UNIVERSITAT DE BARCELONA

Statistical modelling of avalanche observables: criticality and universality

Víctor Navas Portella

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Statistical modelling of avalanche observables: criticality and universality

PhD Dissertation

Víctor Navas Portella

Thesis supervisors

Álvaro Corral Cano

Eduard Vives i Santa Eulalia

Thesis tutor

Josep Vives i Santa Eulalia

PhD program in Mathematics and Computer Science of the University of Barcelona Graduate School (EDUB)

UNIVERSITAT DE BARCELONA

CRM

CENTRE DE RECERCA MATEMÀTICA

Centre de Recerca Matemàtica
Complex Systems Group
Edifici C, Campus Bellaterra,
E-08193 Bellaterra, Catalonia, Spain

Acknowledgments

Tal i com va escriure Miguel de Cervantes a la segona part de la seva obra "El Quijote": "La ingratitud es hija de la soberbia". Dit això, no voldria deixar ningú fora de tot el meu sincer agraïment. Si en algun descuit, sempre lluny de la supèrbia, ometo algun nom (que queda inclòs en algun dels grupúscles aquí anomenats) o, senzillament, oblido algú, demano amb antel·lació disculpes.

Començaré agraïnt a totes aquelles persones que, des de l'àmbit acadèmic han fet possible que aquesta tesi sortís endevant. No podria fer altra cosa que començar agraïnt els meus directors de tesi: a l'Álvaro Corral, per dipositar la teva confiança en mi i en donar-me l'oportunitat d'embarcar-me en aquesta aventura. A l'Eduard Vives, per introduir-me al món de la recerca i donar-me ales i motivació. Què menys que un senzill agraïment als dos per aguantar-me durant tot aquest temps, per esperonar-me dia a dia i intentar fer que la feina sempre fós una mica millor.

Agrair també a tots els components del grup de Complex Systems. A Patricia Paredes por ser un referente de fuerza de voluntad. A l'Isabel Serra per totes les oportunitats que m'has donat, pel teu constant suport i tota l'ajuda incondicional. A en Jordi Baró, per ser una font d'inspiració constant i per tots els bons moments a Pamplona, Nova Orleans i a Les Houches. A Álvaro González, por toda la atención, dedicación y tiempo. Siempre estaré agradecido por la oportunidad que me diste en Viena y los buenos momentos que pasamos. To Niclas Rieger and Mónica Minjares, although we have met at the last stage of my PhD it is always a pleasure work with positive people at your side. Good luck with your PhD Theses. Agradecer también a Abigail Jiménez, compañera de fatigas científicas que finalmente han llegado a buen puerto. Gracias por descubrirme el mundo de la seismología y por todo lo batallado.

To Nicholas Moloney for his attention and all his time and effort during my stay at London Mathematical Laboratory (LML). To all the people from LML for their kindness and hospitality.

A tota la gent del CRM. Començant pels companys: Núria Folguera, Helena Ribera, Gemma Colldeforns, Marc Calvo, Claudia Fanelli, Alberto Debernardi, Carmelo Puliatti, Roberto de la Cruz, Marina Vegué, Carles Barril, Bernat Rovira, Enric Costa, Genís Prat, Nicolás Pollán, Daria Stepanova, Pau Blanco, Citlali Vilar, Elisa Beltran, Francesc Font i Jordi Canela. Gràcies a tots pels bons moments que m'heu fet passar. A tota la gent d'Administració i que fan possible que el CRM funcioni: Pau Varela, Núria Fernández, Alba

Tomás, Vanessa Ramírez, Mari Paz Valero, Consol Roca, Ana García Donás, Raquel Garrido, Luca Tancredi, Arantxa Sanz, Gemma Martínez, Alejandra Marques, David Romero, Claudi Lledyda, Mª Àngels Huertos, José Antonio Fuentes i Lluís Alsedà. Moltes gràcies pel vostre temps, atenció i la vostra confiança.

A tots els companys del Grau i del Màster en Física: Enric Sanmartí, Gerard Pelegrí, Pere Arnan, Josep Ignasi Decarrega, Èric Seró, Marta Conti, Pere Mujal i Xavier Roderic Hoffmann. Per tots aquells cafès a la Masia, els dies estudiant al Raval, les calçotades, les rutes de bars, l'estiu a Menorca... crec que gran part d'aquests moments han contribuït a aquesta tesi.

A tots els companys i companyes del MBA de la Universitat de Vic - Universitat Central de Catalunya que vaig realitzar durant el primer any de tesi. A toda la gente de EAE y sus estudiantes. Gracias por todo lo que he aprendido con y de vosotros.

Aquesta tesi no hauria estat possible sense totes aquelles persones que, fora de l'àmbit acadèmic, m'han recolzat. A totes aquelles persones que, en els moments que he estat al fang, han sabut reconeixe'm, valorar-me, m'han prestat la seva ajuda desinteressadament i amb els qui tindré una gratitud eterna: Raúl López Zamorano, Javi Vengas i Anaïs Martínez. A tots aquells companys i companyes del Cor Camins i del club d'altetisme A4elKm. A tota la gent del teatre. Tornar a agrair al Raúl, per ser com un germà gran. Per totes les bones estones que hem passat i les que passarem. No ets conscient de tota la vida que em dones. Seguint amb el món del teatre, especial menció a l'Albert Vilar, per tota l'estima que et tinc i les bones estones que ens fas passar. A Kuriatin de Txékhov, Currito de Lorca, Leandre de Molière i Christopher Wren de Christie. Personatges que m'han acompanyat durant aquesta tesi. A Ricard Manchón y a Guillermo Caro: A pesar de que cronólogicamente tengo que irme muy hacia atrás en el tiempo, nunca es tarde ni suficiente para acordarme de vosotros y daros vuestro merecido. Os debo toda la gratitud desde aquel momento en que me introdujisteis en aquel instituto y me distéis alas. Ricky, siento lo del placaje, pero, quiero que sepas, que ante todo, iba con cariño. A Mireia Carles i a Eva de la Cruz. Per la vostra gran companyia i sempre lleieltat a la meva germana. A les Núries: Núria Portella, no només et destacaré com a Núria, et destacaré com a mare unes línies mes avall. Ets la persona que més agraïments mereix en aquest humil escrit. Núria Federico, per ajudar-me en moments difícils i per tot el que hem compartit fins ara. A Núria Folguera, perquè mai podré agrair-te tots els riures i totes les tristeses compartides. De no ser per tu, probablement hauria caigut a la meitat d'aquest camí. Saps que al món falten peresones com tu. A Núria Arredondo, senzillament perquè sense tu no estaria aquí escrivint això. Als Cucus: A la Mel i al Conrad. Per tantes estones amb vosaltres, pel vostre afecte i el vostre amor incondicional. Us trobo a faltar en el meu dia a dia. A Nerea Carreras y a Ostrobotnia Central. Clave para la elaboración de esta tesis. Un oasis de hielo que me dio fuerzas para seguir caminando y que me permitió escribir. No hay elemento retórico con más potencia que el hecho de comer fresas congeladas en Ostrobotnia Central.

A la meva família. A tots els meus cosins i cosines: Anna Portella, Alberto Cid, Raquel Guerra, Noelia Galván, Javier Cid, Sergio Guerra, Jaime Galván y Alex Muñoz. Especials agraïments al meu cosí David Portella. Aquell camí de Santiago ha estat una de les millors experiències de la meva vida. No ets conscient de la força mental que ha representat aquest viatge amb tu a l'hora d'elaborar aquesta tesi. A tots els meus oncles i tietes: Joan Portella, José Galván, Manuel Guerra i Daniel Navas, Matilde Asensio, Montse Navas, Charo Navas, Mari Carmen Navas i Cristina Navas. A les meves àvies: Rosario Moreno i Maria Portas. Cómo no mencionar a mi vallisoletana favorita, Ester Redondo. Muchas gracias por tu buen humor y por todas las risas. A la meva família del Sud: Anna, Carmen, Ramon, Joan i Pau. Per tots els bons moments que hem passat i els que ens queden per passar!

A totes aquelles persones que, malauradament, ens han anat deixant en el transcurs d'aquesta tesi: Josep Portas Brunsó i Santiago Tomás Arias. Als meus dos avis, amb els qui no tinc l'honor ni el plaer de parlar: Nemesio Navas Montes i a Vicens Portella Sels.

Finalment, dedicar unes paraules al meu nucli més proper. El nucli dur d'aquesta tesi. Abans d'escriure, ja estic cent per cent segur que les paraules d'agraïment es quedaran curtes. A Sara Martínez Buera, per tot el que has suposat en aquesta etapa final de la tesi. Perquè el teu somriure em dona força dia rere dia. Per tot el que hem viscut i tot el que ens queda per viure. A Santi Tomás, mi hermano no biológico. Por todas las risas y grandes momentos. Por todas las aventuras vividas (Cavalls del Vent o Behobia por mencionar algunas) y las que nos quedan por vivir. Sencillamente por estar siempre ahí. A Sònia Navas, la meva germana. Perquè sempre alegres la vida a tots els qui estem al teu voltant. Senzillament, no podria viure sense tu. A Núria Portella, la meva mare. Per no ser només una mare sino també una amiga amb qui parlar i compartir-ho tot. Perquè, juntament amb la Sònia, sou les meves guies emocionals, brúixoles en tempestes i bonances. A Nemesio Navas, mi padre. Por enseñarme que cada día es un nuevo desafío y una nueva batalla. Por enseñarme también que con trabajo y esfuerzo se pueden lograr muchas cosas.

Finalment, recordar també aquells qui no van creure. Especial mención a María del Carmen Ruiz Zabalza, mi profesora del ciclo superior en la escuela primaria, quien dijo que yo era una persona de letras y que no era alguien dado a los números. És cert. A l'escola primària suspenia matemàtiques... Doncs bé, aquí estem.

Víctor Navas Portella
26 de febrer de 2020

Abstract

Complex systems can be understood as an entity composed by a large number of interactive elements whose emergent global behaviour cannot be derived from the local laws characterizing their constituents. The observables characterizing these systems can be observed at different scales and they often exhibit interesting properties such as lack of characteristic scales and self-similarity. In this context, power-law type functions take an important role in the description of these observables. The presence of power-law functions resembles to the situation of thermodynamic quantities close to a critical point in equilibrium critical phenomena. Different complex systems can be grouped into the same universality class when the power-law functions characterizing their observables have the same exponents.

The response of some complex systems proceeds by the so called avalanche process: a collective response of the system characterized by following an intermittent dynamics, with sudden bursts of activity separated by periods of silence. This kind of out-of-equilibrium systems can be found in different disciplines such as seismology, astrophysics, ecology, finance or epidemiology, just to mention a few of them. Avalanches are characterized by a set of observables such as the size, the duration or the energy. When avalanche observables exhibit lack of characteristic scales, their probability distributions can be statistically modelled by power-law-type distributions. Avalanche criticality occurs when avalanche observables can be characterized by this kind of distributions. In this sense, the concepts of criticality and universality, which are well defined in equilibrium phenomena, can be also extended for the probability distributions describing avalanche observables in out-of-equilibrium systems.

The main goal of this PhD thesis relies on providing robust statistical methods in order to characterize avalanche criticality and universality in empirical datasets. Due to limitations in data acquisition, empirical datasets often only cover a narrow range of observation, making it difficult to establish power-law behaviour unambiguously.

With the aim of discussing the concepts of avalanche criticality and universality, two different systems are going to be considered: earthquakes and acoustic emission events generated during compression experiments of porous materials in the laboratory (labquakes). The techniques developed in this PhD thesis are mainly focused on the distribution of earthquake and labquake sizes, which is known as the Gutenberg-Richter law. However, the methods are much more general and can be applied to any other avalanche observable.

The statistical techniques provided in this work can also be helpful for earthquake forecasting. Coulomb-stress theory has been used for years in seismology to understand how earthquakes trigger each other. Earthquake models that relate earthquake rates and Coulomb stress after a main event, such as the rate-and-state model, assume that the magnitude distribution of earthquakes is not affected by the change in the Coulomb stress. Several statistical analyses are performed to test whether the distribution of magnitudes is sensitive to the sign of the Coulomb-stress increase.

The use of advanced statistical techniques for the analysis of complex systems has been found to be necessary and very helpful in order to provide rigour to the empirical results, particularly, to those problems regarding hazard analysis.

Resum

Els sistemes complexos es poden entendre com entitats compostes per un gran nombre d'elements en interacció on la seva resposta global i emergent no es pot derivar de les lleis particulars que caracteritzen cadascun dels seus constituents. Els observables que caracteritzen aquests sistemes es poden observar a diferents escales i, sovint, mostren propietats interessants tals com la manca d'escales característiques i autosimilitud. En aquest context, les funcions amb lleis de potència prenen un paper important en la descripció d'aquests observables. La presència de lleis de potència s'assimila a la situació dels fenòmens crítics en equilibri, on algunes quantitats termodinàmiques mostren un comportament funcional similar prop d'un punt crític. Diferents sistemes complexos es poden agrupar en la mateixa classe d'universalitat quan les funcions de lleis de potència que caracteritzen els seus observables tenen els mateixos exponents.

Quan són conduïts externament, la resposta d'alguns sistemes complexos segueix el que s'anomonena un procès d'allaus: una resposta col·lectiva del sistema caracteritzada per seguir una dinàmica intermitent, amb sobtats increments d'activitat separats per períodes de silenci. Aquesta mena de sistemes fora de l'equilibri es poden trobar en diferents disciplines tals com la sismologia, astrofísica, ecologia, epidemologia o finances, per mencionar alguns. Les allaus estan caracteritzades per un conjunt d'observables tals com la mida, l'energia o la durada. Quan aquests observables mostren una manca d'escales característiques, les seves distribucions de probabilitat es poden modelitzar estadísticament per distribucions de lleis de potència. S'anomenen allaus crítiques aquelles en que els seus observables es poden caracteritzar per aquestes distribucions. En aquest sentit, els conceptes de criticalitat i universalitat, els quals estan ben definits per fenòmens en equilibri, es poden extendre per les distribucions de probabilitat que descriuen els observables de les allaus en sistemes fora de l'equilibri.

L'objectiu principal d'aquesta tesi doctoral és proporcionar mètodes estadístics robusts per tal de caracteritzar la criticalitat i la universalitat en allaus corresponents a dades empíriques. Degut a les limitacions en l'adquisició de dades, les dades empíriques sovint cobreixen un rang petit d'observació, dificultant que es pugui establir un determinat comportament en forma de llei de potència de manera inequívoca.

Amb l'objectiu de discutir els conceptes de criticalitat i universalitat en allaus, es consideraran dos sistemes diferents: els terratrèmols i els esdeveniments d'emissió acústica que es generen durant experiments de compressió de materials porosos al laboratori (*labquakes*). Les

tècniques desenvolupades en aquesta tesi doctoral estan enfocades principalment a la distribució de la mida dels terratrèmols i labquakes, altrament coneguda com a llei de Gutenberg-Richter. No obstant, aquests mètodes són molt més generals i es poden aplicar a qualsevol observable de les allaus.

Les tècniques estadístistiques proporcionades en aquest treball poden també ajudar al pronòstic de terratrèmols. Durant anys, la teoria d'esforços de Coulomb s'ha utilitzat en sismologia per tal d'entendre com els terratrèmols desencadenen l'ocurrència d'altres de nous. Els models de terratrèmols que relacionen la taxa d'ocurrència de rèpliques i l'esforç de Coulomb després d'un gran esdeveniment, assumeixen que la distribució de la mida dels terratrèmols no està afectada pel canvi en l'esforç de Coulomb. Diverses anàlisi estadístiques s'aplicaran per tal de comprovar si la distribució de magnituds és sensible al signe de l'esforç de Coulomb.

S'ha provat que l'ús de tècniques estadístiques avançades en l'anàlisi de sistemes complexos és útil i necessari per tal d'aportar rigor als resultats empírics i, en particular, a problemes d'anàlisi de riscos.

Preface

This PhD thesis dissertation summarizes the author's contributions to the research on the statistical modelling of avalanche observables. In this context, special attention is paid to power-law distributions, which are a fingerprint of criticality and a first step to establish universality. This PhD thesis has been carried out as a part of the PhD program in Mathematics and Computer Science of the University of Barcelona Graduate School (EDUB) in the period 2016-2020. The research leading to this PhD thesis has received funding from the "La Caixa" Foundation and from the Spanish Ministry of Economy and Competitiveness (MINECO, Spain) through the "María de Maeztu" Programme for Units of Excellence in R&D (Grant No. MDM-2014-0445, Project 16-3, Reference BES-2016-076580).

The aim of this PhD thesis is to provide a robust statistical background to those researchers working with power-law distributions in complex systems and, specifically, to those in the field of avalanches in out-of-equilibrium phenomena. The document is self-contained and pretends to provide an interdisciplinary confluence point for mathematicians, statisticians, physicists and complex-systems researchers to study avalanche criticality and universality.

The PhD thesis is structured as follows:

- **Part I : Introduction**
  This first part pretends to introduce the reader to the topic of avalanche criticality and universality. At the same time, the main goals of this PhD thesis are exposed and motivated. This introduction part is divided into three chapters:
  - **Chapter 1: Introduction**
    The state of the art of avalanche criticality and universality is explained. A review of complex systems and equilibrium critical phenomena is exposed together with a discussion about how the characteristic concepts of this framework can be extended to out-of-equilibrium processes.
  - **Chapter 2: Modelling avalanche criticality**
    A general mathematical description of avalanches is provided together with the difficulties appearing when avalanches are defined in experimental works. Typical features of physics-based models for avalanche processes are exposed. The fundamental concepts of statistical modelling are also explained.

  - **Chapter 3: Earthquakes and labquakes: Earth and Laboratory**
    This chapter presents the most important concepts regarding the main phenomena of interest in this PhD thesis: earthquakes and acoustic emission events generated during compression experiments of porous materials in the laboratory (labquakes).

Chapter 1 is necessary in order to understand Chapter 2. Chapters 2 and 3 are essential in order to understand Chapter 4. Chapter 3 is fundamental in order to understand the results exposed in Chapter 8. The whole Introduction part is needed in order to understand the motivations and the goals of this Thesis.

The second part is intended to present to the reader the main experimental and statistical techniques which are necessary for the development of the results exposed in this PhD thesis.

- **Part II: Methods**

  - **Chapter 4: Experimental techniques**
    Experimental methods for the obtention of labquake observables from the acoustic emission in displacement-driven compression experiments are exposed. Some basic concepts about earthquake detection are also introduced.
  - **Chapter 5: Statistical techniques**
    The already known statistical techniques, fundamental to build up and understand the procedures developed in the Results part are presented.

These chapters are essential in order to understand how the datasets analysed in the results part are obtained and analysed.

- **Part III: Results**
  Novel findings and methodologies are exposed in this part.

  - **Chapter 6 Acoustic emission during displacement-driven compression experiments of porous materials: labquakes**
    Statistical analysis of labquakes energies and force drops found in the acoustic emission during displacement-driven compression experiments of porous glasses. The results presented in this chapter have been published in:

    * Víctor Navas-Portella, Álvaro Corral, and Eduard Vives, Avalanches and force drops in displacement-driven compression of porous glasses, Physical Review E 94, 033005 (2016) [Navas-Portella et al., 2016].
  - **Chapter 7: Universality of power-law exponents by means of maximum likelihood estimation**
    A statistical procedure to merge different datasets is first presented in this chapter with two different separate aims. First, obtaining a broader fitting range for the

statistics of different experiments or observations of the same system. Second, establishing if two or more different systems may belong to the same universality class. The results exposed in this chapter have been published in:

  * Víctor Navas-Portella, Isabel Serra, Álvaro Corral, and Eduard Vives, Increasing power-law range in avalanche amplitude and energy distributions, Physical Review E 97, 022134 (2018) [Navas-Portella et al., 2018].
  * Víctor Navas-Portella, Álvaro González, Isabel Serra, Eduard Vives, and Álvaro Corral. Universality of power-law exponents by means of maximum likelihood estimation, Physical Review E 100, 062106 (2019) [Navas-Portella et al., 2019].

– **Chapter 8: Effect of the Coulomb Stress on the Gutenberg-Ricther law** This chapter pursues statistical evidences that the b-value of the Gutenberg-Richter law is affected by the value of the Coulomb stress. The results in this chapter have been published in:

  * Víctor Navas-Portella, Abigail Jiménez, and Álvaro Corral, No significant effect of Coulomb stress on the Gutenberg-Richter law after the Landers earthquake. Scientific Reports 10, 2901 (2020) [Navas-Portella et al., 2020]

- **Part IV: Conclusions**: A list of concluding remarks are exposed together with a set of suggested works that can be carried out from the results and methodologies provided in this PhD thesis.

- **Part V: Appendixes**: Some codes in C and R languages are provided as appendixes.

- **List of publications**: The publications related to this PhD thesis are presented together with additional works in which the author has contributed in.

- **Nomenclature**: A list of symbols and abbreviations is provided at the end of this PhD thesis.

- **Bibliography**: A reference list ordered by first authors' surname is provided at the end.

Víctor Navas-Portella
Centre de Recerca Matemàtica, February 26, 2020

Contents

Part I

Introduction

Chapter 1

Introduction

> The whole is greater than the sum of its parts.
>
> *Aristotle*

> Pero dos no es igual que uno más uno.
>
> *Joaquín Sabina*

In this introductory chapter, complex systems are firstly presented by exposing their most characteristic features as well as the main approaches used to describe them. Complex-systems observables often exhibit lack of characteristic scales. In this context, power-law functions play an important role. When externally driven, the dynamics of some complex systems is manifested in terms of avalanches. Avalanches are introduced by paying special attention to its definition within the framework of statistical physics. Criticality, understood as the lack of characteristic scales, is discussed by providing a classification depending on the nature of the systems which are studied. The main concepts of equilibrium critical phenomena such as Landau theory, the renormalization group approach and universality are outlined in order to elucidate how they are extended to out-of-equilibrium phenomena. Different frameworks that have appeared along the last decades to describe avalanche criticality are also discussed. Finally the importance of correctly determining critical exponents for the classification of systems into universality classes is highlighted.

1.1 Complex systems: The science of complexity

According to the dictionary [O'Neill and Summers, 2015], *Complexity is the state of having many different parts connected or related to each other in a complicated way.* This definition can be useful to introduce the study of complex systems because it contains important key words to describe them: *many*, *parts* and *connected*. The word *complicated* can be a source of confusion because essentially one does not know exactly what complicated means and how complexity can be quantified [Horgan, 1995].

Generally speaking, a complex system can be understood as a large number of interacting elements whose global behaviour cannot be derived from the local laws that characterize each of its components. The observables characterizing the emergent response of the system can be detected at different scales and the vast number of degrees of freedom makes prediction very difficult [Anderson, 1972; Sherrington, 2010]. In this context, a probabilistic description of the phenomenon is needed in order to reasonably characterize it in terms of random variables.

Taking into account this context, this work will refer to complex systems as those collections of elements which are characterized by the following general features:

- Large number of degrees of freedom.
- Interactions between the elements constituting the system.
- Emergent response of the system that cannot be derived from the local laws that characterize each of its components.

The first pillar of the study of complex systems is devoted to data acquisition and the nature of the data one is dealing with. One distinction usually done is the one regarding how the observer influences the system [Van der Laan and Rose, 2011].

Observational data are referred to data-acquisition processes for which the observer does not perturb the conditions in order to establish a clear relationship between induced causes and effects. The reasons why the observer remains as a passive agent in these cases can be diverse. Often, there is no degree of control over these systems, as it occurs in observational sciences such as astrophysics [Lu and Hamilton, 1991; Regev, 2006], geophysics [Kagan, 2013] or climate sciences [Corral and González, 2019].

Experimental data are those for which the observer influences the system in order to establish a relationship between causes and effects. In this case, there exists a certain degree of control over the system and conditions can be controlled in order to reproduce a phenomenon many times. Examples of experimental data can be extensively found in experimental physics [James, 2006; Roe, 2012] or in experimental biology and medicine [Bernard, 1957; Mead, 2017].

The second pillar of the study of complex system is based on how they can be described.

- **Physics-based modelling** proposes a set of principles, hypotheses or rules which are formulated in mathematical terms containing the key ingredients which are often identified with physical quantities. In order to develop these models, different areas of mathematics play an important role: stochastic processes [Gardiner, 2009], in particular point-processes [Lowen and Teich, 2005; Daley and Vere-Jones, 2008; Ogata, 1988], dynamical systems [Steven, 1994] or graph theory in the development of complex networks [Barabási and Albert, 1999; Albert and Barabási, 2002], just to mention few of them. Examples of physical modelling of complex systems are Ising models [Ising, 1925; Peierls, 1936] (see Secs. 1.3.2 and 2.2.2), models based on cellular automata [Von Neumann et al., 1966; Wolfram, 1983] such as the sand-pile model [Bak et al., 1987] (see Sec. 1.3.3.1), mollecular dynamics [Alder and Wainwright, 1959], fibber bundle models [Daniels and Jeffreys, 1945; Hansen et al., 2015] and so on. It must be mentioned that physics-based models are not only restricted to physical phenomena but they can be mapped in other disciplines such as financial markets [Sornette, 2009], opinion-decision [Grabowski and Kosiński, 2006; Carbone and Giannoccaro, 2015], social networks [Klemm et al., 2003], and so on [Stauffer, 2008]. .
- **Statistical modelling** has to be understood as a data-based description of the complex system without taking into account the underneath physical principles [Wood, 2015; Chave, 2017]. In front of the challenge to deal with empirical datasets, the field of statistics plays an important role and can be considered as the third pillar of the science of complex systems. Data mining algorithms might provide techniques to properly handle with large empirical datasets [James et al., 2014]. In this context, different branches of probability such as time series analysis [Hamilton, 1994] or extreme value theory [De Haan and Ferreira, 2007; Embrechts et al., 1999] take an important role. Despite some parameters characterizing the statistical model might be related to physical quantities, it is not object of the statistical modelling to explain the relationship among them. In this sense, statistical modelling of complex system has to be always understood as data-based modelling.

When the response of complex systems is characterized by observables exhibiting lack of characteristic scales, it can be described in terms of power-law functions, $f(x) \propto x^{-\gamma}$, where x ($x \in (0, \infty)$) corresponds to the values that a certain observable or quantity can take, $\propto$ denotes proportionality and the power-law exponent γ acquires values larger than one. One of the most important features of this kind of functions is their invariance under any scale transformation. This property of scale invariance can be written as

$$f(\lambda x) = \lambda^{-\gamma} f(x) \ .$$

The only solution for any real λ of this functional equation [Christensen and Moloney, 2005] is a power-law $f(x) = kx^{-\gamma}$, where k is a constant. This property of scale invariance gives rise to a description of the response of the system when there are no characteristic scales.

Historically, the study of complex systems has been closely related to the physics and physics-based models through the principles and tools provided by the theory of equilibrium statistical physics [Pathria and Beale, 2011; Huang, 2000]. In general words, statistical physics is the branch of physics which intends to describe thermodynamics from a microscopic point of view. In order to deal with this problem, one could use the procedures of classical mechanics based on considering a Hamiltonian $\mathcal{H}$ and solving the equations of motion for all the components of the system. Among the definitions of the Hamiltonian in physics [Goldstein, 1980; Thornton and Marion, 2004], it can be naively said that the Hamiltonian can be understood as a scalar function that depends on the different degrees of freedom (spatial and momentam coordinates of all the particles which conform the system). Under certain conditions, the Hamiltonian can be identified with the energy of the system and can be separated in different energy contributions: particle terms, interaction terms, external potentials, etc. However, due to the large number of degrees of freedom involved in the system, the usual approach of classical mechanics can be theoretically formulated but not used in practice. Instead of that, one tries to deal with this problem by considering a probabilistic description. The most common approaches to deal with equilibrium statistical physics are the ensemble theory, the ergodic theory and the information theory [Pathria and Beale, 2011; Huang, 2000].

Independently of which approach is used, there are some common concepts: the large number of degrees of freedom implies working with averages, there is always a probability assigned to a micro-state or to a point of the phase space, the Hamiltonian is always involved and provides important information, etc. Particularly, the fact that the observables of some complex systems can be described by power-law functions resembles to the equilibrium critical phenomena, which is actually a branch of statistical physics [Pathria and Beale, 2011; Huang, 2000]. In this context, the theory of critical phenomena and the renormalization group approach appear as a framework to which complex systems can be somehow mapped (see Sec. 1.3.2 for more details).

In summary, the study of complex systems encompasses different areas of study with a philosophy which is somehow provided by the principles of statistical physics and using techniques provided by Mathematics and Statistics.

1.2 Avalanche processes: General description

Sometimes, the response of externally driven complex systems proceeds by means of the so called avalanche process. The concept "externally driven" implies that the occurrence of a phenomenon is triggered by a perturbation caused by an agent which is not part of that system. For example, a magnetic system can be driven by an external magnetic field or fractures can be induced in a material when an external force is applied on it. Note that external driving does not necessarily imply a degree of control over the system. For instance, the tectonic blocks in the Earth crust are externally driven by the inner mechanisms of the Earth and they can hardly be controlled. As real examples of avalanche processes one can think about earthquakes [Turcotte, 1997; Kagan, 2013], the magnetization process of ferromagnetic materials [Perković et al., 1995], snow avalanches [Van Herwijnen et al., 2016], rock-falls [Dussauge et al., 2003], forest fires [Corral et al., 2008], solar flares [Lu and Hamilton, 1991] or rainfall [Peters et al., 2010]. Not only examples from physics can be mentioned but also from neuroscience [Beggs and Plenz, 2003], finance [Sornette, 2009; Iori, 1999], power-grid blackouts [Dobson, 2012], epidemic processes [Pastor-Satorras et al., 2015], social networks [Cheng et al., 2014; Goel et al., 2015].

The word *"avalanche"* is referred within statistical physics and condensed matter physics context. Let us define an avalanche process as a collective response of an externally driven system which is characterized by following an intermittent dynamics, with sudden bursts of activity separated by periods of silence. These bursts of activity or avalanches arise from the interactions among the system constituents and propagate through the spatial degrees of freedom of the system. Several words must be underlined in this definition in order to stress their implications.

First of all, the word *"collective"* implies that the system, which is an entity conformed by a certain number of interacting agents or individuals, exhibits an emergent macroscopic response. Due to these interactions, this global behaviour cannot be derived from the local laws characterizing the individuals. *"dynamics"* implies that the observables are time-dependent. Consequently, the system cannot be considered to be in equilibrium. The word *"sudden"* means that observables exhibit changes in a very short period of time. This implies the existence of different time scales in an avalanche process: events are characterized by short durations D in comparison with the time between them δ. The word *"burst"* implies that the observable exhibits an important change of scale where many orders of magnitude are involved. This also points that there exists a threshold value below which a change in the observable is not considered as an event. Apparently, the word *"activity"* simply reveals that something is happening. One can identify that something is happening if nothing was happening before. These sentences that seem extremely obvious actually hide a deep concept. On the one hand, the fact that something is happening implies the existence of underlying mechanisms which favour the occurrence of that phenomenon. On the other hand, there are other underlying mechanisms which inhibit the phenomenon. Thus, there is a competition

Fig. 1.1: Scene of Kobe, Japan, after the 1995 Hyogo-ken Nanbu earthquake (magnitude 7.2), showing a collapsed section of the Hanshin Express-way. Photography extracted from Pan-Asia Newspaper Alliance.

between mechanisms favouring and inhibiting the occurrence of an avalanche. Therefore, this word cannot be underestimated and it actually takes an important role in the framework of self-organized criticality [Bak et al., 1987] (see Sec. 1.3.3.1). Finally, the word *"silence"* is as important as the word *"activity"* because the same interpretations can be extracted from it. This word also implies that events are separated for a time-scale longer than the internal time-scale of duration of the event. Note that, together with the word *"burst"*, this word also implies the existence of a threshold value.

Note that this general description does not argue about which mechanisms avalanches appear through. It should be useful to know which are the set of conditions which are needed in order to have an avalanche process. This question is discussed in Chpt. 2.

In this general framework of avalanche processes, one individual event or avalanche is characterized by an array of observables such as the size S, the duration D, the amplitude A, the energy E, the shape, etc. Recalling the already exposed general properties of complex systems, avalanche observables are considered as random variables. Given that some avalanche observables may span many orders of magnitude, it is quite common to find that they can be described by power-law distributions. In this context, these avalanches are referred as scale-free or critical avalanches. The fact of observing scale-free behaviour of certain observables in

out-of-equilibrium systems opened the door to study them by using the concepts of equilibrium critical phenomena.

1.3 Criticality and universality

1.3.1 Physical criticality versus Statistical criticality

The concept of criticality is sometimes carelessly used and it is source of confusion and controversies. Let us introduce two different ways of dealing with the term "critical".

On the one hand, one can talk about criticality from the physical point of view. This concept is related to complex systems where it is possible to define equilibrium and non-equilibrium states. Let us remark that equilibrium must be considered from the point of view of thermodynamics [Zemansky and Dittman, 1981]: time independence and absence of fluxes in the system. Originally in physics, the term "criticality" has been associated to the existence of a critical point in an equilibrium second-order phase transition. Close to this critical point, some observables exhibit lack of characteristic scales and can be characterized by power-law functions. Due to its importance as a framework to understand scale invariance in complex systems, some basic concepts about equilibrium critical phenomena are exposed in Sec. 1.3.2. Not only the scale-invariance and self-similarity of some complex systems' observables are interesting. Additionally, completely different systems can be grouped into the same universality class depending on the values of the power-law exponents [Kadanoff et al., 1989]. In this sense, some general concepts about universality are exposed in Sec. 1.3.4.

Often, this conceptual framework is extended when avalanche observables (which are clearly out-of-equilibrium phenomena) are characterized by power-law probability density functions [Perković et al., 1995; Salje and Dahmen, 2014; Bak et al., 1987]. Let us remark that the general concept of scale invariance can be misleading when talking about the probability distributions of avalanche observables. Given that probability density functions must be normalized [Wood, 2015], the presence of cut-offs is necessary. This implies that values outside these cut-offs are excluded thus implying a restriction in the property of scale invariance. Therefore, when power-law distributions of avalanche observables are found, it is more convenient to talk about restricted scale invariance.

On the other hand, the term statistical criticality can be used whenever power-law distributions are found for different complex systems, independently on whether they are physical systems or not [Cubero et al., 2019]. This name is clearly inspired by the concepts of physical criticality but note that there can be situations in which one cannot associate the existence of a critical point. Just as it is possible to distinguish equilibrium from non-equilibrium states in physical systems, the definitions of equilibrium and out-of-equilibrium are ill-defined in other complex systems. For instance, when studying the Zipf's law in texts or songs, one could not say with conviction whether these systems are at equilibrium or not. And still, in these situations, some authors talk about criticality when the mapping with physical criticality might

not be really clear [Jensen, 1998; Schwab et al., 2014]. Statistical criticality understood as the presence of power-law distributions is clearly an abuse of language. However, conceptually it encompasses a broader range of cases of study in complex systems.

1.3.2 Equilibrium critical phenomena and universality

Historically, critical phenomena were firstly found experimentally in 1822 by the French physicist Cagniard De La Tour [1822]. By performing liquefaction experiments of gases he discovered that, above a certain temperature, the system could not experience a discontinuous change from gas to liquid. It was not until 1869, when the chemist and physicist Thomas Andrews [Rowlinson, 1969] named this temperature as the critical temperature. Contrarily to a first order phase transition, where liquid and gas phases can be clearly identified, it is not possible to distinguish them at the critical point (see Fig. 1.2). Above the critical point, there is no phase transition.

For instance, water experiences a transition from liquid to vapour at 1 atm and 100°C where these phases can be clearly differentiated. However, at the critical temperature T_c (374°C) and pressure (218 atm) phases are not distinguishable (see Fig. 1.2). This kind of

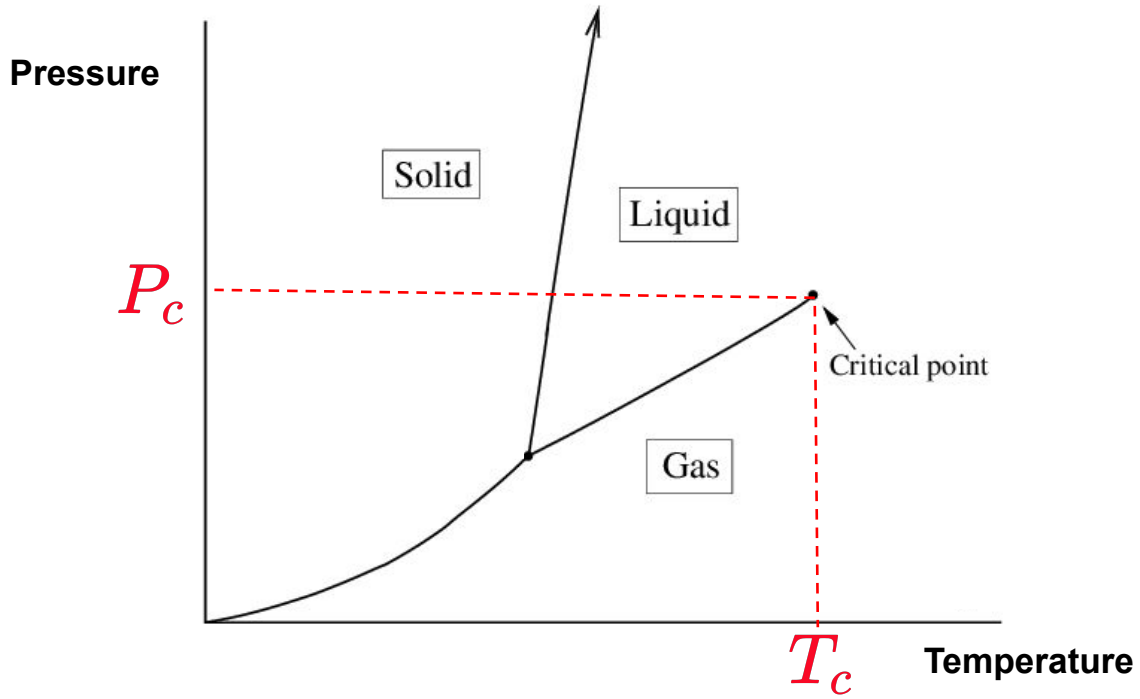

Fig. 1.2: Schematic representation of a phase diagram of an hydrostatic system. The state of the system is determined by a point in this pressure-temperature diagram. Different phases are clearly separated by black solid lines. Above the critical point (P_c, T_c), liquid and gas phases cannot be distinguished.

phase transitions are called continuous or second-order phase transitions. The nomenclature of first-order or second-order phase transitions is referred to the discontinuities that can be found either in the first or second derivatives of the free energy function of the thermodynamic system [Zemansky and Dittman, 1981; Callen, 1985].

Let us consider two important quantities in the study of equilibrium critical phenomena: the free-energy function and the correlation length. The free-energy function $\mathcal{F}$ can be written as a function of an order parameter Λ [Zemansky and Dittman, 1981; Callen, 1985; Pathria and Beale, 2011]. This order parameter depends on every system and it is zero above the critical temperature (disordered phase) and different from zero below it (ordered phase). Examples of order parameter are the magnetization in ferromagnetic system or the difference of densities in liquid-gas systems (see [Callen, 1985] for more examples). At a fixed temperature T, the equilibrium states or macroscopic phases of the systems are represented by those values of Λ for which $\mathcal{F}$ attains a minimum. Contrarily to the free-energy function, the correlation length ξ is a microscopic quantity which can be understood as a characterization of the range of "influence" among the components constituting the system. At the vicinity of a critical point, the free energy and the correlation length are non-analytic. In particular, the correlation length and other thermodynamic quantities such as the susceptibility, exhibit power-law divergences. The exponents characterizing these power-law behaviours are referred in the literature as critical exponents. The fact that the correlation length shows power-law divergence near the critical point implies a lack of characteristic scales in the correlations ruling the system. Given that correlations are independent on the observation scale, the system is said to be self-similar or to exhibit scale invariance. This means that the statistical properties characterizing a system near the critical point are the same independently on the scale.

Ising spin models have had strong influence in modelling equilibrium critical phenomena [Ising, 1925; Peierls, 1936; Pathria and Beale, 2011]. These physics-based models are often treatable analytically and their computational implementation is rather easy. Contrarily to the internal structure of a gas or a liquid, Ising models are inspired in the regular position of atoms in a solid. In this sense, Ising systems can be easily mapped into a magnetic system. In general, one considers an Ising model in a finite regular lattice where a variable called spin takes a certain value at each node (see Fig. 1.3). A wide variety of factors can be taken into account in the Hamiltonian of this model: lattice connectivity, lattice dimensionality, external perturbations, range of interaction between spins, just to mention few of them. Even under very simple conditions and rules, these models are in good agreement with the phenomenology of equilibrium critical phenomena [Christensen and Moloney, 2005]. The importance of these models is not only limited to the field of statistical physics but they have been used in other research areas such as neuroscience [Hopfield, 1982], social sciences [Weidlich, 2006], economics and finance [Sornette, 2009; Molins and Vives, 2016], and biology [Majewski et al., 2001], to name a few.

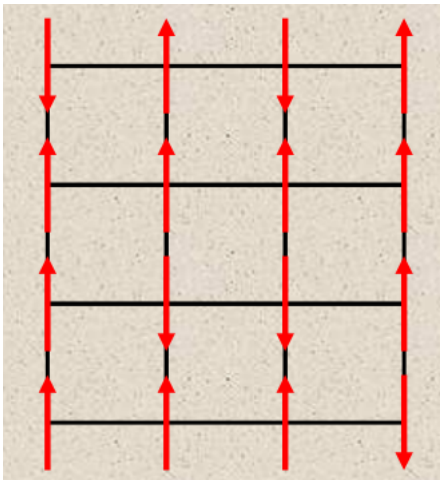

Fig. 1.3: Schematic representation of an Ising model in a two-dimensional regular square lattice. Arrows in red at each node of this lattice correspond to the spins. In this case, spins can take two values: +1 if it points upwards and -1 if points downwards.

In 1936, the Soviet physicist Lev Landau proposed a common framework to explain critical phenomena [Landau, 1936]. Despite this theory has been long debated [Menyhárd, 1970; Hargitai et al., 1972] and extended [Ginzburg and Landau, 1965], it provides a simple explanation to understand equilibrium critical phenomena and can be used to demonstrate the existence of critical exponents. As already mentioned, the different macroscopic phases of a thermodynamic system can be identified by different values of Λ for which the free-energy function attains a minimum. By considering symmetries of Λ, Landau theory proposes that, close to the critical point, the free-energy of the system can be expressed as a series expansion of the order parameter:

$$\mathcal{F}(\Lambda; T, H = 0) = \mathcal{F}_0 + a_2 (T - T_c) \Lambda^2 + a_4 \Lambda^4 , \tag{1.1}$$

where $a_2 \neq 0$, $a_4 \in \mathbb{R}^+$, the external field is considered to be zero and T_c is the critical temperature and T and H are the temperature and external field. Equilibrium states can be found by determining the values Λ^* that minimize $\mathcal{F}(\Lambda, T)$:

$$\left(\frac{d\mathcal{F}}{d\Lambda}\right)_{\Lambda^*} = 0 \Longrightarrow \Lambda^* = \begin{cases} 0 \text{ if } T > T_c \\ \pm \left(\frac{a_2}{2a_4}\right)^{1/2} (T_c - T)^{1/2} \text{ if } T < T_c \end{cases} \tag{1.2}$$

$$\left(\frac{d^2\mathcal{F}}{d\Lambda^2}\right)_{\Lambda^*} = \begin{cases} 2a_2 (T - T_c) \text{ if } T > T_c \\ 4a_2 (T_c - T) \text{ if } T < T_c \end{cases} \tag{1.3}$$

Figure 1.4 shows the free-energy of the system as a function of the order parameter Λ for different regimes in absence of external field ($H = 0$).

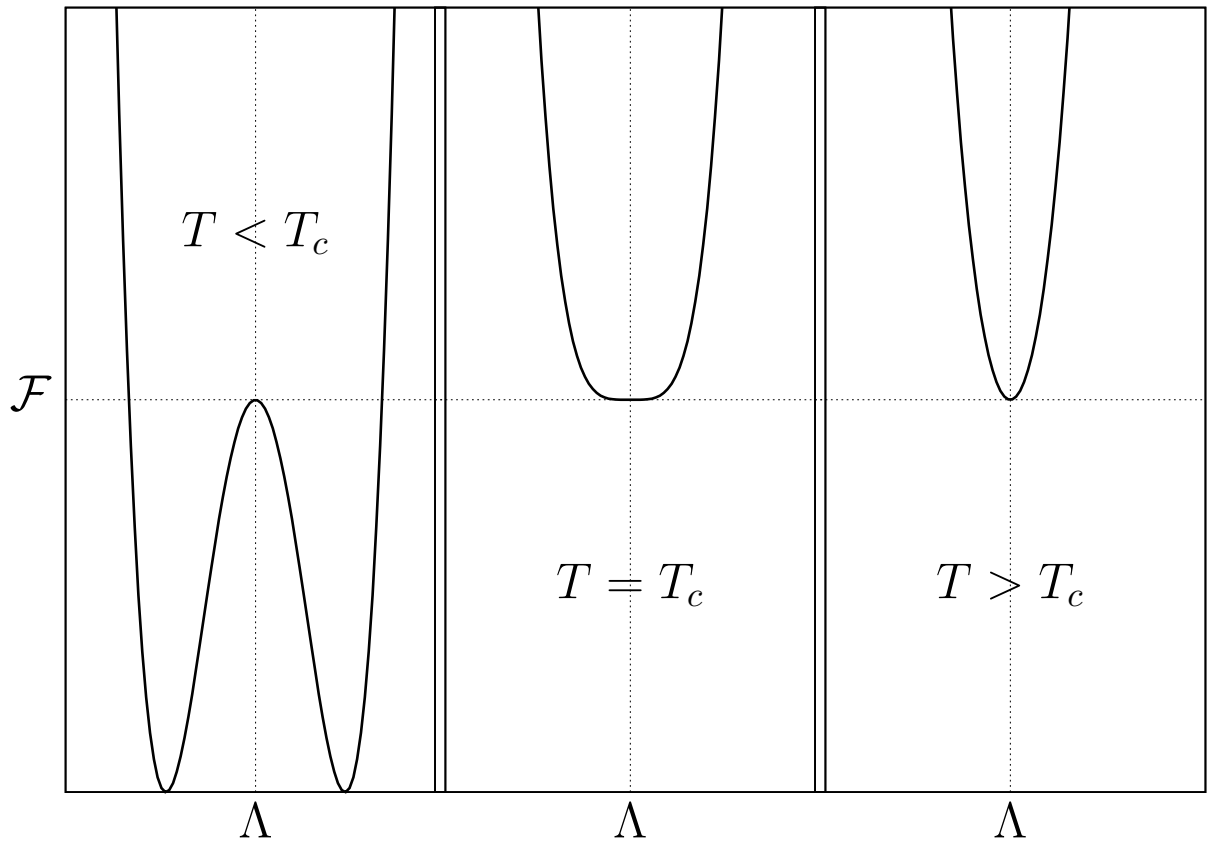

Fig. 1.4: Free-energy $\mathcal{F}$ of a physical system as a function of the order parameter Λ for three different temperatures T: subcritical ($T < T_c$), critical ($T = T_c$) and supercritical ($T > T_c$).

As it can be appreciated from Eqs. 1.2 and 1.3, below the critical temperature T_c, there are two symmetric minima corresponding to two different stables states or phases separated by a maximum (energy barrier). These different phases would correspond to liquid and gas in a hydrostatic system or states with positive and negative magnetization in a ferromagnet. The order parameter characterizing the stable states of the system approaches zero ($\Lambda = 0$) close the critical temperature with an exponent $\beta = \frac{1}{2}$ (see Eq. (1.2)). The susceptibility of the system is inversely proportional to the second derivative of the free-energy $\chi \propto \left(\frac{d^2\mathcal{F}}{d\Lambda^2}\right)^{-1}$ [Callen, 1985; Pathria and Beale, 2011]. Therefore, according to Eq. (1.3), close to T_c, the susceptibility diverges with an exponent $\gamma = 1$. These are two of the critical exponents that can be easily derived from the Landau theory. Other critical exponents can be easily found by adding external fields in Eq. (1.1) and applying other thermodynamical relations [Pathria and Beale, 2011]. This set of critical exponents correspond to the so called "mean-field universality class" [Hohenberg and Halperin, 1977]. For temperatures greater than the

critical temperature, the free-energy exhibits one single minimum at $\Lambda^* = 0$ corresponding to one stable phase.

