagnostics to the SPTM, the complexity-entropy diagram and the recurrence quantification analysis (RQA), which were not used in this context before, but can give important insights into the signal structure.

Chapter 2 will address the mathematical and statistical methods used in this dissertation to analyze turbulence data. In Chapter 3, it will be explained how local turbulent measurements were done with Langmuir probes. After, it will be presented an experimental characterization of turbulent density fluctuations in the TCABR tokamak. Chapter 4 will then describe the stochastic pulse train model with noise. Although Chapters 3 and 4 will already contain some authorial work, our main contributions start on Chapter 5, in which our method for fitting the SPTM with noise will be introduced. Chapter 6 then describes an application of this fit for the TCABR experiment 34132. Chapter 7 presents an improved version of the SPTM model considering stretched bursts and a pulse background, while Chapter 8 applies this new version to TCABR. Finally, Chapter 9 summarizes the results.

[4] For information about the curvature and $\mathbf{E} \times \mathbf{B}$ drifts, the reader is referred to Chapter 2 of Ref. 10.

Chapter 2

Methods for analyzing turbulence data

Turbulent signals often require careful statistical analysis. This chapter presents the main mathematical tools used in the dissertation.

2.1 Probability distributions

2.1.1 Definitions and moments

The probability density function (PDF) is used in countless fields, ranging from sociology to engineering and nature sciences [66, 67]. In the case of turbulent plasma fluctuations, the PDF standardized moments skewness and kurtosis have special roles (e.g. Ref. 50).

The PDF of a random continuous variable Φ measures the probability of Φ to be ranged in a certain interval. Let P be the PDF and Pr the probability. Then

$$\Pr\left(\Phi_1 \leq \Phi \leq \Phi_2\right) = \int_{\Phi_1}^{\Phi_2} P(\Phi) d\Phi$$

The PDF is normalized to one:

$$\int_{-\infty}^{\infty} P(\Phi) d\Phi = 1$$

The expected value of a function $f(\Phi)$ is defined by

$$\langle f(\Phi) \rangle \equiv \int_{-\infty}^{\infty} f(\Phi) P(\Phi) d\Phi$$

whereas the so called raw moments of the distribution are obtained with $f(\Phi) = \Phi^n$, $n \in \mathbb{N}^+$,

$$\langle \Phi^n \rangle \equiv \int_{-\infty}^{\infty} \Phi^n P(\Phi) d\Phi$$

The average value is the first moment,

$$\langle \Phi \rangle = \int_{-\infty}^{\infty} \Phi P(\Phi) d\Phi$$

Moreover, the central moments are defined as

$$\mu_n(\Phi) \equiv \langle (\Phi - \langle \Phi \rangle)^n \rangle = \int_{-\infty}^{\infty} (\Phi - \langle \Phi \rangle)^n P(\Phi) d\Phi$$

and hence the variance is the second central moment,

$$\sigma_\Phi^2 \equiv \left\langle \left(\Phi - \langle\Phi\rangle\right)^2 \right\rangle = \int_{-\infty}^{\infty} \left(\Phi - \langle\Phi\rangle\right)^2 P(\Phi) d\Phi \tag{2.1}$$

The variance measures how spread the distribution is from its average value. The positive square root of the variance is known as the standard deviation: $\sigma_\Phi = \sqrt{\sigma_\Phi^2}$.

Finally, the standardized moments are defined as

$$\frac{\mu_n}{\sigma_\Phi^n} = \frac{1}{\sigma_\Phi^n} \int_{-\infty}^{\infty} \left(\Phi - \langle\Phi\rangle\right)^n P(\Phi) d\Phi$$

The skewness and kurtosis of the distribution are the 3rd and 4th standardized moments:

$$S_\Phi = \frac{\mu_3}{\sigma_\Phi^3}, \quad K_\Phi = \frac{\mu_4}{\sigma_\Phi^4}$$

Some authors also use the excess kurtosis, $K'_\Phi = K_\Phi - 3$ [68].

The skewness measures how asymmetric the distribution is. It can be negative, positive or null. Usually, if the longest tail is on the left side of the distribution (with relation to its mode), then the skewness is negative: $S_\Phi < 0$. Meanwhile, if the longest tail is on the right side, then the skewness is usually positive: $S_\Phi > 0$. Eventually, a symmetric distribution (mode equal to mean and median) has null skewness: $S_\Phi = 0$.[1]

These three cases are exemplified in Fig. 2.1(a), using the beta distribution (e.g. p. 99 of Ref. 66 or Appendix D of Ref. 67):

$$P(\Phi; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \Phi^{\alpha-1} (1 - \Phi)^{\beta-1} \tag{2.2}$$

where $\Phi \in [0, 1]$, $\alpha, \beta > 0$ are called shape parameters and $\Gamma(\alpha)$ is the Gamma function (e.g. Appendix D of Ref. 67):

$$\Gamma(\alpha) \equiv \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \quad \mathrm{Re}(\alpha) > 0 \tag{2.3}$$

As shown in Ref. 69, turbulent density fluctuations in magnetically confined plasmas usually have positive skewness and of order of $S_\Phi \sim 1$. Similar results will be presented in Chapter 6 for the TCABR tokamak.

In its turn, the kurtosis measures the weight of outliers (or extreme events) [68].[2] Here

[1] These are the general ideas behind the skewness concept, but, if a tail happens to be fat, it can make the skewness behave otherwise. For example, a non-symmetric distribution can have null skewness if one tail is longer and the other is fat in such a way that both sides have the same weight. Multimodal distributions are other cases that do not obey the general rule of longest tails.

[2] In plasma physics, the term "flatness" is sometimes used to refer to kurtosis [24, 38, 50, 54, 69]. However, following Ref. 68, the terminologies of "peakedness" or "flatness" are inadequate to designate kurtosis, since distributions with a variety of peaks and degrees of flatness can have the same kurtosis (Fig. 2 of Ref. 68). In addition, distributions with very similar flattening can have astonishingly different kurtosis (Fig. 2 and 3 of Ref. 68).

outliers are understood as points distant from the distribution mean in several standard deviations.[3]

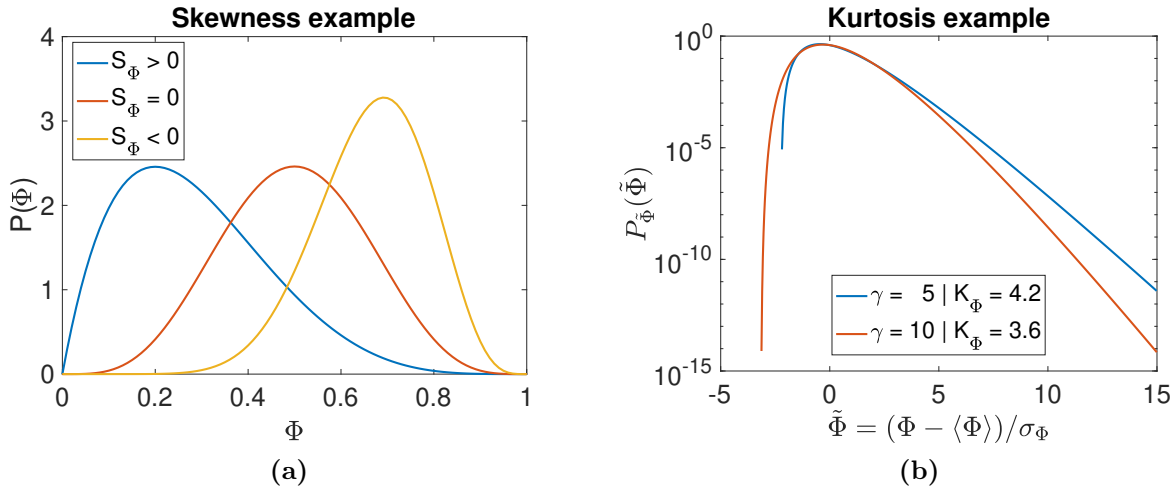


Figure 2.1: **(a)** Examples of positively, null and negatively skewed beta distributions. (Parameters (α, β) used: (2,5) in blue; (5,5) in orange; (10,5) in yellow.) **(b)** PDFs (in log scale) from two normalized Gamma distributions, respectively with $\gamma = 5$ (in blue) and $\gamma = 10$ (in orange).

The Gamma distribution, which will be of great importance for this dissertation (Chapters 4–9), can be used to exemplify the effect of the kurtosis. It is defined by (e.g. Ref. 50 or Appendix D of Ref. 67)

$$P_\Phi(\Phi;\gamma,A) = \frac{1}{\Gamma(\gamma)A}\left(\frac{\Phi}{A}\right)^{\gamma-1}\exp\left(-\frac{\Phi}{A}\right) \tag{2.4}$$

where $\Phi > 0$, $\Gamma(\gamma)$ is the Gamma function and γ and A are respectively the shape and scale parameters. The Gamma distribution has mean, standard deviation, skewness and kurtosis given by [55]

$$\begin{aligned} \langle\Phi\rangle &= A\gamma \\ \sigma_\Phi &= A\gamma^{1/2} \\ S_\Phi &= \frac{2}{\gamma^{1/2}} \\ K_\Phi &= 3 + \frac{6}{\gamma} \end{aligned} \tag{2.5}$$

[3] As a counterexample, a distribution like $P(\Phi;\alpha) = 3/(2a^3)\Phi^2$, defined for $\Phi \in [-a, a]$, would have no outliers, since it has standard deviation $\sigma_\Phi \simeq 0.8a$. That is, this distribution does not have points greater than $a = a\sigma_\Phi/(0.8a) = 1.25\sigma_\Phi$ and therefore there are no Φ distant from the mean $\langle\Phi\rangle = 0$ in several standard deviations. Thus, it is of no surprise that its kurtosis is small, $K_\Phi = 1.25$, closest to the smallest possible value, $K_\Phi = 1$.

The excess kurtosis $K_\Phi - 3$ is inversely proportional to γ and is independent of A.

When comparing different kurtosises, it is especially helpful to consider normalized variables with null mean and unity standard deviation,

$$\tilde{\Phi} = \frac{\Phi - \langle\Phi\rangle}{\sigma_\Phi}$$

A point $\tilde{\Phi} = 10$ is for example 10 standard deviations distant from the mean. With such notation, the kurtosis can be written as $K_\Phi = \left\langle \tilde{\Phi}^4 \right\rangle$. In its turn, the distribution of $\tilde{\Phi}$ is obtained by the change of variables

$$P_{\tilde{\Phi}}(\tilde{\Phi}) = \left| \frac{d\Phi}{d\tilde{\Phi}} \right| P_\Phi(\Phi(\tilde{\Phi}))$$

It is easy to show that, for a Gamma PDF (Eq. (2.4)), the corresponding $P_{\tilde{\Phi}}(\tilde{\Phi})$ is independent of A:

$$P_{\tilde{\Phi}}(\tilde{\Phi}) = \frac{\gamma^{1/2}}{\Gamma(\gamma)} \left(\gamma^{1/2}\tilde{\Phi} + \gamma\right)^{\gamma-1} \exp\left(-\gamma^{1/2}\tilde{\Phi} - \gamma\right)$$

Fig. 2.1(b) illustrates two normalized Gamma distributions $P_{\tilde{\Phi}}(\tilde{\Phi})$ with different γ. The PDFs in blue and orange have shape parameters $\gamma = 5$ and $\gamma = 10$, respectively. As seen considering the left tails, the blue curve produces outliers with a higher probability and hence has a greater kurtosis than the orange distribution.

The kurtosis of any random variable is always greater than the squared skewness by at least 1 (proof in pp. 432–433 of Ref. 70):

$$K_\Phi \geq 1 + S_\Phi^2$$

Thus, the smallest possible value for the kurtosis is $K_\Phi = 1$. In turbulent density fluctuations of magnetized plasmas, usually the following approximate parabolic relation is obeyed [27, 31–34, 36, 50, 51, 69, 71–73]:

$$K_\Phi \simeq 3 + \frac{3}{2} S_\Phi^2$$

From Eq. (2.5), it is easy to derive that the Gamma distribution exactly follows the relation $K_\Phi = 3 + 1.5 S_\Phi^2$.

To finish the discussion, let us highlight the differences between the moments analyzed in this section. The average is the 1st raw moment of a distribution, $\langle\Phi\rangle$, while the variance is the 2nd centralized moment, $\sigma_\Phi^2 = \mu_2 = \left\langle (\Phi - \langle\Phi\rangle)^2 \right\rangle$. Moreover, the skewness and kurtosis are respectively the 3rd and 4th standardized moments, $S_\Phi = \mu_3/\sigma_\Phi^3$ and $K_\Phi = \mu_4/\sigma_\Phi^4$. For the sake of simplicity, in this dissertation these four statistical indicators will be generically called "moments", following previous works [32, 50, 52, 54, 55]. When needed, the adjectives raw, centralized and standardized will be used to avoid ambiguity.

2.1.2 Cumulative distribution functions (CDF)

The CDF $\mathcal{P}_\Phi(\Phi')$ of a random variable Φ, evaluated at a value Φ', is the probability that Φ assumes values smaller or equal than Φ':

$$\mathcal{P}_\Phi(\Phi') = \Pr(\Phi \leq \Phi')$$

The probability that Φ is measured between Φ_1 and Φ_2 is then

$$\Pr(\Phi_1 \leq \Phi \leq \Phi_2) = \mathcal{P}_\Phi(\Phi_2) - \mathcal{P}_\Phi(\Phi_1)$$

Remembering that this same probability is given in terms of the PDF $P_\Phi(\Phi)$ by the following integral,

$$\Pr\left(\Phi_1 \leq \Phi \leq \Phi_2\right) = \int_{\Phi_1}^{\Phi_2} P_\Phi(\Phi) d\Phi$$

then we conclude by the fundamental theorem of calculus that the PDF is the CDF derivative,

$$\mathcal{P}_\Phi(\Phi_2) - \mathcal{P}_\Phi(\Phi_1) = \int_{\Phi_1}^{\Phi_2} P_\Phi(\Phi) d\Phi \implies$$

$$P(\Phi) = \frac{d\mathcal{P}}{d\Phi}(\Phi) \tag{2.6}$$

As it will be shown in Subsection 2.2.2, this expression is useful to deduce the formula for the empirical characteristic function (ECF). The ECF, in its turn, is important to fit turbulence data in plasmas [55].

2.1.3 Convolution of PDFs

Let Φ be given by a sum of the two other random variables X and Y, with distributions P_X and P_Y:

$$\Phi = X + Y$$

Then the PDF of Φ is the convolution of P_X and P_Y, which is defined by

$$P_\Phi(\Phi) = (P_X * P_Y)(\Phi) \equiv \int_{-\infty}^{\infty} P_X(\Phi - \Psi) P_Y(\Psi) d\Psi$$

Even if $P_X(X)$ and $P_Y(Y)$ can be expressed by elementary functions, the convolution $(P_X * P_Y)(\Phi)$ may not have a closed-form.

Let us consider one example. As it will be described in Chapter 4, local measurements of plasma density fluctuations can be modeled by a sum of pulses that follow a Gamma distribution plus a Gaussian noise [32, 34, 37, 41, 48, 54, 55]. As seen in Subsection 2.1.1, the PDF of the Gamma distribution is given by (e.g. Appendix D of Ref. 67)

$$P_X(X) = \frac{1}{\Gamma(\gamma) A} \left(\frac{X}{A}\right)^{\gamma-1} \exp\left(-\frac{X}{A}\right) \tag{2.7}$$

where $X \geq 0$, $\gamma > 0$ and $A > 0$ are respectively the distribution shape and scale parameters, and $\Gamma(\gamma)$ is the Gamma function (Eq. (2.3)). In its turn, the PDF of Gaussian data is (Appendix D of Ref. 67)

$$P_Y(Y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(Y-\mu)^2}{2\sigma^2}\right] \tag{2.8}$$

where $\mu \in \mathbb{R}$ is the distribution mean and $\sigma > 0$ is its standard deviation. The convolution of P_X and P_Y can not be expressed in terms of elementary functions. When the Gaussian noise mean is zero (i.e., $\mu = 0$), then the PDF assumes the following intricate form (Eq. (A6) of Ref. 54)

$$\begin{aligned} P_\Phi(\Phi) = {} & \frac{(\gamma\epsilon)^{\gamma/2-1}}{2^{\gamma/2}\langle A\rangle} \times \exp\left(-\frac{\Phi^2}{2\gamma\epsilon A^2}\right) \times \left\{ \frac{(\gamma\epsilon)^{1/2}}{2^{1/2}\Gamma\left[(1+\gamma)/2\right]} \right. \\ & \times M\left(\frac{\gamma}{2}, \frac{1}{2}; \frac{1}{2\epsilon}\left(\frac{\Phi}{\gamma^{1/2}A} - \gamma^{1/2}\epsilon\right)^2\right) \\ & + \frac{\gamma^{1/2}}{\Gamma(\gamma/2)}\left(\frac{\Phi}{\gamma^{1/2}A} - \gamma^{1/2}\epsilon\right) \\ & \left. \times\, M\left(\frac{1+\gamma}{2}, \frac{3}{2}; \frac{1}{2\epsilon}\left(\frac{I_{sat}}{\gamma^{1/2}\langle A\rangle} - \gamma^{1/2}\epsilon\right)^2\right)\right\} \end{aligned} \tag{2.9}$$

where ϵ is the ratio between the variances from the noise and the Gamma process, $\epsilon = \sigma^2/(\gamma A^2)$, and $M(a, b; z)$ is the confluent hypergeometric function of the first kind, for parameters a and b and argument z [74]. In general, $M(a, b; z)$ can not be written in terms of elementary functions.

Even though the convolution of two distributions often results in very complicated expressions like Eq. (2.9), there is another function that carries the same information as the PDF and has simpler expressions for the sum of random variables. This is the characteristic function (CF) and is the theme of the next section.

2.2 Characteristic functions of PDFs

2.2.1 Definition and properties

Let Φ be a random variable with a probability distribution $P(\Phi)$. The characteristic function (CF) of Φ is defined by (see for example Chapter 4 of Ref. 75)

$$C_\Phi(u) = \mathrm{E}\left[e^{iu\Phi}\right] = \int_{-\infty}^{\infty} e^{iu\Phi} P(\Phi) d\Phi \tag{2.10}$$

where i is the imaginary unit and $u \in \mathbb{R}$ is the CF variable. The CF determines the same information of the PDF. This affirmation can be explained in terms of the characteristic

function logarithm, which is a moment generating function (Section 4.6 of Ref. 75),

$$\ln C_\Phi(u) = \sum_{n=1}^{\infty} \kappa_n \frac{(iu)^n}{n!} \tag{2.11}$$

where κ_n are the n-th cumulants of the PDF. The cumulants are cumulative (hence the name). That is, if Φ is the sum of random independent variables $\Phi = \Phi_1 + \Phi_2 + ... + \Phi_M$, then its cumulants are

$$\kappa_n = \sum_{m=1}^{M} \kappa_n^{(m)} \tag{2.12}$$

where $\kappa_n^{(m)}$ is the n-th cumulant of Φ_m. There is a correspondence between the cumulants and the distribution moments. For example, the mean, variance, skewness and kurtosis of Φ can be written as

$$\begin{aligned} \langle \Phi \rangle &= \kappa_1 \\ \sigma_\Phi^2 &= \kappa_2 \\ S_\Phi &= \frac{\kappa_3}{\kappa_2^{3/2}} \\ K_\Phi &= 3 + \frac{\kappa_4}{\kappa_2^2} \end{aligned} \tag{2.13}$$

The same can be done to other moments of the distribution. Thus, Eq. (2.11) shows that the characteristic function completely determines the PDF moments and therefore its information.

Nevertheless, in a lot of cases the PDF has no closed-form, whereas the CF is usually much simpler and can be put in terms of elementary functions. Therefore the characteristic function is a good candidate to fit data and it has already been used in the context of plasma density fluctuations [55].

Other properties of the CF are also very useful for fitting data and will now be presented. From Eq. (2.10) it is also seen that, by construction, $C_\Phi(0) = 1$. Hence the point at the origin carries no statistical information, since its standard deviation is null.

Also, from Eq. (2.10) it is noticeable that the CF is complex and, for real signals,

$$C_\Phi(u) = C_\Phi(-u)^*$$

where the star means the complex conjugate. Separating $C_\Phi(u)$ in real and imaginary parts yields

$$\begin{aligned} \text{Re}\, C_\Phi(u) &= \text{Re}\, C_\Phi(-u), \\ \text{Im}\, C_\Phi(u) &= -\text{Im}\, C_\Phi(-u) \end{aligned}$$

Thus the CF real and imaginary parts are respectively symmetric and anti-symmetric. Consequently, to fit real data, only $u > 0$ needs to be considered (as done in Refs. 55 and 76).

Some characteristic functions that are important for this dissertation are the ones from Gamma and Gaussian distributions (Eq. (2.7) and (2.8)). The Gamma characteristic function is (Table 4.2 of Ref. 75)

$$C_\Phi(u) = (1 - iAu)^{-\gamma}$$

where, again, $\gamma > 0$ and $A > 0$ are respectively the shape and scale parameters of the distribution. The Gaussian CF is (as in Table 4.2 of Ref. 75)

$$C_\Phi(u) = \exp\left(i\mu u - \frac{1}{2}\sigma^2 u^2\right)$$

where $\mu \in \mathbb{R}$ is the distribution mean and $\sigma > 0$ is its standard deviation.

The CF of a sum of random variables is the product of the original CFs [75]. That is,

$$\Phi = X + Y \implies$$
$$C_\Phi(u) = C_X(u)C_Y(u)$$

Therefore the CF for a sum of Gamma and Gaussian random variables is

$$C_\Phi(u) = (1 - iAu)^{-\gamma} \exp\left(i\mu u - \frac{1}{2}\sigma^2 u^2\right)$$

which is simpler than the PDF of $X + Y$ (Eq. (2.9)).

2.2.2 The empirical characteristic function

When treating experimental data, it is important to evaluate an estimative of the CF. This is normally done using the Empirical Characteristic Function (ECF) (well explained in Ref. 77 and also present in Refs. 55 and 76). The definition of the CF of a random variable Φ is given by Eq. (2.10),

$$C_\Phi(u) = \int_{-\infty}^{\infty} e^{iu\Phi} P(\Phi) d\Phi$$

Using that the PDF is the derivative of the commutative distribution function (Eq. (2.6)),

$$P(\Phi) = \frac{d\mathcal{P}}{d\Phi}(\Phi),$$

then it is also possible to write the CF in terms of the CDF,

$$C_\Phi(u) = \int_0^1 e^{iu\Phi} d\mathcal{P}(\Phi) \tag{2.14}$$

As seen on Ref. 77, in a frequentist description the empirical CDF is simply given by $\mathcal{P}_N(\phi) = n(\phi)/N$, where N is the total number of data points and $n(\phi)$ is the number of Φ_j such that $\Phi_j < \phi$, with $1 \leq j \leq N$. Therefore, in this discrete formulation, $\Delta\mathcal{P}_N(\Phi) = 1/N$, and the integral (2.14) can be approximated by the sum

$$C_N(u) = \frac{1}{N}\sum_{j=1}^{N} e^{iu\Phi_j}$$

This is the empirical characteristic function (ECF), which can be used to fit plasma density fluctuations [55].

2.3 Covariance and correlation

While the variance of a random variable is defined by Eq. (2.1),

$$\sigma_\Phi^2 \equiv \left\langle \left(\Phi - \langle\Phi\rangle\right)^2 \right\rangle = \int_{-\infty}^{\infty} \left(\Phi - \langle\Phi\rangle\right)^2 P(\Phi) d\Phi$$

the *covariance* between Φ and another random variable Ψ is, as seen for ex. in Section 4.3 of Ref. 66,

$$\mathrm{cov}(\Phi, \Psi) \equiv \left\langle \left(\Phi - \langle\Phi\rangle\right)\left(\Psi - \langle\Psi\rangle\right) \right\rangle$$

This means that

$$\left\langle \left(\Phi - \langle\Phi\rangle\right)\left(\Psi - \langle\Psi\rangle\right) \right\rangle = \int_{-\infty}^{\infty} \left(\Phi - \langle\Phi\rangle\right)\left(\Psi - \langle\Psi\rangle\right) P_{\Phi\Psi}(\Phi, \Psi) d\Phi d\Psi$$

where $P_{\Phi\Psi}(\Phi, \Psi)$ is the joint probability of Φ and Ψ (see for ex. Chapter 5 of Ref. 67). The covariance can be re-written as

$$\mathrm{cov}(\Phi, \Psi) = \langle\Phi\Psi\rangle - \langle\Phi\rangle\langle\Psi\rangle$$

If the variables are independent, $\langle\Phi\Psi\rangle = \langle\Phi\rangle\langle\Psi\rangle$, and then the covariance is null.

The correlation coefficient (or simply correlation) of Φ and Ψ is

$$\rho(\Phi, \Psi) \equiv \frac{\mathrm{cov}(\Phi, \Psi)}{\sigma_\Phi \sigma_\Psi}$$

ρ is bounded between -1 and 1 (theorem 5.4 of Ref. 67). If $\rho \approx 1$, the variables are strongly correlated, meaning that if Φ varies positively, Ψ tends to do the same. If $\rho \approx -1$, the variables are strongly anti-correlated and they tend to have opposite behaviors (when one increases, the other tends to decrease).

If Φ is measured N times in an experiment (or simulation), then its variance can be estimated as

$$\sigma_\Phi^2 = \frac{1}{N-1} \sum_{j=1}^{N} \left(\Phi_j - \overline{\Phi}\right)^2$$

where the sample average is

$$\overline{\Phi} = \frac{1}{N} \sum_{j=1}^{N} \Phi_j$$

The sample covariance between Φ and Ψ is, by its turn,

$$\mathrm{cov}(\Phi, \Psi) = \frac{1}{N-1} \sum_{j=1}^{N} \left(\Phi_j - \overline{\Phi}\right)\left(\Psi_j - \overline{\Phi}\right)$$

A correlation matrix between the points of a random-valued vector $\mathbf{x} = (x_1, x_2, ..., x_N)$ can be constructed with elements

$$\rho_{j,k} = \frac{\mathrm{cov}(\Phi_j, \Psi_k)}{\sigma_{\Phi_j} \sigma_{\Psi_k}} \tag{2.15}$$

for $j, k \in [1, N]$. The diagonal of such matrix is 1 by definition. The correlation matrix will be used for example in Sections 5.2 and 5.6.

2.4 Fourier transforms and power spectra

In general, a Fourier transform (FT) can be understood as a mathematical transformation that decomposes space or time functions into spatial or temporal frequency functions. In this dissertation, time series will be analyzed.

The Fourier transform of a temporal signal $\Phi(t)$ is given by (see for example Chapter 2 of Ref. 78, or 1 of Ref. 79)[4]

$$\hat{\Phi}(f) = \int_{-\infty}^{\infty} \Phi(t) e^{-2\pi i f t} dt$$

whereas its inverse is

$$\Phi(t) = \int_{-\infty}^{\infty} \hat{\Phi}(f) e^{2\pi i f t} df$$

$i = \sqrt{-1}$ is the imaginary unit and f is the frequency. When only positive time is defined, it is helpful to use the one-sided Fourier transform,

$$\hat{\Phi}(f) = \int_{0}^{\infty} \Phi(t) e^{-2\pi i f t} dt$$

Experimental data are in general discrete. Thus it is important to define a discrete Fourier transform (DFT) and its inverse (e.g. Chapter 2 of Ref. 78 or 1 of Ref. 79),[5]

$$\hat{\Phi}_j = \frac{1}{N} \sum_{n=0}^{N-1} \Phi_n \exp\left(-i \frac{2\pi j n}{N}\right), \quad j = 0, 1, 2, \ldots, N-1 \tag{2.16}$$

$$\Phi_n = \sum_{j=0}^{N-1} \hat{\Phi}_j \exp\left(i \frac{2\pi j n}{N}\right), \quad n = 0, 1, 2, \ldots, N-1 \tag{2.17}$$

where j is the index for frequency, n for time and N is the total number of points. The function $\Phi_n = \Phi(t_n)$ is thus decomposed into a linear combination of exponentials $\exp(-i2\pi jn/N)$ with coefficients $\hat{\Phi}_j = \hat{\Phi}(f_j)$.

To express the time t in terms of n, one writes

$$t_n = n\Delta t = n \frac{T}{N-1}, \quad n = 0, 1, 2, \ldots, N-1$$

[4] From Eq. (2.10), one notes that the characteristic function is the Fourier transform of the PDF.

[5] Some authors define the discrete Fourier transform as

$$\hat{\Phi}_j = \frac{1}{N} \sum_{m=-N/2+1}^{N/2} \Phi_m \exp\left(-i \frac{2\pi j m}{N}\right), \quad j = -N/2+1, -N/2+2, \ldots, N/2$$

Nevertheless, this notation is equivalent to Eq. (2.16). Also, Eq. (2.16) holds for complex Φ. For details see Sections 2.3 and 3.1 of Ref. 78.

where $\Delta t = T/(N-1)$ is the time interval between points and $T = (N-1)\Delta t$ is the total time analyzed.[6] In its turn, to express the frequency f in terms of the index j, one writes

$$f_j = \frac{j}{(N-1)\Delta t} = \frac{j}{T}, \quad j = 0, 1, 2..., N-1$$

Another important quantity based on frequency is the power spectral density (PSD, as in chapter 10 of Ref. 67), or simply power spectrum,

$$\Omega_\Phi(f) = \lim_{T\to\infty} \frac{1}{2T}\hat{\Phi}(f)\cdot\hat{\Phi}^*(f)$$

where $\hat{\Phi}^*(f)$ is the complex conjugate of $\hat{\Phi}(f)$. $\Omega_\Phi(f)$ is a non-negative real function, whereas the Fourier transform $\hat{\Phi}(f)$ is a complex one. Sometimes it is practical to write the Fourier transform and the PSD as a function of the angular velocity $\omega = 2\pi f$.

Spikes and hills in the graph of Ω_Φ indicate the prevalence of frequencies on the signal. For example, the discrete power spectral density of a sine function $s(t) = A\sin(2\pi bt)$ is

$$\Omega_s(f) = \frac{1}{4}A^2\left(\delta_{f,\,b} + \delta_{f,\,f_{max}-b}\right) \tag{2.18}$$

where $A \in \mathbb{R}$, $f_{max} = 1/\Delta t$ and $\delta_{f,b}$ is the Kronecker delta function:

$$\delta_{f,b} = \begin{cases} 1, & f = b \\ 0, & f \neq b \end{cases}$$

From Eq. (2.18) follows that a finite linear combination of sines must show spikes in its PSD graph. For example, if

$$\Phi(t) = \sin(2\pi\cdot 50t) + 2\sin(2\pi\cdot 80t) \tag{2.19}$$

then the PSD is

$$\Omega_\Phi(f) = \frac{1}{4}\left(\delta_{f,\,50} + \delta_{f,\,f_{max}-50}\right) + \left(\delta_{f,\,80} + \delta_{f,\,f_{max}-80}\right)$$

If we choose $\Delta t = 0.005$, then $f_{max} = 1/\Delta t = 200$. Fig. 2.2 presents the discrete PSD of Φ. Four spikes can be seen, at $f = 50$, $f = 80$, $f = 120$ and $f = 150$. From Eq. (2.18), it is clear that the last two are just mirrored versions of the first two frequencies, with the symmetry axis located at the so-called Nyquist frequency, $f_{\mathrm{Ny}} = f_{max}/2 = 1/(2\Delta t)$. In fact, phenomena with frequencies higher than f_{Ny} cannot be properly detected (pp. 95-97 of Ref. 78). Thus, for practical reasons, the region with $1/(2\Delta t) < f < 1/(\Delta t)$ is usually omitted in PSD graphs of real data.

[6] As an example, in the TCABR tokamak the time series of plasma steady fluctuations have: $\Delta t = 0.5\,\mu$s and $N < 12\cdot 10^4 + 1$ for the total number of points (that is, $T = (N-1)\Delta t < 60$ ms).

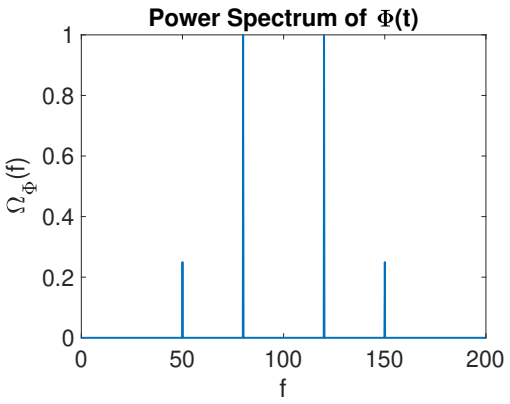


Figure 2.2: Power spectrum of Eq. (2.19), with $\Delta t = 0.005$. The spikes indicate that the signal is dominated by frequencies $f = 50$ and $f = 80$, while $f = 120$ and $f = 150$ are "mirrored frequencies". Inspiration of the example: MathWorks® fft documentation.

Sampling the PSD can create an effect known as spectral leakage (see pp. 98-99 of Ref. 78, or Ref. 81), which distorts the power spectrum. To avoid this, the signal is divided into segments and smoothing windows are applied [81–83]. Appendix A.1 details this process.

The power spectral density is a powerful tool to analyze plasma turbulent signals. With the PSD, it is possible to: identify unexpected phenomena in the frequency domain [48], estimate parameters from models [48, 80] and compare simulations and experiments [35, 48].

2.5 Autocorrelation function

The autocorrelation of a complex function Φ can be defined as (e.g. p. 397 of Ref. 84)

$$R_\Phi(\tau) = \langle \Phi(t+\tau)\Phi^*(t)\rangle = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{+T} \Phi(t+\tau)\Phi^*(t)dt$$

It is clear that the autocorrelation of a real function is even—that is, $R_\Phi(\tau) = R_\Phi(-\tau)$. τ is called the lag of the autocorrelation. When $\tau = 0$, then $R_\Phi(0) = \langle \Phi^2\rangle$. For this reason, sometimes the autocorrelation is normalized by $\langle \Phi^2\rangle$ so that $R_\Phi(0) = 1$.[7] The sample autocorrelation function is (section 4.2 of Ref. 85)

$$R_\Phi(\tau) = \begin{cases} \frac{1}{N_{pt}-\tau}\sum_{n=1}^{N_{pt}-\tau} \Phi_{n+\tau}\Phi_n^* & \text{for } \tau = 0, 1, ..., (N_{pt}-1) \\ \frac{1}{N_{pt}-\tau}\sum_{n=1-\tau}^{N_{pt}} \Phi_{n+\tau}\Phi_n^* & \text{for } \tau = -1, -2, ..., -(N_{pt}-1) \end{cases} \tag{2.20}$$

where N_{pt} is the total number of signal points. The power spectral density (PSD) is given by the Fourier transform of the autocorrelation (see for example Ref. 53 or pp. 395–398 of

[7] MATLAB has an in-built function that calculates the sample autocorrelation, called xcorr. By default it uses the normalization $R_\Phi(0) = 1$.

Ref. 84),

$$\Omega_\Phi(\omega) = \int_{-\infty}^{\infty} R_\Phi(\tau) e^{-i\omega\tau} d\tau \tag{2.21}$$

The autocorrelation function can be used to deduce the PSD formula (as in Ref. 53), and can describe some of the signal properties. For example, different colors of Gaussian noise (Subsection 2.6.1) have the same PDF, but different PSD and autocorrelation.

The autocorrelation can be generalized to analyze two signals at the same time, in which case it is called the cross-correlation or cross-covariance (see Ref. 86 and section 4.2 of Ref. 85). Also, some authors [85, 86] define the autocorrelation so that $R_\Phi(0) = 1$ and the terms are calculated with respect to $(\Phi_{n+\tau} - \langle\Phi\rangle)(\Phi^* - \langle\Phi\rangle^*)$ instead of $\Phi_{n+\tau}\Phi_n^*$. We choose to follow the definitions used in our field [53, 54].

The power spectrum of the sum of independent real variables can be obtained via the autocorrelation. Let

$$\Phi(t) = X(t) + Y(t)$$

Then

$$R_\Phi(\tau) = \langle\langle [X(t+\tau) + Y(t+\tau)]\,[X(t) + Y(t)]\rangle\rangle \implies$$
$$R_\Phi(\tau) = R_Y(\tau) + R_Z(\tau) + 2\,\langle X\rangle\,\langle Y\rangle$$

From the relation between the frequency spectrum and the autocorrelation (Eq. (2.21)), results

$$\Omega_\Phi(\omega) = \Omega_X(\omega) + \Omega_Y(\omega) + 4\pi\,\langle X\rangle\,\langle Y\rangle\,\delta(\omega) \tag{2.22}$$

where I have used a identity for the Dirac distribution,

$$\int_{-\infty}^{\infty} e^{-i\omega\tau} d\tau = 2\pi\delta(\omega)$$

2.6 Correlated Gaussian noise

In this section, we introduce three of the four noise backgrounds that will be used in the dissertation to analyze the structure of plasma density fluctuations: colored Gaussian noise, Ornstein-Uhlenbeck noise, and uncorrelated Gaussian noise (which is a special case of the previous ones). The fourth background is the pulse background, but its description will be postponed to Chapter 7.

2.6.1 Colored noise

A colored noise (CN) can be defined as a time signal $\Phi_\mathcal{N}(t)$ with a power spectrum that follows the relation

$$\Omega_\mathcal{N} \propto f^{-\eta}$$

at least for high frequencies below the Nyquist frequency, $f < 1/(2\Delta t)$ [87,88]. We call $\eta \in \mathbb{R}$ the noise exponent. Usually $\eta \in [-1, 2]$. Gaussian colored noise is the one that follows the

relation $\Omega_{\mathcal{N}} \propto f^{-\eta}$ and is also normally distributed. The color of the Gaussian noise does not change its PDF,

$$P_{\mathcal{N}}(\Phi_{\mathcal{N}}) = \frac{1}{\sqrt{2\pi\sigma_{\mathcal{N}}^2}} \exp\left[-\frac{(\Phi_{\mathcal{N}} - \langle\Phi_{\mathcal{N}}\rangle)^2}{\sigma_{\mathcal{N}}^2}\right]$$

The most used Gaussian colored noise is the white noise (WN), which has exponent $\eta = 0$. This means that the WN signal is uncorrelated and has a constant power spectrum. However, experimental time signals are often correlated, in which case colored noises may be a good option. All noises with $\eta \neq 0$ have a non-vanishing correlation between their points. Ref. 48 presents one example of CN applied in the context of plasma density fluctuations.

In this dissertation the MATLAB® dsp.ColoredNoise function will be employed to generate colored noise. In the region where the relation $\Omega_{\mathcal{N}} \propto f^{-\eta}$ is obeyed, the MATLAB CN follows

$$\Omega_{\mathcal{N}}(\omega) \approx \sigma_{\mathcal{N}}^2 \Delta t \frac{C(\eta)}{\omega^{\eta}} \tag{2.23}$$

where $\sigma_{\mathcal{N}}$ is the noise standard deviation, Δt is the interval between points and $\omega = 2\pi f$ is the angular velocity. We found an "empirical" expression for C,

$$C(\eta) = 10^{5.73\eta^{0.82}} \tag{2.24}$$

as shown in the fit of Fig. 2.3(a). Fig. 2.3(b) exhibits an example of the power spectrum for pink noise ($\eta = 1$), along with the fit of the Eq. (2.24).

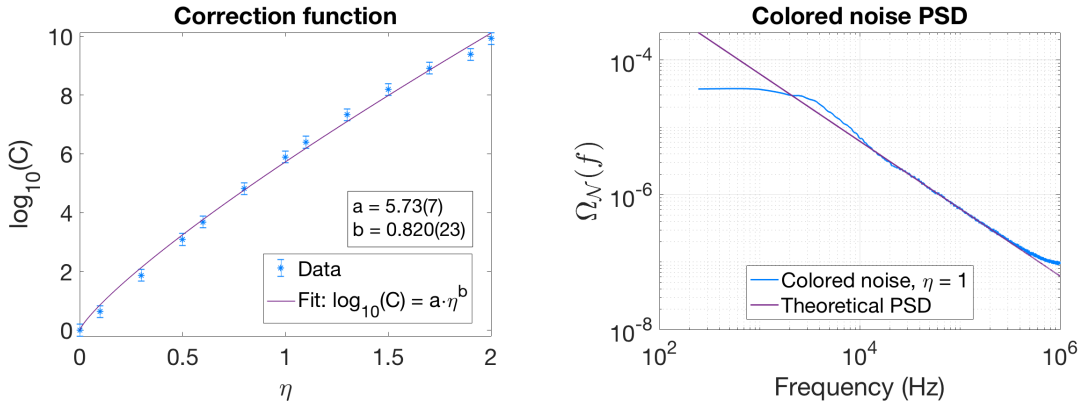


Figure 2.3: **(a)** Correction function $C(\eta)$ for the straight part of the colored noise PSD: $\Omega_{\mathcal{N}}(\omega) = \sigma_{\mathcal{N}}^2 \Delta t C(\eta)/\omega^{\eta}$. **(b)** Example of fit for $\eta = 1$.

2.6.2 Ornstein-Uhlenbeck noise

Another very useful correlated Gaussian noise is the Ornstein-Uhlenbeck (OU) process [89–91]. Following the notation of Ref. 89, it can be defined by

$$\Phi_{\mathcal{N}}(t+dt) = \Phi_{\mathcal{N}}(t) - \frac{1}{\tau_{\mathcal{N}}}\Phi_{\mathcal{N}}(t)dt + c^{1/2}g(t)(dt)^{1/2} \tag{2.25}$$

$\tau_{\mathcal{N}}$ is called the relaxation or correlation time [89,90]. $\Phi_{\mathcal{N}}$ here is the OU process (or noise), t is a continuous-time variable, c is the diffusion coefficient and $g(t)$ is an uncorrelated Gaussian distributed variable with vanishing mean and unity standard deviation. dt is understood as an infinitesimal quantity, that is, $dt \in [0, \delta]$, where δ is a real positive number and $\delta \to 0$.

The OU noise is a continuous Markov process, meaning that the current state of the system at $t + dt$ only depends on the previous state at t, for $dt \to 0$. The initial motivation for the OU process was to describe the velocity of a Brownian particle [89,91]. Later, it has been applied to model some types of colored noise [89] and different phenomena, including plasma physics [54].

For initial time $t_0 = 0$, the Ornstein-Uhlenbeck noise (OUN) has mean and variance

$$\langle \Phi_{\mathcal{N}}(t) \rangle = \Phi_0 e^{-t/\tau_{\mathcal{N}}} \tag{2.26}$$

$$\sigma^2_{\mathcal{N}}(t) = \frac{1}{2}c\tau_{\mathcal{N}}\left(1 - e^{-2t/\tau_{\mathcal{N}}}\right) \tag{2.27}$$

Thus the OU process is "asymptotically stationary", in the sense that its mean and variance asymptotically tend to be constants for $t/\tau_{\mathcal{N}} \to \infty$. For this dissertation, the transient term $e^{-2t/\tau_{\mathcal{N}}}$ can be neglected, as very rapidly in the modeled signals $t/\tau_{\mathcal{N}} \gg 1$.[8] Neglecting the transient regime, it is possible to write the diffusion coefficient in terms of the signal variance and correlation time,

$$c = 2\frac{\sigma^2_{\mathcal{N}}}{\tau_{\mathcal{N}}}$$

A non-vanishing average can also be added to the signal, transforming $\Phi_{\mathcal{N}}(t)$ into $\Phi_{\mathcal{N}}(t) + \langle \Phi_{\mathcal{N}} \rangle$. Thus, neglecting the transient regime, the OU process can be described in terms of $\langle \Phi_{\mathcal{N}} \rangle$, $\sigma_{\mathcal{N}}$ and $\tau_{\mathcal{N}}$.