Curiously, the values of all the critical exponents are the same for the liquid-gas transition, for magnetic systems and for the three-dimensional Ising model [Pathria and Beale, 2011; Sethna, 2006]. Regardless of their differences in structure and underlying physical phenomena, the behaviour of these systems is the same near the critical point. In this sense, these systems are considered to belong to the same universality class. The idea of universality in equilibrium relies on classifying different systems whose power-law divergences have the same critical exponents (and scaling functions [Stanley, 1999]).

The theoretical framework that encompasses criticality and universality for equilibrium systems is the renormalization group approach. This theory is much more general than Landau theory and proposes a semi-group transformation based on the homogeneity hypothesis and the scale invariance near a critical point [Kadanoff, 1966; Kadanoff et al., 1967; Wilson, 1971, 1975]. Let us suppose that the free energy can be written as a function of two dimensionless variables x,y corresponding to relative distances to the critical point of the control parameters, let us say, for instance, temperature T and external field H:

$$x = \frac{T - T_c}{T_c} ,$$
$$y = \frac{H}{k_B T} ,$$

where k_B is the Boltzmann constant providing the denominator with the same units as the field H. Note that these reduced variables at the critical point are $(x, y) = (T = T_c, H = 0) = (0, 0)$. The free-energy function can be decomposed into an analytic part $\mathcal{F}_a$ and singular (or non-analytic) part $\mathcal{F}_s$: $\mathcal{F}(x, y) = \mathcal{F}_a(x, y) + \mathcal{F}_s(x, y)$ [Christensen and Moloney, 2005]. The singular part is the one containing information about the critical exponents. The homogeneity hypothesis states that, near the critical point, this quantity can be expressed as:

$$\mathcal{F}_s \left(\lambda x, \lambda^{\vartheta} y\right) = \lambda^{-\nu} \mathcal{F}_s \left(x, y\right) , \tag{1.4}$$

where λ is an arbitrary scale factor and ϑ and ν are known as critical exponents. In other words, the singular part of the free-energy function is a generalized homogeneous function. If one proposes $\lambda \sim x^{-1}$, then Eq. (1.4) can be written as:

$$\tilde{\mathcal{F}}_s \left(1, x^{-\vartheta} y\right) = x^{\nu} \mathcal{F}_s \left(x, y\right) \Longrightarrow \mathcal{F}_s \left(x, y\right) = x^{-\nu} \tilde{\mathcal{F}}_s \left(x^{-\vartheta} y\right) , \tag{1.5}$$

where the function $\tilde{\mathcal{F}}_s$ is usually called scaling function. By a convenient choice of the scaling factor λ, any thermodynamic quantity close to the critical point can be written as a function of one single variable.

Due to the self-similarity of the correlations near the critical point, some degrees of freedom can be obviated without loosing the important information of the system. The renormalization group transform is based on a coarse-graining or decimation process, where some degrees

of freedom of the system are eliminated according to a certain rule [Kadanoff, 1966]. Once this reduction of degrees of freedom has been carried out, the variables of the decimated system are rescaled by a factor λ. By applying this transformation many times, the system can flow to a fixed point characterized by a set of critical exponents [Wilson, 1983]. The same critical exponents are found for different systems when the renormalization group transform is recursively applied. In this sense, the renormalization group represents a theoretical framework through which universality can be explained [Stanley, 1999].

1.3.3 Criticality in out-of-equilibrium phenomena

Experimental conditions in a laboratory can be controlled in order to achieve states of a physical system that can be considered to be at equilibrium. Similarly, one can consider a certain simulation algorithm for Ising models where an equilibrium state can be reached [Metropolis et al., 1953]. However, this degree of control can hardly be achieved in natural systems. Let us consider, for instance, the field of Earth Sciences and it will be extremely difficult to find a phenomenon that can be considered to be at equilibrium.

Criticality and universality are concepts which are theoretically well established for equilibrium systems. Unfortunately, there is still a theoretical gap in out-of-equilibrium theories for critical phenomena. This must be taken into account because this PhD thesis is devoted to the study of avalanche processes which are clearly out-of-equilibrium. Some works [Perković et al., 1995; Salje and Dahmen, 2014] dealing with scale-free avalanches claim that systems exhibit criticality just because the presence of power-law distributions in avalanche observables. This is in contradiction with the classical definition of criticality at equilibrium, which always considers the existence of a critical point.

Let us suppose a complex system that can be studied in equilibrium and out-of-equilibrium. This situation can be easily reproduced in Ising models [Liu and Dahmen, 2009; Lancia and Scoppola, 2013]. Given a critical point in equilibrium, the first question that comes up is: "Is this critical point altered when the system is brought out-of-equilibrium?". It is generally assumed [Salje and Dahmen, 2014] that the critical properties in equilibrium are so robust that are valid in an out-of-equilibrium scenario. However, there are still open debates for Ising models regarding whether a critical point which occurs in out-of-equilibrium conditions can be destroyed at equilibrium [Aizenman and Wehr, 1989; Spasojević et al., 2011b].

Independently on whether it exists an out-of-equilibrium critical point or not, the systems are said to be critical or at criticality when avalanche observables can be described by power-law distributions. As it occurs with the term statistical criticality, the use of the term "criticality" in out-of-equilibrium physical complex systems is a misnomer. However, taking into account this language abuse, the term "avalanche criticality" is used along this work when avalanche observables exhibit power-law distributions.

Along the last decades, different frameworks have appeared with the aim of explaining avalanche criticality. The focus of these frameworks can be diverse depending on whether they are supported by physics-based models, a plausible theoretical framework or experimental

results. It must be mentioned that the following classification does not intend to provide a strong separation and the different scenarios presented here are not mutually exclusive. The first three frameworks (self-organized criticality, disorder induced criticality and sweeping of an instability) are devoted to explain through which mechanisms avalanche criticality is originated. The rest of frameworks (driven interfaces in random potentials and failure of highly heterogeneous materials) are based upon experimental works and physics-based models. In this sense, this classification can be considered to be based on the findings of the last decades and can be easily altered and modified.

1.3.3.1 Self-organized Criticality (SOC)

Self-Organized Criticality (SOC) [Bak et al., 1987, 1988] is a theory that was developed in the late 80's of the past century, which claimed that out-of-equilibrium systems tend to a critical point by themselves. As it was explained in Sec. 1.2, there are underlying mechanisms favouring and inhibiting the presence of avalanches. SOC considers that the system is externally driven towards an active phase but there are dissipative mechanisms bringing the systems to an inactive phase. The competition between these two mechanisms is balanced in such a way that the system remains in a critical state. In this critical state, power-law distributions in the different avalanche observables are found. In this sense, criticality is reached without the necessity to tune external or internal parameters.

The sand-pile model is proposed to exhibit SOC [Bak et al., 1987, 1988; Christensen and Moloney, 2005; Pruessner, 2012] (see Fig. 1.5 for an illustration of the sand-pile metaphor). This physics-based model supposes that sand grains are added to a pile. At a certain point of this process, some parts of the pile will reach a value of the slope implying the fall of a

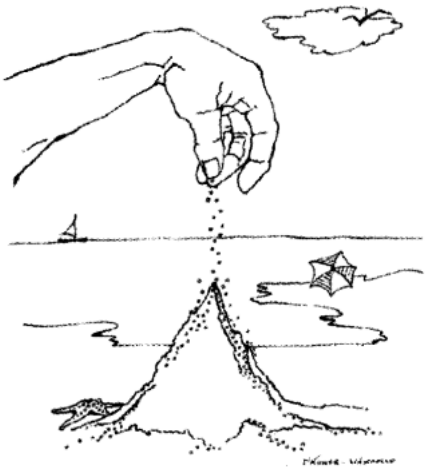

Fig. 1.5: Sand-pile metaphor. Drawing of Mrs. Elain K. D. Wiesenfeld, U.S.A extracted from Christensen and Moloney [2005].

certain number of grains due to gravity. The fall may imply that high enough values of the slope occur also in the vicinity which induces further displacements. This collective motion of grains is considered as an avalanche. Adding sand grains to the pile constitutes the driving mechanism which favours the presence of avalanches in this system. However, the friction between grains in the pile avoids avalanches to occur constantly and flatten the pile. The avalanche size S distribution follows a power-law behaviour truncated by the finite-size effects of the pile. Note that the response of the system is not linear and a perturbation can induce an event at any scale. The presence of avalanches of different sizes avoids having a very steep pile but they will not destroy the sand pile with an event comparable with the length scale of the system. In this sense, the pile is said to self-organize into a critical state.

One interesting property of SOC is that it enables the mapping between avalanches and branching processes [Alstrøm, 1988; Christensen and Moloney, 2005; Pruessner, 2012]. This mapping confers an interesting point of view studied by several authors [Corral et al., 2016; Garcia-Millan et al., 2018; Corral et al., 2018; Bordeu et al., 2019] which try to explain avalanches from the point of view of stochastic processes.

One inconvenience about SOC is that it exhibits self-censorship when one tries to apply the concept to real systems. Let us suppose earthquakes occurring in the Earth as a SOC system. The Olami-Feder-Christensen model [Olami et al., 1992; Christensen and Moloney, 2005; Pruessner, 2012] is one of the most accepted models trying to explain earthquakes in the SOC context. The Earth crust is in a critical state of stress due to two different mechanisms. One the one hand, tectonic driving tends to accumulate stress in some regions and provokes earthquakes whenever a rupture occurs. The release of stress after each earthquake occurrence prevents from very large events. On the other hand, friction between tectonic plates is a mechanism that slows down the rupture process and attenuates earthquake occurrence. In this case, it is clear that these mechanisms cannot be controlled in such a way to destabilize this so-called critical state. In other words, an empirical SOC system is said to be at a critical point although this point cannot be identified in a phase diagram, simply because it is in general not possible to bring the system out of its current state.

The SOC framework has been applied in a lot of complex systems such as atmospheric sciences [Dessai and Walter, 2000], rainfall [Deluca et al., 2015], Barkhausen noise [Cote and Meisel, 1991], seismology [Olami et al., 1992; Frigg, 2003], and so on [Jensen, 1998; Pruessner, 2012].

1.3.3.2 Disorder-induced criticality (DIC)

In the early 90's of the last century, another framework through which avalanche dynamics can be understood appeared. In contrast to SOC, disorder induced criticality (DIC) considers that the systems reach a critical state under a certain set of values of the variables that control the system. In this sense, this framework can be closely related to the principles of the equilibrium critical phenomena [Perković et al., 1995]. Different phases can be explored and an out-of-equilibrium critical point can be identified in a phase diagram. Similarly to equilibrium critical phenomena, this framework can be easily modelled by Ising models. In particular, the athermal metastable random field Ising model represents the paradigmatic model to explain DIC. For more details about this model see Sec. 2.2.2.

Let us assume an out-of-equilibrium physical system undergoing a first-order phase transition induced by an external perturbation or field H. Under these circumstances, the trajectory of the order parameter Λ as a function of the external driving field H is an hysteresis curve [Bertotti and Mayergoyz, 2006]. Below the critical temperature (or any other control variable), the Landau theory models the free-energy landscape in pure systems as a double-well with two stable states (see Fig. 1.4). For a given value of the external field, the system is able to overcome the energy barrier between the two wells and changes from one phase (one stable well) to the other one. However, this abrupt change rarely occurs in real systems and there is a coexistence in some regions where some local domains have experienced the phase transition whereas other regions have not. The presence of these domains is due to the impurities in the system that can be understood as disorder. This disorder distorts the free-energy landscape by introducing a sequence of local minima. In these local wells, the system is in a metastable state, and very small variations of the external driving field H might bring the system from one local minimum to another one. This change of the system from one metastable state to another is known as an avalanche. As it occurs in SOC, any small change of the external driving H can induce an avalanche at any scale. Temperature can also be considered as an ingredient contributing in this dynamics. In some situations, thermal fluctuations can be neglected and the external driving H is the unique factor that may provoke avalanches. In this situation, this kind of phase transitions are called athermal first-order phase transitions (aFOPT) [Pérez-Reche et al., 2001] .

The amount of disorder in the system can be characterized by a parameter R. Depending on its values, the observables characterizing avalanches appearing during the aFOPT are different. Let us assume that the disorder is quenched and it is not affected by the dynamics. Let us also assume that the system is driven infinitely slowly in such away that avalanches do not overlap each other. This regime, known as adiabatic regime, ensures that the field is varied in such a way that the system explores all the local minima of the free-energy introduced by the disorder. In this way, there is a clear separation of time scales: the one corresponding to the duration D of the avalanche, and the time δ between events ($D << \delta$). According to the classification done by Sethna et al. [2006], the avalanche dynamics of the system can be classified in the following three regimes depending on the values of R:

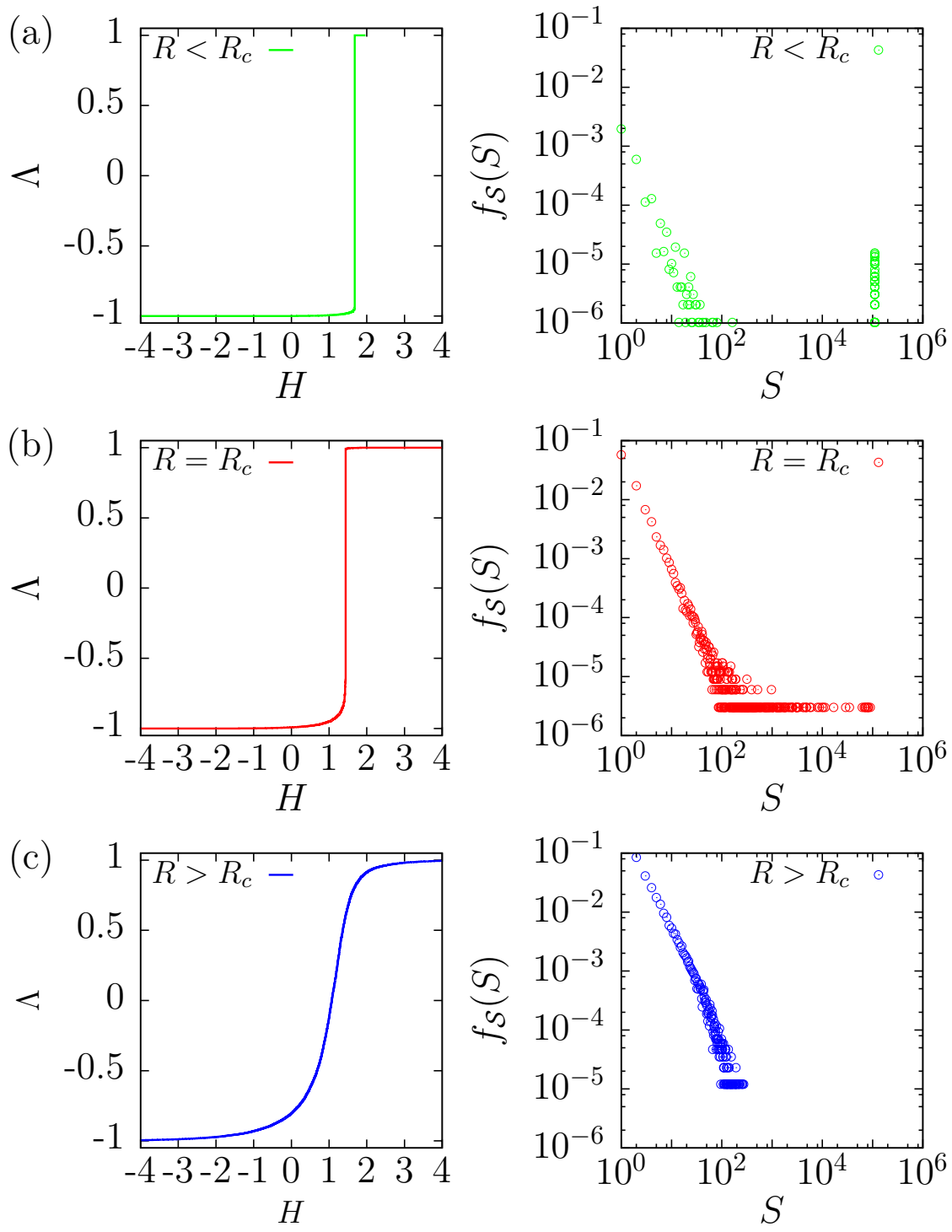

Fig. 1.6: Left panels illustrate the three regimes through which an aFOPT occurs. The order parameter Λ in this case is the particular magnetization and the phase transition goes from a phase with negative magnetization through a phase with positive magnetization driven by the external field H. Right panels, corresponding distribution of avalanche sizes S obtained from different realizations. Simulations have been done using the athermal three-dimensional Random Field Ising Model with $N = 48^3$ spins (for more details, see Sec. 2.2.2).

- $R < R_c$: Snapping noise: The amount of disorder is so small that very few distortions in the free-energy landscape appear between the two phases. The dynamics of the system evolves by means of macroscopic avalanches. The distribution of avalanche sizes will be characterized by a peak very close to the system size (see Fig. 1.6(a)).

- $R = R_c$: Crackling noise: The amount of disorder attains a critical value for which avalanche observables exhibit lack of characteristic scales. In this case, the distribution of avalanche sizes is modelled by a power-law distribution (see Fig. 1.6(b)). Note that this power-law distribution exhibits an upper cut-off due to finite-size effects.

- $R > R_c$: Popping noise: The amount of disorder is so large that a lot of local-minima have distorted the free-energy landscape. The system evolves from one phase to another by means of microscopic avalanches. In this case, the distribution of avalanche sizes can be roughly characterized by an exponential distribution (see Fig. 1.6(c)).

In spite of dealing with out-of-equilibrium systems, conceptually, DIC shows a lot of similarities with the theory of equilibrium critical phenomena. The mapping between these two theories is well supported by renormalization group approach [Dahmen and Sethna, 1996] and thus, it confers a strong theoretical background to DIC [Sethna, 2006].

This framework has been very useful in the study experimental systems such as the Barkhausen noise [Perković et al., 1995; Spasojević et al., 1996], capillary condensation of gases in porous materials [Lilly et al., 1993, 1996], structural transitions in solids [Vives et al., 1994; Carrillo et al., 1998], or superconductivity in thin films [Wu and Adams, 1995].

1.3.3.3 Sweeping of an instability (SOI)

One of the main sources of discussion between SOC and DIC is the fact of tunning of some parameter to achieve a critical state. Some authors consider that some degree of tunning is unavoidable and the tunning in a SOC model can be attributed to the nature of the boundary conditions as well as the driving [Jensen, 1998]. The DIC framework assumes that one must tune a parameter in order to observe avalanche criticality. However, SOC covers a lot of examples in nature where avalanche criticality appears without the necessity of tunning any parameter. With the aim of supporting DIC as the true mechanism to generate avalanche scale-invariance, one could think that these systems are found to be at the critical state by chance.

Sornette [1994] proposed an alternative explanation about how criticality is reached: sweeping of an instability (SOI). This framework proposes that avalanche criticality emerges from the slow sweeping of a control parameter towards an instability. Contrarily to DIC, this mechanism does not assume an underlying critical parameter which is fixed during the avalanche process. Instead of that, there exists a time-dependent interplay between interactions and disorder resulting in a non-stationary underlying critical point. This approach is somehow in between SOC and DIC. On the one hand, there are mechanisms affecting disorder

and interactions bringing the system into a critical state similarly to SOC. On the other hand, this non-stationary critical point enables to explore different parts of the phase diagram and the integrated distribution of avalanche observables along this process should be a power-law [Dahmen and Sethna, 1993]. This framework has been useful to explain several empirical data where a progressive convergence to criticality is found [Salje and Dahmen, 2014].

1.3.3.4 Driven interfaces in random potentials (DIRP)

In parallel to DIC, another scenario appeared in the mid 90s of the 20th century. Under the assumption that a d-dimensional system is divided in two different regions clearly separated by a $d-1$-dimensional geometric object called interface, this framework defends the idea that avalanches can be associated with the problem of depinning of that driven interface [Narayan and Fisher, 1993]. In contrast to DIC, note that the regions separated by the interface do not necessarily correspond to different thermodynamic phases. In this sense, the concepts of order parameter or hysteresis do not usually apply. Another difference with DIC, which is supported by Ising models, is that the description of this interface is usually supported by continuous models [Narayan and Fisher, 1993; Kardar, 1998; LeBlanc et al., 2012].

In the simplest case of a two-dimensional system, the driven interface is a continuous line. This interface is driven by an external field H in a medium where there are irregularities or pinning centres halting its advance. These pinning centres can be understood as disorder and are modelled according to random potentials. For this reason, this framework will be referred as driven interface in random potentials (DIRP) from here on out. There exists a value H_c of the external field, which is known as depinning threshold, below which the velocity of the interface is zero and it is finite for $H > H_c$. Leschhorn and Tang [1994] and Falk et al. [1994] discovered that, close to the depinning threshold H_c, the interface advances by following an avalanche dynamics with lack of characteristic scales.

Experimental examples in which the DIRP framework can be applied are fluid fronts moving in disordered media [Bramwell et al., 1998; Planet et al., 2009], crystal plasticity [Tsekenis et al., 2013] or the Barkhausen effect [Urbach et al., 1995; Narayan, 1996; Zapperi et al., 1998]. It is important to mention the case when interfaces are identified in fracture of materials with disorder [Caldarelli et al., 1996; Zapperi et al., 1997; Alava et al., 2006; Raischel et al., 2005; Kun et al., 2006; Bonamy et al., 2008] as well as the works relating DIRP and seismology [Barés et al., 2019].

1.3.3.5 Failure of highly heterogeneous materials (FHHM)

The phenomena of interest in this PhD thesis can be classified into a different framework called Failure of highly heterogeneous materials (FHHM). As it happens in DIRP, this framework is rooted in experimental works, specifically, in the failure of highly heterogeneous materials. Let us consider two different classes of materials. The first class is the one related with metallic glasses [Wang et al., 2004]. The second class encompasses materials that can be reduced to a set of physical links: foams and highly porous materials. In some cases, these materials are classified as soft-matter [Jones et al., 2002].

The term failure can be understood in several ways depending on the experimental case. One the one hand, failure can be understood as a topological rearrangement of the elements conforming the system. Given an external perturbation, the material changes its internal configuration thus minimizing the internal energy. As an example, a granular material can change the distribution of its elements under stress [Lherminier et al., 2019]. On the other hand, failure might occur at different sites where the material is breaking. When failure has occurred in this second case, the material either experiences an internal rearrangement or gets destroyed. In any of these situations, a certain amount of energy is released in each failure event and it can be experimentally detected.

The failure can be caused either by compression or shear stress [Rodney et al., 2011]. There are several ways in which failure can be detected. Failure occurs in different localized regions of the material and exhibits an avalanche-like behaviour. Whenever a failure nucleates in a region of the material, it is manifested as a sudden burst of activity for a short time. How the avalanches corresponding to failure events are detected is a problem related with the different experimental methods. These avalanches can be detected in the trajectories of the force curve exhibited as a mechanical response of the material [Antonaglia et al., 2014; Navas-Portella et al., 2016; Zreihan et al., 2018], beams of light from photo-elastic granular materials [Barés et al., 2017], particle tracking-techniques [Garcimartín et al., 2009; Garcimartín et al., 2011], or by means of acoustic emission techniques [Scruby, 1987; Salje et al., 2011; Baró et al., 2013]. In most of these experimental works, avalanche criticality is detected.

Some physics-based models such as continuous damage fibber bundle models [Hidalgo et al., 2009], micro-fracturing models [Zapperi et al., 1997] or discrete element models [Kun and Herrmann, 1999; Kun et al., 2014] have been able to describe avalanche process with power-law distributions characterizing their observables. There is a set of physics-based models which are related to the presence of a critical point [Dahmen et al., 2009, 2011]. In these models, failure of materials is considered to occur at a certain value of the stress and the distribution of avalanche observables is power-law close to this value. Clearly, these kind of models can be easily related with the concepts of equilibrium critical phenomena. However, there are situations in which the material may experience several failures and the existence of a critical point is ill-defined (see Sec. 3.3).

Despite the FHHM framework shares a lot of properties with SOC, DIC and DIRP, there are enough reasons to consider it as an independent category. Contrarily to DIRP, in general,

it is not possible to identify a clear rupture front or interface for FHHM. Given that failure might nucleate anywhere in the material, one can be tempted to think that FHHM could be understood as a nucleation process in which different domains coexist, grow and coalesce. This would place FHHM in scenario much similar to the DIC framework. However, so far, it has not been possible to identify an underlying phase diagram where a first-order phase transition can occur. One could also be tempted to think of FHHM as SOC systems in which a dissipative mechanism prevents a material from destruction and a external driving mechanism activates avalanches. However, the external driving mechanism clearly wins the competition once the material has been totally destroyed and SOC might not be an accurate description in these situations.

The phenomenology related to FHHM is the one which better fits in the context of the experimental works developed in this PhD thesis. Avalanches corresponding to failure events are detected in serrated trajectories of the force curve of the material and through acoustic emission in compression experiments of porous materials [Navas-Portella et al., 2016]. The first technique is convenient for measuring avalanches in macroscopic variables, such the force opposed by a material, where a time resolution about $\Delta t \sim 10^{-1}$ is required. Acoustic emission techniques allow to detect events at a very good time resolution ($\Delta t \sim 10^{-6}$ s) exhibiting thus a broad range of values and accessibility to different scales of avalanche observables. See Chpt.4 for more details about the experimental techniques used in this thesis.

1.3.4 Universality in out-of-equilibrium phenomena

As it occurs in equilibrium critical phenomena, different real systems and models exhibiting avalanche behaviour share the same power-law exponents. For instance, earthquakes, compressed nanocrystals [Uhl et al., 2015] and neuronal avalanches in control cultures [Yaghoubi et al., 2018] have been suggested to belong to the same universality class. Although some efforts have been done in order to characterize out-of-equilibrium universality in lattice systems [Hinrichsen, 2000; Ódor, 2004; Marro and Dickman, 2005], there is still a lack of a robust theoretical background to explain phase transitions and universality out-of-equilibrium.

It has been broadly studied [Stanley, 1999; Stanley et al., 2000; Clauset et al., 2009; Ódor, 2004] that different complex systems have a common value of the power-law exponent and can be grouped into the same universality class. Therefore, it is important to determine rigorously these exponents not only to properly characterize phenomena but also to provide a good classification into these universality classes.

When a physical system can be described in terms of a Hamiltonian $\mathcal{H}$, this contains information about the system dimensionality d and the order-parameter symmetry. These quantities determine the universality class of a system in equilibrium [Stanley, 1999]. Nevertheless, in practice, exponents are difficult to measure empirically. Some experimental works regarding physical complex systems confirm the presence of scale invariance in data by assuming power-law behaviour for which data scarcely covers few orders of magnitude [Friedman et al., 2012; Papanikolaou et al., 2011; Uhl et al., 2015].

As it has been already mentioned, despite the concept of universality can be applied for critical exponents of any power-law function, in this thesis, the concept of universality is focused on probability density functions. In this context, the broader the distribution range the more reliable the property of restricted scale invariance in experimental data. Typically the existence of noise and/or under-counting effects affects the smallest observable values, whereas saturation and/or lack of statistics due to under-sampling limits the largest observable values. In most cases, these experimental limitations are not sharp due to experimental uncertainties. These limitations distort the power-law behaviour and the property of scale invariance can only be measured in a limited range.

Chapter 2

Modelling avalanche criticality

> The method of science depends on our attempts to describe the world with simple theories: theories that are complex may become untestable, even if they happen to be true. Science may be described as the art of systematic over-simplification—the art of discerning what we may with advantage omit.
>
> *Karl Raimund Popper*

> Even the largest avalanche is triggered by small things.
>
> *Vernor Vinge*

Some generalities about avalanche processes were exposed in Chapter 1 without paying attention to the details of their description. The first section of this chapter presents how avalanche criticality can be mathematically modelled by presenting avalanche processes as arising from time series. Avalanche observables can be obtained through a transformation from time series to temporal marked point processes. As explained in the previous chapter, physics-based models have been very useful when trying to explain avalanche criticality. A set of physical features which are common in some physics-based models of avalanche processes are presented. Among the different physics-based models trying to explain avalanche criticality, the athermal Random Field Ising model is presented within the DIC framework. Finally, the main concepts of statistical modelling which are needed to understand all the techniques exposed along this PhD Thesis are exposed.

2.1 Mathematical description of avalanche criticality

The aim of this section is to provide an operational definition of avalanche processes and avalanche criticality. By operational is meant that this definition allows to determine avalanches and their corresponding observables in practice. Let us consider the framework in which the avalanche process is obtained from a time series. Let us start by exposing three different concepts:

1. Time series $\mathbb{X}(t)$
2. Avalanche process: Temporal marked point process obtained from $\mathbb{X}(t)$. The marks are characterized by a set of observables $\{\mathcal{X}\}$.
3. Statistical modelling of the observables.

The first two points correspond to a mathematical description of avalanches, i.e. how avalanches and their corresponding observables are defined. In this sense, the theory of stochastic processes, time series and point processes are specially relevant. It must be mentioned that these two points can be applied to any avalanche process independently on whether the observables are characterized by power-law distributions or not. Avalanche criticality is determined by means of statistical methods, which provide information about whether it is reasonable or not to consider power-law models for the description of avalanche observables.

2.1.1 Time series $\mathbb{X}(t)$

Let us consider a time series of a random variable $\mathbb{X}(t)$, which can take values X conforming a stochastic process or signal. The random variable can be a spatial trajectory [Villegas et al., 2019], a voltage [Scruby, 1987], the value of an asset in a financial market [Sornette, 2009] or any other quantity which can be sampled in time. The probability $\mathcal{P}$ that a system has exhibited a certain signal $\mathbb{X}(t)$ reaching a value X_t at a time t is given by the joint probability of all the intermediate steps $\{\mathbb{X}\} = \{\mathsf{X}_1, \mathsf{X}_2, ..., \mathsf{X}_{t-1}, \mathsf{X}_t\}$ it performed:

$$\begin{aligned}\mathcal{P}\left[\mathbb{X}(t)\right] &= \mathcal{P}\{[\mathbb{X}_1 = \mathsf{X}_1] \cap [\mathbb{X}_2 = \mathsf{X}_2] \cap \cdots \cap [\mathbb{X}_t = \mathsf{X}_t]\} \\ &= \mathcal{P}\left[\mathbb{X}_1 = \mathsf{X}_1\right] \cdot \mathcal{P}\left[\mathbb{X}_2 = \mathsf{X}_2 | \mathbb{X}_1 = \mathsf{X}_1\right] \cdot \cdots \\ &\cdots \mathcal{P}\left[\mathbb{X}_t = \mathsf{X}_t | \mathbb{X}_{t-1} = \mathsf{X}_{t-1}, \mathbb{X}_{t-2} = \mathsf{X}_{t-2}, ..., \mathbb{X}_2 = \mathsf{X}_2, \mathbb{X}_1 = \mathsf{X}_1\right] .\end{aligned} \tag{2.1}$$

Contrarily to the usual procedures for stochastic processes, independence and Markovian properties cannot be assumed Henceforth, all the values taken by the signal $\mathbb{X}(t)$ from $t = 0$ to t, are expressed as a new random variable called history $\mathbb{H}$:

$$\mathcal{P}\left[\mathbb{X}(t)\right] = \mathcal{P}\left[\mathbb{X}_t = \mathsf{X}_t | \mathbb{H}\right] . \tag{2.2}$$

When dealing with experimental data, it is important to stress the fact that the aim of these approaches is to describe the measured signal rather than the process as a phenomenon itself. Let us remark that this measured signal can be affected by many factors: the instrumental response of the experimental set-up, background and experimental noises, or dissipative effects.

The operational definition of avalanche can hardly be separated from the experimental techniques and data analysis that are used for their detection. Two different approaches can be considered regarding how physical systems are driven.

The first approach is based on the experimental detection of the process by assuming that signal values $\{\mathsf{X}\}$ are detected at a certain sampling rate. This sampling rate can be regular (equally-spaced times) or not and its value depends on the temporal resolution of the experimental device. Both situations can be mathematically modelled by time series. A simple random walk would be a reasonable model for a regular sampling rate [Papanikolaou et al., 2011] whereas continuous-time random walk [Klafter and Sokolov, 2011] would be an appropriate model for a non-regular sampling rate. Recall that these models are useful as illustrations but not in practice, since the history $\mathbb{H}$ of the process has an influence in the probability of the next step.

The second approach can be achieved in some experimental systems [Rosinberg and Vives, 2004], where the variable $\mathbb{X}(t)$ is obtained for a set of values of the external driving field H, not necessarily equispaced. Therefore, this approach assumes that time can be substituted by the external driving field H.

2.1.2 Avalanche process: Obtaining a temporal marked point process from $\mathbb{X}(t)$

As it has been already exposed, an avalanche process corresponds to a temporal marked point process obtained from $\mathbb{X}(t)$: a set of points distributed along the real temporal axis each of them characterized by a set of observables or marks. Each point of this temporal marked point process corresponds to an avalanche which is characterized by a set of avalanche observables $\{\mathcal{X}\}$.

The operational definition of avalanche is thus intimately tied to the collapsing procedure from a time series into a temporal marked point process. How is this collapse achieved? The answer to this question depends on the nature of the signal $\mathbb{X}(t)$.

As stated in Sec. 1.2, the words *"activity"* and *"silence"* in the definition of avalanche imply the existence of a value below which one considers that nothing is occurring and above which one considers that something is happening. The existence of a threshold is thus key in the definition of avalanche process. Two different thresholding procedures can be defined:

- **One-side thresholding** [Laurson et al., 2009; Font-Clos et al., 2015; Villegas et al., 2019]: This procedure is useful when the signal exhibits relevant features for one sense of the values that $\mathbb{X}(t)$ can acquire: let us say positive or negative values. From here

on out, let us write the values that the time series can take as $\mathsf{X}(t)$. The j-th avalanche starts at time t_j when $\mathsf{X}(t)$ crosses a threshold (THR) which is set in the preferred direction. At this point, the end of the avalanche can be defined in two ways: the avalanche finishes when the process no longer surpasses the threshold (see Fig. 2.1(a)) or at the moment when the process has undergone the threshold and it remains below it for a hit definition time (HDT). This second approach is quite common in experimental studies where it is often much suitable to talk about hits instead of avalanches (see next Sec. 2.1.4 and Chpt. 4). Once the avalanche has finished, one can define the duration D, which constitutes the first immediate observable.

The avalanche amplitude of the j-th event corresponds to the maximum of the absolute value acquired by the process along the avalanche duration D_j:

$$A_j = \max\left[|\mathsf{X}(t)|\right]_{t_j \leq t \leq t_j + D_j} \,. \tag{2.3}$$

The avalanche energy corresponds to the area under the absolute value of the process along the avalanche duration:

$$E_j = C \int_{t_j}^{t_j + D_j} |\mathsf{X}(t)| dt \,, \tag{2.4}$$

where C is a constant of proportionality providing energy units. It must be mentioned that it is also usual to define the energy by considering the square of the values of the signal (see Chpt. 4). These observables correspond to an operational definition of avalanche that takes into account not only the part of the process exceeding the THR (dark blue regions in Fig. 2.1(a)) but also the corresponding part below it (light blue regions in Fig. 2.1 (a)). An alternative definition of these observables can be stated if one considers that the avalanche corresponds just to that part of the process exceeding the threshold (dark blue areas in Fig. 2.1 (a)). Note that, for this case of one-side thresholding procedure, amplitudes and energies in this alternative operational definition of avalanche can be obtained by just subtracting the value of THR and $D_j \cdot THR$ to Eqs. (2.3) and (2.4) respectively.

- **Two-side thresholding** [Lebyodkin et al., 2013; Baró et al., 2013; Lebyodkin et al., 2017]. This situation is applied when the process $\mathbb{X}(t)$ acquires positive and negative values and any large event (in absolute value) is relevant. This situation can be typical of waveform signals such as electric voltages [Scruby, 1987]. In this case, two symmetrical thresholds are set for positive and negative exceedances of the thresholds and the process $\mathbb{X}(t)$ is also considered when it occurs in between these two values. In this case one cannot consider that the event has finished when it has undergone the THR because it can experience an excursion to an extreme value of the opposite sign. In this procedure, the presence of the HDT is crucial in order to define when the avalanche finishes. See Sec. 4.1.3 for a discussion of the dependence of the THR and the HDT in the avalanche

process in a two-side thresholding process. Note that, if one considers the absolute value of the signal, this procedure is equivalent to the one-sided trhesholding where an HDT is considered.

Independently on which thresholding procedure is used, at the end one has reduced the degrees of freedom of the original process by collapsing the information into a marked point process of N avalanches characterized by a set of observables $\{\mathcal{X}\}$:

- A time of occurrence t
- A duration D
- An amplitude A
- An energy E

One avalanche is separated by the next one by an inter-event time or waiting time δ which accepts different definitions depending on the approach one is considering. Two different approaches can be assumed [Paczuski et al., 2005]. One the one hand, avalanches can be considered to have a finite duration D. In this situation, waiting-times are defined as $\delta_j = t_{j+1} - (t_j + D_j)$. On the other hand, avalanches in the temporal marked point process can be assumed to be instantaneous ($D \to 0$). In this case, the waiting-time is defined as the difference between occurrence times: $\delta_j^* = t_{j+1} - t_j$. Note that both definitions yield the same result $\delta \simeq \delta^*$ as long as $\frac{D}{\delta} \to 0$. This situation in which time scales can be clearly separated is known as the adiabatic regime [White and Dahmen, 2003].

At this point, one has a temporal marked point process which can be described by three different specifications [Lowen and Teich, 2005; Daley and Vere-Jones, 2008]:

- **Interval specification**: When the point process is defined through the distribution of waiting times δ.
- **Counting specification**: Let us assume that the real time axis can be divided in time intervals Δt. A marked point process is defined by the distribution of the number of events N that appear in a certain time interval Δt.
- **Intensity specification**: When the point process is determined by an intensity $\varrho\left(t|\mathbb{H}\right)$ which is defined through the probability of having a point a certain time interval $t + \Delta t$. The intensity of a temporal point process is given by:

$$\varrho\left(t|\mathbb{H}\right) = \mathbb{E}\left[\lim_{\Delta t \to 0} \frac{\mathcal{P}\left(N\left(t + \Delta t|\mathbb{H}\right) = 1\right)}{\Delta t}\right], \tag{2.5}$$

where $\mathbb{E}$ corresponds to the expected value. let us remark that the this quantity can be experimentally estimated by the rate of occurrence:

$$r(t) = \frac{N(t + \Delta t) - N(t)}{\Delta t}. \tag{2.6}$$

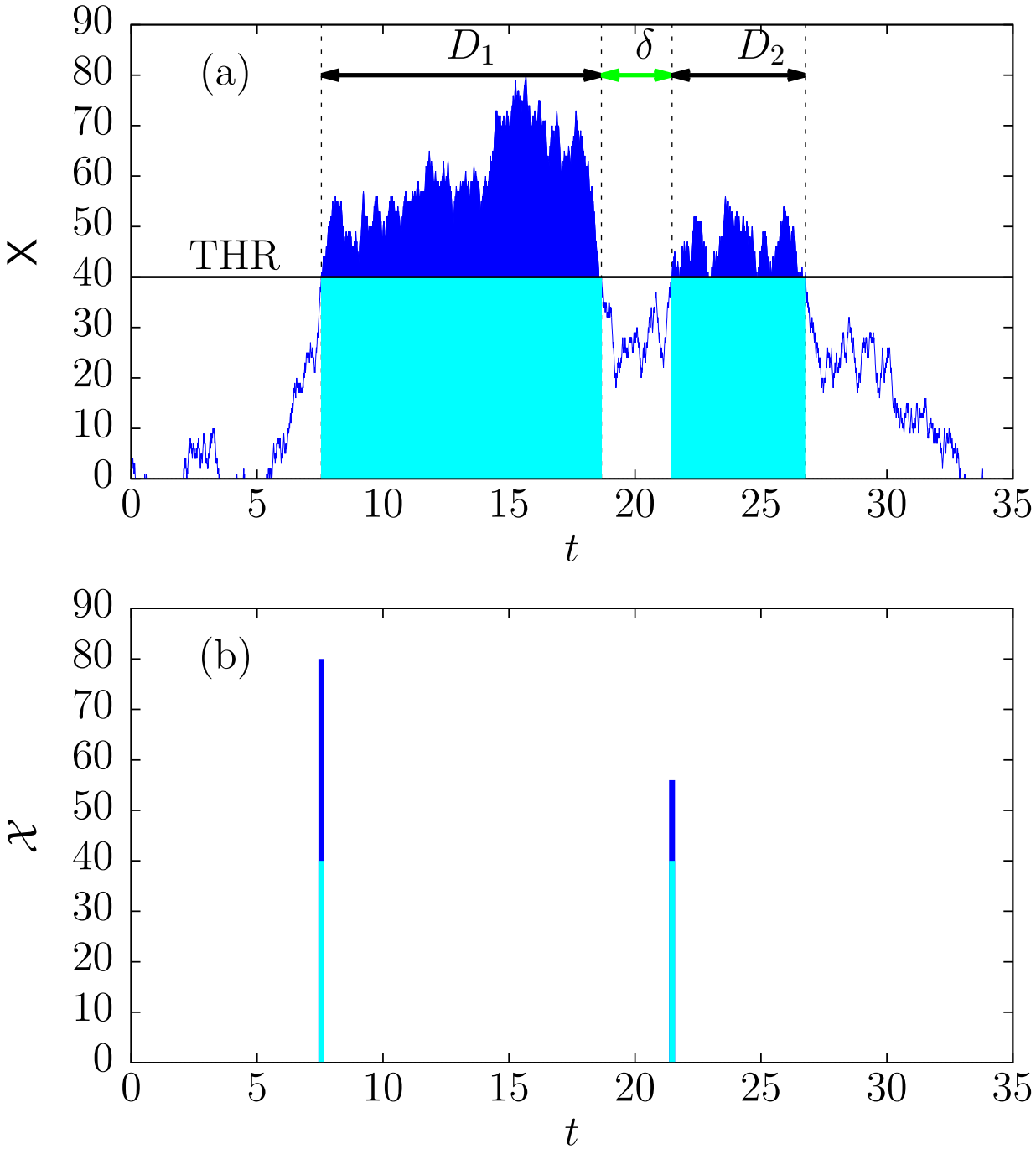

Fig. 2.1: (a) Realization of a random walk $\mathbb{X}(t)$ in order to illustrate the one-side thresholding procedure of a stochastic process. Areas in blue correspond to two different events of duration D_1 and D_2 respectively separated by a waiting-time δ. (b) Collapsing of the stochastic process in (a) into a marked point process. In this case, the marks $\mathcal{X}$ correspond to the avalanche amplitude. According to this representation on can adopt two different definitions of avalanche: (i) an avalanche is considered as that part of the signal exceeding the threshold (dark blue area) (ii) an avalanche is considered as that part of the signal which is below $\mathbb{X}(t)$ when it takes values above the threshold (THR) (the sum of dark and light blue areas).

Although it is true that the expected values $\mathbb{E}$ are the same: $\mathbb{E}[r(t)] = \mathbb{E}[\varrho(t)]$ [Daley and Vere-Jones, 2008], one must take into account that the experimental rate of occurrence can exhibit a different distribution from the intensity ϱ.

According to the interval specification, history-independent temporal marked point processes can be classified in two different families:

- Homogeneous Poisson processes: The waiting-time distribution is an exponential. Homogeneous Poisson processes assume that event occurrences are independent and identically distributed (i.i.d) and the rate of occurrence does not depend on the past history $\mathbb{H}$ and thus $\varrho(t|\mathbb{H}) = \varrho$.

- Renewal processes: When the waiting-time distribution is different from an exponential. Renewal point processes also assume that waiting times are i.i.d and independent on the history $\mathbb{H}$.

To obtain a marked point process from an arbitrary time series is a procedure that can always be applied independently on whether power-law distributions are found for avalanche observables or not.

2.1.3 Operational definition of avalanche processes and avalanche criticality

At this point, the main ingredients needed to provide an operational definition of avalanche processes and avalanche criticality have been already explained.

Operational definition of avalanche process and avalanche criticality: A driven physical system is said to exhibit an avalanche process when a temporal marked point process can be obtained from a time series $\mathbb{X}(t)$ through a thresholding procedure. Avalanche criticality occurs when the marks or avalanche observables $\{\mathcal{X}\}$ can be described by means of statistical models exhibiting lack of characteristic scales.

In contrast with the definition of avalanche process provided in the previous chapter, this definition is much more basic and less restrictive.

2.1.4 Experimental detection of avalanches

Let us remark that the operational definition of avalanche process and avalanche criticality is considered in a framework where avalanches are obtained from a time series. Note that the previous definition of avalanche process does not take into account which are the mechanisms inducing changes in the time series. One could apply the previous definition to any time series independently on the underlying phenomenon. In fact, this is what was done in Fig. 2.1, where the thresholding procedure was applied to a random walk and avalanches were defined. However, this definition presents some disadvantages when avalanches are detected in real systems. When the system is externally driven, avalanches arise as a collective behaviour and, therefore, they correspond to the phenomenon of interest. Despite the previous definition is mathematically consistent, avalanches obtained through it have no physical sense unless the system overcomes a physical threshold thus triggering the phenomenon. Moreover, not only a contribution from the avalanche is present in the time series but also noise can alter the observed $\mathbb{X}(t)$. Contributions as a consequence of the noise need to be eliminated but also different problems can appear due to driving rate effects and the detection devices.

Ideally, one could think about a driving regime where avalanches do not superpose each other, that is, avalanches do not coincide in time. Independently on whether the external driving can be controlled or not, different problems can appear when this adiabatic regime is not achieved. For instance, let us suppose that different avalanches are triggered at different spatial locations but very close in time in such a way that, when they spread along the system, they superpose both in time and space. This effect is known as avalanche overlapping [White and Dahmen, 2003]. Note that in this framework where avalanches are studied from time series, spatial information is not necessarily known. Unless one can ensure that the system is driven in the adiabatic regime, avalanche overlapping cannot be clearly identified by solely studying a time series. The measured signal can be thus a superposition of avalanches and the operational definition of avalanche provided in the previous section might be inconsistent. If this definition is applied to this signal, it is more convenient to talk about hits rather than avalanches. This is done in order to stress the fact that one measured hit does not necessarily correspond to one avalanche.

Additionally, the use of HDT in the experimental procedure (see Sec. 4.1.3) can induce two different effects: hit splitting and hit overlapping. Let us suppose that measured signal has remained below threshold for a certain time and a sudden re-intensification occurs. On the one hand, hit splitting occurs when this re-intensification takes place once HDT is fulfilled and it is considered as a separated hit. On the other hand, hit overlapping occurs when the re-intensification occurs before HDT is fulfilled and it is considered as a part of the previous hit. However, note that it is not possible to discern whether this re-intensification corresponds to a new avalanche or it is still part of the previous one. Clearly, the value of HDT is influencing the number of hits one is able to detect (see Sec. 4.1.3 for more details).

2.2 Physics-based models for avalanche criticality

2.2.1 Usual features for avalanche response in physics-based models

The aim of physical models is based in unveiling the essential factors that are needed to understand the behaviour of a physical system. Let us discuss a set of common physical features that appear in physics-based models describing avalanche processes. Despite some of these qualities are focused on the modelling avalanche processes they are somehow related with the nature of the complex system in which they take place.

- **Spatially extended systems**: A system lies in a space in which avalanches can spread out. From the physical point of view, it is clear that avalanches take place in a physical space. For instance, earthquakes occur in the Earth crust or failure occurs at a certain region of a material. However, one could also think of avalanche process that might not need physical spatial degrees of freedom: market crashes or meme diffusion in social networks [Gleeson et al., 2014] for instance. Despite these are not physical systems, they can be modelled by means of physical models [Sornette, 2009; Molins and Vives, 2016]. As long as a graph or network in which the nodes represent system constituents and the links or edges correspond to the interactions among them, the avalanche is able to spread out.

- **Interactions**: Interactions are necessary to propagate an avalanche. It is not possible to have an emergent behaviour in a complex system without interaction between the elements.

- **External driving**: The transitions between states are induced by the external driving. Ideally, the driving has to be considered in the adiabatic limit. However, it must be mentioned that one might have avalanche dynamics when this condition is not fulfilled [Corral and Paczuski, 1999]. This is what actually happens experimentally, where the system can be driven very slowly but never reaching the adiabatic limit where avalanches can be clearly discriminated. In other words, avalanche overlapping may affect the statistics but, if this effect is considerably reduced, it is possible to have avalanche processes when the temporal scale of the avalanches and the waiting-times are clearly separated. The effects of the driving have also been studied in physics-based models where avalanche criticality is found [Pérez-Reche et al., 2004].

- **Existence of a physical threshold**: At this point, it is important to highlight that there exist two thresholds. On the one hand, the external driving has to overcome a physical threshold in order to trigger an avalanche. On the other hand, due to the presence of noise, thresholding procedures are necessary in order to define avalanches from time series. Avalanches defined in this way have no physical meaning unless the

physical threshold has been overcome by the system. Therefore, the physical threshold is the one intimately related with the occurrence of an avalanche whereas the other threshold is an auxiliary tool .

It must be mentioned that these features are present in systems exhibiting avalanches, independently on whether they are critical or not. It is still not clear which are the necessary conditions in order to ensure avalanche criticality.

The first feature can be easily satisfied in all the frameworks that try to model critical avalanches explained in Chpt. 1. Additionally, systems have to be large enough so as to observe avalanche observables at different scales.

There are particular features characterizing physics-based models which are specifically focused on the different frameworks exposed in Chpt. 1. For instance, due to the similarities with the theory of equilibrium critical phenomena, some thermodynamical quantities are usually defined in models within the DIC framework. Similarly to the Landau theory of equilibrium critical phenomena, physics-based models for DIC often define an order parameter Λ characterizing the different states in which the system can be. This order parameter can be either discrete or continuous. In experimental systems, the order parameter can be measured: for instance, the magnetization in a ferromagnetic material or the densities in hydrostatic systems. Although it might not be mathematically formulated in some situations, an underlying free-energy landscape with different stable, metastable and unstable states is present in physics-based models for DIC. The existence of free-energy instabilities enables transitions from different metastable states in which avalanches can be identified. Clearly, the features of order parameter and free-energy function can be easily defined in the DIC framework due to its correspondence with the theory of equilibrium critical phenomena. However, it is not clear if these thermodynamical concepts can be extended in the rest of scenarios. Moreover, as its name indicates, disorder is an important feature in DIC and represents the system inhomogeneities and might act either by favouring the occurrence and propagation of avalanches or inhibiting them.

One of the most controversial points that faces SOC, DIC and SOI is the fact of tunning or not some parameters or variables in order to observe avalanche criticality. As it has been already mentioned, DIC exhibits avalanche criticality for a given value of the disorder whereas SOC and SOI do not need the tunning of any parameter. SOC assumes that the system self-balances the dissipative and the active mechanisms thus remaining in a critical state [Bak et al., 1987] whereas SOI assumes that the interplay between disorder and interactions evolves in time by approaching to a region close to criticality [Sornette, 1994].

For the different frameworks exposed in Chpt. 1, there exist several models reproducing avalanche criticality. Among the different models that can be attributed to the different frameworks, the athermal Random Field Ising model is explained as a paradigm of avalanche processes by paying attention in whether some of the mentioned features are fulfilled or not.

2.2.2 Athermal metastable Random Field Ising Model (RFIM)

The athermal (T=0) metastable Random Field Ising Model (RFIM) is a variant of the Ising model that has been commonly used to explain avalanche processes within the DIC framework [Sethna, 2006]. The model, originally thought to explain the Barkhausen noise in ferromagnetic materials, can be applied to describe physical systems such as structural transitions, capillary condensation of gases in porous solids (see [Sethna, 2006; Rosinberg and Vives, 2004] and references therein).

As indicated in Sec. 1.3.2, Ising models consist of a collection of N interactive spins s_i situated at the nodes of a d-dimensional lattice. Spins can take the values $s_i = \pm 1$. The order parameter Λ in this model is given by the sum of all the spin values:

$$\Lambda = \sum_{i=1}^{N} s_i \,. \tag{2.7}$$

The model contains three basic ingredients: external field, H, coupling between spins, J ($J > 0$), and random fields h_i at each node of the lattice. The Hamiltonian describing this model is:

$$\mathcal{H} = -\sum_{\langle ij \rangle} J s_i s_j - \sum_i (H + h_i) s_i. \tag{2.8}$$

The first term accounts for interaction with nearest neighbours denoted by $\langle ij \rangle$ and the second one for the interplay of spins with the external field and disorder. For simplicity, let us consider $J = 1$. Quenched random fields $\{h_i\}$ are independent and Gaussian distributed according to:

$$f_h(h) = \frac{1}{\sqrt{2\pi R^2}} exp\left(-\frac{h^2}{2R^2}\right) , \tag{2.9}$$

where the mean is zero and the standard deviation R characterizes the amount of disorder present in the system. In order to generate a metastable dynamics, it is necessary to establish a criterion to determine under which conditions the system remains in a local minimum of the energy landscape. Following Sethna's rule [Sethna et al., 1993], a spin is stable when it is aligned with its effective field:

$$h_i^{eff} = \sum_j s_j + H + h_i \,, \tag{2.10}$$

where the sum extends over the nearest neighbours of the i-th spin. Following this deterministic rule, a spin flips when its local field changes sign. A flipping event changes the effective field of its nearest neighbours and any of these could become unstable. The process continues in the same way analysing successive shells and originating an avalanche of flipping spins. The size of the avalanche corresponds to the number of spins that have changed their state for a fixed value of the external field H:

$$S = \frac{1}{2}\Delta\Lambda \,,$$

where Δ corresponds to the difference experienced by the order parameter after the avalanche occurred at H. The duration corresponds to the number of steps (shells in the RFIM literature) needed to complete the avalanche [Dahmen and Sethna, 1993; Frontera and Vives, 1999]. Avalanches can be identified in the discontinuities in Λ for a fixed value of H. The properties of these avalanches strongly depend on the quenched disorder, determining which sites of the system are more favourable to nucleate or not.

This model can be easily interpreted in the context of the Landau theory of phase transitions. Consider the case $T < T_c$ in Fig. 1.4 and let the system be in one of the two stable minima. In absence of disorder, the system will experience a first-order phase transition when the value of the external field H is large enough to overcome the energy barrier separating both minima. However, disorder plays an important role by distorting the free-energy landscape and introducing a sequence of local minima separated by high energy barriers between them. The reason why a normal distribution with zero mean is chosen for the random fields is due to the fact that disorder is not supposed to favour any of the stable states. During the field driven process, these free energy barriers are so large that thermally activated events are negligible. Therefore, the external field is the only responsible for the behaviour of the system. In the adiabatic limit, the external field H drives the system from one stable state to the other by jumping through the different local minima. In each of these minima, the system remains in a metastable state and the jumps between them are identified as avalanches.

The model exhibits three different regimes as a function of the value of the disorder R. Let us consider a two-dimensional RFIM to explain the different regimes but remark that it also applies for higher dimensions. The RFIM exhibits a subcritical regime for $R < R_c$, where there is an infinite avalanche which is usually spans the system in both vertical and horizontal directions (see Fig. 8.11 sequence a-b-c that corresponds to increasing field values). When disorder approaches its critical value R_c, there is a peak in the average number of avalanches spanning one system direction [Spasojević et al., 2011a] (see Fig. 8.11 sequence d-e-f). Power-law distributions are found in some magnitudes related to avalanches in this situation (sizes and durations). For the supercritical regime $R > R_c$, the avalanche process takes place by nucleation of small domains that grow and coalesce and spanning avalanches hardly appear (see Fig. 8.11 sequence g-h-i).

Let us check whether the set of conditions stated in Sec. 2.2.1 are fulfilled by this model:

- **A spatially extended system** ✓In general, this model is usually thought in a d-dimensional regular lattice although many other configurations can be considered.
- **Interactions** ✓They are represented by the first term in the Hamiltonian in Eq. (2.8).
- **External driving** ✓The external field H.
- **Existence of a threshold** ✓Each spin has its particular threshold given by its effective field in Eq. (2.10). The effects of the threshold on RFIM avalanches have been studied in [Janićević et al., 2018].

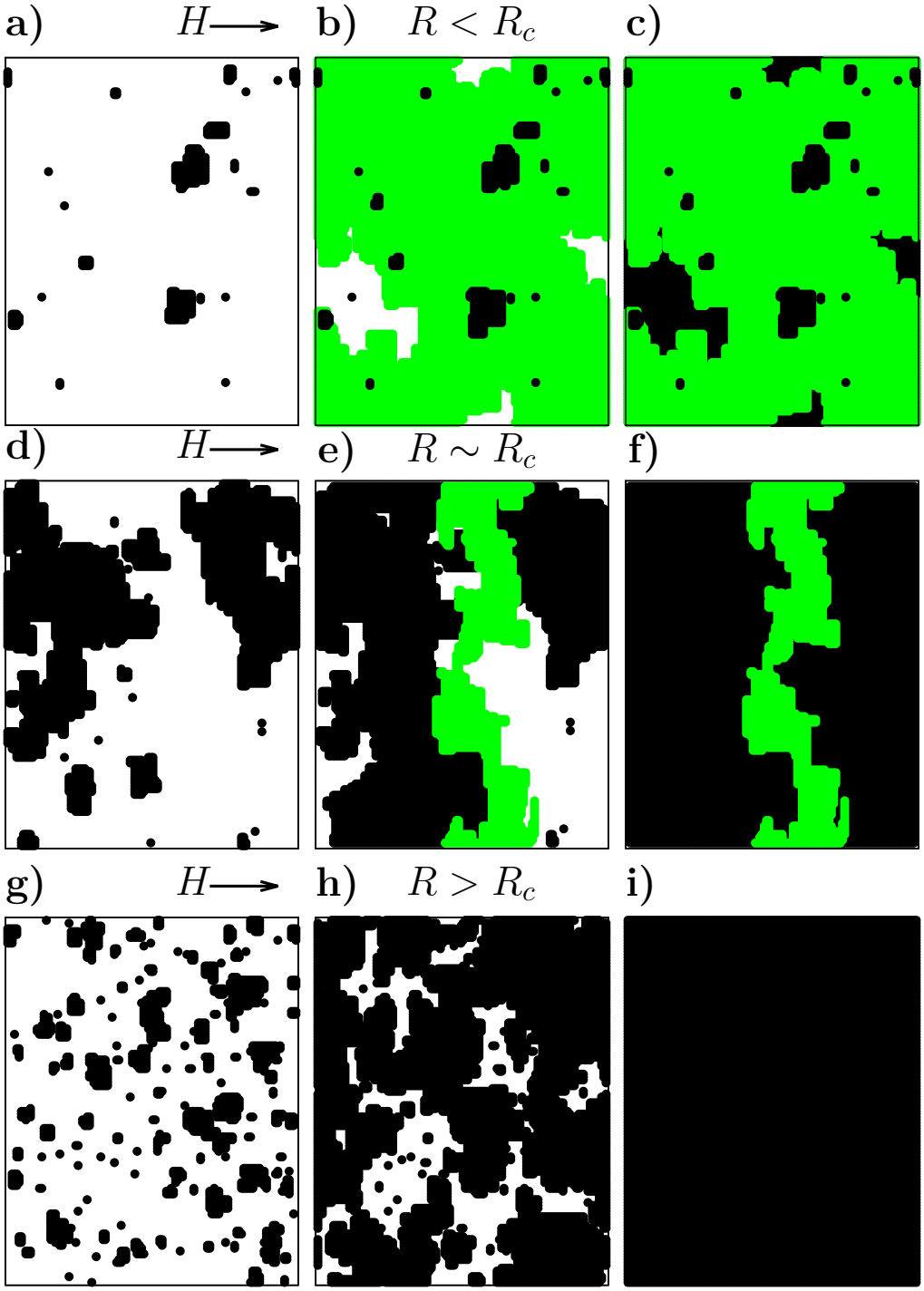

Fig. 2.2: Sequence of configurations during the avalanche process in different disorder regimes for the RFIM with metastable dynamics for a system ($N = 128 \times 128$) with periodic boundary conditions. The values of the disorder are $R = 0.80, 0.95, 1.20$ for sequences (a-b-c),(d-e-f) and (g-h-i) respectively. External field H is increased from left to right bringing the system from negative magnetization (white regions) to positive magnetization (coloured regions). Black coloured spins correspond to regions that have been transformed due to avalanches that have not spanned the system in any direction. Green regions correspond to spanning avalanches.

The model is relatively easy to implement computationally by means of the Brute force and the Sorted list algorithms [Kuntz et al., 1999]. Simulations start in one of the stable minima (for instance, with all the spins pointing down $\{s_i = -1\}$) and external field set at an extremely large value ($H = -\infty$ if all spins point down at the beginning). The external field is increased until a spin triggers an avalanche. Then the external field is kept constant until there are no more unstable spins and the system reaches a metastable state again. The dynamical process finishes when the system reaches the other stable well (all the spins are pointing up $\{s_i = +1\}$ for a very large positive value of the field H).

This model is rich enough to address different studies regarding avalanche criticality. The effect of the external driving [Pérez-Reche et al., 2004], the range of the interactions [Pérez-Benitez et al., 2009], the connectivity [Sabhapandit et al., 2000; Detcheverry et al., 2005; Shukla and Thongjaomayum, 2017; Janićević et al., 2017], the effect of the finite temperature [Illa et al., 2006] are some examples of studies that can be performed by means of the RFIM. The author of this PhD thesis has contributed to the research topic of the RFIM by studying the influence of the aspect ratio and boundary conditions on finite-size scaling functions [Navas-Portella and Vives, 2016] and the crossover from two-dimensional to three-dimensional systems [Spasojević et al., 2018][1].

2.3 Statistical modelling of avalanche criticality

The confluence point of physics-based models and the mathematical description of avalanches exposed in Sec. 2.1 is the statistical modelling of avalanche observables. Determining whether these observables are characterized by power-law distributions is the definitive step to check whether a system exhibits avalanche criticality or not. Furthermore, as it has been claimed in Chpt. 1, the accurate determination of critical exponents is crucial in order to perform a correct classification of complex systems into different universality classes.

Let us introduce the common situation in which one wishes to determine the power-law exponent of a probability distribution characterizing an avalanche observable. Let us assume that this observable can be considered as a continuous random variable $\mathcal{X}$ characterized by an univariate probability density function (PDF) $f_{\mathcal{X}}(x)$, where the sub-index denotes the random variable to which it is referred and x are the values that the random variable can take. One can obtain an estimation of $f_{\mathcal{X}}(x)$ by plotting the histogram of x (see Fig. 2.3 and Sec. 5.3). A power-law PDF looks like a straight line when it is plotted in log-log scale [Christensen and Moloney, 2005]. Figure 2.3 shows the estimated PDF of the observable $\mathcal{X}$ sampled from a certain probability distribution. As it can be seen, this estimated PDF looks like a straight line when it is plotted in log-log scale. In this situation, one could be tempted to fit a power-law distribution by means of the linear least squares method and identifying the power-law exponent with the slope of the straight line. Although this has been a common

[1] This work was performed in collaboration with the University of Belgrade.

practice for years within physics and other scientific disciplines [Goldstein et al., 2004], a power-law distribution is not guaranteed by the simple presence of a straight line in log-log scale. In fact, data in Fig. 2.3 have not been sampled from a power-law distribution but from a truncated log-normal. Moreover, in case of really dealing with a power-law distribution, the exponent measured as the slope of the straight line could be biased [Bauke, 2007; Clauset et al., 2009; Deluca and Corral, 2013]. Note that this approach is based on the graphical estimation of the PDF and the value of the estimated parameters might depend on the binning procedure that has been used to construct the PDF. In conclusion, one can be easily fooled when power-law exponents are estimated from graphical representations and more rigorous statistical techniques are required.

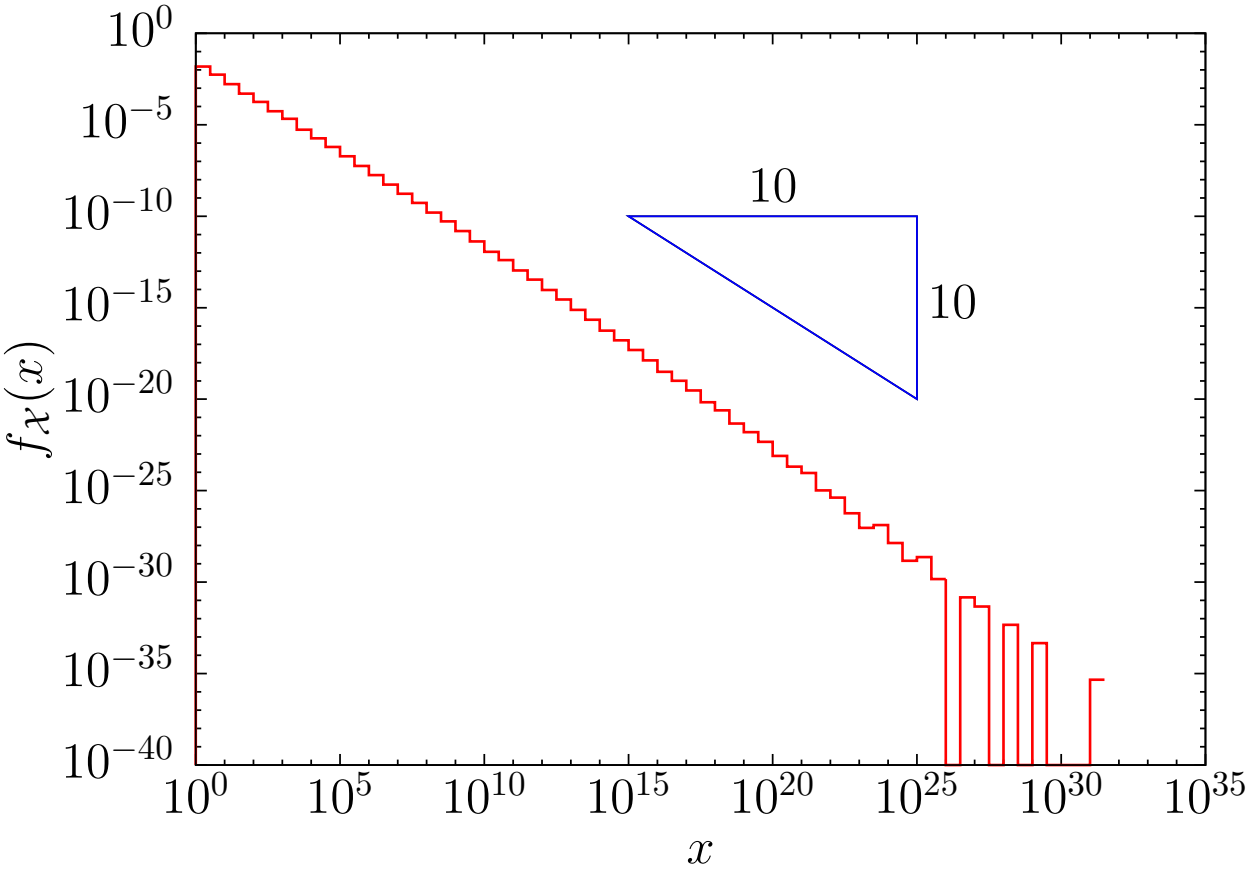

Fig. 2.3: Probability density function (PDF) for simulated values x of the random variable $\mathcal{X}$ in log-log scale. Similarly to what some authors do when estimating power-law exponents [Corominas-Murtra et al., 2015; Uhl et al., 2015; Scholz, 2019], a triangle is plot in order to illustrate the apparent slope of the PDF. Despite one could identify a straight line, data have not been sampled from a power-law distribution but from a log-normal with parameters $\mu = 10$ and $\sigma = 15$ truncated at $x_{min} = 1$.

One of the most accepted ways of fitting power-law distributions is through the maximum likelihood estimation and hypothesis testing [Knopoff and Kagan, 1977; Clauset et al., 2009; Deluca and Corral, 2013]. Contrarily to the least-squares method, it is not based in any graphical representation of the PDF. Maximum likelihood estimation provides a rigorous statistical technique supported by a mathematical background. Additionally to the estimation of the power-law exponent, goodness-of-fit tests are necessary in order to determine whether data have been really sampled from a power-law distribution or not.

In this section, a general overview of the different stages of statistical modelling is presented. Let us remark that, although the distributions of interest in avalanche processes are power-laws, the concepts explained below are general.

2.3.1 Statistical inference

2.3.1.1 General definitions

Prior to expose the different steps of statistical modelling, let us define some important concepts and approaches that are going to be recurrently used along this PhD Thesis. The starting point is rooted on a given dataset containing avalanche observables. This dataset has been sampled from an experiment, a computer simulation or any other data acquisition technique. Henceforth, this set of observations is referred as dataset or sample indistinctly.

An usual framework to understand statistical modelling is based on the existence of an idealized "true" probability distribution generating data [Burnham and Anderson, 2002; Pawitan, 2013] characterized by an array of true parameters $\tilde{\boldsymbol{\theta}}$. Let us suppose a sample of n i.i.d observations of the random variable $\mathcal{X}$ from this population. If not otherwise stated, let us assume that observables $\mathcal{X}$ are continuous random variables. The statistical model $F_{\mathcal{X}}\left(x;\boldsymbol{\theta}\right)$ is a mathematical object characterized by a set of estimators $\boldsymbol{\theta}$ that establishes a link between the sample and the population. The object of the estimation theory explained in the following Sec. 2.3.1.3 is devoted to find the values of $\boldsymbol{\theta}$ from the available information in the sample in order to evaluate the true parameters $\tilde{\boldsymbol{\theta}}$.

As it has been mentioned, one of the most important mathematical characterizations of a statistical model is, if it exists, the probability density function (PDF) $f_{\mathcal{X}}(x,\boldsymbol{\theta})$. Let us remark that this description is reduced to uni-variate distributions. From the uni-variate PDF one can define the cumulative distribution function (CDF) $F_{\mathcal{X}}(x,\boldsymbol{\theta})$ as:

$$F_{\mathcal{X}}\left(x;\boldsymbol{\theta}\right)=\mathcal{P}\left(\mathcal{X}<x\right)=\int_{-\infty}^{x}f_{\mathcal{X}}\left(x';\boldsymbol{\theta}\right)dx'\,, \tag{2.11}$$

which will be named as distribution or cumulative distribution indistinctly. One can also define the complementary cumulative distribution function (CCDF) written as:

$$S_{\mathcal{X}}\left(x;\boldsymbol{\theta}\right)=\mathcal{P}\left(\mathcal{X}\geq x\right)=1-F_{\mathcal{X}}\left(x\right)=\int_{x}^{\infty}f_{\mathcal{X}}\left(x';\boldsymbol{\theta}\right)dx'\,, \tag{2.12}$$

where S is a notation quite used in the literature which refers to the CCDF as survivor function [Kleinbaum and Klein, 2012].

2.3.1.2 Statistical models and likelihood

The concept that links data and the statistical model is the likelihood function. The likelihood function can be naively understood as the probability that the actual data can be described in terms of a statistical model. In other words, if the following question is addressed: "What is the probability that data are described by a certain statistical model?", the answer to this question is given by the likelihood function $\mathcal{L}(\boldsymbol{\theta})$ [Pawitan, 2013; Wood, 2015]. This function depends on the array of estimators, which can be obtained from the sample values. Therefore, by assuming a statistical model $F_{\mathcal{X}}$ characterized by a set of $n_{\mathcal{L}}$ unknown parameters $\boldsymbol{\theta}$, the likelihood function is the probability of the observed data $\{x\}$ considered as a function of $\boldsymbol{\theta}$. For more details about the likelihood function see Sec. 5.1.

Let us define the likelihood ratio between two statistical models (1) and (2):

$$\mathfrak{R} = \frac{\mathcal{L}_2(\boldsymbol{\theta}_2)}{\mathcal{L}_1(\boldsymbol{\theta}_1)}\,, \tag{2.13}$$

which measures the relative probability of models (1) and (2). The models might be represented by the same function but with different values of the parameters or their functional behaviour might be completely different. This quantity is crucial when doing model comparison (see next Sec. 2.3.1.4).

2.3.1.3 Estimation theory and maximum likelihood

For any model one is working with, one can evaluate the likelihood function which is tied to the statistical model. This model is indeed characterized by a set of values $\boldsymbol{\theta}$ which are intended to be similar to those from the population $\tilde{\boldsymbol{\theta}}$. Estimation theory deals with the different methods through which this relationship can be well established.

Let us define a statistic or sample statistic z as the value returned by an application $\mathfrak{Z}$ that receives the sample as an array of size n:

$$\mathfrak{Z} : \mathbb{R}^n \to \mathbb{R}\,. \tag{2.14}$$

Note that the statistic z is indeed a random variable which is characterized by a PDF $f_{\mathcal{Z}}(z)$. Descriptive statistics is generally focused on determining different statistics trying to characterize a sample. The sample mean, the sample median or the sample standard deviation are well-known examples of statistics.

Let us define an estimator $\hat{\theta}_n$ as an statistic which does not depend on any other unknown parameter of the statistical model and it is entirely determined by the n values of the sample.

An estimator is unbiased if its expected value (denoted by $\mathbb{E}$) corresponds to the value of the true parameter:

$$\mathbb{E}\left[\hat{\theta}_n\right] = \tilde{\theta}\,.$$

An estimator is said to be consistent as long as:

$$\lim_{n\to\infty} \hat{\theta}_n \to \tilde{\theta}\,,$$

where $\tilde{\theta}$ is the true parameter and the convergence is in probability. Independently on the methodology used to calculate them, it is desirable that estimators fulfil both properties.

The method of moments assumes equalities between population and sample moments at successive orders [Chave, 2017]. Sample and population moments are equated order by order until there exists a solvable system of equations that can be uniquely determined by sample moments. This method is very useful since some analytic results can be easily found but it leads to consistent but often biased estimators [Chave, 2017].

The maximum likelihood estimation (MLE) is a methodology which considers that the estimators correspond to those values of $\boldsymbol{\theta}$ that maximize the likelihood function $\mathcal{L}$. This methodology is broadly accepted in statistics and provides asymptotically unbiased and consistent estimators, at least for models from the exponential family (see [Pawitan, 2013] p.240). The theory of MLE is explained in Sec. 5.1.

2.3.1.4 Model comparison

At this point, one has the set of estimators referred to one or two statistical models $F_{(1)}\left(x;\boldsymbol{\theta}_1 = \hat{\boldsymbol{\theta}}_1\right)$ and $F_{(2)}\left(x;\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2\right)$ maximizing their respective likelihood functions $\mathcal{L}_1$ and $\mathcal{L}_2$. These statistical models are referred to as fits. Determining which of both models is more suitable to fit data can be considered either before or after estimating the parameters $\hat{\boldsymbol{\theta}}$. Methodologies to determine the model *a priori* usually use the fact that a different functional behaviour of the probability distribution is detected. As an example, one can think about a sample exhibiting a power-law distribution in a certain range but, at a certain value, deviations from this functional behaviour appear. In this case, more parameters have to be added in order to exclude these deviations.

Once the parameters maximizing the likelihood function of both models have been determined, one has two different descriptions of the sample from which different properties from the population can be inferred. There are several techniques trying to determine which model is more suitable to describe data. Among the most popular, the Akaike information criterion (AIC) and the likelihood ratio test (LRT) are explained in Secs. 5.7 and 5.6 respectively. Specifically, the likelihood ratio test is a useful tool when one desires to compare a simpler against a more complex model through the likelihood ratio in Eq. (2.13).

As an example, let us consider a sample with the height of a group of people. One could consider a normal distribution as a plausible and simple statistical model. This PDF is

characterized by two parameters: the mean and the standard deviation. However, for this particular example, one could reasonably argue that extra parameters are needed in order to truncate this distribution. One the one hand, it is not possible to have people with zero or negative height. On the other hand, let us say that it is very unlikely to observe a human being whose height is 10 metres. In this sense, one could add more complexity to the statistical model by adding more parameters. Remark that complexity in this sense implies adding more parameters to the statistical model and its meaning is rather different from the one used in the last chapter. Let us consider a simpler model $F_{Sim}\left(x;\boldsymbol{\theta}_{Sim}\right)$ characterized by $n_{\mathcal{L}}^{(Sim)}$ parameters and a more complex model $F_{Com}\left(x;\boldsymbol{\theta}_{Com}\right)$ characterized by $n_{\mathcal{L}}^{(Com)}$ parameters ($n_{\mathcal{L}}^{(Com)} > n_{\mathcal{L}}^{(Sim)}$). As it occurs in this example, it is usual that the simpler model is a limit case of the more complex one. In this case, it is said the the simpler model is nested into the more complex one. It is obvious that the more parameters in the statistical model, the higher will be the likelihood. However, at a certain point, adding more parameters will not have a significant improvement. The limit case results in overfitting: the statistical model is in perfect agreement with data but it is not useful to infer information from the population. For instance, if the probability (given by the simpler statistical model) to observe an individual higher than 5 metres is extremely small, the fact of adding an upper-truncation at this value to the simpler model will not bring a significant improvement. Despite the likelihood might be higher, the cost of adding one more parameter does not offer a better explanation. As a rule of thumb, the Occam's razor criterion is applied and, if there are two ways to describe data, the simpler will be the preferred one. For more details about the likelihood ratio test, see Sec. 5.6.

At the end of these procedures, independently on whether model selection has been done previously or after parameter estimation, one has a unique fit as a candidate. Note that this does not imply that this model corresponds to that from the population from which the dataset has been sampled from. In order to determine whether it is reasonable or not, on has to perform a goodness-of-fit test (see next section and Sec. 5.4 for more details).

2.3.1.5 Statistical hypothesis testing

Statistical hypothesis testing is a procedure within the context of statistical inference which tries to test the validity of a statement. One can assume that the population distribution or some of its parameters are unknown. Let us consider a common framework in which hypothesis testing deals with statements referred to these unknown parameters.

The general procedure relies on a statement, usually called null hypothesis H_0. The main purpose of statistical hypothesis testing is to determine whether this statement or hypothesis is reasonable or not against an alternative hypothesis H_1 based on the information in the sample or samples. There are two different modalities of hypothesis tests depending on how the null hypothesis is formulated. A test is one-sided when the alternative hypothesis is uni-directional. For instance, let us consider as a null hypothesis H_0 that the population mean $\tilde{\mu}$ is greater or equal than a certain value μ_0, $\mathrm{H}_0 : \tilde{\mu} \geq \mu_0$. The alternative hypothesis can just

take the opposite direction, which is $\mathrm{H}_1 : \tilde{\mu} < \mu_0$. A test is two-sided as long as the alternate hypothesis can take two different directions. For instance, in this case let us slightly change the previous null hypothesis by this statement: $\mathrm{H}_0 : \tilde{\mu} = \mu_0$. In this case, the alternative hypothesis can take two different directions: $\mathrm{H}_1 : \tilde{\mu} < \mu_0$ or $\mathrm{H}_1 : \tilde{\mu} > \mu_0$.

Once the null and alternative hypothesis are established, one has to impose a significance level α to the statistical test. Rejecting H_0 when it is actually true is known as type I error or false positive [Chave, 2017]. The significance level corresponds to the probability of obtaining a false positive:

$$\alpha = \mathcal{P}\left(\text{Rejecting } \mathrm{H}_0 | \mathrm{H}_0 \text{ is true}\right) . \tag{2.15}$$

A significance $\alpha = 0.05$ implies that the investigator risks that the null hypothesis will be rejected in just one out of twenty samples when it was actually true. This corresponds to a confidence level $CL = 1 - \alpha = 0.95$.

One could also be in a situation in which no statistical evidences are found to reject H_0 when it is actually false. The non-rejection of the null hypothesis when it is false is known as type II error or false negative, whose probability β is:

$$\beta = \mathcal{P}\left(\text{Non-Rejecting } \mathrm{H}_0 | \mathrm{H}_0 \text{ is false}\right) . \tag{2.16}$$

The power of the test is the ability of rejecting the null hypothesis when it is false. The power is thus given by the complementary probability of β:

$$1 - \beta = \mathcal{P}\left(\text{Rejecting } \mathrm{H}_0 | \mathrm{H}_0 \text{ is false}\right) . \tag{2.17}$$

The desirable situation for statistical testing is that in which the power is large and the significance level is small. In this context, the Neyman and Pearson lemma [Neyman et al., 1933] states that, given a sample $\{x\}$ from a statistical model F characterized by the parameters $\boldsymbol{\theta}$ and likelihood function $\mathcal{L}(\boldsymbol{\theta})$, one can formulate the next null and alternative hypotheses:

$$\mathrm{H}_0 : \ \boldsymbol{\theta} = \boldsymbol{\theta}_1 \tag{2.18}$$

$$\mathrm{H}_1 : \ \boldsymbol{\theta} = \boldsymbol{\theta}_2 \tag{2.19}$$

Both hypotheses are thus characterized by two likelihood functions $\mathcal{L}_1(\boldsymbol{\theta}_1)$ and $\mathcal{L}_2(\boldsymbol{\theta}_2)$ and a likelihood ratio $\mathfrak{R}$ can be computed. The hypothesis test with the highest power for a given significance level α must fulfil:

$$\mathfrak{R} > C \text{ for each sample inside the rejection region} , \tag{2.20}$$

and

$$\mathfrak{R} \leq C \text{ for each sample outside the rejection region} , \tag{2.21}$$

where C is a positive constant which fulfils $\mathcal{P}\left(\mathfrak{R} > C | \mathrm{H}_0 \text{ is true}\right) = \alpha$. In other words, among all the statistical tests assuming a significance level α, the likelihood ratio test is the one

maximizing the power. Therefore, the best rejection region of size α is provided by this lemma.

Provided that the best rejection region is defined, the next step in the statistical hypothesis-testing consists in choosing a test statistic. At this point, the sample or samples play the important role. The test statistic can be computed independently on whether its distribution is known or not. However, it is preferable to choose a statistic whose theoretical distribution is known. Nevertheless, sometimes the conditions are not supported by any theoretical background and the distribution of the statistic may be estimated by means of Monte Carlo simulations (see Sec. 5.4.1).

Once the statistic has been determined from the sample or samples information, a decision must be taken in order to reject or not the null hypothesis. Given a significance level α, the test statistic can acquire a value either inside or outside the rejection region, where the null hypothesis is either rejected or not rejected. The rejection region is delimited by the critical values of the statistic, which can be obtained by the percentiles of its distribution. Note that the word critical in this context has nothing to the with the concepts of criticality exposed in the previous chapter. This decision is taken either based on the critical value of the statistic or on the p-value (p_{value}). Under the assumption that the null hypothesis is true, the p-value is the probability to find a test statistic as extreme as the one that has been observed. In this sense, the p-value corresponds to a quantitative measure of statistical significance. If the distribution of the test statistic is known, the p-value can be easily found. A p-value smaller than the significance level α implies the rejection of H_0. In this case, it is said that the result is statistically significant. Coming back to the example of $\alpha = 0.05$, assuming that H_0 is true, finding a $p_{value} < \alpha$ implies that the probability to find a value of the statistic as extreme as the one in the sample is less than a 5%. Therefore, the value of the statistic obtained from that sample is considered unlikely unlikely and the null hypothesis is rejected.

Note that the non rejection of the null hypothesis does not imply in any case that it is a true statement. The non-rejection means that there is no enough statistical evidence to reject H_0. In any case, the non-rejection implies that the null hypothesis is accepted. The only conclusive result is the one rejecting the null hypothesis and accepting the alternative. The same null hypothesis H_0 can be tested by using different statistics leading to different p-values. If one of them is below the significance level α, this is enough statistical evidence to declare that H_0 is rejected.

The hypothesis testing can be summarized by the following points:

1. Formulate the null and alternative hypothesis.
2. Establish a significance level α.
3. Select and compute a statistic z.
4. Compute the p-value (p_{value}) either from its analytical distribution or by Monte-Carlo simulations.

5. Decision based on p_{value}:

   - $p_{value} < \alpha \Longrightarrow \mathrm{H}_0$ rejectable.
   - $p_{value} \geq \alpha \Longrightarrow \mathrm{H}_0$ non-rejectable.

One important application of statistical hypothesis testing along this PhD Thesis relies on the goodness-of-fit test. This test has a null hypothesis which states that the fit $F_{\mathcal{X}}\left(x;\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\right)$ which has been proposed to describe data is actually the same as the statistical model characterizing the entire population. If not enough evidences are found to reject this null hypothesis, then $F_{\mathcal{X}}\left(x;\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\right)$ is considered to be a plausible statistical model. For more details about the goodness-of-fit test, see Sec. 5.4.

Hypothesis testing is also applied to compare different power-law exponents (Sec. 5.9), whether different datasets have been sampled from the same population (Sec. 5.8) and it is fundamental piece for the tools for model comparison in Sec. 2.3.1.4.

2.3.2 Summary of the general procedure to characterize the distribution of an avalanche observable

Let us assume a sample of size n of an avalanche observable $\mathcal{X}$, e.g. avalanche energy. The distribution $F_{\mathcal{X}}\left(x;\boldsymbol{\theta}\right)$ characterizing this observable is unknown. Due to the relevance of power-law distributions that has been discussed so for, the parameters in $\boldsymbol{\theta}$ should give information about the power-law exponent and the range in which this exponent holds. The range is important since the property of scale invariance for empirical datasets is stronger when the range with lack of characteristic scales is broader. The importance of the exponent is capital because it determines the universality class to which the system belongs to.

Let us expose the general procedure:

1. **Model proposals**: A set of statistical models exhibiting lack of characteristic scales is proposed. The first model is simpler and is characterized by a number $n_{\mathcal{L}}^{(Sim)}$ of parameters.

   $$F_{Sim}\left(x;\boldsymbol{\theta}_{Sim}\right) . \tag{2.22}$$

   The second model is more complex in the sense that it is characterized by $n_{\mathcal{L}}^{(Com)}$ (with $n_{\mathcal{L}}^{(Com)} > n_{\mathcal{L}}^{(Sim)}$):

   $$F_{Com}\left(x;\boldsymbol{\theta}_{Com}\right) . \tag{2.23}$$

   In general, the two models exhibiting pure power-law behaviour are the untruncated and truncated power-law distributions (see Sec. 5.2 for more details). In this case, the untruncated power-law is a limiting case of the truncated power-law distribution.

2. **Estimation of parameters**: Maximum likelihood estimation (MLE) of the parameters. One obtains the set of parameters $\hat{\boldsymbol{\theta}}$ that maximize the likelihood function. At the end

of this procedure, one might have two different fits:

$$F_{Sim}\left(x;\hat{\boldsymbol{\theta}}_{Sim}\right) . \tag{2.24}$$

$$F_{Com}\left(x;\hat{\boldsymbol{\theta}}_{Com}\right) . \tag{2.25}$$

3. **Model comparison**: One can decide which model is good enough to describe data by means of the likelihood ratio test (LRT) or the Akaike information criterion (AIC). At the end of this procedure one might have either the simpler or the more complex model. From here to the end of this summary, let us consider that the fit is $F_{\mathcal{X}}\left(x;\hat{\boldsymbol{\theta}}\right)$.

4. **Statistical goodness-of-fit hypothesis testing**: Formulate the null and alternative hypothesis:

   - H_0: Data have been sampled from $F_{\mathcal{X}}\left(x;\hat{\boldsymbol{\theta}}\right)$.
   - H_1: Data have not been sampled from $F_{\mathcal{X}}\left(x;\hat{\boldsymbol{\theta}}\right)$.

   By means of the goodness-of-fit hypothesis testing procedure, one can determine with a significance level α whether the null hypothesis is rejectable or not.

5. **Conclusion**: If the statistical hypothesis testing yields non rejection of the null hypothesis, that means that there is not enough statistical evidence to reject the hypothesis that the proposed statistical model with the set of estimators $\hat{\boldsymbol{\theta}}$ is suitable to describe data.

Chapter 3

Earthquakes and Labquakes: Earth and Laboratory

In this chapter, the basic concepts that are needed in order to motivate and explain the two phenomena of interest for the works developed in this thesis are introduced. A general introduction to the physics of materials is presented by including the main definitions of Continuum Mechanics as well as the Mohr-Coulomb failure criterion. After an introductory overview of physics of materials is presented, one is able to understand the main concepts about physics of earthquakes. The main statistical laws in seismology are also presented. Finally, the state of the art of labquakes (synthetic earthquakes generated in laboratory compression experiments) is presented by exposing the similarities and differences with real earthquakes.

3.1 Mechanics of materials

Rheology is a sub-field of Continuum Mechanics that studies the response of materials when they are subjected to stresses. The main purpose of this section is to give a brief overview about the main definitions about physics of materials so as to present the basic concepts of rheology that will be useful to understand labquake and earthquake physics.

3.1.1 General concepts in Continuum Mechanics

Let us assume that the continuum hypothesis for materials [Lautrup, 2011] is valid and let us assume that materials are isotropic but not necessarily homogeneous. In this context, a material can be modelled as an object which is composed by cubic blocks. The coordinate system is denoted by the orthonormal vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$, which are normal to the planes 1, 2 and 3 respectively (see Fig. 3.1). Let us focus on one of these cubic blocks and let us suppose that it is subjected to a force per unit area, known as stress, which is acting over its faces. Given that the force can acquire three components for each of the faces of the cube, the stress $\boldsymbol{s}$ has to be understood as a tensorial magnitude whose components s_{kl} correspond to the stress in the k-th direction ($k = 1, 2, 3$) applied on the l-th plane of the cube ($l = 1, 2, 3$). A schematic representation of the stress tensor and its components in a cube is shown in Fig. 3.1.

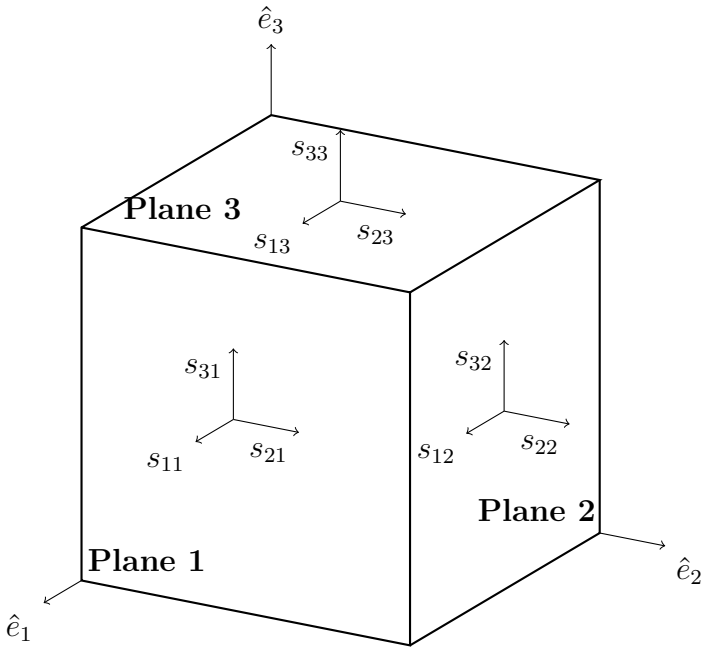

Fig. 3.1: Representation of the stress tensor $\boldsymbol{s}$ for a cubic element of a material. Note that within this choice of axes compression is negative.

A stress applied on a surface can be decomposed in normal σ_n and shear τ components which are perpendicular and parallel to the surface respectively. Normal and shear components correspond to the diagonal and non-diagonal terms of the stress tensor. By the principle of conservation of angular momentum [Lautrup, 2011], it can be shown that the stress tensor is symmetric. Therefore, there will always exist a coordinate system in which it is diagonal. The eigenvalues of the stress tensor are usually called principal stresses s_1, s_2, s_3 and correspond to forces which are perpendicular among them. Taking into account that compression is negative (in this convention), these eigenvalues are usually called maximum, intermediate and minimum principal stresses $|s_1| \geq |s_2| \geq |s_3|$.

Let us now suppose that a material may experience a deformation under stress. Let us consider that $\boldsymbol{r}$ is a vector that describes the position of a point element of this material and let us define the displacement field $\boldsymbol{u}(\boldsymbol{r})$ as the vector that characterizes the displacement of this point. The k-th component of the displacement of a nearby point originally placed at $\boldsymbol{r} + d\boldsymbol{r}$ can be approximated at first order by:

$$u_k(\boldsymbol{r} + d\boldsymbol{r}) \simeq u_k(\boldsymbol{r}) + \sum_{l=1}^{3} \frac{\partial u_k(\boldsymbol{r})}{\partial r_l} dr_l = u_k(\boldsymbol{r}) + du_k\,.$$

The last term in this equation can be modified:

$$du_k = \frac{1}{2}\left(\frac{\partial u_k(\boldsymbol{r})}{\partial r_l} + \frac{\partial u_l(\boldsymbol{r})}{\partial r_k}\right) dr_l + \frac{1}{2}\left(\frac{\partial u_k(\boldsymbol{r})}{\partial r_l} - \frac{\partial u_l(\boldsymbol{r})}{\partial r_k}\right) dr_l = (\varepsilon_{kl} + \Phi_{kl})\, dr_l\,,$$

in such a way that two different tensors can be defined in this last expression: $\boldsymbol{\Phi}$ is a rigid body rotation [Stein and Wysession, 2003] and $\boldsymbol{\varepsilon}$ corresponds to the strain tensor, whose components are:

$$\varepsilon_{kl} = \frac{1}{2}\left(\frac{\partial u_k(\boldsymbol{r})}{\partial r_l} + \frac{\partial u_l(\boldsymbol{r})}{\partial r_k}\right)\,. \tag{3.1}$$

The strain tensor is, by construction, symmetric and thus, there exists a basis where it is diagonal. This system of axes is usually known as principal strain axes and the sum of the eigenvalues or principal strains (trace of the strain tensor) corresponds to a relative volumetric change of the material. Therefore, no volume variations will occur if the trace of the strain tensor is zero.

A material may exhibit different mechanical regimes depending on the applied stress, its composition, internal structure, etc. (see Fig. 3.2). An isotropic material is said to be in elastic regime when the stress and strain tensors are proportional through a set of constants called elastic moduli. The elastic regime is characterized by the Hooke's law:

$$s_{kl} = 2\mu\varepsilon_{kl} + \lambda\delta_{kl}\sum_j \varepsilon_{jj}\,, \tag{3.2}$$

where μ and λ are the Lamé elastic moduli and have units of Pa=Nm^{-2}. In this regime the deformation is reversible: if the applied stress is removed, the material will not exhibit any

permanent deformation (see Fig. 3.2(a) before fracture). A material may overcome the elastic regime at the yielding point s_c. From that point, the material can exhibit different non-linear regimes. A material is said to be perfectly brittle when the yielding stress coincides with the failure stress s_f and a big event breaks the material (see Fig. 3.2(a)). A perfectly-plastic behaviour is achieved above the yielding point if the stress is constant for all the strains. A general plastic behaviour is achieved when there exists a functional relationship (linear or not) between the applied stress and the strain above the yielding point (see Fig. 3.2(c)). In any of the plastic regimes the deformation is irreversible: if the applied stress is removed, the material will exhibit a permanent deformation (see Fig. 3.2(c)). A brittle-plastic material might overcome the plastic regime at $s_f > s_c$ and experience failure (see Fig. 3.2(b)).

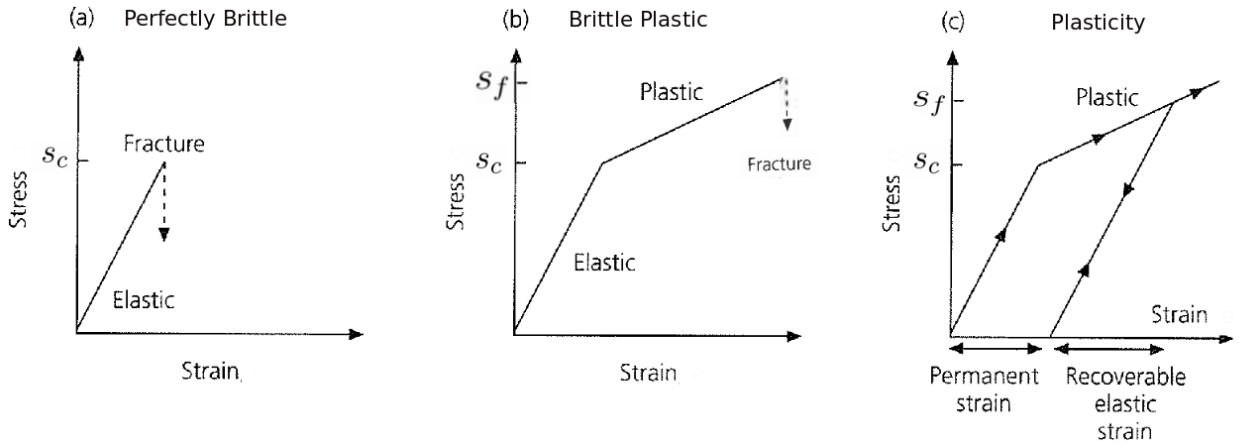

Fig. 3.2: Mechanical response of materials exhibiting perfectly brittle (a), brittle plastic (b) and plasticity (c) regimes. Figure extracted and modified from Stein and Wysession [2003].