The correlation time is easily understood in terms of the auto-correlation (Section 2.5 and Eq. (A17) of Ref. 54) of the OU process,

$$R_{\mathcal{N}}(\tau) = \langle \Phi_{\mathcal{N}}(t+\tau)\Phi_{\mathcal{N}}(t) \rangle = \sigma^2_{\mathcal{N}} \exp\left(-\frac{|\tau|}{\tau_{\mathcal{N}}}\right) + \langle \Phi_{\mathcal{N}} \rangle^2 \tag{2.28}$$

Thus the correlation time $\tau_{\mathcal{N}}$ is the time where the autocorrelation decays in $1/e$. The power spectrum of the OUN, on the other hand, is (Eq. (A14) of Ref. 54)[9]

[8] In the results of this dissertation, the maximum $\tau_{\mathcal{N}}$ will be about 10 μs. For $t > 5\tau_{\mathcal{N}}$, the transient effect $e^{-t/\tau_{\mathcal{N}}}$ is already smaller than 1%. Thus, after no more than 50 µs the signal can be considered stationary. On the other hand, the total analysis time considered for the signals in this dissertation will be $T \approx 40$ ms. Hence the noise non-stationary effects would be noticeable at most only at 50 µs/40 ms = 0.125% of the analysis.

[9] Such functions of the form $y(x) \propto 1/(1 + cx^2)$ are called Lorentzian functions [54,80].

$$\Omega_{\mathcal{N}}(\omega) = \frac{2\tau_{\mathcal{N}}\sigma_{\mathcal{N}}^2}{1+\tau_{\mathcal{N}}^2\omega^2} + 2\pi\delta(\omega)\left\langle\Phi_{\mathcal{N}}\right\rangle^2 \tag{2.29}$$

where $\omega = 2\pi f$ is the angular velocity and $\delta(\omega)$ is the Dirac distribution.

The Ornstein-Uhlenbeck process can be easily simulated using Gillespie's algorithm, which gives an exact updating formula (Eq. (1.10) of Ref. 89):

$$\Phi_{\mathcal{N}}(t+\Delta t) = \Phi_{\mathcal{N}}(t)e^{-\Delta t/\tau_{\mathcal{N}}} + \left[\frac{c\tau_{\mathcal{N}}}{2}\left(1-e^{-2\Delta t/\tau_{\mathcal{N}}}\right)\right]^{1/2} g(t) \tag{2.30}$$

The non-vanishing mean $\langle\Phi_{\mathcal{N}}\rangle$ can be added to $\Phi_{\mathcal{N}}(t)$ after the implementation of the above equation.

2.6.3 Comparing CN and OUN

Some comparisons between colored noise (CN, Subsection 2.6.1) and the Ornstein-Uhlenbeck noise (OUN) can be made. Both are correlated Gaussian noises—meaning that each point $\Phi_{\mathcal{N}}(t_i)$ is normally distributed, but can be correlated with another $\Phi_{\mathcal{N}}(t_j)$. Both CN and OUN can be expressed in terms of three parameters: the mean $\langle\Phi_{\mathcal{N}}\rangle$, standard deviation $\sigma_{\mathcal{N}}$ and a correlation parameter. In the case of colored noise, the correlation parameter is the noise exponent η, from $\Omega \propto 1/f^{\eta}$. For the OU process, it is $\tau_{\mathcal{N}}$. The correlation of both processes increases with their correlation parameter. For $\eta = 0$ and $\tau_{\mathcal{N}} \to 0$, the signals tend to the uncorrelated Gaussian noise (white noise).[10] When $\tau_{\mathcal{N}} \to \infty$, the OUN tends to brown noise ($\eta = 2$), as seen in Eq. (2.29).

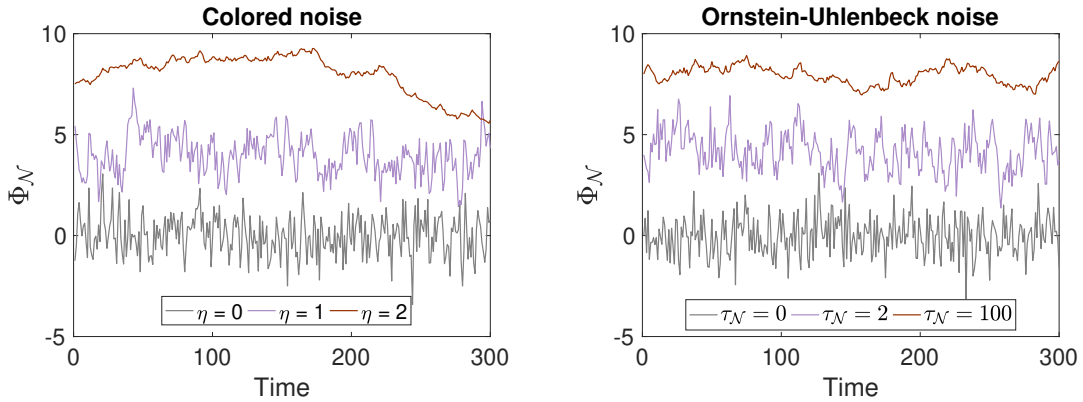


Figure 2.4: **(a)** Examples of colored noise with different exponent η ($\Omega_{\mathcal{N}} \propto 1/f^{\eta}$). $\eta = 0$ corresponds to the so-called white noise; $\eta = 1$, to pink noise; $\eta = 2$, to brown noise. **(b)** Examples of OU noise with different time correlations $\tau_{\mathcal{N}}$.

[10] In the OUN case, the condition $\tau_{\mathcal{N}} \to 0$ has to be accompanied with $\sigma_{\mathcal{N}}^2 = c\tau_{\mathcal{N}}/2 = cte$.

Fig. 2.4(a) shows three different characteristic signals of CN. All cases have unity standard deviation. Each has a different average value so that the curves can be distinguished in the same plot. For $\eta = 0$, white noise is achieved. For $\eta = 1$, results in the so-called pink noise. $\eta = 2$ corresponds to brown noise, in reference to the Brownian motion [87]. Fig. 2.4(b) depicts three OU processes with similar structures from the CN signals of Fig. 2.4(a). $\tau_{\mathcal{N}}$ and t have the same unit. For $\tau_{\mathcal{N}} \to 0$, white noise is recovered. For $\tau_{\mathcal{N}} \approx 2$, the signal resembles pink noise. For $\tau_{\mathcal{N}} \gg 1$ (e.g., $\tau_{\mathcal{N}} = 100$), the process resembles brown noise ($\Omega \propto 1/\omega^2$), since for $\omega \neq 0$ the power spectrum of the OU process is given by $\Omega_{\mathcal{N}}(\omega) = 2\tau_{\mathcal{N}}/(1 + \tau_{\mathcal{N}}^2\omega^2)$ (Eq. (2.29)) and thus, for $\tau_{\mathcal{N}}\omega \gg 1$, then $\Omega_{\mathcal{N}}(\omega) \approx 1/(\tau_{\mathcal{N}}\omega^2)$.

In resume, the colored and Ornstein-Uhlenbeck noises are very similar. However, the latter has some advantages. The PSD of CN is distorted as $\omega \to 0$. Additionally, there is no exact expression for the power spectrum of colored noise (Eq. (2.23)). The PSD of the OU process, however, is well defined (Eq. (2.29)). Furthermore, the Ornstein-Uhlenbeck process is easily computed with Gillespie's algorithm (Eq. (2.30)), while colored noise implementation is not unique and is not straight-forward [87, 88].

In this dissertation, we will compare the results from white noise, colored noise and the Ornstein-Uhlenbeck process to describe the background of plasma density fluctuations. The three have been used in the literature of magnetized plasmas [34, 48, 54]. However, until now no comparison between them was employed with experimental data.

2.7 Complexity-Entropy diagram

2.7.1 Bandt-Pompe probability

The Bandt-Pompe (BP) probability [92] is used to create the complexity-entropy diagram [93], which is very useful to distinguish noise from chaos and to characterize a signal structure. In this section will be explained how to obtain the BP probability from a signal $\mathbf{\Phi}$ (for example $\mathbf{\Phi} = I_{sat}$, Chapter 3). The notations and explanations here are mainly based on Refs. 94 and 95.

Let $\mathbf{\Phi}$ be a time series with N elements,[11]

$$\mathbf{\Phi} = (\Phi_1, \Phi_2, ..., \Phi_N)$$

An embedding dimension d is chosen to create overlapped d-tuples of $\mathbf{\Phi}$. There are $D = N - d + 1$ d-tuples in total, and they are elements of the set

$$\mathbf{Y} = \{(\Phi_j, ..., \Phi_{j+d-1}) \mid j = 1, ..., D\}$$

For example, if $N = 6$ and $d = 3$,

$$\mathbf{\Phi} = \{\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6\},$$

[11] In this section 2.7, bold letters will be reserved to arrays and sets.

$$D = 6 - 3 + 1 = 4$$

and the 3-tuples are

$$\mathbf{y}^{(1)} = (\Phi_1, \Phi_2, \Phi_3),\ \mathbf{y}^{(2)} = (\Phi_2, \Phi_3, \Phi_4)$$
$$\mathbf{y}^{(3)} = (\Phi_3, \Phi_4, \Phi_5),\ \mathbf{y}^{(4)} = (\Phi_4, \Phi_5, \Phi_6)$$

The Bandt-Pompe probability is obtained with permutations of the d-tuples. Let $\mathbf{y} = (y_1, ..., y_d)$ be a d-tuple of $\mathbf{Y}$. Then exists a permutation $\pi = \pi_1 ... \pi_d$, where $\pi_j \in \mathbb{N}_d = \{1, 2, ..., d\}$, such that $\tilde{y} = (y_{\pi_1}, y_{\pi_2}, ..., y_{\pi_d})$ is in descending order. This π is called the ordinal state of y. For example, if $d = 3$ and

$$\mathbf{\Phi} = \{9.1, 0.3, 4.5, 6.7, 3.2, 3.8\}$$

then the 3-tuples of $\mathbf{\Phi}$ are

$$\mathbf{y}^{(1)} = (9.1, 0.3, 4.5),\ \mathbf{y}^{(2)} = (0.3, 4.5, 6.7) \tag{2.31}$$
$$\mathbf{y}^{(3)} = (4.5, 6.7, 3.2),\ \mathbf{y}^{(4)} = (6.7, 3.2, 3.8) \tag{2.32}$$

The ordinal state of $\mathbf{y}^{(1)}$ is $\pi^{(1)} = 132$, because the corresponding permutation of $\mathbf{y}^{(1)}$ is in descending order,

$$\tilde{\mathbf{y}}^{(1)} = (y_1^{(1)}, y_3^{(1)}, y_2^{(1)}) = (9.1, 4.5, 0.3)$$

For $\mathbf{y}^{(2)}$, $\mathbf{y}^{(3)}$ and $\mathbf{y}^{(4)}$, the ordinal states are $\pi^{(2)} = 321$, $\pi^{(3)} = 213$ and $\pi^{(4)} = 132$.

The Bandt-Pompe probability can now be defined. Let N_π be the number of times the d-tuples $\mathbf{y} \in \mathbf{Y}$ have ordinal state π. As stated before, the total number of d-tuples is $D = N - d + 1$. Then the BP-probability p_π is defined as

$$p_\pi = \frac{N_\pi}{D} \tag{2.33}$$

p_π is the probability that a randomly chosen d-tuple $\mathbf{y}$ have the ordinal state π. In the example under appreciation, $D = 4$ and the possible ordinal states are $K = 3! = 6$,

$$\pi^{(a)} = 123,\ \pi^{(b)} = 132,\ \pi^{(c)} = 312$$
$$\pi^{(d)} = 321,\ \pi^{(e)} = 231,\ \pi^{(f)} = 213$$

The Bandt-Pompe probability distribution is defined as the set of all p_π,

$$\mathbf{P} = \{p_\pi : \pi \in \mathbf{\Pi}\} \tag{2.34}$$

where $\mathbf{\Pi}$ is the set of all possible $K = d!$ permutations. In the present case,

$$\mathbf{P} = \{p_{123}, p_{132}, p_{312}, p_{321}, p_{231}, p_{213}\}$$

For the signal $\mathbf{\Phi} = \{9.1, 0.3, 4.5, 6.7, 3.2, 3.8\}$, there are four 3-tuples (Eq. (2.31) and (2.32)) and they respectively have the ordinal states $\pi^{(1)} = 132$, $\pi^{(2)} = 321$, $\pi^{(3)} = 213$ and $\pi^{(4)} = 132$. Thus, in this case, the BP probability distribution is

$$\mathbf{P} = \{0,\ 1/2,\ 0,\ 1/4,\ 0,\ 1/4\}$$

In this thesis the time series of interest will not have 6 points, but rather $N_{pt} = 8 \times 10^4$.

2.7.2 Complexity and entropy

With the Bandt-Pompe probability distribution (Eq. (2.34)), one can proceed to construct the Complexity-Entropy (CH) diagram [93]. The CH diagram is a two-dimensional plane, having a complexity measure C_{JS} on the y-axis and an normalized Shannon entropy H_S on the x-axis. Both C_{JS} and H_S are functionals of the BP distribution $\mathbf{P}$.

The Shannon entropy is defined as

$$S(\mathbf{P}) = -\sum_{j=1}^{K} p_{\pi^{(j)}} \log_2(p_{\pi^{(j)}})$$

where $K = d!$ is the number of permutations of the d-tuple and the probability $p_{\pi^{(i)}}$ is given by Eq. (2.33). In the CH diagram, the entropy is normalized by the maximum entropy,

$$S_{max} = S(\mathbf{P}_e)$$

where $\mathbf{P}_e$ is a Bandt-Pompe equilibrium distribution, with equally probable states $p_\pi = 1/K$, for all possible permutations π [96].[12] Thus

$$S(\mathbf{P}_e) = -\sum_{j=1}^{K} \frac{1}{K} \log_2\left(\frac{1}{K}\right) = \log_2(K)$$

Consequently, the normalized Shannon entropy, used in the CH diagram, is

$$H_S(\mathbf{P}) = \frac{S(\mathbf{P})}{S(\mathbf{P}_e)} = -\frac{1}{\log_2(K)} \sum_{j=1}^{K} p_{\pi^{(j)}} \log_2\left(p_{\pi^{(j)}}\right)$$

The possible values of H_S are between 0 and 1. $H_S = 0$ implies that $p_{\pi^{(j)}} = 1$ for one j, meaning that the system has only one state and is deterministic. $H_S = 1$ implies that $p_{\pi^{(j)}} = 1/K$ for all j, yielding the maximum entropy.

In its turn, the Jensen-Shannon complexity is defined by

$$C_{JS}(\mathbf{P}) = Q_{JS}(\mathbf{P}) H_S(\mathbf{P})$$

where the disequilibrium $Q_{JS}(\mathbf{P})$ is

$$Q_{JS}(\mathbf{P}) = Q_0 \cdot D_{JS}(\mathbf{P})$$

and D_{JS} is the divergence

$$D_{JS}(\mathbf{P}) = S\left(\frac{\mathbf{P} + \mathbf{P}_e}{2}\right) - \left[\frac{S(\mathbf{P}) + S(\mathbf{P}_e)}{2}\right]$$

[12] For example, a signal $\mathbf{\Phi}$ that is uniformly distributed or a Gaussian white noise will have a BP distribution $\mathbf{P}_e$, because in these two cases all possible ordinal states are equally probable. In this thesis, the chosen embedding dimension is $d = 6$, yielding $K = d! = 720$ possible ordinal states and $S(\mathbf{P}_e) \simeq 9.5$.

Finally, the normalization constant on the $Q_{JS}(\mathbf{P})$ equation is[13]

$$Q_0 = \frac{K+1}{K}\log_2(K+1) - 2\log_2(2K) + \log_2(K)$$

Writing in terms of $S(\mathbf{P})$ and constants, the normalized Shannon entropy and the Jensen-Shannon complexity are given by

$$H_S(\mathbf{P}) = \frac{S(\mathbf{P})}{S(\mathbf{P}_e)}$$

$$C_{JS}(\mathbf{P}) = Q_0 \left\{ S\left(\frac{\mathbf{P}+\mathbf{P}_e}{2}\right) - \left[\frac{S(\mathbf{P})+S(\mathbf{P}_e)}{2}\right] \right\} \frac{S(\mathbf{P})}{S(\mathbf{P}_e)}$$

This Shannon entropy quantifies the number of ordering states a system has. On the other hand, the Jensen-Shannon complexity measures how much a signal deviates from maximum entropy. In other words, C_{JS} quantifies the preferred ordering states of a system [93].

2.7.3 The CH diagram

With the information provided, it is now possible to present a CH diagram. For a given Shannon entropy H_S, the complexity C_{JS} can not be arbitrary. Rather, C_{JS} is constrained between a minimum C_{min} and maximum C_{max} [96]. This is illustrated in the red and blue curves of Fig. 2.5, which depicts a complexity-entropy diagram with embedding dimension $d_{CH} = 6$.[14] (Following Ref. 95, $d_{CH} = 6$ will be used in all CH diagrams presented in this thesis.)

In the complexity-entropy diagram, between C_{min} and C_{max} is located the fractional Brownian motion curve (fBm), seen in green in Fig. 2.5 for different Hurst exponents H ranging from 0 to 1 [93]. Following Ref. 97, the fBm represents a family of processes $B_H(t)$ with null mean and autocorrelation

$$\langle B_H(t_1)B_H(t_2)\rangle = \frac{1}{2}\left(t_1^{2H} + t_2^{2H} - |t_1 - t_2|^{2H}\right)$$

where $H \in [0,1]$ is the Hurst exponent of the process. The usual interpretation [93] states that signals situated above the fBm curve in the CH diagram have a predominantly chaotic behavior, whereas signals located on the fBm curve or below it are mainly stochastic.

A Gaussian white noise signal has maximum entropy $H_S = 1$ and minimum complexity, $C_{JS} = 0$. Other types of colored noise and Ornstein-Uhlenbeck noise are depicted in Fig. 2.5. It is seen that, as the correlation parameter (η or $\tau_{\mathcal{N}}$) increases, the corresponding CH points climb the fBm curve, decreasing the data entropy and enhancing the complexity. Nevertheless, pure OU noises can not go further than the position with fBm Hurst exponent $H = 0.5$.

[13] When $K \gg 1$, $Q_0 \simeq -2$. Example: for $K = d! = 720$ ordinal states, $Q_0 \simeq -1.98$.

[14] From now on, instead of d, the symbol d_{CH} will be used to denote the immersion dimension of the CH diagram, to avoid confusion with the embedding dimension from the recurrence plots and the recurrence quantification analysis (Section 2.8).

In summary, the CH diagram quantifies the ordering of states of different signals. It can be used to indicate if a time series has strong deterministic and/or stochastic components. For these reasons, the CH diagram will be employed in Chapters 6 and 8 to compare different categories of background signals which can model density fluctuations in the TCABR tokamak. In special, we aim to verify whether the stochastic pulse train model (to be introduced in Chapters 4 and 7) has compatible results with experiments in the CH plane. This question was proposed in Ref. 80.

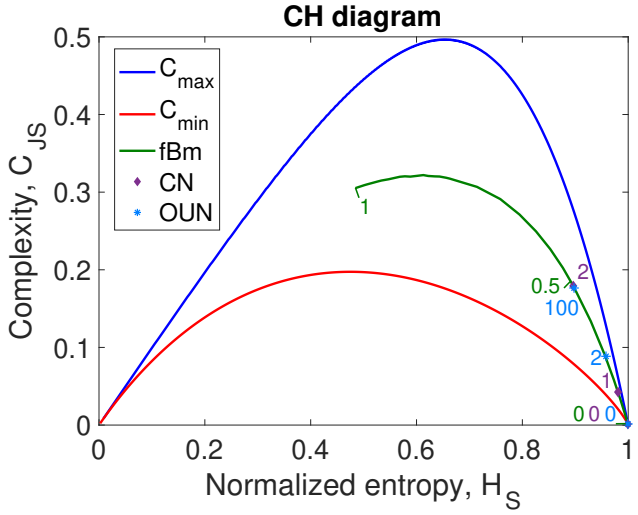


Figure 2.5: Example of a Complexity-Entropy diagram, showing the curves of minimum and maximum complexity, C_{min} and C_{max}. The green curve corresponds to fractional Brownian motion and the points with Hurst exponents $H \in \{0, 0.5, 1\}$ have been highlighted in green. Also shown are CH points of the signals of Fig. 2.4: in purple, are depicted CH points of colored noise signals with exponents $\eta \in \{0, 1, 2\}$ (where $\Omega \propto 1/f^{\eta}$); in light blue, are exhibited CH points of Ornstein-Uhlenbeck processes with correlation time $\tau_{\mathcal{N}}/\Delta t \in \{0, 2, 100\}$. The points with $H = \eta = \tau_{\mathcal{N}}/\Delta t = 0$ coincide, as they correspond to white noise. Moreover, the points with $H = 0.5$, $\eta = 2$ and $\tau_{\mathcal{N}}/\Delta t = 100$ almost coincide.

2.8 Recurrence plots and RQA determinism

2.8.1 Recurrence plots

Recurrence plots were proposed in Ref. 98 with the purpose of "measuring the time constancy of dynamical systems". Since then, the theory of recurrence has evolved, finding applications in various fields, such as plasma physics [99–103] and biophysics [104]. This section follows Ref. 101, in which the concepts of recurrence plots (RPs) and recurrence quantification analysis (RQA) are didactically elucidated.

Suppose an experiment measures a one-dimensional signal $\Phi_j = \Phi(t_j)$, where the j-th instant is given by $t_j = j\Delta t$, with $\Delta t > 0$, $j \in [1, 2, ..., N_{pt}]$ and $N_{pt} \in \mathbb{N}$. The signal $\Phi(t)$ could be for example the ion saturation current measured in a tokamak (Chapter 3).

An embedding space can be generated using vectors of the kind

$$\mathbf{x}_j = (\Phi_j, \Phi_{j+\tau}, \Phi_{j+2\tau}, ..., \Phi_{j+(d-1)\tau})$$

where d is the embedding dimension and τ is the delay. Usually, τ is set as the value at which the autocorrelation function of $\tilde{\Phi} = (\Phi - \langle\Phi\rangle)/\sigma_\Phi$ reaches 0 for the first time, $R_{\tilde{\Phi}}(\tau) = 0$ [99, 102, 105]. For practical purposes, it is enough to set τ such that $R_{\tilde{\Phi}}(\tau) = 10\%$. The embedding dimension d, on the other hand, in general does not have such a straightforward method to be evaluated. Some details about this will be given in Subsection 2.8.2.

The embedding vectors $\mathbf{x}_j$ can be understood as elements of a d-dimensional reconstructed phase space. Embedding theorems guarantee that this phase space $\mathbf{X} = \{\mathbf{x}_j | j = 1, 2, ..., N_{pt}\}$ conserves dynamical proprieties of the original signal $\Phi(t)$, as long as d is large enough. Nonetheless, even if d is small, important information about the system can be recovered.

The time evolution of the vectors $\mathbf{x}_j$ can be represented by a trajectory in the phase space with monotonically increasing j: $\mathbf{x}_j \to \mathbf{x}_{j+1} \to \mathbf{x}_{j+2} \to ...\mathbf{x}_{N_{pt}-(d-1)\tau}$. In this framework, two states $\mathbf{x}_i$ and $\mathbf{x}_j$ are said to be recurrent if the trajectory has approximately the same d-dimensional state for both of them, "$\mathbf{x}_i \approx \mathbf{x}_j$". Formally, the recurrence is defined as

$$R_{ij} = \begin{cases} 1, & \|\mathbf{x}_i - \mathbf{x}_j\| \leq \epsilon \\ 0, & \|\mathbf{x}_i - \mathbf{x}_j\| > \epsilon \end{cases}$$

where ϵ is a threshold and $\|...\|$ stands for some norm (usually the Euclidian norm).

The $N_{pt} \times N_{pt}$ matrix with elements R_{ij} is known as the recurrence matrix (RM). A recurrence plot (RP) is an illustration of this matrix, drawing dots for $R_{ij} = 1$ and letting the space empty for $R_{ij} = 0$. By construction, the RM is symmetric ($R_{ij} = R_{ji}$) and the main-diagonal elements are one, $R_{ii} = 1$, meaning that a point is always recurrent to itself. In fact, diagonal structures are important features of recurrence plots, as they are often related to deterministic characteristics of the original signal.

Another important quantity is the recurrence rate, which is the fraction of points in the RM that are recurrent with each other:

$$\mathrm{RR} = \frac{1}{N_{pt}(N_{pt}-1)} \sum_{i,j=1, i\neq j}^{N_{pt}} R_{ij}$$

N_{pt}^2 is the total of points of the recurrence matrix. In RR, the main diagonal isn't accounted for, since $R_{ii} = 1$ by definition. The higher the threshold ϵ is, the more points are recurrent with others. The threshold is in general chosen as a small fraction of the signal variance [99, 101] or set so that the recurrence rate is fixed at some small value, usually ranging from $\mathrm{RR} = 1\%$ to $\mathrm{RR} = 10\%$ [101, 102].

With recurrences, it is possible to analyze the dynamical behavior of systems even when their governing equations are unknown. This applies for example to turbulent systems, such as the ones found in the boundary of tokamaks. Besides, magnetized plasmas are associated both with stochastic and chaotic behaviors [50, 80] and can even alternate between regimes of bursts and Gaussian noise [49]. At the same time, the plasma signal can have periodic and stochastic components [48]. The recurrence theory can analyze all these types of signals.

As examples, Fig. 2.6 depicts recurrence plots of synthetic generated series of white noise (WN, $\eta = 0$ in Subsection 2.6.1), brown noise (BN, $\eta = 2$), and a periodic motion (PM),

$$\Phi_i = \sin(2\pi \cdot 0.02i) + \sin\left(2\pi \cdot \frac{0.1}{3} i\right) \tag{2.35}$$

so that $t = i\Delta t$, with $i = 1, 2, ..., 1000$ and $\Delta t = 1$. The noises had vanishing mean and unit standard deviation. The recurrence rate was fixated at RR $= 10\%$, the embedding dimension was chosen as $d = 4$ and the delay τ was fitted by the autocorrelation of the signals, $R_{\tilde{\Phi}}(\tau) = 10\%$. The values obtained were $\tau_{WN} = 1$, $\tau_{BN} = 75$ and $\tau_{PM} = 18$.

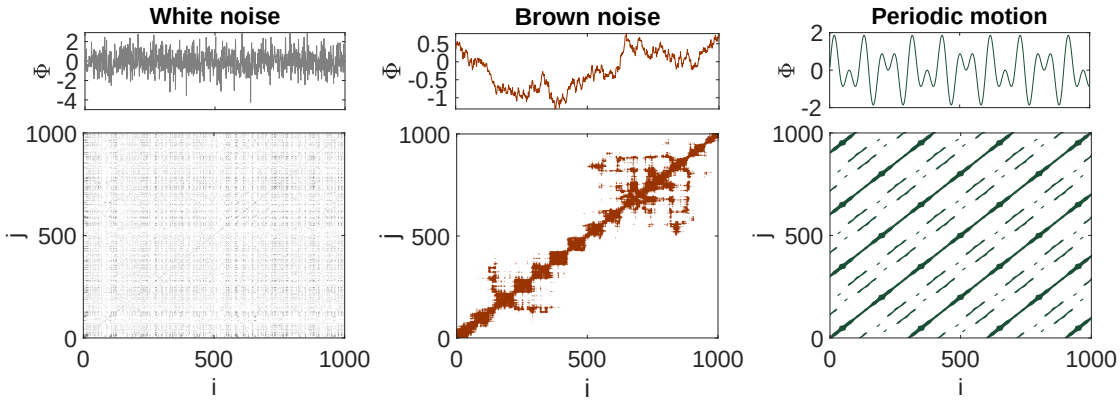


Figure 2.6: Examples of recurrence plots, with the original signals in the upper panels, and the RP in the lower ones. Time index used: $t = i\Delta t$, with $\Delta t = 1$. **(a)** RP of a white noise with zero mean and unity variance. **(b)** RP of a brown noise with vanishing mean and asymptotically unitary variance. **(c)** RP of a sum of two sines (Eq. (2.35)).

For the white noise (Fig. 2.6(a)), the recurrence plot is filled with isolated dots. This absence of structures and diagonal lines indicates that recurrences occur by chance, as expected for an uncorrelated process.

For the brown noise (Fig. 2.6(b)), the recurrences are concentrated near the main diagonal. This means that one state of the system is mainly recurrent with its neighbors, which is expected for highly correlated noise. At times, the recurrences points form squares of length $L \approx \tau_{BN}$. Regions around $i \in [100, 300]$ and $i \in [600, 800]$ also were shown to be recurrent because for these segments the signal fluctuated less.

Lastly, Fig. 2.6(c) depicts the recurrence plot of the sum of sines (Eq. (2.35)). All recurrences are grouped into diagonals (smalls or large), reflecting that the signal is indeed periodic, since a state is revisited by the trajectory in a predictable and periodic manner. Small circles are seen near the extrema points, where the derivative $d\Phi/dt$ is close to zero.

We see that recurrence plots treat uncorrelated noise, correlated noise and periodic motion

very differently.[15] In the next section, it will be explained how to quantify these recurrences in one quantity called determinism.

2.8.2 RQA determinism and embedding dimension

The recurrence plot can be quantified by various functions of the field known as recurrence quantification analysis (RQA), which was introduced by Ref. 104 and is well explained for example in Ref. 99. The most important quantifier of the RP is the so-called determinism, defined as the fraction of recurrent points (*i.e.*, $R_{i,j} = 1$) which form a diagonal in the RP,

$$\mathrm{DET} = \frac{\sum_{\ell=\ell_{min}}^{\ell_{max}} \ell P(\ell)}{P(1) + \sum_{\ell=\ell_{min}}^{\ell_{max}} \ell P(\ell)} \tag{2.36}$$

The name comes from the fact that deterministic behavior is linked to the presence of diagonal lines in the recurrence plot [101].

In Eq. (2.36), $\ell_{max} \leq N - 1$ and $\ell_{min} \geq 2$ are respectively the lengths of the biggest and smallest diagonals found in the given RP (disregarding the identity line, which has $\ell = N$ by construction). ℓ is the length of an arbitrary diagonal. $P(\ell)$ is the frequency distribution of ℓ in the recurrence plot, so that $\ell P(\ell)$ counts how many recurrent points are contained in diagonals of length ℓ. $P(1)$ is the frequency of isolated points.[16]

The RQA determinism of the signals of Fig. 2.6 are respectively $\mathrm{DET}_{WN} = 0.760(1)$, $\mathrm{DET}_{BN} = 0.958(4)$ and $\mathrm{DET}_{PM} = 0.99945(2)$ (with uncertainties in parenthesis). As expected, the DET from the white noise series was the lowest. On the other hand, the periodic motion has determinism practically compatible with 1—which makes sense since the PM is deterministic. The brown noise case, being highly correlated, also has a high determinism.

Increasing the embedding dimension also increases the RQA determinism (e.g Fig. 5 of Ref. 99). Because of this, DET can be employed to estimate a practical embedding dimension for different systems. Fig. 2.7 depicts the profiles of DET $\times\, d$ for the three signals considered in this section (Fig. 2.6). For $d \geq 2$, the DET from the periodic motion saturates at DET = 1. Thus, $d = 2$ would be enough in this case. Deterministic signals such as this one have well-defined embedding dimensions (see for example Sections 3.2 and 3.3 of Ref. 105).

Purely stochastic signals, however, can be regarded as having infinite embedding dimensions (as stated in Section 6.5 of Ref. 105). This is illustrated by the white noise curve in

[15] Furthermore, Appendix B.5 shows RPs for the saturation current signals measured in TCABR. Nevertheless, to fully understand what are these signals and their simulations, Chapters 3 to 8 are needed.

[16] There are other ways to write the determinism, such as in Ref. 99,

$$\mathrm{DET} = \frac{\sum_{\ell=\ell_{min}}^{\ell_{max}} \ell P(\ell)}{\sum_{\ell=1}^{\ell_{max}} \ell P(\ell)}$$

The representation of Eq. (2.36) was chosen to highlight that $P(1)$ is the difference between the numerator and denominator.

Fig. 2.7—the RQA determinism keeps increasing even for $d = 20$. Thus, in practice what can be done is to choose an embedding dimension above which DET does not change much.

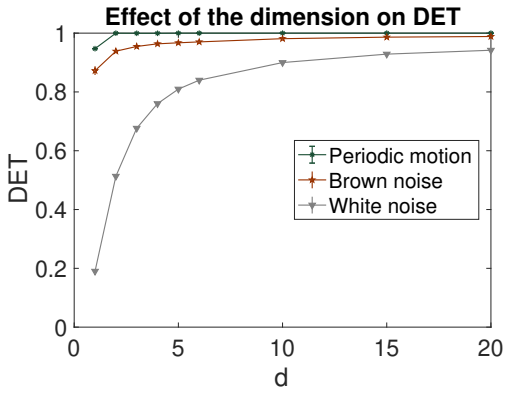


Figure 2.7: Profiles of the RQA determinism versus the embedding dimension, for the three synthetic signals considered in this section.

A similar problem occurs with turbulent data, like the ones measured in magnetized plasmas [99–102]. In practice, $d = 4$ was found to be an adequate dimension to analyze the signals, because for higher values the determinism saturated, as discussed in Refs. 99 and 101. Thus, following these results, we will use $d = 4$ in the present dissertation. Higher dimensions would also be adequate, but would demand more computational effort.

Table 2.1 stores the values that will be used in this work for the parameters of the recurrence quantification analysis. We highlight that moderate changes in the parameters wouldn't alter the results. From now on, a subindex RQA will be added to N, τ and d to avoid confusion with other quantities (such as the embedding dimension from the CH diagram, d_{CH}, Section 2.7).

Table 2.1: Parameters from the recurrence quantification analysis. Values chosen in analogy to other works about magnetized plasmas [99–102].

Parameter	Value	Name
RR	10%	Recurrence rate
N_{RQA}	10^3	Window length of the recurrence plot
τ_{RQA}	5 μs	Time delay
d_{RQA}	4	Embedding dimension

2.9 Chi-square

Let $\mathbf{y} = (y_1, y_2, ..., y_N)$ be a vector of N independent data points which obey a function $\mathbf{f}(\boldsymbol{\theta}_0) = (f_1, f_2, ..., f_N)$, where $\boldsymbol{\theta}_0 = (\theta_1, \theta_2, ..., \theta_M)$ is the vector of M unknown true parameters. The normalized residuals between y_i and f_i are (see for example Eq. (E23) of Ref.

106)

$$R_j(\boldsymbol{\theta}) = \frac{y_j - f_j(\boldsymbol{\theta})}{\sigma_j} \quad j \in \{1, 2, ..., N\}$$

where σ_i is the uncertainty associated with y_j. An important quantity is the sum of squared weighted residuals,

$$S(\boldsymbol{\theta}) = \sum_{j=1}^{N} \left(\frac{y_j - f_j(\boldsymbol{\theta})}{\sigma_j} \right)^2$$

When $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, S assumes its smallest value and is called chi-square,

$$\chi^2 = \sum_{j=1}^{N} \left(\frac{y_j - f_j(\boldsymbol{\theta}_0)}{\sigma_j} \right)^2 \tag{2.37}$$

In this thesis, χ^2 maps will be used to fit functions without an analytical form. As in the least-squares method, the goal of χ^2 maps is to estimate the vector $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ which minimizes S. In the context of χ^2 maps, usually χ^2 is also used to denote S (see for example Chapter 8 of Ref. 107 or pp. 268–278 of Ref. 108). This convention will be used in Chapter 5.

As stated in Section 1.3.6.6.6 of Ref. 109, when the residuals R_i are independent and normally distributed, the χ^2 PDF is a Gamma distribution (Eq. (2.7)) with scale parameter $A = 2$ and shape parameter $\gamma = \nu/2$, where ν is the number of freedom degrees. That is,

$$P_{\chi^2}(\chi^2; \nu) = \frac{1}{2\Gamma(\nu/2)} \left(\frac{\chi^2}{2} \right)^{\nu/2-1} \exp\left(-\frac{\chi^2}{2} \right)$$

The expected value of this Gamma distribution is the number of freedom degrees,

$$\langle \chi^2 \rangle = A\gamma = 2 \cdot \frac{\nu}{2} = \nu \tag{2.38}$$

while its standard deviation is

$$\sigma_{\chi^2} = A\sqrt{\gamma} = 2\sqrt{\frac{\nu}{2}} = \sqrt{2\nu} \tag{2.39}$$

Another important quantity is the reduced chi-square (e.g. p. 278 of Ref. 108),

$$\chi_r^2 = \frac{\chi^2}{\nu} \tag{2.40}$$

In the case of Gaussian data, from Eq. (2.38), (2.39) and (2.40) it is immediate that the reduced chi-square has an expected value and a standard deviation of

$$\langle \chi_r^2 \rangle = 1 \quad \text{and} \quad \sigma_{\chi_r^2} = \sqrt{\frac{2}{\nu}}$$

Even though experimental data is not always Gaussian, the above results are useful as approximate reference values.

Chapter 3

Measuring turbulent density fluctuations with the saturation current

This chapter briefly describes the theory of Langmuir probes, focusing on the measurement of saturation current (I_{sat}), related to the plasma density. After, some of the principal diagnostics decurrent of I_{sat} will be outlined, such as its distribution, power spectra and bursts analysis.

3.1 Langmuir probes and saturation current

Langmuir probes are electrodes widely used in both laboratories and space to measure the density and temperature of plasmas (see for ex. Section 11.5 of Ref. 11, Chapter 3 of Ref. 15 and Chapter 4 of Ref. 110). When immersed in plasmas, a sheath is formed around the probe, because of the plasma shielding effect discussed in Subsection 1.2.1. In other words, the plasma shields itself from the disturbing field caused by the electrode. The sheath extent is of the order of the Debye length.

Since electrons are much lighter than ions, the velocity of the first is higher in the plasma. Thus, the immersed electrode rapidly accumulates a negative charge. This potential then repels electrons and attracts ions, becoming less negative. At some point, the rate at which electrons and ions hit the probe surface will be equal, and the electrode achieves the so-called floating potential ϕ_w (Section 11.1 of Ref. 11). At ϕ_w, the probe is in equilibrium and the net current measured in it is 0.

An external potential ϕ can also be applied in the Langmuir probe. When ϕ is more negative than ϕ_w, an electrical field is created around the electrode, repelling electrons. Decreasing ϕ even more, eventually the contribution of the electrons to the probe current will be negligible and the total current density will approach a constant value J_i. This value is known as the ion saturation current density, since only ions are measured (Section 11.5 of Ref. 11).[1]

If, on the other way around, $\phi > \phi_w$, then the induced electric field will repel ions. For a large enough potential, the ion contribution to the electrode current density will be near

[1] There is an expansion effect of the sheath which slightly deforms the current-potential asymptote [45,111]. Still, for the proposes of this thesis, this correction can be neglected.

zero, letting the current density reach a constant value J_{e0}, only due to electron collision. J_{e0} is known as the electron saturation current density. The characteristic current-potential curve of a plane probe immersed in a plasma is disposed at Fig. 3.1.

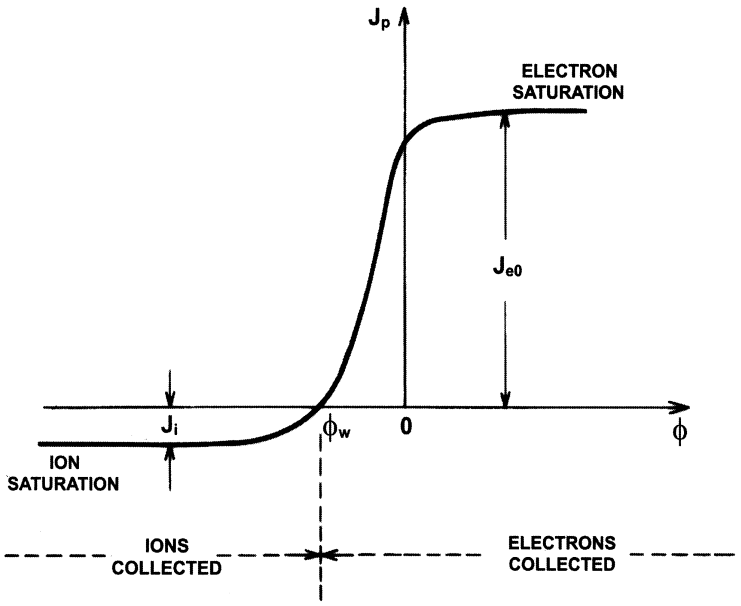


Figure 3.1: Current-potencial curve for a Langmuir probe immersed in a plasma. J_p is the plasma current density. Source: Fig. 11.4 of Ref. 11.

The current measured in the probe is equal to the current density times the area of the Langmuir probe pin, A_{LP}. The ion saturation current is positively defined,

$$I_{sat} \equiv -A_{LP} J_i$$

In this dissertation, measurements of I_{sat} performed in the TCABR tokamak will be analyzed.

The relation between I_{sat}, density and temperature at the sheath is given by (Eq. (2.81) of Ref. 112)

$$\boxed{I_{sat} \approx 0.5 A_{LP} e n_0 c_s}$$

where n_0 is the estimated plasma density that would exist without the probe perturbation. Due to quasi-neutrality, n_0 is approximately equal to ions and electrons. In its turn, c_s is the sound speed in the plasma (as used in Eq. (2.56) of Ref. 112),

$$c_s = \sqrt{k_B \frac{T_e + T_i}{m_i}}$$

k_B is the Boltzmann constant, T_e and T_i are the electron and ion temperatures and m_i is the mass of ions (protons in TCABR).

The relation

$$I_{sat} \propto n_0 \left(T_e + T_i\right)^{1/2}$$

shows that I_{sat} is a measure of density, for given temperatures. The electron temperature can be obtained with the characteristic curve ($I_{LP} \times V_{LP}$) of the Langmuir probe (as in Eq. 2.80 of Ref. 112, or Ref. 45). However, it is difficult to measure the ion temperature (Section 2.9 of Ref. 112) and it is even harder to simultaneously obtain I_{sat}, T_e and T_i to locally evaluate the density [38,45]. Nevertheless, for the analysis presented in this dissertation, the values of I_{sat} are sufficient to characterize turbulent fluctuations, similarly to what was done in previous works which also employed the saturation current [30–32,34,39,40,48,49].

Sometimes (as for example in Ref. 38) the ion temperature is neglected in the sound speed expression, $c_s = \sqrt{k_B T_e / m_i}$. However, in the SOL, T_i can be compatible to or higher than T_e (e.g. Fig. 2.21 of Ref. 112) and thus the assumption $T_e \gg T_i$ is not adequate for our analysis.

For this thesis, a rake Langmuir probe was used and detailed information about the instrument is given in Chapter 4 of Ref. 110. It has 18 pins divided into two rows, as seen in Fig. 3.2(a). The distance between the center of two consecutive pins in the same row is $\Delta r = 5$ mm (much greater than the Debye length, which is smaller than 0.06 mm in TCABR, as seen in Fig. 4.1 of 110). Each tip has a cylindric shape, with a diameter of $d = 0.8$ mm and height of $h = 3.0$ mm. Fig. 3.2(b) shows a schematic of the probe inserted in the plasma. Because of the disposition used in Fig. 3.2(b), the effective area A_{LP} that interacts with the plasma is about $A_{LP} \approx d \times h = 2.4$ mm^2.