3.1.1.1 Mohr-Coulomb failure criterion

Determining when a material is going to collapse under applied stress is one of the cornerstones in science of materials. There are several ways to establish a failure criterion $s_f = f(\boldsymbol{s})$ as a function of the components of the stress tensor. The Mohr-Coulomb failure criterion is a mathematical model which aims to determine failure of brittle materials ($s_f = s_c$). The Mohr-Coulomb criterion considers that the intermediate principal stress s_2 does not have influence on the failure criteria. There are other failure criteria taking into account the influence of the intermediate principal stress and there is still debate on whether they have an influence or not on the failure criteria [Bernt Aadnøy, 2010]. Nevertheless, due to its mathematical simplicity and easy physical interpretation, the Mohr-Coulomb failure criterion is one of the most accepted in literature.

The Mohr-Coulomb failure criterion assumes that the failure will occur in a planar surface of the material which is named as rupture surface. Given that the important phenomenon

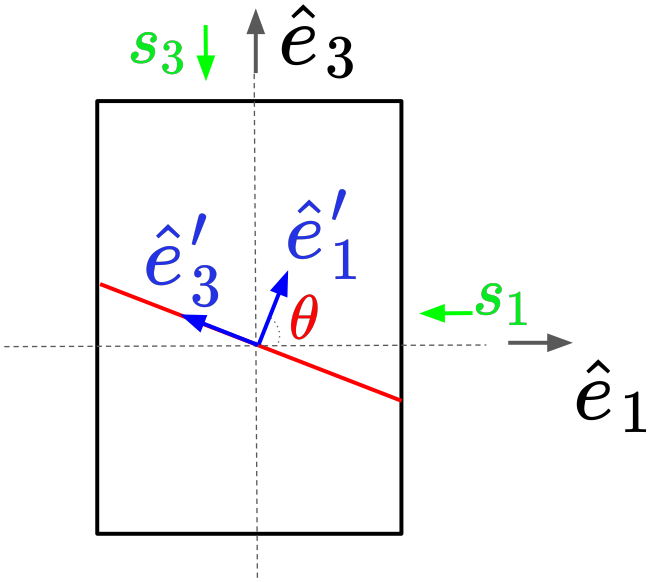

Fig. 3.3: Geometric configuration of the Mohr-Coulomb failure criterion on the section of a rectangular sample. $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_3\}$ correspond to the principal axes where the stresses s_1 and s_3 (in green) are acting. The red line corresponds to the rupture surface and $\{\hat{\boldsymbol{e}}'_1, \hat{\boldsymbol{e}}'_3\}$ (in blue) are normal and tangential to this surface.

occurs in a surface and the influence of the intermediate stress is assumed to be negligible, the problem can be solved in a two-dimensional system. Let us consider that, in the principal axes coordinate system, the largest and the smaller principal stress are s_1 and s_3 with $|s_1| \geq |s_3|$. Note that, under the convention that compressive stresses are negative, these principal stresses are considered in absolute value ($|s_1| \geq |s_3|$). Let us denote as $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_3\}$ the principal axes system, where the stress tensor is diagonal, and let us consider the angle θ between the vector $\hat{\boldsymbol{e}}'_1$ normal to the rupture surface and the direction of the principal stress $\hat{\boldsymbol{e}}_1$ (see Fig. 3.3). In order to find the components of the stress tensor in the coordinate system given by $\{\hat{\boldsymbol{e}}'_1, \hat{\boldsymbol{e}}'_3\}$, one needs to perform a rotation of angle θ:

$$\begin{pmatrix} s'_{11} = \sigma_n & \tau \\ \tau & s'_{33} \end{pmatrix} = \boldsymbol{s}' = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} s_1 & 0 \\ 0 & s_3 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3.3}$$

$$= \begin{pmatrix} s_1\cos^2\theta + s_3\sin^2\theta & (s_1 - s_3)\sin\theta\cos\theta \\ (s_1 - s_3)\sin\theta\cos\theta & s_3\cos^2\theta + s_1\sin^2\theta \end{pmatrix} \tag{3.4}$$

$$= \begin{pmatrix} \frac{s_1+s_3}{2} + \frac{s_1-s_3}{2}\cos 2\theta & \frac{s_3-s_1}{2}\sin 2\theta \\ \frac{s_3-s_1}{2}\sin 2\theta & \frac{s_1+s_3}{2} + \frac{s_3-s_1}{2}\cos 2\theta \end{pmatrix}, \tag{3.5}$$

where the normal component s'_{11} corresponds to σ_n.

The normal σ_n and shear τ components are thus represented by:

$$\sigma_n = \frac{s_1 + s_3}{2} + \frac{s_1 - s_3}{2} \cos 2\theta \,, \tag{3.6}$$

$$\tau = \frac{s_3 - s_1}{2} \sin 2\theta \,. \tag{3.7}$$

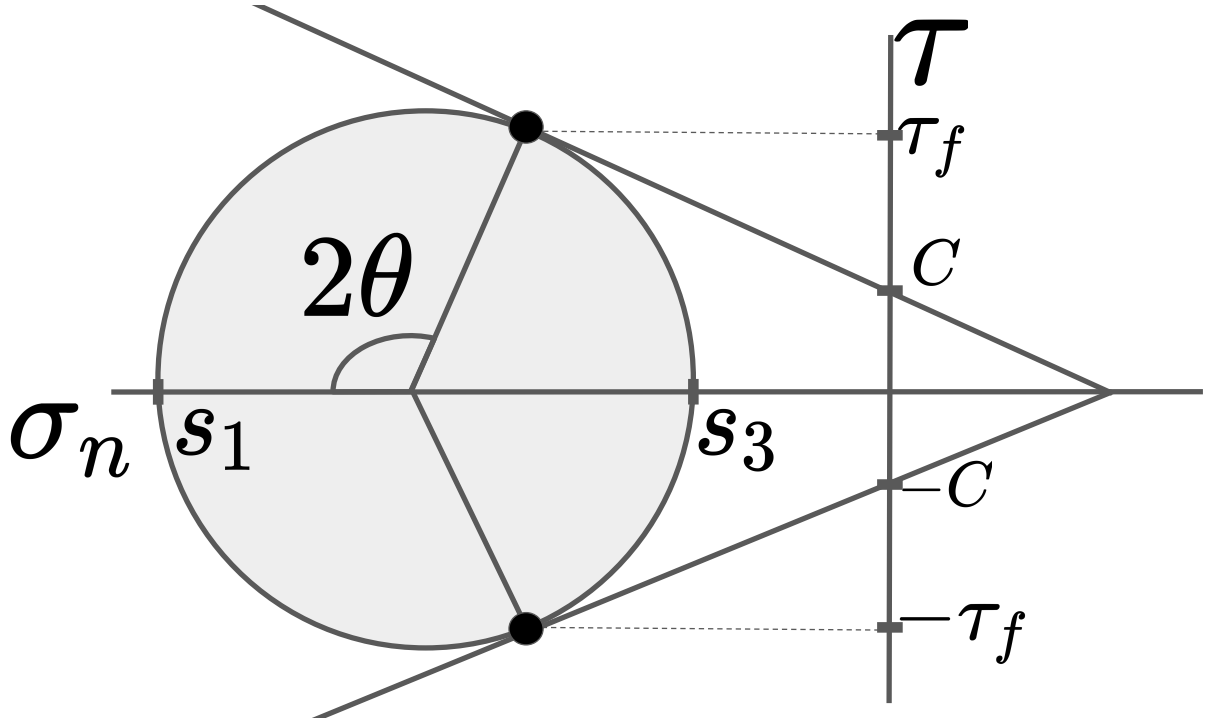

Fig. 3.4: Mohr-circle: Given a material whose stress tensor has principal components s_1 and s_3, all the possible rupture planes are represented by this Mohr-circle. For a rupture plane whose normal vector forms an angle θ with the principal axis of the material, the normal and critical shear stresses σ_n and τ_f will lie in the point of the circle determined by the angle 2θ measured clockwise with respect the $-\sigma_n$ axis.

Note that these equations correspond to the parametric equations of a circumference with radius $\frac{s_3 - s_1}{2}$ centred at $\left(\frac{s_1+s_3}{2}, 0\right)$. This circumference in the σ_n-τ plane is usually known as the Mohr circle (see Fig. 3.4). In other words, all the possible orientations of the rupture planes are represented in the Mohr circle. The values of σ_n and τ over the rupture plane are given by the point in the circle placed at angle 2θ, measured clockwise from the $-\sigma_n$ axis (see Fig. 3.4).

In this framework, the Mohr-Coulomb criterion states that failure will occur when the shear stress on the rupture planes surpasses the critical value τ_f given by:

$$\tau_f = C - \mu' \sigma_n \,, \tag{3.8}$$

where C is the cohesion and μ' the effective fault friction coefficient (including the contribution of the pore pressure [Cocco and Rice, 2002; King et al., 1994]). Care must be taken with the convention of signs in the normal stress, which is not the same in geophysics than in solid mechanics. Remark that convection in this work is the one from geophysics, which takes the negative sign for compression, this is the reason for the negative sign before μ'. From this failure criterion it is natural to define the Coulomb stress as

$$\mathfrak{C} = \tau + \mu' \sigma_n \,, \tag{3.9}$$

which signals failure by $\mathfrak{C} > C$. Therefore, the Coulomb stress has to be understood as a linear combination of the normal and shear stresses. It must be mentioned that the nomenclature used in seismology for the Coulomb stress, CFS, has been omitted in order to avoid confusion with other symbols used along this PhD thesis. The Coulomb stress is relevant when studying models for earthquake forecasting (see Sec. 3.2.2.5), specially in the case of physics-based models, where the forecasted rates of events depend on the Coulomb stresses calculated in the region of interest (see Sec. 3.2.2.5 for more details).

3.2 Seismology

The aim of this section is to expose general concepts of seismology that are relevant to understand the works exposed in this PhD thesis.

3.2.1 Physics of Earthquakes

3.2.1.1 Faulting geometry

Let us consider as "*common*" earthquakes those that have been induced either by tectonic driving or by precedent shocks [Stein et al., 1992; King et al., 1994; Steacy et al., 2005; Nandan et al., 2016; Ishibe et al., 2015]. The geometrical spaces where these *common* earthquakes occur are called faults. Although the real geometry of a fault can be very complex, the usual approach is to consider faults as planar surfaces separating tectonic blocks. In this sense, a fault can be understood as the rupture surface in the Mohr-Coulomb theory exposed in Sec. 3.1.1.1. Due to the internal dynamics of the Earth, stress is accumulated in some regions. These regions experience failure whenever that stress surpasses the strength of the material and tectonic blocks move with respect to each other. That failure is the earthquake. The materials that conform the Earth crust are constantly under stresses and suffer deformations with respect their original length, area or volume. In addition to tectonic driving, it must be mentioned that there are other mechanisms through which the stress changes in a region: by volcanic activity [Segall et al., 2013] or by artificial means such as injection of fluids [Cesca et al., 2014] or aquifer withdrawal [Shirzaei et al., 2016].

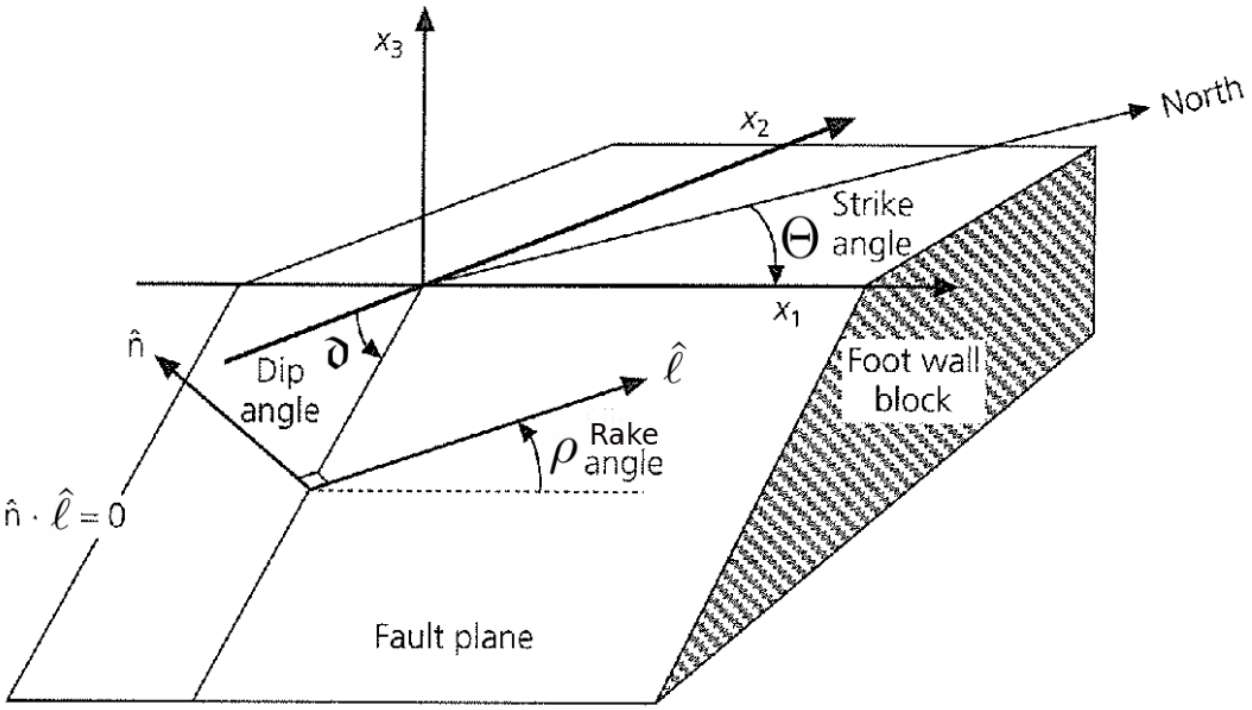

Fig. 3.5: Fault geometry shown for the foot wall block. The fault plane is defined by the normal vector $\hat{n}$ and the sense of the movement is determined by the slip vector $\hat{\ell}$. The vector pointing to the geographic north forms an angle Θ with the axis x_1, x_3 points upwards and x_2 is perpendicular to x_1 and x_3. The focal mechanism is defined by the angles Θ, $\mathfrak{d}$ and ρ. Figure modified from Stein and Wysession [2003].

In a general case in which the tectonic blocks are not at the same height, the block situated at the upper side of the fault is known as the hanging-wall block whereas the one placed at the lower side is called foot-wall block. Figure 3.5 shows the foot wall block in order to illustrate the standard configuration of fault geometry. The fault plane is determined by the normal vector to the fault $\hat{\boldsymbol{n}}$. The slip vector $\hat{\boldsymbol{\ell}}$ is contained in the fault lane ($\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{\ell}} = 0$) and determines the direction in which the hanging wall block moved with respect to the foot wall block. The system of axes $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}\}$ is chosen in such a way that $\boldsymbol{x_1}$ is defined by the intersection of the fault plane with the Earth surface, the axis $\boldsymbol{x_3}$ is normal to the Earth surface and $\boldsymbol{x_2}$ is perpendicular to $\boldsymbol{x_1}$ and $\boldsymbol{x_3}$ (see Fig. 3.5). Three different angles can be defined in order to determine the fault geometry: the strike Θ, dip $\mathfrak{d}$, and rake ρ. The strike angle Θ takes values from 0° to 360° and is defined by $\boldsymbol{x_1}$ and the vector pointing to the geographic North. The dip angle $\mathfrak{d}$ acquires values from 0° to 90° and is defined by $\boldsymbol{x_2}$ and the orientation of the fault plane. The dip $\mathfrak{d}$ is measured from the $-\boldsymbol{x_2}$ axis and determines the sense of the axis $\boldsymbol{x_1}$ in such away that $\mathfrak{d} \leq 90°$. The rake angle ρ takes values from $-180°$ to 180° and arises from the intersection of the slip vector $\hat{\boldsymbol{\ell}}$ with the axis $\boldsymbol{x_1}$. By considering $\boldsymbol{x_1}$ as a reference, the rake ρ is measured counter-clockwise whereas the strike Θ is measured clockwise. In seismology,

this set of angles is known as the focal mechanism of the earthquake.

Assuming a coordinate system in which the stress tensor is diagonal and one axis is normal to the Earth surface, different styles of faulting can be determined depending on the distribution of the stresses. The general classification considers three styles of faulting: normal, strike-slip and thrust (see Fig. 3.6). A sub-classification can be done for strike-slip faults depending on the relative movement of both sub-blocks. By considering the frame of reference

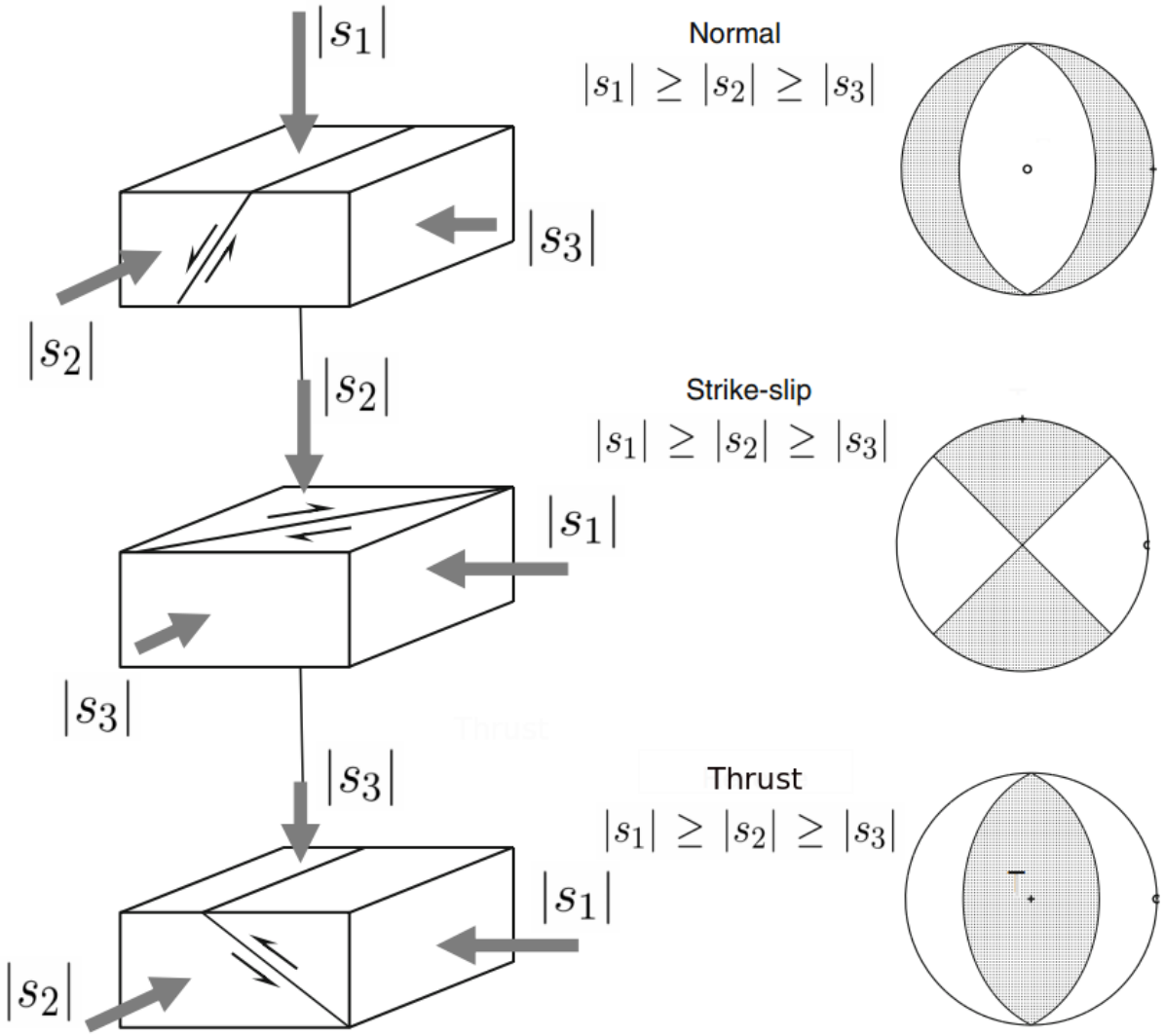

Fig. 3.6: Different faulting regimes: (top) normal, (center) strike-slip and (down) thrust. The parallelepiped at left corresponds to two different tectonic blocks subjected to the different compressing stress $|s_1| \geq |s_2| \geq |s_3|$ and moving with respect each other. Beach-ball diagrams are presented for each case: compressional quadrants are shadowed and dilatational quadrants are white. Figure modified from Vavryčuk [2015].

at one tectonic block facing the other block, right-lateral strike-slip faults are those where the other tectonic block moves to the right, whereas left-lateral strike-slip faults are those where it moves to the left. It is interesting to note that faulting regimes can be also determined by the rake angle: $\rho = 90°$ for pure thrust faults, $\rho = -90°$ for pure normal faults, $\rho = 0°$ for pure left-lateral strike slip and $\rho = 180°$ for pure right-lateral strike slip. In reality, this classification is more flexible and each style of faulting can be characterized by a range of rake angles ρ:

- Thrust: $45° \leq \rho \leq 135°$
- Normal: $-135° \leq \rho \leq -45°$
- Strike-slip left-lateral: $-45° \leq \rho \leq 45°$
- Strike-slip right-lateral: $135° \leq \rho \leq 180°$ and $-180° \leq \rho \leq -135°$

The normal $\hat{\boldsymbol{n}}$ and slip $\hat{\boldsymbol{\ell}}$ vectors as a function of the strike, dip and rake angles are presented in Sec. 8.1.1.

3.2.1.2 Nodal planes

The stored elastic energy before an earthquake is released in the form of pressure waves and heat. Given the nature of the process through which these pressure waves have been generated, they are commonly referred as seismic waves. In this context, an earthquake can be understood as a point source radiating seismic waves when the surrounding stress field has experienced some variations. Part of these seismic waves is detected as a perturbation by means of seismometers, from which the location of the source where the earthquake has occurred can be inferred. This point source is called hypocenter and is determined by three spatial coordinates: latitude, longitude and depth. The projection of the hypocenter to the surface is called epicenter. More details about the detection of earthquakes are exposed in Sec. 4.2.

The first waves detected by a seismometer are longitudinal waves which compress and dilatate the medium in the propagation direction of the perturbation. Given that these waves are the first ones arriving, they are called Primary (P) waves. The seismometer detects Secondary (S) shear waves which are slower than P waves and are transversal to the propagation direction of the perturbation. All these perturbations can be detected by different seismometers at different sites. Each seismometer generates a seismogram which is crucial to determine the fault geometry and the focal mechanism.

Depending on the location of the seismometer, the first detected motions corresponding to P waves arriving to the seismometer from below will bring the material towards the position of the seismometer (compression) or away from it (dilatation). This movement generates a vertical component in the seismometer pointing upwards or downwards when the material is either compressing or dilating. In this sense, the first motions detected in an earthquake

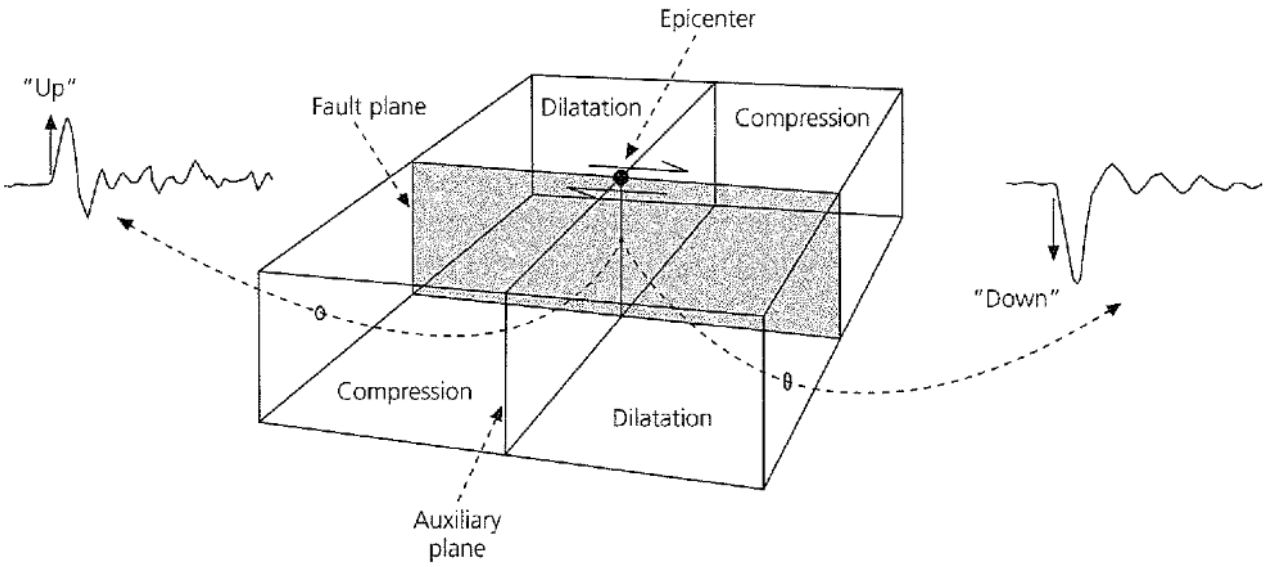

Fig. 3.7: Simple example of a right-lateral strike-slip earthquake in a vertical fault whose first motions are detected at two different stations. For the station placed in the left, the first motion points upwards ("Up") whereas for the one placed in the left it points downwards ("Down"). The nodal planes separate compressive from dilatational quadrants. The shadowed nodal plane corresponds to the real fault plane. Note that it is not possible to distinguish which of them is the actual fault plane given these first motions data. Figure extracted from Stein and Wysession [2003].

define four quadrants (see Fig. 3.7) [Stein and Wysession, 2003]. These quadrants are defined by the intersection of two planes which are perpendicular to each other. These planes, which are called nodal planes, separate the quadrants where the material is compressed from those where it experiences dilatation. One of these nodal planes corresponds to the real fault plane. Nevertheless, given that the first motions can be very similar for both nodal planes, there usually exists an ambiguity in their association to the real and the fictitious one. Despite additional geological information can be used in order to clearly discern which is the actual fault plane, it is not always possible to determine it unambiguously [Stein and Wysession, 2003].

A useful tool to graphically illustrate the nodal planes and the focal mechanism of an earthquake are beach-ball diagrams (see Fig. 3.6). Beach ball diagrams correspond to the stereographic projection of the nodal planes onto a circle. One can roughly understand that the material is moving from white to shadowed regions of the circle. Depending on each focal mechanism, each of these quadrants will be compressional or dilatational. The beach balls associated to each focal mechanism are shown in Fig. 3.6 and are a very useful pictorial resource to see which is the rupture tendency for an earthquake sequence occurring in a fault.

As already mentioned, it is not possible to determine the fault plane from the first motions unambiguously and it should be necessary to have also the focal mechanism in the auxiliary plane [Stein and Wysession, 2003]. It is thus necessary to establish a way to compute the

focal mechanism in the complementary plane from the focal mechanism in the other one and vice-versa. For more details about how to compute the focal mechanism in the other plane, see Sec. 8.2.4.

3.2.2 Statistical seismology

Whenever an earthquake occurs, some observables characterizing it, such as the spatial location or the moment magnitude, can be reported in an earthquake catalog for further statistical analysis. As a general approach, these observables are considered as random variables that follow a certain probability distribution. This is the starting point of statistical seismology [Kagan, 2013]. In this framework, it is important to determine the existence of statistical patterns in order to associate a probability of occurrence to a certain event. These statistical patterns can reflect the existence of temporal correlations, a scale-free behaviour of the probability distribution of some observables, etc. The knowledge of these statistical patterns is a fundamental step for the development of earthquake forecasting.

Two main different approaches can be assumed when studying statistical seismology [Sornette and Werner, 2011]:

- Holistic approach: The statistical analysis of large catalogs is performed without much focusing on the geophysical details of every earthquake [Kagan, 1999; Godano and Pingue, 2000]. There is no reason to assign a hierarchy of events and there are no earthquakes which are more important than other ones. In this sense, the holistic approach defends a universal point of view of the statistical seismological laws.

- Reduccionist approach: Detailed modelling and analysis for the understanding of particular faults. Much attention is paid in the geophysical details: classification of different environments, stress conditions, fault types, rock properties, temperature, underwater earthquakes, clusters, etc. [Zaliapin and Ben-Zion, 2013].

In statistical terms, one can consider that a phenomenon is described by a population which is power-law distributed with a certain exponent. One can select samples from that population in order to perform statistical inference. However, these samples are not necessarily power-law distributed with exactly the same exponent. Depending on the sample sizes, there exists a range of values for which exponents are compatible (see Sec. 5.9). Undoubtedly, it is important to determine whether these differences are statistically significant or not in order to see whether universality can be rejected or not.

3.2.2.1 The Gutenberg-Richter law

One of the most important laws in statistical seismology is the one that tries to characterize the probability distribution of earthquake sizes. The knowledge of this distribution is crucial in models for earthquake forecasting [Scholz, 2019] since it associates probabilities to the different earthquake sizes.

First of all, as it was explained in the previous section, one needs a convenient way to measure earthquake sizes. Among the different scales that one can choose for this purpose, the one which is considered along this PhD thesis is the moment magnitude scale. This scale is widely accepted to represent the size of an earthquake since it is related with a physical quantity, the seismic moment. Moreover, contrarily to the Richter magnitude [Stein and Wysession, 2003], this scale does not saturate for large earthquakes. The scalar seismic moment M [Brune, 1968; Yeats et al., 1997; Stein and Wysession, 2003; Scholz, 2019] is defined as:

$$M = \mu \bar{\ell} A_r \,, \tag{3.10}$$

where μ is the Lamé elastic shear modulus from Eq. (3.2), $\bar{\ell}$ is the mean slip originated by the earthquake and A_r is the rupture area. The moment magnitude has energy units and can be easily interpreted since it keeps the idea that, the larger the rupture is (the average slip and the ruptured area caused by the earthquake) the larger the size one can associate to that earthquake. The shear modulus depends on the medium and it is reasonable to think that the process to originate the rupture on a material will also depend on its properties.

At this point one could wonder: Is the moment magnitude somehow related with the energy released by the earthquake? Whether the ratio of seismically radiated energy over the seismic moment is independent on the moment magnitude is still an unsolved question [Kanamori and Brodsky, 2004]. This ratio may depend on different earthquake parameters such as moment magnitude and the depth of the source [Brodsky and Kanamori, 2001; Vassiliou and Kanamori, 1982; Izutani and Kanamori, 2001; Kanamori and Heaton, 2013]. On the contrary, the seismically radiated energy is in some cases underestimated such as the ratio may be considered as constant [Ide and Beroza, 2001; Wang, 2015; Bilek et al., 2004; Zollo et al., 2014]. The value of this unique factor seems to be $\frac{E_r}{M} = 10^{-4.6}$ [Bormann, 2015], where E_r is the energy radiated by the earthquake.

The moment magnitude m is related to the seismic moment according to:

$$m = \frac{2}{3}\left(\log_{10}\frac{M}{M_0} - 9.1\right) \,, \tag{3.11}$$

where the moment magnitude m is dimensionless and M_0 is a reference moment magnitude of 1 Nm. More details about the moment magnitude scale are exposed in Sec. 4.2.2.

As soon as there is a way of quantifying earthquake sizes, one is in a position to consider earthquake magnitude as a random variable $\mathfrak{m}$ which follows a certain probability distribution.

The distribution of earthquake magnitudes is known as the Gutenberg-Richter (GR) law [Utsu, 1999]. This law states that, above a lower cut-off value, earthquake magnitudes follow an exponential distribution; in terms of the magnitude PDF

$$f_{\mathfrak{m}}(m) = (b \ln 10) 10^{-b(m-m_{min})} \propto 10^{-bm} , \tag{3.12}$$

defined for $m \geq m_{min}$, with m the moment magnitude, m_{min} the lower cut-off in magnitude and b the so called b-value which turns out to be close to one ($b \simeq 1$). As the $b-$value is directly related to the log-ratio between the number of small and large earthquakes, variations in b can be associated with the ability of an earthquake rupture to propagate (more large events, low b) or not (less large events, high b). A straightforward property of the exponential distribution leads to the fact that the rate (the number per unit time of earthquakes above a certain magnitude m) is also a decreasing exponential function of the magnitude, with the same b-value.

The GR law is a power-law distribution when it is written as a function of the seismic moment M as can be derived from Eqs. (3.11) and (3.12):

$$f_{\mathcal{M}}(M) = \frac{2}{3}\frac{b}{M_{min}}\left(\frac{M}{M_{min}}\right)^{-\left(1+\frac{2}{3}b\right)} = \frac{\gamma-1}{M_{min}}\left(\frac{M}{M_{min}}\right)^{-\gamma_M} , \tag{3.13}$$

where it has been conveniently defined $\gamma_M = 1 + \frac{2}{3}b$ and M_{min} corresponds to the value of the seismic moment when the cut-off magnitude m_{min} is introduced in Eq. (3.11). Note that this PDF has support $M \in [M_{min}, +\infty)$ as $b > 0$, then $\gamma_M > 1$. One of the most interesting properties of this law is that, given the measured values of the exponent γ_M, the first moment of the seismic moment diverges (see Sec. 5.2.1). This fact implies that the mean is not well defined and the central limit theorem cannot be applied [Corral and Font-Clos, 2013].

This genuine description of the GR law does not take into account all the deviations at the tail from the expected behaviour. Corral and González [2019] checked that a different behaviour appeared for moment magnitudes larger than 7.4 in global earthquakes. The authors proposed that the tail could be described by another power-law behaviour with a b-value close to $b \simeq 1.5$. Serra and Corral [2017] proposed several statistical models with the aim of describing the tail of the GR law. If one desires a unique power-law behaviour, one can exclude all these big events that deviate from the power-law simply by imposing an upper cut-off M_{max} to this statistical law and properly normalizing the PDF (see Sec. 5.2.2). Note that the fact of modelling the tail of the GR law by an exponential tail or working with an upper truncation eliminates the divergence of the first moment but the restricted scale invariance property can just be considered in a narrower range.

Independently of the deviations appearing in the tail of the GR law, it has been long debated [Kamer and Hiemer, 2015] whether the b-value is essentially universal [Kagan, 2013] (accordingly to the holistic approach) or, on the contrary, it is affected by different geophysical conditions (as the reductionist approach would support). Some studies [Schorlemmer et al., 2005; Narteau et al., 2009] have correlated the $b-$value with the focal mechanism [Wyss and

Wiemer, 2000]. These studies indicate that (at least for California, for a long time period) $b \approx 1.03$ for normal events, $b \approx 0.87$ for strike-slip events, and $b \approx 0.79$ for thrust events [Schorlemmer et al., 2005].

As it has been already mentioned, this PhD thesis will discuss whether this seismological law can be considered to be universal (see Chpt. 7) or whether it is affected by geophysical quantities, such as the Coulomb stress increase (see Chpt. 8).

3.2.2.2 The Omori law

The Omori law [Omori, 1894; Utsu et al., 1995] is the main seismological law which reflects the temporal clustering of earthquakes. The mere fact of considering clusters of earthquakes is somehow making differences among them; which is somewhat in contradiction with the holistic approach of assigning the same relevance to all the events.

Let us consider a big earthquake, which will be named as mainshock, occurring at a time t_{MS} and a sequence of smaller subsequent earthquakes appearing at a time t ($t > t_{MS}$) after this main event. These smaller earthquakes are known as aftershocks and are specially relevant in earthquake forecasting for being potentially hazardous [Scholz, 2019]. The aftershocks have been somehow triggered by the mainshock and this implies the existence of a temporal dependence. The hallmark of these temporal dependence can be mathematically expressed by means of the Omori law:

$$dn_{AS} = r_{AS}(t)dt = \frac{\mathsf{C}}{(\mathsf{K} + t - t_{MS})^{\pi}} dt \,, \tag{3.14}$$

where $r_{AS}(t)$ corresponds to the number n_{AS} of aftershocks per unit time (also called activity rate) at a time $t > t_{MS}$. π is known as the Omori exponent and is usually around 1, and C and K are constants with units. This phenomenological law states that the seismic activity is very large for small times after a mainshock has occurred and decreases for larger times as a power-law with exponent π. Therefore, the Omori law reflects that earthquakes tend to cluster in time.

3.2.2.3 The Productivity law

The productivity law [Helmstetter, 2003] reflects the relationship between earthquake sizes and the number of aftershocks in a sequence. This law states states that the number of aftershocks scales with the mainshock size as:

$$n_{AS}\,(m) \propto 10^{\alpha_p m} \longrightarrow n_{AS}\,(M) \propto M^{\alpha_p \frac{2}{3}} \tag{3.15}$$

where $\propto$ denotes proportionality. The number of aftershocks generated by a mainshock is related to its magnitude (or seismic moment) through the productivity exponent $\alpha_p > 0$. Despite the measured value of α_p clearly depends on which definition of mainshocks and

aftershocks is considered, this value is found to be robust [Helmstetter, 2003]. This law thus implies that earthquakes tend to form larger clusters for larger mainshocks.

Under the assumption that the Omori law and the productivity law are valid, the rate of aftershocks as a function of time and the energy of the mainshock can be written as [Baró i Urbea, 2015]:

$$\frac{dn_{AS}}{dt}(t,M)\,dt = r_{AS}(t,M)\,dt = \mathtt{C}\frac{M^{\frac{2}{3}\alpha_p}}{(\mathsf{K}+t-t_{MS})^{\pi}}dt\,. \tag{3.16}$$

This term is relevant in the development of the Epidemic Type Aftershock Sequence (ETAS) model [Ogata, 1998]. Note that this expression implies that, if different Omori curves for different mainshock energies are graphically represented, they will collapse into one single curve when times and the rate of aftershocks are conveniently rescaled [Helmstetter, 2003].

3.2.2.4 The Unified Scaling law of waiting-times

Not only earthquake size is an observable which is of special interest in earthquake forecasting but one also is concerned about when an event is going to occur. In this sense, one must also pay attention in the distribution of the time between consecutive events (or waiting-time δ).

Let us consider a certain threshold in magnitudes m_{Th} or seismic moment M^{Th} and a number n_{Th} of earthquakes whose magnitude is larger or equal than this threshold. The effect of thresholding results in a new sequence of earthquakes separated the one from another by a waiting-time δ. By taking into account the spatial degrees of freedom, a regular grid of cells of size L is considered and waiting-times δ are computed for those consecutive events whose magnitudes are larger than the threshold and occur in the same cell. According to this procedure, the four dimensional problem (latitude, longitude, time and earthquake size) has been reduced to a one dimensional problem where solely the temporal degrees of freedom are considered.

Bak et al. [2002] proposed the Unified law of waiting-times. By averaging $L \times L$ regions for different thresholds M_{Th}, the unified scaling law of waiting times can be written as:

$$f_\delta\left(\delta;\langle\delta_{M_{Th},L}\rangle\right) = \langle\delta_{M_{Th},L}\rangle^{-1}\mathcal{G}\left(\frac{\delta}{\langle\delta_{M_{Th},L}\rangle}\right), \tag{3.17}$$

where $\mathcal{G}$ is an universal scaling function and $\langle\delta_{M_{Th},L}\rangle$ is the mean waiting time for a given threshold M_{Th} and region sizes L. This equation is true for any value of the threshold M_{Th} and the size of the region [Corral, 2005a,b].

The waiting-time distribution in earthquakes has been extensively studied and it has been long debated about its universal features [Corral, 2003, 2004a; Davidsen and Goltz, 2004; Lindman et al., 2005; Corral and Christensen, 2006; Touati et al., 2009]. The waiting-time distribution exhibits power-law regimes spanning several orders of magnitude in time [Corral, 2003, 2004a]. This fact is indicative that seismology understood as a marked point process cannot be described in terms of Poisson processes. It must be mentioned that some authors

have debated this holistic point of view of the unified scaling law of waiting-times [Davidsen and Goltz, 2004]. It has been also argued that the power-law distributed waiting-times are artifacts that appear when a threshold is imposed [Janićević et al., 2016].

3.2.2.5 Earthquake forecasting

One of the main purposes of statistical seismology is to develop models to forecast earthquakes and thus try to reduce the human and material damages as a consequence of this natural hazard. Without the aim of reviewing all the kinds of models and methodologies [Scholz, 2019], one of the most common procedures of earthquake hazard forecasts usually comprises two stages: in the first one, the rate of events is forecasted, while in the second one, the Gutenberg-Richter law is applied to those rates in order to obtain the probabilities of occurrence for each magnitude threshold. In the case of physics-based models, the forecasted rates of events depend on the Coulomb stresses calculated in the region of interest. These models are variants of the rate-and-state model by Dieterich [1994]

$$r(t)dt = r_0 \left[1 + \left(e^{-\Delta\mathfrak{C}/B} - 1\right) e^{-t/t_a}\right]^{-1} dt\,, \tag{3.18}$$

where $r(t)$ is the rate of events (i.e., aftershocks) at any given time t after a mainshock, r_0 is the rate of background seismicity, $\Delta\mathfrak{C}$ is the increase in Coulomb stress induced by the mainshock, B is a constant, for our purposes, and t_a is the characteristic relaxation time [Dieterich, 1994]. Note that, for short times $t << t_a$ and $\Delta\mathfrak{C} > 0$, the rate-and-state model is in agreement with an Omori law with exponent $\pi = 1$:

$$r(t)dt \simeq \frac{\left(1 - e^{-\Delta\mathfrak{C}/B}\right)^{-1} r_0 t_a}{\left(e^{\Delta\mathfrak{C}/B} - 1\right)^{-1} t_a + t} dt\,.$$

Note that in the application of the Gutenberg-Richter law to the forecasted rate $r(t)$ given by the previous expression it is implicit that the Coulomb-stress change caused by a mainshock does not alter the fulfilment of the GR law for the aftershocks, in particular, this law remains the same no matter whether $\Delta\mathfrak{C}$ is positive or negative. In some sense, $r(t)$ inherites the dependence of the background rate r_0 with the magnitude. Therefore, the rate-and-state formulation [Dieterich, 1994; Parsons et al., 2000; Toda et al., 2005; Chan et al., 2012; Cattania et al., 2014, 2015] assumes the fulfillment of the GR law for the incoming events (aftershocks), with no change in the $b-$value. This assumption is made when inverting earthquake rates to obtain stress changes [Seeber and Armbruster, 2000; Dieterich et al., 2000; Segall et al., 2013]. Physics-based models also assume the magnitude distribution does not depend on the stress values, so that forecasted rates can be translated into probabilities of occurrence for different magnitudes.

According to Mohr-Coulomb theory [Narteau et al., 2009; Kanamori and Brodsky, 2004], thrust faults rupture at much higher Coulomb stress than normal faults (with strike-slip faults

in between, assuming the same value for the effective fault friction coefficient μ'). When the stress required to initiate a rupture is higher, stress interactions are enhanced and cracks can propagate faster in many different directions, yielding larger earthquakes [Narteau et al., 2009], consistent with the empirically observed $b-$values for thrust faulting [Schorlemmer et al., 2005]. Conversely, for lower rupture thresholds, one should find indeed the large $b-$values characterizing normal faulting. Despite the threshold for triggering might be different for the different styles of faulting, the rupture or not of a fault also depends on its previous state. Chapter 8 is focused on investigating, with rigorous statistical tools, whether the GR law is affected by the binary choice between positive and negative increases of the Coulomb stress, using the sequence of events after the 1992 Landers earthquake.

3.3 Earthquakes in the Laboratory: Labquakes

One of the main difficulties in studying earthquakes is the large spatial, temporal and energy scales involved in these phenomena [Ben-Zion, 2008; Rundle et al., 2003]. Not only these vast scales are problematic but it is almost impossible to control the Earth as a system to generate an earthquake at a given time. In this sense, earthquake catalogs are considered as observational data, where the observer cannot perturb the conditions of the Earth in order to establish a clear relationship between causes and effects. It has been extensively discussed whether there exists an experimental system which can generate earthquakes in such a way that experimental conditions can be somewhat controlled. Given that some of the main statistical laws in seismology were reproduced in rock fracture experiments [Mogi, 1967; Scholz, 1968; Davidsen et al., 2007], brittle fracture of materials has been found to be a good experimental mechanism to represent earthquakes at laboratory scale .

Power-law distributions were found for the different definitions of fracture-size [Main, 1996; Alava et al., 2006; Davidsen et al., 2007]. Despite the differences between the reported exponents with those found in real earthquakes, it is quite accepted in literature that a GR law is found for laboratory experiments of rock-fracture. Not only the exponents were different but they were not stationary and exhibited temporal variations as the materials got damaged [Alava et al., 2006].

The Omori law was also found in rock-fracture experiments [Mogi, 1967; Hirata, 1987]. Note that this law turns out to be impossible to analyse in rock-fracture experiments if one considers the big fracture as a mainschock. The different Omori laws analysed in laboratory experiments suggested that the Omori exponent π was far from one and non-stationary [Hirata, 1987]. The productivity law had not been found for any brittle-fracture laboratory experiments but it was in plastic deformation [Weiss and Miguel, 2004]. The waiting-time distribution had also been measured in laboratory experiments [Mogi, 1967; Alava et al., 2006]. However, the unified scaling law of waiting-times did not exhibit the functional form with two power-law branches, as it occurs in real earthquakes [Corral and Christensen, 2006].

A remarkable advance was done in 2011 when acoustic emission during compression experiments of porous materials was studied [Salje et al., 2011]. The acoustic emission (AE) events or avalanches appeared as a consequence of the fractures during the compression of materials [Scruby, 1987; Baró et al., 2013; Main, 2013]. By now, let us consider this simple definition of AE avalanches and a deeper insight on how they are defined will be explained in Chpt. 4. Uni-axial compression experiments without lateral confinement of porous samples placed between two compressional plates were performed by controlling the applied force over the samples. This framework is referred to as force-driven compression. The AE was detected by placing a piezoelectric transducer embedded in one or two of the compressional plates.

Some authors consider a big failure in the material as a critical point in order to explain the universal features of compression experiments [Uhl et al., 2015]. However, the compression experiments of porous material does not finish when the sample is destroyed. This implies that there is not a unique failure point in order to associate it to a "critical point". The great advantage about the existence of multiple failure points is that it enables for better statistics [Baró et al., 2013].

A typical output from one of these experiments is shown in Fig. 3.8. The number of AE avalanches occurring in a certain time window (also referred as activity rate $r(t)$) as well as the cumulative number of AE avalanches are presented in panel (a). The energy of AE avalanches as well as the evolution of the sample height as a function of the time is shown in panel (b). It is important to stress different facts that can be immediately observed from this output. First of all, it must be noticed that thousands of AE events appear all along the compression process and the rate of activity exhibits very large changes (note the logarithmic scale in Fig. 3.8(a)). The steep steps in the number of AE avalanches can also be observed when the activity dramatically increases. Secondly, the energy of these AE events can acquire values in a range about eight orders of magnitude. Note that, although the largest energy events are associated to the largest drops in the sample height, there are also high energetic events without an apparent macroscopic change in height. This fact is particularly evidenced in the last part of the experiment (for times larger than 16000 seconds) where there is still AE activity when the sample height has been enormously reduced and just tiny fragments are compressed. Note also that the largest macroscopic change in the sample height does not correspond with the period with largest activity rate of all the experiment (occurring around 12000 s).