(a) **(b)**

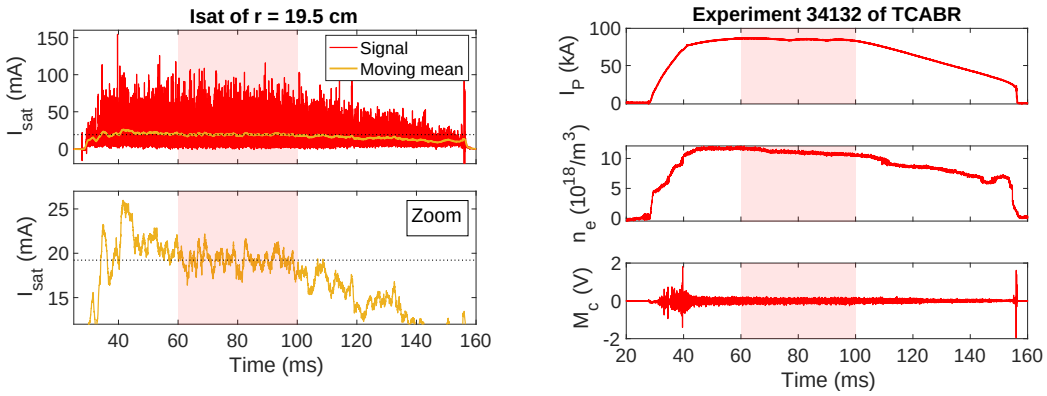
Figure 3.2: **(a)** Photo of a rake probe with 18 tips used in TCABR. Source: Fig. 4.4 of Ref. 110. **(b)** Scheme of the TCABR poloidal section, with the rake probe inserted. Also shown are the tokamak major and minor radii, $R_0 = 61.5$ cm and $a = 18.0$ cm, and the radial position r.

3.2 Steady state in TCABR

Although a discharge in TCABR lasts around 150 ms, only a part of it can be identified as a stationary state for the plasma. Usually, the steady regime duration is about $T = 30 \sim 60$ ms in TCABR, as described in Table 1.1. For this dissertation, experiment 34132 was

analyzed (Chapters 6 and 8), and the stationary regime was identified between $t_i = 60$ ms and $t_f = 100$ ms of the discharge, yielding $T = 40$ ms. To help recognize this state, a moving average of the saturation current was computed (yellow curve in Fig. 3.3(a)). For $t \in [0, 40]$ ms, I_{sat} increases on average. For $t \in [40, 60]$ ms, the moving mean decreases, and for $t \in [60, 100]$ ms it is approximately in steady state. For $t > 100$, the current decreases. Therefore, the interval $t \in [60, 100]$ ms was chosen for the analysis. In all further I_{sat} graphs of this work, only the stationary regime will be considered.

As seen in Fig. 3.2(b), in this experiment the plasma current and the line-averaged electron density were about $I_p \approx 85$ kA and $n_e \approx 1.1 \times 10^{19}$ m^{-3} in the steady regime. In the same figure, the Mirnov coil signal shows that this discharge had low magnetohydrodynamical activity.



(a) **(b)**

Figure 3.3: **(a)** In red, the saturation current for the position $r = 19.5$ cm of the TCABR experiment 34132. In yellow, the moving mean with windows of 1 ms. The highlighted area for $t \in [60, 100]$ ms corresponds to the interval detected as the steady regime. The dotted black line indicates the average of I_{sat} for this interval. **(b)** Diagnostics of experiment 34132: plasma current I_p, line-averaged electron density n_e, and Mirnov coil signal M_c.

3.3 The conditional average of bursts

To understand the idea underlying the stochastic pulse train model (Chapter 4), firstly it is useful to explain a conditional averaging technique [15, 25, 113] widely used in plasma physics to describe the average properties of extreme events [30–35, 40, 41]. Density signals of magnetically confined plasmas are generally turbulent. An example is given in Fig. 3.4(a), showing a saturation current measurement[2] at the scrape-off layer of TCABR.

[2]Other local measurements of density used in magnetically confined plasmas are gas puff imaging (GPI) [33, 35–37, 39] and Lithium-Beam Emission Spectroscopy (Li-BES) [41]. As the three aim to measure the same quantity (plasma density), they exhibit a similar turbulent behavior.

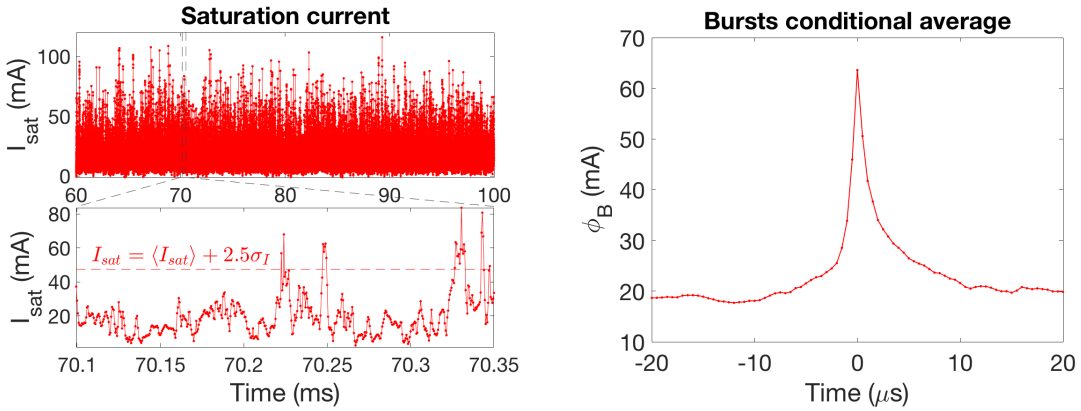


Figure 3.4: **(a)** An example of saturation current measurement made in TCABR (experiment 34132, position $r = 19.0$ cm). In the bottom panel, a zoom showing 4 bursts. **(b)** Conditionally averaged waveform made with the 493 bursts detected in the signal of (a).

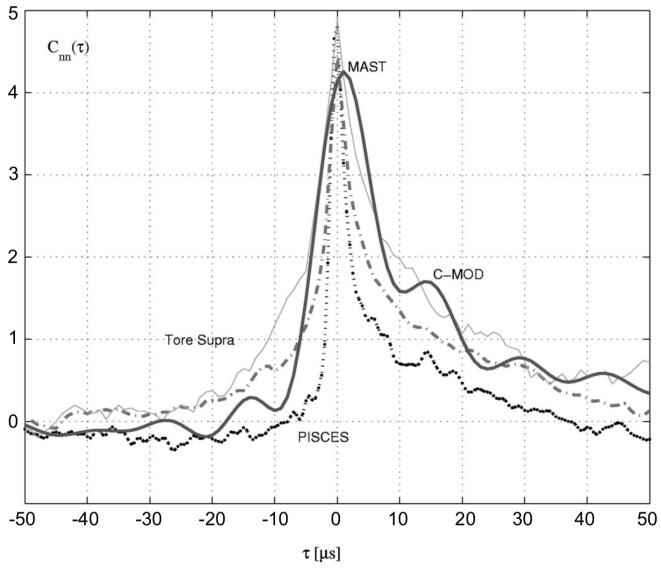


Figure 3.5: The conditional averaged burst for four different machines (Tore Supra, Alcator C-Mod, MAST and PISCES). The signal was normalized by $\tilde{I}_{sat} = \left(I_{sat} - \langle I_{sat} \rangle\right)/\sigma_I$, with threshold $\mathcal{T} = 3$ for bursts detection. Source: Ref. 25.

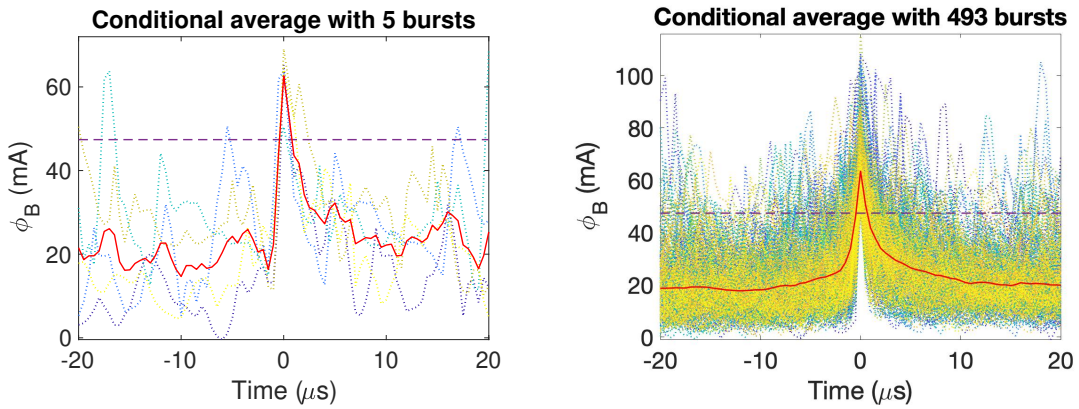


Figure 3.6: Conditional average of **(a)** 5 and **(b)** 493 bursts detected in the I_{sat} signal of Fig. 3.4(a). In red it is shown the conditional average, while the purple dashed line indicates the threshold for bursts detection, $I_{sat} > \langle I_{sat} \rangle + 2.5\sigma_I$.

Despite this noisy appearance of the density signals, the conditionally averaged burst shapes of several plasma devices have similar waveforms and are reproducible within each machine (given the same radial position and plasma conditions). Thus, this indicates a universal feature of large-amplitude fluctuations in plasmas. Fig. 3.5 shows the conditional average of bursts for four different devices, while Fig. 3.4(b) depicts an example for TCABR.

The procedure to construct the conditional average of bursts is the following. Bursts are detected as peaks with intensity higher than the average of the signal by $\mathcal{T}$ times the signal standard deviation. That is, a pulse is identified as a burst if its peak is bigger than $\langle I_{sat} \rangle + \mathcal{T}\sigma_I$. $\mathcal{T}$ is called the threshold and is usually defined as 2 [48], 2.5 [15, 30–35, 40, 41], or 3 [25]—so that bursts are indeed seen as extreme events with intensity much stronger than the plasma background fluctuations. The bottom panel in Fig. 3.4(a) exhibits a zoom with four detected bursts.[3] After detecting these extreme events, they are put together in a time axis with each peak at $t = 0$ µs, as in Fig. 3.6. Then the average of these pulses is made, and this is called a conditional average (or an auto-conditional average [25]). An example is given in Fig. 3.4(b). The average of all the 493 bursts results in Fig. 3.4(b).

Table 3.1: Fixed parameters used in experiments of I_{sat} analyzed in this dissertation.

Symbol	Meaning	Value(s)
t_i	Initial time for the I_{sat} analysis	60 ms
t_f	Final time for the I_{sat} analysis	100 ms
Δt	Time step of the I_{sat} measurements	0.5 µs
Δt_B	Minimum waiting time between bursts	15 µs
$\mathcal{T}$	Threshold for bursts detection	2.5

[3] To avoid detecting the same burst two times, it is determined a minimum waiting time between these extreme events so that only the most intense peak is considered. The value used here is $\Delta t_B = 15$ µs.

3.4 Characterization of density fluctuations in TCABR

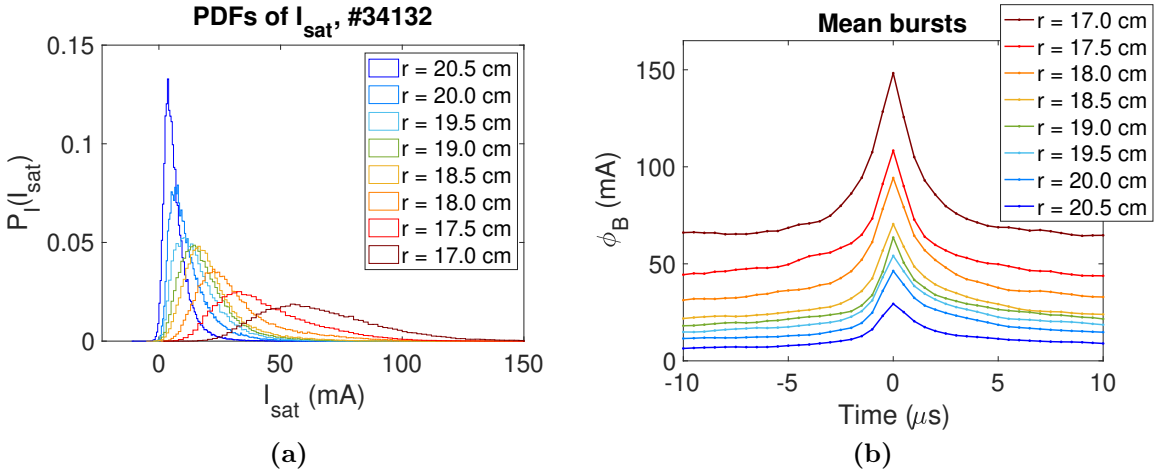
This subsection describes some features of the I_{sat} measurement. For this dissertation, experiment 34132 of TCABR is used as an example. It was coordinated by G. G. Grenfell [13,110]. In this experiment, the plasma had low magnetohydrodynamic (MHD) activity and, in the discharge steady phase (Section 3.2), the average electron density and the plasma current in the core were respectively $n_e \approx 1.1 \cdot 10^{19}$ m^{-3} and $I_p \approx 85$ kA (Fig. 3.3(b)). With the rake Langmuir probe (Section 3.1), I_{sat} signals at eight radial positions were simultaneously measured.

Table 3.1 exhibits the fixed parameters used for I_{sat} measurement and burst detection. The total duration of the analysis was $T = t_f - t_i = 40$ ms. The method used for identifying t_i and t_f was described in Section 3.2.

All probability distributions of I_{sat} measured in the experiment 34132 are displayed in Fig. 3.7(a). As it will be shown in Chapter 6, the analyzed I_{sat} PDFs of TCABR can be fitted as a convolution of Gamma and Gaussian distributions (presented in Subsection 2.1.3). In Fig. 3.7(a), the PDFs become thinner as r increases and their average value decreases. This expresses the reduction in plasma density with the radial position.

The conditionally averaged bursts also shrink with the increase of the radial position, as seen in Fig. 3.7(b). Their asymmetry increases with r too, having a quicker ascension and a slower fall as the position rises. Moreover, all appear to have a double-exponential shape.

On the other hand, Fig. 3.8 exhibits the radial profile of the number of detected bursts. The number of extreme events varied between 400 and 560. The profile has a well-defined peak at the last closed flux surface, $r = a = 18.0$ cm, indicating that bursts/blobs may be mostly produced at this position in TCABR.



(a) (b)

Figure 3.7: **(a)** I_{sat} PDFs for the 8 radial positions, r, measured in the discharge 34132. **(b)** Conditionally averaged bursts of the same experiment showing that the amplitude decreases with r. This is in accordance with Fig. 1.9, which shows a blob dissipation through the SOL.

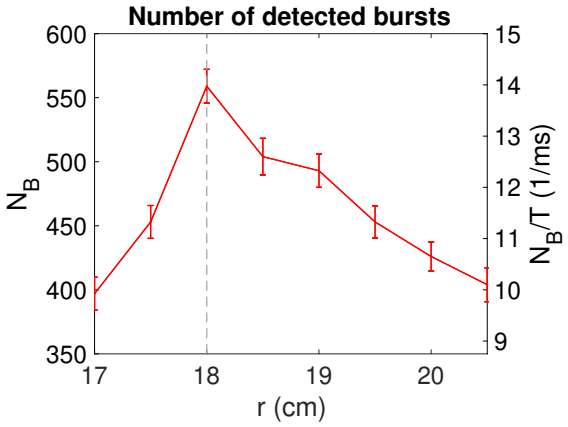


Figure 3.8: Radial profile of the number of detected bursts, in absolute values (left axis) and divided by the total analysis time, $T = 40$ ms (right axis). Threshold for burst detection: $\mathcal{T} = 2.5$. The profile has a peak at the minor radius $r = a = 18.0$ cm.

Another important diagnostic obtained from turbulence measurements in magnetized plasmas is the power spectral density (PSD, Section 2.4). Over different devices (tokamaks, helimaks and stellarators), the spectra of gas puffing imaging, saturation current and floating potential show a universal behavior: a flat low-frequency region, and for high-frequencies a steep power-law tail ($1/f^{2\sim3}$) [35, 39, 48, 114]. Fig. 3.9(a) shows a I_{sat} log-log spectrum for the TCABR tokamak, in agreement with the results from other machines. The inset of Fig. 3.9(a) depicts the same PSD but with a linear x-axis.

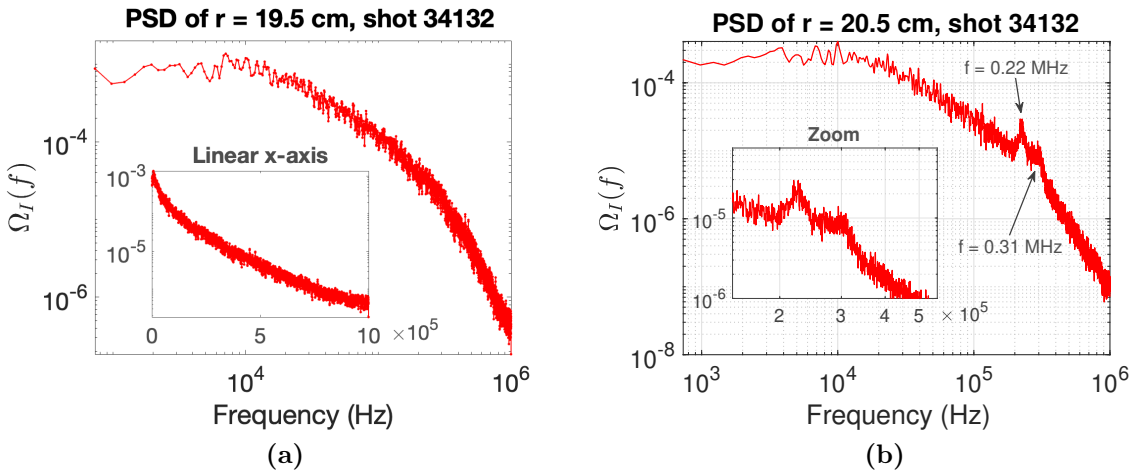
Power spectra can be used to detect periodical and correlated phenomena. In Ref. 48, the PSD was used to identify correlated bursts. Another example can be depicted in TCABR. At $r = 20.5$ cm, the position with the smallest density measured in the experiment 34132, two small peaks were observed in the power spectrum (Fig. 3.9(b)), one at $f_1 = 0.22$ MHz and another at $f_2 = 0.31$ MHz. In the time domain, they could correspond to periodic phenomena with a time scale of $1/f_1 = 4.5$ μs and $1/f_2 = 3.2$ μs. Oscillations of this order can indeed be seen in the saturation current and in the background of the conditionally averaged burst (Fig. 3.10).

In the context of magnetized plasmas, the parameter $\sigma_I / \langle I_{sat} \rangle$ has been called the relative fluctuation level [50]. In probability theory, it is also known as the coefficient of variation (see for ex. p. 380 of Ref. 75). Because turbulence is associated with fluctuations in density [20], $\sigma_I / \langle I_{sat} \rangle$ may be also referred to as the turbulence level, as done in Ref. 115. Fig. 3.11(a) shows its radial profile for experiment 34132 of TCABR. The turbulence level increases with the position, as in Alcator C-Mod [33] or Texas Helimak [115].[4]

A parabolic relation between kurtosis and skewness has been observed for various turbulent signals of magnetized plasmas, in special for density data [32, 33, 36, 50, 69, 71]. The most

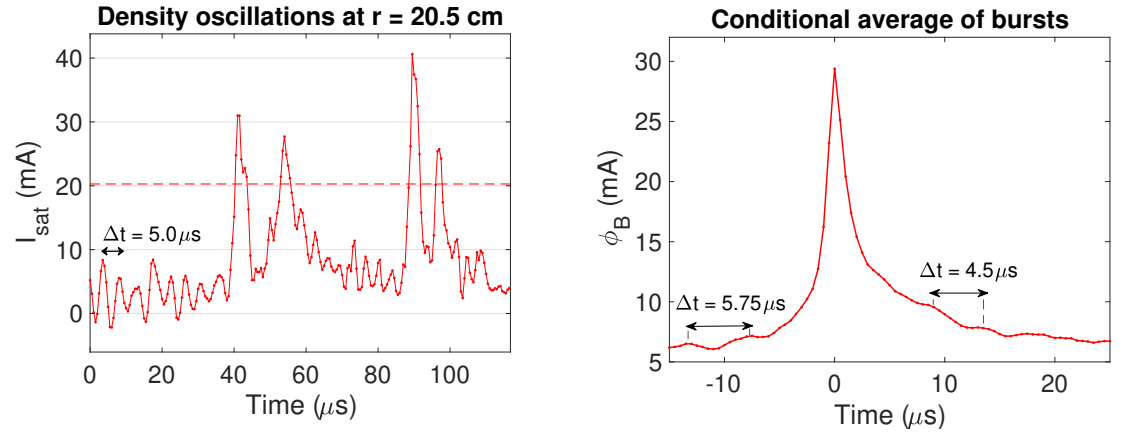
[4]The turbulence level has also been associated with the gradient of density [115], which is proportional to $d \langle I_{sat} \rangle / dr$. To see the profile of $\langle I_{sat} \rangle$, the reader is referred to Fig. 6.4(a).

common expression is $K_I = 3 + 1.5S_I^2$, where S_I and K_I are respectively the 3rd and 4th standardized moments of I_{sat}, as defined in Subsection 2.1.1. Fig. 3.11(b) shows reasonable agreement between this equation and data from the TCABR experiment 34132.



(a) (b)

Figure 3.9: **(a)** PSD of $r = 19.5$ cm. **(b)** I_{sat} power spectrum for position $r = 20.5$ cm in the experiment 34132 of TCABR. There are two unexpected bumps, one at $f = 0.22$ MHz and another at $f = 0.31$ MHz.



(a) (b)

Figure 3.10: **(a)** 120 μs of I_{sat} measured in the position $r = 20.5$ cm of the experiment 34132 of TCABR. The dashed line corresponds to the threshold $I_{sat} = \langle I_{sat} \rangle + 2.5\sigma_I$. We can see oscillations of period around $\Delta t = 5.0$ μs, compatible with the bump of $f = 0.22$ MHz in the power spectrum (Fig. 3.10(a)). **(b)** The respective conditional average of the bursts exhibits small oscillations with period comparable to the ones referred in the PSD, since $(4.5\ \mu\text{s})^{-1} = 0.22$ MHz.

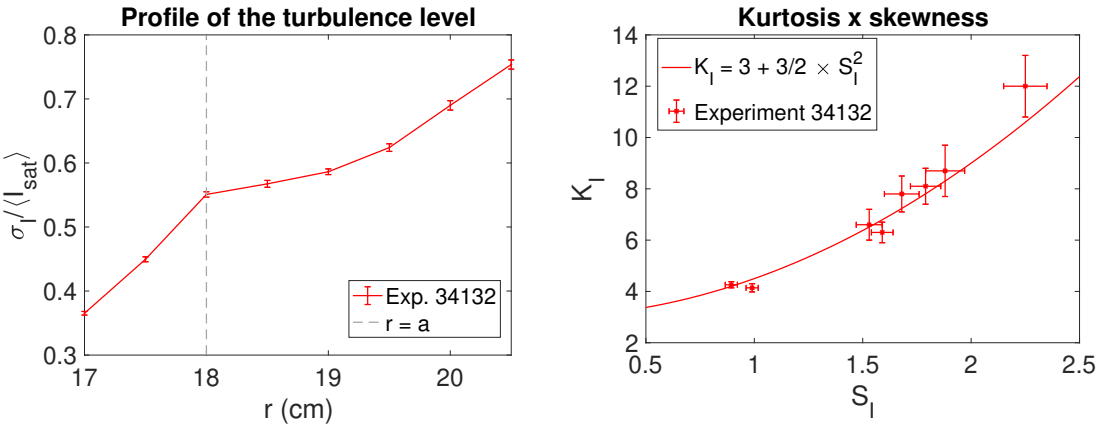


Figure 3.11: **(a)** Radial profile of the turbulence level, $\sigma_I / \langle I_{sat} \rangle$. The dashed line indicates the minor radius position, $a = 18.0$ cm. **(b)** Parabolic relation between kurtosis and skewness from I_{sat}. Uncertainties were estimated with repeated simulations (Chapters 4–6). Experiment 34132 was used.

Chapter 4

Stochastic pulse train model with correlated noise

In this Chapter, the stochastic pulse train model with correlated noise will be introduced. Central for this dissertation, it will be used in all following chapters.

4.1 The model

4.1.1 Background-less case

In the past decade, a stochastic model was developed [50–55] to describe density fluctuations in single-point measurements of magnetically confined plasmas. The model is referred by many names, such as: Garcia Model [40, 41], shot noise process [50, 54–57, 118], Filtered Poisson Process (FPP) [54–57], train of filaments [56, 57], model of bursts-train turbulence [48] and stochastic pulse train model (SPTM) [49].[1] In this dissertation, we shall adopt the last one. The SPTM is central for this work and will be used in the following sections.

The general idea of the stochastic pulse train model consists of superposing pulses with randomly distributed amplitudes, durations and arrival times, in order to reproduce average proprieties of the experimental density fluctuations. In its most used form, the pulses are assumed to be uncorrelated, with the same shape of the conditionally averaged burst (Fig. 3.4(b)). That is, in one of its simplest descriptions, the SPTM for I_{sat} is

$$I_{sat}(t) = \sum_{j=1}^{N_P} A_j \phi\left(\frac{t - t_j}{\tau_d}; \lambda\right) \tag{4.1}$$

where N_P is the number of pulses spawned in the time interval considered, A_j and t_j are respectively the amplitude and time occurrence of the j-th pulse, τ_d is the characteristic duration of the pulses, $\lambda \in [0, 1]$ is the pulse asymmetry parameter and $\phi(\theta; \lambda)$ is the pulse shape function. τ_d and λ are assumed to be the same for all pulses in this description.

In Fig. 4.1 a fit of the conditional average of bursts is shown with a double-exponential

[1] Also, following Refs. 51 and 116, the model ancestors correspond to Refs. 117–119.

waveform for $\phi(\theta;\lambda)$. This suggests that the function adopted for (4.1) should be

$$\phi(\theta;\lambda) = \begin{cases} \exp\left(\theta/\lambda\right) & , \ \theta < 0 \\ \exp\left[-\theta/\left(1-\lambda\right)\right] & , \ \theta \geq 0 \end{cases} \tag{4.2}$$

When $\lambda = 0.5$, the pulses are symmetric, and usually in TCABR this is the case for radial positions inside the plasma column ($r < a = 18.0$ cm), as seen in Fig. 5.4 and 5.5 of Ref. 15. If $\lambda \to 0$, the rise tends to be instantaneous, and such approximation is considered in Refs. 32, 50 and 53. In the scrape-off layer ($r > a$), the conditionally averaged burst shape is usually asymmetric in a lot of devices, with $\lambda \in [0.10, 0.35]$ [15, 30, 31, 33–35, 40]. Sometimes it is convenient to use τ_r and τ_f instead of λ and τ_d, where τ_r is the characteristic duration of the pulse rise and τ_f, of the fall. That is, $\tau_d = \tau_r + \tau_f$ and $\lambda = \tau_r/\tau_d$.

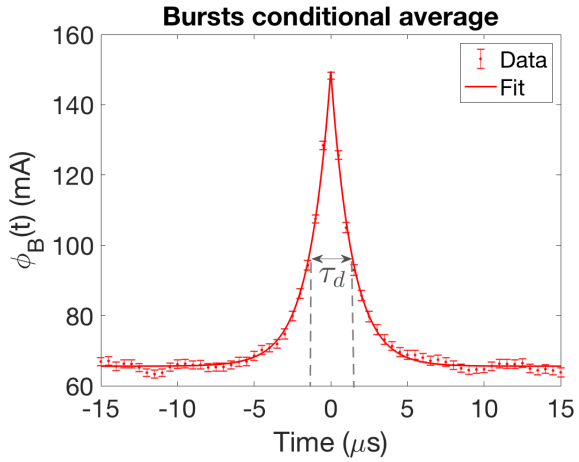


Figure 4.1: Fit of the conditionally averaged burst (position $r = 17.0$ cm), using the model of two exponentials: $\phi_B = A_0 + A_1\phi(t/\tau_d;\lambda)$. A_0 is the background, A_1 is the mean burst amplitude and ϕ is given by Eq. (4.2). The fit yielded $\tau_d = 2.85(5)$ µs and $\lambda = 0.506(14)$.

Supported by vast experimental evidence [30, 32, 36, 37, 40, 41, 48], the distribution of amplitudes, A_j in Eq. (4.1), is assumed to be exponential (Eq. (14) of Ref. 50):

$$P_A(A_j) = \frac{1}{\langle A\rangle}\exp\left(-\frac{A_j}{\langle A\rangle}\right) \tag{4.3}$$

where $\langle A\rangle$ is the average amplitude. On the other hand, provided with the same references, the arrival times t_j of the pulses are supposed to be uncorrelated and uniformly distributed—that is, $P_t(t_j) = 1/T$, where T is the time interval considered in the analysis.

Since the time occurrence of the pulses is uniform and uncorrelated, the resultant number of pulses K in an interval T follows a Poisson distribution [51],

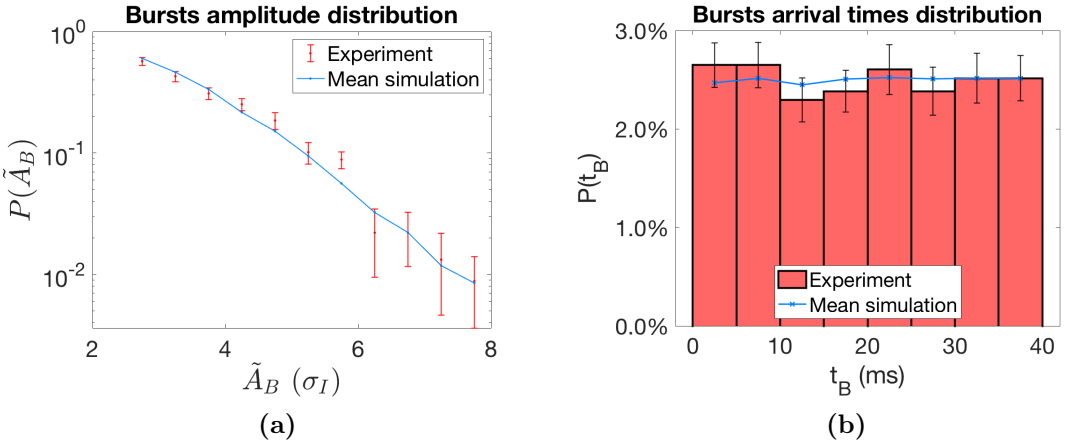
$$P_K(K) = \frac{1}{K!}\left(N_P\right)^K e^{-N_P}$$

where N_P is the expected value, $N_P = \langle K\rangle$. In the simulations we will fix $K = N_P$, to diminish statistical fluctuations without losing relevant information.

The fact that the pulses are uncorrelated and uniformly distributed also implies that the waiting time Δt_w between pulses follows an exponential distribution (Eq. (4) of Ref. 50):

$$P_w(\Delta t_w) = \frac{1}{\tau_w} \exp\left(-\frac{\Delta t_w}{\tau_w}\right) \tag{4.4}$$

where $\tau_w = \langle \Delta t_w \rangle = T/N_P$ is the average waiting time.



(a) **(b)**

Figure 4.2: Experimental and simulated distributions of **(a)** burst amplitude and **(b)** burst arrival time. The 493 bursts that composed the distributions were detected as events with peaks higher than $\tilde{I}_{sat} > 2.5$. Position used: $r = 19.5$ cm.

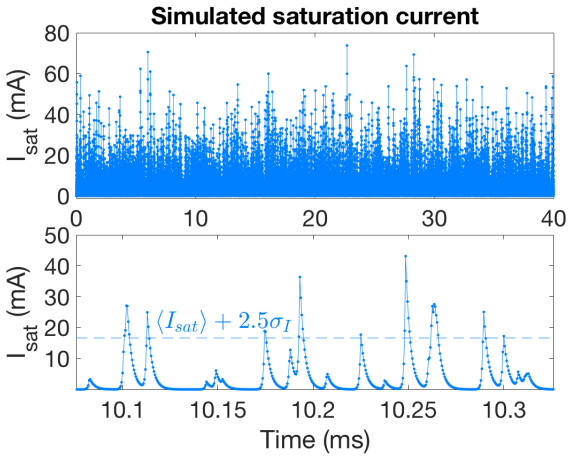


Figure 4.3: 40 ms of a simulated saturation current made with the stochastic pulse train model (Eq. (4.1) and (4.2)). The bottom panel exhibits a 0.25 ms zoom of the signal.

The assumptions about the pulse amplitude and arrival time also hold for TCABR. Fig. 4.2 compares experimental and simulated distributions of burst amplitudes and arrival times for this device. The blue lines seen in Fig. 4.2 correspond to the mean of 50 simulations, just

to diminish the statistical fluctuation. (Further information about how the simulations are obtained will be given in Chapters 5 and 6.) It is seen in Fig. 4.2(a) that the experimental and simulated burst amplitude distributions are consistent with an exponential (a line in the logarithm scale). Moreover, Fig. 4.2(b) indicates that the burst arrival times agree with a uniform distribution.

Fig. 4.3 depicts an example of simulation with the background-less stochastic pulse train model from Eq. (4.1) and (4.2). The parameters used were $N_P/T = 100$ ms^{-1}, $\langle A \rangle = 10$ mA, $\tau_d = 3$ µs and $\lambda = 0.2$.

4.1.2 Adding a background of correlated noise

Previous works [35,41,48,49,54,55] reported the need of adding some kind of noise background in the SPTM to correctly fit the signal PDF and to account for diffusive transport. The most common choice is to use Gaussian noise. With this, the saturation current becomes

$$I_{sat}(t) = I_B(t) + I_{\mathcal{N}}(t) \tag{4.5}$$

where $I_B(t)$ is the pulse train, which gives rise to the bursts, and is given by Eq. (4.1),

$$I_B(t) = \sum_{j=1}^{N_P} A_j \phi\left(\frac{t - t_j}{\tau_d}; \lambda\right) \tag{4.6}$$

In its turn, $I_{\mathcal{N}}(t)$ is the Gaussian noise background, having the probability distribution

$$P_{\mathcal{N}}(I_{\mathcal{N}}(t)) = \frac{1}{\sqrt{2\pi\sigma_{\mathcal{N}}^2}} \exp\left[-\frac{\left(I_{\mathcal{N}}(t) - \langle I_{\mathcal{N}}(t)\rangle\right)^2}{2\sigma_{\mathcal{N}}^2}\right] \tag{4.7}$$

The use of correlated Gaussian noise was considered in Refs. 48 and 54, in the first with colored noise (CN), and in the second with an Ornstein–Uhlenbeck (OU) process. In this dissertation, the results of both noises will be compared. They were briefly described in Section 2.6. In resume, the correlation of the Gaussian noise does not alter the PDF of the signal. As stated in Subsection 2.6.1, colored noise can be defined by the relation of proportionality between the power spectrum and frequency,

$$\Omega_{\mathcal{N}}(f) \propto \frac{1}{f^{\eta}}$$

where $\eta \in [0, 2]$ the noise exponent. The case of uncorrelated noise ($\eta = 0$) has been used with the SPTM in some cases [34, 40, 41, 49, 54].

Ornstein-Uhlenbeck processes, on the other hand, have power spectra (Eq. (2.29))

$$\Omega_{\mathcal{N}}(\omega) = \frac{2\tau_{\mathcal{N}}\sigma_{\mathcal{N}}^2}{1 + \tau_{\mathcal{N}}^2\omega^2} + 2\pi\delta(\omega)\langle I_{\mathcal{N}}\rangle^2$$

where $\omega = 2\pi f$, $\delta(\omega)$ is the Dirac distribution and $\tau_{\mathcal{N}}$ is the correlation time. Thus, colored and OU noise can be described by three parameters. The mean and variance, $\langle I_{\mathcal{N}}\rangle$ and $\sigma_{\mathcal{N}}^2$, and the correlation parameter (η for CN and $\tau_{\mathcal{N}}$ for OUN).

In this thesis, we shall deepen the discussion of Refs. 48 and 54 and present evidence of why correlated noises are better than uncorrelated ones to describe the background of plasma density fluctuations in tokamaks. Furthermore, it will also be argued that OUN has some advantages over CN.

Table 4.1 groups the parameters of the pulse-train model with correlated (colored or OU) noise. Sometimes, instead of presenting N_P/T and $\sigma_\mathcal{N}$, the model is also given in terms of the parameter $\gamma = \tau_d N_P/T$ and the noise level $\epsilon = \sigma^2_\mathcal{N}/\sigma^2_B = \sigma^2_\mathcal{N}/(\gamma \langle A\rangle^2)$, as in Ref. 54.

Table 4.1: Parameters of the stochastic pulse train model with correlated noise (OUN or CN), with reference equations and values range obtained in TCABR.

Symbol	Meaning	Eq.	Values range
N_P/T	Number of pulses per time	(4.6)	$[50, 10^3]$ ms^{-1}
τ_d	Characteristic duration of the pulses	(4.6)	$[2, 6]$ µs
λ	Pulse asymmetry parameter	(4.2)	$[0.1, 0.6]$
$\langle A\rangle$	Average amplitude of the main pulses	(4.3)	$[5, 19]$ mA
$\langle I_\mathcal{N}\rangle$	Average of the Gaussian noise	(4.7)	$[2, 18]$mA
$\sigma_\mathcal{N}$	Standard deviation of the Gaussian noise	(4.7)	$[1, 6]$ mA
$\tau_\mathcal{N}$	Correlation time of the OU noise	(2.28)	$[0, 6]$ µs
η	Noise exponent of the colored noise	(2.23)	$[0, 1]$

4.2 Probability distribution

For the background-less model of I_{sat} (Subsection 4.1.1), the probability distribution function is well known to be a Gamma distribution [50, 52, 55]:

$$P_I(I_{sat}) = \frac{1}{\langle A\rangle \Gamma(\gamma)} \left(\frac{I_{sat}}{\langle A\rangle}\right)^{\gamma-1} \exp\left(-\frac{I_{sat}}{\langle A\rangle}\right) \tag{4.8}$$

where $\langle A\rangle$ and γ are respectively the scale and shape parameters, and Γ is the Gamma function:

$$\Gamma(\gamma) \equiv \int_0^\infty x^{\gamma-1} e^{-x} dx \tag{4.9}$$

Further details about the Gamma distribution were given in Subsection 2.1.1.

The parameter γ is the characteristic duration of the pulses multiplied by the number of pulses per time (Table 4.1),

$$\gamma \equiv \tau_d N_P/T = \tau_d/\tau_w \tag{4.10}$$

γ is known as the intermittency parameter [38, 49–52, 116]. "Overlap parameter" might be a more appropriate name for this quantity, because the overlapping of the pulses increases

with γ, while the intermittency decreases.[2] Nevertheless, we chose to follow the conventional nomenclature.

The Gamma distribution (Eq. (4.8)) has mean, variance, skewness and kurtosis given respectively by [50]

$$\begin{aligned} \langle I_{sat} \rangle &= \gamma \langle A \rangle \\ \sigma_I^2 &= \gamma \langle A \rangle^2 \\ S_I &= \frac{2}{\gamma^{1/2}} \\ K_I &= 3 + \frac{6}{\gamma} \end{aligned} \tag{4.11}$$

By adding Gaussian noise to the Gamma-distributed pulses, the above moments become [54]

$$\begin{aligned} \langle I_{sat} \rangle &= \gamma \langle A \rangle + \langle I_{\mathcal{N}} \rangle \\ \sigma_I^2 &= \gamma \langle A \rangle^2 + \sigma_{\mathcal{N}}^2 \\ S_I &= \frac{2}{\gamma^{1/2} (1+\epsilon)^{3/2}} \\ K_I &= 3 + \frac{6}{\gamma (1+\epsilon)^2} \end{aligned} \tag{4.12}$$

where ϵ is the noise parameter,

$$\epsilon \equiv \frac{\sigma_{\mathcal{N}}^2}{\sigma_B^2} = \frac{\sigma_{\mathcal{N}}^2}{\gamma \langle A \rangle^2}$$

The SPTM with Gaussian noise ($I_{sat} = I_B + I_{\mathcal{N}}$) also has a known probability distribution, for the case of null noise mean ($\langle I_{\mathcal{N}} \rangle = 0$). The answer is Eq. (A6) of Ref. 54:

$$\begin{aligned} P_I(I_{sat}) = \frac{(\gamma\epsilon)^{\gamma/2-1}}{2^{\gamma/2}\langle A \rangle} \times \exp\left(-\frac{I_{sat}^2}{2\gamma\epsilon\langle A \rangle^2}\right) \\ \times \left\{ \frac{(\gamma\epsilon)^{1/2}}{2^{1/2}\Gamma\left((1+\gamma)/2\right)} \times M\left(\frac{\gamma}{2}, \frac{1}{2}; \frac{1}{2\epsilon}\left(\frac{I_{sat}}{\gamma^{1/2}\langle A \rangle} - \gamma^{1/2}\epsilon\right)^2\right) \right. \\ \left. + \frac{\gamma^{1/2}}{\Gamma(\gamma/2)}\left(\frac{I_{sat}}{\gamma^{1/2}\langle A \rangle} - \gamma^{1/2}\epsilon\right) \times M\left(\frac{1+\gamma}{2}, \frac{3}{2}; \frac{1}{2\epsilon}\left(\frac{I_{sat}}{\gamma^{1/2}\langle A \rangle} - \gamma^{1/2}\epsilon\right)^2\right) \right\} \end{aligned} \tag{4.13}$$

where $M(a, b; x)$ is the confluent hypergeometric function of the first kind, with parameters a and b and variable x [74]:

$$M(a, b; x) \equiv \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b)}{\Gamma(b+n)\Gamma(a)} \frac{x^n}{n!}$$

[2] One may ask what "intermittency" exactly means. Quoting the *Encyclopedia of Nonlinear Science* [121], if a signal $I(t)$ "exhibits segments of relative constant values (laminar phase) interspersed by erratic bursts, we say the system dynamics is intermittent". So, following Refs. 50 and 51, a strong intermittency regime occurs when the pulse overlap $\gamma = \tau_d/\tau_w$ is small, since in this case there will be long calm phases, interrupted by sudden bursts with peak much higher than the signal average.

The probability distribution for the correlated noise model (Eq. (4.13)) is exceedingly complicated and hence is not of much use. Instead, it is preferable to work with the PDF characteristic function, as it will be shown in the next section.

4.3 Characteristic function

The characteristic function (CF) was defined in Section 2.2. Following Ref. 50, the CF from the Gamma distribution (Eq. (4.8)) is

$$C_I(u) = (1 - iu\,\langle A\rangle)^{-\gamma} \tag{4.14}$$

while the CF from the Gaussian distribution (Eq. (4.7)) is

$$C_{\mathcal{N}} = \exp\left(i\,\langle I_{\mathcal{N}}\rangle\,u - \frac{1}{2}\sigma_{\mathcal{N}}^2 u^2\right) \tag{4.15}$$

As stated in Section 2.2, the CF of a sum of random variables is the product of the original CFs. That is, for the SPTM with correlated noise (Eq. (4.5)),

$$I_{sat}(t) = I_B(t) + I_{\mathcal{N}}(t) \implies$$

$$C_I(u) = C_B(u)C_{\mathcal{N}}(u)$$

Therefore the CF for a sum of Gamma and Gaussian random variables is

$$C_I(u) = (1 - i\,\langle A\rangle\,u)^{-\gamma} \exp\left(i\,\langle I_{\mathcal{N}}\rangle\,u - \frac{1}{2}\sigma_{\mathcal{N}}^2 u^2\right) \tag{4.16}$$

which is simpler than the PDF of $I_B + I_{\mathcal{N}}$ (Eq. (4.13)). Following Ref. 55, a fit of Eq. (4.16) will be used in this work to estimate the parameters γ, $\langle A\rangle$, $\langle I_{\mathcal{N}}\rangle$ and $\sigma_{\mathcal{N}}$ from a given experimental I_{sat}.

4.4 Power spectral density

In this section will be listed the power spectral density (PSD) for the stochastic pulse train model (SPTM) with colored noise (CN) and Ornstein-Uhlenbeck noise (OUN). Details about the derivations are given in Appendix B.1.

As in Subsection 4.1.2, the saturation current is modeled as a sum of burst pulses plus correlated Gaussian noise,

$$I_{sat} = I_B(t) + I_{\mathcal{N}}(t)$$

Then, from Subsection 2.6.2 and Refs. 53 and 54, results that the PSD for the SPTM with OUN is given by

$$\Omega_I^{OUN}(\omega) = \frac{2\tau_d\sigma_B^2}{\left(1+\tau_r^2\omega^2\right)\left(1+\tau_f^2\omega^2\right)} + \frac{2\tau_{\mathcal{N}}\sigma_{\mathcal{N}}^2}{1+\tau_{\mathcal{N}}^2\omega^2} + 2\pi\,\langle I_{sat}\rangle^2\,\delta(\omega)$$

where $\tau_r = \lambda\tau_\mathrm{d}$ and $\tau_f = (1-\lambda)\tau_\mathrm{d}$ are respectively the characteristic durations of the pulse rise and fall. $\omega = 2\pi f$ is the angular velocity and $\delta(\omega)$ is the Dirac distribution.

For the case with colored noise, the power spectrum is ill-defined and just an approximate relation for $f \gg 10^2$ Hz can be given (again, using Subsection 2.6.1 and Refs. 53 and 54),

$$\Omega_I^{CN}(\omega) \simeq \frac{2\tau_d\sigma_B^2}{\left(1+\tau_r^2\omega^2\right)\left(1+\tau_f^2\omega^2\right)} + \frac{C(\eta)}{\omega^\eta}\Delta t\cdot\sigma_\mathcal{N}^2 + 2\pi\left\langle I_{sat}\right\rangle^2\delta\left(\omega\right)$$

where $C(\eta) \simeq 10^{5.73\cdot\eta^{0.820}}$.

4.5 Mean simulations

With simulations, we aim to understand concepts underlying experimental data. Thus, it is of no interest to have large statistical fluctuations in the functions of these synthetic realizations—since this noise would only difficult the analysis. For this reason, in Chapter 6 the experimental functions of I_{sat} will be compared to averaged simulated functions. This means that simulated diagnostics of PDFs, PSDs and others will be presented as the average of several realizations with the same inputs.

Nevertheless, it is important to show the difference between unitary and averaged simulated diagnostics at least once. This is the purpose of Fig. 4.4, which exhibits a comparison of the conditional average of bursts, the PDF, the power spectrum, and the bursts distributions of waiting time and amplitude. 100 repetitions were used. This averaging procedure was also employed in Fig. 4.2.

The inputs used in the simulations of Fig. 4.4 were the following. For the main pulses,

$$N_P/T = 0.4\ \mu\mathrm{s}^{-1}, \quad \langle A\rangle = 7\ \mathrm{mA}, \quad \tau_d = 5\ \mu\mathrm{s}, \quad \lambda = 0.1$$

For the Ornstein-Uhlenbeck process,

$$\langle I_\mathcal{N}\rangle = 2\ \mathrm{mA}, \quad \sigma_\mathcal{N} = 1\ \mathrm{mA}, \quad \tau_\mathcal{N} = 0.5\ \mu\mathrm{s}$$

These values are similar to the ones found in the scrape-off layer of TCABR (more specifically in the position $r = 19.5$ cm of experiment 34132). The meaning of each parameter is described in Table 4.1. The uncertainties of the graphs in Fig. 4.4 were estimated with the standard deviations of the simulations from each diagnostic.

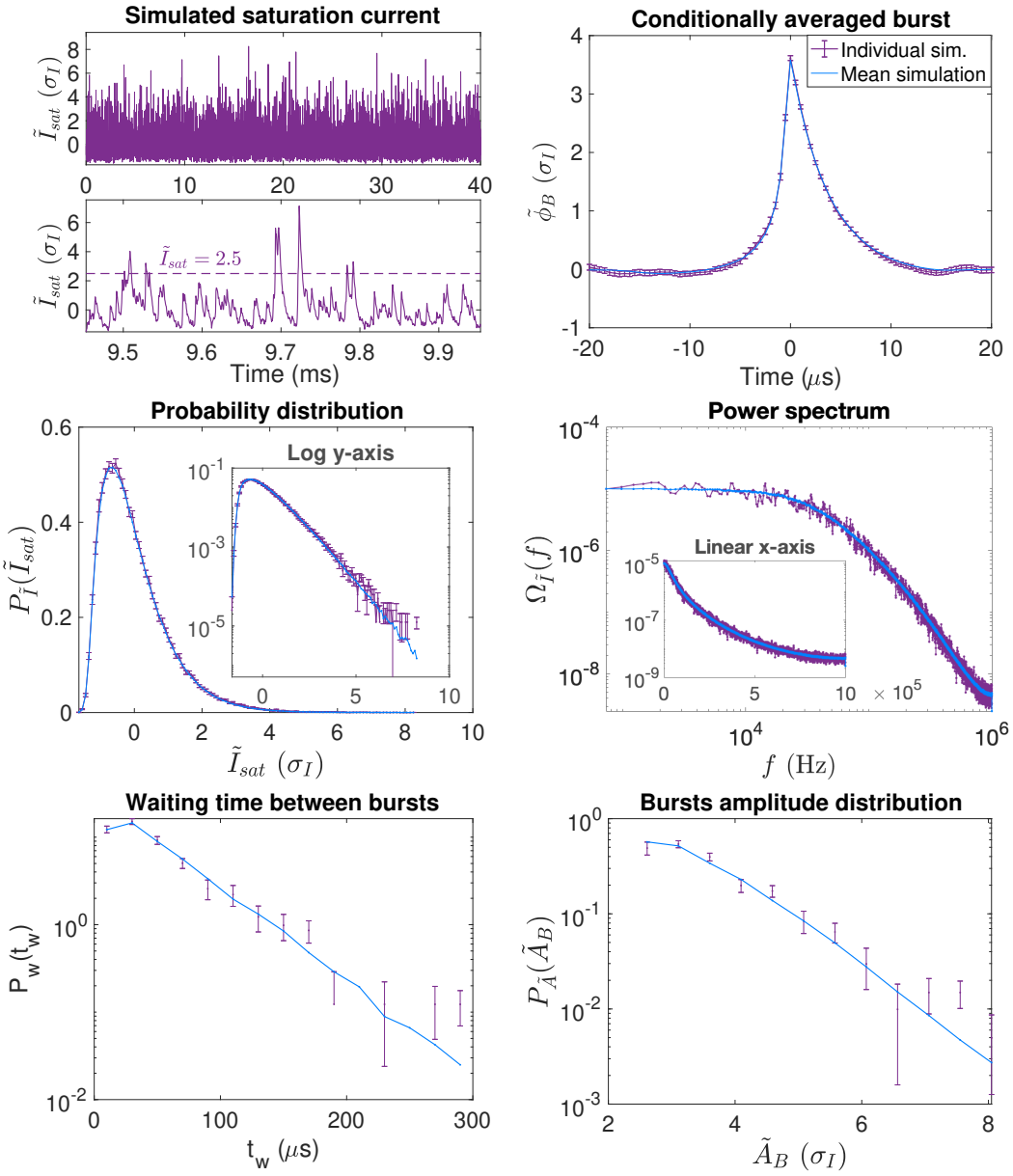


Figure 4.4: Comparison of unitary and averaged simulated diagnostics (respectively in purple and blue). Normalization: $\tilde{I}_{sat} = (I_{sat} - \langle I_{sat} \rangle) / \sigma_I$.

Chapter 5

Fitting the SPTM with correlated noise

This chapter introduces the fitting method developed for the stochastic model with noise.

5.1 Introduction

This section initiates the thesis main results. Here will be described the fitting method we developed to adjust the stochastic pulse train model (SPTM) with Gaussian noise—the noise being uncorrelated, colored or made with Ornstein-Uhlenbeck processes. The method consists of fitting the seven SPTM parameters (Table 5.1) in three steps.

Table 5.1: Parameters adjusted by the CF-CAB-PSD fit.

Symbol	Name	Eq.	Fitted by
γ	Intermittency parameter	(4.10)	CF
$\langle A \rangle$	Mean amplitude of the pulses	(4.3)	
$\langle I_{\mathcal{N}} \rangle$	Noise average	(4.7)	
$\sigma_{\mathcal{N}}$	Noise standard deviation	(4.7)	
λ	Pulse asymmetry parameter	(4.2)	CAB
τ_d	Pulse characteristic duration	(4.2)	
$\tau_{\mathcal{N}}$	Noise correlation time	(2.28)	PSD

First, a covariant least-squares fit[1] of the characteristic function of I_{sat} evaluates the parameters of the PDF (γ, $\langle A \rangle$, $\langle I_{\mathcal{N}} \rangle$, $\sigma_{\mathcal{N}}$), taking advantage of the fact that the CF does not depend on the other three parameters. Then, the shape parameters of the bursts (τ_d and λ) are adjusted with a χ^2 map of the conditionally averaged waveform (Section 3.3). The conditional average of bursts (CAB) is independent of the correlation parameter—since this is a background feature, while the CAB only detects extreme events. Lastly, the correlation parameter (η or $\tau_{\mathcal{N}}$) is fitted with a χ^2 map of the frequency spectrum of I_{sat} (Section 2.4). The whole three-step procedure will be referred to as the CF-CAB-PSD fit.

The position $r = 19.5$ cm of the TCABR experiment 34132 will be used as an example of the fit. To fit the stochastic pulse train model with correlated noise, the data must have similar behavior to the one expected by the model (Section 4.1). In special, a convolution

[1] For details about the generalized least-squares fit, see for example Appendix B of Ref. 15.

of a Gamma and a Gaussian distribution should adequately describe the PDF of the density signal.

5.2 Fit of the characteristic function

As stated in 4.2, the PDF of a convolution of a Gamma and a Normal distribution (i.e., PDF of $I_{sat}(t) = I_B(t) + I_{\mathcal{N}}(t)$), is too intricate to fit. Thus, it is preferable to adjust the signal characteristic function, as proposed in Ref. 55. The CF has the same information as the PDF (Section 2.2). Following Section 4.3, the CF of the stochastic pulse train model with Gaussian noise is (Eq. (4.16)),

$$C_I(u) = (1 - i\,\langle A\rangle\,u)^{-\gamma} \exp\left(i\,\langle I_{\mathcal{N}}\rangle\,u - \frac{1}{2}\sigma_{\mathcal{N}}^2 u^2\right)$$

The CF of a discrete signal can be estimated by the empirical characteristic function [55,77],

$$C_I(u) = \frac{1}{N_{pt}} \sum_{j=1}^{N_{pt}} e^{iuI_j} \tag{5.1}$$

where I_j is the j-th point of I_{sat}.

The CF is a complex-valued function. However, the parts with $u > 0$ and $u < 0$ carry the same information, since $C_I(u) = C_I(-u)^*$. Thus, to use a least-squares routine, we defined the effective characteristic function

$$\mathcal{C}_I(u) \equiv \begin{cases} \text{Re } C_I(u), & u \leq 0. \\ \text{Im } C_I(u), & u > 0. \end{cases}$$

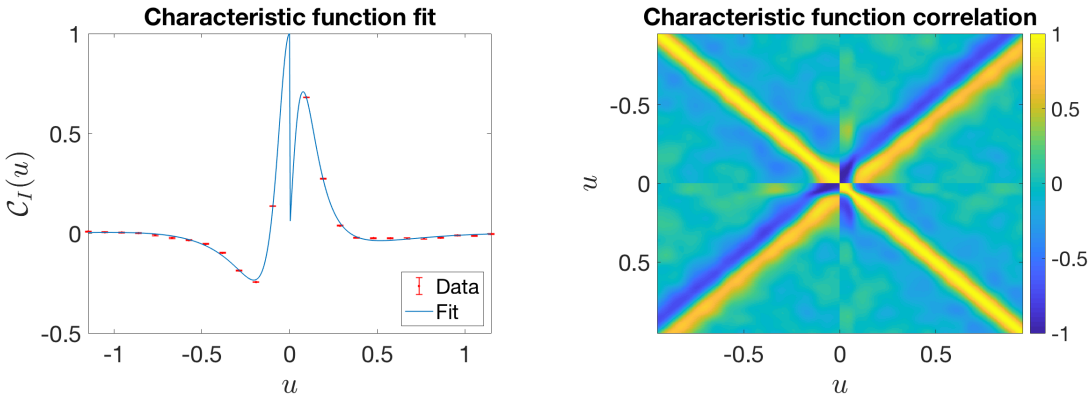
$\mathcal{C}_I(u)$ is real-valued. The real part of the CF is put in the negative part, taking advantage of Re $C_I(u) =$ Re $C_I(-u)$. Fig. 5.1(a) shows an example of fit of the $\mathcal{C}_I(u)$, for the radial position $r = 19.5$ cm in TCABR.

The characteristic function has non-negligible correlated points, as can be seen in the "continuous" correlation matrix of Fig. 5.1(b).[2] For this reason, it is mandatory to choose just a few points u to construct the empirical CF of Eq. (5.1), as otherwise correlated points near the diagonals of the correlation matrix can make it non-invertible (determinant 0), precluding the use of the least-squares [122]. We chose an array with 12 points,

$$\vec{u} = \frac{u_{1\%}}{12}(1, 2, ..., 12)$$

where $u_{1\%}$ means $|C_I(u_{1\%})| = 1\%$. From the point of view of $\mathcal{C}_I(u)$, there are 24 points in total (12 for the real part and 12 for the imaginary part).

[2]The terms of the correlation matrix are given by Eq. (2.15), where in this case $\mathbf{x} = \mathcal{C}_I(u)$. The correlations were estimated using repeated simulations of $\mathcal{C}_I(u)$.



(a) (b)

Figure 5.1: Fit of the characteristic function and its correlation matrix (with high resolution) for the position $r = 19.5$ cm in TCABR.

5.3 Chi-square map of the conditionally averaged burst

Having obtained γ, $\langle A \rangle$, $\langle I_{\mathcal{N}} \rangle$ and $\sigma_{\mathcal{N}}$ with the CF fit, it is possible to adjust τ_d and λ with the conditional average of bursts (CAB). However, as shown in Appendix B.2, a direct fit of the burst conditional average may distort the true values of τ_d and λ, due to effects of pulse overlap and finite time resolution. For this reason, we chose to compare experiments and simulations, as both are influenced by these effects.

To make this comparison, we used χ^2 maps (Section 2.9). The chi-square can be defined as (Eq. (2.37))

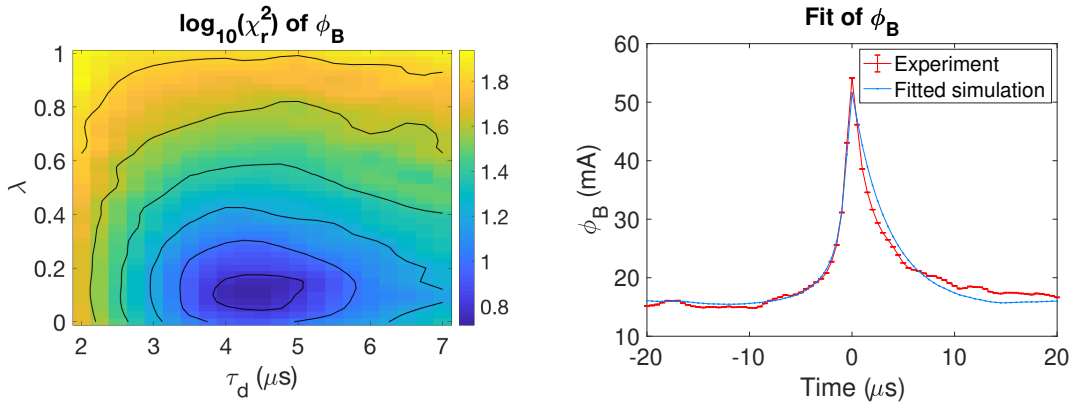
$$\chi^2 = \sum_{j=1}^{N} \left(\frac{y_j - f_j(\boldsymbol{\theta})}{\sigma_j} \right)^2 \tag{5.2}$$

where y_j are the points from the experimental CAB, f_j are the points from the simulated CAB, σ_j is the total uncertainty associated with y_j and f_j and $\boldsymbol{\theta} = (\tau_d, \lambda)$ is the vector of parameters to be fitted. We want to find the pair $\boldsymbol{\theta} = (\tau_d, \lambda)$ which minimizes the χ^2. $N = 81$ points (40 μs) were used to make the conditional average in TCABR. The uncertainties σ_j are estimated with repeated Monte Carlo simulations.

Fig. 5.2(a) shows a χ^2 map of the CAB for the position $r = 19.5$ cm in TCABR.[3] The map has a parabolic format. An estimate of the true parameters is obtained interpolating the χ^2 values near the minimum. For this case, the obtained parameters were $\tau_d = 4.57(32)$ μs and $\lambda = 0.105(13)$, where the uncertainties (in parenthesis) were estimated as the standard deviation of repeated fits. With the blue curve, Fig. 5.2(b) presents the corresponding adjust of the CAB. Although the experimental conditional average is not perfectly fitted by a double

[3] In general, the asymmetry parameter λ ranges from 0 to 0.5 in tokamaks [30–42]. In Fig. 5.2, $\lambda \in [0, 1]$ was displayed just to highlight the parabolic format of the map.

exponential, it is at least a reasonable approximation.



(a) (b)

Figure 5.2: **(a)** χ^2 map for the conditional average of bursts, ϕ_B, showing a minimum around the region $(\tau_d, \lambda) = (4.25\ \mu\text{s}, 0.10)$. The black contours highlight the parabolic shape of the map. (Log scale was used to improve visualization, and the χ^2 was reduced by the degrees of freedom, $\chi^2_r = \chi^2/N$.) **(b)** The corresponding fit of the conditionally averaged burst. Position in TCABR: $r = 19.5$ cm.

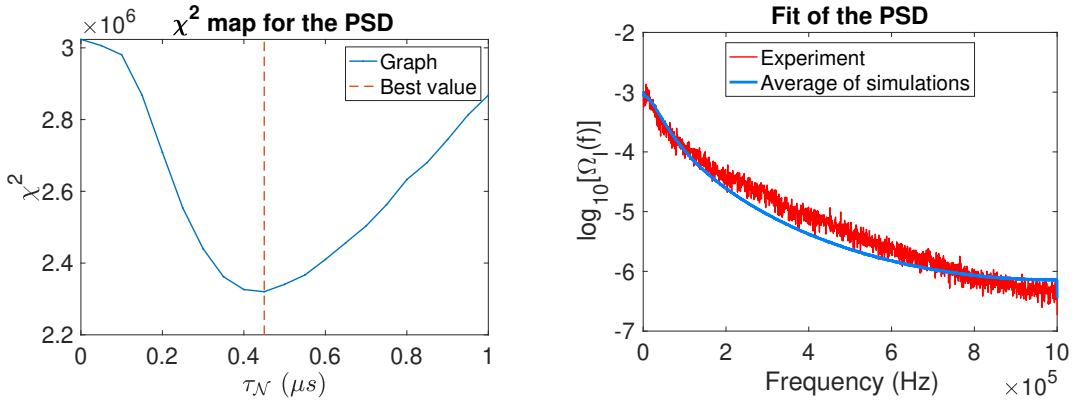
5.4 Chi-square map of the power spectrum

With γ, $\langle A \rangle$, $\langle I_\mathcal{N} \rangle$ and $\sigma_\mathcal{N}$ obtained by the characteristic function (Section 5.1) and τ_d and λ, by the conditional average of bursts (Section 5.3), it only lasts to adjust the correlation parameter. In the case of signals modeled with Ornstein-Uhlenbeck noise, this quantity is the relaxation time $\tau_\mathcal{N}$, whereas for the colored noise it is the noise exponent η (Sections 2.6 and 4.1.2). The correlation parameter can be fitted by the power spectral density (Sections 2.4 and 4.4).

As stated in Ref. 54, trying to fit the experimental PSD directly can yield biased results, since the sample PSD differs from the analytical one near the Nyquist frequency $f_{\text{Ny}} = 1/(2\Delta t)$. Furthermore, the colored noise model doesn't even have an exact expression for the PSD (Subsection 2.6.1). For these reasons, it is preferable to fit the frequency spectrum with χ^2 maps made from experiments and simulations, since in this case only sample spectra are compared and analytical expressions are not needed.

The chi-square adjustment was already explained in Section 5.3. Now, the data to be fitted is the logarithm of the PSD, $y = \log_{10} \Omega_I(\omega)$ and the function to adjust is the corresponding PSD from the simulated signal. The fitted parameter is the correlation time $\theta = \tau_\mathcal{N}$ (or the noise exponent η for the colored noise case). Each frequency spectra was set to have

$N = 2^{12} = 4096$ points.[4] Fig. 5.3(a) shows an example of adjusted χ^2 curve, while Fig. 5.3(b) compares the experimental PSD and its simulated counterpart that was fitted via the chi-square.



(a) **(b)**

Figure 5.3: **(a)** χ^2 map of the PSD diagram, as a function of the noise correlation time $\tau_{\mathcal{N}}$ and with a minimum around $\tau_{\mathcal{N}} = 0.45$ µs. **(b)** The corresponding fitted power spectrum. Position in TCABR: $r = 19.5$ cm.

Table 5.2: True parameters used in the synthetic realizations of I_{sat} to test the the consistency of the CF-CAB-PSD fit for different positions.

	CF fit				CAB fit		PSD fit
r (cm)	γ	$\langle A \rangle$ (mA)	$\langle I_{\mathcal{N}} \rangle$ (mA)	$\sigma_{\mathcal{N}}$ (mA)	λ	τ_d (µs)	$\tau_{\mathcal{N}}$ (µs)
17.0	4	12	10	3	0.5	2	0.3
18.5	0.7	14	12	5	0.3	4	1.3
19.5	2	7	2	1	0.1	5	0.5

5.5 Consistency of the CF-CAB-PSD fit

This section analyses whether the CF-CAB-PSD method is consistent or not. For this purpose, we simulated positions $r = 17.0$ cm, $r = 18.5$ cm and $r = 19.5$ cm of the TCABR tokamak. The first is in the plasma edge ($r < a = 18.0$ cm), having large pulse overlap γ and symmetrical pulses, $\lambda = 0.5$. The second is right after the last closed flux surface (LCFS, approximately at $a = 18.0$ cm) and it had large noise parameters ($\langle I_{\mathcal{N}} \rangle$, $\sigma_{\mathcal{N}}$ and $\tau_{\mathcal{N}}$) and small pulse overlap γ in comparison to other positions. Lastly, $r = 19.5$ cm is in the scrape-off layer of the tokamak (SOL, $r > a$) and is far from the LCFS. It has large pulse

[4] The PSDs were computed with MATLAB's *pwelch* function, which uses Welch's method [82]. Windows of $N = 2^{12}$ points were used, with an overlap of 50% (Appendix A.1).

duration τ_d and very asymmetric pulses (small λ). All parameters used in the simulations are in Table 5.2.

To then verify the consistency of the CF-CAB-PSD fit, various simulated signals of I_{sat} were created, with the same initial conditions for each position (Table 5.2), and the fitting procedure was applied to each one of them. The evaluated parameters were stored in histograms (Fig. 5.4, 5.5 and 5.6). The parameters are well distributed around the true values.

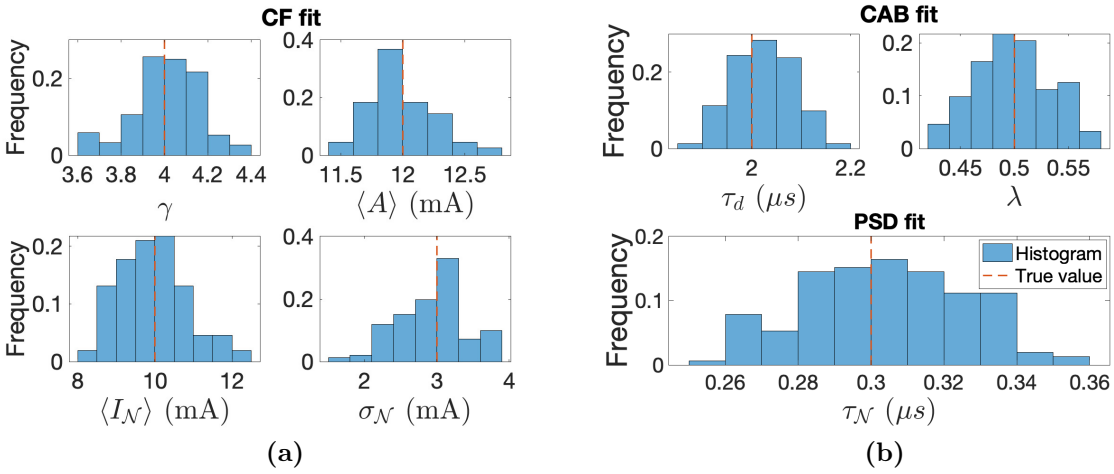


Figure 5.4: Histograms of the CF-CAB-PSD fit, showing the results for 150 realizations of a synthetic I_{sat} signal of $r = 17.0$ cm, with true values given by the first row of Table 5.2 and highlighted in orange in the figures.

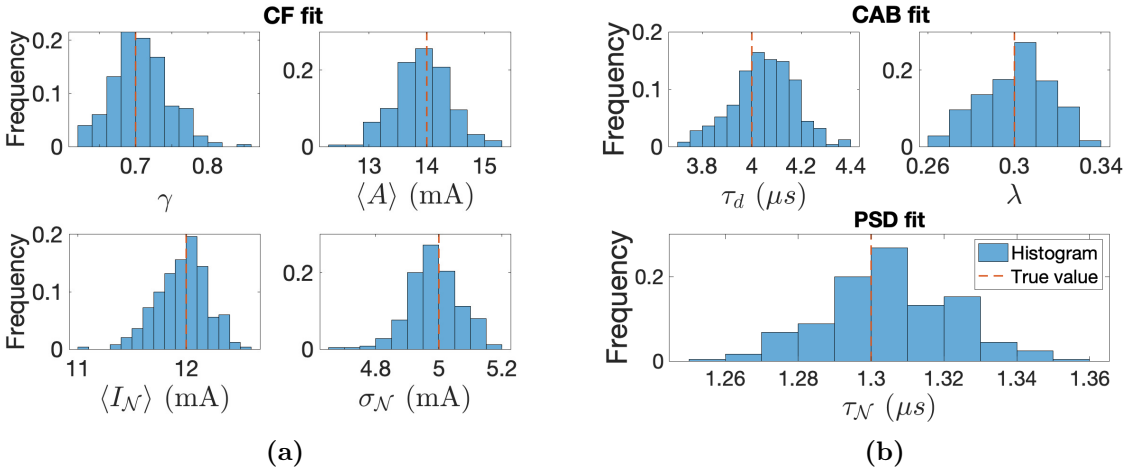
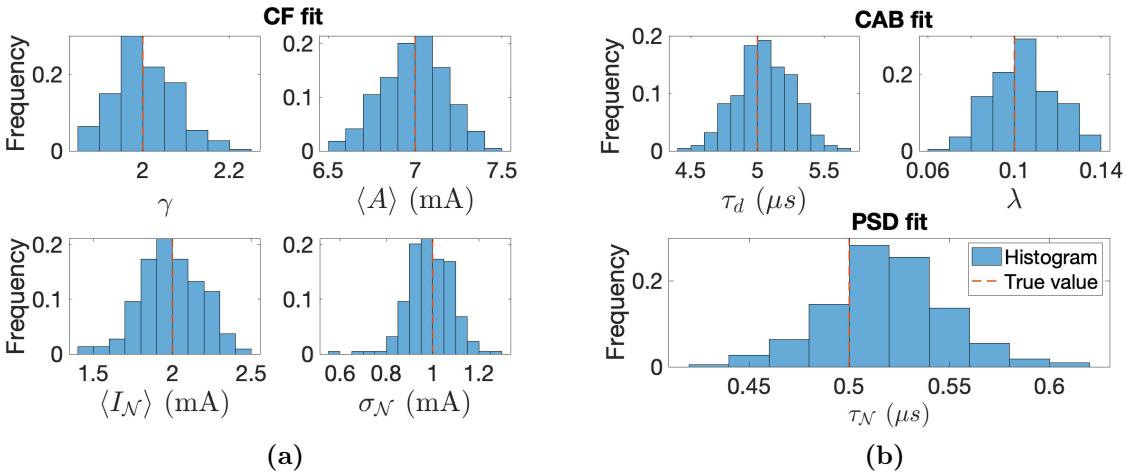
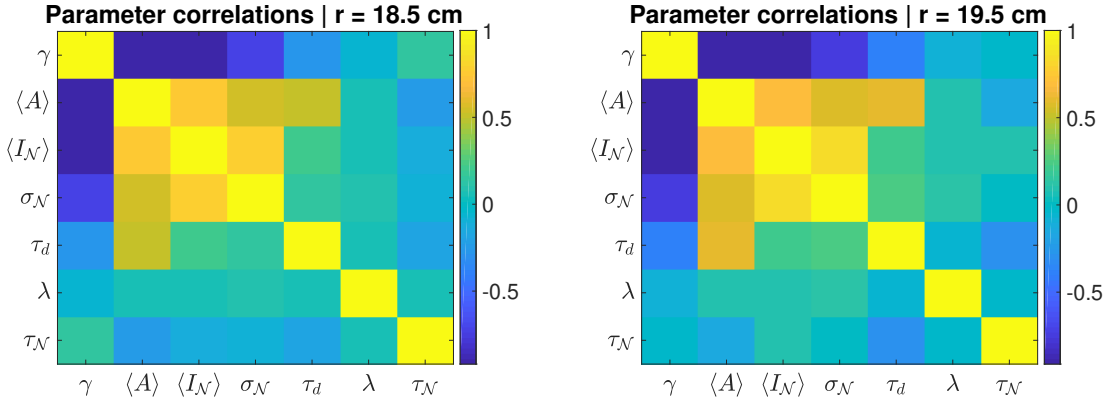


Figure 5.5: Histograms of the CF-CAB-PSD fit, showing the results for 250 realizations of a synthetic I_{sat} signal of $r = 18.5$ cm, with true values given by the middle row of Table 5.2.



(a) (b)

Figure 5.6: Histograms of the CF-CAB-PSD fit, showing the results for 200 realizations of a synthetic I_{sat} signal of $r = 19.5$ cm, with true values given by the last row of Table 5.2.



(a) (b)

Figure 5.7: Correlations of the seven fitted parameters for the simulations of positions $r = 18.5$ cm and $r = 19.5$ cm.

5.6 Correlations between parameters

With the repeated simulations of Section 5.5, it is possible to estimate the correlations (Section 2.3) between the seven fitted parameters of the CF-CAB-PSD method (Table 5.1). Fig. 5.7 shows the correlations for the simulated positions of $r = 18.5$ cm and $r = 19.5$ cm of Section 5.5. In both cases the pattern is almost the same: the intermittency parameter γ is strongly anti-correlated with the other parameters adjusted by the characteristic function—$\langle A \rangle$, $\langle I_\mathcal{N} \rangle$ and $\sigma_\mathcal{N}$. Moreover, these last three have a high positive correlation between them. We underline that this correlation is not necessarily physical, but statistical. This means

the present analysis does not indicate that one parameter can be omitted in favor of another because of physical reasons.

As illustrated in Fig. 5.7, all the other parameters (τ_d, λ and $\tau_{\mathcal{N}}$) do not have important correlations, with exception to $\rho(\langle A \rangle, \tau_d)$, which is close to 0.5. The correlation matrix for the $r = 17.0$ cm is also very similar to the ones shown here. Table 5.3 presents the numerical values for the case $r = 19.5$ cm.

Table 5.3: Correlations of the parameters obtained in Subsection 5.5 for the simulation of $r = 19.5$ cm. In bold, the correlations with absolute values above 0.5, which correspond to the parameters of the CF fit.

	γ	$\langle A \rangle$	$\langle I_{\mathcal{N}} \rangle$	$\sigma_{\mathcal{N}}$	τ_d	λ	$\tau_{\mathcal{N}}$
γ	1	**-0.91**	**-0.91**	**-0.74**	-0.28	-0.12	-0.01
$\langle A \rangle$	**-0.91**	1	**0.72**	**0.55**	0.46	0.18	0.11
$\langle I_{\mathcal{N}} \rangle$	**-0.91**	**0.72**	1	**0.85**	0.17	0.09	0.10
$\sigma_{\mathcal{N}}$	**-0.74**	**0.55**	**0.85**	1	0.18	0.06	0.02
τ_d	-0.28	0.46	0.17	0.18	1	0.15	0.07
λ	-0.12	0.18	0.09	0.06	0.15	1	0.04
$\tau_{\mathcal{N}}$	-0.01	0.11	0.10	0.02	0.07	0.04	1

Chapter 6

Simulations of the SPTM with correlated noise to describe density fluctuations in TCABR

This chapter compares a experiment of TCABR with stochastic simulations with noise, in order to describe local measurements of plasma fluctuations.

6.1 Introduction

As shown in Chapter 5, we developed a routine to fit the stochastic pulse train model with correlated noise. Now, this CF-CAB-PSD fit will be applied to the discharge 34132 of the TCABR tokamak. This experiment was already introduced in Sections 3.2–3.4. Eight radial positions were measured with a rake Langmuir probe, at the outboard mid-plane of the tokamak. The first tip was mounted at $r = 17.0$ cm and the eighth, at $r = 20.5$ cm.

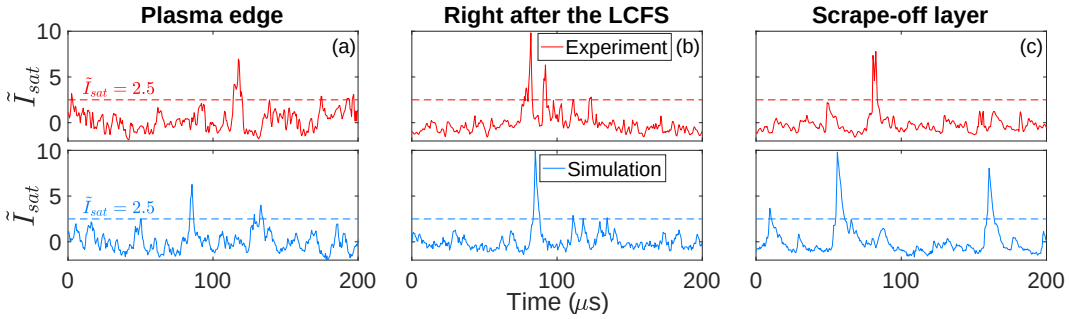


Figure 6.1: The top panels show excerpts of saturation current for three positions in the TCABR tokamak: (a) in the plasma edge, $r = 17.0$ cm; (b) right after the last closed flux surface (LCFS), $r = 18.5$ cm; (c): at the scrape-off layer, $r = 19.5$ cm. The bottom panels show simulations of the corresponding experimental data. In each panel, at least one large-amplitude burst is seen, with a peak above $\tilde{I}_{sat} = 2.5$.

Following the usual convention [30–36, 40, 48], the results of this chapter are displayed using the normalized saturation current,

$$\tilde{I}_{sat} = \frac{I_{sat} - \langle I_{sat} \rangle}{\sigma_I}$$

such that $\tilde{I}_{sat}$ has zero mean and unit standard deviation. The top panels of Fig. 6.1 depict some excerpts of $\tilde{I}_{sat}$ for three positions measured in TCABR. The bottom panels show the corresponding simulations made with the parameters obtained with the CF-CAB-PSD fit presented in Chapter 5.

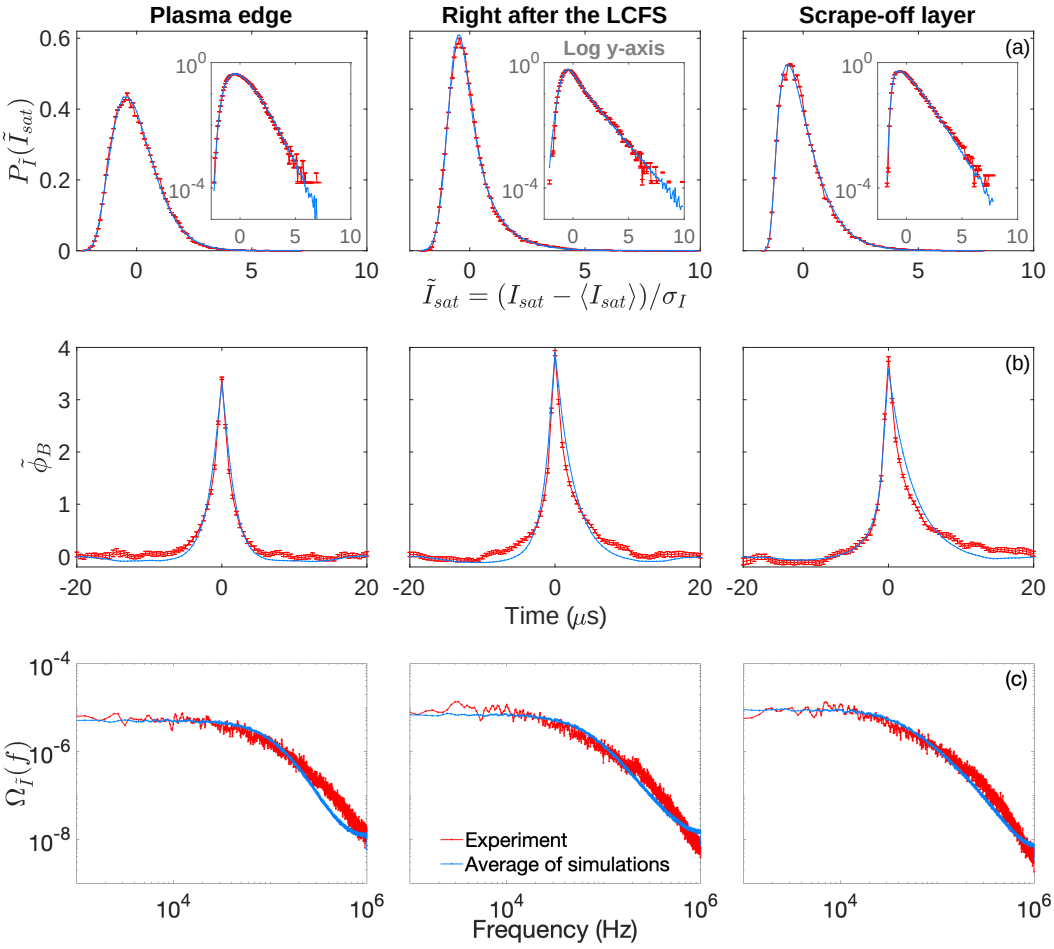


Figure 6.2: In red, (a) PDFs (in linear and log-y scales), (b) conditionally averaged bursts and (c) power spectra of $\tilde{I}_{sat}$ for three positions—plasma edge ($r = 17.0$ cm), right after the last closed flux surface ($r = 18.5$ cm) and at the scrape-off layer ($r = 19.5$ cm). In light blue, graphs of simulations resultant from the CF-CAB-PSD fit (Chapter 5).

6.2 Transition from the plasma edge to the SOL

In this section are shown the results of the CF-CAB-PSD fit for three characteristic positions of TCABR—at the plasma edge ($r = 17.0$ cm), readily after the last closed flux surface ($r = 18.5$ cm) and at the scrape-off layer ($r = 19.5$ cm). With this, it is possible to analyze

whether the stochastic pulse train model can describe the transition between the plasma edge, $0.8 < r/a < 1$, to the scrape-off layer (SOL), $r/a > 1$ (where $a = 18.0$ cm is TCABR's minor radius). For the results of all eight positions, the reader is referred to Appendix B.3.1.

Each column of Fig. 6.2 corresponds to a position in TCABR, and each row depicts graphs with the same axes. The simulated graphs in blue were averaged with 30 Monte Carlo simulations, to diminish statistical fluctuations. Excerpts of $\tilde{I}_{sat}$ for these same positions can be seen in Fig. 6.1.

An increase of intermittency is observed with the probability distribution functions (PDFs), Fig. 6.2(a). At the plasma edge panel, the maximum value is about $\tilde{I}_{sat} = 7$, whereas after the last closed flux surface (LCFS), $\tilde{I}_{sat} = 10$. At the SOL, the maximum is $\tilde{I}_{sat} = 8$. Furthermore, it is noticeable that inside the plasma edge the PDF is closer to a Gaussian, indicating that the pulse overlap is high and the intermittency, low. Outside the confinement region, the PDF becomes more positively asymmetric, and its kurtosis and intermittency increase. The simulated distributions, obtained with the fitting procedure of Chapter 5, are in excellent agreement with the experimental ones.

Fig. 6.2(b) shows the conditionally averaged burst shapes. On average the bursts are symmetric in the plasma edge, where the fitting method obtained $\lambda = 0.509(81)$ for the pulse asymmetry parameter (Eq. (4.2)). As the radial position goes from the plasma edge to the SOL, the conditionally averaged waveforms become more asymmetric, with $\lambda = 0.293(30)$ at $r = 18.5$ cm and with $\lambda = 0.105(13)$ at $r = 19.5$ cm (uncertainties in parenthesis). The pulse characteristic duration also increases with r, going from $\tau_d = 2.240(67)$ µs to $\tau_d = 3.81(12)$ µs and $\tau_d = 4.57(32)$ µs, at $r = 17.0$ cm, $r = 18.5$ cm and $r = 19.5$ cm, respectively.

Finally, Fig. 6.2(c) shows the comparison between experimental and simulated power spectral densities (PSD). The fit is not perfect, but the model reproduces the general behavior of the frequency spectrum—that is, a plane region for low frequencies followed by a power-law spectrum for high frequencies.

6.3 Radial profiles

Using the procedure described in Chapter 5, it is possible to fit the eight positions measured in TCABR. A radial profile of the obtained parameters is depicted in Fig. 6.3. (For the exact values, see Appendix B.4.) The parameters γ, $\langle A \rangle$, $\langle I_{\mathcal{N}} \rangle$ and $\sigma_{\mathcal{N}}$, which define the probability distribution of I_{sat}, generally decay with r. This is expected since the density decreases with the radial profile ($\langle I_{sat} \rangle$ in Fig. 6.4(a)).

For these four parameters shown in Fig. 6.3(a), it is possible to see a regime change in $r = a = 18.0$ cm, where a indicates the beginning of the limiter and thus the end of the plasma confinement region. Before $r = a$ the intermittency parameter $\gamma = \tau_d N_P / T$ achieves its biggest values, demonstrating that the pulse overlap is higher inside the plasma column than in the scrape-off layer. On the other hand, the average amplitude of the pulses $\langle A \rangle$ starts to decrease rapidly at $r = a = 18.0$ cm. The same happens to the mean and standard deviation of the noise, $\langle I_{\mathcal{N}} \rangle$ and $\sigma_{\mathcal{N}}$.

As for the graphs in Fig. 6.3(b), one sees that the pulse characteristic duration τ_d increases

with r, going from $\tau_d = 2.2$ µs at $r = 17.0$ cm to $\tau_d = 5.0$ µs at $r = 20.5$ cm. The pulse asymmetry parameter λ, on the other hand, is compatible with $\lambda = 0.5$ in the plasma edge (indicating symmetric pulses) and then decays to $\lambda = 0.1$ in the scrape-off layer (indicating asymmetric pulses). The behavior of τ_d and λ was predicted in the discussion of Fig. 6.2(b).