One thing which is not evident in this output but it must also be mentioned is that AE avalanches have a duration typically ranging from microseconds to milliseconds. It must be taken into account that the experimental set-up does not provide a correct information about the location of the AE events. This implies that, contrarily to what happens in earthquakes, spatial degrees of freedom are not available in this experimental set-up. In spite of this lack of information, one could speculate that the fracture sizes in the sample might range from micrometres to few millimetres. Therefore the space, time and energy scales differ enormously with respect those of earthquakes.

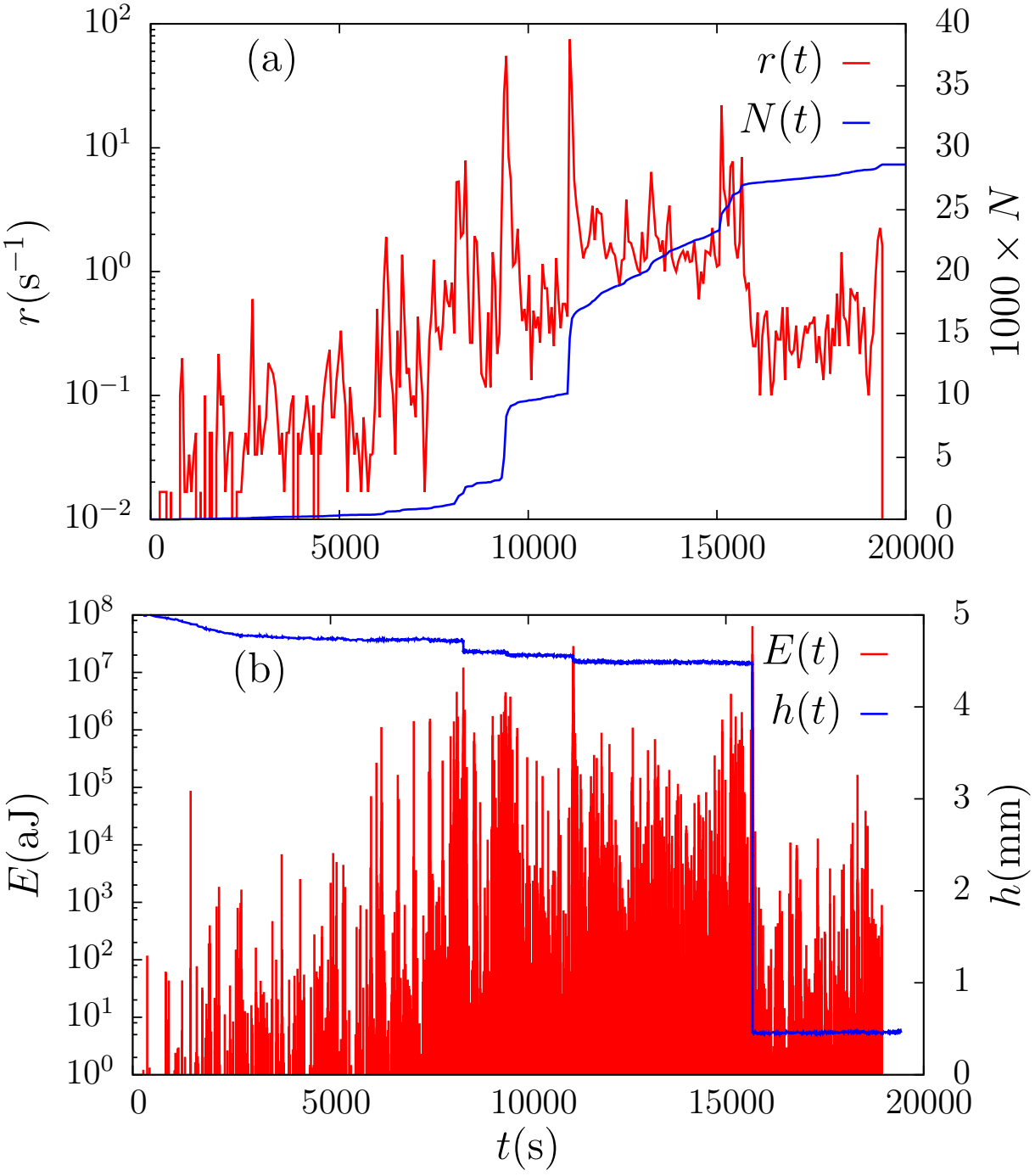

Fig. 3.8: Output from a force-driven compression experiment of a Vycor sample performed at a rate of 1.6 kPa/s. (a) activity rate of labquakes as a function of the time $r(t)$ in time intervals of 20 seconds (in red) and the temporal evolution of the number of labquakes $N(t)$. (b) energy of labquakes occurred at a time t (in red) and temporal evolution of the height of the sample $h(t)$. (Extracted from Baró et al. [2013]).

In spite of these differences in scales, the main seismological laws can also be found when the observables of these AE are studied.

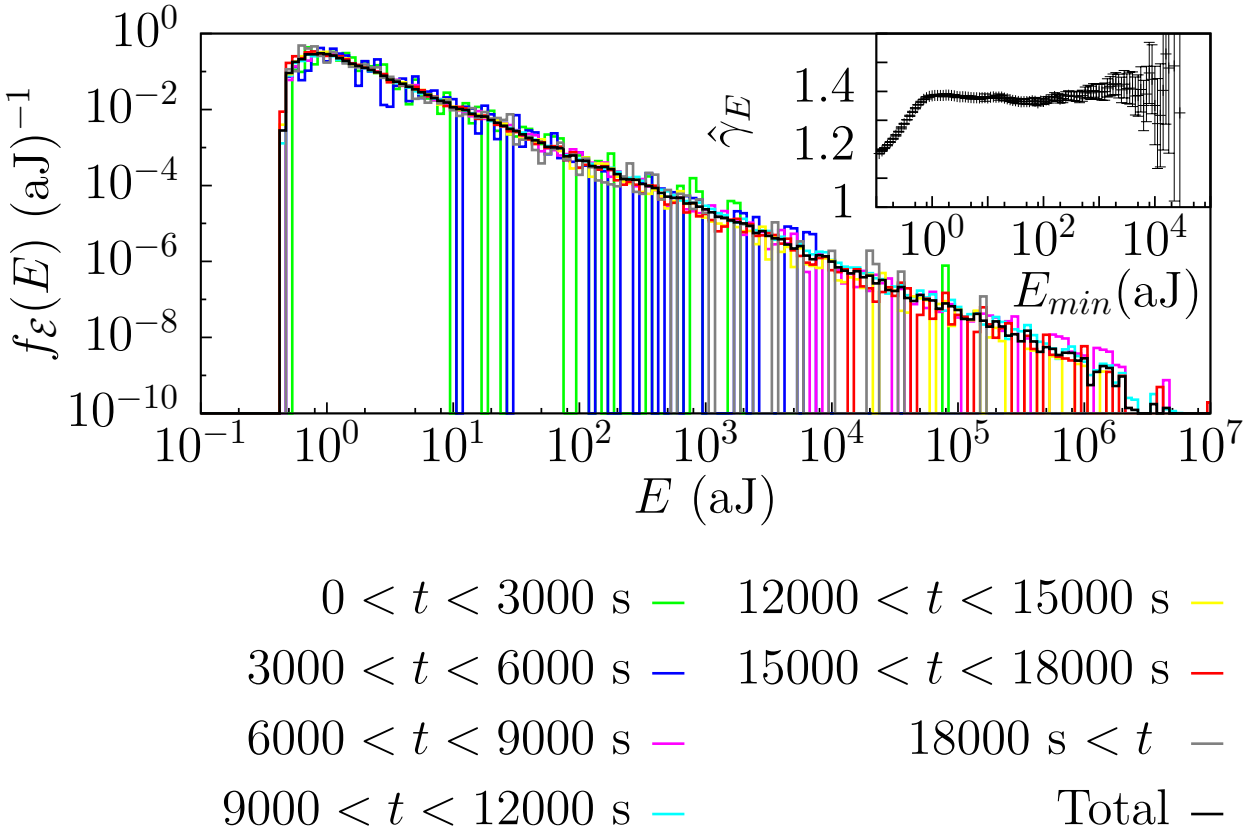

Fig. 3.9: Estimated probability density function for the energy of labquakes occurred in different time windows of 3000 seconds during a force-driven compression experiment of a Vycor sample performed at a rate of 1.6 kPa/s from Baró et al. [2013] (see Fig. 3.8(b)). Inset: the evolution of the fitted exponent $\hat{\gamma}_E$ as a function of the lower cut-off E_{min} of the power-law distribution.

A stationary GR law was found along more than six orders of magnitude (see Fig. 3.9). The power-law shape of the PDF was unaltered independently on the time window in which AE avalanches were observed. However, the b-value of the GR law was somewhat smaller ($\hat{\gamma}_E = 1.4 \rightarrow \hat{b}_{lab} \simeq 0.6$) than the the one reported for earthquakes ($b_{Earth} \simeq 1$). Despite this difference, this power-law model for earthquake and labquakes with these exponents implies that it is not possible to define a finite first moment. In other words, a mean value for earthquake or labquake sizes cannot be defined. Although the power-law distribution for the event size in fracture experiments was already known [Main, 1996; Alava et al., 2006; Davidsen et al., 2007], the relevant improvement with respect previous works was the fact of finding a stable exponent along time and valid for an important number of orders of magnitude in

energy (as it can be observed in the inset of Fig. 3.9).

There is still an open debate to explain this difference in the GR law exponent. On the one hand, Pérez-Reche et al. [2008] suggested that the driving mechanism might have an influence in the power-law exponents of the different avalanche observables. AE avalanches found in Baró et al. [2013] were generated during force-driven compression experiments. Nevertheless, the compressive mechanism can also be carried out by controlling the displacement and measuring the force opposed by the material. Such a framework is referred as displacement-driven compression. This is a good starting point to wonder whether the energy exponent is the same or not for a different compressional driving mechanism (see Chpt. 6). On the other hand, this difference could be also attributed to the nature of the material and its internal structure [Xu et al., 2019]. The composition of the Earth crust is highly heterogeneous [Goff and Holliger, 2003] and the mechanical response of the materials composing it can be very different from the samples that have been studied in the laboratory (see Chpts. 6 and 7).

As it happens when studying earthquakes, these AE avalanches can also be studied as a marked point process represented by Fig. 3.8(b). The Omori law was also found with an exponent somewhat smaller ($\pi \simeq 0.75$) than the one of earthquakes. An important achievement of these experiments was that the productivity law was also fulfilled with a value $\alpha_p \simeq 0.5$, which is somewhat smaller than the one reported for earthquakes [Helmstetter, 2003]. In spite of this smaller value, it is still positive thus implying that, the larger the mainshock is the larger the number of triggered aftershocks will be. The unified scaling law of waiting times was also found. The universal scaling function has a double power-law structure with exponents very close to those found for earthquakes [Corral, 2004b; Baró et al., 2013].

Because of the fulfilment of these main seismological laws, AE events occurring during the uni-axial compression experiments of materials will henceforth be named as labquakes [Main, 2013]. Given that labquakes were found to be a reasonable experimental system to reproduce earthquakes in a laboratory scale, multiple experiments generating labquakes were subsequently studied for different materials such SiO_2-based minerals [Nataf et al., 2014], wood [Mäkinen et al., 2015] or porcine cortical bones [Baró et al., 2016]. The work from Baró et al. [2013] also inspired other researchers to study if these seismological laws are also present in different systems showing avalanche behaviour: ethanol-dampened charcoal [Ribeiro et al., 2015], confined-granular matter under continuous shear [Lherminier et al., 2019] or AE of crumpled plastic sheets [Costa et al., 2016], to mention few of them.

3.3.1 Open questions in labquakes

Similarly to astrophysics and other observational sciences, it is not possible to study the causes and effects under controlled protocols in earthquakes. Some non-tectonic earthquakes can be induced through other mechanisms such fluid injection [Cesca et al., 2014; Baró et al., 2016] or bomb explosions [Plenkers et al., 2011; Kornei, 2019]. However, the 'experimental conditions' in these cases are exceptional and can hardly be controlled. Despite the enormous difference in the characteristic scales (see Sec. 4.3 for more details), the main seismological laws are also fulfilled in labquakes. Therefore, labquakes can be considered as a benchmark to study earthquakes in a laboratory where the experimental conditions can be easily controlled. Experiments can be repeated many times under different conditions and with very good statistics. This can be a good starting point in order to develop, for instance, forecasting algorithms.

Nevertheless, as it was exposed in the previous Sec. 3.3, some values of the exponents characterizing the main seismological laws in labquakes are found to be quite different from those for earthquakes. Therefore, one can realize that there are different open questions regarding their comparison with real earthquakes. Just to mention few of them:

- Does the exponent of the GR law for labquakes depend on the experimental conditions? For instance on the compressional driving mechanism?
- Does the exponent of the GR law for labquakes depend on the internal structure of the compressed material? Is there a material which exhibits an exponent closer to the one of earthquakes?
- If the previous question has an affirmative answer, are the other laws of statistical seismology fulfilled with the same parameters as for earthquakes?
- To which extent it is possible to strengthen the scale-invariance property by finding a validity for the GR law holding for more orders of magnitude?
- Is it possible to obtain information about the spatial degrees of freedom?

As mentioned in Chpt. 1, universality implies that power-law exponents characterizing avalanche observables can be considered to be the equal. But, what does exactly mean that power-law exponents are equal? How can this statement be quantified? The answer to this question needs to be rigorously studied with the proper statistical techniques. However, very few works provide robust statistical methods to check avalanche criticality and universality [Kagan, 2013; Deluca and Corral, 2013; Deluca et al., 2016]. One of the main goals of this PhD thesis is to provide rigorous statistical techniques in order to characterize avalanche criticality and universality in empirical data. By means of different statistical tools and always considering their limitations, the works presented in the Results part of this PhD thesis try to answer the first, second and fourth questions.

In Chpt. 6, displacement-driven compression of the porous glass Vycor (the same material as the one used by Baró et al. [Baró et al., 2013]) are performed with the aim of studying whether the exponent of the GR law is altered or not [Navas-Portella et al., 2016]. With the aim of answering the second question, a displacement-driven compression experiment of charcoal is performed by finding an exponent of the GR law much closer to the one found for earthquakes.

With the aim of answering the fourth question, more displacement-driven compression experiments of the porous glass Vycor have to be performed by changing the experimental parameters. These changes in the experimental conditions enable to study the distribution of the AE energy in different energy windows. A statistical technique for merging different datasets is used in order to check the validity of the GR along the different observation windows. This technique is also useful in order to try to merge charcoal labquake with real earthquake catalogs. All these results are shown in Chpt. 7.

The rest of questions are still open but this thesis will try to provide the needed tools so as to answer them in future works (see the Conclusions part for more details).

Part II

Methodology

Chapter 4

Experimental Techniques

In this chapter, the experimental techniques used to record labquakes and their corresponding observables are defined. It is crucial for the further statistical analysis to understand how avalanches are defined in the acoustic emission experiments.

Earthquake detection and how the observables characterizing them are defined has been studied since long ago. It is not the intention of this chapter to expose all the details that seismology has been dealing with along decades. Some basic concepts about earthquake detection are introduced together with a brief explanation on how to obtain the moment magnitude scale.

4.1 Labquake detection

In general terms, a labquake is an acoustic emission (AE) event occurred during a compression experiment of a porous material. As it was exposed in Chpt. 3, in spite of the difference in length, time and energy scales, these AE events exhibit some statistical similarities with real earthquakes [Salje et al., 2011; Baró et al., 2013]. In this sense, labquakes are often considered as synthetic earthquakes that can be reproduced in a laboratory. Some questions need to be addressed:

1. Which materials generate labquakes? What is their internal structure?
2. Which is the mechanism generating labquakes?
3. Which is the experimental set-up for labquakes detection?
4. How are the characteristic observables of a labquake defined?
5. Which experimental limitations exist in labquake detection?

The first question is considered in Sec. 4.1.6, where the porous materials which are used to generate the labquake catalogs analysed in this PhD thesis are presented.

The second question can be answered by stating that labquakes are generated during uni-axial compression experiments without lateral confinement of porous samples. The compression can be performed in two different ways: either by controlling the applied force on the sample and measuring the displacement (force-driven compression), or by imposing a displacement rate over the sample and measuring the force opposed by the material (displacement-driven compression). Labquakes generated by means of force-driven compression experiments have been broadly studied [Salje et al., 2011; Baró et al., 2013; Nataf et al., 2014]. The displacement-driven compression mechanism is the object of study in this PhD thesis.

The experimental set-up through which displacement-driven compression experiments of porous materials are performed is shown in Fig. 4.1. All the parts of the experimental set-up are detailed in Secs. 4.1.1, 4.1.2 and 4.1.4. The two fundamental pieces of the experimental set-up are the compression mechanism and the AE recording. The compression is the external mechanical driving which induces failures in different regions of the sample. These failures are detected as AE sources by one or more transducers. The detected AE signal is processed to generate the labquake catalog. In order to answer the last three questions, this experimental set-up together with the definition of labquakes are presented in Sec. 4.1.2.

It must be mentioned that, contrarily to what happens in earthquake detection, the experimental setup used in this work does not allow labquake location in the sample because only one AE transducer is used.Therefore, spatial degrees of freedom are not accessible for the labquakes studied in this work. All the experiments exposed in the Results part have been performed in the Laboratory of Materials of the Condensed Matter Physics Department at the University of Barcelona.

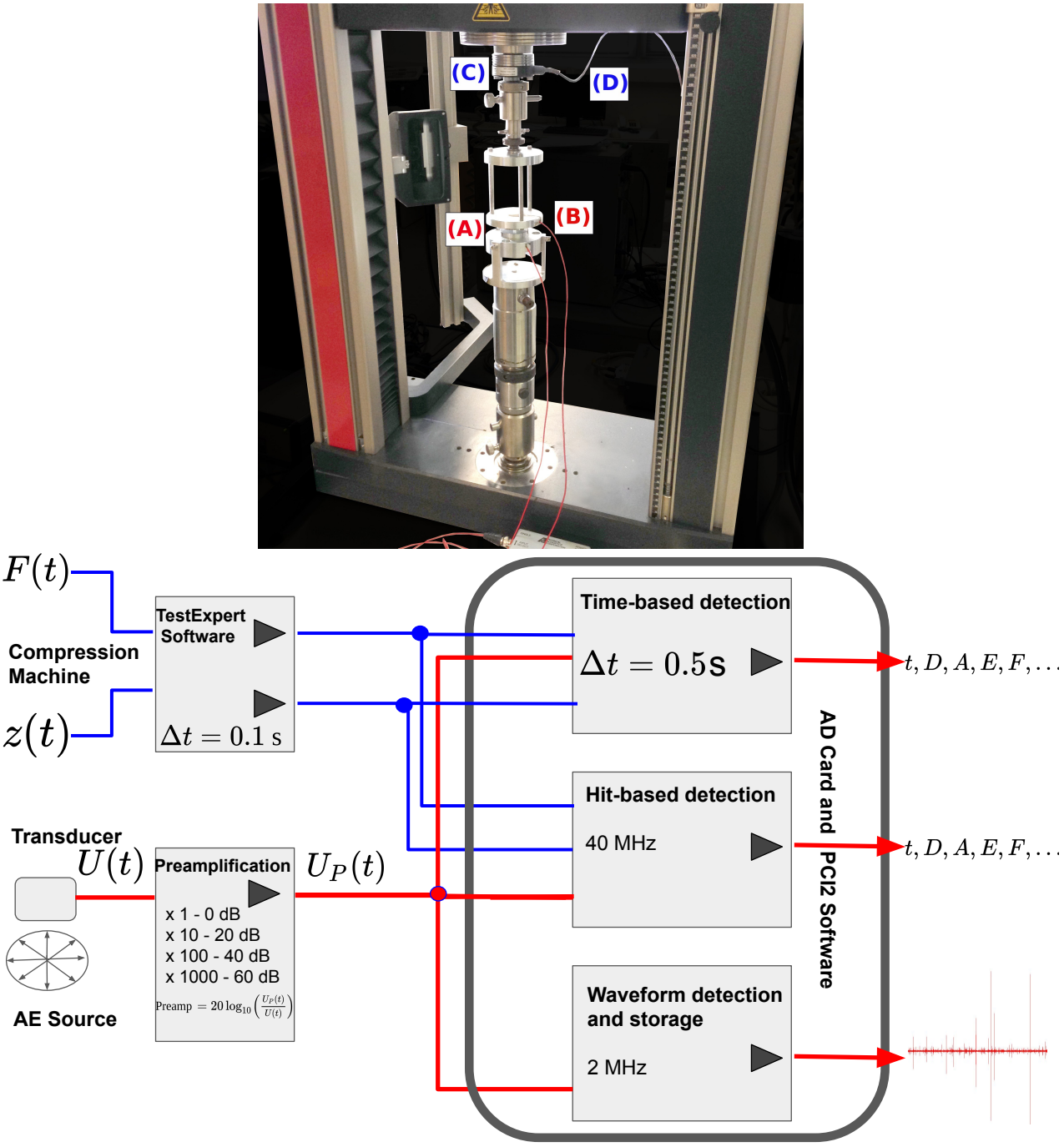

Fig. 4.1: Experimental set-up for the compression mechanism. (A) Compressional plates. Transducers are imbedded in the plates. (B) AE chanel connecting the transducer with the preamplification. (C) Load cell measuring the force $F(t)$. (D) Output chanels of the force and the displacement $z(t)$ with respect the original position to the TestExpert software. Photo edited by V.N.P. From this compression mechanism one can obtain the value of the force $F(t)$ opposed by the sample and the distance between the two plates $z(t)$. A piezoelectric transducer is placed in one of the compressional plates where the porous sample is placed. The transducer detects the AE signal $U(t)$. Bottom, schematic representations of the different stages through wich an the AE signal (in red) as well as the additional measured parameters (in blue) ($F(t)$ and $z(t)$) are recorded.

4.1.1 Compression and associated parameters

The main components of the experimental setup for the compression of the samples are:

- Conventional test machine Z005 (Zwick/Roell): Test machine which represents the external driving over the sample.
- Load Cell Xforce P (Zwick/Roell): Measures the force oposed by the sample with a maximal nominal force of 5 kN.
- Test Expert Software [Zwick-Roell, 2019]: Processes and records the force and the displacement with respect the original position.

Uni-axial compression experiments of porous samples are performed in a conventional test machine Z005 (Zwick/Roell). Samples with no lateral confinement are placed between two plates that approach each other at a certain constant rate $\dot{z}$ (see top panel in Fig. 4.1). With the aim of compressing close to the adiabatic regime in which avalanches do not overlap each other [Sethna et al., 2001], the compression rate is chosen to be as low as possible. Note that, strictly speaking, the adiabatic regime cannot be achieved experimentally because the relaxation time scale of some events is too long in comparison with the driving rate. Therefore, one must assume that not all the avalanches can be completely discriminated and some overlap will occur (see Sec. 4.1.2).

The displacement $z(t)$ with respect the original position as well as the force $F(t)$ can be measured at this stage and output for further analysis (see blue links in Fig. 4.1 bottom). For more details about the detection of this additional experimental parameters, see Sec. 4.1.4.

4.1.2 Acoustic emission (AE) recording

AE is a non-destructive technique that have been broadly used in order to study properties of materials [Scruby, 1987] such as damage evaluation, structural phase transitions [Vives et al., 1995], domain-wall motion [Kusanagi et al., 2005], etc. [Scruby, 1987]. These techniques are based on placing one or more sensors in the material in order to detect a wide range of events taking place in it. The equipment used in this work consists of:

- Piezoelectric transducers (micro-80 from Mistras Group France) working in a frequency range of 0.1 Hz-10 kHz. They detect pressure changes and convert them into an electric voltage $U(t)$.
- Pre-amplifiers: Amplification of the original signal $U(t)$ into $U_P(t)$ for further analysis.
- Analog-Digital (AD) card with 18 bits resolution: Receives and processes the electric signal U_P with a sampling rate of 40 MHz. The card can also detect external analog parameters at a sampling rate of 10 kHz.

- AE-Win PCI2 Software (Version 1.30) [Mistras, 2004] Real-time analysis of the AE signal: definition of AE hits, determination of avalanche observables, etc.

During the compression experiment, recording of an AE signal is performed by using a piezoelectric transducer embedded in one of the compression plates. Failures occurring in the sample release energy in form of pressure waves [Scruby, 1987] which is detected by the piezoelectric transducer and converted into a voltage due to the piezoelectric effect [Gautschi, 2002].

It is very common in AE literature to work with dB units instead of voltages [Scruby, 1987]. The relationship between this two units is given by:

$$U(t) \text{ (in dB)} = \left[20 \log_{10} \left(\frac{|U(t)|}{U_0} \right) \right] , \tag{4.1}$$

where $U_0 = 1\mu$V is a reference voltage. Note that, in this logarithmic scale, an increase of 20 dB corresponds to a difference of one order of magnitude in voltages.

The detected AE signal $U(t)$ is then pre-amplified and band filtered (between 20 kHz and 2 MHz) in order to be ready for the following step, in which signal analysis is performed. Signal pre-amplification is necessary if one wants to record small AE events. The largest pre-amplification is of 60dB, which actually means to multiply the value of the signal by 1000 whereas the case without pre-amplification corresponds to 0 dB (see Fig. 4.1). The pre-amplified signal is denoted by $U_P(t)$.

The pre-amplified signal $U_P(t)$ arrives to the PCI-2 acquisition system (see bottom panel in Fig. 4.1). The AD card together with the AE-Win PCI2 software acts as a signal processing center detecting the AE signal (red links in Fig. 4.1) as well as the force $F(t)$ and displacement $z(t)$ from the compression setup (blue links in Fig. 4.1). This signal processing device can work in three different modes:

- Hit-based detection mode: Reads the pre-amplified AE signal $U_p(t)$ at 40 MHz in order to detect individual AE hits. Given that this is the best resolution at which this signal processing device can work, this signal has to be understood as the time series $\mathbb{X}(t)$ exposed in Sec. 2.1. Note that this time series is affected by many external factors: noise in the transducer, in the AD card, in the wires through which the signal travels, the labquake itself, rebounds from the same failure, etc. Thus, a thresholding procedure is needed in order to define AE hits from the time series $\mathbb{X}(t)$. The main purpose of this thresholding procedure relies in establishing a correspondence between AE hits and labquakes. For more details about how to obtain a labquake catalog see next Sec. 4.1.3.

- Time-based detection mode: Records the values of external experimental parameters ($F(t)$ and $z(t)$, see next Sec. 4.1.4) as well as the cumulated energy every Δt (usually Δt seconds). The energy computed in this mode corresponds to the cumulated energy in Δt. This detection mode is important in order to have an additional recording of the

force and the displacement apart from the one provided by the TestExpert software (see Sec. 4.1.4).

- Waveform detection mode: Reads the pre-amplified AE signal $U_p(t)$ and records the waveform at 2 MHz. This detection mode is activated manually when the user is interested in recording the AE waveform in certain time window of the experiment. For an example of a waveform obtained during a experiment, see Fig. 4.2.

The three detection modes can work simultaneously. However, care must be taken when using the waveform detection mode because, although the sampling rate has been reduced with respect the hit-based detection, it is easy to run-out of memory in the computer and stop the recording of all the detection modes.

All the experimental measurements are always subjected to the resolution of the experimental device and are thus discontinuous. Independently on the sampling rate, the AD works with 18 bits precision [Mistras, 2004]. This should be kept in mind when considering some of the measures as real values (energy or voltage).

4.1.3 From the AE signal to a labquake catalog

4.1.3.1 Definition of AE hits

Working in the hit-based detection mode, the main purpose of this section is to establish a thresholding procedure to the signal $U(t)$ in order to define AE hits.According to this thresholding procedure the information of the signal will be expressed as a temporal marked point process. As mentioned in Sec. 2.1.4, it is not always possible to associate an AE hit to a labquake unambiguously. Depending on the set of parameters chosen for the thresholding procedure, rebounds of the same labquake can be considered as a part of an AE hit, several labquakes can be considered as one single AE hit or labquake overlapping may occur. As it will be seen in next Sec. 4.1.3.1.3, it is possible to have a number of AE hits which is stable for a wide range of values of the parameters involved in the thresholding procedure (essentially the HDT). This would indicate that, problems regarding hit overlapping and splitting would be reduced. Nonetheless, since overlapping of labquakes cannot be resolved just from the AE signal (see Sec. 2.1.4), working in this set of parameters involved in the thresholding procedure would not ensure a clear correspondence between AE hits and labquakes. Taking into account these limitations and, as an abuse of language, the terms labquake, avalanche and AE hit are used indistinctly along this PhD thesis.

Provided that the AE signal is a voltage and can take positive and negative values, the two-sided thresholding procedure exposed in Sec. 4.1.3 is the most convenient in order to define avalanches. An AE hit starts at the time t when the signal $U(t)$ crosses a fixed detection threshold (THR) and finishes at time $t+D$ when the signal remains below threshold ($|U(t)| < THR$) for a hit definition time (HDT) from $t+D$ to at least $t+D$+HDT (see Fig. 4.2).

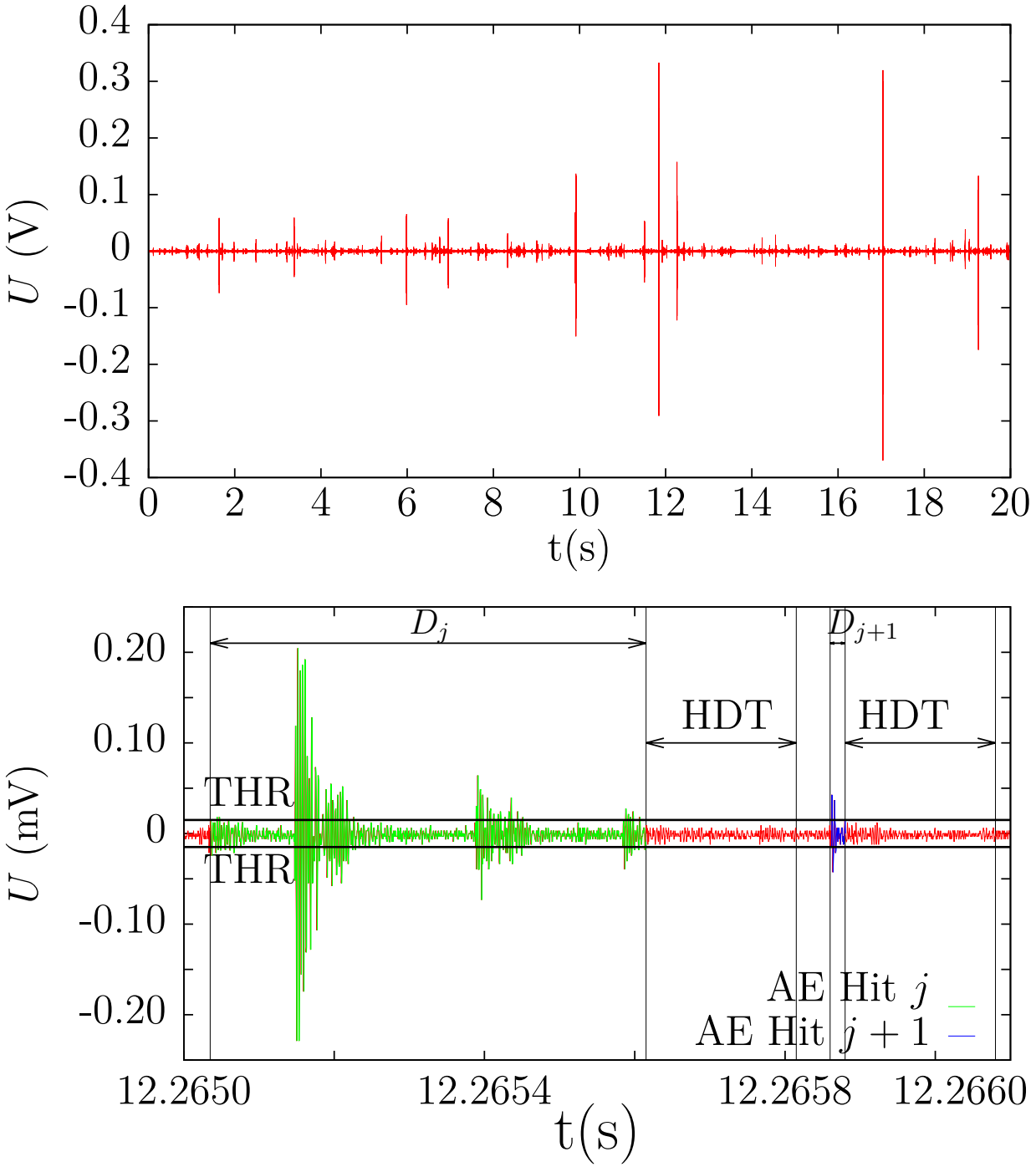

Fig. 4.2: (a) Real signal $U(t)$, i.e. the one received in the transducer (not the preamplified) measured in V during $20s$ at 2 MSPS in the experiment performed with a charcoal sample. The largest event has an amplitude of 70 dB. (b) Zoom of the AE signal in a time interval of 1.1 ms around 12.265 s. In this time intervals, hits j (highlighted in green) and $j+1$ (highlighted in blue) are detected with durations D_j and D_{j+1} respectively. The threshold THR is set at 23 dB (14.1μV) and the HDT=200μs. The HLT is not represented because this time is so small (HLT=2 μs) that cannot be appreciated at the present resolution.

Experimental protocols as well as limitations in the definitions of AE hits are exposed in the following paragraphs.

4.1.3.1.1 Blank measurements

Conditions in the laboratory can vary due to several factors thus altering the electric background noise in the electronic devices. This is the main reason why a previous blank measurement must be performed before any experiment. This blank measurement provides information about how the values of the threshold should be for a given pre-amplification in order to ensure a good recording of the AE signal. The background noise level will be a reference in order to set the value of the threshold. Blank measurements are performed without compressing sample but the compressional mechanism can be either working or not. These experiments are also useful in order to check the standard deviation (error bars) of some quantities, such as the force, when no labquakes are occurring.

4.1.3.1.2 Threshold

As it has been already mentioned, the threshold must be high enough so as to avoid the background noise from the experimental devices, electronic lectures and in the wires. There are two major source of noise affecting the AE signal at different stages of the experimental set-up. The first electronic noise is present in the transducer as well as in the channels carrying the AE signal. Values of the the voltage below a threshold $U_{THR}^{(Trans)} \simeq 23$ dB (14.1 μV approximately) are affected by it. The supper-script $(Trans)$ refers to the threshold in the transducer. This implies that the electric signal $U(t)$ detected in the transducer must be above this value in order to avoid this noise. As it can be appreciated in the bottom panel of Fig. 4.2, where the THR has been set at this level of noise 14.1 μV, no clearly defined AE events can be identified below this THR and the AE signal $U(t)$ fluctuates around zero.

The second noise is due to the AD card lecture and disables records below a certain value $U_{P,THR}^{(AD)}$.The supper-script (AD) refers to the threshold in the AD card. This implies that, pre-amplified voltages below $U_{P,THR}^{(AD)}$ are not be correctly detected. Consequently, when the pre-amplification is changed, the value of the threshold for the signal in the transducer $U(t)$ must be accordingly varied. Both values of the threshold $U_{THR}^{(Trans)}$ and $U_{P,THR}^{(AD)}$ can be sensible to the laboratory conditions and this is the reason why performing blank measurements is important in order to avoid spurious lectures.

4.1.3.1.3 Hit definition time (HDT) and hit overlapping

As it was discussed in Sec. 4.1.3, the existence of the hit definition time (HDT) can favour hit overlapping as well as hit splitting. In this sense, the definition of an AE hit is also intimately tied to HDT. Note that the compression driving rate can influence which HDT is selected. If the adiabatic regime could be reached, one would just have to choose an HDT in order to avoid the rebounds of the same labquake. Nevertheless, strictly speaking, the adiabatic

regime cannot be experimentally achieved and a labquake can nucleate in other sample regions with a very short inter-event time. If the HDT is too large in comparison with this interevent-time, different labquakes will be considered in the same AE hit thus causing hit overlapping. This obscuration effect might appear after big events and it can affect the statistical analysis implying undercounting effects [Lippiello, E. et al., 2010]. One could be tempted to select a smaller HDT in order to avoid this effect. However, if an AE event has not really finished and the AE signal has remained below threshold for more than the predefined HDT, a single AE hit might be splitted into several ones. This effect is known as hit splitting. Therefore, a compromise between these two effects must be achieved in order to avoid them as much as possible.

In order to study these effects, the waveform of the AE signal has been recorded during 20 seconds in the waveform detection mode (see Fig. 4.2). Despite the sampling rate in this case (2 MHz) is significantly smaller than the one in the hit-based detection (40 MHz), this AE waveform has been useful in order to study the effects of the threshold as well as the HDT in

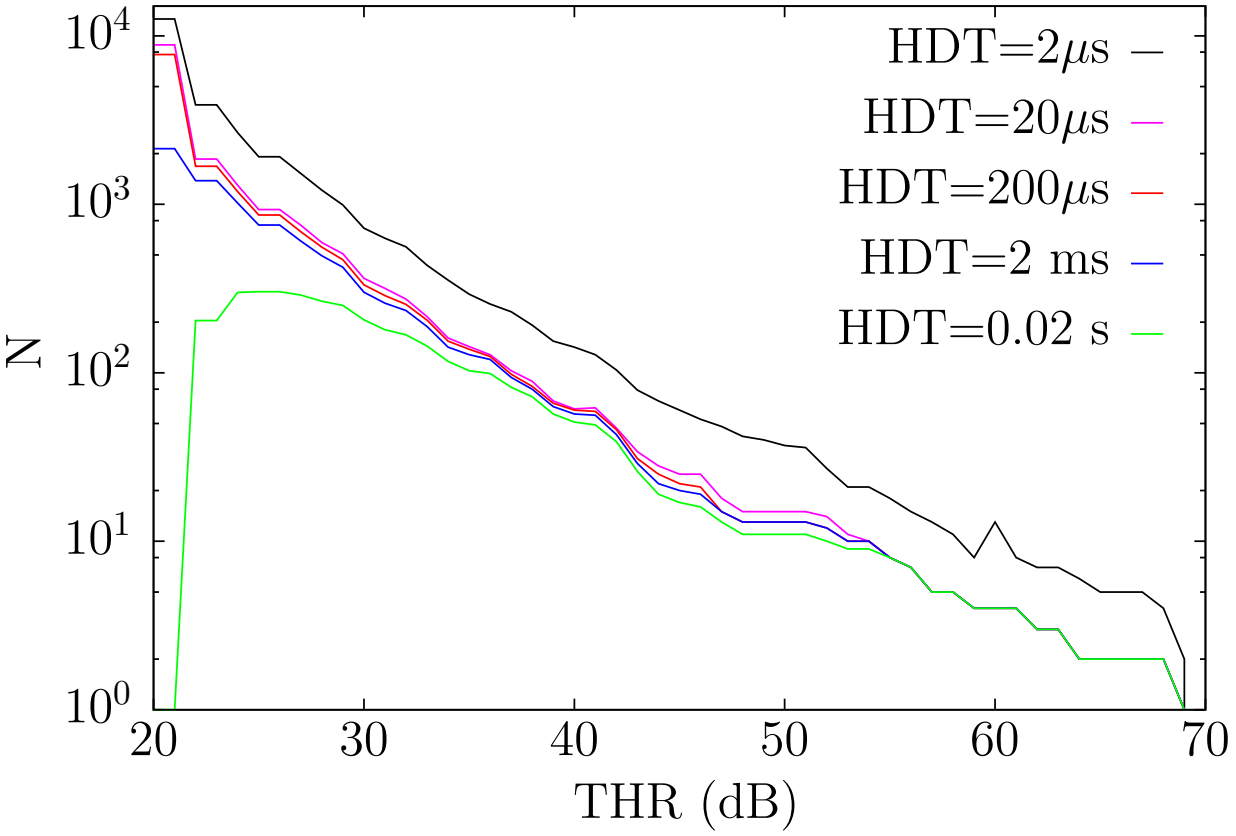

Fig. 4.3: Number of avalanches as a function of the threshold THR of the original signal in dB for different values of the HDT for the signal in Fig. 4.2. HLT= 2μs

the AE hits. The number of detected AE hits as a function of the threshold in dB for different values of HDT is shown in Fig. 4.3.

Hit overlapping can be clearly observed in Fig. 4.3, where the curve with the largest HDT is always accounting for less events. This fact can also be appreciated in hit A in the bottom panel in Fig. 4.2, where the AE signal has apparently fulfilled a relaxation time and a re-intensification of the signal crosses the THR before the HDT has finished. Let us remark that it is not possible to distinguish whether this corresponds to a reflection of the same labquake or it is actually a new one. As mentioned in Sec. 2.1, spatial information is needed in order to determine if different avalanches actually overlap or not. Hit splitting can also be appreciated for the curve with the smallest HDT in Fig. 4.3. It can be seen that, over a threshold of 23 dB, there exists a regime along two orders of magnitude in HDT in which the number of events does not show remarkable differences.

Figure 4.4 shows different CCDFs of the AE hit durations for different values of HDT for a fixed value of the threshold of THR= 25 dB. As it can be observed, the smaller the HDT

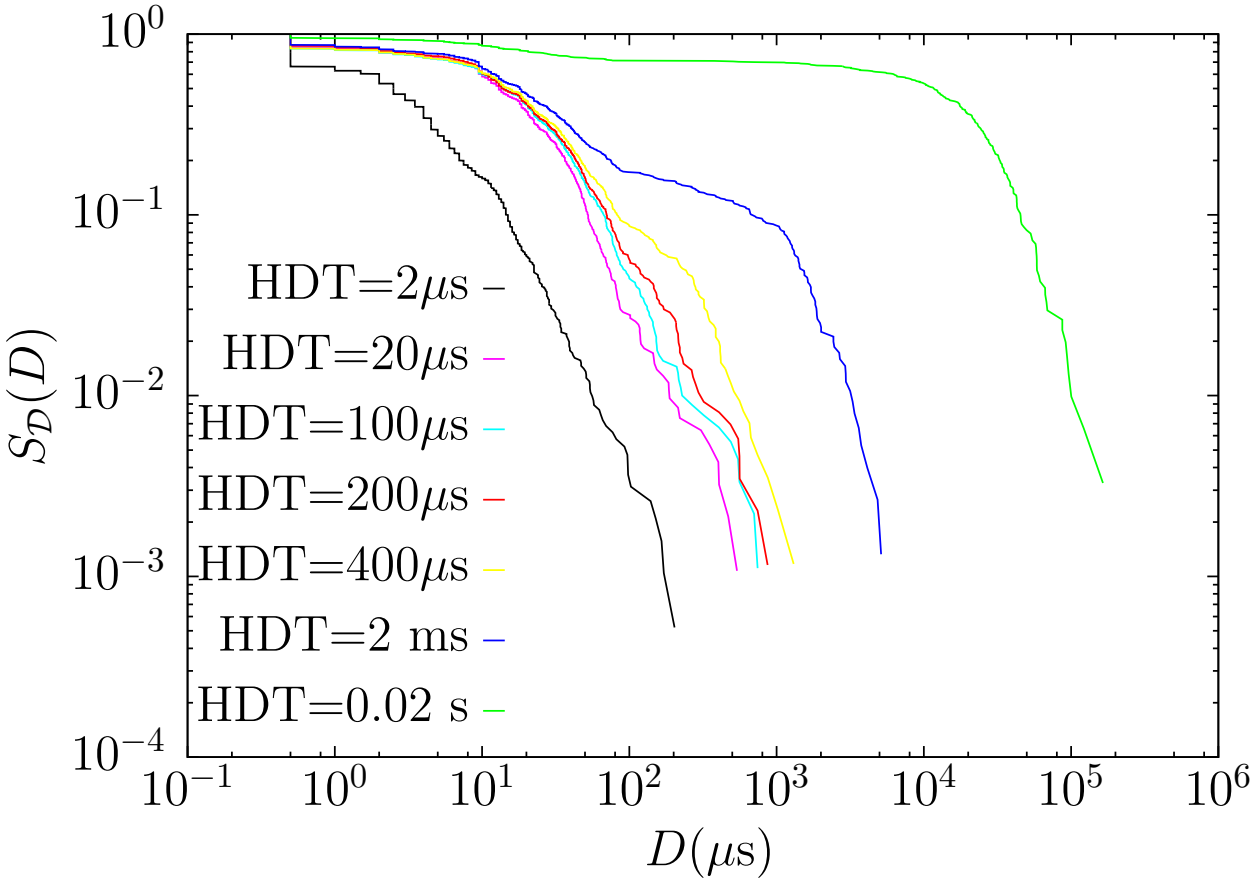

Fig. 4.4: CCDF of avalanche durations for THR=25 dB for different values of the HDT for the signal in Fig. 4.2. HLT= 2μs

the shorter the durations. However, there exists a regime from 20μs to 400μs in which the distributions of AE hit durations do not exhibit relevant differences. This implies that the distribution of durations, as well as amplitudes and energies, are quite independent of HDT in this regime. Lebyodkin et al. [Lebyodkin et al., 2013] shown that the effect of the HDT on the energy distribution is not very relevant. Taking into account this stability regime the value of HDT which has been used for all the experiments is HDT= 200μs.

4.1.3.1.4 Hit Lockout Time (HLT)

No labquakes are considered in a Hit Lockout Time (HLT) immediately after $D + HDT$. This time is crucial if one is interested in the first stages of an AE hit when the avalanche process reaches the highest values. In this case, in order to obviate the avalanche decay or coda, the HDT is set as low as possible and the HLT as high as possible in order to obviate the coda and not to account for reflections. The existence of this time might involve a loss of information if a new real AE event occurs before this time has finished. The presence of this time is thus quite meaningless if one desires to study the entire avalanche. For this reason, the HLT is an irrelevant parameter in the experiments exposed in this work and is set as low as possible: HLT= 2μs.

4.1.3.1.5 Saturation effects

Pre-amplification can be useful when one is interested in detecting small AE hits. However, some values of the pre-amplified signals are so large that cannot be correctly detected by the acquisition system. The AD card can detect signals up to 140 dB (after pre-amplification). This implies that, for pre-amplifications of 60, 40 and 20 dB, the largest values of the AE signal $U(t)$ that can be detected are of 80, 100 and 120 dB respectively (see Table 4.1). This fact leads to a saturation in the amplitude and, consequently, an underestimated energy of the AE event which exceeds these values under these pre-amplifications (See next Sec. 4.1.3.2). This effect can be immediately observed in the distributions of the amplitude, Fig. 4.5, where there is an excess of AE hits in the last bin for an experiment with pre-amplification 60 dB and threshold 23 dB.

Preamp (dB)	U_{top}
0	140
20	120
40	100
60	80

Table 4.1: Maximum values U_{top} of the AE signal $U(t)$ that can be recorded by the AD card for each pre-amplification.

4.1.3.2 AE hit observables

4.1.3.2.1 Labquake amplitude

The AE hit, avalanche or labquake amplitude A recorded in dB follows the expression:

$$A = \left[20 \log_{10}\left(\frac{|V|}{V_0}\right)\right] \text{ in dB}, \tag{4.2}$$

where $V = \max|U(t)|$ is the peak voltage achieved by the AE signal $U(t)$ during the AE hit, $V_0 = 1\mu$V is a reference voltage, and the brackets round the value to its nearest integer in dB. Note that the terminology A refers to the amplitude in dB whereas the peak voltage V will be referred to simply as amplitude, in agreement with previous literature [Scruby, 1987]. The amplitude A is an observable which can take continuous values but it is recorded in dB. From the values of A one can obtain the values ψ of the discretized peak-voltage:

$$\psi = g(A) = V_0 10^{A/20} \text{ in } \mu V. \tag{4.3}$$

As the values of A are integer, the values of ψ will no longer be integer but they will collapse into a set of values $\{\psi_1, \psi_2, ..., \psi_j, ..., \psi_k\}$ measured in μV.

Special attention must be paid to correct the preamplification. Let us rewrite the amplitude read by the AD card in Eq. (4.2) as:

$$A = \left[20 \log_{10}\left(\frac{|V|}{V_0}\right)\right] = \left[20 \log_{10}\left(\frac{|V_P|}{V_0}\right)\right] - \text{Preamp in dB} \tag{4.4}$$

where V is the peak-voltage read by the transducer, i.e. the one corresponding to the original signal $U(t)$, whereas $V_P = \max|U_P(t)|$ corresponds to the pre-amplified peak-voltage from the $U_P(t)$ signal read by the AD card.