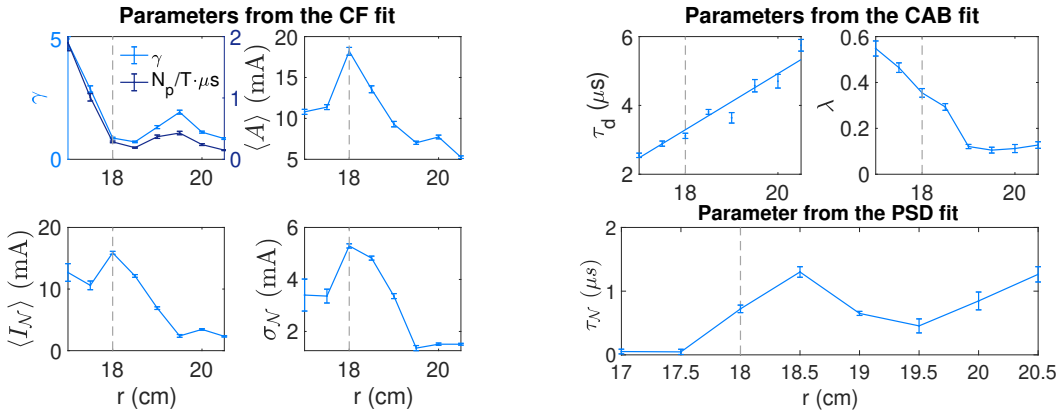


Figure 6.3: **(a)** Radial profile of the parameters fitted with the characteristic function for the eight radial positions measured in TCABR. Also included is the number of pulses per time, evaluated as $N_P/T = \gamma/\tau_d$. **(b)** Radial profile from the parameters fitted with χ^2 maps from the CAB and PSD. The dashed line indicates the plasma minor radius, $a = 18.0$ cm.

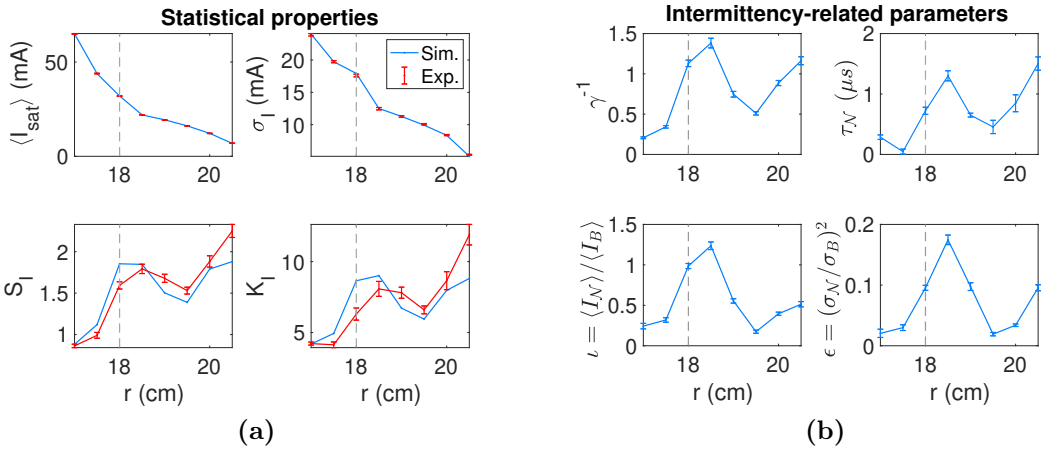


Figure 6.4: **(a)** Four statistical properties of I_{sat}: mean, standard deviation, skewness and kurtosis. The dashed line in $r = a = 18.0$ cm indicates the limiter position. **(b)** The radial profile of the parameters with a similar trend as the kurtosis.

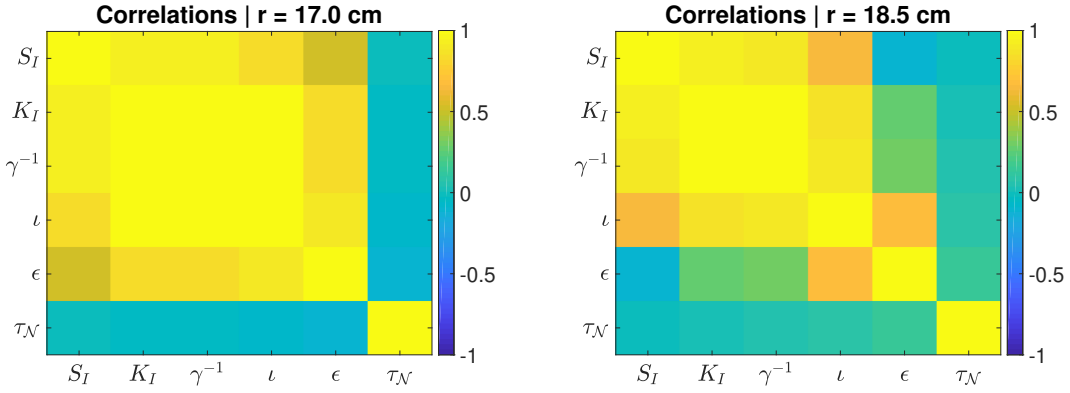
It lasts to analyze noise correlation time $\tau_{\mathcal{N}}$ in Fig. 6.3(b). First of all, it is clear that the $\tau_{\mathcal{N}}$ profile is different from the pulse characteristic duration τ_d, unlike what was used in Ref. 54. In the second place, the noise correlation time seems to have a peculiar profile in

Fig. 6.3(b)—it increases until $r = 18.5$ cm, then decreases up to $r = 19.5$ cm and thereafter increases again. Surprisingly, this behavior can be accounted for by analyzing the radial profile of the kurtosis.

One sees in Fig. 6.4 that, for this experiment, S_I, K_I, γ^{-1}, $\iota = \langle I_{\mathcal{N}} \rangle / \langle I_B \rangle$, $\epsilon = \sigma^2_{\mathcal{N}}/\sigma^2_B$ and $\tau_{\mathcal{N}}$ have a similar profile—namely, a local maximum and minimum respectively at $r = 18.5$ cm and $r = 19.5$ cm. As mentioned in Sections 2.1 and 3.4, there is a well-known parabolic relation between excess kurtosis and skewness for density fluctuations in magnetized plasmas: $K_I - 3 \approx 1.5 S_I^2$ [32,33,36,50,69,71].[1] From Eq. (4.12), it is seen the connection between K_I, γ^{-1} and ϵ,

$$K_I - 3 = \frac{6\gamma^{-1}}{(1+\epsilon)^2}$$

For small noise levels ϵ, then the excess kurtosis is approximately equal to the intermittency level γ^{-1} times a constant, $K_I - 3 \approx 6\gamma^{-1}$. Furthermore, when the pulse overlap $\gamma = \tau_d/\tau_w$ decreases (such as for $r = 18.5$ cm in Fig. 6.3(a)), an increase is expected for the relative mean and variance of the noise, $\iota = \langle I_{\mathcal{N}} \rangle / (\gamma \langle A \rangle)$ and $\epsilon = \sigma^2_{\mathcal{N}} / \left(\gamma \langle A \rangle^2\right)$. Finally, when $\epsilon = \sigma^2_{\mathcal{N}}/\sigma^2_B$ increases, then the noise correlation $\tau_{\mathcal{N}}$ must also do so, to mimic the signal's true complexity. This explains the similar trend between the parameters S_I, K_I, γ^{-1}, ι, ϵ and $\tau_{\mathcal{N}}$ in Fig. 6.4.



(a) **(b)**

Figure 6.5: Correlations between the intermittency-related parameters of Fig. 6.4, for two characteristic cases: **(a)** $r = 17.0$ cm, with low noise level $\epsilon < 5\%$. **(b)** $r = 18.5$ cm, with high noise level, $9\% < \epsilon$.

6.3.1 Correlations of parameters

Fig. 6.5 depicts the correlations between the intermittency-related parameters, S_I, K_I, γ^{-1}, ι, ϵ and $\tau_{\mathcal{N}}$. As examples, the $r = 17.0$ cm and $r = 18.5$ cm cases are shown respectively in

[1] Using Eq. (4.12), one sees that, for the stochastic pulse train model with correlated noise, the exact expression is $K_I - 3 = 1.5(1+\epsilon)S_I^2$, which corresponds to $K_I - 3 \approx 1.5 S_I^2$ for small ϵ.

the (a) and (b) panels. For both scenarios, the correlations between S_I, K_I, γ^{-1} and ι are above 0.5. In contrast, the correlations between $\tau_{\mathcal{N}}$ and the other parameters are near to 0.0. The main difference between the two positions occurs for the noise level ϵ. For $r = 17.0$ cm, the correlations of ϵ are above 0.5, whereas for $r = 18.5$ cm the correlations between ϵ and S_I, K_I, γ^{-1} are close to 0.0.

Positions with low noise level ϵ (i.e., $r \in [17.0, 17.5, 19.5, 20.0]$ cm, following Fig. 6.4(b)) have correlations similar to the $r = 17.0$ cm case. On the other hand, positions with high noise level ϵ (i.e., $r \in [18.0, 18.5, 19.0, 20.5]$ cm) are similar to the $r = 18.5$ cm case.

6.4 Comparison between noises

6.4.1 Determinism

In this section, three different choices for the correlation time $\tau_{\mathcal{N}}$ will be compared. Some authors use $\tau_{\mathcal{N}} \to 0$ µs (corresponding to uncorrelated or white noise) [34, 40, 41], while recently the dynamical noise $\tau_{\mathcal{N}} = \tau_d$ was proposed [37, 54]. In this thesis, we considered the case where $\tau_{\mathcal{N}}$ is fitted with the PSD (Section 5.4), and thus the correlation time is not necessarily equal to the pulse duration, $\tau_{\mathcal{N}} \neq \tau_d$. Apart from $\tau_{\mathcal{N}}$, all the other parameters are equal in the three scenarios (see Fig. 6.3 or Tables B.1 and B.2 for the values).

To make this comparison, we propose the use of the determinism (DET) from the recurrence quantification analysis (RQA)—see for example Refs. 98 and 99 or Section 2.8. DET quantifies the percentage of diagonal lines that are recurrent in the recurrence plot [99].[2] Chaotic signals tend to have higher RQA determinism than stochastic ones.

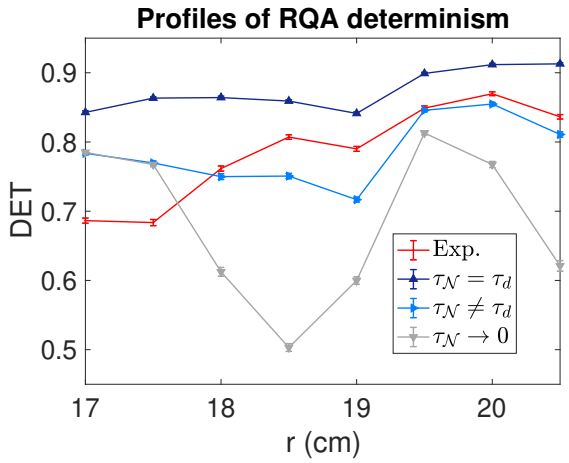


Figure 6.6: Radial profiles of determinism obtained with recurrence quantification analysis. In red, the data from the TCABR experiment; in dark blue, light blue and gray are the data from the SPTM simulations. All have the same parameters from Fig. 6.3, with the exception of the correlation time $\tau_{\mathcal{N}}$. The dashed line marks the LCFS approximate position.

[2] For recurrence plots of the experimental and simulated I_{sat} of TCABR, see Appendix B.5.

Fig. 6.6 shows the radial profile of the determinism of I_{sat} for the experimental and simulated data. An embedding dimension of $d_{\mathrm{RQA}} = 4$ was used, as well as a delay of $\tau_{\mathrm{RQA}} = 10$ points (5 μs) and repeated windows with $N_{\mathrm{RQA}} = 10^3$ points to evaluate the average and standard deviation of the determinism. The recurrence rate was fixated at 10%. These parameter values were chosen following previous works [99–102]. Moderate changes in them don't alter the results.

The red curve in Fig. 6.6 represents the experimental RQA determinism. It slightly increases with r. In gray is shown the data obtained with Gaussian white noise ($\tau_{\mathcal{N}} \to 0$ µs). There is an obvious structure with a minimum in $r = 18.5$ cm and maximum at $r = 19.5$ cm. Therefore, for the signal with white noise, we infer that the determinism and the noise level ϵ follow opposite trends, since the profile of this last one has a maximum at $r = 18.5$ cm and minimum at $r = 19.5$ cm (Fig. 6.4(b)).

In its turn, the dark blue curve in Fig. 6.6 shows the determinism obtained with the correlation time equal to the characteristic duration of the pulses. This condition $\tau_{\mathcal{N}} = \tau_d$ was proposed in Ref. 54 to model dynamic noise. However, the dark blue curve in Fig. 6.6 is systematically above the experimental data.

Lastly, the light blue curve in Fig. 6.6 shows the RQA determinism obtained with the simulated signal with correlation time different from the pulse duration ($\tau_{\mathcal{N}} \neq \tau_d$). Although not perfect, the simulated data is closer to the experimental points than in the other two scenarios ($\tau_{\mathcal{N}} \to 0$ µs and $\tau_{\mathcal{N}} = \tau_d$).

6.4.2 CH diagram

In addition to determinism, we propose to use the complexity-entropy diagram to differentiate which noise best describes density fluctuations in tokamaks. The CH plane was presented in Section 2.7. It characterizes a signal I in terms of a pair of complexity and entropy (C_{JS}, H_S), which describes the local structure of signals in terms of tuples of dimension d_{CH}. Following previous works [80, 93–95], we adopted $d_{CH} = 6$, meaning that each embedding vector had a duration of $d_{CH}\Delta t = 3$ µs.

Fig. 6.7 presents CH diagrams for the eight measured positions in TCABR and the corresponding simulations with dynamical OU noise ($\tau_{\mathcal{N}} = \tau_d$), fitted OU noise ($\tau_{\mathcal{N}} \neq \tau_d$) and uncorrelated noise ($\tau_{\mathcal{N}} \to 0$). The experiment, illustrated in all four panels of Fig. 6.7, has its points above the fractional Brownian motion (fBm) curve. The points from the three different simulations are in general below the experimental data and close to the fBm curve. The model with dynamical noise is the closest to the experiment in the CH plane, followed up by the fitted noise case. The uncorrelated noise model has the majority of its points close to $(C_{JS}, H_J) = (0, 1)$, which corresponds to a purely white noise signal. In all three scenarios, the simulations with the lowest noise levels ϵ (i.e. $r \in [17, 17.5, 19.5, 20]$ cm, see Fig. 6.4) have a higher complexity. In contrast, signals with high noise levels (such as the one of $r = 18.5$ cm) tend to have the lowest complexity and highest entropy.

It can be elucidating to investigate the entropy and complexity separately, as done in Fig. 6.8. For the described signals, H_S and C_{JS} are highly anti-correlated. This means that it suffices to examine one, since the results of the other will be the opposite—a minimum

in H_S will be a maximum in C_{JS}, and vice-versa. The white noise model has the worse results, whereas in this case the dynamical noise models better the experimental complexity and entropy.

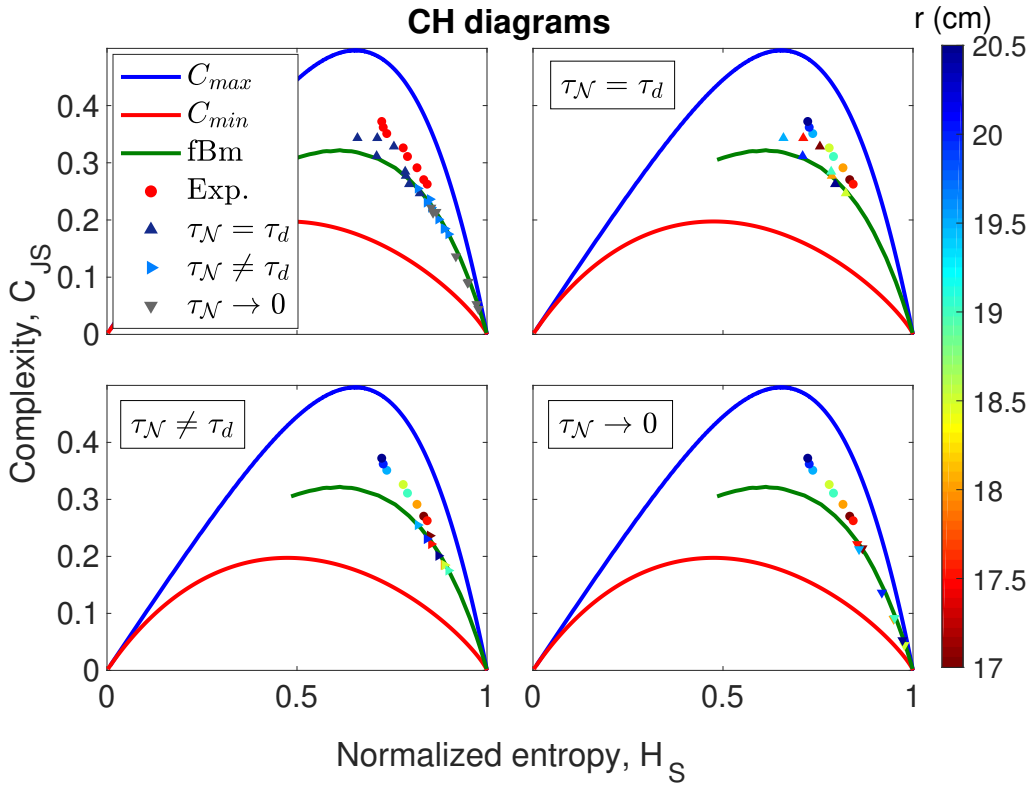


Figure 6.7: Complexity-entropy diagrams for the eight experiments measured in TCABR (in circles), together with the simulations from three different Ornstein-Uhlenbeck noises ($\tau_\mathcal{N} = \tau_d$, $\tau_\mathcal{N} \neq \tau_d$ and $\tau_\mathcal{N} \to 0$ µs, in triangles). In the top-left panel, the points of all signals are shown. For the other panels, the experimental data is compared to one type of simulation at a time, and each position r is represented by a color in the rainbow scale.

The simulated entropies depicted in Fig. 6.8(a) have a local maximum and minimum at $r = 18.5$ cm and $r = 19.5$ cm, respectively, as seen before for the intermittency-related parameters (Fig. 6.4).[3] The experimental entropy, on the other hand, mainly decays with the radial position. This evident discrepancy between experiment and simulation indicates that the stochastic pulse train model with Gaussian Ornstein-Uhlenbeck noise does not describe well the local structure of density fluctuations for the TCABR tokamak.

[3]For the case of fitted noise ($\tau_\mathcal{N} \neq \tau_d$), the maximum at $r = 18.5$ cm is flattened, as seen in Fig. 6.8(a).

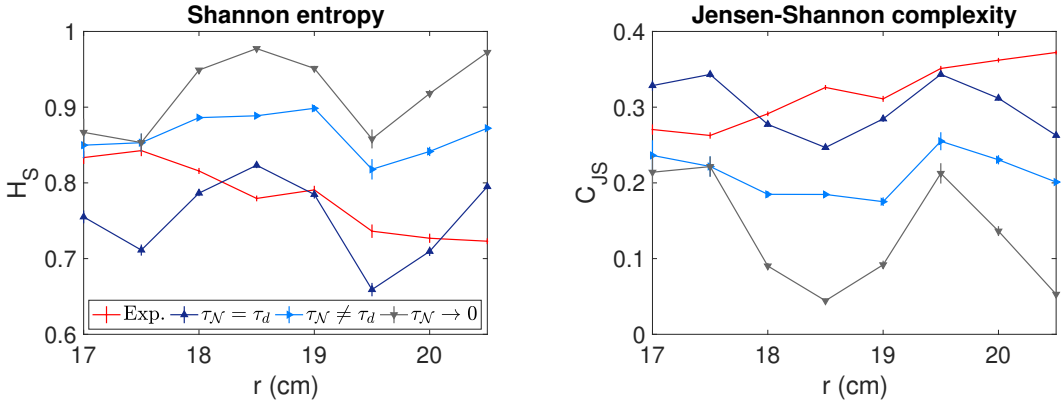


Figure 6.8: Shannon entropy and Jensen-Shannon complexity for the experimental positions (in red) and the corresponding simulations with different correlation times $\tau_{\mathcal{N}}$.

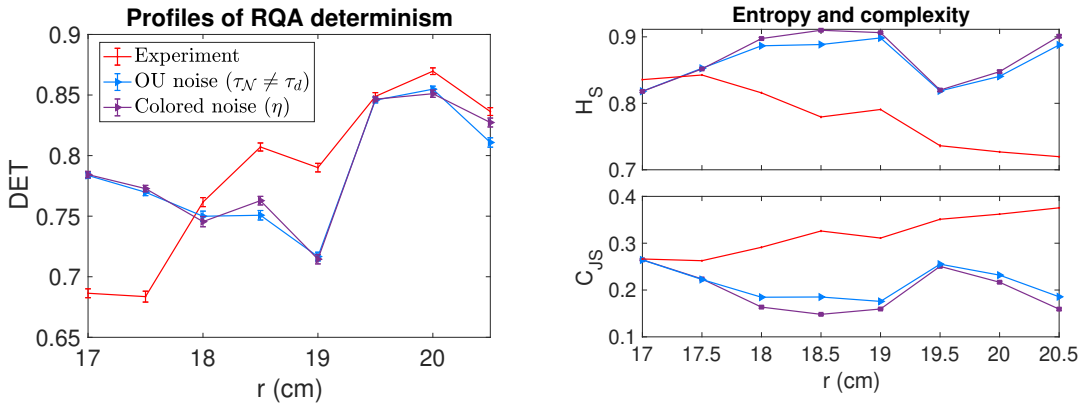
6.4.3 Colored noise

In this subsection, the results from colored and Ornstein-Uhlenbeck noises are compared. The model with colored noise (CN)—characterized by the relation $\Omega_I(f) \propto f^{-\eta}$ in the frequency spectrum—was used in Ref. 48, whereas the OU noise case was introduced in 54 and applied in 37. The parameters from both models are the same: γ, $\langle A \rangle$, $\langle I_{\mathcal{N}} \rangle$, $\sigma_{\mathcal{N}}$, λ and τ_d (as in Table 5.1). The only exception is the correlation parameter, which is η for the CN case and $\tau_{\mathcal{N}}$ for the OU case. See Section 2.6 for more details.

Because of the similarities, both can be fitted with the CF-CAB-PSD procedure (Chapter 5). The only difference is that, for the colored noise scenario, the exponent η is fitted instead of the correlation time $\tau_{\mathcal{N}}$. With this in mind, we created simulations for the CN model, fitting the experimental I_{sat} with the CF-CAB-PSD fit, just as done for the OUN in Chapter 5. Hence, the parameters γ, $\langle A \rangle$, $\langle I_{\mathcal{N}} \rangle$, $\sigma_{\mathcal{N}}$, λ and τ_d were already evaluated. They can be seen in Appendix B.4, together with the values of the noise exponent η.

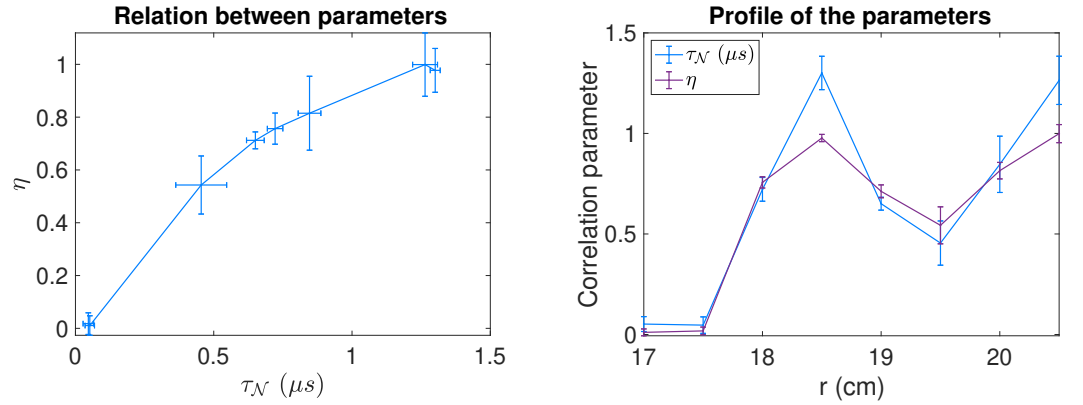
Results comparing CN and OUN simulations are presented in Fig. 6.9. One notices that both have very similar determinism, Shannon entropy and Jensen-Shannon complexity. In fact, a monotonic relation was found between the two, as indicated by Fig. 6.10(a). The radial profile of the parameters η and $\tau_{\mathcal{N}}$ is shown in Fig. 6.10(b).

The results from the CN and OUN simulations are almost identical. Still, as discussed in Section 2.6, the power spectral density (PSD) of the colored noise is ill-defined, especially for frequencies approaching zero, $f \to 0$. On the contrary, the PSD of the OUN is well defined and is given in terms of analytical functions. Thereafter, it is preferable to use the Ornstein-Uhlenbeck noise.



(a) **(b)**

Figure 6.9: **(a)** Profile of the determinism for the experiment (in red) and simulations with Ornstein-Uhlenbeck noise (in blue) and colored noise (in purple). **(b)** Profiles of the Shannon entropy and Jensen-Shannon complexity for the same signals.



(a) **(b)**

Figure 6.10: **(a)** Graph of $\eta \times \tau_{\mathcal{N}}$. **(b)** Profiles of these same correlation parameters $\tau_{\mathcal{N}}$ and η, obtained with the CF-CAB-PSD fit for the models with Ornstein-Uhlenbeck noise and colored noise, respectively.

6.5 Summary and discussion

A lot of information was given in this chapter and hence it is worthwhile to summarize and organize the main results. We applied the CF-CAB-PSD fit to the experiment 34132 of TCABR, in which I_{sat} was measured at eight radial positions. With the fit, simulations with bursts and correlated noise were generated and compared to experimental data.

In Section 6.2, the distributions of the model were shown to be in excellent agreement with the experiment. The conditionally averaged burst (CAB) was also adequately adjusted,

at least as a first approximation. The power spectral density (PSD), on the other hand, was poorly fitted. This indicates that the adopted pulse train model with Gaussian Ornstein-Uhlenbeck noise does not reproduce so well the structure of density fluctuations in TCABR.

In Section 6.3, radial profiles of the parameters were presented. The distribution parameters mainly decay with the radial position, which makes sense since the plasma density also decreases with r. Also, we found a relation between the intermittency-related parameters, such as γ, ϵ and $\tau_{\mathcal{N}}$. A decrease in the pulse overlap γ tends to be accompanied by an increase in the noise level ϵ, so that the absence of pulses is replaced by noise with higher fluctuation levels. However, to mimic the pulse shapes, the noise correlation $\tau_{\mathcal{N}}$ must also grow.

In Section 6.4, four different noise cases used in the literature were compared: three with Ornstein-Uhlenbeck noise (OUN), defined by the correlation time $\tau_{\mathcal{N}}$, and one with colored noise (CN), which has the noise exponent η. In Subsection 6.4.3, the CN model proved to be very similar to the OUN scenario with $\tau_{\mathcal{N}}$ fitted by the PSD ($\tau_{\mathcal{N}} \neq \tau_d$ case). However, as the power spectrum of the CN is ill-defined, it is preferable to use the OUN.

In Subsections 6.4.1 and 6.4.2 we introduced the use of the determinism and the CH diagram to distinguish which OUN better models density fluctuations. Between white ($\tau_{\mathcal{N}} \to 0$ µs), dynamical ($\tau_{\mathcal{N}} = \tau_d$) and fitted noise ($\tau_{\mathcal{N}} \neq \tau_d$), the last performed better, from the point of view of the determinism. On the other hand, from the perspective of the CH diagram, the dynamical noise had better results. In both scenarios, the white noise model had the worst outcomes. Thus, from the four noise models, only the dynamical and fitted OU cases are indicated to describe I_{sat} measurements in TCABR.

Nevertheless, the analysis in the CH diagram showed that all of the considered models with noise couldn't accurately reproduce the experimental entropy and complexity. Even the dynamical noise case, which had the better results for this graph, showed a biased profile (Fig. 6.8).

Certainly, the problem is not with the fitting procedure, which is unbiased, as seen in Section 5.5. Hence, our analysis with the CH diagram, the RQA determinism and the power spectrum indicate that the stochastic pulse train model with noise may not describe faithfully the local structure of density fluctuations in the TCABR tokamak. These results motivated us to analyze another type of background fluctuations. This new background will be subject of Chapters 7 and 8.

Chapter 7

Stochastic pulse train model with pulse background

In Chapters 4–6, we considered the stochastic pulse train model with correlated noise (SPTM-N). As analyzed in Sections 6.4.2 and 6.5, the SPTM-N could not faithfully reproduce the structure of the background of density fluctuations in TCABR. In this chapter, another type of SPTM will be introduced, changing correlated noise to a pulse background. The SPTM with a pulse background was first considered in Ref. 48, to analyze correlated bursts. The burst shape will also be generalized here, to better fit the experimental averaged burst.

7.1 The model

The stochastic pulse train model with pulse background (SPTM-PB) is given by

$$I_{sat}(t) = I_B(t) + I_b(t) \tag{7.1}$$

where both terms correspond to a pulse train (as in Eq. (4.6)),

$$I_B(t) = \sum_{j=1}^{N_P} A_j \phi\left(\frac{t - t_j}{\tau_d}; \lambda, \zeta_r, \zeta_f\right) \tag{7.2}$$

and

$$I_b(t) = \sum_{k=1}^{N_P^{(b)}} A_k^{(b)} \phi\left(\frac{t - t_k^{(b)}}{\tau_d^{(b)}}; 0.5, 1, 1\right) \tag{7.3}$$

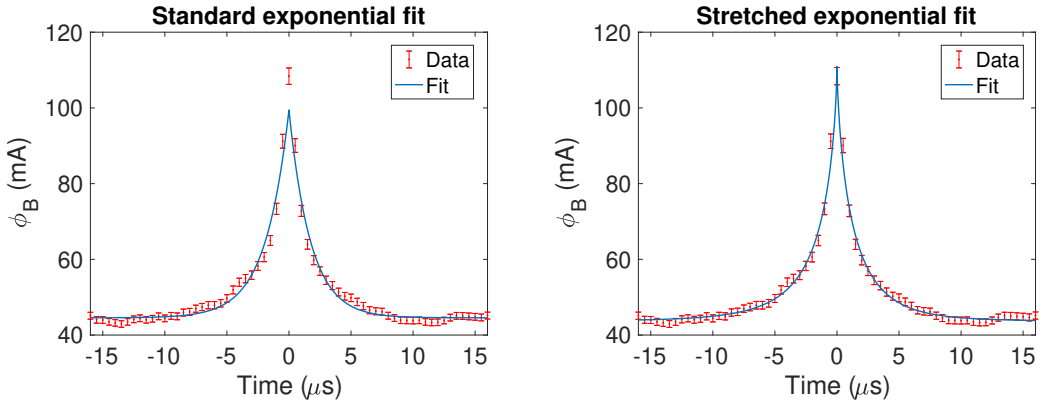
The supper script $^{(b)}$ will be used to denote the background. The distributions of amplitude and interval between events remain exponential and uncorrelated, as in the SPTM-N case (Eq. (4.3) and (4.4)).

Also, here we are also generalizing the pulse shape to

$$\phi(\theta; \lambda, \zeta_r, \zeta_f) = \begin{cases} \exp\left(-\left|\theta/\lambda\right|^{\zeta_r}\right), & \theta < 0 \\ \exp\left\{-\left[\theta/\left(1-\lambda\right)\right]^{\zeta_f}\right\}, & \theta \geq 0 \end{cases} \tag{7.4}$$

The exponents ζ_r (for the pulse rise) and ζ_f (for the fall) transform the standard exponentials in stretched exponentials. They were first used to fit conditionally averaged bursts in Ref.

15. As it will be shown in Chapter 8, in TCABR these stretching exponents ζ range from 0.5 to 1.0.[1] The case $\zeta_r = \zeta_f = 1$ corresponds to standard exponentials.



(a) **(b)**

Figure 7.1: Conditionally averaged burst of the experiment 34132, position $r = 17.5$ cm, in TCABR. **(a)** Least-squares (LS) fit with standard exponential. **(b)** LS fit with generalized stretched exponentials, yielding $\tau_r = 1.51(4)$ µs, $\tau_f = 1.40(4)$ µs, $\zeta_r = 0.812(22)$ and $\zeta_f = 0.837(24)$.

Table 7.1: Parameters of the stochastic pulse train model with pulse background and their possible values found for TCABR.

Symbol	Meaning	Eq.	Possible value
N_P/T	Number of burst-related pulses, per time	(7.2)	$[0.05, 1]$ µs^{-1}
$\langle A \rangle$	Average amplitude of the main pulses	(4.3)	$[9, 20]$ mA
τ_d	Characteristic duration of the main pulses	(7.2)	$[2, 5]$ µs
λ	Asymmetry parameter of the main pulses	(7.4)	$[0.1, 0.5]$
ζ_r	Stretching exponent of the rise of the main pulses	(7.4)	$[0.5, 1.0]$
ζ_f	Stretching exponent of the fall of the main pulses	(7.4)	$[0.5, 1.0]$
$N_P^{(b)}/T$	Number of background pulses, per time	(7.3)	$[1, 5]$ µs^{-1}
$\langle A^{(b)} \rangle$	Average amplitude of the background pulses	(4.3)	$[2, 7]$ mA
$\tau_d^{(b)}$	Characteristic duration of the background pulses	(7.3)	$[0.8, 1.8]$ µs

The motivation for using stretched exponentials comes from the fact that the conditionally averaged burst was not perfectly fitted by the model with unstretched exponentials (as seen

[1] To optimize the creation of simulations, it is interesting to truncate the exponentials after their amplitude becomes sufficiently small. Details are given in Appendix B.7.

for example in Fig. 6.2(b)). Fig. 7.1 exemplifies the difference between the two fittings. Panel (a) shows a fit with the standard exponential ($\zeta_r = \zeta_f = 1$). Panel (b) exhibits the stretched fit ($\zeta_r, \zeta_f \in]0, 1]$). The model used in (b) is clearly better for this data since the one in (a) is unable to adjust the peak.

For the pulse background we shall consider symmetric unstretched pulses ($\lambda^{(b)} = 0.5$ and $\zeta_r^{(b)} = \zeta_f^{(b)} = 1$). With this consideration, the SPTM with pulse background has nine parameters, as depicted in Table 7.1.

7.2 Characteristic function of the SPTM-PB

The pulse background $I_b(t)$ from Eq. (7.3) is a train of unstretched bi-exponential pulses which have exponentially distributed amplitudes and waiting time between events. As in Section 4.2, a signal such as this is Gamma distributed [50, 52, 55],

$$P_b(I_b(t)) = \frac{1}{\langle A^{(b)} \rangle \Gamma(\gamma^{(b)})} \left(\frac{I_b(t)}{\langle A^{(b)} \rangle} \right)^{\gamma^{(b)}-1} \exp\left(-\frac{I_b(t)}{\langle A^{(b)} \rangle} \right) \tag{7.5}$$

where $\left\langle A^{(b)} \right\rangle$ is the mean amplitude of the pulses of Eq. (7.3), $\gamma^{(b)} = \tau_d^{(p)} N_P^{(p)}/T$ is the intermittency parameter of the pulse background (PB) and Γ is the Gamma function (Eq. (4.9)). A Gamma distribution has the characteristic function [55]

$$C_b(u) = \left(1 - iu \left\langle A^{(b)} \right\rangle\right)^{-\gamma^{(b)}} \tag{7.6}$$

However, with stretched pulses, the PDF of $I_B(t)$ is not Gamma distributed anymore. In fact, in this $\zeta \neq 1$ case, neither the PDF nor the characteristic function (CF) can be given in terms of elementary functions. Nevertheless, we found out that the derivation of the CF, now expressed in terms of a nonelementary integral, continues to be straightforward.

Following Eq. (A13) of Ref. 55, the characteristic function logarithm for the SPTM with independent random variables and uniform pulses occurrence is

$$\ln C_B(u) = \gamma \int_{-\infty}^{\infty} d\theta \left[C_A \left(u\phi(\theta) \right) - 1 \right]$$

where $\gamma = \tau_d N_P/T$ is the same intermittency parameter from the SPTM-N model and, for stretched pulses (Eq. (7.4)),

$$\phi(\theta) = \begin{cases} \exp\left(-\left| \theta/\lambda \right|^{\zeta_r} \right), & \theta < 0 \\ \exp\left\{ -\left[\theta/\left(1-\lambda\right) \right]^{\zeta_f} \right\}, & \theta \geq 0 \end{cases}$$

C_A is the CF of the amplitude distribution. Since A is exponentially distributed,

$$C_A(u) = \frac{1}{1 - iu \langle A \rangle}$$

and thus

$$C_A(u\phi(\theta)) = \frac{1}{1 - iu\,\langle A\rangle\,\phi(\theta)}$$

Returning to the integral,

$$\ln C_B(u) = \gamma \int_{-\infty}^{\infty} d\theta \left[\frac{1}{1 - iu\,\langle A\rangle\,\phi(\theta)} - 1\right]$$

As detailed in Appendix B.8, results that this expression is equal to

$$\boxed{\ln C_B(u) = \gamma\lambda \cdot L\left(iu\,\langle A\rangle\,;\zeta_r\right) + \gamma(1-\lambda) \cdot L\left(iu\,\langle A\rangle\,;\zeta_f\right)} \tag{7.7}$$

where we defined the nonelementary integral

$$L(z;\zeta) \equiv \int_0^{\infty} dx \frac{z}{\exp\left(x^{\zeta}\right) - z}$$

such that

$$L(iu\,\langle A\rangle\,;\zeta) = \int_0^{\infty} dx \frac{iu\,\langle A\rangle}{\exp\left(x^{\zeta}\right) - iu\,\langle A\rangle}$$

$L(z;\zeta)$ converges for $\mathrm{Re}(1/\zeta) > 0$ and for all $z \in \mathbb{C}$ except for real z such that $z \geq 1$. Since (by the definition of the stretching exponent) $0 < \zeta \leq 1$ and $iu\,\langle A\rangle$ is imaginary, it results that $L(iu\,\langle A\rangle\,;\zeta)$ always converges—and thus $\ln C_B(u)$ always exists.

The characteristic function of a sum of random variables is the product of the individual random variables. Hence, from

$$I_{sat}(t) = I_B(t) + I_b(t)$$

follows that the characteristic function of the pulse train model with pulse background is

$$C_I(u) = C_B(u)C_b(u)$$

with $C_B(u)$ and $C_b(u)$ given respectively by Eq. (7.7) and (7.6).

7.3 Moments of the SPTM-PB

As seen in Eq. (7.7), the characteristic function logarithm of the stretched bursts $I_B(t)$ is rather complicated and is given in terms of the nonelementary integral $L(iu\,\langle A\rangle\,;\zeta)$. Nevertheless, with the CF it is possible and easy to deduce the moments of the PDF of $I_B(t)$. As detailed in Appendix B.8, for $|iu\,\langle A\rangle\,| < 1$, the integral $L(iu\,\langle A\rangle\,;\zeta)$ can be written as[2]

$$L(iu\,\langle A\rangle\,;\zeta) = \Gamma\left(1 + \frac{1}{\zeta}\right) \sum_{n=1}^{\infty} \frac{(iu\,\langle A\rangle)^n}{n^{1/\zeta}}$$

[2]For $\zeta = 1$, this infinite sum is equal to

$$L(iu\,\langle A\rangle\,;1) = \sum_{n=1}^{\infty} \frac{(iu\,\langle A\rangle)^n}{n} = -\ln\left(1 - iu\,\langle A\rangle\right)$$

Therefore, in this case, the logarithm of $C_B(u)$ (Eq. (7.7)),

$$\ln C_B(u) = \gamma\lambda \cdot L\left(iu\left\langle A\right\rangle; \zeta_r\right) + \gamma(1-\lambda) \cdot L\left(iu\left\langle A\right\rangle; \zeta_f\right)$$

becomes

$$\ln C_B(u) = \gamma \sum_{n=1}^{\infty} \left(iu\left\langle A\right\rangle\right)^n \left[\frac{\lambda}{n^{1/\zeta_r}}\Gamma\left(1+\frac{1}{\zeta_r}\right) + \frac{1-\lambda}{n^{1/\zeta_f}}\Gamma\left(1+\frac{1}{\zeta_f}\right)\right] \tag{7.8}$$

As observed in Section 2.2, the logarithm of the CF is a cumulant generating function,

$$\ln C_B(u) = \sum_{n=1}^{\infty} \kappa_n \frac{(iu)^n}{n!}$$

Hence, from Eq. (7.8) the cumulants can be written as

$$\kappa_n = s_n\left(\lambda, \zeta_r, \zeta_f\right) \times (n-1)!\, \gamma \left\langle A\right\rangle^n$$

where we defined the stretch coefficient of order n as

$$s_n\left(\lambda, \zeta_r, \zeta_f\right) \equiv \frac{\lambda}{n^{1/\zeta_r - 1}}\Gamma\left(1+\frac{1}{\zeta_r}\right) + \frac{1-\lambda}{n^{1/\zeta_f - 1}}\Gamma\left(1+\frac{1}{\zeta_f}\right) \tag{7.9}$$

Note that for unstretched pulses $s_n\left(\lambda, 1, 1\right) = 1$.

With the cumulants and Eq. (2.13), it is possible to obtain the mean, variance, skewness and kurtosis of the train of stretched pulses $I_B(t)$ (Eq. (7.2)),

$$\begin{aligned}
\left\langle I_B\right\rangle &= \kappa_1 \\
\sigma_B^2 &= \kappa_2 \\
S_B &= \frac{\kappa_3}{\kappa_2^{3/2}} \\
K_B &= 3 + \frac{\kappa_4}{\kappa_2^2}
\end{aligned}$$

(Continuation of the footnote of the previous page.) Hence, for this case of unstretched pulses, the $C_B(u)$ from Eq. (7.7) is the characteristic function of a Gamma distribution,

$$\ln C_B(u) = \gamma\lambda \cdot L\left(iu\left\langle A\right\rangle; 1\right) + \gamma(1-\lambda) \cdot L\left(iu\left\langle A\right\rangle; 1\right) \implies$$

$$\ln C_B(u) = -\gamma \ln\left(1 - iu\left\langle A\right\rangle\right) \implies$$

$$C_B(u) = \left(1 - iu\left\langle A\right\rangle\right)^{-\gamma}$$

just as in Eq. (7.6), for example.

Therefore

$$\langle I_B \rangle = \gamma \langle A \rangle \left[\lambda \Gamma \left(1 + \frac{1}{\zeta_r} \right) + (1 - \lambda) \Gamma \left(1 + \frac{1}{\zeta_f} \right) \right]$$

$$\sigma_B^2 = \gamma \langle A \rangle^2 s_2(\lambda, \zeta_r, \zeta_f)$$

$$S_B = \frac{2}{\gamma^{1/2}} \frac{s_3(\lambda, \zeta_r, \zeta_f)}{s_2^{3/2}(\lambda, \zeta_r, \zeta_f)}$$

$$K_B = 3 + \frac{6}{\gamma} \frac{s_4(\lambda, \zeta_r, \zeta_f)}{s_2^2(\lambda, \zeta_r, \zeta_f)}$$

For the stochastic pulse train model with pulse background (SPTM-PB), the saturation current is given by Eq. (7.1),

$$I_{sat}(t) = I_B(t) + I_b(t)$$

Since the cumulants are cumulative (Eq. (2.12)), then the mean, variance, skewness and kurtosis of the SPTM-PB are respectively

$$\boxed{\begin{aligned} \langle I_{sat} \rangle &= s_1(\lambda, \zeta_r, \zeta_f) \gamma \langle A \rangle + \gamma^{(b)} \left\langle A^{(b)} \right\rangle \\ \sigma_I^2 &= s_2(\lambda, \zeta_r, \zeta_f) \gamma \langle A \rangle^2 + \gamma^{(b)} \left\langle A^{(b)} \right\rangle^2 \\ S_I &= \frac{2}{\sigma_I^3} \left[s_3(\lambda, \zeta_r, \zeta_f) \gamma \langle A \rangle^3 + \gamma^{(b)} \left\langle A^{(b)} \right\rangle^3 \right] \\ K_I &= 3 + \frac{6}{\sigma_I^4} \left[s_4(\lambda, \zeta_r, \zeta_f) \gamma \langle A \rangle^4 + \gamma^{(b)} \left\langle A^{(b)} \right\rangle^4 \right] \end{aligned}} \tag{7.10}$$

where the stretch coefficient $s_n(\lambda, \zeta_r, \zeta_f)$ is defined in Eq. (7.9).

7.4 Fitting the SPTM-PB

As seen in Section 7.2, with stretched pulses, the characteristic function can no longer be given in terms of elementary functions. Because of this, the CF-CAB-PSD fitting method of Chapter 5 cannot be used to adjust the stochastic pulse train model with pulse background ($I_{sat} = I_B + I_b$). Moreover, even with unstretched pulses ($\zeta = 1$), it is difficult to implement the CF-CAB-PSD fit for the SPTM-PB, because in this case both I_B and I_b are a train of unstretched pulses, and then the characteristic function fit is not sensible to choose between the parameters of I_B and I_b.

It results that the fitting procedure adopted was to manually adjust the simulated mean and standard deviation to the experimental I_{sat} and also analyze the reduced χ^2 from various diagnostics—PDF, PSD, conditionally averaged burst, CH point, and distributions of bursts amplitude and waiting time. For this reason, the obtained profiles for the SPTM-PB parameters are not entirely objective and it would be desirable to improve the fitting method for the SPTM-PB in future works. Nevertheless, as will be shown in Chapter 8, these preliminary results already present important advantages with respect to the SPTM-N model (which was described in Chapters 4, 5 and 6).

Chapter 8

Simulations of the SPTM with pulse background to describe density fluctuations in TCABR

This chapter compares simulations of the SPTM-PB (Chapter 7) with the experiment 34132 of TCABR (described in Sections 3.2–3.4), in a similar manner to what was done to the SPTM-N case (Chapter 6). First, the transition from the plasma edge to the scrape-off layer will be presented (Section 8.1). Then radial profiles of the SPTM-PB parameters will be depicted in Section 8.2. Finally, a comparison between backgrounds will be made in Section 8.3, using the CH diagram and the RQA determinism (Sections 2.7 and 2.8).

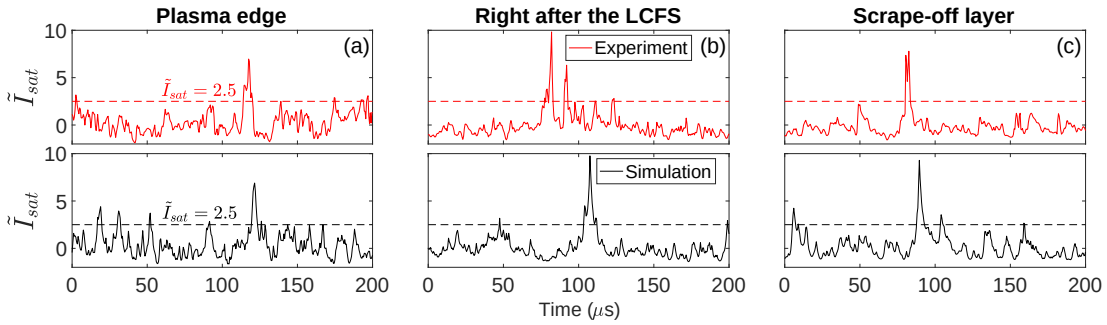


Figure 8.1: The top panels show excerpts of saturation current for three positions in the TCABR tokamak: (a) in the plasma edge, $r = 17.0$ cm; (b) right after the last closed flux surface (LCFS), $r = 18.5$ cm; (c): at the scrape-off layer, $r = 19.5$ cm. The bottom panels show SPTM-PB simulations. In each panel, at least one large-amplitude burst is seen, with a peak above $\tilde{I}_{sat} > 2.5$.

8.1 Transition from the plasma edge to the SOL

In this section, three positions of the TCABR will be compared (as done in Section 6.2 for the SPTM-N): the plasma edge ($r = 17.0$ cm), right after the last closed flux surface ($r = 18.5$ cm $> a$) and at the scrape-off layer ($r = 19.5$ cm $> a$). For the results of all eight positions, the reader is referred to Appendix B.3.2.

The normalized saturation current will be used, as in Chapter 6,

$$\tilde{I}_{sat}(t) = \frac{I_{sat} - \langle I_{sat} \rangle}{\sigma_I}$$

Fig. 8.1 shows excerpts of the experimental and simulated saturation current, in red and black, respectively. The simulations seem to reproduce well the behavior of the experiment.

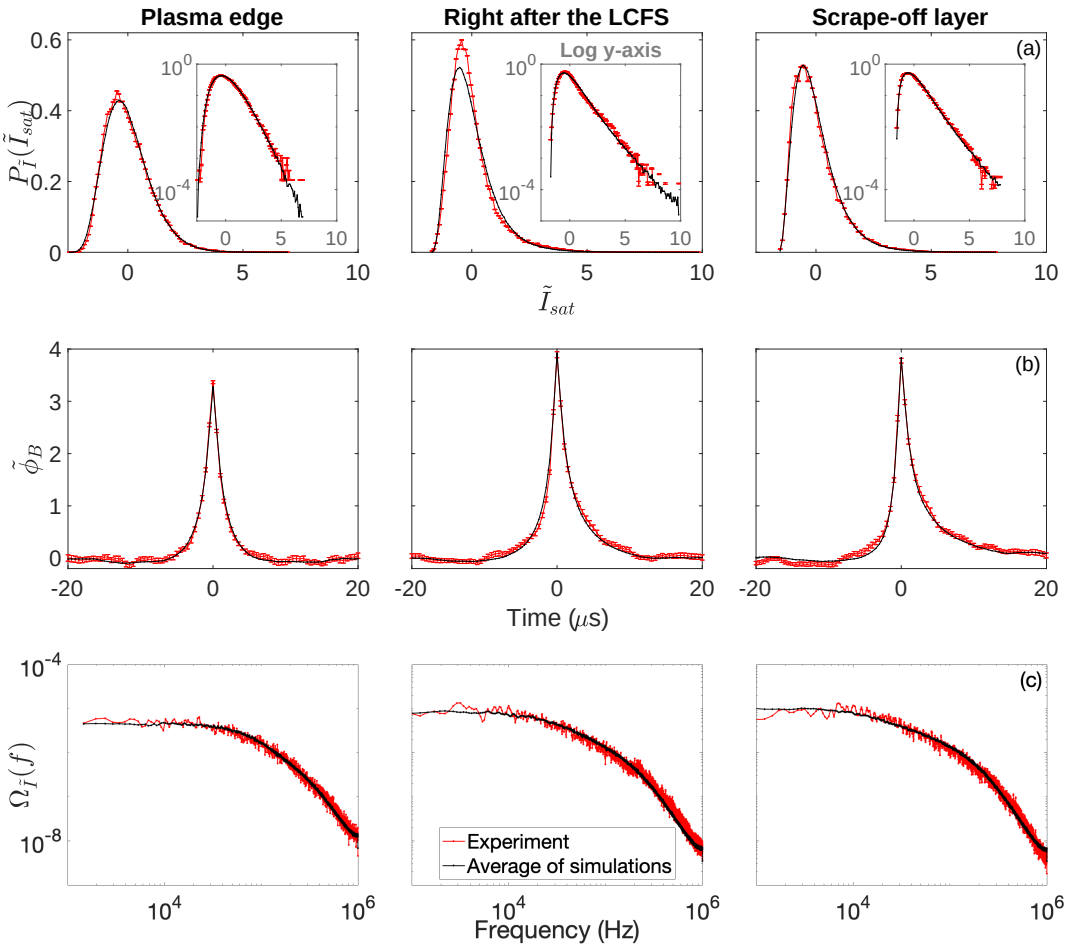


Figure 8.2: In red, (a) PDFs (in linear and log-y scales), (b) conditionally averaged bursts and (c) power spectra of $\tilde{I}_{sat}$ for three positions—plasma edge ($r = 17.0$ cm), right after the last closed flux surface ($r = 18.5$ cm) and at the scrape-off layer ($r = 19.5$ cm). In black, simulations of the SPTM-PB (Chapter 7).

Fig. 8.2 compares the distributions, conditionally averaged bursts and power spectra of $\tilde{I}_{sat}$ for the same positions of Fig. 8.1. In opposition to what was seen for the SPTM with Gaussian noise Fig. 6.2, the simulated averaged bursts and frequency spectra are in great agreement with the experiment. The first occurred because of the use of stretched bursts ($\zeta \neq 1$), as described in Section 7.1. The use of pulse background, in this case, makes no difference with respect to the noise background (Chapters 4 and 6), because the background structure is not important to the conditionally averaged burst. In its turn, we infer that

the power spectra of Fig. 8.2 were better fitted because of the pulse background. This last statement is in accordance with the prediction of Appendix B.6: that the SPTM with a pulse background fits the I_{sat} PSDs in TCABR better than the SPTM with Gaussian noise.

Nevertheless, as seen in the middle panel of Fig. (8.2), the SPTM-PB couldn't fit the I_{sat} distribution so well as the SPTM-N. (For all the other five fitted PDFs, see Appendix B.3.2.) This indicates that the stochastic pulse train model with pulse background isn't able to perfectly reproduce the Gaussian background. Thus, with the PB model, a small noise term may still be needed to accurately describe small fluctuations.

In conclusion, the SPTM-PB described well the transition from the plasma edge to the scrape-off layer of TCABR, from the perspective of conditionally averaged bursts and the power spectra. In contrast, the PDF was not so well fitted.

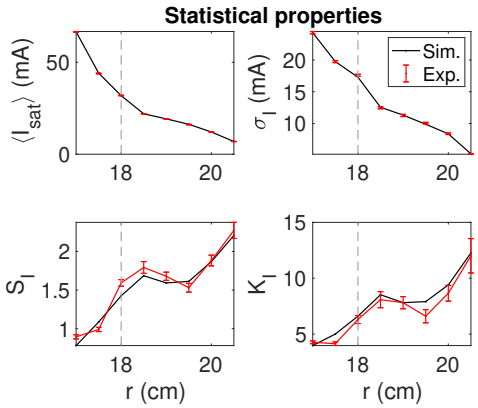


Figure 8.3: (a) Four main statistical properties of I_{sat}: mean, standard deviation, skewness and kurtosis. The dashed line in $r = a = 18.0$ cm indicates the limiter position.

8.2 Parameters and moments

Fig. 8.3 shows the average, standard deviation, skewness and kurtosis of the I_{sat} measured in TCABR, compared to the values predicted by the SPTM-PB model. As expected, the simulated mean and standard deviation are in excellent agreement with the experiment, since these two were used to fit the SPTM-PB (Section 7.4). The skewness and kurtosis are also well reproduced, although the predicted values of the kurtosis are in general slightly above the experimental ones.

Fig. 8.4 presents the nine parameters of the SPTM-PB fitted for the TCABR experiment. More specifically, Fig. 8.4(a) depicts the principal parameters of the SPTM-PB distribution.[1] The average amplitude of the pulses, $\langle A \rangle$, presented a global maximum around the minor

[1] As discussed in Section 7.2, in this case, the PDF (which has the same information as the CF) depends on γ, $\langle A \rangle$, $\gamma^{(b)}$, $\langle A^{(b)} \rangle$, λ, ζ_r and ζ_f. Nevertheless, the dependence on the last three occurs only through the stretch coefficients $s_n(\lambda, \zeta_r, \zeta_f)$. For this reason, here γ, $\langle A \rangle$, $\gamma^{(b)}$ and $\langle A^{(b)} \rangle$ are referred to as the main parameters of the distribution.

radius $a = 18.0$ cm. This was also obtained for the SPTM-N (Fig. 6.3(a)). All the other parameters on Fig. 8.4(a) in general decay with the radial position, which makes sense since $\langle I_{sat} \rangle$ and σ_I decay with r (Fig. 8.3). Fig. 8.4(b) depicts the profiles of the pulse parameters λ, ζ_r, ζ_f, τ_d and $\tau_d^{(b)}$.

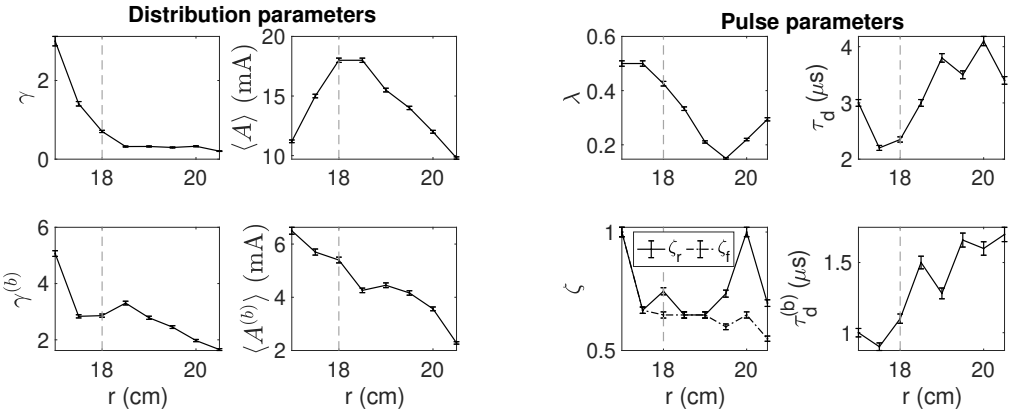


Figure 8.4: Parameters of the SPTM-PB fitted for the TCABR experiment. The exact values are stored in Tables B.4 and B.5.

The parameters from the main pulses, λ, ζ_r, ζ_f and τ_d, are difficult to fit manually (and thus should not be taken as an ultimate answer). This happens because they are highly correlated, which can be more easily understood in the terms of the rise and fall characteristic durations, $\tau_r = \lambda\tau_d$ and $\tau_f = (1-\lambda)\tau_d$. If τ_r is overestimated, then also would be ζ_r, so that the overestimated duration is compensated by a less stretched pulse ($\zeta_r \to 1$). The same argument holds for τ_f and ζ_f.

Nevertheless, the effective duration $\tau_{ef} = s_1(\lambda, \zeta_r, \zeta_f)\tau_d$ is well behaved, since the stretching and duration effects are combined. It is τ_{ef} who matters for the average of I_{sat}, as can be noted using Eq. (7.10),

$$\langle I_{sat} \rangle = s_1(\lambda, \zeta_r, \zeta_f)\,\gamma\,\langle A \rangle + \gamma^{(b)}\left\langle A^{(b)} \right\rangle \implies$$

$$\langle I_{sat} \rangle = \gamma_{ef}\,\langle A \rangle + \gamma^{(b)}\left\langle A^{(b)} \right\rangle$$

where the effective intermittency parameter was defined as $\gamma_{ef} = s_1\gamma = s_1\tau_d N_P/T = \tau_{ef}N_P/T$. The dashed line in Fig. 8.5(a) depicts the values obtained for the effective duration $\tau_{ef} = s_1\tau_d$. It is seen that its profile is better behaved than the one of τ_d. The effective duration is stable for $r \leq 18.0$ cm. Then it rapidly increases until $r = 19$ cm, a position in which it becomes stable again.

Lastly, Fig. 8.5(b) exhibits the profiles of the stretch coefficients s_1, s_2, s_3 and s_4, which influence the I_{sat} average, standard deviation, skewness and kurtosis in the SPTM-PB. We

recall that the stretch coefficient of order n was defined in Eq. (7.9),

$$s_n\left(\lambda,\zeta_r,\zeta_f\right) \equiv n\left[\frac{\lambda}{n^{1/\zeta_r}}\Gamma\left(1+\frac{1}{\zeta_r}\right)+\frac{1-\lambda}{n^{1/\zeta_f}}\Gamma\left(1+\frac{1}{\zeta_f}\right)\right]$$

Note that $s_n(\lambda,1,1)=1$. For $n=1$,

$$s_1\left(\lambda,\zeta_r,\zeta_f\right) = \left[\lambda\Gamma\left(1+\frac{1}{\zeta_r}\right)+(1-\lambda)\,\Gamma\left(1+\frac{1}{\zeta_f}\right)\right]$$

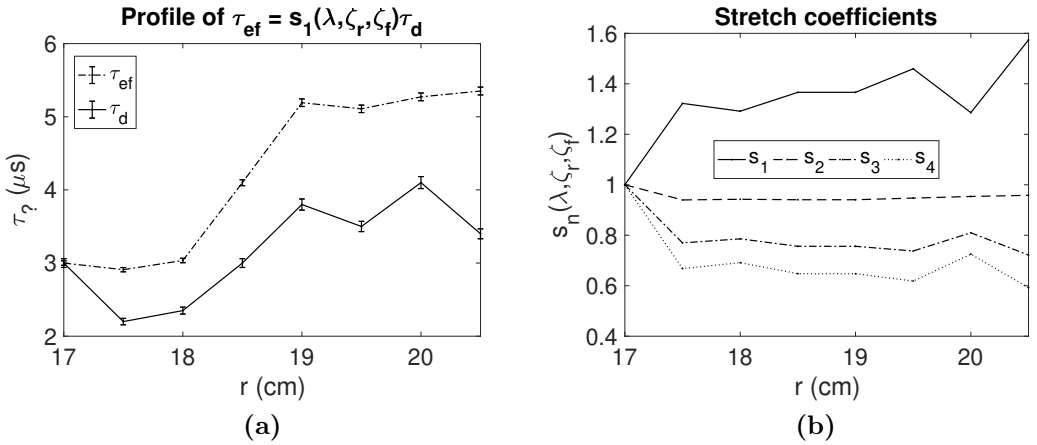
Hence, s_1 increases with $1/\zeta_r$ and $1/\zeta_f$ and thus $s_1(\lambda,\zeta_r,\zeta_f)\geq 1$ (for ζ_r,ζ_f and λ defined between 0 and 1). Conversely, for $n=2$ and $0.5\leq\zeta\leq 1$ it results that

$$0.94 < \frac{2}{2^{1/\zeta}}\Gamma\left(1+\frac{1}{\zeta}\right)\leq 1$$

Therefore s_2 is really close to 1 and the variance of I_{sat}, $\sigma_I^2 = s_2\gamma\left\langle A\right\rangle^2+\gamma^{(b)}\left\langle A^{(b)}\right\rangle$, is not much affected by the stretch coefficient. For $n>2$, contrastingly, the factor $n/n^{1/\zeta}\Gamma(1+1/\zeta)$ rapidly decreases with $1/\zeta$ and hence $s_n\leq 1$ for $n>2$. In resume, the stretch coefficients behave like

$$\begin{aligned} s_1\left(\lambda,\zeta_r,\zeta_f\right) &\geq 1\\ s_2\left(\lambda,\zeta_r,\zeta_f\right) &\approx 1\\ s_{n>2}\left(\lambda,\zeta_r,\zeta_f\right) &\leq 1 \end{aligned}$$

for $\zeta_r,\zeta_f\in[0.5,1]$ and $\lambda\in[0,1]$. Their behavior is shown in Fig. 8.5(b).



(a) **(b)**

Figure 8.5: **(a)** Profile of the effective pulse duration $\tau_{ef}=s_1\tau_d$, compared to the profile of the pulse duration τ_d. **(b)** Profiles of the stretch coefficients s_1, s_2, s_3 and s_4 obtained for the TCABR experiment using the SPTM-PB.

8.3 Comparison between background models

In this section, the pulse background model (SPTM-PB) will be compared to the TCABR experiment and Gaussian noise models (SPTM-N), using nonlinear tools, namely the Complexity-Entropy diagram (Section 2.7) and the determinism from recurrence quantification analysis (Section 2.8).

The complexity-entropy points of the SPTM-PB are shown in the squares of Fig. 8.6. Unlike what was obtained for the SPTM-N (Fig. 6.4.2), the points from the pulse background case are in great agreement with the experimental data, in circles.

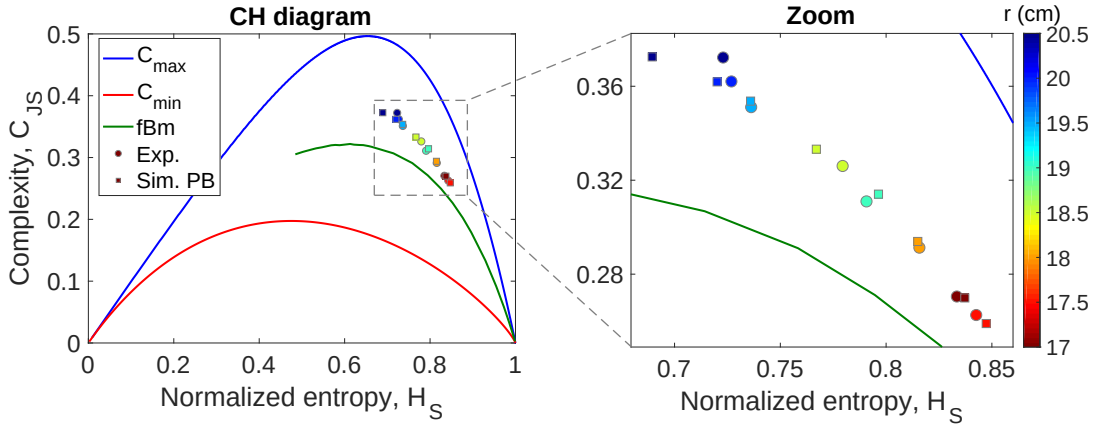
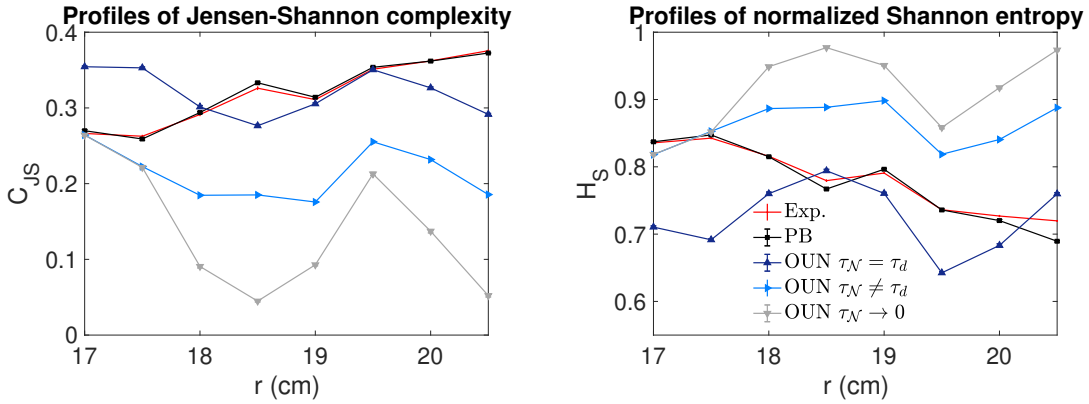


Figure 8.6: Complexity-Entropy diagram and its points for the TCABR experiment and the SPTM-PB data.



(a) **(b)**

Figure 8.7: Profiles of the Jensen-Shannon complexity (a), and the Shannon normalized complexity (b). In red are the data from the TCABR experiment. In other colors are the synthetic data corresponding to dynamical Ornstein-Uhlenbeck noise (OUN), in dark blue; fitted OUN, in light blue; white noise, in grey; pulse background, in black.

It is elucidating to analyze the entropy and complexity separately, as done in Fig. 8.7(a)

and (b). The PB simulations have Shannon entropy and Jensen-Shannon complexity in excellent agreement with the experiment, in opposition to the Gaussian noise simulations. Fig. 8.8 shows the profile of the RQA determinism.[2] The PB profile is similar to the fitted OU noise case ($\tau_{\mathcal{N}} \neq \tau_d$) and the experiment.

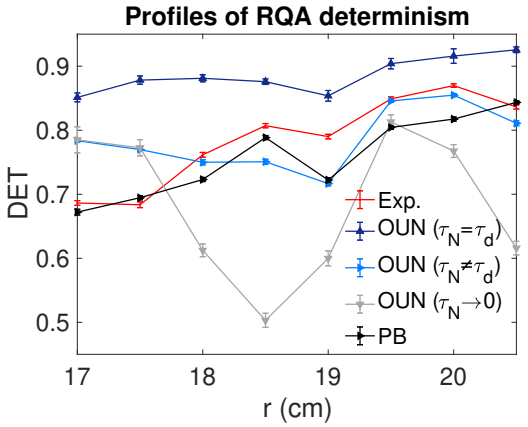


Figure 8.8: Profiles of the RQA determinism from the same signals of Fig. 8.7.

8.4 Summary and discussion

In this chapter, the stochastic pulse train model with pulse background (SPTM-PB) was applied to local measurements of density fluctuations in the TCABR tokamak. The model reproduced well the experimental conditionally averaged bursts and the power spectral densities from I_{sat} (Fig. 8.2). The I_{sat} PDF was adequately fitted for $r = 17.0$ cm and $r = 19.5$ cm in Fig. 8.2(a). For $r = 18.5$ cm, on the other hand, the PDF was not so well adjusted, and we found out that the same occurred for $r \in [18.0, 20.0, 20.5]$ cm (Fig. B.6). These results for the CAB, PSD and PDF are in contrast with the obtained for the SPTM-N (Section 6.2), since this last model reproduced well the PDF for all positions, but not the CAB and PSD.

In Section 8.2, radial profiles of the obtained parameters were shown. Lastly, in Section 8.3 two non-linear diagnostics were used to compare the experiment with the SPTM-N and the SPTM-PB. Data of the SPTM-PB obtained from the CH diagram and the RQA determinism showed good agreement with the experiment, in contrast to the SPTM-N data.

Furthermore, Ref. 80 questioned the possibility of the stochastic model to figure inside the chaotic region in the CH diagram (i.e., above the fBm curve). It is therefore a striking result that the model with pulse background not only managed to be in this region, but was also extremely close to the experimental data. The fact that the PB data are similar to the experimental ones for the CH diagram and the RQA determinism indicates that this model and the experimental signal have a similar dynamic behavior.

[2]The RQA determinism is obtained from recurrence plots. For recurrence plots of the experimental and simulated I_{sat} of TCABR, see Appendix B.5.

Chapter 9

Conclusion

As stated in Chapter 1, turbulence in the edge of magnetized plasmas transports high-amplitude bursts to the boundary of tokamaks, deteriorating the particle confinement [25, 26, 65]. Our work focus on modeling local measurements of density fluctuations with bursts in the TCABR tokamak. With a stochastic pulse train model (SPTM) [50, 54], different types of backgrounds are used to describe these signals: four different Gaussian noises (three of them portrayed in Ref. 48 and 54) and a pulse background [48]. They are summarized in Table 9.1.

Table 9.1: Different background signals used with the SPTM for this dissertation. OU stands for Ornstein-Uhlenbeck. All four noises are Gaussian and the $\tau_{\mathcal{N}} \neq \tau_d$ case was proposed in the present work. For details, see Chapters 4, 5 and 7.

Background	Correlation parameter	How is set
Colored noise	η	η is fitted
White noise	0	$\tau_{\mathcal{N}} = 0$ or $\eta = 0$
Dynamical OU noise	$\tau_{\mathcal{N}}$	$\tau_{\mathcal{N}} = \tau_d$
Fitted OU noise	$\tau_{\mathcal{N}}$	$\tau_{\mathcal{N}}$ is fitted ($\tau_{\mathcal{N}} \neq \tau_d$)
Pulse background	$\tau_d^{(b)}$	$\tau_d^{(b)}$ is fitted

However, these models hadn't yet been compared simultaneously to an experiment. Furthermore, there wasn't a fitting method for the SPTM that could adjust all its parameters in a unified and objective way. In this dissertation, we aimed to tackle these two gaps.

To do the analysis, we introduced the CF-CAB-PSD fitting method for this stochastic model, considering correlated noise (Chapter 5). The fit is consistent and can be applied to different tokamaks.

In Chapter 6, the fit and model were applied to the TCABR tokamak. Between the four noise models considered, the white noise case had the worst results when concerning the RQA determinism and the CH diagram. Also, the results from the model with colored noise were very similar to the ones of the fitted Ornstein-Uhlenbeck noise. However, since the power spectrum of the colored noise is ill-defined, it is preferable to work with OU processes. Therefore, out of the four noise models for the SPTM, we only suggest the use of two: the dynamical and the fitted Ornstein-Uhlenbeck processes.

The model distributions were in excellent agreement with the experimental ones. However, the simulated conditionally averaged bursts (CABs), power spectral densities (PSDs), RQA

determinism and CH points weren't in so good accordance with the experimental values, even for the best noise models.

With this in mind and inspired by Refs. 15 and 48, we introduced a pulse train model with stretched bursts and a pulse background (Chapters 7 and 8). Surprisingly, this SPTM-PB had worse distributions than the SPTM-N, but better CAB, PSD, DET and CH. In special, not only its points were really close to the experimental ones in the CH diagram, but they also appeared in the chaotic region of the plane. This remarkable result was thought to be unlikely [80], since the model used is stochastic. Hence, our analysis indicates that the SPTM with pulse background and the density fluctuations in TCABR have a similar dynamic nature. For further works, it would be interesting to improve the fitting method for the SPTM-PB and apply it to more experiments in TCABR and other plasma devices.

Bibliography

[1] Smil V & BP Statistical Review of World Energy. *Energy Transitions: Global and National Perspectives.* (2017). Published online at Our World In Data. Cited on pages 19 and 20.

[2] *BP Statistical Review of World Energy 2016.* Published online at Our World In Data. Cited on page 19.

[3] Smith CL. *The need for fusion.* Fusion Engineering and Design Vol. 74, Issues 1–4, pag. 3–8 (2005). Cited on page 19.

[4] Ikeda K. *ITER on the road to fusion energy.* Nuclear Fusion **50**, 014002 (2010). Cited on pages 19, 20, and 23.

[5] ITER newsline. *Will Fusion Run Out of Fuel?* Question of the week on ITER newsline of 9/7/18 (2018). Cited on page 19.

[6] Chen FF. *An Indispensable Truth: How Fusion Power Can Save the Planet.* Springer International Publishing, 1st Ed. (2011). Cited on pages 19, 20, 23, and 25.

[7] Magaud P, Marbach G and Cook I. Cleveland CJ (ed.). *Encyclopedia of Energy.* Vol. 4. (2004) Elsevier. "Nuclear Fusion Reactors", pp. 365–381. Cited on page 19.

[8] Waldrop MM. *Plasma physics: The fusion upstarts.* Nature, **511**, p. 398–400 (2014). Cited on page 21.

[9] Greenwald M. *Status of the SPARC physics basis.* Journal of Plasma Physics, vol. **86**, 861860501 (2020). Cited on pages 21 and 23.

[10] Chen FF. *Introduction to Plasma Physics and Controlled Fusion.* Springer International Publishing, 3rd Ed. (2016). Cited on pages 21, 22, 25, and 29.

[11] Bittencourt JA. *Fundamentals Of Plasma Physics.* Springer International Publishing, 3rd Ed. (2004). Cited on pages 21, 22, 56, and 57.

[12] Galvão RMO et al. *Report on recent results obtained in TCABR.* Journal of Physics: Conference Series **591**, 012001 (2015). Cited on pages 12, 23, 24, 25, and 26.

[13] Grenfell GG *et al. H-mode access and the role of spectral shift with electrode biasing in the TCABR tokamak.* Physics of Plasmas **25**, 072301 (2018). Cited on pages 23 and 62.

[14] Cheetham AD et al. *The TCA tokamak project report 1979.* EPFL technical report (1980). Cited on pages 23 and 24.

[15] Baquero WAH. *Turbulent temperature fluctuations in TCABR tokamak.* Doctoral thesis, Physics Institute of the University of São Paulo. São Paulo (2017). Retrieved from teses.usp.br. Cited on pages 12, 26, 27, 56, 59, 61, 67, 75, 95, and 109.

[16] Ronchi GG. *Estudo de Perfis de Pressão no Tokamak TCABR.* Doctoral thesis, Physics Institute of the University of São Paulo. São Paulo (2017). Retrieved from teses.usp.br. Cited on pages 12 and 26.

[17] Artsimovich LA and Kurchatov IV. *Tokamak devices.* Nuclear Fusion Vol. 12, N. 2, p. 215 (1972). Cited on page 26.

[18] Wootton AJ, Carreras BA, Matsumoto H, McGuire K, Peebles WA *et al. Fluctuations and anomalous transport in tokamaks.* Physics of Fluids B vol. 2, p. 2879 (1990). Cited on page 26.

[19] Krasheninnikov S, Smolyakov A, Kukushkin *A. On the Edge of Magnetic Fusion Devices.* Springer, 1st Ed. (2020). s 7. Cited on pages 26 and 29.

[20] Conway GD. *Turbulence measurements in fusion plasmas.* Plasma Physics and Controlled Fusion 50 124026 (2008). Cited on pages 26 and 63.

[21] Connor JW. *Tokamak Turbulence—Electrostatic or Magnetic?* Plasma Physics and Controlled Fusion vol. 35, p. B293–B305 (1993). Cited on page 26.

[22] LaBombard B *et al. Particle transport in the scrape-off layer and its relationship to discharge density limit in Alcator C-Mod.* Physics of Plasmas **8**, 2107 (2001). Cited on page 26.

[23] Pigarov AY, Krasheninnikov SI, Rognlien TD, Schaffer MJ, and West WP. *Tokamak edge plasma simulation including anomalous cross-field convective transport.* Physics of Plasmas **9**, 1287 (2002). Cited on page 26.

[24] Antar GY, Krasheninnikov SI, Devynck P, Doerner RP, Hollmann EM, Boedo JA, Luckhardt SC and Conn RW. *Experimental Evidence of Intermittent Convection in the Edge of Magnetic Confinement Devices.* Physical Review Letters **87**, 065001 (2001). Cited on pages 26, 27, and 31.

[25] Antar GY, Counsell G, Yu Y, LaBombard B and Devynck P. *Universality of intermittent convective transport in the scrape-off layer of magnetically confined devices.* Physics of Plasmas **10**, 419 (2003). Cited on pages 26, 27, 59, 60, 61, and 108.

[26] LaBombard B, Boivin RL, Greenwald M, Hughes J, Lipschultz B, Mossessian D, Pitcher CS, Terry JL, Zweben SJ, and Alcator Group. *Particle transport in the scrape-off layer and its relationship to discharge density limit in Alcator C-Mod.* Physics of Plasmas, Vol. 8, No. 5, (2001). Cited on pages 26 and 108.

[27] D'Ippolito DA, Myra JR, and Zweben SJ. *Convective transport by intermittent blob-filaments: Comparison of theory and experiment.* Physics of Plasmas **18**, 060501 (2011). Cited on pages 26, 27, 28, and 33.

[28] Nielsen AH, Juul Rasmussen J, Madsen J , Xu GS , Naulin V, Olsen JMB, Løiten M, Hansen SK , Yan N, Tophøj L, and Wan BN. *Numerical simulations of blobs with ion dynamics*. Plasma Physics and Controlled Fusion **59**, 025012 (2017). Cited on page 27.

[29] Pereira FAC, Hernandez WA, Toufen DL, Guimarães-Filho ZO, Caldas IL, and Gentle KW. *Burst temperature from conditional analysis in Texas Helimak and TCABR tokamak*. Physics of Plasmas **25**, 042301 (2018). Cited on page 27.

[30] Garcia OE, Horacek J and Pitts RA. *Intermittent fluctuations in the TCV scrape-off layer*. Nuclear Fusion **55**, 062002 (2015). Cited on pages 27, 58, 59, 61, 67, 77, and 83.

[31] Theodorsen A, Garcia OE, Horacek J, Kube R and Pitts RA. *Scrape-off layer turbulence in TCV: evidence in support of stochastic modelling*. Plasma Physics and Controlled Fusion **58**, 044006 (2016). Cited on pages 27, 33, 58, 59, 61, 67, 77, and 83.

[32] Garcia OE, Kube R, Theodorsen A, Bak JG, Hong S H, Kim HS, the KSTAR Project Team and Pitts RA. *SOL width and intermittent fluctuations in KSTAR*. Nuclear Materials and Energy **12**, p. 36–43 (2017). Cited on pages 27, 33, 34, 58, 59, 61, 63, 67, 77, 83, and 87.

[33] Garcia OE, Fritzner SM, Kube R, Cziegler I, LaBombard B and Terry JL. *Intermittent fluctuations in the Alcator C-Mod scrape-off layer*. Physics of Plasmas **20**, 055901 (2013). Cited on pages 27, 33, 59, 61, 63, 67, 77, 83, and 87.

[34] Kube R, Theodorsen A, Garcia OE, LaBombard B and Terry JL. *Fluctuation statistics in the scrape-off layer of Alcator C-Mod*. Plasma Physics and Controlled Fusion **58**, 054001 (2016). Cited on pages 27, 33, 34, 46, 58, 59, 61, 67, 69, 77, 83, and 88.

[35] Theodorsen A, Garcia OE, Kube R, LaBombard B and Terry JL. *Relationship between frequency power spectra and intermittent, large-amplitude bursts in the Alcator C-Mod scrape-off layer*. Nuclear Fusion **57**, 114004 (2017). Cited on pages 27, 41, 59, 61, 63, 67, 69, 77, and 83.

[36] Garcia OE, Kube R, Theodorsen A, LaBombard B, Terry JL. *Intermittent fluctuations in the Alcator C-Mod scrape-off layer for ohmic and high confinement mode plasmas*. Physics of Plasmas **25**, 056103 (2018). Cited on pages 27, 28, 33, 59, 63, 67, 77, 83, and 87.

[37] Theodorsen A, Garcia OE, Kube R, LaBombard B and Terry JL. *Universality of Poisson-driven plasma fluctuations in the Alcator C-Mod scrape-off layer*. Physics of Plasmas **25**, 122309 (2018). Cited on pages 27, 28, 34, 59, 67, 77, 88, and 91.

[38] Kube R, Garcia OE, Theodorsen A, D Brunner D, Kuang AQ, LaBombard B and Terry JL. *Intermittent electron density and temperature fluctuations and associated fluxes in the Alcator C-Mod scrape-off layer*. Plasma Physics and Controlled Fusion **60**065002 (2018). Cited on pages 27, 31, 58, 70, and 77.

[39] Kube R, Theodorsen A, Garcia OE, Brunner D, B. LaBombard B and Terry JL. *Comparison between mirror Langmuir probe and gas-puff imaging measurements of intermittent fluctuations in the Alcator C-Mod scrape-off layer.* Journal of Plasma Physics vol. **86**, 905860519 (2020). Cited on pages 27, 58, 59, 63, and 77.

[40] Walkden NR, Wynn A, Militello F, Lipschultz B, Matthews G, Guillemaut C, Harrison J, Moulton D and JET Contributors. *Statistical analysis of the ion flux to the JET outer wall.* Nuclear Fusion **57**, 036016 (2017). Cited on pages 27, 58, 59, 61, 66, 67, 69, 77, 83, and 88.

[41] Bencze A, Berta M, Buzás A, Hacek P, Krbec J, Szutyányi M and the COMPASS Team. *Characterization of edge and scrape-off layer fluctuations using the fast Li-BES system on COMPASS.* Plasma Physics and Controlled Fusion **61**, 085014 (2019). Cited on pages 27, 34, 59, 61, 66, 67, 69, 77, and 88.

[42] Greenwald M et al. *20 years of research on the Alcator C-Mod tokamak.* Physics of Plasmas **21**, 110501 (2014). Cited on pages 27 and 77.

[43] Zweben SJ and Gould RW. *Structure of edge-plasma turbulence in the Caltech tokamak.* Nuclear Fusion **25**, 171 (1985). Cited on page 27.

[44] Zweben SJ et al. *Blob structure and motion in the edge and SOL of NSTX.* Plasma Physics and Controlled Fusion **58**, 044007 (2016). Cited on page 27.

[45] Hernandez WA, Guimarães-Filho ZO, Grenfell GG and Nascimento IC. *Spatial inhomogeneity effects on burst temperature estimation using a triple probe configuration in Tokamak Chauffage Alfvén Brésilien tokamak.* Journal of Plasma Physics **85**, 905850407 (2019). Cited on pages 27, 56, and 58.

[46] Ferreira AA, Heller MVAP, Caldas IL, Lerche EA, Ruchko LF and Baccalá LA. *Turbulence and transport in the scrape-off layer TCABR tokamak.* Plasma Physics and Controlled Fusion **46**, 669 (2004). Cited on page 27.

[47] Fuchert G, Birkenmeier G, Nold B, Ramisch M and Stroth U. *The influence of plasma edge dynamics on blob properties in the stellarator TJ-K.* Plasma Physics and Controlled Fusion **55**, 125002 (2013). Cited on page 27.

[48] Pereira FAC, Sokolov IM, Toufen DL, Guimarães-Filho ZO, Caldas IL and Gentle KW. *Statistical properties of intermittent bursts in the Texas Helimak.* Physics of Plasmas **26**, 052301 (2019). Cited on pages 27, 29, 34, 41, 43, 46, 51, 58, 61, 63, 66, 67, 69, 70, 83, 91, 94, 108, 109, and 139.

[49] Pereira FAC, Toufen DL, Guimarães-Filho ZO, Caldas IL, Viana RL and Gentle KW. *Coexistence of Turbulence Regimes in the Texas Helimak.* Physics of Plasmas **28**, 032301 (2021). Cited on pages 27, 51, 58, 66, 69, and 70.

[50] Garcia OE. *Stochastic Modeling of Intermittent Scrape-Off Layer Plasma Fluctuations.* Physical Review Letters **108**, 265001 (2012). Cited on pages 27, 28, 29, 30, 31, 32, 33, 51, 63, 66, 67, 68, 70, 71, 72, 87, 96, and 108.

[51] Garcia OE, Kube R, Theodorsen A and Pécseli HL. *Stochastic modelling of intermittent fluctuations in the scrape-off layer: Correlations, distributions, level crossings, and moment estimation.* Physics of Plasmas **23**, 052308 (2016). Cited on pages 27, 33, 66, 67, 70, and 71.

[52] Theodorsen A and Garcia OE. *Level crossings, excess times, and transient plasma–wall interactions in fusion plasmas.* Physics of Plasmas **23**, 040702 (2016). Cited on pages 27, 33, 66, 70, and 96.

[53] Garcia OE and Theodorsen A. *Auto-correlation function and frequency spectrum due to a super-position of uncorrelated exponential pulses.* Physics of Plasmas **24**, 032309 (2017). Cited on pages 27, 41, 42, 66, 67, 72, 73, and 123.

[54] Theodorsen A, Garcia OE and Rypdal M. *Statistical properties of a filtered Poisson process with additive random noise: distributions, correlations and moment estimation.* Physica Scripta **92**, 054002 (2017). Cited on pages 27, 29, 31, 33, 34, 35, 42, 44, 46, 66, 69, 70, 71, 72, 73, 78, 86, 88, 89, 91, 108, and 124.

[55] Theodorsen A and Garcia OE. *Probability distribution functions for intermittent scrape-off layer plasma fluctuations.* Plasma Physics and Controlled Fusion **60**, 034006 (2018). Cited on pages 27, 28, 29, 32, 33, 34, 36, 37, 66, 69, 70, 72, 76, 96, and 141.

[56] Militello F and Omotani JT. *Scrape off layer profiles interpreted with filament dynamics.* Nuclear Fusion **56**, 104004 (2016). Cited on pages 28 and 66.