4.1.3.2.2 Labquake energy

The energy E of an AE hit, avalanche or labquake occurred at time t is determined as

$$E = \frac{1}{R}\int_t^{t+D} U^2(t)dt, \tag{4.5}$$

where R is a reference resistance of 10 kΩ related to the electrical circuit. Let us recall that this integral is computed for a $\Delta t = 2.5 \times 10^{-8}$s, which corresponds to the sampling rate of 40 MHz of the hit-based detection mode. By considering that the characteristic voltages and durations of the smallest AE hits are of the order of μV and μs, the order of magnitude associated to the energy is about 10^{-22}J. Therefore, a convenient scale to measure the energy of AE hits is the aJ= 10^{-18}J.

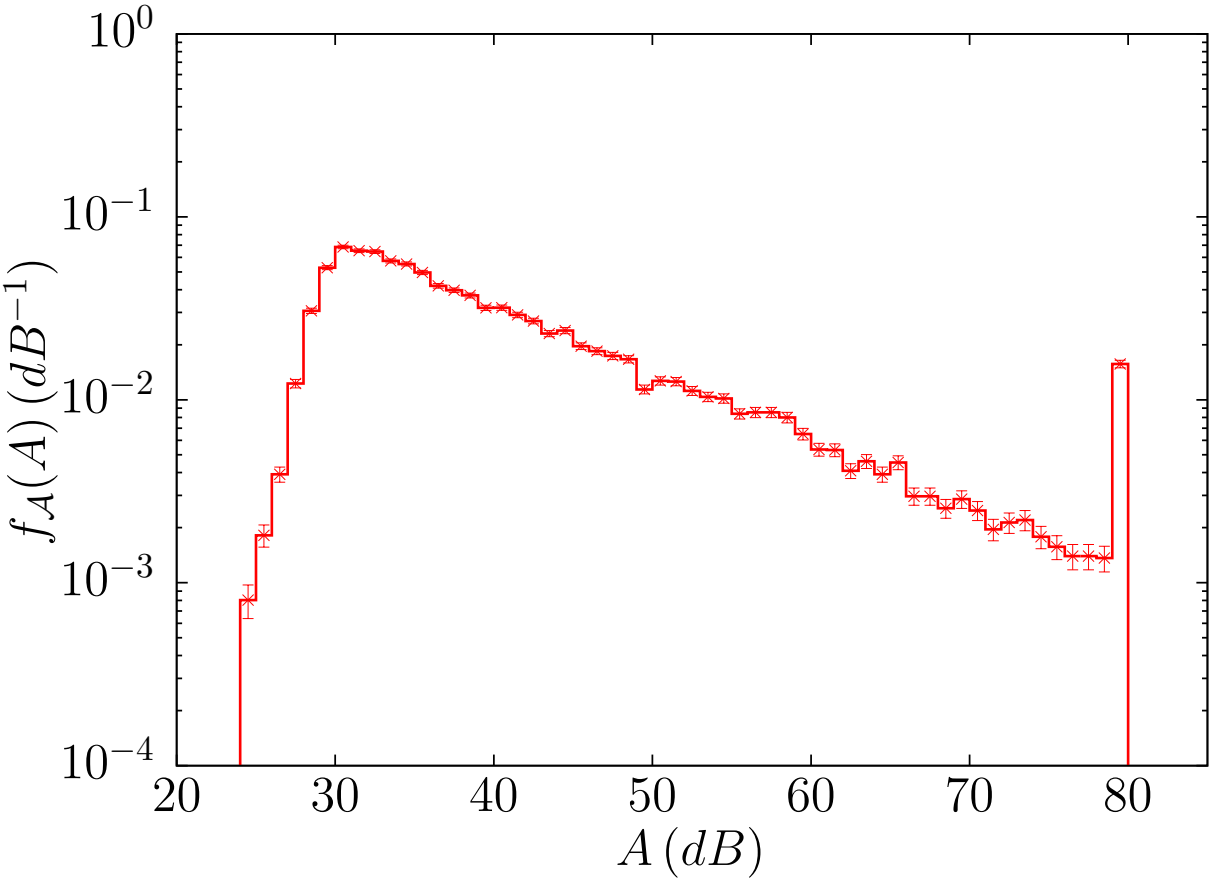

Fig. 4.5: Estimated probability mass function of the amplitude in dB for a experiment performed at Preamp of 60 dB and Threshold THR of 23. Details about the experiment are exposed in Table 4.2. Error bars are estimated as the standard deviation for each bin.

4.1.3.2.3 Modulation of AE hit observables

An additional effect that must be taken into account is the modulation of the AE wave when it travels from its origin to the AE sensor [Stehly et al., 2008; Scruby, 1987]. Modulation means that the medium can attenuate the radiated energy of the AE waves by absorbing certain frequencies. Modulation effects can also appear due to the response of the transducer to certain frequencies, thus detecting better certain frequencies than other ones. These effects result in a distortion of the shape of the measured AE signal. Given that the distances from the labquakes source to the transducer are small, one can assume that this modulation does not affect the measured value of the avalanche amplitudes and energies [Baró i Urbea, 2015]. This implies that there are not preferent ranges and the measured amplitudes and energies are proportional to those associated to the original source. In other words, if different AE hits whose amplitudes are A and $2A$ respectively are detected, one expects that the original

sources that have generated these hits keep the same proportion.

4.1.4 Additional experimental parameters

As it was mentioned in Sec. 4.1.1, the compressional set-up can also measure the values of the displacement $z(t)$ with respect the original position as well as of the force $F(t)$ opposed by the sample. On the one hand, the displacement $z(t)$ is measured by means of the test machine Z005 (Zwick/Roell) which simply calculates the distance travelled since the beginning of the experiment by the rotation of the mechanical engine. On the other hand, the force opposed by the material is measured by means of the load cell. Both values are measured by the testXpert Testing software [Zwick-Roell, 2019] where are processed and recorded every $\Delta t = 0.1$ s. Blank measurements were performed in order to evaluate the electronic noise when measuring the force. These experiments indicate that the standard deviation of the force is $\sigma = \sqrt{\langle F^2 \rangle - \langle F \rangle^2} = 9,4 \times 10^{-3}$N $\sim 10^{-2}$ N. Consequently, if a force change below this value is found, it will not be considered in the analysis.

Two channels connect the force and the displacement outputs recorded in the testXpert testing software to the AD-card which detects the AE. These channels need to be previously calibrated and are recorded in a time step of $\Delta t = 0.5$s. Therefore, displacement and force are recorded twice: one in the test machine ($\Delta t = 0.1$s) and the other one in the AE recording set-up ($\Delta t = 0.5$s). It must be noted that there might be electrical noise and losses in the wires which connect both machines. This might imply that the values read by the AD card are slightly smaller than the ones measured in the compression machine although calibration is correctly done.

Finally, despite the time resolution of these parameters is much lower than the one from the AE set-up, let us remark that these mechanical quantities can be also useful for the study of avalanches.

4.1.5 Final labquake catalog

A displacement-driven compression experiment can last few hours ($10^3 - 10^4$s) before the sample height has been enormously reduced. At the end of one experiment, one has a catalog or collection of events each of them characterized by a time of occurrence t, duration D, amplitude A in dB, energy E, and the values of the compression parameters F and z (see Fig. 4.6).

```
##########################################################################################################################
LABQUAKE CATALOG
##########################################################################################################################
Fri   Sep   9     11:07:26    2016

t                F           z           DURATION    AMP   THR   ENERGY
..........................................................................................................................
..........................................................................................................................
4844.2969078     118.6560    0.1091      16          64    63    756782
4844.3053780     118.0456    0.1091      261         68    63    25034000
4844.3401467     112.4302    0.1086      341         78    63    111020000
4844.8228182     100.2228    0.1062      189         72    63    18804000
4846.1006395     100.2228    0.1077      44          67    63    3577000
4846.1285223     100.1007    0.1076      33          68    63    4312000
..........................................................................................................................
..........................................................................................................................
5031.0206005     117.3131    0.1176      304         76    63    82599000
5031.0206015     117.3131    0.1176      1021        95    63    6004000000
5037.6222330     117.8015    0.1193      290         79    63    114428000
5046.6569832     118.4118    0.1190      48          66    63    3019000
5114.5233880     125.7363    0.1241      81          69    63    11286000
5128.3760515     127.5674    0.1256      28          65    63    1876000
5140.8452725     128.9102    0.1270      327         80    63    160816000
5141.4466025     128.7881    0.1265      241         74    63    39855000
5141.4466033     128.7881    0.1265      748         91    63    2376000000
5141.8007880     128.7881    0.1259      283         75    63    49884000
```

Fig. 4.6: Snapshot of a labquake catalog with the different measured labquake observables for two different time periods of the same experiment. The time is measured in s, force F is measured in N, the displacement z with respect the original position in mm, the duration D measured in μs, the amplitude A as well as the threshold THR are measured in dB and the energy E in aJ.

4.1.6 Samples

4.1.6.1 Vycor

Vycor is a synthetic nanoporous silica (SiO_2) glass with 40% porosity [Bentz et al., 1998] produced by Corning Incorporated [Corning, 2020]. Two different series of experiments with Vycor at room temperature were performed looking for different goals.

The first series consists of the compression of 13 different Vycor cylinders with different diameters and heights as well as different rates. A cleaning protocol before compression can be performed. It consist in cleaning the samples with a 30% solution of H_2O_2 during 24 h and dried at 130°C. It has been checked that the different cleaning protocols before the experiment do not alter the results. Such a number of experiments was performed in order to evaluate the variability of the energy exponent ϵ for these experiments. With the goal of checking whether the driving mechanism can alter the value of the energy exponent, this first series of experiments is analysed in Chpt. 6. With the aim of having the same conditions for all the experiments, the second series consists of four Vycor cylinders with the same diameters $\Phi = 4.45$mm and heights $H = 8$mm, and the compression rate is fixed at $\dot{z} = 0.005$ mm/min. In order to obtain catalogs spanning different observation windows, these experiments have been performed for different values of the pre-amplification and the detection threshold. In this case, the four experiments have been performed with the following pre-amplification values: 60 dB, 40 dB, 20 dB and 0 dB, and the respective values of the detection threshold 23 dB, 43

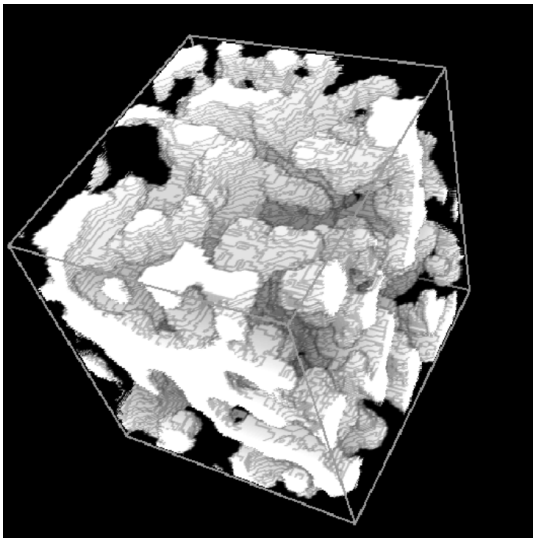

Fig. 4.7: Reconstructed image of Vycor. The sizes of the cube are about 100 μm. Figure extracted from Bentz et al. [1998]

dB, 63 dB and 83 dB referring to the signal $U(t)$, not the pre-amplified signal (in such a way that after pre-amplification the threshold always moves to 83 dB). This value of the threshold is above the level of noise of 63 dB in the AD card. This second series of experiments is analysed in Chpt. 7.

4.1.6.2 Charcoal

Charcoal is generally originated by the combustion of wood or any other organic matter in a limited supply of air [Harris, 1999]. One experiment with a prismatic charcoal sample (basis of 8.47 mm $\times$ 5.96 mm and height $H = 14.22$ mm) was performed at a constant compression rate $\dot{z} = 0.005$ mm/min with a pre-amplification of 40 dB and a value of 24 dB for the detection threshold (in such a way that after pre-amplification the threshold moves to 64 dB). This value of the threshold was set as low as possible in order to avoid parasitic noise. The sample corresponded to commercially available fine art fusains (HB 5 mm, NITRAM, Canada) [Nitram, 2020].

As an important improvement with respect the experiments with Vycor samples, a ceramic compression plate was used for this experiment instead of the aluminium one. This change of material improved the acoustic coupling between the transducer and the compression plate. Consequently, a larger number of AE events was found in this experiment. An important difference between these two materials is the degree of heterogeneity. The synthetic mesoporous silica structure of Vycor is much more homogeneous than the one of the charcoal which has been formed through different natural processes and it may contain voids and macropores. These structural differences may lead to differences in the energy exponents [Xu et al., 2019].

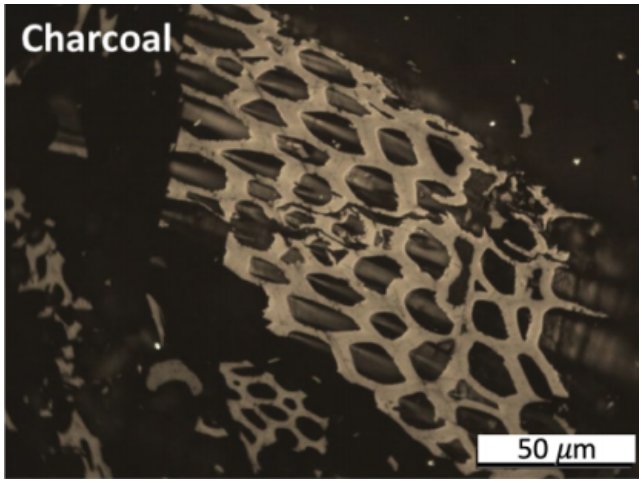

Fig. 4.8: Optical microscope image of charcoal taken on reflected light oil immersion. Figure extracted from Xu et al. [2019].

All the details related to experiments are listed in Table 4.2.

Sample	**Material**	**Φ(mm)**	H**(mm)**	$\dot{z}$ **(mm/min)**	Preamp (dB)	Threshold (dB)
V105	Vycor	1	0.5	2×10^{-3}	60	23
V11		1	1	2×10^{-3}	60	23
V115		1	1.5	2×10^{-3}	60	23
V12		1	2	2×10^{-3}	60	23
V125		1	2.5	2×10^{-3}	60	23
V205		2	0.5	1×10^{-2}	60	23
V21		2	1	1×10^{-2}	60	23
V22		2	2	1×10^{-2}	60	23
V23		2	3	1×10^{-2}	60	23
V26		2	6	1×10^{-2}	60	23
V28		2	8	1×10^{-2}	60	23
V212		2	12	1×10^{-2}	60	23
V24		2	4	5×10^{-2}	60	23
VP60		4.45	8	5×10^{-3}	60	23
VP40		4.45	8	5×10^{-3}	40	43
VP20		4.45	8	5×10^{-3}	20	63
VP00		4.45	8	5×10^{-3}	0	83
CP40	Charcoal	Prismatic	Sample	5×10^{-3}	40	24

Table 4.2: Summary of sample dimensions (Φ diameter and H the height), compression rates $\dot{z}$, pre-amplifications and thresholds for the different experiments reported in this work.

4.2 Earthquake catalogs

As it has been already exposed in Sec. 3.3, the main seismological laws fulfilled by real earthquakes are also reproduced by labquakes. One of the main goals of this PhD Thesis is to check differences in the distributions of earthquake and labquakes observables, mainly in their sizes. In this sense, it is also important to know how earthquakes are registered in catalogs. The similarities between these systems are such that the same set of questions addressed in the last section can also be formulated for earthquakes by simply changing the word labquakes by earthquakes. Despite some of these questions are still an open debate today, it is not the intention to answer them exhaustively.

1. Which materials are involved in earthquake occurrence? How is their internal structure characterized?
2. Through which mechanisms are earthquakes generated?
3. Which is the experimental set-up through which earthquakes are detected?
4. How are earthquake observables defined?
5. Which experimental limitations exist in earthquake detection?

The first question could be answered by stating that earthquakes occur in the Earth crust. It is composed by several materials whose spatial distribution depends on whether one is considering continents or ocean basins. These continental and oceanic crusts are separated from the dense Earth mantle by the sharp Mohorovicic discontinuity. Broadly speaking, continents are composed by silica based materials whereas oceanic basins are composed by basaltic rocks which are generally denser [Yeats et al., 1997].

The Earth crust is divided in tectonic blocks which are separated by locus called faults. In order to answer the second question, one could say that stress is accumulated in a region until a rupture occurs. Despite close the fault edges the mechanics of the rupture process can be very difficult to describe, the natural accepted driving mechanism is that tectonic plates are driven at constant velocity at far enough distances from the faults [Larson et al., 1997]. The failure criterion is explained in Sec. 3.1.1.1. This rupture in which a certain amount of energy is released in the form of seismic waves is known as earthquake. The third question can thus be answered by saying that these seismic waves are detected in the Earth surface as ground motion which is recorded by means of a seismometer. From a network of seismometers one can obtain the location of the earthquake, i.e. spatial determination of the hypocenter. The observable taking more relevance along this PhD thesis is related to earthquake sizes. Although the definition of earthquake size has been extensively studied and discussed in seismology, a unique definition does not exist. Along all the variables and scales, the moment magnitude scale is one of the most accepted ways of measuring earthquake size. The definition as well as the way it is measured are exposed in Sec. 4.2.2.

Given this definition in which an earthquake can be considered as the original pulse that has been generated from the rupture process (see Fig. 4.9), there are many experimental limitations that may influence its detection: the attenuation of the seismic waves due to the scattering, reflection of seismic waves, noise from the electronic devices detecting the ground motion, energy dissipation, etc. The signal recorded in a seismogram is thus affected by of all these factors. In fact, the underlying principles and limitations found in earthquake detection are very similar to those for labquake detection in Sec. 4.1. However, additional limitations are found in earthquake detection. The first important limitation is due to the small size of the detecting device with respect the characteristic size of an earthquake. This implies that a very small fraction of the earthquake radiated energy is detected by seismometers. A second limitation arises from the fact that soil conditions can vary from the earthquake source to the seismometer emplacement and isotropy and homogeneity assumptions are quite far from reality. The third important limitation is due to the differences in seismometers, their calibration due to different environmental factors such as the barometric pressure, the temperature, the magnetic field, among others [Wielandt, 2002]. These conditions make earthquake information very difficult to compare when it is obtained obtained from different seismometers.

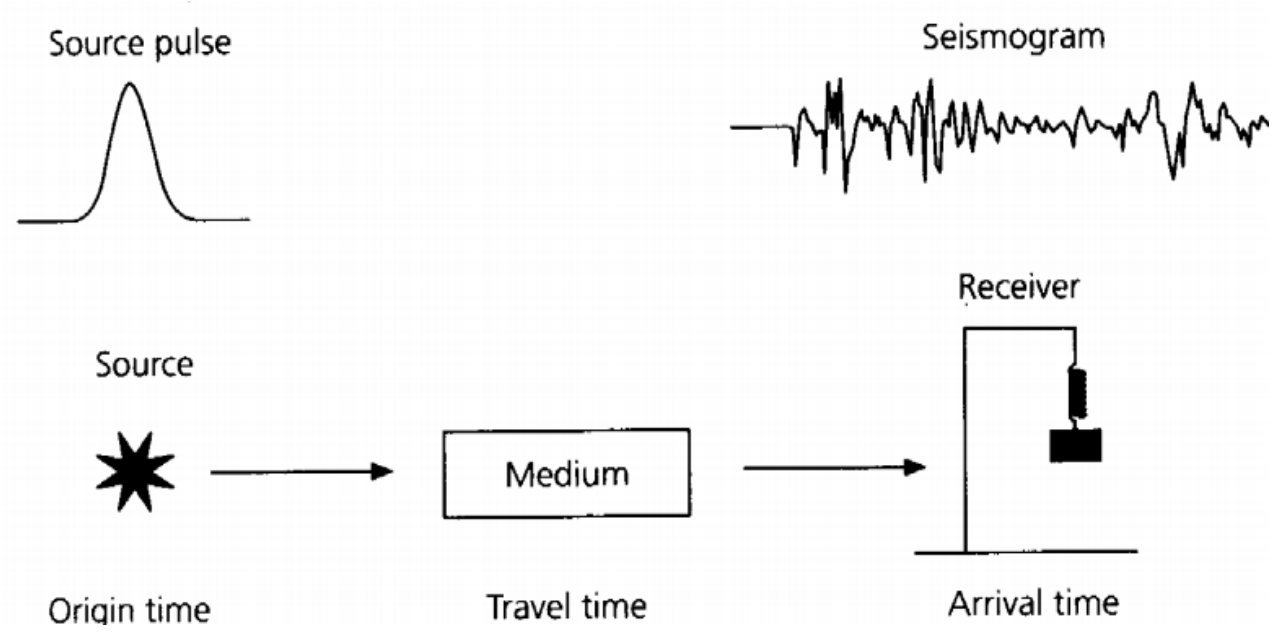

Fig. 4.9: Schematic representation of the original earthquake rupture (source pulse) and the proccess through which it is recorded in a seismogram. Figure extracted from Stein and Wysession [2003]

4.2.1 Recording earthquakes

Earthquakes are recorded by means of seismometers which detect ground motions. Seismometers generate seismograms that can be analyzed in order to extract information from the earthquake. In general, a seismogram can either measure the ground displacement or the velocity. A classic seismometer can be easily understood as a damped harmonic oscillator whose design is controlled by the nature of the damping material as well as its natural frequency [Yeats et al., 1997; Stein and Wysession, 2003]. Seismometers have constantly evolved in the last century but their basic operation is based on a coil attached to the mass of the damped harmonic oscillator that moves in a magnetic field thus generating an electric current. This electric signal can be digitally analysed in order to generate the seismogram [Stein and Wysession, 2003].

One of the most revolutionary advances in earthquake detection was the break-in of broad band seismometers [Yeats et al., 1997]. These devices allow to register earthquakes in a wide range of frequencies. As it is explained below, the frequency detected by the seismometer is intimately tied to the size of the earthquake which is able to detect. In order to provide the orders of magnitude involved in earthquake detection, one must take into account that an earthquake can last from seconds to few minutes ($\sim 10^{-1} - 10^{2}$s) [Shoji et al., 2005; Goodfellow and Young, 2014]. This implies working in a frequency range typically of $0.01 - 10$ Hz. For extremely large earthquakes on can also use strong-motion accelerographs, which measure the acceleration of the ground as a fraction of the Earth gravity g.

From a seismogram one can extract information about the observables characterizing an earthquake. Let us recall that P waves are the first ones that can be identified in a seismogram at a time t_P whereas S waves are slower and are identified afterwards at a time t_S. By using a network of seismometers one can locate the hypocenter by assuming that the velocities of P and S waves as a function of the distance to the source are known. By taking these into account, one can compute the distance to the source by means of the time difference between the arrival of P and S waves $\Delta t = t_S - t_P$. Then one can assume that the source is placed in the surface of a sphere with that radius. By computing these spheres from different seismograms, the location of the hypocenter is thus reduced to a geometrical problem [Stein and Wysession, 2003].

4.2.2 Earthquake size: the moment magnitude scale

The moment magnitude scale [Hanks and Kanamori, 1979; Lay and Wallace, 1995; Stein and Wysession, 2003] is one of the most accepted ways of determining earthquake size because it is defined through a directly measurable physical magnitude: the seismic moment. The direct dependence with a physical quantity has some advantages with respect other magnitude scales based on empirical assumptions and conditioned to experimental parameters [Stein and Wysession, 2003]. Therefore, in order to determine the moment magnitude, one has to compute the seismic moment previously.

As it was explained in Sec. 3.2.2.1, the seismic moment is a quantity that has energy units and can be thus related to the energy released by an earthquake. The simplest physical definition of the scalar seismic moment is given by its proportionality to the rupture area and earthquake slip through the shear modulus (see Eq. (3.10) in Sec. 3.2.2.1). Although the procedure to compute the seismic moment can be complex to derive in some cases, a sketch is presented here in order to show how it can be determined in order to obtain the moment magnitude.

The seismic moment can be theoretically determined by analysing the fault geometry and the earthquake rupture process. However, this information cannot be directly measured and it must be obtained from the seismogram records. As a general approach to determine the seismic moment, one can consider that in the far field approximation (seismic waves detected at large distance from the seismic source, $r \to \infty$) [Aki and Richards, 2002; Lay and Wallace, 1995; Stein and Wysession, 2003; Stork and Verdon, 2014] the displacement u detected in a

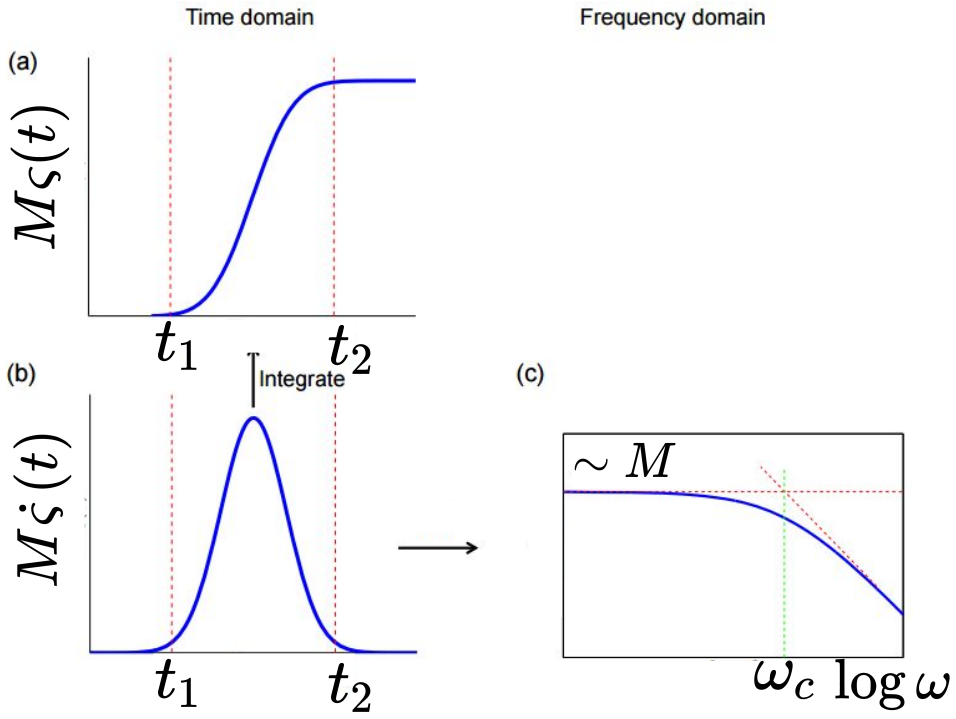

Fig. 4.10: (a) Model of the time evolution of the seismic moment $M(t) = M\varsigma(t)$ where M is the total seismic moment and $\varsigma(t)$ is the source time function. (b) Seismic moment rate function $\dot{M}(t) = M\dot{\varsigma}(t)$. As stated in the text, the seismic moment rate function is proportional to the far-field displacement pulse detected in a seismometer (see Eq. (4.6)). (c) Logarithm of the moment rate spectrum as a function of the logarithm of the frequency ω. For frequencies below the corner frequency ω_c, the spectral amplitude of the seismogram is prpoortional to the seismic moment M. Figure modified from Stork and Verdon [2014].

seismogram can be written as:

$$u(r,t) = \frac{1}{v_{eff}\mu}\frac{1}{r}\dot{M}(t)\,, \tag{4.6}$$

where $v_{eff} = \frac{4\pi}{\mathcal{R}_p}\frac{\rho}{\mu}v_{wave}^3$ is an effective velocity that depends on a geometric factor $\mathcal{R}_p$ that takes into account the radiation pattern, the medium shear modulus μ and density ρ and the velocity of the seismic wave v_{wave}. Note that the displacement of a seismic wave decays as the inverse of the distance r and is proportional to the seismic moment rate $\dot{M}(t)$, which is usually known as source pulse [Stein and Wysession, 2003; Stork and Verdon, 2014]. Note also that, in general, the seismic moment is a quantity with tensorial character and in this case is considered, for simplicity, as a scalar magnitude. In general the temporal evolution of the seismic moment can be expressed as [Madariaga, 1989; Schubert, 2015]:

$$M(t) = M\varsigma(t) \tag{4.7}$$

where M is the total seismic moment and $\varsigma(t)$ is the source time function that expresses its temporal dependence. This function can be understood as the shape of body-wave pulses due to faulting process. In other words, the source time function contains information about the time dependence of the rupture process. Different source models can be proposed in order to obtain a proper description of the source-time function. One of the most accepted is the Haskell source model [Haskell, 1966; Madariaga, 1978] which assumes a one-dimensional fault with unilateral rupture and a constant velocity of propagation of the seismic wave but other models are also widely accepted [Stork and Verdon, 2014]. A typical model of this pulse is shown in Fig. 4.10 (a) and (b). In any case, the source time function must fulfil $\lim_{t\to\infty}\varsigma(t)$ =constant. It must be mentioned that the characterization of the source has been also studied in acoustic emission [Yu and Clapp, 1987], which would correspond to the case of interest in labquakes.

By taking into account the time evolution of the seismic moment, Eq. (4.6) can be written:

$$u(r,t) = \frac{1}{v_{eff}\mu}\frac{M}{r}\dot{\varsigma}(t) \Longrightarrow \int_0^\infty u(r,t)dt \propto M\int_0^\infty \dot{\varsigma}(t)dt\,, \tag{4.8}$$

where in the last step integration over time is performed. Note that $\propto$ denotes proportionality It is obvious that the rupture does not take infinite time and the integral can be performed in a finite time interval (see Fig. 4.10(a)). At this point, one could find the total seismic moment M by simply converting the seismogram into a displacement (if it is recorded as a velocity) and integrating it in the time domain. However, in practice this procedure has a lot of limitations that avoid a good estimation. The main factors affecting this calculation are the difficulties in correctly determine t_1 and t_2, the poor signal-to-noise ratio and the possible artefacts that can appear in the conversion to a displacement seismogram, among others [Stork and Verdon, 2014]. However, Eq. (4.8) can be written in the frequency domain [Bender and Orszag, 1999]:

$$M\lim_{\omega\to 0}|\tilde{\dot{\varsigma}}(\omega)| \propto \lim_{\omega\to 0}|\tilde{u}(r,\omega)|\,, \tag{4.9}$$

where, in the limit of small frequencies ω, the moment magnitude is proportional to the spectral amplitude of the displacement through the value of the limit of the Fourier transform of the derivative of the source-time function. Independently on the source model used, the finite duration of the rupture process implies the existence of a corner frequency ω_c below which the spectral amplitude of the displacement is constant and proportional to the seismic moment M [Doornbos, 1982; Aki and Richards, 2002; Lay and Wallace, 1995; Stein and Wysession, 2003; Stork and Verdon, 2014] (see Fig. 4.10 (c)).

The existence of the corner frequency has implications related to the experimental set-up used for earthquake detection. One the one hand, in general it can be stated that the smaller the earthquake the shorter its duration. This implies that these small events will have an associated higher corner frequency. Therefore, sample rates need to be high enough in order to capture this corner frequency. On the other hand, larger earthquakes imply larger times and, consequently, smaller corner frequencies. In order to correctly record and associate a moment magnitude M to these bigger earthquakes, the frequency at which the experimental device works should be smaller than the corner frequency. If the experimental device cannot register frequencies smaller than ω_c, the corner frequency cannot be detected and thus the relationship with the seismic moment cannot be established for that earthquake. This is the reason why some magnitude scales saturate for large earthquakes. Surface wave, body wave or local magnitude scales are defined by means of the amplitude of the waves with a predefined frequency detected in the seismogram [Stein and Wysession, 2003; Yeats et al., 1997]. As already mentioned, modern broad-band seismometers try to cover a range of frequencies typically from few mHz to 5 Hz [Ahern, 2003] in order to detect a reasonable range of earthquake sizes.

Once the seismic moment is determined, the moment magnitude m scale [Hanks and Kanamori, 1979] can be found according to:

$$m = \frac{2}{3}\left(\log_{10}\left(\frac{M}{M_0}\right) - 9.1\right), \tag{4.10}$$

where M_0 is a reference seismic moment of 1 N·m and M is measured in Nm. Contrarily to other scales [Stein and Wysession, 2003; Yeats et al., 1997] such as the Richter scale, the local magnitude or the surface wave magnitude, the moment magnitude does not exhibit saturation effects. Provided that the moment magnitude scale depends on a physical magnitude rather than on experimental parameters (such as the characteristic frequency of the seismometer),it is one of the most accepted scales in seismology [Lay and Wallace, 1995; Stein and Wysession, 2003; Yeats et al., 1997].

4.2.3 Earthquake catalogs used in this thesis

Earthquakes catalogs covering different magnitude ranges have been chosen for this thesis with the aim of investigating the scale-invariance properties of the size distribution of earthquakes. A convenient combination of these datasets will give us a larger range of validity of the GR law with a unique exponent. Let us briefly describe the catalogs that have been used in this work (see also Fig. 4.11):

- Global Centroid Moment Tensor (CMT) catalog: It comprehends earthquakes worldwide from 1977 to 2017 [Dziewonski and Woodhouse, 1981; Ekström and Dziewonski, 2012]. This catalog reports the values of the moment magnitude as well as the seismic moment. Given that the seismic moment is provided with three significant digits, the resolution of the magnitude in the catalog is approximately $\Delta m \simeq 10^{-3}$.

- Yang–Hauksson–Shearer (YHS) A catalog: It records the earthquakes in Southern California with $m \geq 0$ in the period 1981-2010 [Yang et al., 2012]. This catalog does not report the seismic moment but a preferred magnitude that is converted into seismic moment according to Eq. (3.11). The resolution of the magnitude in the catalog is $\Delta m = 0.01$.

- Yang–Hauksson–Shearer (YHS) B catalog: It is a subset of YHS A that contains the earthquakes in the region of Los Angeles in the period 2000-2010 [Yang et al., 2012]. LA region is defined by the following four vertices in longitude and latitude: (119°W, 34°N),(118°W, 35°N) ,(116°W, 34°N), and (117°W, 33°N) (see Fig. 4.11). This region has been selected because it is among the best monitorized ones [Hutton et al., 2010; Schorlemmer and Woessner, 2008]. Furthermore, we selected this time period due to the better detection of smaller earthquakes than in previous years [Hutton et al., 2010]. This should reduce the completeness magnitude of the catalog [Schorlemmer and Woessner, 2008]. The resolution of this catalog is the same as YHS A.

It must be mentioned that all these catalogs report the focal mechanism of the earthquakes. This is important when computing the Coulomb stress in Chpt. 8.

Figure 4.12 shows a snapshot of some relevant earthquake observables provided in the CMT catalog for two different time windows. Note that in this catalog the seismic moment M is provided in units of dyn·cm (1dyn = 10^{-5}N) and Eq. (4.3) is rewritten as:

$$m = \frac{2}{3}\left(\log_{10}\left(\frac{M}{M_0}\right) - 16.1\right) , \tag{4.11}$$

where M_0 is a reference seismic moment of 1 dyn·cm.

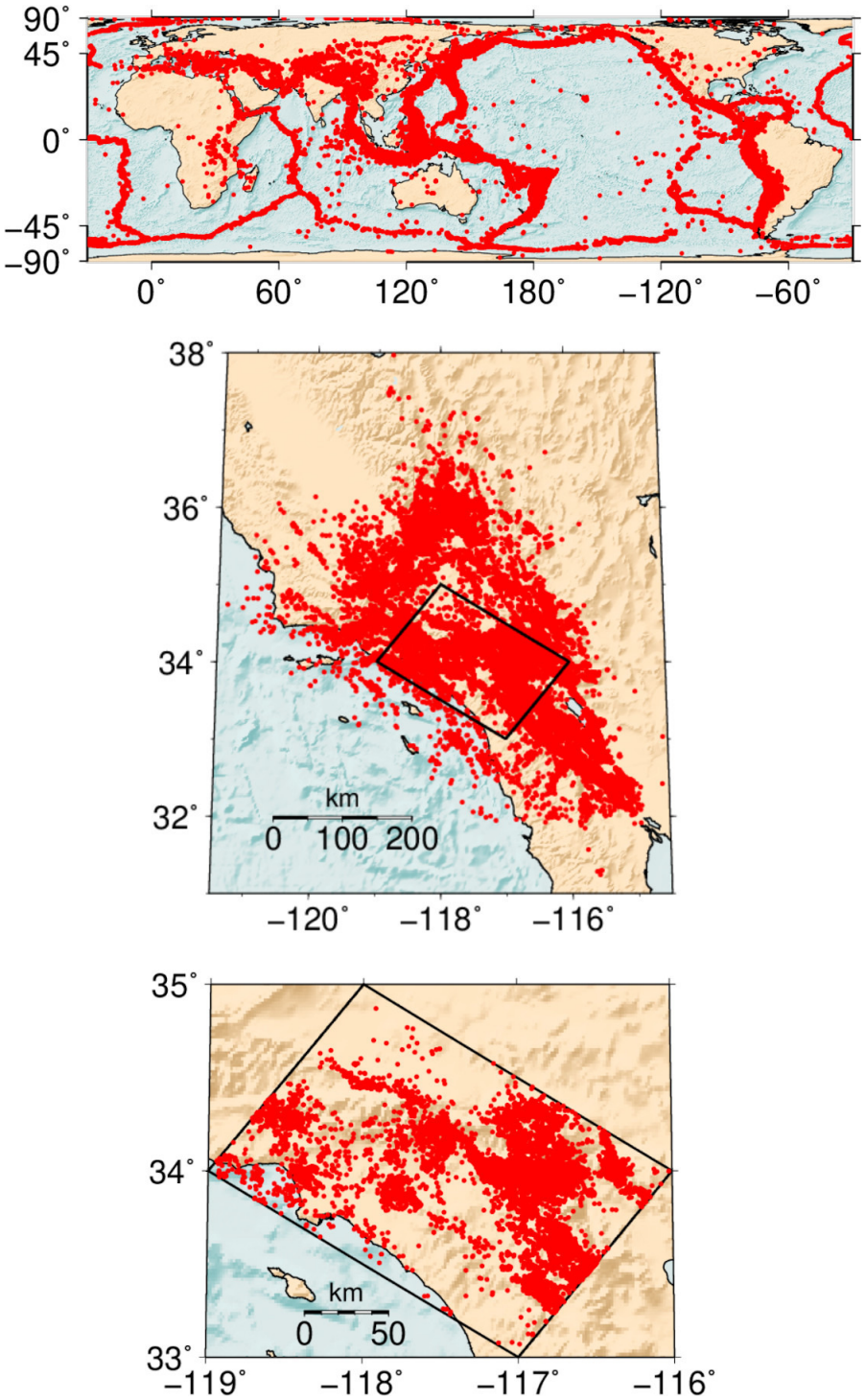

Fig. 4.11: Earthquake epicenters of the different catalogs used in this work. Horizontal and vertical axes correspond to longitude and latitude in degrees respectively. Top, full CMT catalog for the period 1977-2017 [Dziewonski and Woodhouse, 1981; Ekström and Dziewonski, 2012]. Middle, entire YHS A Catalog [Yang et al., 2012] of Southern California for the period 1981-2010. Bottom, region (YHS B) of the YHS catalog for Los Angeles area for the period 2000-2010. Maps in cylindrical equal-area projection (top) and sinusoidal projection (middle and bottom), produced with Generic Mapping Tools [Wessel et al., 2019].

```
################################################################################
CMT EARTHQUAKE CATALOG
################################################################################

Year  Month Day   Time        Latitude    Longitude   Depth       Seismic Moment Magnitud
.....................................................................................
.....................................................................................
.....................................................................................
1997  9     26    33:12.2     43.05       12.88       10          3.82E24     5.6547089086
1997  9     26    40:26.3     43.08       12.81       10          1.14E25     5.9712699009
1997  9     26    00:05.8     0.12        -16.92      10          1.92E24     5.4555341525
1997  9     26    16:00.9     -3.14       139.08      33          1.299E24    5.3424061007
1997  9     26    48:34.3     -5.39       128.99      253.8       1.34E25     6.0180698656
1997  9     26    44:27.2     10.06       126.01      33          4.51E23     5.0361176946
1997  9     27    22:51.0     -29.56      -176.91     33          1.018E24    5.271831852
.....................................................................................
.....................................................................................
.....................................................................................
.....................................................................................
2016  9     9     0:03:45     -31.38      -65.82      157.6       2.832E24    5.568062166
2016  9     9     11:53:15    36.39       141.13      32.1        2.342E24    5.5130579272
2016  9     9     17:17:54    22.13       -45.88      10          1.322E24    5.3474876368
2016  9     10    0:25:16     11.69       -86.84      36.4        3.519E23    4.9642795112
2016  9     10    0:46:18     30.22       82.68       35          1.835E23    4.7757573791
2016  9     10    2:49:54     -9.86       111.65      10          2.31E23     4.8424079866
.....................................................................................
.....................................................................................
.....................................................................................
```

Fig. 4.12: Snapshot of the CMT earthquake catalog with the different observables characterizing earthquakes for two different time periods. The first four columns correspond to the earthquake occurrence time, whereas the fifth, sixth and seventh columns correspond to spatial coordinates: latitude, longitude and depth. The next two columns represent the seismic moment measured in dyn·cm and the moment magnitude.

4.3 Earthquakes and labquakes: Experimental similarities and differences

Given that earthquake and labquake catalogs together with the definition of their respective observables have been already exposed, let us explain the main differences and similarities between these phenomena.

First of all, let us consider the snapshots of Figs. 4.6 and 4.12 in order to illustrate the most evident differences. The most remarkable difference is the lack of spatial degrees of freedom in labquake catalogs. Although it would be interesting to study the spatial distribution of labquakes, they are analysed indistinctly of their occurrence site. Apart from placing the statistical analysis in an easier scenario, the fact of finding similarities between the probability distributions of labquake and earthquake observables reinforces the postulates of the holistic approach.

The second difference is related to the scales involved in these phenomena [Rundle et al., 2003]:

- Characteristic time scales: The first main difference is the time scale at which these phenomena occur. The time scale of a short labquake is about $\sim 10^{-6}$ s whereas for a short earthquake it is typically about 1 s. The total duration of a compression experiment is about 2-3 hours ($\sim 10^4$s) whereas earthquake catalogs available nowadays record earthquakes about 50 years ($\sim 10^9$) (see for instance the catalogs in Sec. 4.2). Note that, in both cases, there is a difference of six orders of magnitude between earthquakes and labquakes temporal scales. The ratio between the typical time scale for a short labquake with respect the duration of the total experiment is about 10^{-10}. If the same ratio is assumed for the shortest earthquakes, a complete displacement-driven compression experiment would correspond to a time scale of $10^{10} - 10^{11}$s, which is about thousands of years.

- Characteristic energy scales: At this point, one could wonder whether a seismic moment (an energy) could be associated to labquakes (see Sec. 4.1.2). Recall that the reasonable energetic quantity to compare AE and earthquakes is the radiated energy. As stated in Sec. 3.2.2.1, the value of the seismic moment is about 10^4 times larger than the radiated energy. This implies that the AE energies of labquakes, which are of the order of 10^{-18}

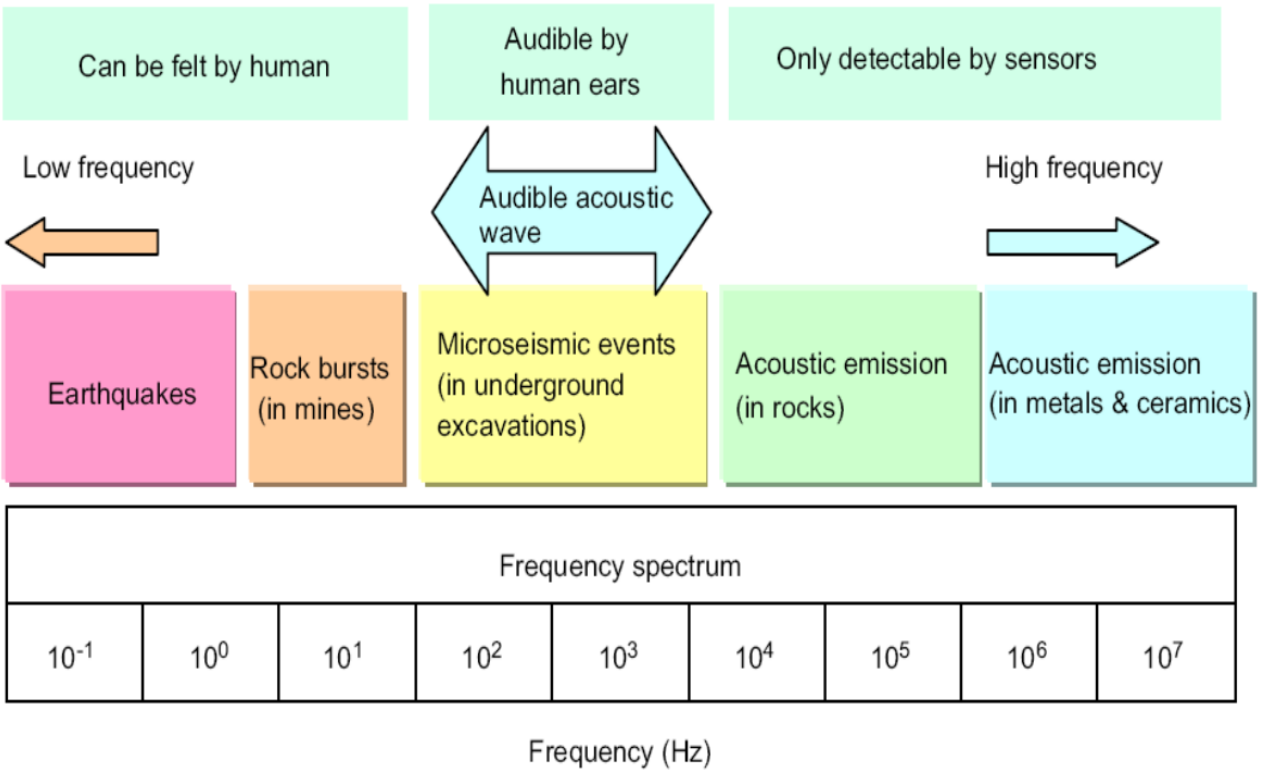

Fig. 4.13: Different parts of the frequency spectrum and their associated phenomena. Figure extracted from Park et al. [2013].

J, are translated into seismic moments of the order of 10^{-14} Nm. A big earthquake with moment magnitude $m = 9$ would have a seismic moment about $M \simeq 4 \times 10^{22}$ Nm corresponding to 4×10^{18} J. There is a difference around 30 orders of magnitude between earthquake and labquake radiated energy.

- Characteristic frequencies: The characteristic time scale is directly tied to the frequency range of the detection device. By taking into account that characteristic frequencies can be obtained from the inverse of the characteristic durations, earthquakes have a frequency range about $0.01 - 10$ Hz whereas labquakes cover a frequency range about $10^3 - 10^6$ Hz. Fig. 4.13 shows the different ranges in the frequency spectrum and their associated phenomena.

  The concept of corner frequency ω_c can also be extended for AE [Goodfellow and Young, 2014]. Figure 4.14 shows the dependence of the corner frequency with the seismic moment as well as for the moment magnitude for different scales of seismicity reported in the literature [Goodfellow and Young, 2014]. Most of the data points with negative moment magnitude correspond to AE experiments. Valuable information can be extracted from this figure. For instance, an earthquake with moment magnitude $m \simeq 5$ has an associated corner frequency about 1 Hz. This means that a seismometers working at higher frequencies would not be able to correctly measure this event. An event with seismic moment $M \sim 1$ Nm has an associated corner frequencies of the order of MHz, which is indeed the frequency range at which the AE equipment works. This means that the AE set-up could not correctly detect the observables of AE hits whose associated energy is larger than 1 Nm.

- Characteristic spatial scales: Ruptures associated to earthquakes can span from few meters to kilometres ($\sim 1 - 10^3$ m) whereas failure events corresponding to labquakes have a characteristic size of the order of 10^{-6} m. As it has been already mentioned, one important difference between earthquake and labquake detection is the difference in the size of the detection devices with respect the characteristic spatial scale of the phenomenon. For labquakes, the surface of the transducer is very large with respect to the characteristic labquake size. In this case, an important fraction of the labquake energy can be detected. On the contrary, the size of a seismometers is very small with respect the characteristic size of the rupture caused by an earthquake. In this case, a tiny fraction of an earthquake is detected.