[57] Militello F, Dudson B, Easy L, Kirk A and Naylor P. *On the interaction of scrape off layer filaments.* Plasma Physics and Controlled Fusion **59**, 125013 (2016). Cited on pages 28 and 66.

[58] Walkden NR, Wynn A, Militello F, Lipschultz B, Matthews G, Guillemaut C, Harrison J, Moulton D and JET Contributors. *Interpretation of scrape-off layer profile evolution and first-wall ion flux statistics on JET using a stochastic framework based on filamentary motion.* Plasma Physics and Controlled Fusion **59**, 085009 (2017). Cited on page 28.

[59] Maqueda RJ, Stotler DP, Zweben SJ and the NSTX team. *Intermittency in the scrape-off layer of the National Spherical Torus Experiment during H-mode confinement.* Journal of Nuclear Materials **415** S459–S462 (2011). Cited on page 28.

[60] Maqueda RJ and Wurden GA. *Edge turbulence measurements in NSTX by gas puff imaging.* Review of Scientific Instruments **72**, 931 (2001). Cited on page 28.

[61] Grulke O, Terry JL, LaBombard B and Zweben SJ. *Radially propagating fluctuation structures in the scrape-off layer of Alcator C-Mod.* Physics of Plasmas **13**, 012306 (2006). Cited on page 28.

[62] Theodorsen A. *Statistical properties of intermittent fluctuations in the boundary of fusion plasmas.* Doctoral thesis, Arctic University of Norway (2018). Cited on page 28.

[63] Krasheninnikov SI. *On the origin of plasma density blobs.* Physics Letters A vol. 380, p. 3905–3907 (2016). Cited on page 28.

[64] Fasoli A, Furno I and Ricci P. *The role of basic plasmas studies in the quest for fusion power.* Nature Physics **15**, 872–875 (2019). Cited on page 29.

[65] Krasheninnikov SI, D'Ippolito DA and Myra JR. *Recent theoretical progress in understanding coherent structures in edge and SOL turbulence.* Journal of Plasma Physics, vol. 74, part 5, pp. 679–717 (2008). Cited on pages 29 and 108.

[66] Stark H and Woods JW. *Probability, Statistics, and Random Processes for Engineers.* Pearson, 4th Ed. (2012). Cited on pages 30, 31, and 38.

[67] Miller S and Childers D. *Probability and random processes*. Academic Press, 2nd Ed. (2012). Cited on pages 30, 31, 32, 34, 35, 38, and 40.

[68] Westfall PH. *Kurtosis as Peakedness, 1905–2014. R.I.P.* The American Statistician **68:3**, 191-195 (2014). Cited on page 31.

[69] Sattin F *et al. On the statistics of edge fluctuations: comparative study between various fusion devices.* Plasma Physics and Controlled Fusion 51, 055013 (2009). Cited on pages 31, 33, 63, and 87.

[70] Pearson K. *IX. Mathematical contributions to the theory of evolution.—XIX. Second supplement to a memoir on skew variation.* Philosophical Transactions of the Royal Society of London. Series A, 216: p. 429–457 (1916). Cited on page 33.

[71] Krommes JA. *The remarkable similarity between the scaling of kurtosis with squared skewness for TORPEX density fluctuations and sea-surface temperature fluctuations.* Physics of Plasmas 15, 030703 (2008). Cited on pages 33, 63, and 87.

[72] Bergsaker AS, Fredriksen Å, Pécseli HL and Trulsen JK. *Models for the probability densities of the turbulent plasma flux in magnetized plasmas.* Physica Scripta **90**, 108005 (2015). Cited on page 33.

[73] Banerjee S, Zushi H, Nishino N, Hanada K, Sharma SK, Honma H, Tashima S, Inoue T, Nakamura K, Idei H, Hasegawa M and Fujisawa A. *Statistical features of coherent structures at increasing magnetic field pitch investigated using fast imaging in QUEST.* Nuclear Fusion **52**, 123016 (2012). Cited on page 33.

[74] Muller KE. *Computing the confluent hypergeometric function, $M(a, b; x)$.* Numerische Mathematik **90**, p. 179-196 (2001). Cited on pages 35 and 71.

[75] Gut A. *Probability: A Graduate Course*. Springer, 4th edition (2013). Cited on pages 35, 36, 37, and 63.

[76] Tran KC. *Estimating mixtures of normal distributions via empirical characteristic function.* Econometric Reviews, **17:2**, p. 167-183 (1998). Cited on pages 36 and 37.

[77] Feuerverger A and Mureika RA. *The Empirical Characteristic Function and its applications.* Annals of Statistics **vol. 5**, p. 88 (1977). Cited on pages 37 and 76.

[78] Briggs WL, Henson VE. *The DFT: An Owner's Manual for the Discrete Fourier Transform.* Siam (1995). Cited on pages 39, 40, 41, and 120.

[79] Kammler DW. *A First Course in Fourier Analysis.* Cambridge University Press, 2nd Ed. (2008). Cited on page 39.

[80] Zhu Z, White AE, Carter TA, Baek SG and Terry JL. *Chaotic edge density fluctuations in the Alcator C-Mod tokamak.* Physics of Plasmas **24**, 042301 (2017). Cited on pages 41, 44, 50, 51, 89, 107, and 109.

[81] Smith DE, Powers EJ and Caldwell GS. *Fast-Fourier-Transform spectral-analysis techniques as a plasma fluctuation diagnostic tool.* IEEE Transactions on Plasma Science, **vol. 2** no. 4, p. 261-272 (1974). Cited on pages 41 and 120.

[82] Welch P. *The Use of Fast Fourier Transformfor the Estimation of Power Spectra: A Method Based on Time Aver. aging Over Short, Modified Periodograms.* IEEE Transactions on audio and electroacoustics. **vol. AU-15** no. 2, p. 70-73 (1967). Cited on pages 41, 79, 120, and 121.

[83] Podder P, Khan TZ, Khan MH, and Rahman MM. *Comparative Performance Analysis of Hamming, Hanning and Blackman Window.* International Journal of Computer Applications (0975–8887) vol. 96–No.18 (2014). Cited on pages 41 and 121.

[84] Gubner JA. *Probability and Random Processes for Electrical and Computer Engineers.* Cambridge University Press, 1st Ed. (2006). Chap. 10. Cited on pages 41 and 42.

[85] Fernandes T. *MHD instabilities in TCABR tokamak.* Master's dissertation, Physics Institute of the University of São Paulo. São Paulo (2016). Retrieved from teses.usp.br. Cited on pages 41 and 42.

[86] Basse NP, Zoletnik S, Antar GY, Baldzuhn J, Werner A, and the W7-AS Team. *Characterization of turbulence in L- and ELM-free H-mode Wendelstein 7-AS plasmas.* Plasma Physics and Controlled Fusion, **vol. 45**, no. 4, p. 439 (2003). Cited on page 42.

[87] Kasdin NJ. *Discrete Simulation of Colored Noise and Stochastic Processes and* $1/f^{\alpha}$ *Power Law Noise Generation.* Proceedings of the IEEE – volume: 83; issue: 5; pages 802-827 (1995). Cited on pages 42 and 46.

[88] Kłosek-Dygas MM , BJ Matkowsky BJ, Schuss Z. *Colored noise in dynamical systems.* SIAM Journal on Applied Mathematics **48**(2), 425–441 (1988). Cited on pages 42 and 46.

[89] Gillespie DT. *Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral.* Physical Review E vol. 54, p. 2084 (1996). Cited on pages 43, 44, and 45.

[90] Schuss Z. *Theory and Applications of Stochastic Processes: An Analytical Approach.* Springer, 1st edition (2010). Chap 1. Cited on pages 43 and 44.

[91] Uhlenbeck GE and Ornstein LS. *On the theory of the Brownian motion.* Physical Review vol. 36 p. 823 (1930). Cited on pages 43 and 44.

[92] Bandt C and Pompe B. *Permutation Entropy: A Natural Complexity Measure for Time Series.* Physical Review Letters **88**, 174102 (2002). Cited on page 46.

[93] Rosso OA, Larrondo HA, Martin MT, Plastino A and Fuentes MA. *Distinguishing Noise from Chaos.* Physical Review Letters **99**, 154102 (2007). Cited on pages 46, 48, 49, and 89.

[94] Crepaldi C and Guimarães-Filho ZO. *Dynamical nature of plasma density fluctuations in TCABR tokamak.* To be submitted (2022). Cited on pages 46 and 89.

[95] Crepaldi C. *Busca por evidências de dinâmica caótica nas flutuações turbulentas de densidade em Plasmas.* Master's dissertation, Physics Institute of the University of São Paulo. São Paulo (2021). To be added in teses.usp.br. Cited on pages 46, 49, and 89.

[96] Calbet X and López-Ruiz R. *Tendency towards maximum complexity in a nonequilibrium isolated system.* Physical Review E **63** 066116 (2001). Cited on pages 48 and 49.

[97] Zunino L, Pérez DG, Martín MT, Plastino A, Garavaglia M, and Rosso OA. *Characterization of Gaussian self-similar stochastic processes using wavelet-based informational tools.* Physical Review E **75**, 021115 (2007). Cited on page 49.

[98] Eckmann JP, Oliffson Kamphorst S and Ruelle D. *Recurrence Plots of Dynamical Systems.* Europhysics Letters **4**973 (1987). Cited on pages 50 and 88.

[99] Guimarães-Filho ZO, Caldas IL, Viana RL, Kurths J, Nascimento IC and Kuznetsov YK. *Recurrence quantification analysis of electrostatic fluctuations in fusion plasmas.* Physics Letters A **vol. 372**, issue 7, p. 1088-1095 (2008). Cited on pages 12, 50, 51, 53, 54, 88, 89, and 135.

[100] Guimarães-Filho ZO, Caldas IL, Viana RL, Nascimento IC, Kuznetsov YU and Kurths J. *Recurrence quantification analysis of turbulent fluctuations in the plasma edge of Tokamak Chauffage Alfvén Brésilien tokamak.* Physics of Plasmas **17**, 012303 (2010). Cited on pages 12, 50, 54, 89, and 135.

[101] Guimarães-Filho ZO, Santos Lima GZ, Caldas IL, Viana RL, Nascimento IC, and Kuznetsov YK. *Characterizing electrostatic turbulence in tokamak plasmas with high MHD activity.* Journal of Physics: Conference Series **246**, 012014 (2010). Cited on pages 12, 50, 51, 53, 54, 89, 135, and 136.

[102] Viana RL, Toufen DL, Guimarães-Filho ZO, Caldas IL, Gentle KW and Nascimento IC. *Recurrence Analysis of Turbulent Fluctuations in Magnetically Confined Plasmas.*

In: Webber Jr C, Ioana C, Marwan N (eds) *Recurrence Plots and Their Quantifications: Expanding Horizons*. Springer Proceedings in Physics, vol. 180 (2016). Cited on pages 12, 50, 51, 54, 89, and 135.

[103] Toufen DL, Guimarães-Filho ZO, Caldas IL, Szezech JD, Lopes S, Viana RL and Gentle KW. *Analysis of the influence of external biasing on Texas Helimak turbulence.* Physics of Plasmas **20**, 022310 (2013). Cited on pages 50 and 135.

[104] Webber CL and Zbilut JP. *Dynamical assessment of physiological systems and states using recurrence plot strategies.* Journal of Applied Physiology **76**965 (1994). Cited on pages 50 and 53.

[105] Kantz H and Schreiber T. *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, 2nd Ed. (2003). Chap. 3. Cited on pages 51 and 53.

[106] Mohr PJ, Taylor BN. *CODATA recommended values of the fundamental physical constants: 1998*. Rev. Mod. Phys. 72 (2) 351 (2000). Cited on page 55.

[107] Bevington PR, Robinson DK. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, 3rd Ed. (2002). Chap. 8. Cited on page 55.

[108] Taylor, JR. *An Introduction to Error Analysis*. University Science Books, 2nd Ed. (1997). Pages 268–274. Cited on page 55.

[109] NIST. *e-Handbook of Statistical Methods*. Chi-square distribution (2003). Cited on page 55.

[110] Grenfell GG. *Estudo de Efeitos da Polarização Eletrostática Periférica no Tokamak TCABR*. Master's dissertation, Physics Institute of the University of São Paulo. São Paulo (2016). Retrieved from teses.usp.br. Cited on pages 56, 58, and 62.

[111] Desideri D and Serianni G. *Four parameter data fit for Langmuir probes with nonsaturation of ion current*. Review of Scientific Instruments **69**, 2354 (1998). Cited on page 56.

[112] Stangeby PC. *The Plasma Boundary of Magnetic Fusion Devices*. IOP Publishing (2000). Cited on pages 57 and 58.

[113] Huld T, Nielsen HA, Pecseli HL, and Juul Rasmussen J. *Coherent structures in two-dimensional plasma turbulence*. Physics of Fluids B: Plasma Physics **3**, 1609 (1991). Cited on page 59.

[114] Pedrosa MA *et al. Empirical Similarity of Frequency Spectra of the Edge-Plasma Fluctuations in Toroidal Magnetic-Confinement Systems.* Physical Review Letters vol. 82, n. 18, p. 3621 (1999). Cited on page 63.

[115] Toufen DL, Pereira FAC, Guimarães-Filho ZO, Caldas IL, Gentle KW. *Plasma Turbulence driven by density gradient in Texas Helimak.* (To be published.) Cited on page 63.

[116] Garcia OE and Theodorsen A. *Power law spectra and intermittent fluctuations due to uncorrelated Lorentzian pulses.* Physics of Plasmas **24**, 020704 (2017). Cited on pages 66 and 70.

[117] Campbell N. *The study of discontinuous phenomena.* Proceedings of the Cambridge Philosophical Society, Mathematical and physical sciences **15**, p. 117–137 (1909). Cited on page 66.

[118] Rice SO. *Mathematical Analysis of Random Noise.* Bell System Technical Journal **23**, 282 (1944). Cited on page 66.

[119] Pécseli HL and Trulsen JK. *On the interpretation of experimental methods for investigating nonlinear wave phenomena.* Plasma Physics and Controlled Fusion **35**, 1701 (1993). Cited on page 66.

[120] Toufen DL, Pereira FAC, Guimarães-Filho ZO, Caldas IL, and Gentle KW. *Electrostatic turbulence intermittence driven by biasing in Texas Helimak.* Physics of Plasmas **21**, 122302 (2014).

[121] Ding M. Scott A (ed.). *Encyclopedia of Nonlinear Science.* Routledge, 1st Ed. (2004). "Intermittency", p. 463. Cited on page 71.

[122] Rockwell RC. *Assessment of Multicollinearity: The Haitovsky Test of the Determinant.* Sociological Methods & Research **3**: 308 (1975). Cited on page 76.

[123] Olver FWJ, Lozier DW, Boisvert RF and Clark CW (editors). *NIST Handbook of Mathematical Functions.* Cambridge University Press, Cambridge, 1st Ed. (2010). p. 611. Cited on page 142.

Appendix A

Methods details

A.1 Windowing the power spectrum

Sampling the PSD of a signal (simulated or experimental) is subjected to a phenomenon called spectral leakage, which distorts the power spectrum (see pp. 98-99 of Ref. 78, or Ref. 81). To suppress this leakage, the signal is often multiplied by a function that gradually falls to zero at the edges [81]. The sampled PSD can also be very noisy, and thus it is interesting to apply smoothing procedures, dividing the signal into segments and after taking their average [81, 82]. These procedures to decrease the leakage and the signal noise can be unified, as described by the windowing method of Welch [78, 82].

In Welch's method, the original signal X with length N is divided into K overlapping segments, as seen in Fig. A.1. Each segment has L points and is separated from its neighbors by D points. Usually, $D = L/2$ is chosen, meaning that the overlap is 50%. If $D = L$, then the overlap is 0%.

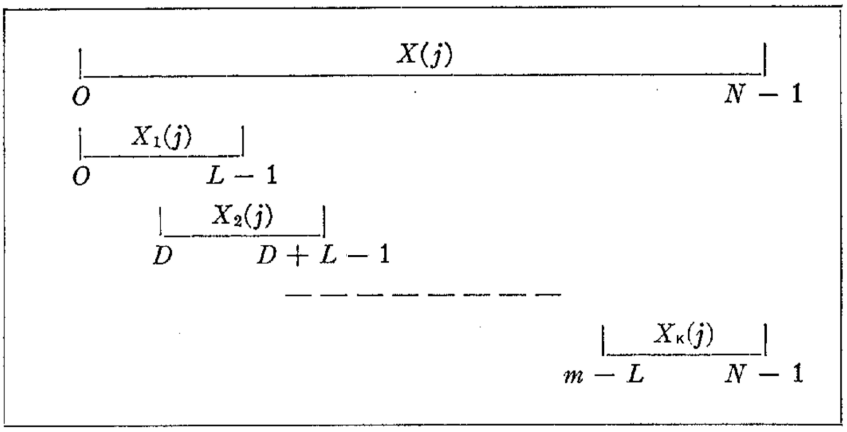


Figure A.1: Illustration of data segmentation in Welch's method. Source: Fig. 1 of Ref. 82 (with $D + L + 1$ corrected to $D + L - 1$ in the 2nd segment).

After the splitting, the segments are windowed. That is, in the time domain, the segments X_k $(k = 1, 2, ..., K)$ are multiplied by a window function W. Widespread used windows are

the Hann, Hamming and Blackman [83]. In this thesis, the Hamming window was used,

$$W(j) = \frac{25}{46} - \left(1 - \frac{25}{46}\right) \cos\left(\frac{2\pi j}{L-1}\right), \quad 0 \leq j \leq L-1$$

One can see that this function W afford more influence on the center of the interval than on the edges. To prevent loss of information and to decrease statistical fluctuations, Welch's method uses a segment overlap of 50% [82].

After windowing X_k, a discrete Fourier transform is applied in the corresponding functions $X_k W$:

$$A_k(n) = \frac{1}{L} \sum_{j=0}^{L-1} X_k(j) W(j) e^{-2\pi i j n / L}, \quad n = 0, 1, ..., L-1$$

The squared magnitude S_k is then computed:

$$S_k = \frac{L}{U} |A_k|^2$$

where

$$U = \frac{1}{L} \sum_{j=0}^{L-1} W^2(j)$$

Finally, the estimated PSD of the signal X is given by the mean of the S_k, where $k = 1, 2, ..., K$. With this average, the PSD is smoothed, generating a graph with lower statistical fluctuation and less leakage.

In resume, the steps of Welch's method are:

1. Divide the signal X into K segments with length L and distance D between the beginning of neighboring segments;
2. Multiply each segment X_k by a window function W;
3. Compute the discrete Fourier transform of each modified segment $X_k W$;
4. Compute the PSD S_k of each segment;
5. Evaluate the average of all S_k, generating a smoothed power spectrum.

Fig. A.2 illustrates the effect of windowing a signal with the Hamming function. The leakage effect is reduced.

Welch's method is available in several programming languages, such as Python™ (*welch* function of the scipy.signal library) and MATLAB® (*pwelch* function). In this dissertation, the MATLAB version was used, with a normalization factor of $\pi \Delta t$ to match theoretical spectra.

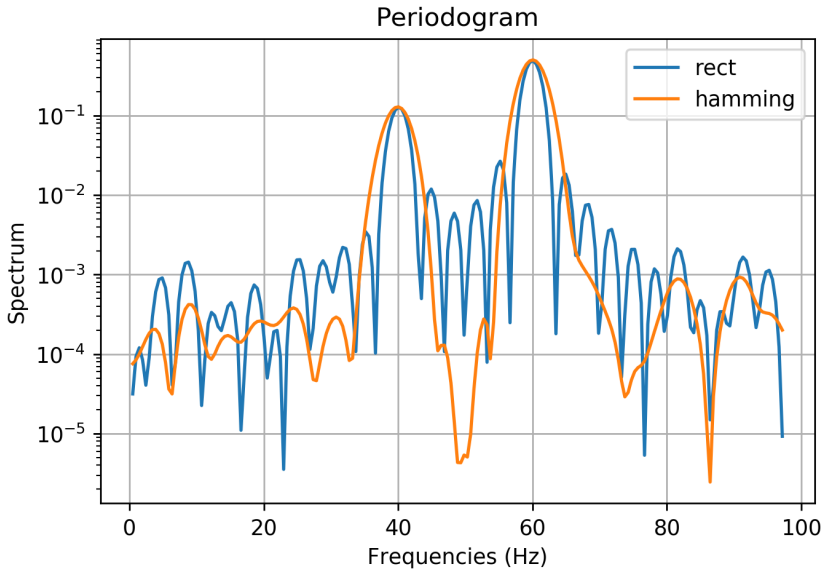


Figure A.2: Periodograms (i.e., PSD estimates) of a signal, with the use of a rectangular window ($W = 1$), in blue, and with a Hamming window, in orange. The Hamming window reduces the effect of leakage, which distorts the PSD estimation. Figure taken from periodogram article in Wikipedia.

Appendix B

Details of the stochastic pulse train model

B.1 PSD with correlated noise

B.1.1 Colored noise case

In this appendix, some derivations of the power spectrum are presented, for the stochastic pulse train model with colored noise. Let I_{sat} be given by the stochastic pulse train model with unstretched waveform and white Gaussian noise (see Chapter 4 for details),

$$I_{sat} = I_B(t) + I_{\mathcal{N}}(t),$$

$$I_B(t) = \sum_{j=1}^{N_P} A_j \phi\left(\frac{t - t_j}{\tau_d}; \lambda\right),$$

$$\Omega_{\mathcal{N}} = \sigma_{\mathcal{N}}^2 \Delta t$$

Δt is the time interval between measurements, N_P is the number of pulses, ϕ their shape, τ_d their characteristic duration and λ the asymmetry parameter. A_j and t_j are respectively the amplitude and time occurrence of the j-th pulse. The normalized I_{sat} is

$$\tilde{I}_{sat} = \frac{I_{sat} - \langle I_{sat} \rangle}{\sigma_I}$$

According to Eq. (C8) of Ref. 53, the power spectrum of this $\tilde{I}_{sat}$ is

$$\Omega_{\tilde{I}}(\omega) = \frac{2\tau_d}{1+\epsilon}\left\{\left(1 + \lambda^2\tau_{\mathrm{d}}^2\omega^2\right)^{-1}\left[1 + (1-\lambda)^2\tau_{\mathrm{d}}^2\omega^2\right]^{-1} + \frac{\epsilon\Delta t}{2\tau_d}\right\}$$

where ϵ is the ratio of the noise and burst variances (Eq. (C2) of Ref. 53),

$$\epsilon \equiv \frac{\sigma_{\mathcal{N}}^2}{\sigma_B^2}$$

and ω is the angular frequency, $\omega = 2\pi f$. Thus

$$1 + \epsilon = \frac{\sigma_B^2 + \sigma_{\mathcal{N}}^2}{\sigma_B^2} = \frac{\sigma_I^2}{\sigma_B^2}$$

and, using $\lambda\tau_{\rm d} = \tau_r$ and $(1-\lambda)\tau_{\rm d} = \tau_f$,

$$\Omega_{\tilde{I}}(\omega) = 2\tau_d \frac{\sigma_B^2}{\sigma_I^2}\left(1+\tau_r^2\omega^2\right)^{-1}\left(1+\tau_f^2\omega^2\right)^{-1} + \frac{\sigma_{\mathcal{N}}^2}{\sigma_I^2}\Delta t$$

An interesting question is what happens when the color of the noise is changed from white ($\Omega_{\mathcal{N}} = \text{cte}$) to another color type ($\Omega_{\mathcal{N}} \propto 1/f^{\eta}$). Following Section 2.6, an approximate answer for the PSD of the MATLAB colored Gaussian noise generator is

$$\Omega_{\mathcal{N}} \simeq \Delta t \sigma_{\mathcal{N}}^2 \frac{C(\eta)}{\omega^{\eta}}$$

where

$$C(\eta) \simeq 10^{5.73\cdot\eta^{0.820}}$$

Hence the PSD of the normalized saturation current is approximately

$$\Omega_{\tilde{I}}(\omega) \simeq 2\tau_d \frac{\sigma_B^2}{\sigma_I^2}\left(1+\tau_r^2\omega^2\right)^{-1}\left(1+\tau_f^2\omega^2\right)^{-1} + \Delta t \frac{\sigma_{\mathcal{N}}^2}{\sigma_I^2}\frac{C(\eta)}{\omega^{\eta}} \tag{B.1}$$

Furthermore, the normalized PSD formula is (Eq. (17) of [54]):

$$\Omega_{\tilde{I}}(\omega) = \frac{\Omega_I(\omega) - 2\pi\left\langle I_{sat}\right\rangle^2\delta(\omega)}{\sigma_I^2}$$

where $\delta(\omega)$ is the Dirac distribution and $\sigma_I^2 = \sigma_B^2 + \sigma_{\mathcal{N}}^2$ is the total variance. Thus, for the non-normalized PSD,

$$\Omega_I(\omega) = \sigma_I^2\Omega_{\tilde{I}}(\omega) + 2\pi\left\langle I_{sat}\right\rangle^2\delta\left(\omega\right)$$

and hence the stochastic pulse train model with colored noise has the power spectrum

$$\Omega_I(\omega) \simeq 2\tau_d\sigma_B^2\left(1+\tau_r^2\omega^2\right)^{-1}\left(1+\tau_f^2\omega^2\right)^{-1} + \Delta t\sigma_{\mathcal{N}}^2\frac{C(\eta)}{\omega^{\eta}} + 2\pi\left\langle I_{sat}\right\rangle^2\delta\left(\omega\right) \tag{B.2}$$

The CN part is approximated. While $\sigma_{\mathcal{N}}^2$ is a model parameter, the burst variance is given by Eq. (4.11):

$$\sigma_B^2 = \frac{N_P}{T}\tau_d\left\langle A\right\rangle^2$$

Fig. B.1(a) is depicts a fit of the power spectrum of the SPTM with colored noise. The data is in blue and corresponds to a mean of 10 SPTM simulations, while the curve in purple is the fit from Eq. (B.1). It is clear that the PSD has a flat behavior for low frequencies and has a power law at high frequencies. The model parameters for the bursts were $\tau_r = 1.0$ µs, $\tau_f = 2.5$ µs, $N_P/T = 200\ \text{ms}^{-1}$, $\langle A\rangle = 10$ mA; for the colored noise, $\langle I_{\mathcal{N}}\rangle = 0$ $\sigma_{\mathcal{N}} = 1$ mA and $\eta = 1$. The simulation of Fig. B.1(b) has exactly the same inputs, except for $\sigma_{\mathcal{N}} = 3$ mA. The increase of the noise standard deviation shows that the model is not perfect for the colored noise part, as the new fit is not adequate for low frequencies.

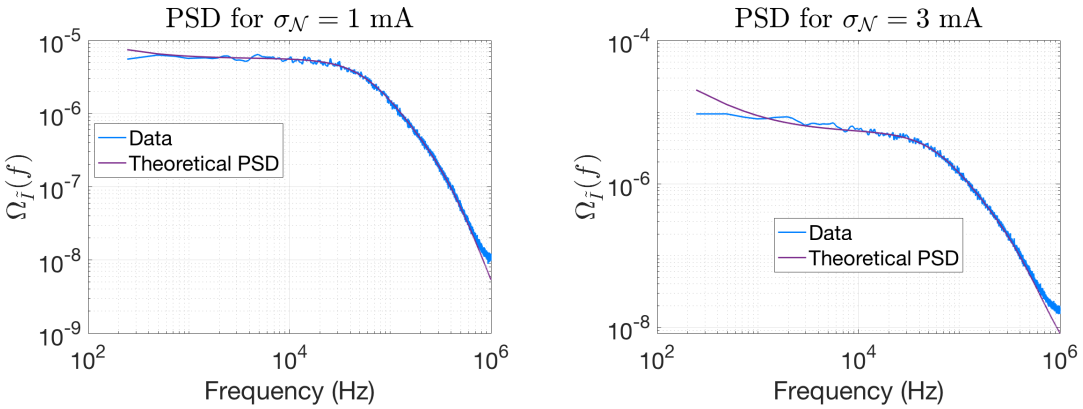


Figure B.1: **(a)** Example of a simulated PSD (in blue) and its fit (in purple) with the pulse background model. **(b)** Same simulation but with $\sigma_{\mathcal{N}} = 3$ mA instead of $\sigma_{\mathcal{N}} = 1$ mA.

B.1.2 Ornstein-Uhlenbeck noise case

As seen in the previous subsection, the power spectrum due to the burst pulses is

$$\Omega_B(\omega) = \frac{2\tau_d\sigma_B^2}{(1+\tau_r^2\omega^2)\left(1+\tau_f^2\omega^2\right)} + 2\pi \left\langle I_B\right\rangle^2 \delta\left(\omega\right)$$

Furthermore, the PSD of the Ornstein-Uhlenbeck noise (OUN) is (Eq. (2.29)),

$$\Omega_{\mathcal{N}}(\omega) = \frac{2\tau_{\mathcal{N}}\sigma_{\mathcal{N}}^2}{1+\tau_{\mathcal{N}}^2\omega^2} + 2\pi\delta(\omega)\left\langle\Phi_{\mathcal{N}}\right\rangle^2$$

Hence, using the expression for the power spectrum of the sum (Eq. (2.22)), results

$$\Omega_I(\omega) = \Omega_B(\omega) + \Omega_{\mathcal{N}}(\omega) + 4\pi\delta(\omega)\left\langle I_B\right\rangle\left\langle I_{\mathcal{N}}\right\rangle$$

Thus the PSD of the SPTM with OUN ($I_{sat} = I_B + I_{\mathcal{N}}$) is

$$\Omega_I(\omega) = \frac{2\tau_d\sigma_B^2}{(1+\tau_r^2\omega^2)\left(1+\tau_f^2\omega^2\right)} + \frac{2\tau_{\mathcal{N}}\sigma_{\mathcal{N}}^2}{1+\tau_{\mathcal{N}}^2\omega^2} + 2\pi\delta\left(\omega\right)\left\langle I_{sat}\right\rangle^2$$

where I used

$$\left\langle I_B\right\rangle^2 + \left\langle I_{\mathcal{N}}\right\rangle^2 + 2\left\langle I_B\right\rangle\left\langle I_{\mathcal{N}}\right\rangle = \left(\left\langle I_B\right\rangle + \left\langle I_{\mathcal{N}}\right\rangle\right)^2 = \left\langle I_{sat}\right\rangle^2$$

B.2 Least-squares fit of the conditionally averaged burst

The dark blue data in Fig. B.2 corresponds to the values that would be obtained for the bursts parameters τ_d and λ using a least-squares fitting of the conditionally averaged burst. The results are different from the ones obtained with the CF-CAB-PSD fit (Chapter 5 and

Fig. 6.3(b)). Partially because of the pulse overlapping, the conditionally averaged burst appears to be, in general, longer and less asymmetric than the original pulses. This indicates that the conditional averaging analysis may distort the original pulse shape. If the rising duration $\tau_r = \lambda\tau_d$ is close to the temporal resolution of measurements, $\Delta t = 0.5$ µs, then the pulse shape in the CAB may also be distorted.

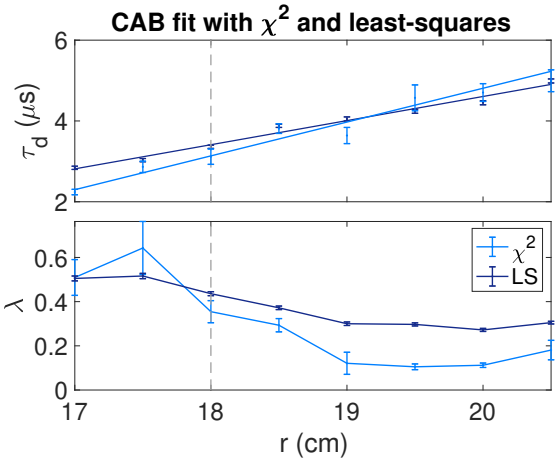


Figure B.2: Radial profile from the parameters fitted with the conditionally averaged burst (CAB). In light blue are the results from the χ^2 fitting. In dark blue are the results from a least-squares (LS) fitting. The dashed line in $r = a = 18.0$ cm indicates the LCFS approximate position.

B.3 All fitted PDFs, CABs and PSDs

B.3.1 Stochastic model with noise

This section presents all the PDFs (Fig. B.3), CABs (Fig. B.4), and PSDs (Fig. B.5) fitted with the stochastic model with Ornstein-Uhlenbeck noise and the CF-CAB-PSDs method (Chapters 4 to 6), for the experiment 34132 of TCABR. The parameters used for the simulations are shown in Appendix B.4.1.

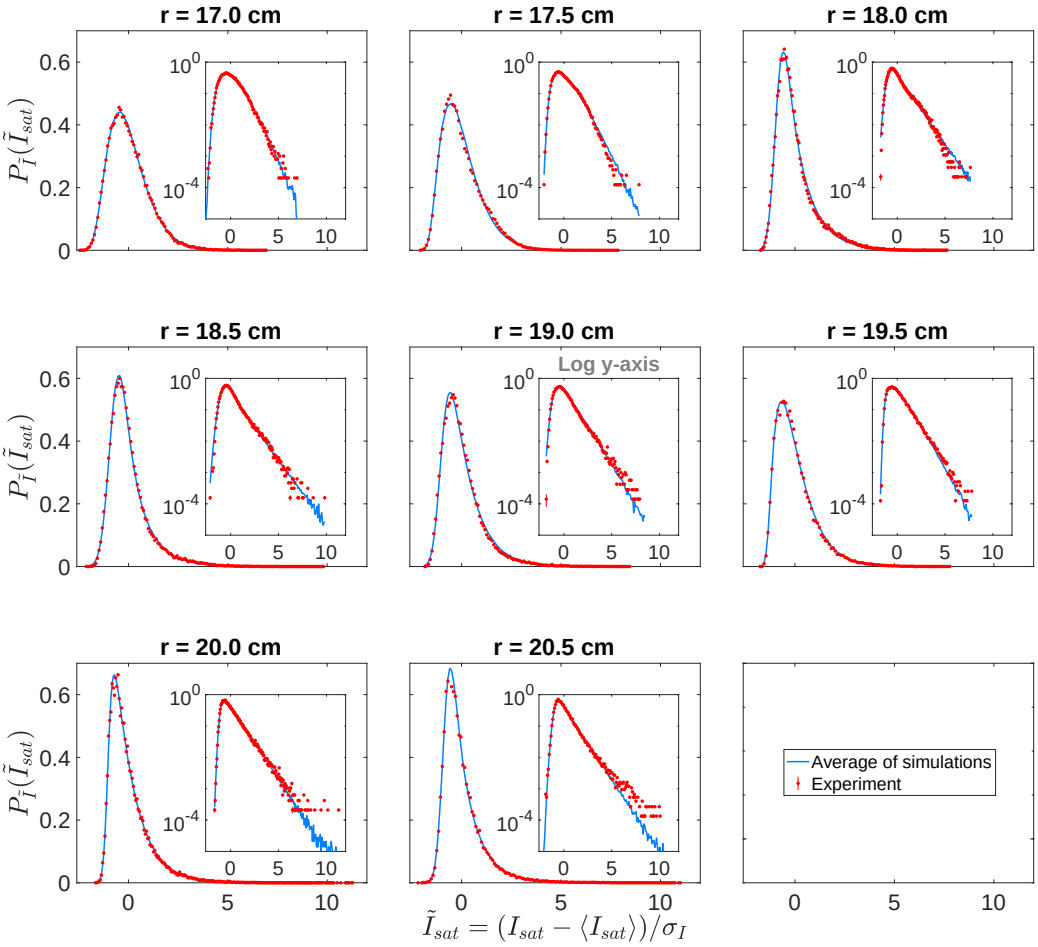


Figure B.3: In red, the experimental $\tilde{I}_{sat}$ distributions in linear and log-y scales. In blue, the corresponding simulated distributions, made with the stochastic model with OU noise (Chapters 4 to 6). The simulated PDFs in general are in great agreement with the experimental ones. When $r \in [18.0, 20.0, 20.5]$ cm, the fit is not ideal for high $\tilde{I}_{sat}$. This indicates a difference in the simulated and experimental kurtosis, as confirmed by the $K_I \times r$ panel of Fig. 6.4(a).

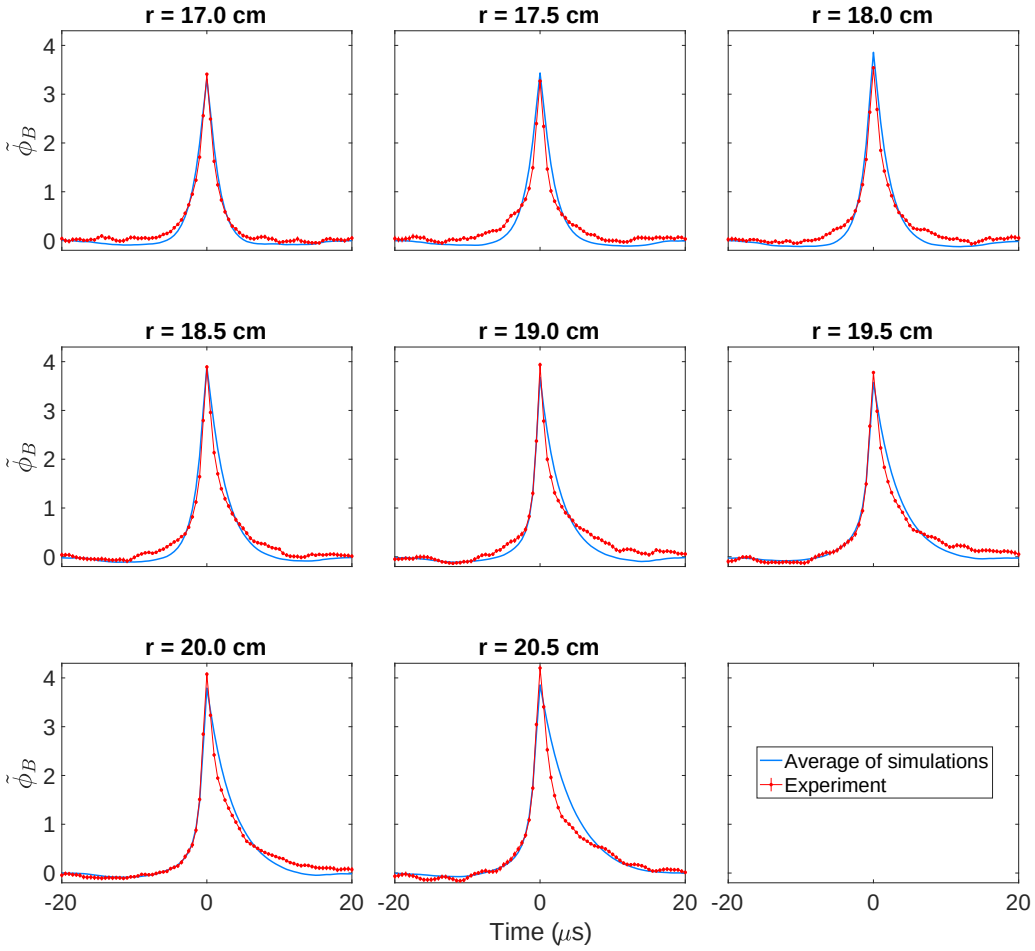


Figure B.4: In red, the conditionally averaged bursts (CABs) of the eight positions measured in the experiment. In blue, the corresponding simulated CABs, made with the stochastic model with Ornstein-Uhlenbeck noise (Chapters 4 to 6). Apart from $r = 17.0$ cm, the model recovered only approximations of the averaged waveforms. Nevertheless, for $r \in \{19.5\text{ cm}, 20.0\text{ cm}, 20.5\text{ cm}\}$, the CABs rise was well fitted.

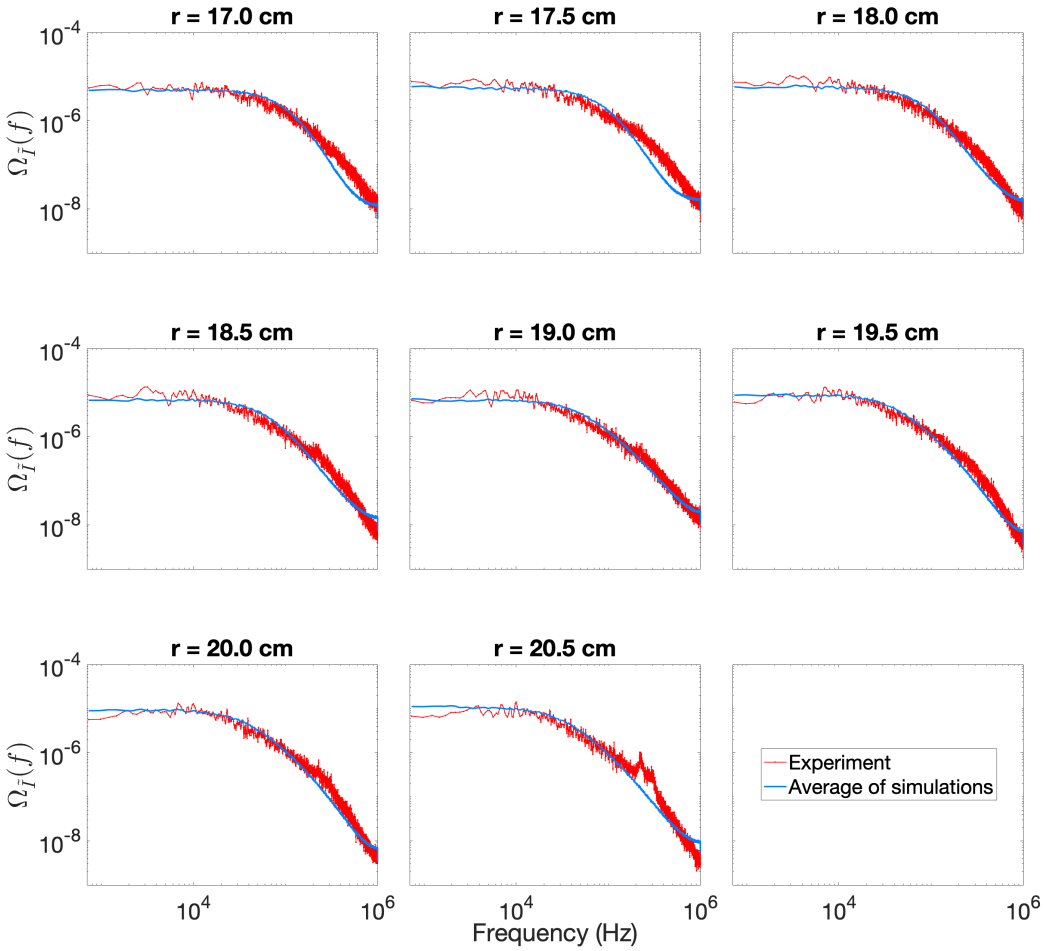


Figure B.5: In red, the power spectral densities (PSDs) of the eight positions measured in the experiment. In blue, the corresponding simulated PSDs, made with the stochastic model with Ornstein-Uhlenbeck noise (Chapters 4 to 6). The model recovered only approximations of the functions. For $r = 20.5$ cm the experimental spectrum had two peaks, one around $f = 0.22$ MHz and the other around $f = 0.31$ MHz (as described in Section 3.4).

B.3.2 Stochastic model with pulse background

This section presents all the PDFs (Fig. B.6), CABs (Fig. B.7), and PSDs (Fig. B.8) fitted manually with the stochastic model with stretched bursts and pulse background (Chapters 7 and 8), for the experiment 34132 of TCABR. In opposition to the case of unstretched bursts and Gaussian background (Subsection B.3.1), the present model fitted well the CABs and PSDs, but not the PDFs. The parameters used for the simulations are shown in Appendix B.4.2.