Other differences and similarities can be related to the mechanics of both phenomena. As it was explained in Sec. 3.2.1, the mechanics associated to a rupture can be extremely complex and it cannot be limited to the compressive stress through which labquakes are induced. Furthermore, as it has been already mentioned, the mechanical properties of the samples where labquakes occur can be very different from those in the Earth crust.

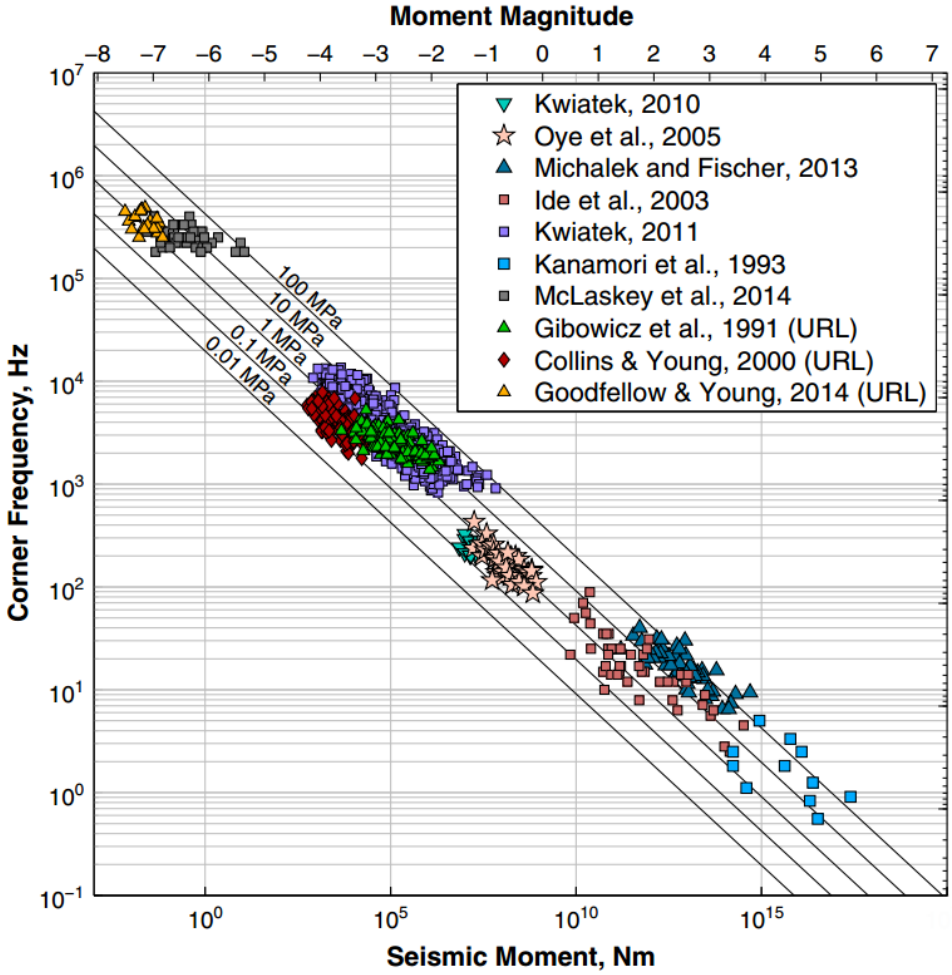

Fig. 4.14: Dependence of the corner frequency with the seismic moment (and its translation to moment magnitude) for AE data in microseismic events (Goodfellow & Young, 2014 (URL) in the legend) m in comparison with manmade and natural seismicity reported in the literature. Figure extracted from Goodfellow and Young [2014].

It is remarkable that, despite the difference in scales and in the underlying physical mechanisms in which earthquake and labquakes take place, similar statistical behaviours are found for their observables.

Chapter 5

Statistical Techniques

In this chapter, the statistical techniques, fundamental to build up and understand the procedures developed in the Results part of this PhD thesis, are exposed. The Maximum Likelihood Estimation (MLE) is presented as a method to estimate the parameters of the statistical model. Afterwards, this technique is focused on the particular case of interest of power-law distributions. Graphical estimation of the probability density function as well as the cumulative distribution function from an empirical dataset are exposed by paying special attention to the power-law case. Afterwards, a series of statistical tests that are recurrently used along this PhD thesis are explained. The likelihood-ratio test is presented as a statistical version of the Occam's razor criterion that will be useful in order to discern between a simpler and a more complex statistical model. The Akaike information criterion is explained as an alternative and complementary procedure to the likelihood-ratio test. The two-sample Kolmogorov-Smirnov (KS) test and the permutational test are useful tools in order to state whether two power-law distributions significantly differ between each other or not. Finally, the residual Coefficient Variation test is presented as a tool to know whether the tail of a power-law distribution deviates from the expected behaviour or not.

5.1 Maximum Likelihood Estimation (MLE)

5.1.1 Concepts and definitions

5.1.1.1 The log-likelihood function and main goal of the MLE

As stated in Sec. 2.3.1.3, the link between data and the statistical model is the likelihood function $\mathcal{L}(\boldsymbol{\theta})$. This likelihood function is characterized by a $n_{\mathcal{L}}$-dimensional array of parameters $\boldsymbol{\theta}$ which are inherently related with the proposed statistical model to describe data. The likelihood function is thus a real function that receives an array of dimensions $n_{\mathcal{L}}$ and returns a real value:

$$\mathcal{L} : \mathbb{R}^{n_{\mathcal{L}}} \to \mathbb{R} \, .$$

The general framework to understand the maximum likelihood estimation (MLE) is to consider a sample of size n whose values can be described in terms of a random variable $\mathcal{X}$ that follows a certain univariate probability distribution $F_{\mathcal{X}}(\boldsymbol{\theta})$. The values that the random variable takes are represented as $x_1, x_2, ..., x_i, ..., x_n$. By assuming independent and identically distributed (i.i.d) random variables, the likelihood function describes the joint probability density that all these values occur:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_{\mathcal{X}}\left(\mathcal{X} = x_i; \boldsymbol{\theta}\right) \, , \tag{5.1}$$

where $f_{\mathcal{X}}(x; \boldsymbol{\theta})$ is the probability density function (PDF) conveniently normalized in its domain. Unless otherwise stated, PDFs are assumed to be continuous. By taking the natural logarithm of the likelihood function, one can define the log-likelihood function as:

$$\log \mathcal{L}(\boldsymbol{\theta}) = \log \left(\prod_{i=1}^{n} f_{\mathcal{X}}\left(\mathcal{X} = x_i; \boldsymbol{\theta}\right) \right) = \sum_{i=1}^{n} \log f_{\mathcal{X}}\left(\mathcal{X} = x_i; \boldsymbol{\theta}\right) \, , \tag{5.2}$$

which is useful when developing calculations.

The main purpose of MLE is to find the array $\hat{\boldsymbol{\theta}}$ whose values of the estimators maximize the likelihood function for a given sample. Therefore, MLE is reduced to a problem of maximization of a function of $n_{\mathcal{L}}$ parameters.

At this point it is important to mention that, as its name indicates, MLE provides an estimation of the set of parameters that maximize the likelihood that the dataset has been sampled from a population characterized by a certain statistical model. The array of parameters (true parameters) is denoted as $\tilde{\boldsymbol{\theta}}$. As it will stated in the following sections, the MLE does not ensure whether the population from which the dataset has been sampled from is suitably characterized by that statistical model. In the next section, the consistency of the MLE is illustrated in the sense that the estimators $\hat{\boldsymbol{\theta}}$ are arbitrary close to the true value of the parameters $\tilde{\boldsymbol{\theta}}$ for large enough samples.

5.1.1.2 Regularity conditions

The results and properties of the maximum likelihood estimators explained in this section are true as long as the following regularity conditions are fulfilled ([Pawitan, 2013] p.241 and [Wood, 2015] p.85):

1. The PDF characterized by the array of parameters $\boldsymbol{\theta}$ is different for different values of the components of $\boldsymbol{\theta}$.

2. The array of parameters $\boldsymbol{\theta}$ is contained inside the space of parameters. If $\boldsymbol{\theta}$ is placed at a boundary of the space of parameters, regularity is not ensured.

3. The first three derivatives with respect the parameters of the log-likelihood function in a vicinity of $\hat{\boldsymbol{\theta}}$ exist and are finite.

4. The matrix $\boldsymbol{I}$ of the second derivatives of the log-likelihood function evaluated at $\hat{\boldsymbol{\theta}}$ is negative semi-definite. This condition ensures that the eigenvalues are negative and that $\mathcal{L}\left(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\right)$ is indeed a maximum.

5. The expansion of the log-likelihood function up to second order is sufficient in order to fulfil consistency.

5.1.1.3 Score function and score equations

Given the likelihood function $\mathcal{L}(\boldsymbol{\theta})$, one can define a $n_{\mathcal{L}}$-dimensional Score function $\boldsymbol{\mathcal{S}}$:

$$\boldsymbol{\mathcal{S}} : \mathbb{R}^{n_{\mathcal{L}}} \to \mathbb{R}^{n_{\mathcal{L}}} .$$

The j-th component of this array is defined by the partial derivative of the log-likelihood with respect the j-th parameter:

$$S_j\left(\boldsymbol{\theta}\right) = \sum_{i=1}^{n} \frac{1}{f_{\mathcal{X}}\left(x_i, \boldsymbol{\theta}\right)} \frac{\partial f_{\mathcal{X}}\left(x_i, \boldsymbol{\theta}\right)}{\partial \theta_j} . \tag{5.3}$$

It can be easily shown ([Pawitan, 2013] p.216 [Wood, 2015] p.79) that the expected value of the j-th component of the Score function array evaluated at the real value of the set of parameters $\tilde{\boldsymbol{\theta}}$ is identically zero:

$$\mathbb{E}S_j\left(\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}\right) = 0 . \tag{5.4}$$

This result will be important when discussing about the consistency and the distribution of the ML estimators.

The log-likelihood function attains a maximum when the Score array leads the null vector $\mathbf{0}$:

$$\boldsymbol{S}\left(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\right) = \frac{\partial \log \mathcal{L}\left(\hat{\boldsymbol{\theta}}\right)}{\partial \boldsymbol{\theta}} = \mathbf{0} .$$

This set of equations is known as score equations. In some particular cases, the solution of score equations can be solved analytically and a closed expression for the values of the estimators can be easily found. For other statistical models, it might not be possible to find an analytical expression for the components of $\hat{\boldsymbol{\theta}}$. In these cases, one has to find the estimators for which the maximum is attained by means of numerical techniques.

5.1.1.4 The matrix of the second derivatives

Once the values of the estimators maximizing the likelihood are found, one could be tempted to question whether this array of estimators corresponds to a maximum of the log-likelihood function or not. One can define a matrix of second derivatives $\boldsymbol{I}$ evaluated at the array of estimators $\hat{\boldsymbol{\theta}}$ whose components can be written as:

$$I_{jk}\left(\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\right)=-\frac{\partial^2 \log \mathcal{L}(\hat{\boldsymbol{\theta}})}{\partial\theta_j\partial\theta_k}=-\sum_{i=1}^{n}\frac{1}{f_{\mathcal{X}}\left(x_i,\hat{\boldsymbol{\theta}}\right)}\frac{\partial^2 f_{\mathcal{X}}\left(x_i,\hat{\boldsymbol{\theta}}\right)}{\partial\theta_j\partial\theta_k}\,, \tag{5.5}$$

with indices $j=1,...,n_{\mathcal{L}}$ and $k=1,...,n_{\mathcal{L}}$. This matrix is usually named as the Observed Fisher Information [Pawitan, 2013; Wood, 2015]. If the regularity conditions of the log-likelihood function are fulfilled, this matrix is symmetric and thus it exists a basis where it is diagonal. Note that the matrix is defined with a negative sign and the sign of the components of this matrix is entirely determined by the sign of the second derivatives of the PDF. For the case in which PDFs are from the exponential family, such as the normal, the exponential or the gamma PDF, the sign of the second derivative is always negative and it can be ensured that the log-likelihood function evaluated at $\hat{\boldsymbol{\theta}}$ corresponds to a maximum. Given that the power-law PDF is from the exponential family, the same reasoning can be applied.

By following a similar reasoning to the one done with the Score statistic, the expected information matrix $\boldsymbol{\mathcal{I}}$, whose components correspond to the expected values of the components of the matrix of the second derivatives, can be defined as:

$$\mathcal{I}_{jk}=\mathbb{E}I_{jk}\left(\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}\right)\,. \tag{5.6}$$

It is important to remark that the expected information is not the same as the real value of I_{jk} extracted from the dataset and computed at $\hat{\boldsymbol{\theta}}$. The expected information has to be understood as an average quantity obtained along many samples from the original PDF with the real parameters $\tilde{\boldsymbol{\theta}}$. Another important result that can be easily shown is that the elements of the covariance matrix conformed by the different components of the Score array correspond to the elements of the expected information matrix [Pawitan, 2013] p.217 and [Wood, 2015] p.80 :

$$cov\left[S_j\left(\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}\right),S_k\left(\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}\right)\right]=\mathcal{I}_{jk}\left(\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}\right)\,. \tag{5.7}$$

5.1.2 Consistency and Distribution of the ML Estimator

In order to present the consistency and how the ML estimators are distributed, it is useful to previously consider some important results. Let us split the different contributions of the Score function array by considering that for the i-th point in the sample ($i = 1, ..., n$) the $n_{\mathcal{L}}$-dimensional array can be defined as:

$$\boldsymbol{l}'_{(i)} \equiv \frac{1}{f_{\mathcal{X}}\left(x_i, \boldsymbol{\theta}\right)} \frac{\partial f_{\mathcal{X}}\left(x_i, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}} \, .$$

Note that

$$\sum_{i=1}^{n} \boldsymbol{l}'_{(i)} = \frac{\partial \log \mathcal{L}\left(\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}} = \boldsymbol{S}\left(\boldsymbol{\theta}\right) \, .$$

By using this notation, one can consider a sample $\boldsymbol{l}'_{(1)}, \boldsymbol{l}'_{(2)}, ..., \boldsymbol{l}'_{(n)}$ where, according to Eqs. (5.4) and (5.7), fulfils:

$$\mathbb{E} l'_{(i)j}\left(\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}\right) = 0 \, ,$$

$$cov\left[l'_{(i)j}\left(\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}\right), l'_{(i)k}\left(\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}\right)\right] \equiv \mathcal{I}_{(i)jk} \, .$$

These results, together with the Central Limit Theorem imply that:

$$\sqrt{n}\left(l'_{(i)j} - 0\right) \rightarrow N\left(0, \mathcal{I}_{(i)jk}\right) \Longrightarrow \frac{S_j\left(\boldsymbol{\theta}\right)}{\sqrt{n}} \rightarrow N\left(0, \sqrt{\mathcal{I}_{(i)jk}}\right) \, , \tag{5.8}$$

where the arrow $\rightarrow$ implies convergence in distribution to the normal distribution. Therefore, for large samples, the j-th component of the Score function array is normally distributed with zero mean and standard deviation $\mathcal{I}_{(i)jk}$.

Regularity condition number 5 in Sec. 5.1.1.2 is useful to introduce the concept of consistency of the MLE as well as to investigate the behaviour of the ML estimator in the asymptotic limit. Let us suppose a very large dataset ($n \rightarrow \infty$) sampled from a distribution characterized by the array of real parameters $\tilde{\boldsymbol{\theta}}$. Let us expand the Score function around the vector of the real parameters:

$$\boldsymbol{S}\left(\boldsymbol{\theta}\right) \simeq \boldsymbol{S}\left(\tilde{\boldsymbol{\theta}}\right) + \boldsymbol{\mathcal{I}}\left(\tilde{\boldsymbol{\theta}}\right)\left(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\right) + h.o.t. \, ,$$

where $h.o.t$ is the abbreviation of higher order terms. This function can be evaluated at $\hat{\boldsymbol{\theta}}$, so that $\boldsymbol{S}\left(\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}\right) = 0$, and can also be multiplied at both sides by $\sqrt{n}$:

$$-n\boldsymbol{\mathcal{I}}^{-1}\left(\tilde{\boldsymbol{\theta}}\right)\frac{\boldsymbol{S}\left(\tilde{\boldsymbol{\theta}}\right)}{\sqrt{n}} = \sqrt{n}\left(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}\right) \, .$$

By taking into account Eq. (5.8) and the Slutsky Theorems which ensure the convergence in distribution and probability (see [Pawitan, 2013] p.234) it can be said that, for the j-th component:

$$\hat{\theta}_j \to N\left(\tilde{\theta}_j, \sqrt{\mathcal{I}^{-1}_{(i)jk}}\right) . \tag{5.9}$$

Note that $\mathcal{I}^{-1}_{(i)jk} \neq \left(\mathcal{I}_{(i)jk}\right)^{-1}$ but the jk-th component of the inverse matrix $\boldsymbol{\mathcal{I}}^{-1}$ (which is not the same unless $\boldsymbol{\mathcal{I}}$ is diagonal). The distribution of the ML estimators are normally distributed with mean $\tilde{\boldsymbol{\theta}}$ and standard deviation $\mathcal{I}^{-1}_{(i)jk}$. A large value of $\mathcal{I}_{(i)jk}$ would indicate that the normal distribution shows a very small dispersion around the mean value and thus the ML estimators $\hat{\boldsymbol{\theta}}$ are very close to the real ones $\tilde{\boldsymbol{\theta}}$. This result confirms the consistency of the ML estimators under the regularity conditions and is giving information about how these estimators are distributed. Therefore, one way to estimate the dispersion or the error of the MLEs is through the matrix of second derivatives.

It is important to remark that this is just a first step in the inference process. In order to know whether the proposed statistical model is reasonable or not to describe data, one needs more information which is not given by the MLE. Given a statistical model, MLE provides the array of parameters which maximizes the log-likelihood for a certain sample. For instance, one could apply MLE to a sample under the assumption that data are normally distributed and obtain the estimates which maximize the log-likelihood function. However, this is not giving any information about how reasonable is to consider that data are sampled from a normal distribution or from any other distribution. One could just compare two different statistical models by means of a likelihood ratio test (see Sec. 5.6) and check which of them is more likely to describe data. Nevertheless, not even in this case one is able to determine whether these statistical models are reasonable to describe data or not. If one just wishes to check whether it is reasonable or not to state that data are sampled from a certain probability distribution with the estimates obtained via MLE, one might use the goodness-of-fit test explained in Sec. 5.4.

In order to fix ideas about a general problem that can be solved by using MLE, one can consider a sample of size n of the random variable $\mathcal{X}$ to which one wants to link with a particular statistical model: the normal distribution. This distribution is characterized by a two-dimensional array of parameters $\boldsymbol{\theta} = (\mu, \sigma)$ and its PDF is $f_{\mathcal{X}}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. The log-likelihood and the equations defining the Score array in this case are:

$$\log \mathcal{L}\left(\mu, \sigma\right) = -\frac{n}{2} \log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2 , \tag{5.10}$$

$$S_\mu = \frac{\partial \log \mathcal{L}\left(\hat{\mu}, \hat{\sigma}\right)}{\partial \mu} = \frac{1}{\hat{\sigma}^2} \left(\sum_{i=1}^{n} x_i - n\hat{\mu} \right) = 0 \Longrightarrow \boxed{\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i} , \tag{5.11}$$

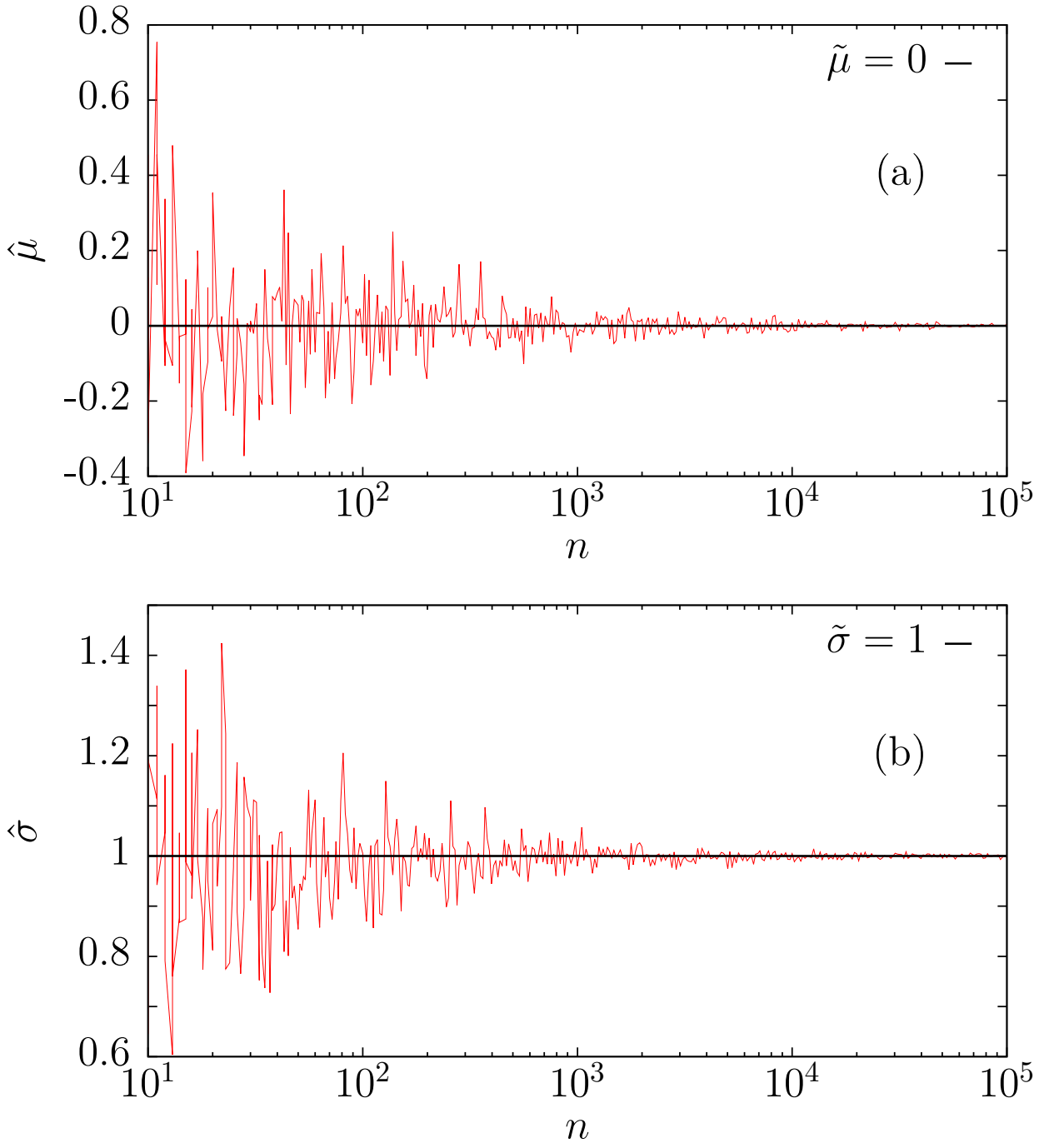

Fig. 5.1: Consistency of ML estimators for the mean and the standard deviation respectively $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma})$ for increasing sample sizes obtained from simulated datasets sampled from a $N(0, 1)$.

$$S_\sigma = \frac{\partial \log \mathcal{L}\left(x; \hat{\mu}, \hat{\sigma}\right)}{\partial \sigma} = -\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3} \sum_{i=1}^{n} \left(x_i - \hat{\mu}\right)^2 = 0\,, \tag{5.12}$$

$$\Longrightarrow \boxed{\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(x_i - \hat{\mu}\right)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n} \sum_{i=1}^{n} x_i\right)^2}\,.} \tag{5.13}$$

The ML estimators $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma})$ for the normal distribution correspond to the sample mean and the unadjusted sample standard deviation. By following the procedure explained before, one might be able to construct the matrix of the second derivatives as:

$$\boldsymbol{I}\left(\hat{\boldsymbol{\theta}}\right) = \begin{pmatrix} \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & 2\frac{n}{\hat{\sigma}^2} \end{pmatrix}\,, \tag{5.14}$$

whose eigenvalues are positive, thus indicating that the log-likelihood exhibits a maximum at $\hat{\boldsymbol{\theta}}$ (note that $\boldsymbol{I}$ was defined with a negative sign). In order to illustrate the consistency of the MLE estimators for the normal distribution, one can simulate n numbers sampled from a known $N\left(0, 1\right)$ distribution and check the value of the MLEs for increasing sample sizes. Figure 5.1 shows how the values of the estimates $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma})$ obtained through the Eqs. (5.11) and (5.13) converge to the real parameters $\tilde{\boldsymbol{\theta}} = (\tilde{\mu}, \tilde{\sigma})$ as the sample size increases.

5.2 Fitting power-law distributions to empirical data via MLE

The PDF of main interest in this PhD thesis is the power-law distribution. Two different power-law behaviours are going to be considered depending on whether there is an upper truncation (truncated power-law) or not (untruncated power-law). There are PDFs, such as the Gamma or the truncated Gamma distributions [Serra and Corral, 2017], that exhibit power-law regimes together with other behaviours that may appear at a certain characteristic scale. The main goal is to study pure power-law regimes searching for a statistical model describing the lack of characteristic scales in physical systems. Truncations are considered in order to either normalize or to exclude deviations from the power-law behaviour. Such distortions need to be discussed in detail for each particular dataset. One of the goals in this work is to defend the idea that these deviations come from irregularities in the data acquisition processes or finite-size effects and are not intrinsically related to the physical phenomenon itself.

5.2.1 Untruncated power-law distributions

The untruncated power-law is the simplest form of power-law PDF in which the random variable $\mathcal{X}$ has support $x \in [x_{min}, \infty)$. The value x_{min} ($x_{min} > 0$) corresponds to the lower cut-off of the power-law and, in this case, this PDF is said to be untruncated because it is not upper-bounded. The untruncated power-law is written as:

$$f_{\mathcal{X}}(x)dx = Cx^{-\gamma}dx\,, \tag{5.15}$$

where γ is the power-law exponent ($\gamma > 1$) and C is a normalizing constant that can be easily found by imposing:

$$\int_{x_{min}}^{\infty} f_{\mathcal{X}}(x)dx = 1 \Longrightarrow C = \frac{\gamma - 1}{x_{min}^{1-\gamma}}\,.$$

Note that, for $\gamma < 1$, the integral diverges and normalization cannot be achieved. Therefore, the correct expressions of the untruncated power-law PDF and Cumulative Distribution Function (which is denoted indistinctly as F or CDF) are:

$$f_{\mathcal{X}}\left(x; x_{min}, \gamma\right) dx = \frac{\gamma - 1}{x_{min}^{1-\gamma}} x^{-\gamma}dx\,, \tag{5.16}$$

$$F_{\mathcal{X}}\left(x; x_{min}, \gamma\right) = \int_{x_{min}}^{x} f_{\mathcal{X}}(x')dx' = 1 - \left(\frac{x}{x_{min}}\right)^{1-\gamma}\,. \tag{5.17}$$

Note that the moments of this distribution diverge if $\gamma < 2$.

Let us suppose a sample of size n from which one wishes to estimate the parameters of an untruncated power-law distribution. As explained in the previous section, in order to perform MLE in this context, the first step consists in writing the log-likelihood function:

$$\log \mathcal{L} = \log \prod_{i=1}^{n} f_{\mathcal{X}}\left(x_i; x_{min}, \gamma\right) = \sum_{i=1}^{n} \log f_{\mathcal{X}}\left(x_i; x_{min}, \gamma\right)\,, \tag{5.18}$$

where n is the number of data between x_{min} and ∞, the set of parameters characterizing the distribution are $\boldsymbol{\theta} = (x_{min}, \gamma)$ and $n_{\mathcal{L}} = 2$. If Eq. (5.16) is introduced into this equation one obtains:

$$\log \mathcal{L} = n \log \left(\frac{\gamma - 1}{x_{min}^{1-\gamma}}\right) - \gamma \sum_{i=1}^{n} \log x_i\,. \tag{5.19}$$

At this point, it is important to remark that x_{min} is considered here as a parameter that will not be estimated by means of MLE. Instead of that, x_{min} is fixed and the exponent γ is the only estimator computed via MLE. This is the reason why, although the PDF is written as $f_{\mathcal{X}}\left(x_i; x_{min}, \gamma\right)$ as if the array of parameters had two components, it actually has just one free parameter to be computed via MLE. Calculations are thus simplified since the array of

parameters consists solely in a scalar. Hence, the Score function array just consists in a single component and the unique Score equation in this case can be read:

$$\mathcal{S}\left(x_{min}, \gamma = \hat{\gamma}\right) = \left(\frac{\partial \log \mathcal{L}}{\partial \gamma}\right)_{\gamma=\hat{\gamma}} = \frac{n}{\hat{\gamma} - 1} + n \log x_{min} - \sum_{i=1}^{n} \log x_i = 0\,, \tag{5.20}$$

where the estimation of the exponent $\hat{\gamma}$ can be easily found analytically by:

$$\hat{\gamma} = 1 + \frac{n}{\sum_{i=1}^{n} \log\left(\frac{x_i}{x_{min}}\right)}\,. \tag{5.21}$$

Hence, for a given value of x_{min}, the set of parameters that maximize the log-likelihood-function are:

$$\hat{\boldsymbol{\theta}} = \left(1 + \frac{n}{\sum_{i=1}^{n} \log\left(\frac{x_i}{x_{min}}\right)}\right)\,.$$

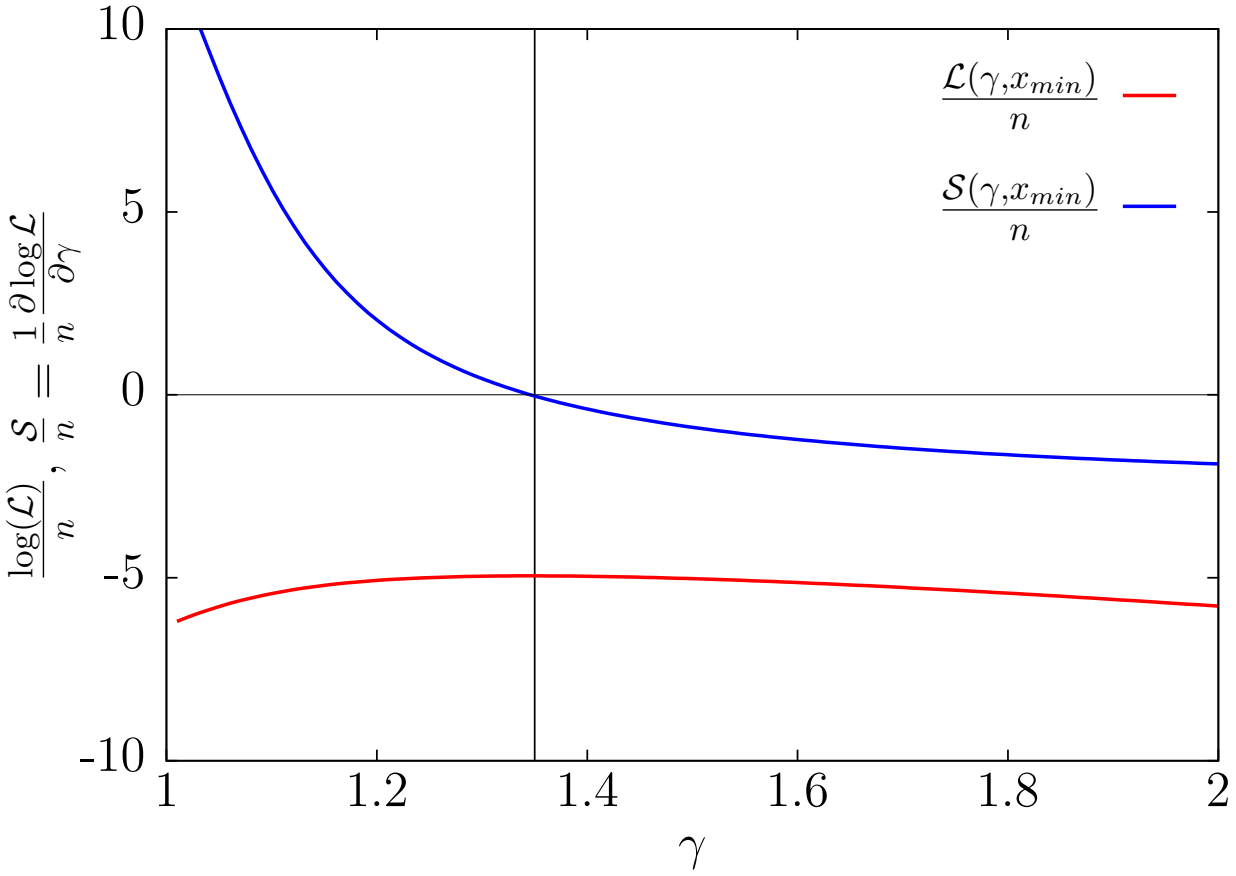

Fig. 5.2: Log-likelihood $\log \mathcal{L}(\gamma)$ (in red) and Score function $\mathcal{S}(\gamma)$ (in blue) per datum for synthetic untruncated power-law data with exponent $\gamma = 1.35$, $n = 10^4$ and $x_{min} = 1$.

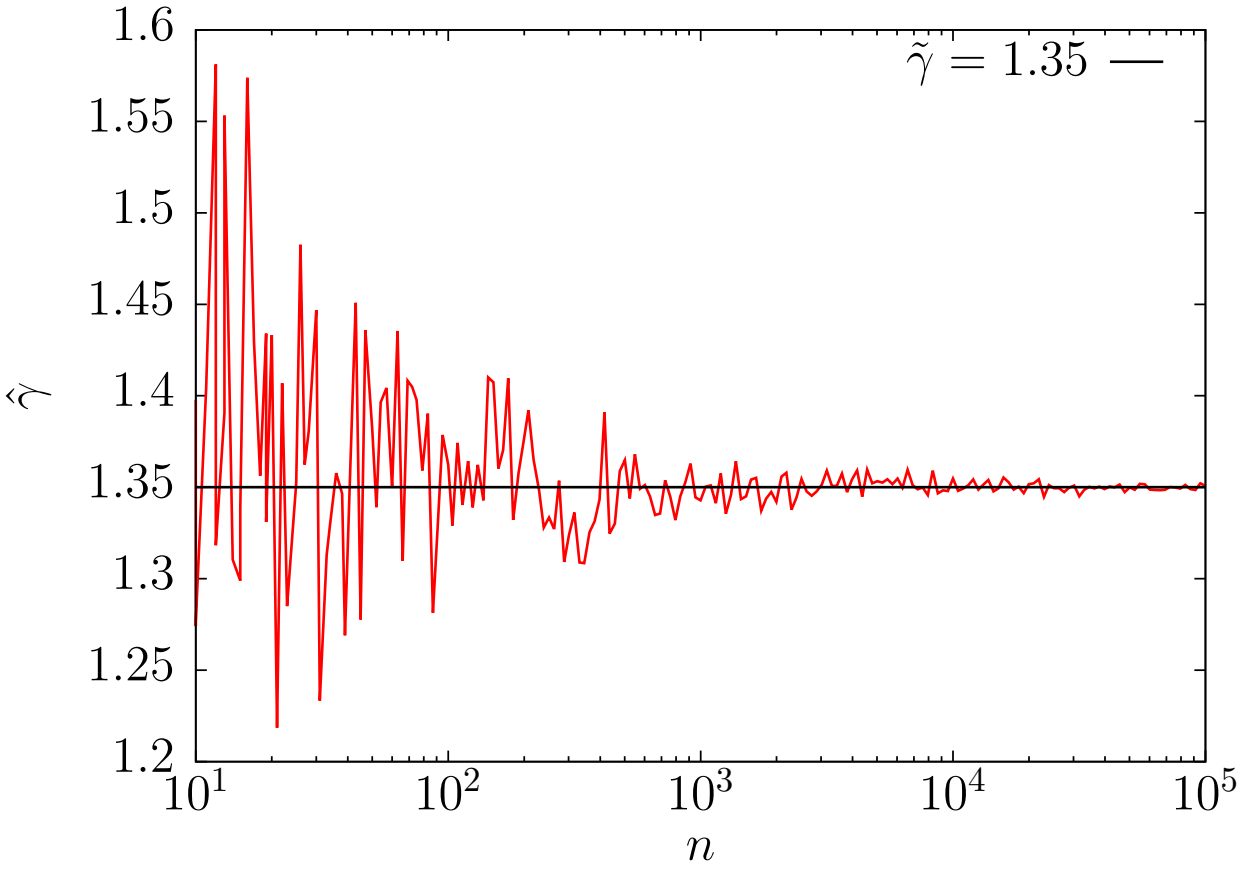

Fig. 5.3: Consistency of ML estimator $\hat{\gamma}$ for the power-law exponent for synthetic untruncated power-law data with exponent $\tilde{\gamma} = 1.35$, $x_{min} = 1$ and increasing sample size n.

Figure 5.3 shows the consistency of the MLE of the power-law exponent for increasing sample sizes. The log-likelihood and the Score function per datum for a simulated power-law sample with $\tilde{x}_{min} = 1$, exponent $\tilde{\gamma} = 1.35$ and sample size $n = 10^4$ are shown in Fig. 5.2. In this case, the matrix of the second derivatives (defined in Sec. 5.1 with a negative sign) is composed by a unique element:

$$I = \frac{n}{(\hat{\gamma} - 1)^2}\,, \tag{5.22}$$

which is positive, confirming thus the fact that the log-likelihood evaluated at $\hat{\gamma}$ is a maximum. As it was explained in Sec. 5.1.2, this ML estimate is normally distributed about the true value of the exponent $\tilde{\gamma}$ with a standard deviation $\sqrt{I^{-1}}$ for large enough samples. This standard deviation is giving an estimation of the error when estimating the exponent. Consequently, it can be written that the standard deviation or error of the estimated exponent is:

$$\sigma_\gamma = \frac{\hat{\gamma} - 1}{\sqrt{n}}\,. \tag{5.23}$$

5.2.2 Truncated power-law distributions

The untruncated power-law PDF assumes that the power-law behaviour is kept from the cut-off value x_{min} to infinity. However, there is a wide casuistry that might cause deviations from the power-law behaviour for large values: finite-size effects, the presence of an exponential tail, a change of the power-law exponent, etc. In order to exclude these deviations from a pure power-law behaviour, one can consider that the power-law is restricted to a limited range from a lower cut-off x_{min} to an upper cut-off x_{max}. This procedure is used along this PhD thesis in order to discriminate the power-law regimes in the global distribution. For instance, if the tail of the distribution can be modelled by an exponential decay with a characteristic scale, one can set the x_{max} in order to exclude this behaviour and approximate the rest of the distribution by a truncated power-law.

The truncated power-law PDF $f_{\mathcal{X}}\left(x;\gamma,x_{min},x_{max}\right)$ is thus characterized by $\boldsymbol{\theta}=(x_{min},x_{max},\gamma)$: two values for the lower x_{min} and upper x_{max} truncations and the exponent γ ($\gamma\neq 1$). The the random variable $\mathcal{X}$ has support $x\in[x_{min},x_{max}]$. Therefore the normalizing constant is:

$$\int_{x_{min}}^{x_{max}} f_{\mathcal{X}}(x)dx=1\Longrightarrow C=\frac{1-\gamma}{x_{max}^{1-\gamma}-x_{min}^{1-\gamma}}\,. \tag{5.24}$$

By introducing this normalizing constant into Eq. (5.15) one obtains:

$$f_{\mathcal{X}}\left(x;x_{min},x_{max},\gamma\right)dx=\frac{1-\gamma}{x_{max}^{1-\gamma}-x_{min}^{1-\gamma}}x^{-\gamma}dx\,, \tag{5.25}$$

$$F_{\mathcal{X}}(x;x_{min},x_{max},\gamma)=\int_{x_{min}}^{x} f_{\mathcal{X}}\left(x';x_{min},x_{max},\gamma\right)dx'=\frac{x^{1-\gamma}-x_{min}^{1-\gamma}}{x_{max}^{1-\gamma}-x_{min}^{1-\gamma}}\,. \tag{5.26}$$

Note that, contrarily to untruncated power-law distributions, the first moment is well defined:

$$\langle x\rangle=\int_{x_{min}}^{x_{max}} x f_{\mathcal{X}}(x)dx=\begin{cases}\frac{\gamma-1}{\gamma-2}\frac{x_{min}^{2-\gamma}-x_{max}^{2-\gamma}}{x_{min}^{1-\gamma}-x_{max}^{1-\gamma}} & \text{for } \gamma\neq 1 \text{ and } \gamma\neq 2\\ \frac{1-\gamma}{x_{min}^{1-\gamma}-x_{max}^{1-\gamma}}\log\left(\frac{x_{max}}{x_{min}}\right) & \text{for } \gamma=2\end{cases},$$

In order to obtain by means of MLE the parameters which link a truncated power-law PDF to an empirical dataset of n values $x_1,...,x_n$, one has to write the log-likelihood function as:

$$\log\mathcal{L}=\log\prod_{i=1}^{n} f_{\mathcal{X}}\left(x_i;\gamma,x_{min},x_{max}\right)=\sum_{i=1}^{n}\log f_{\mathcal{X}}\left(x_i;\gamma,x_{min},x_{max}\right)\,. \tag{5.27}$$

At this point, as it happens for the MLE of untruncated power-law PDFs, it must be mentioned that the cut-offs x_{min} and x_{max} are not estimated via MLE but are fixed instead. The only parameter that is estimated by means of MLE is the power-law exponent γ. Again, the log-likelihood, the Score function and the matrix of second derivatives are simply scalar functions:

$$\log\mathcal{L}=n\left[\log\left(|1-\gamma|\right)-\log\left(|x_{max}^{1-\gamma}-x_{min}^{1-\gamma}|\right)\right]-\gamma\sum_{i=1}^{n}\log x_i\,. \tag{5.28}$$

The Score equation in this case can be read:

$$\begin{aligned}\mathcal{S}\left(x_{min}, x_{max}, \gamma = \hat{\gamma}\right) &= \left(\frac{\partial \log \mathcal{L}}{\partial \gamma}\right)_{\gamma=\hat{\gamma}} \\ &= n\left[\frac{1}{\hat{\gamma}-1} + \frac{x_{max}^{1-\hat{\gamma}} \log x_{max} - x_{min}^{1-\hat{\gamma}} \log x_{min}}{x_{max}^{1-\hat{\gamma}} - x_{min}^{1-\hat{\gamma}}}\right] - \sum_{i=1}^{n} \log x_i \\ &= 0\,,\end{aligned} \tag{5.29}$$

where the exponent $\hat{\gamma}$ can be found by a numerical method to determine the root of this equation. Two different numerical methods are going to be used along this thesis: the function *"optimize"*, which is already implemented in R programming language [R Documentation, 2017; Brent, 2002] and the false-position method [W. H. Press and Flannery., 1992]. The matrix of the second derivatives is a scalar function that can be separated into two different terms [Aban et al., 2012]:

$$I = \left(\frac{\partial \mathcal{S}}{\partial \gamma}\right)_{\hat{\gamma}} = \frac{n}{\left(\hat{\gamma}-1\right)^2} - n\frac{\left(\frac{x_{min}}{x_{max}}\right)^{\hat{\gamma}-1} \log^2\left(\frac{x_{min}}{x_{max}}\right)}{\left(1 - \left(\frac{x_{min}}{x_{max}}\right)^{\hat{\gamma}-1}\right)^2}\,. \tag{5.30}$$

This separation is done in order to evidence that the first term is exactly the same that appeared for the untruncated case in Eq. (5.22), and the second term is the one keeping information about the upper cut-off. Note that this second term vanishes for the untruncated case, which occurs at the limit $\frac{x_{min}}{x_{max}} \to 0$. Taking into account that the standard deviation of the estimated exponent $\sigma_\gamma = \sqrt{I^{-1}}$, it can be written:

$$\sigma_\gamma = \frac{1}{\sqrt{n}}\left(\frac{1}{\left(\hat{\gamma}-1\right)^2} - \frac{\left(\frac{x_{min}}{x_{max}}\right)^{\hat{\gamma}-1} \log^2\left(\frac{x_{min}}{x_{max}}\right)}{\left(1 - \left(\frac{x_{min}}{x_{max}}\right)^{\hat{\gamma}-1}\right)^2}\right)^{-\frac{1}{2}}\,. \tag{5.31}$$

Note that Eq. (5.23) is recovered in the limit $\frac{x_{min}}{x_{max}} \to 0$. It must be mentioned that the standard deviation of the exponent can also be estimated by means of Monte Carlo simulations (see Sec. 5.4.1).

5.3 Graphical representation of the estimated probability density function and the cumulative distribution

Along this thesis, many samples of different nature are analysed with the aim of fitting the parameters $\boldsymbol{\theta}$ that characterize the statistical model which they have been generated from and also to check whether it is statistically reasonable to state that a certain dataset has been sampled from that statistical model. Statistical techniques such MLE shown in Sec. 5.1 or the hypothesis tests that will be shown in the forthcoming sections are very important tools to show quantitative evidences of these goals. However, plotting data in a convenient way can offer an intuitive point of view that can be helpful in order to illustrate how a sample is distributed in a visual way.

The procedure through which one can graphically estimate a PDF from an empirical dataset as well as the CDF, is exposed in this section. The graphically estimated PDF and CDF are denoted as $f_{\mathcal{X},(e)}(x)$ and $F_{\mathcal{X},(e)}(x)$ respectively. Henceforth, the subscript e refers to empirical.

The classical approach to deal with this problem is to perform a binning procedure of the variable $\mathcal{X}$. Let us suppose a sample of size N that takes the values $x_1, ..., x_i, ..., x_n, ..., x_N$. Note that n ($n \leq N$) corresponds to the number of data-points in the sample that follow a certain PDF. Let us define $x_{low} = \min\{x\}$ and $x_{top} = \max\{x\}$ and consider a partition of the real axis given by a set of $d+1$ points such that $x_{[0]} = x_{low} < x_{[1]} \leq ... < x_{top} < x_{[d]}$. Note that the notation $x_{[k]}$ does not denote the k-th value in the sample but the value of the real axis of the variable $\mathcal{X}$ in which the k-th point ($k = 0, ..., d$) in the partition is placed. This partition is thus made of consecutive semi-open intervals $\left[x_{[k]}, x_{[k+1]}\right)$ called bins, which include the closed interval $[x_{low}, x_{top}]$. This partition is called linear binning if these d points are equally-spaced in linear scale ($x_{[k+1]} - x_{[k]}$ =constant), whereas if the points are equally-spaced in logarithmic scale ($\left(\frac{x_{[k+1]}}{x_{[k]}}\right)$ =constant), the partition is called logarithmic binning [Hergarten, 2002; Christensen and Moloney, 2005; Pruessner, 2012].

The main goal is to build up an estimation of the PDF from the sample histogram. The standard approach to deal with this problem is to consider that the different intervals in which the horizontal axis has been splitted are called bins and the number of these objects is $n_{bins} = d$. At this point, it is important to remark the fact that bins can be graphically represented by steps or boxes whose width corresponds to the semi-open interval $\left[x_{[k]}, x_{[k+1]}\right)$. However, one could be interested in representing this box or step by a point. This approach has more sense not only from the mathematical point of view of estimating a PDF but also to have a reasonably overlap with the theoretical curve of the PDF. The point corresponding to the k-th bin is placed at the midpoint $\overline{x}_{[k]}$ of $x_{[k]}$ and $x_{[k+1]}$. The midpoint corresponds to the arithmetic mean of $x_{[k]}$ and $x_{[k+1]}$ in linear binning and to the geometric mean of $x_{[k]}$ and $x_{[k+1]}$ in logarithmic binning (see below). There are other choices to associate a point to a certain bin [Hergarten, 2002]. However, the reason why the midpoint of the interval is chosen

is because it is not necessary to assume that data are distributed according to a certain PDF. Consequently, the estimated PDF $f_{\mathcal{X},(e)}$ is a function which receives $\overline{x}_{[k]}$ with $k = 0, ..., n_{bins}-1$ as arguments.