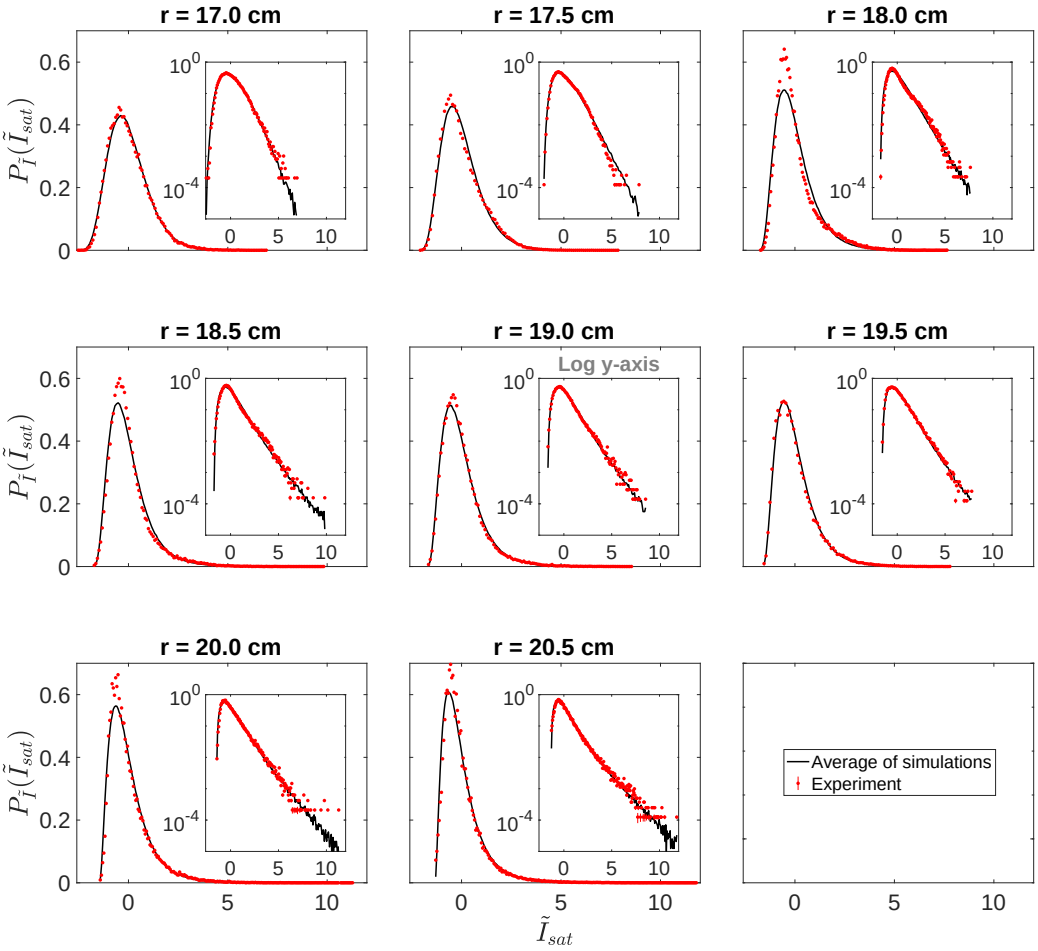


Figure B.6: In red, the experimental $\tilde{I}_{sat}$ distributions (in linear and log-y scales). In black, the corresponding simulated distributions, made with the stochastic model with stretched bursts and a pulse background (Chapters 7 and 8). In general, the fitted model couldn't fit the peak of the experimental PDFs, in opposition to Fig. B.3.

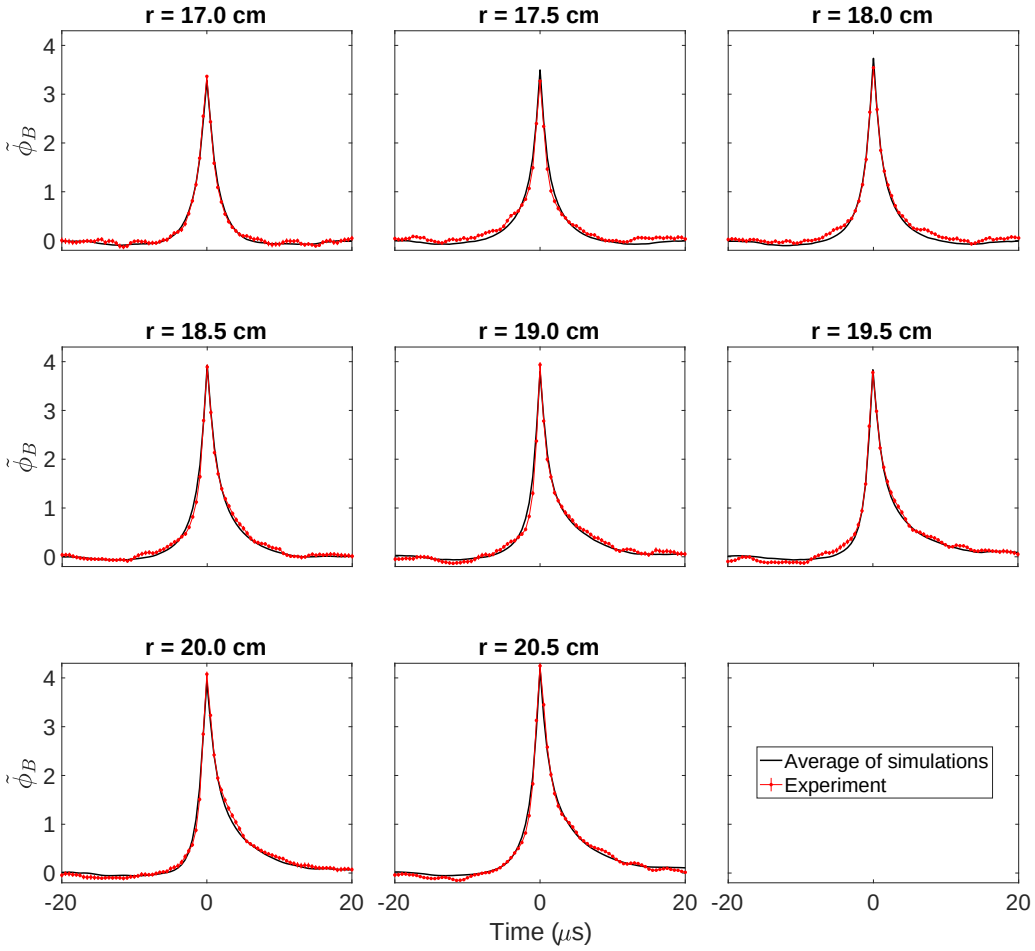


Figure B.7: In red, the conditionally averaged bursts (CABs) of the eight positions measured in the experiment. In black, the corresponding simulated CABs, made with the stochastic model with stretched bursts and a pulse background (Chapters 4 to 6). Because of the stretched bursts ($\zeta_r, \zeta_f \neq 1$), the model fitted really well the experimental waveforms, in opposition to Fig. B.4.

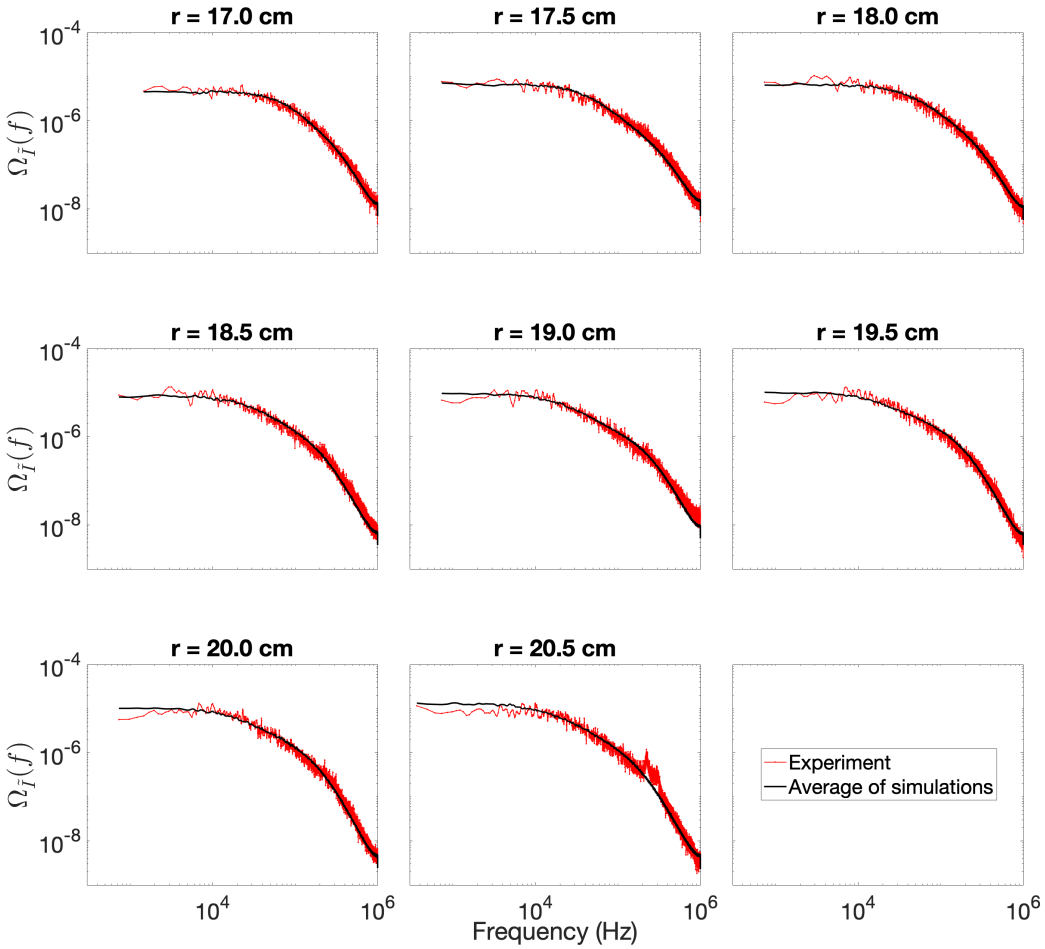


Figure B.8: In red, the power spectral densities (PSDs) of the eight positions measured in the experiment. In black, the corresponding simulated PSDs, made with the stochastic model with stretched bursts and a pulse background (Chapters 4 to 6). Contrary to Fig. B.5, the model fitted very well the experimental spectra.

B.4 Fitted parameters

B.4.1 Parameters for the SPTM with noise

Tables B.1 and B.2 show the parameters fitted for experiment 34132 of TCABR (Chapter 6), with the CF-CAB-PSD method (Chapter 5) and the stochastic pulse train model with noise (SPTM-N, Chapter 4). Both the Ornstein-Uhlenbeck noise and the colored noise cases are considered in Table B.2, respectively with the parameters $\tau_{\mathcal{N}}$ and η. For the meaning of each parameter, the reader is referred to Table 5.1.

Table B.3 shows the values and uncertainties for the relative mean and variance of the noise, $\iota = \langle I_{\mathcal{N}} \rangle / \langle I_B \rangle = \langle I_{\mathcal{N}} \rangle / (\gamma \langle A \rangle)$ and $\epsilon = \sigma^2_{\mathcal{N}} / \sigma^2_B = \sigma^2_{\mathcal{N}} / (\gamma \langle A \rangle^2)$. Also included is the number of pulses per time, $N_P/T = \gamma/\tau_d$. To estimate the uncertainties of ι, ϵ and N_P/T, one must beware of the correlations between the UP-OUN parameters. Table 5.3 gives a good

approximation for these correlations. To evaluate them more precisely, the same procedure of Section 5.5 can be employed. For each position r, 100 repeated fits were made to adjust synthetic signals with true values of Tables B.1 and B.2. The correlations from the SPTM-N parameters can then be estimated numerically from these sets of 100 fits. This procedure was used for example in Fig. 6.5, where the correlations between the intermittency-related parameters were evaluated for two positions.

Table B.1: Parameters fitted with the characteristic function for the model with noise (Chapters 4, 5 and 6). Uncertainties in parenthesis.

r (cm)	γ	$\langle A \rangle$ (mA)	$\langle I_{\mathcal{N}} \rangle$ (mA)	$\sigma_{\mathcal{N}}$ (mA)
17.0	4.84(26)	10.80(31)	12.7(14)	3.4(64)
17.5	2.92(13)	11.38(29)	10.63(70)	3.36(26)
18.0	0.886(32)	18.20(44)	15.88(25)	5.286(88)
18.5	0.724(35)	13.56(42)	12.13(24)	4.820(76)
19.0	1.340(62)	9.30(26)	6.90(26)	3.364(85)
19.5	1.956(82)	7.02(19)	2.39(24)	1.365(96)
20.0	1.129(42)	7.73(22)	3.45(12)	1.511(44)
20.5	0.86(41)	5.26(16)	2.32(10)	1.51(28)

Table B.2: Parameters obtained for the SPTM-N with the χ^2-map fits—λ and τ_d obtained with the conditionally averaged burst and $\tau_{\mathcal{N}}$ with the PSD (Chapters 4, 5 and 6). Also included is the exponent η for the colored noise case.

r (cm)	λ	τ_d (µs)	$\tau_{\mathcal{N}}$ (µs)	η
17.0	0.548(33)	2.550(63)	0.052(37)	0.21(11)
17.5	0.465(21)	2.890(75)	0.047(41)	0.018(19)
18.0	0.355(19)	3.116(76)	0.722(59)	0.757(28)
18.5	0.293(15)	3.811(71)	1.301(83)	0.977(18)
19.0	0.121(9)	3.64(15)	0.651(32)	0.712(32)
19.5	0.105(13)	4.57(18)	0.46(11)	0.543(92)
20.0	0.112(18)	4.70(20)	0.85(14)	0.815(41)
20.5	0.127(14)	5.75(17)	1.26(12)	1.063(17)

Table B.3: Values and uncertainties obtained for $\iota = \langle I_N \rangle / \langle I_B \rangle$, $\epsilon = \sigma_N^2/\sigma_B^2$ and $N_P/T = \gamma/\tau_d$, using the CF-CAB-PSD fit for the UP-OUN model.

r (cm)	ι	ϵ	N_P/T (1/µs)
17.0	24.3(34)%	2.05(68)%	1.90(13)
17.5	31.9(28)%	2.99(47)%	1.011(61)
18.0	98.6(31)%	9.55(41)%	0.284(13)
18.5	123.6(47)%	17.46(79)%	0.1899(86)
19.0	55.4(27)%	9.75(64)%	0.368(28)
19.5	17.4(19)%	1.92(28)%	0.429(28)
20.0	39.5(20)%	3.38(21)%	0.240(15)
20.5	51.4(32)%	9.57(50)%	0.1496(77)

B.4.2 Parameters for the SPTM with pulse background

Tables B.4 and B.5 show the parameters used in Chapter 8, fitted manually for the experiment 34132 of TCABR with the stochastic pulse train model with pulse background (SPTM-PB). In this model nine parameters are fitted: γ, $\langle A \rangle$, $\gamma^{(b)}$, $\langle A^{(b)} \rangle$, ζ_r, ζ_f, λ, τ_d and $\tau_d^{(b)}$. For the parameter meanings and names, the reader is referred to Table 7.1 (in which one can interchange N_P/T for $\gamma = \tau_d N_P/T$). The uncertainty of a parameter p was estimated as a percentage of its value, δp. For example, the uncertainty chosen for the intermittency parameter was $\delta\gamma = 4\%\gamma$ and, for the pulse average amplitude, $\delta \langle A \rangle = 1\% \langle A \rangle$.

The distribution of the SPTM-PB depends on γ, $\langle A \rangle$, $\gamma^{(b)}$, $\langle A^{(b)} \rangle$, ζ_r, ζ_f and λ. However, the dependence on the last three exists only through the stretch coefficients $s_n(\lambda, \zeta_r, \zeta_f)$, as described in Section 7.2. With this in mind, Table B.4 shows the values obtained for the main parameters of the distribution: γ, $\langle A \rangle$, $\gamma^{(b)}$, $\langle A^{(b)} \rangle$. On the other hand, Table B.5 stores the values for the pulse parameters ζ_r, ζ_f, λ, τ_d and $\tau_d^{(b)}$.

Table B.4: SPTM-PB parameters mostly related to the I_{sat} distribution, manually fitted for the TCABR experiment. Also included is the number of pulses per time, N_P/T.

r (cm)	γ	$\langle A \rangle$ (mA)	$\gamma^{(b)}$	$\langle A^{(b)} \rangle$ (mA)	N_P/T (1/ms)	$N_P^{(b)}/T$ (1/ms)
17.0	3.00(12)	11.20(11)	5.07(10)	6.50(13)	1000(30)	5070(99)
17.5	1.41(6)	15.00(15)	2.84(6)	5.70(11)	640(19)	3150(60)
18.0	0.705(28)	18.00(18)	2.86(6)	5.40(11)	300(9)	2600(50)
18.5	0.321(13)	18.00(18)	3.31(7)	5.26(9)	107.0(32)	2204(44)
19.0	0.323(12)	15.50(16)	2.78(6)	4.45(9)	85.0(26)	2174(44)
19.5	0.293(12)	14.00(14)	2.45(5)	4.16(8)	85.0(26)	1477(30)
20.0	0.328(13)	12.00(12)	1.974(39)	3.56(7)	80.0(24)	1234(25)
20.5	0.204(8)	9.80(10)	1.653(33)	2.28(5)	60.0(18)	972(20)

Table B.5: Pulse parameters of the SPTM-PB, manually fitted for the experiment.

r (cm)	ζ_r	ζ_f	λ	τ_d (µs)	$\tau_d^{(b)}$(µs)
17.0	1.000(20)	1.000(20)	0.500(10)	3.00(6)	1.000(30)
17.5	0.670(13)	0.670(13)	0.500(10)	2.20(5)	0.900(27)
18.0	0.750(15)	0.650(13)	0.426(8)	2.35(5)	1.100(33)
18.5	0.650(13)	0.650(13)	0.333(7)	3.00(6)	1.500(45)
19.0	0.650(13)	0.650(13)	0.2105(42)	3.80(8)	1.280(38)
19.5	0.740(15)	0.600(12)	0.1500(30)	3.50(7)	1.66(5)
20.0	1.000(20)	0.650(13)	0.2195(44)	4.10(8)	1.60(5)
20.5	0.700(14)	0.550(11)	0.294(6)	3.40(7)	1.70(5)

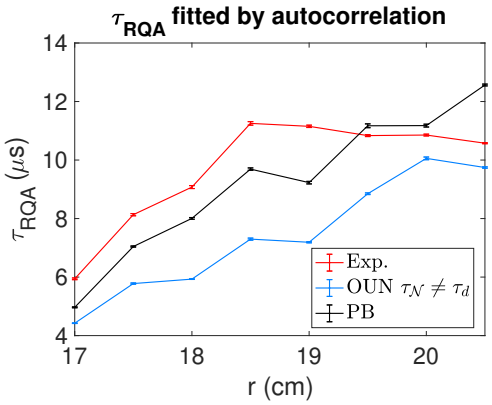


Figure B.9: Profiles of the delays τ_{RQA} fitted as the values at which the autocorrelation function reached 10%, $R_{\tilde{I}}(\tau_{RQA}) = 10\%$. The red, blue and black curves correspond respectively to the experiment 34132 (Sections 3.4 and 6.1), the simulations with fitted Ornstein-Uhlenbeck noise (Chapter 6) and the simulations with pulse background (Chapter 8).

B.5 Recurrence plots of the saturation current

In this appendix, we show recurrence plots (Section 2.8) for the saturation current measured in the TCABR experiment 34132 (Section 3.4) and for simulations of the stochastic model (fitted OU noise and pulse background, Chapters 4 to 8). Other RPs of I_{sat} can be seen in Refs. 102 and 103, but for the Texas Helimak. Here the recurrence rate was fixated at 10% (in analogy to Refs. 100–102), the immersion dimension used was $d_{RQA} = 4$ (as in Refs. 99–102), and each delay τ_{RQA} was chosen as the value at which the signal autocorrelation reached 10% (in analogy to 99 and 102). Fig. B.9 shows the obtained τ_{RQA} for each signal considered in this section.

In its turn, Fig. B.10 exhibits the recurrence plots for the experimental I_{sat} and its simulations. For $r = 17.0$ cm, the RPs slightly resemble the one of white noise (Fig. 2.6(a)), which is dominated by isolated and uniformly distributed points. In contrast, for $r = 20.5$ cm, there is more empty space, caused by high-amplitude bursts that aren't in general recurrent

with other segments. This is in accordance with Ref. 101, where it is stated: "Abrupt changes or rare events manifest themselves as white bands".

As the reader may recall from Fig. 8.7(c), for $r = 17.0$ cm and $r = 20.5$ cm, the RQA determinism respectively assumes its lowest and greater values, DET ≈ 0.70 and DET ≈ 0.85. Also, the same occurs for the kurtosis (and skewness) as seen in Fig. 8.3, or to the turbulence level $\sigma_I / \langle I_{sat} \rangle$ (Fig. 3.11(a)). These facts are evidence that strongly intermittent plasma fluctuations have a higher RQA determinism.

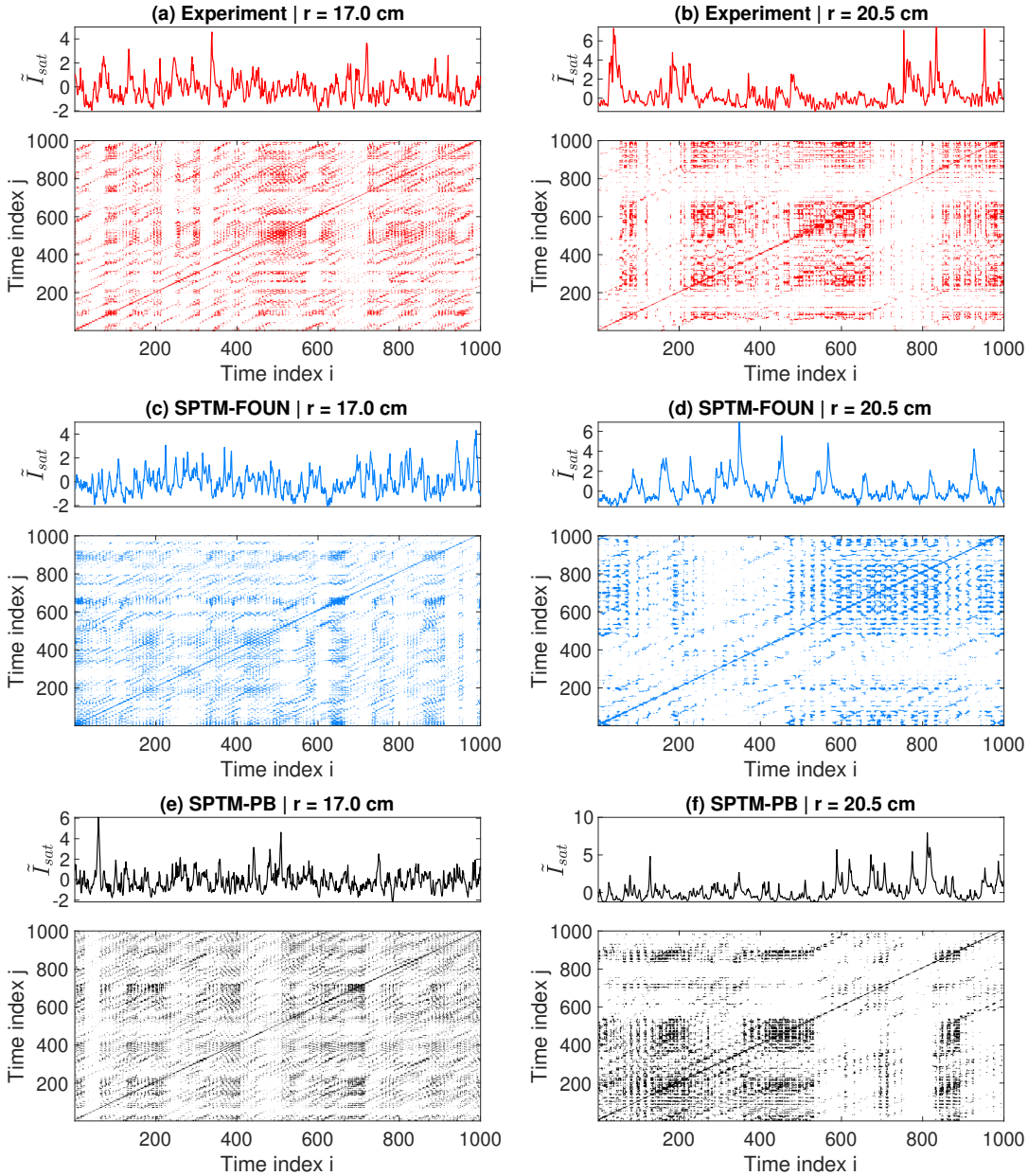
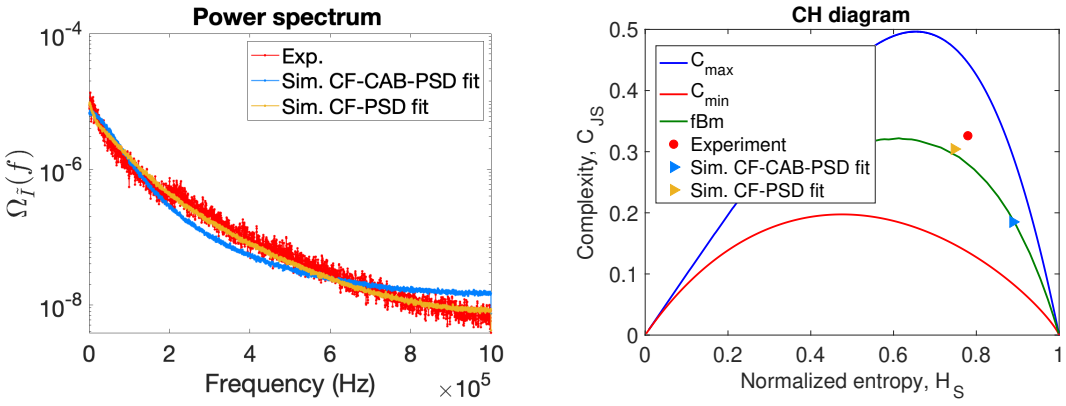


Figure B.10: Recurrence plots of the saturation current, for the experiment 34132 (red, top panels), and for the SPTM simulations with fitted OU noise (SPTM-FOUN, blue, center panels) and with pulse background (SPTM-PB, black, bottom panels).

B.6 CF-PSD fit for the SPTM-N

As noted in Section 6.2, the SPTM-N can not reproduce the experimental I_{sat} power spectrum seen in TCABR, if the CF-CAB-PSD fit is used. However, the power spectrum can be better adjusted if, instead of fitting τ_d and λ with the CAB, these two are adjusted with a chi-square map of the PSD, together with the parameter $\tau_{\mathcal{N}}$. This would yield a CF-PSD fit (the CF part is equal in both methods).



(a) **(b)**

Figure B.11: **(a)** Adjustment of the power spectrum with the CF-CAB-PSD fit (blue) and CF-PSD fit (yellow) for the position $r = 18.5$ cm in TCABR, experiment 34132, and SPTM-N simulations. **(b)** Corresponding points in the CH diagram for the same experiment and simulations.

Fig. B.11(a) shows a comparison between one experimental PSD and two simulated PSDs fitted with the CF-CAB-PSD method (in blue) and with the CF-PSD method (in yellow). The CF-PSD fit is clearly more adequate, and the same is observed for the CH diagram (Fig. B.11(b)). The obtained parameters λ, τ_d and $\tau_{\mathcal{N}}$ for both methods were

$$\boldsymbol{\theta}_{\text{CF-CAB-PSD}} = (\lambda, \tau_d, \tau_{\mathcal{N}}) = (0.29(3), 3.8(1)\ \mu\text{s}, 1.3(1)\ \mu\text{s})$$

and

$$\boldsymbol{\theta}_{\text{CF-PSD}} = (\lambda, \tau_d, \tau_{\mathcal{N}}) = (0.15(3), 2.5(2)\ \mu\text{s}, 15(3)\ \mu\text{s})$$

The asymmetry and duration parameters of the pulses, λ and τ_d, are smaller for the CF-PSD fit.[1] This may indicate that the plasma density in TCABR would be better described by the superposition of two types of structures: the bursts, with higher characteristic duration ($\tau_d \simeq 3.8$ µs in the present position and experiment); and by smaller pulses, not detected by the conditional averaging analysis, with smaller characteristic duration ($\tau_d \simeq 2.5$ µs).

[1] Moreover, the noise correlation parameter $\tau_{\mathcal{N}} = 15(3)$ µs obtained with the CF-PSD method was much higher than the one adjusted by the CF-CAB-PSD, $\tau_{\mathcal{N}} = 1.3(1)$ µs. However, we investigated that the results wouldn't be much different if a lower value was used for the first simulation, for example, $\tau_{\mathcal{N}} = 5$ µs.

This hypothesis is corroborated by the simulated conditionally averaged bursts (Fig. B.12). The fit of the averaged burst made by the CF-PSD method is worse than the one made with the CF-CAB-PSD method since the first has a smaller pulse duration τ_d.

The results of Fig. B.11 and B.12 led us to investigate whether a stochastic pulse train model with a background of small and short pulses would reproduce the experiment better than the SPTM with noise. The results of the SPTM with pulse background (PB) are depicted in Chapters 7 and 8. In special, a glance at Fig. 6.2(c) and Fig. 8.2(c) shows that the I_{sat} power spectrum is better fitted by the SPTM-PB than by SPTM-N.

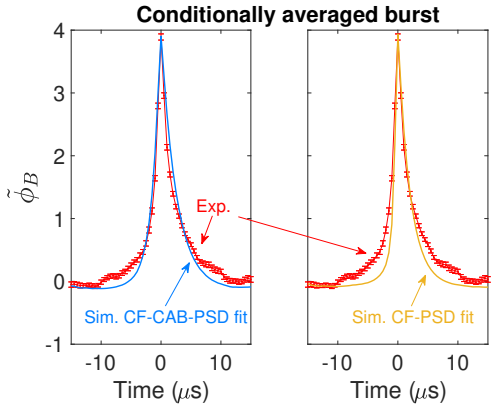


Figure B.12: In red is shown the conditionally averaged burst (CAB) from the TCABR position $r = 18.5$ cm and shot 34132. In blue and yellow, are the simulated CABs fitted with the CF-CAB-PSD and CF-PSD methods, respectively.

B.7 Truncated exponential

In the stochastic pulse train model, the form of the pulses is exponential (see Chapters 4 and 7 for details). Because of this, there is a ingenious way to optimize the simulation computation time, truncating the exponential function.[2] On one side, we know that

$$\int_0^\infty \exp\left(-\frac{t}{\tau_f}\right) dt = \tau_f$$

But, on the other hand,

$$\int_0^{15\tau} \exp\left(-\frac{t}{\tau_f}\right) dt = \tau_f \left[1 - \exp\left(-15\right)\right] = 99.99997\% \cdot \tau_f$$

That is, if the exponential is truncated at $15\tau_f$, the error committed in the pulse integral with relation to the exact value would be only of $\exp(-15) = 3 \cdot 10^{-5}\%$ of the pulse amplitude. The same argument is valid to the pulses rise, τ_r.

[2]The idea for the method is due to F. A. C. Pereira.

Furthermore, the exact average of the background-less pulse train is (Eq. (4.11), where $\gamma = (\tau_r + \tau_d)N_P/T$)

$$\langle I_{sat} \rangle = \langle A \rangle \frac{N_P}{T} (\tau_r + \tau_f)$$

and the truncated mean would be

$$\langle I_{sat} \rangle_{trunc} = \frac{N_P}{T} \langle A \rangle \left(1 - 3 \cdot 10^{-7}\right) (\tau_r + \tau_f)$$

So the error committed in the mean of I_{sat} due the truncation of a symmetric pulse ($\tau_r = \tau_f$) would be

$$\langle I_{sat} \rangle - \langle I_{sat} \rangle_{trunc} = \exp(-15) \langle I_{sat} \rangle = 3 \cdot 10^{-7} \langle I_{sat} \rangle$$

Hence, as this difference is too small compared to the statistical fluctuation of the average ($\sim 10^{-2} \langle I_{sat} \rangle$), it is helpful and secure to adopt this truncation procedure.

For comparison, in the computer used for this work, a simulation with 40 thousand pulses and 80 thousand points takes one minute to be done. When truncating the exponential on $15\tau_r$ and $15\tau_d$, however, the same simulation is computed in just two seconds, which is 30 times faster. Individually, one minute may not seem of great importance, but when dealing with thousands of simulations (to estimate fluctuations, for example), one day of simulation could become one hour. This truncation technique was also adopted in Ref. 48.

We can also extend the truncation concept to the stretched-waveform case (Eq. (7.4)). In this scenario, the shape contribution for the fall is, in units of the fall characteristic duration ($\theta = t/\tau_f$),

$$\int_0^\infty \exp\left(-\theta^\zeta\right) d\theta = \Gamma\left(1 + \frac{1}{\zeta}\right) = \frac{1}{\zeta}\Gamma\left(\frac{1}{\zeta}\right)$$

Truncating the integral by a length l yields

$$\int_0^l \exp\left(-\theta^\zeta\right) d\theta = \frac{1}{\zeta}\Gamma\left(\frac{1}{\zeta}\right) - \frac{1}{\zeta}\Gamma\left(\frac{1}{\zeta}, l^\zeta\right)$$

where $\Gamma(a, x)$ is the incomplete Gamma function:

$$\Gamma(a, x) \equiv \int_x^\infty y^{a-1} e^{-y} dy$$

i.e.,

$$\Gamma\left(\frac{1}{\zeta}, l^\zeta\right) = \int_{l^\zeta}^\infty y^{1/\zeta - 1} e^{-y} dy$$

The difference between the exact and the truncated integral is thus

$$\int_0^\infty \exp\left(-\theta^\zeta\right) d\theta - \int_0^l \exp\left(-\theta^\zeta\right) d\theta = \frac{1}{\zeta}\Gamma\left(\frac{1}{\zeta}, l^\zeta\right)$$

We want too find a truncation length l such that the error committed in the approximation is less than 10^{-3},

$$\varepsilon(\zeta, l) \equiv \frac{\Gamma\left(1/\zeta, l^\zeta\right)}{\Gamma\left(1/\zeta\right)} < 10^{-3} \tag{B.3}$$

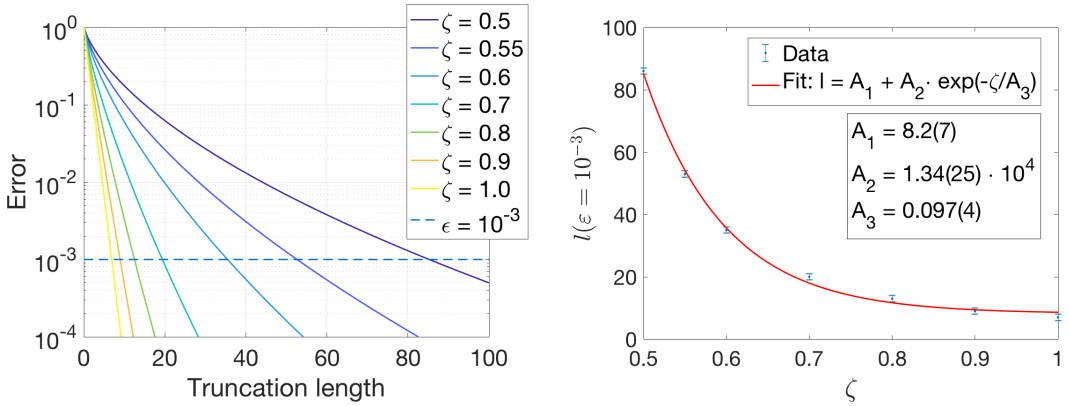
Fig. B.13(a) shows the error (B.3) as a function of the truncation length l for various stretching exponents. Adopting a truncation length of $l = 90$ will suffice up to $\zeta = 0.5$. In terms of the characteristic duration, this means that the pulse fall exponential will be truncated in $90\tau_f$. This time interval is still much smaller than the total time in consideration: $T \sim 40$ ms $\gg 90\tau_f \sim 0.2$ ms.

The same argument is valid for the rise, and it is practical to define a truncation vector $\mathbf{l} = [l_r, l_f]$ such that the rise and the fall are truncated in $\tau_r l_r$ and $\tau_f l_f$, respectively.

The truncation length for $\varepsilon = 10^{-3}$ can be approximately fitted by an exponential, as seen in Fig. B.13(b). The result is

$$l(\zeta; \varepsilon = 10^{-3}) = 8.2(7) + 1.34(25) \cdot 10^4 \exp\left(-\frac{\zeta}{0.097(4)}\right)$$

This fit was employed in the dissertation so that, for each value of ζ, a truncation length was chosen to speed up the simulation, without compromising the accuracy of the model.



(a) **(b)**

Figure B.13: **(a)** Error vs truncation length for various stretching exponents ζ. **(b)** In blue, truncation length as a function of ζ for error $\varepsilon = 10^{-3}$. In red, fit with an exponential.

B.8 Characteristic function of the train of stretched pulses

B.8.1 Derivation

In this section we will derive the expression for the characteristic function of $I_B(t)$, which was used in Section 7.2. In its turn, $I_B(t)$ is the train of stretched pulses (Eq. (7.2)),

$$I_B(t) = \sum_{j=1}^{N_P} A_j \phi\left(\frac{t - t_j}{\tau_d}; \lambda, \zeta_r, \zeta_f\right)$$

The amplitudes A_j and the waiting times $\Delta t_j = t_{j+1} - t_j$ between events follow exponential distributions. The pulse shape ϕ is given by stretched exponentials (Eq. (7.4)),

$$\phi(\theta; \lambda, \zeta_r, \zeta_f) = \begin{cases} \exp\left(-|\theta/\lambda|^{\zeta_r}\right) & , \ \theta < 0 \\ \exp\left\{-\left[\theta/\left(1-\lambda\right)\right]^{\zeta_f}\right\} & , \ \theta \geq 0 \end{cases}$$

Following Eq. (A13) of Ref. 55, the logarithm of the characteristic function for a train of independent pulses with uniform occurrence is

$$\ln C_B(u) = \gamma \int_{-\infty}^{\infty} d\theta \left[C_A\left(u\phi(\theta)\right) - 1\right]$$

where $\gamma = \tau_d N_P / T$ is the same intermittency parameter from the SPTM-N model (Chapter 4) and, in this case, $\phi(\theta) = \phi(\theta; \lambda, \zeta_r, \zeta_f)$. $C_A(u)$ is the characteristic function of the amplitude PDF. Since A is exponentially distributed,

$$C_A(u) = \frac{1}{1 - iu\left\langle A\right\rangle}$$

and thus

$$C_A(u\phi(\theta)) = \frac{1}{1 - iu\phi(\theta)\left\langle A\right\rangle}$$

Returning to the integral,

$$\ln C_B(u) = \gamma \int_{-\infty}^{\infty} d\theta \left[\frac{1}{1 - iu\phi(\theta)\left\langle A\right\rangle} - 1\right]$$

The integral can be split into rise and fall parts,

$$\ln C_B(u) = J_r + J_f$$

where

$$J_r = \gamma \int_{-\infty}^{0} d\theta \left[\frac{1}{1 - iu\exp\left(-|\theta/\lambda|^{\zeta_r}\right)\left\langle A\right\rangle} - 1\right]$$

and

$$J_f = \gamma \int_{0}^{\infty} d\theta \left[\frac{1}{1 - iu\exp\left\{-\left[\theta/\left(1-\lambda\right)\right]^{\zeta_f}\right\}\left\langle A\right\rangle} - 1\right]$$

Defining $\lambda_f = 1 - \lambda$,

$$J_f = \gamma \int_{0}^{\infty} d\theta \left[\frac{1}{1 - iu\left\langle A\right\rangle\exp\left[-\left(\theta/\lambda_f\right)^{\zeta_f}\right]} - 1\right]$$

Since

$$\frac{1}{1-iu\left\langle A\right\rangle \exp\left[-\left(\theta/\lambda_f\right)^{\zeta_f}\right]}-1=\frac{1-1+iu\left\langle A\right\rangle \exp\left[-\left(\theta/\lambda_f\right)^{\zeta_f}\right]}{1-iu\left\langle A\right\rangle \exp\left[-\left(\theta/\lambda_f\right)^{\zeta_f}\right]}$$

then

$$J_f=\gamma\int_0^\infty d\theta\frac{iu\left\langle A\right\rangle \exp\left[-\left(\theta/\lambda_f\right)^{\zeta_f}\right]}{1-iu\left\langle A\right\rangle \exp\left[-\left(\theta/\lambda_f\right)^{\zeta_f}\right]}$$

or even

$$J_f=\gamma\int_0^\infty d\theta\frac{iu\left\langle A\right\rangle }{\exp\left[\left(\theta/\lambda_f\right)^{\zeta_f}\right]-iu\left\langle A\right\rangle }$$

Let $x=\theta/\lambda_f$. So $dx=d\theta/\lambda_f$ and

$$J_f=\gamma\lambda_f\int_0^\infty dx\frac{iu\left\langle A\right\rangle }{\exp\left(x^{\zeta_f}\right)-iu\left\langle A\right\rangle }$$

This is a nonelementary integral. Defining

$$L(z;\zeta)\equiv\int_0^\infty dx\frac{z}{\exp\left(x^{\zeta}\right)-z} \tag{B.4}$$

then

$$J_f=\gamma\left(1-\lambda\right)\cdot L(iu\left\langle A\right\rangle ;\zeta_f)$$

and the analogous is obtained for J_r,

$$J_r=\gamma\lambda\cdot L(iu\left\langle A\right\rangle ;\zeta_r)$$

Thus the natural logarithm of C_B is

$$\ln C_B(u)=J_r+J_f\implies$$

$$\boxed{\ln C_B(u)=\gamma\lambda\cdot L(iu\left\langle A\right\rangle ;\zeta_r)+\gamma\left(1-\lambda\right)\cdot L(iu\left\langle A\right\rangle ;\zeta_f)}$$

B.8.2 Relation to the polylogarithm

The function $L(z;\zeta)$ is closely related to the polylogarithm (p. 611 of Ref. 123),[3]

$$\mathrm{Li}_s(z)\equiv\frac{1}{\Gamma\left(1+s\right)}\int_0^\infty dx\frac{z}{\exp\left(x^{1/s}\right)-z} \tag{B.5}$$

[3]Actually, in Ref. 123 the integral representation of the polylogarithm is expressed as

$$\mathrm{Li}_s(z)=\frac{z}{\Gamma(s)}\int_0^\infty\frac{y^{s-1}}{e^y-z}dy$$

With the change of variables $y=x^{1/s}$, it is easy to show that the above equation is equivalent to Eq. (B.5).

That is,

$$L(z;\zeta) = \Gamma(1+1/\zeta)\mathrm{Li}_{1/\zeta}(z)$$

The polylogarithm converges for $\mathrm{Re}(s) > 0$ and all $z \in \mathbb{C}$, except for real z such that $z \geq 1$. It also converges for $\mathrm{Re}(s) > 1$ and $z = 1$. Therefore, $L(z;\zeta)$ converges for the same conditions, keeping in mind that $s = 1/\zeta$. Since in the case analyzed for this dissertation $\zeta \in]0,1]$ and $z = iu\langle A\rangle$, with $u, \langle A\rangle \in \mathbb{R}$, then $L(iu\langle A\rangle;\zeta)$ converges always.

Furthermore, for $z \in \mathbb{C}$ and $|z| < 1$, the polylogarithm can be written as

$$\mathrm{Li}_s(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^s}$$

Thus, in this case,

$$L(z;\zeta) = \Gamma\left(1+\frac{1}{\zeta}\right)\sum_{n=1}^{\infty}\frac{z^n}{n^{1/\zeta}}$$

With this expression, it is possible to find the cumulants and moments of the train of stretched pulses, I_B. This was done in Section 7.3.