In the classical linear binning approach, the horizontal axis is splitted into different bins with a bin-width Δ determined by

$$\Delta = \frac{x_{[d]} - x_{[0]}}{n_{bins}}.$$

Each of these bins covers a certain range of data and are indexed from $k = 0$ to $k = n_{bins} - 1$. Let us define the function $\mathcal{C}(k)$ as the number of points from data (counts) contained in the k-th bin. What is, the i-th point from the sample is placed at the k-th bin with $k = \lfloor \frac{x_i - x_{low}}{\Delta} \rfloor$, where the brackets $\lfloor$ and $\rfloor$ indicate the floor function (closest integer smaller or equal than the part inside the brackets), and contributes to the number of counts in that bin $\mathcal{C}(k) \rightarrow \mathcal{C}(k)+1$. Once this counting procedure has been completed for all the points in the sample, one has an histogram with n_{bins} that has to be correctly normalized in order to have a PDF. The normalization is carried out by dividing the number of counts $\mathcal{C}(k)$ in each bin by the product of the bin-width Δ and the sample size N:

$$f_{\mathcal{X},(e)}(\overline{x}_{[k]}) = \frac{\mathcal{C}\,(k)}{N\Delta}\,.$$

Since this is an estimated PDF, each value of $f_{\mathcal{X},(e)}(x_k)$ needs to be provided by an error bar. The error bars correspond to the standard deviation of $f_{\mathcal{X},(e)}(\overline{x}_{[k]})$. In order to compute the error bars for each bin, let us consider that data points are independent and that the probability that the k-th bin contains $\mathcal{C}(k)$ counts is given by a binomial distribution $\binom{N}{\mathcal{C}(k)} (1-q)^{\mathcal{C}(k)} (q)^{N-\mathcal{C}(k)}$, where q is the probability that an occurrence does not correspond to the k-th bin. Given that the estimated PDF is proportional to $\mathcal{C}(k)$, the coefficient of variation of $\mathcal{C}(k)$ (understood as a binomial random variable) will also hold for the estimated PDF:

$$\frac{\sigma_f}{\mu_f} = \sqrt{\frac{q}{N\,(1-q)}} \simeq \frac{1}{\sqrt{\mathcal{C}(k)}}\,, \tag{5.32}$$

where in the last step it has been considered that $q \simeq 1$, and the mean number of occurrences in the k-th bin (appearing in the denominator inside the square-root) has been replaced by the number of occurrences in that bin. Therefore, the standard deviation σ_f or errorbar corresponding to the k-th bin can be approximated by:

$$\sigma(\overline{x}_{[k]}) \simeq \frac{f_{\mathcal{X},(e)}(\overline{x}_{[k]})}{\sqrt{\mathcal{C}(k)}}\,. \tag{5.33}$$

However, the paradigmatic distribution of interest in the works presented along this thesis is the power-law and some considerations are needed in order to properly plot it. Since these

distributions span several orders of magnitude, it is not convenient to plot them in linear scale. A power-law distribution looks like a straight line when it is represented in log-log scale. Nonetheless, when the classical linear binning in which bins are equally spaced along the linear axis is performed in logarithmic scale, the bin-width depends on the value they are placed and the bins will have a distorted appearance. In order to avoid this distortion of the bin-width, the binning procedure needs to be performed in such a way that bins are equally spaced in logarithmic scale. This binning procedure is called logarithmic binning.

Logarithmic binning is based on splitting each decade by a certain number of bins n_{bdec}. Once the number of bins per decade is chosen, it is necessary to define a minimal index k_{min} defined by $k_{min} = \lfloor n_{bdec} \log_{10}(x_{low}) \rfloor$. According to this definition, the bin $k = 0$ corresponds to the first bin whereas the last bin is placed at $k_{max} = \left\lceil n_{bdec} \log_{10}\left(\frac{x_{top}}{x_{low}}\right)\right\rceil$. The i-th point x_i in the sample is assigned to the k-th bin according to $k = \lfloor n_{bdec} \log_{10}(x_i) \rfloor - k_{min}$. Consequently, the number of counts in that k-th bin will imply an increase in one unit $\mathcal{C}(k) \to \mathcal{C}(k) + 1$. When all points from the sample are placed according to this criterion, one has an histogram of counts that needs to be normalized by dividing each bin by its corresponding bin-width and the sample size N.

An important question to take into account is to decide to which value of the horizontal axis the k-th bin corresponds to. If the same criterion of placing the x_k at the midpoint of x_k and x_{k+1} considered for linear binning is applied in logarithmic scale, one obtains: $\overline{x}_{[k]} = \sqrt{x_{k+1}x_k} = 10^{\frac{k+k_{min}}{2n_{bdec}}}\sqrt{10^{\frac{1}{n_{bdec}}} - 1}$ [Deluca and Corral, 2013]. This implies that the binwidth Δ depends on the position of the bin:

$$\Delta(k) = x_{[k+1]} - x_{[k]} = 10^{\frac{k+k_{min}}{n_{bdec}}}\left(10^{\frac{1}{n_{bdec}}} - 1\right).$$

Taken into account all these considerations, the estimated PDF for the k-th bin is then written:

$$f_{\mathcal{X},(e)}(\overline{x}_{[k]}) = \frac{\mathcal{C}(k)}{N\Delta(k)} = \frac{\mathcal{C}(k)}{N10^{\frac{k+k_{min}}{n_{bdec}}}\left(10^{\frac{1}{n_{bdec}}} - 1\right)}.$$

The errorbars are constructed in the same way as in Eq. (5.33) but, in this case, it is important to remark that, given that the PDF is plotted in log-log scale, errorbars will not be graphically symmetric. Consequently, errorbars for a estimated power-law PDF will appear distorted although they are numerically symmetric.

Let us remark that this graphical representation of the PDF is done without assuming any underlying statistical model. One can perform the same procedure by assuming that data follow a certain PDF $f_{\mathcal{X}}\left(x;\hat{\boldsymbol{\theta}}\right)$ whose parameters have been estimated $\hat{\boldsymbol{\theta}}$. In this case, the graphical representation can be improved if one takes into account that $\overline{x}'_{[k]}$ corresponds to those values of x whose corresponding density gives the probability of being in the k-th bin

divided by its width [Hergarten, 2002; Deluca and Corral, 2013]:

$$f_{\mathcal{X},(e)}(\overline{x}'_{[k]}) = \frac{1}{\overline{x}'_{[k]} - \overline{x}'_{[k+1]}} \int_{\overline{x}'_{[k]}}^{\overline{x}'_{[k+1]}} f_{\mathcal{X}}\left(x;\hat{\boldsymbol{\theta}}\right) = \frac{F_{\mathcal{X}}\left(\overline{x}'_{[k+1]};\hat{\boldsymbol{\theta}}\right) - F_{\mathcal{X}}\left(\overline{x}'_{[k]};\hat{\boldsymbol{\theta}}\right)}{\overline{x}'_{[k]} - \overline{x}'_{[k+1]}}\,. \tag{5.34}$$

The solution for a power-law distribution with exponent $\hat{\gamma}$ is given by [Deluca and Corral, 2013]:

$$\begin{aligned}\overline{x}'_{[k]} &= \overline{x}_{[k]}\left[(\hat{\gamma}-1)\,\frac{10^{\frac{\hat{\gamma}}{2n_{bdec}}-1}\left(10^{1/n_{bdec}}-1\right)}{10^{\frac{\hat{\gamma}}{n_{bdec}}-1}-1}\right]^{1/\hat{\gamma}} \\ &= 10^{\frac{k+k_{min}}{2n_{bdec}}}\sqrt{10^{\frac{1}{n_{bdec}}}-1}\left[(\hat{\gamma}-1)\,\frac{10^{\frac{\hat{\gamma}}{2n_{bdec}}-1}\left(10^{1/n_{bdec}}-1\right)}{10^{\frac{\hat{\gamma}}{n_{bdec}}-1}-1}\right]^{1/\hat{\gamma}}.\end{aligned} \tag{5.35}$$

This selection of $\overline{x}'_{[k]}$ provides better matching with the fit but, as it has been already mentioned, needs a previous estimation of the parameters and a goodness-of-fit test and for this reason it will not be used in this PhD thesis.

The graphical representation of the empirical CDF is a very well known procedure that is carried out by sorting the sample in ascending order: $x_{(1)}, x_{(2)}, ..., x_{(i)}, ..., x_{(N)}$. Parentheses in the subindex indicate that this index is denoting order in such a way that $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(i)} \leq ... \leq x_{(N)}$. When the sample is sorted, the empirical CDF can be easily constructed by:

$$F_{\mathcal{X},(e)}(x_{(i)}) = \frac{i-1}{N}\,. \tag{5.36}$$

The complementary cumulative distribution function (CCDF) can be constructed by simply $S_{\mathcal{X},(e)}(x_{(i)}) = 1 - F_e(x_{(i)})$. Since no binning procedure is needed in this process, the $F_{\mathcal{X},(e)}(x_{(i)})$ does not exhibit distortions that can lead to misleading interpretations when it is plotted in log-log scale.

As a final remark, it is important to mention that the sub-sample of size n is the one of interest because it exhibits a certain regime, for instance a power-law behaviour, which is distorted in the rest of the bigger sample of size N. However, it is interesting to show the complete sample with its distortions in order to highlight the interesting regime that has been fitted. Let us denote as $f_{\mathcal{X}}(x,\hat{\boldsymbol{\theta}})$ and $F_{\mathcal{X}}(x,\hat{\boldsymbol{\theta}})$ to the PDF and the CDF to which the sample of size n has been fitted. Along this thesis, these functions are referred as fits. Overlapping is not achieved if one tries to plot the fits together with the estimated PDF and CDF referred to the bigger sample. Therefore, some corrections need to be done in order to see a matching between fits and estimated PDF and CDF. Since the PDF is referred to the bigger sample, thus normalized for a number of data N, overlapping is not achieved due to a normalization problem. In that case, the fit has to be multiplied by a factor $\frac{n}{N}$. The same reasoning would be applied for the CDF when it is not upper-bounded. Nevertheless, when the CDF exhibits

an upper cut-off x_{max} (for instance, when dealing with a truncated power-law), the fit has to be also shifted by a factor: $F_{\mathcal{X},(e)}(x = x_{max})$. Consequently, the fit of the CDF has to be corrected by firstly multiplying by $\frac{n}{N}$ and then adding $S_{\mathcal{X},(e)}(x_{max})$ in order to match with the empirical CDF.

5.4 Goodness-of-fit test

As stated in Sec. 5.1, MLE is a technique to estimate the parameters that characterize the statistical model from which a dataset has been "supposedly" sampled from. The word "supposedly" needs to be highlighted in the previous sentence because MLE does not provide information about how reliable is the assumption that data has been really generated according to that statistical model. MLE assumes the PDF *a priori* and estimates the parameters characterizing this statistical model that maximize the likelihood function. However, this does not imply that data are really sampled from a population characterized by this PDF. In order to provide quantitative information about the reliability that data has been sampled from a certain PDF, one needs to perform a goodness-of-fit test [Maydeu-Olivares and García-Forero, 2010; Chave, 2017].

Let us consider a sample of the random variable $\mathcal{X}$ as $x_1, ..., x_n$ that has been really generated by the probability distribution $\tilde{F}_{\mathcal{X}}(x, \boldsymbol{\theta})$ which is actually unknown. Let us apply MLE so that the parameters that maximize the log-likelihood function are $\hat{\boldsymbol{\theta}}$. In this context, a goodness-of-fit test is a non-parametric hypothesis test whose null and alternate hypothesis are:

$$\begin{cases} H_0 : \text{Data are drawn from } \tilde{F}_{\mathcal{X}}\left(x, \hat{\boldsymbol{\theta}}\right) \\ H_1 : \text{Data are not drawn from } \tilde{F}_{\mathcal{X}}\left(x, \hat{\boldsymbol{\theta}}\right) . \end{cases} \tag{5.37}$$

Once the null and alternate hypothesis are established, one has to impose a significance level α and to choose a test statistic whose distribution may be known or not. When the test statistic has been computed for the empirical dataset, a p-value can be computed either from its analytical distribution or by means of Monte Carlo simulations.

5.4.1 Monte Carlo Simulations of power-law distributed data

The standard procedure to simulate power-law distributed data [Devroye, 1986; Deluca and Corral, 2013] is explained in this section. Although there are packages in R [Gillespie, 2015] and other programming languages which have already implemented the generation of random power-law distributed numbers, it is worth explaining how to obtain them from random uniform numbers, which can be easily generated in most of high-level programming languages.

The inverse transform sampling method [Devroye, 1986] can be applied in order to obtain power-law distributed random numbers. The purpose is to obtain a sample of n i.i.d. power-law distributed with exponent γ random numbers. Let us consider the general case

where the power-law exhibits an upper truncation x_{max}. By means of the probability integral transform theorem [Angus, 1994], the CDF in Eq. (5.26) can be considered as random variable $\mathcal{U}$ uniformly distributed acquiring values u between 0 and 1:

$$\frac{x^{1-\gamma} - x_{min}^{1-\gamma}}{x_{max}^{1-\gamma} - x_{min}^{1-\gamma}} = u\,.$$

By isolating x from this last expression one obtains:

$$x = x_{min}\left[1 - \left(1 - \left(\frac{x_{max}}{x_{min}}\right)^{1-\gamma}\right)u\right]^{\frac{1}{1-\gamma}}\,, \tag{5.38}$$

which is a simple equation to generate a power-law distributed number x with exponent γ between x_{min} and x_{max} from a uniform random number u. Untruncated power-law numbers can be generated by taking the limit $x_{max} \to \infty$ in this expression (note that $\gamma > 1$).

$$x = x_{min}\,(1-u)^{\frac{1}{1-\gamma}}\,. \tag{5.39}$$

As a general procedure to estimate the p-value, one can compute N_{sim} simulations of n power-law distributed random numbers and compute the fraction of simulations $N_{sim}^{>}$ in which the test statistic computed for simulated data is greater or equal than the one found for the empirical dataset.

$$p_{value} = \frac{N_{sim}^{>}}{N_{sim}}\,. \tag{5.40}$$

In order to provide an errorbar to the estimated p-value, one can understand it as a binomial random variable. Therefore, the errorbar can be considered as the standard deviation of a binomial random variable:

$$\sigma_{p_{value}} = \sqrt{\frac{p_{value}\,(1-p_{value})}{N_{sim}}}\,. \tag{5.41}$$

5.4.2 Kolmogorov-Smirnov (KS) Goodness-of-fit test

The main test statistic used in this thesis is the well-known Kolmogorov-Smirnov (KS) statistic [Kolmogorov, 1933; Concise, 2008; W. H. Press and Flannery., 1992]. The KS statistic measures the supremum of the set of distances between the empirical CDF, F_e, and its analytic counterpart F:

$$d_e = \sup\left|F_{\mathcal{X},(e)}\,(x) - F_{\mathcal{X}}\,(x;\hat{\gamma},x_{min},x_{max})\,\right|. \tag{5.42}$$

Note that, in order to obtain the value of this test statistic, the exponent $\hat{\gamma}$ needs to be previously computed via MLE. An important point to take into account when computing

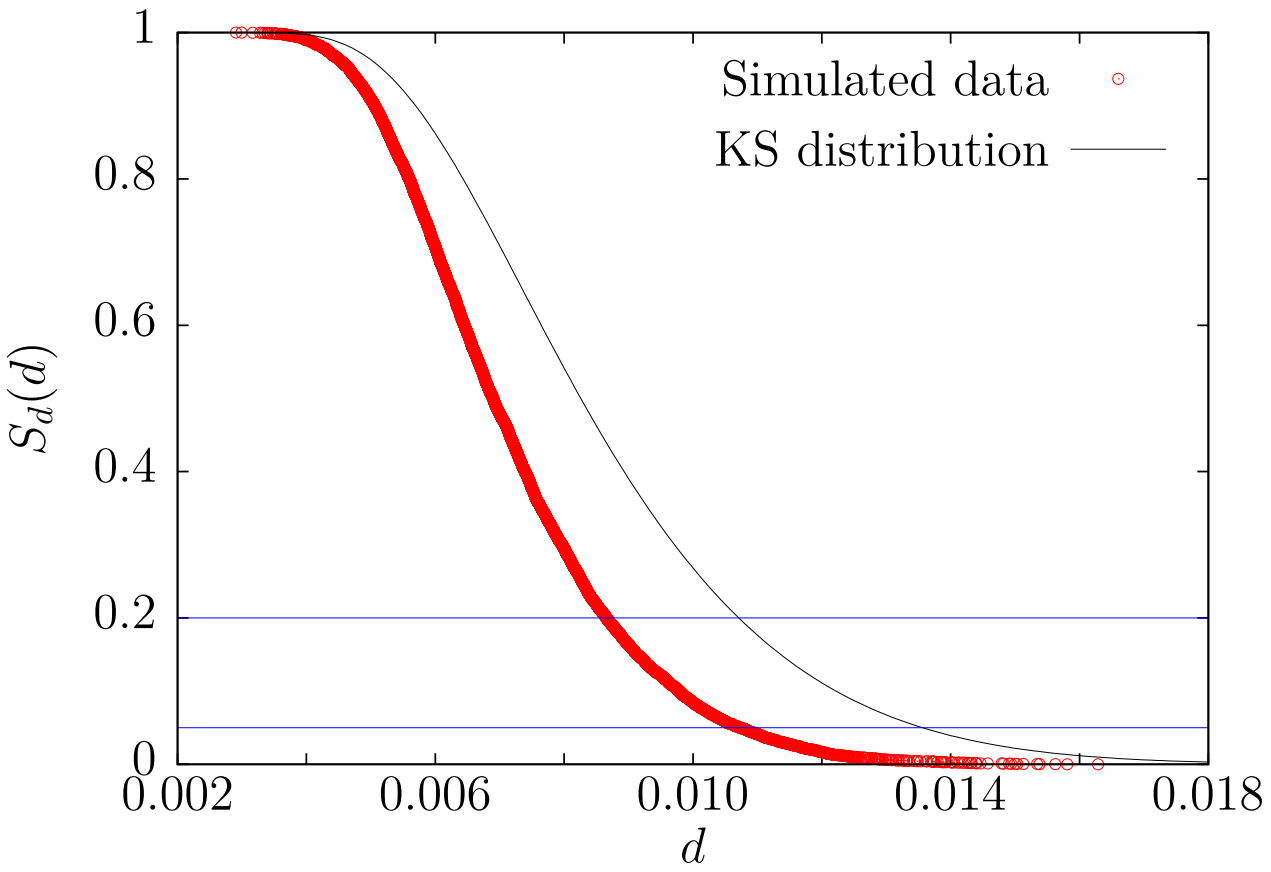

Fig. 5.4: CCDF S_d of the KS Statistic for simulated untruncated power-law data $n = 10^4$ and $\gamma = 1.35$ for 10^4 realizations. Black solid line corresponds to the KS distribution for $n = 10^4$. Blue horizontal solid lines correspond to the 5 and 20 percentiles. As it can be appreciated, the p-value is always larger if one considers the KS distribution instead of the distribution of the statistic.

the KS distance is that the emprical CDF $F_{\mathcal{X},(e)}$ is constructed as a stepwise function. This implies that, for each point in the sample, two sets of distances between the empirical and the theoretical CDF can be defined:

$$d^{+}_{e,i} = \left| F_{\mathcal{X},(e)}\left(x_i\right) - F_{\mathcal{X}}\left(x_i; \hat{\gamma}, x_{min}, x_{max}\right) \right| ,$$

$$d^{-}_{e,i} = \left| F_{\mathcal{X},(e)}\left(x_{i-1}\right) - F_{\mathcal{X}}\left(x_i; \hat{\gamma}, x_{min}, x_{max}\right) \right| .$$

Each point has a unique distance which corresponds to $d_{e,i} = \sup\{d^{+}_{e,i}, d^{-}_{e,i}\}$. Consequently, Eq. (5.42) can be rewritten:

$$d_e = \sup\left(d_{e,i}\right) \text{ for } i = 1, .., n\,. \tag{5.43}$$

Given that the distribution of the KS statistic is well known in the asymptotic limit [Feller, 1948; Smirnov, 1948; W. H. Press and Flannery., 1992], one can easily compute the p-value

as:

$$\lim_{n\to\infty} \mathcal{P}\left(\sqrt{n}d_e \leq z\right) = \frac{\sqrt{2\pi}}{z} \sum_{k=1}^{\infty} e^{-\frac{\pi^2(2k-1)^2}{8z^2}} . \tag{5.44}$$

Note that the factor $\sqrt{n}$ enables the comparison between the empirical KS statistic and the KS distribution for finite n. Inflated p-values are found if the value of the estimator is used in the goodness-of-fit test. Therefore, the p-value computed from the KS distribution is biased when the KS distance is used as a test statistic if one uses Eq. (5.42) with the empirical estimator $\hat{\gamma}$ [Deluca and Corral, 2013]. Consequently, this procedure leads to inflated p-values and can derive in misleading conclusions. In order to avoid this problem, one can estimate the p-value by means of Monte Carlo simulations as explained in Sec. 5.4.1. Figure 5.4 shows how the p-value obtained from the KS distribution is higher than the one computed through Monte Carlo simulations for power-law distributed data.

If the p-value of the fit exceeds a certain threshold p_c (typically 0.05) then one considers that the null hypothesis is accepted (in the sense that it cannot be rejected).

5.5 Fitting a power-law distribution to empirical data

Let us now suppose a sample of size N which exhibits a power-law distributed regime for a certain range of data but some deviations from this behaviour may appear for small or large values. It could also be the case that several power-law regimes with different exponents may appear in the same sample. In any of these cases, one would desire to find this limited power-law regime by fitting a truncated power-law distribution which does not include all the sample but a sub-sample of size n that acquires values in the range $x \in [x_{min}, x_{max}]$. In order to know where the power-law regime (or regimes) is, one needs to know where the cut-offs x_{min} and x_{max} are. Unfortunately, there is not an efficient procedure to know them in advance and they have to be computed by trying different combinations of x_{min} and x_{max} ($x_{min} < x_{max}$). The dependence of the power-law exponent as function of the imposed cut-offs can be studied by means of MLE maps [Baró and Vives, 2012]. Despite this representation does not provide information about the reliability of the fitted exponent, it can be useful in order to check its stability. This technique is used in different works of this PhD thesis for exploratory issues. For all the sub-samples generated by the imposition of the cut-offs, one can perform MLE and goodness-of-fit test in order to check if data are compatible with a power-law distribution in that range and with which value of the exponent γ. In order to put together all the concepts explained so far, a summary of how to fit a truncated power-law distribution is shown in a flow diagram in Fig. 5.5. This task can imply an excessive computational effort if the cut-offs are imposed at each point of the sample (assuming that there are not repeated values in the sample $\binom{N}{2}$ sub-samples should be checked). In order to reduce the computational effort, one can sample equally spaced values in logarithmic scale of x_{min} and x_{max}, thus reducing the number of combinations. At the end of this procedure, one might have obtained several

subsets for which the power-law hypothesis is not rejected. In this case of multiplicity of sub-samples with a non-rejectable p-value, one has to decide which of them is the definitive fit depending on the nature of the problem.

- Maximizing statistics: One could be interested in determining the power-law range covering the largest fraction of data. Moreover, for a given value of the exponent $\hat{\gamma}$, the larger the number of data the better the accuracy will be. In order to obtain such a measure, one needs to reduce the standard deviation of the MLE by selecting the non-rejectable fit containing the maximum number of data: $\max(n)$.

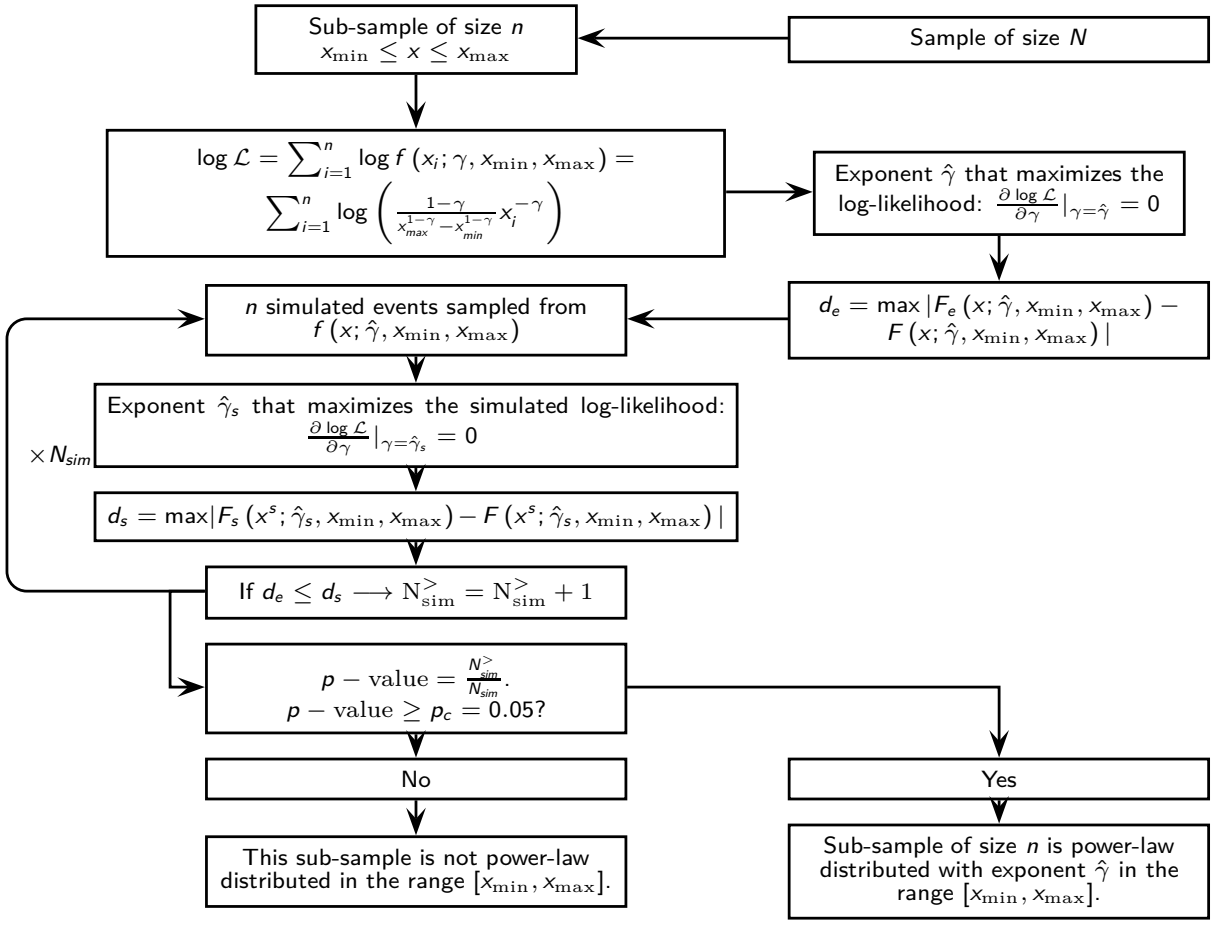

Fig. 5.5: Flow diagram for the procedure to fit a truncated power-law for the variable $\mathcal{X}$ from a sample of size N.

- Maximizing the fitting range: If one is interested in a power-law regime as broad as possible, one should choose the fit that covers the broadest range: $\max\left(\frac{x_{max}}{x_{min}}\right)$. A broader range of validity of the power-law regime implies a stronger evidence of scale-invariance.

Note that both criteria are equivalent as long as $x_{max} \to \infty$. The procedure explained in Fig. 5.5 can be implemented in a programming language. The source code for C language is provided in Appendix 9.3.

5.6 Likelihood-ratio test

The likelihood-ratio test (LRT) is a statistical procedure used to check whether it is worth using a more complex statistical model Com (characterized by a set of $n_{\mathcal{L}}^{(Com)}$ parameters) rather than a simpler one Sim (characterized by a set of $n_{\mathcal{L}}^{(Sim)}$ parameters). Usually, the simpler model is nested into the more complex one. The null hypothesis of this test states that the simpler model Sim is good enough to describe the data and a more general model Com does not provide more information. In order to perform this test, the statistic which can be used is the log-likelihood ratio:

$$\mathcal{R} = \log\left(\frac{\mathcal{L}_{Com}\left(\hat{\boldsymbol{\theta}}_{\boldsymbol{Com}}\right)}{\mathcal{L}_{Sim}\left(\hat{\boldsymbol{\theta}}_{\boldsymbol{Sim}}\right)}\right) = \log\mathcal{L}_{Com}\left(\hat{\boldsymbol{\theta}}_{\boldsymbol{Com}}\right) - \log\mathcal{L}_{Sim}\left(\hat{\boldsymbol{\theta}}_{\boldsymbol{Sim}}\right) . \tag{5.45}$$

Where the likelihoods are evaluated at those values of the set of parameters for which they reach their maximum values.

Under the regularity conditions exposed in Sec. 5.1.1 and for nested models, it is well established that the test statistic $2\mathcal{R}$ is distributed according to a chi-squared distribution with $n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)} > 0$ degrees of freedom [Pawitan, 2013]. Once the significance level α of the test has been fixed, one is able to determine the critical value $2\mathcal{R}_c$. If the empirical value of the statistic is below this threshold, then one can consider that there is not enough statistical evidence to say that the more complex model Com is needed in order to describe the data. In this situation, due to its simplicity, model Sim is preferable to fit data. This procedure can be naively understood as a statistical version of the Occam's razor criterion.

It must be mentioned that the p-value can be computed analytically by taking into account:

$$p_{value} = 1 - F_{\chi^2}\left(2\mathcal{R}\right) , \tag{5.46}$$

where F_{χ^2} is the CDF of a chi-squared distribution with $n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)} > 0$ degrees of freedom.

Note that the rejection of the null hypothesis does not imply the rejection of model Sim as a fit to data and, on the contrary, the non-rejection of model Sim does not imply that it is a good model to describe the data. The test is just a relative comparison between two models which are candidates to describe data.

As it is done for the goodness-of-fit test, one could also compute an estimation of the p-value of the LRT by means of Monte Carlo simulations when not all the regularity conditions exposed in Sec. 5.1.1 are fulfilled. If not all the regularity conditions are ensured, it can be numerically checked wheter the statistic $2\mathcal{R}$ actually follows a chi-squared distribution with $n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)} > 0$ degrees of freedom or not [Serra and Corral, 2017]. An example of numerical results to check this are shown in Fig. 7.6 in Sec. 7.1.4.

5.7 Akaike information criterion

A different comparison comes from the application of the Akaike information criterion (AIC) [Akaike, 1973]. The AIC for a particular model is defined as:

$$\mathcal{A} = 2n_{\mathcal{L}} - 2\log\mathcal{L}\left(\hat{\boldsymbol{\theta}}\right) . \tag{5.47}$$

Two different contributions can be clearly identified in this equation: the first term keeps the information about the number of parameters of the statistical model whereas the second term contains the information of the log-likelihood function of the statistical model evaluated at $\hat{\boldsymbol{\theta}}$, where it is maximum. In general, the larger the number of parameters describing the model the larger will be the likelihood. However, it is reasonable to think that at a certain point, adding more parameters to the statistical model does not imply a significant improvement in describing data and may cause overfitting problems. This is the main reason why the AIC criterion is commonly used for model comparison since it keeps a compromise between the likelihood and the cost of introducing more parameters to describe data. Therefore, for a given set of different candidate models to describe a dataset, the one with smallest AIC is preferred.

Let us consider a simpler model (Sim) characterized by $n_{\mathcal{L}}^{(Sim)}$ and a more complex model (Com) characterized by $n_{\mathcal{L}}^{(Com)}$ parameters. The simpler model is nested into the more complex one. Let us compute the difference between AIC referred to these models:

$$\begin{aligned}\Delta\mathcal{A} &= \mathcal{A}_{Com} - \mathcal{A}_{Sim} \\ &= 2\left(n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)}\right) - 2\left(\log\mathcal{L}_{Com}\left(\hat{\boldsymbol{\theta}}_{\boldsymbol{Com}}\right) - \log\mathcal{L}_{Sim}\left(\hat{\boldsymbol{\theta}}_{\boldsymbol{Sim}}\right)\right) .\end{aligned} \tag{5.48}$$

The AIC difference and the LRT statistic can be easily related by:

$$\Delta\mathcal{A} = 2\left(n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)}\right) - 2\mathcal{R} . \tag{5.49}$$

Given that the distribution of the LRT statistic is known, one could also relate the AIC with a p-value through the following expression:

$$p_{value} = 1 - F_{\chi^2}\left(2\left(n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)}\right) - \Delta\mathcal{A}\right) . \tag{5.50}$$

where F_{χ^2} is the CDF of a chi-squared distribution with $n_{\mathcal{L}}^{(Com)} - n_{\mathcal{L}}^{(Sim)} > 0$ degrees of freedom. In this way, the LRT and AIC provide exactly the same information about which model is preferred [Burnham and Anderson, 2002; Murtaugh, 2014].

5.8 Two-sample Kolmogorov-Smirnov test

The two-sample KS test is a statistical test which aims to determine whether two different datasets have been sampled from the same population or not. Let us consider two datasets (1) and (2) with sizes n_1 and n_2 which have been sampled from the distributions $F_1(x)$ and $F_2(x)$ respectively. The null and alternative hypotheses of the two-sample Kolmogorov-Smirnov test state that:

$$\begin{cases} H_0 : F_1(x) = F_2(x) \\ H_1 : F_1(x) \neq F_2(x) \end{cases} \tag{5.51}$$

The test statistic can be constructed by computing the empirical CDF for each sample (see Sec. 5.3) and finding:

$$d_{n_1,n_2} = \max \left| F_e^{(1)}(x) - F_e^{(2)}(x) \right|. \tag{5.52}$$

For large n_1 and n_2, the p-value of the test can be easily calculated by replacing n by $\frac{n_1 n_2}{n_1+n_2}$ in Eq. (5.44) [Feller, 1948; W. H. Press and Flannery., 1992].

Note that this statistical test does not take into account the particular shape of the probability distribution data come from. Let us consider, for instance, two samples whose underlying distributions are already known, the functional form of both distributions is the same but some parameters characterizing them may differ. One question that could be interesting to determine is whether the parameters characterizing both distributions are significantly different or not. In that case, this test would provide a good solution to know whether these differences are significant or not without estimating any of the parameters.

5.9 Compatibility of power-law exponents

Let us consider two samples (1) and (2) with sizes n_1 and n_2 which are power-law distributed with exponents and standard deviations $\hat{\gamma}_1$, σ_1 and $\hat{\gamma}_2$,σ_2 respectively, all of them computed by means of MLE. In this framework, the objective lies in knowing whether these exponents are significantly different or not.

The null and alternative hypotheses for this test are:

$$\begin{cases} H_0 : \hat{\gamma}_1 = \hat{\gamma}_2 \\ H_1 : \hat{\gamma}_1 \neq \hat{\gamma}_2 \end{cases} \tag{5.53}$$

As it was shown in Sec. 5.1.2, power-law exponents estimated by means of MLE are normally distributed around their real value in the asymptotic limit. This implies that the two-sample z-statistic for the test comparing two means can be used [Kagan, 1999]:

$$z = \frac{\hat{\gamma}_2 - \hat{\gamma}_1}{\sqrt{\sigma_2^2 + \sigma_1^2}}, \tag{5.54}$$

where the population variance is approximated by those from the samples (1) and (2) and zero covariance between $\hat{\gamma}_1$ and $\hat{\gamma}_2$ is assumed.

An additional assumption is that z is normally distributed, which is supported by the theory in the asymptotic limit (n_1 and n_2 going to infinity [Pawitan, 2013]). By assuming normality the null hypothesis is tested by just comparing the value of z with the standard normal distribution and the hypothesis is rejected if the value of z is too extreme for a given significance level. In quantitative terms this is given by a $p-$value, called p_{norm}, smaller than the significance level α (for instance 0.05 thus corresponding to 0.95 confidence).

If one does not want to believe that the asymptotic regime has been reached the best option is to use a permutation test [Good, 2011]. Under the null hypothesis, one is allowed to aggregate both samples (1) and (2) and take, without repetition, two random sub-samples of size n_1 and n_2; note that this is equivalent to take a permutation of the aggregated sample and separate it into two parts ((1) and (2)). One proceeds in the same way as in the original data, calculating (by MLE) $\hat{\gamma}_1^*$, $\hat{\gamma}_2^*$, and from here σ_1^*, σ_2^*, and z^*, where the asterisk marks that the values are those from the permuted data. Repeating the permutation procedure many times, the distribution of z^* can be found so as to compare to the original value z. The $p-$value of the permutation test, p_{perm}, is given by the fraction of permutations for which $|z^*|$ is larger than $|z|$ (the empirical value).

Let us highlight the fact that this technique is less restrictive than the two-sample KS test since not all the parameters characterizing both power-law distributions have to be necessarily compatible. This test can be applied when the power-law distributions are defined in different ranges, whereas the two-sample KS will reject the null hypothesis in this case even if the power-law exponents are the same.

5.10 Residual coefficient of variation (CV) test

The Residual Coefficient of Variation (CV) test [del Castillo and Puig, 1999] was originally thought as a statistical test to distinguish between exponential and truncated normal distributions. However, by considering the logarithm of data, this test can check whether an untruncated power-law distribution is a reasonable model to fit the tail [Malevergne et al., 2011]. Despite talking about CV test can be a misnomer with respect the original test, the CV test is used under this name along this thesis with the purpose of identifying power-law tails.

Let us suppose that we have a variable $\mathcal{X}$ that acquires n values $\{x\}$. We define $x_{(j)}$ as the j-th value of the variable when it is sorted in ascending order: $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(j)} \leq ... \leq x_{(n)}$. The null hypothesis of the test states that there exists a power-law tail in the range $x > x_{(k)}$. The alternative hypothesis can be formulated in two different ways. The hypothesis test would be one- sided if one considers that the alternative hypothesis is that there exists a truncated log-normal in the range $x > x_{(k)}$. If the alternative hypothesis is simply that that it does not exist a power-law distribution in the range $x > x_{(k)}$, then the hypothesis test is two-

sided. Nevertheless, the original test [del Castillo and Puig, 1999] deals with the one-sided case (truncated log-normal alternative) The test is based on the idea that for an untruncated power-law distribution the logarithmic CV is close to 1:

$$CV_l = \frac{s_l}{m_l}\,, \tag{5.55}$$

where m_l corresponds to the mean of the logarithm of the rescaled variable $l = \log(x/x_{(k)})$:

$$m_l = \frac{1}{n-k}\sum_{j=k+1}^{n} \log\left(\frac{x_{(j)}}{x_{(k)}}\right)\,, \tag{5.56}$$

and s_l^2 is its adjusted variance:

$$s_l^2 = \frac{1}{n-k-1}\sum_{j=k+1}^{n} \left(\log\left(\frac{x_{(j)}}{x_{(k)}}\right) - m_l\right)^2\,. \tag{5.57}$$

Given a value of k or $x_{(k)}$, in order to check whether the value of the residual CV is close to one or not for a particular number of remaining data, one simulates many samples power-law distributed with the same number of data $n-k$ in order to extract a distribution for CV_l. It is important to remark that the distribution of the statistic does not depend on the power-law exponent and, therefore, it is not necessary to estimate it previously. Once one determines the significance level α for the test, one can obtain just one critical value (one-sided version of the test) or both upper and lower critical values of the statistic (in the two-sided version of the test). The one-sided version of the test would simply deal with one critical value that would be obtained from the percentile $100\cdot\alpha$ of the distribution of simulated values. If the empirical value of CV_l is below this value, the power-law model is rejected in favour of a truncated log-normal [del Castillo and Puig, 1999]. For the two-sided version of the test, if the empirical value of CV_l lies between the critical values (percentiles $100\cdot\alpha/2$ and $100\cdot(1-\alpha/2)$), one cannot reject the null hypothesis of power-law tail. Otherwise, if the empirical value is below percentile $100\cdot\alpha/2$, the power-law is rejected; if the empirical value is above percentile $100\cdot(1-\alpha/2)$ the power-law is rejected but there is no alternative.

One proceeds by computing the residual CV_l for increasing values of the index k, thus analysing the remaining values in the tail of distribution. Figure 5.6 shows three CV plots for untrucated power-law, truncated power-law and a truncated power-law regime with an untruncated tail. As it can be appreciated, the fat tail of the untruncated case remains inside the confidence interval whereas for the truncated case the CV plot remains below the percentile 2.5 because the test detects the upper truncation. The mixed case where an untruncated power-law emerges at $x_{max} = 10^4$ with an exponent $\gamma_2 = 2$ detects this fat tailed with 120 events. This last mixed case is a useful illustration to justify that the CV plot can give information about where a truncation in a power-law regime exists and a different power-law behavior emerges from that point.

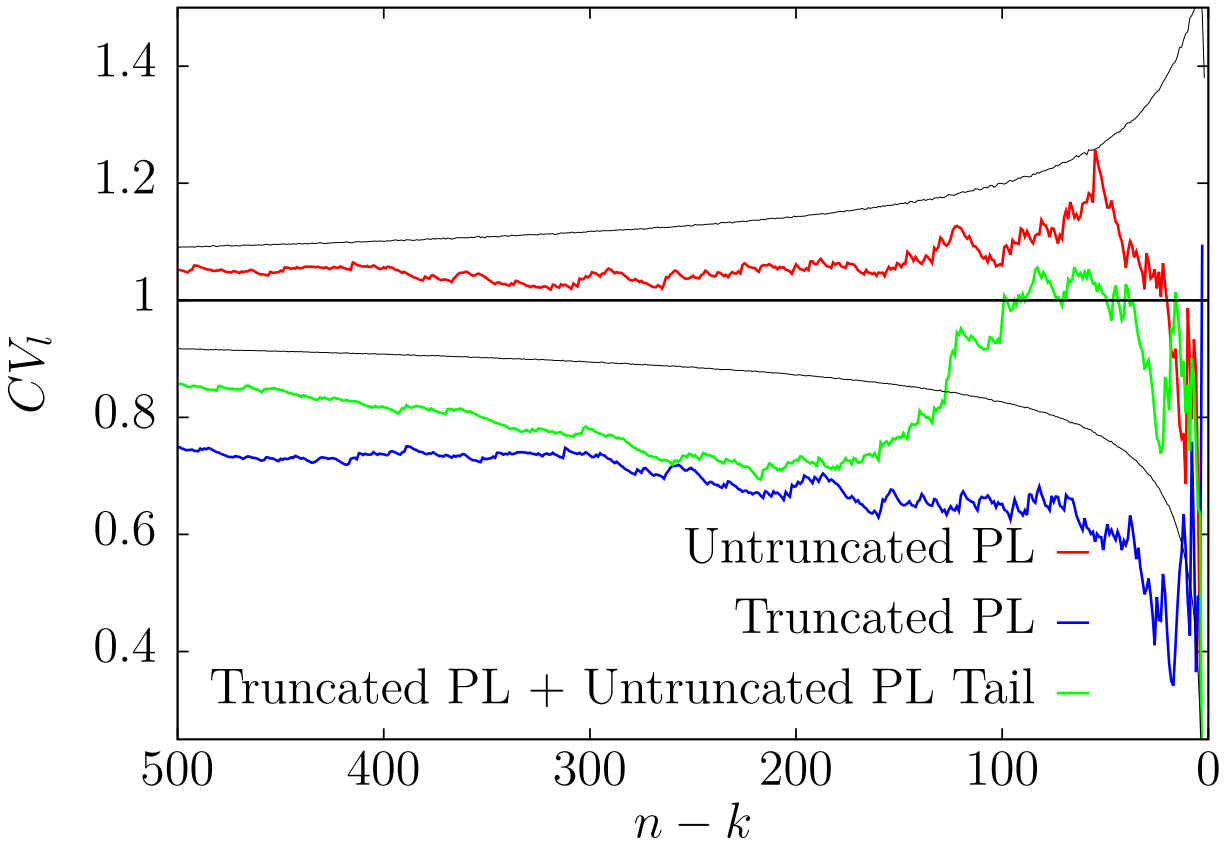

Fig. 5.6: CV plots for synthetic power-law distributed data. CV plot for untruncated power-law distributed data with exponent $\gamma = 1.5$, $x_{min} = 1$ $n = 10^4$ in red, CV plot for truncated power-law distributed data with exponent $\gamma = 1.5$, $x_{min} = 1$, $x_{max} = 10^4$ and $n = 10^4$ in blue, and CV plot for a truncated power-law from $x_{min} = 1$, $x_{max} = 10^4$, $n_1 = 9880$ with an untruncated power-law tail with exponent $\gamma_2 = 2$ arising from $x'_{min} = x_{max} = 10^4$ with $n_2 = 120$. 95% confidence interval shown in black dashed lines.

Part III

Results

Chapter 6

Acoustic emission during displacement-driven compression experiments of porous materials: labquakes

Novel findings in the statistical study of labquakes are presented in this chapter. The results obtained from the AE data recorded during displacement-driven compression experiments of Vycor are presented in Sec. 6.1 and analysed by using the techniques explained in Chpt. 5. One important goal of the statistical analysis is to determine whether the exponent of the distribution of AE energy found in displacement-driven compression experiments is significantly different from those found in force-driven compression experiments [Salje et al., 2011; Baró et al., 2013]. The statistics of force drops in displacement-driven compression experiments of Vycor as well as their relationship with the AE energy are investigated in Sec. 6.2. Results exposed in Secs. 6.1 and 6.2 correspond to the paper:

- Víctor Navas-Portella, Álvaro Corral, and Eduard Vives, Avalanches and force drops in displacement-driven compression of porous glasses. Physical Review E 94, 033005 (2016) [Navas-Portella et al., 2016].

By means of the methods explained in Sec. 5.5, the energy distribution of labquakes generated during the displacement-driven compression of charcoal is studied in Sec. 6.3.

6.1 Independence on the driving mechanism for the energy distribution of labquakes in Vycor

As it was exposed in Chpt. 3, Baró et al. [2013] found statistical similarities between earthquakes and the AE during compression experiments of porous materials. In that case, the experiments were performed using the applied force as a driving parameter, which means that the force increases linearly in time (force-driven compression). Within the context of structural phase transitions, it has been shown that avalanche scale-invariance manifests in different ways depending on the driving mechanism [Pérez-Reche et al., 2008]. If the control variable for the driving is the force, disorder plays an important role leading to a dominant nucleation process and the criticality is of the order-disorder type. However, if the driving mechanism consists in the control of the displacement, the critical state is reached independently of the disorder and by means of a self-organized criticality mechanism. These results were experimentally confirmed [Vives et al., 2009; Planes et al., 2013] based on the study of amplitude and energy distributions in AE experiments of martensitic transformations. The influence of the driving mechanism has been studied in the slip events occurring in compressed microcrystals [Maaß et al., 2015]. One important question is whether the driving mechanism will influence or not the distributions of AE events in the case of failure under compression experiments. This question is important because when comparing with earthquakes, the natural accepted mechanism is that tectonic plates are driven at constant velocity at far enough distances from the faults [Larson et al., 1997].

In force-driven compression experiments of porous glasses [Salje et al., 2011; Baró et al., 2013] it was found that the energy PDF $f_{\mathcal{E}}(E)$ of AE events follows an untruncated power-law with exponent $\hat{\gamma}_E = 1.39 \pm 0.05$ independently of the loading rate (0.2 kPa/s - 12.2 kPa/s),

$$f_{\mathcal{E}}(E) = (\gamma_E - 1)\, E_{min}^{\gamma_E - 1} E^{-\gamma_E}, \tag{6.1}$$

where $E_{min} \sim 1$ aJ is the lower bound required for the normalization of the PDF (see Chpt. 3 for more details about force-driven experiments).

The first series of experiments in Table 4.2 in Chpt. 4 were performed in order to study the value of the energy exponent γ_E for displacement driven experiments. Figure 6.1 shows a typical experimental output for the sample V12. Panel (a) displays the sequence of energies of the AE events and the evolution of the force as a function of the time. The acoustic activity rate r has been computed as the number of events per unit time recorded along windows of 20 seconds. Its behaviour is shown in Figure 6.1(b) together with the cumulative number of events as a function of the time. It must be noticed that the values of the energy acquire values along five or six orders of magnitude whereas force drops occur along the whole curve and clearly show variability on three or four orders of magnitude. In general, the largest force drops coincide with AE events with very large energy. Figure 6.2 shows an example of histogram of the energy of AE events for the sample V12 in one displacement-driven experiment. As can be seen, data seems to follow the Gutenberg-Richter law for more than 6 decades.

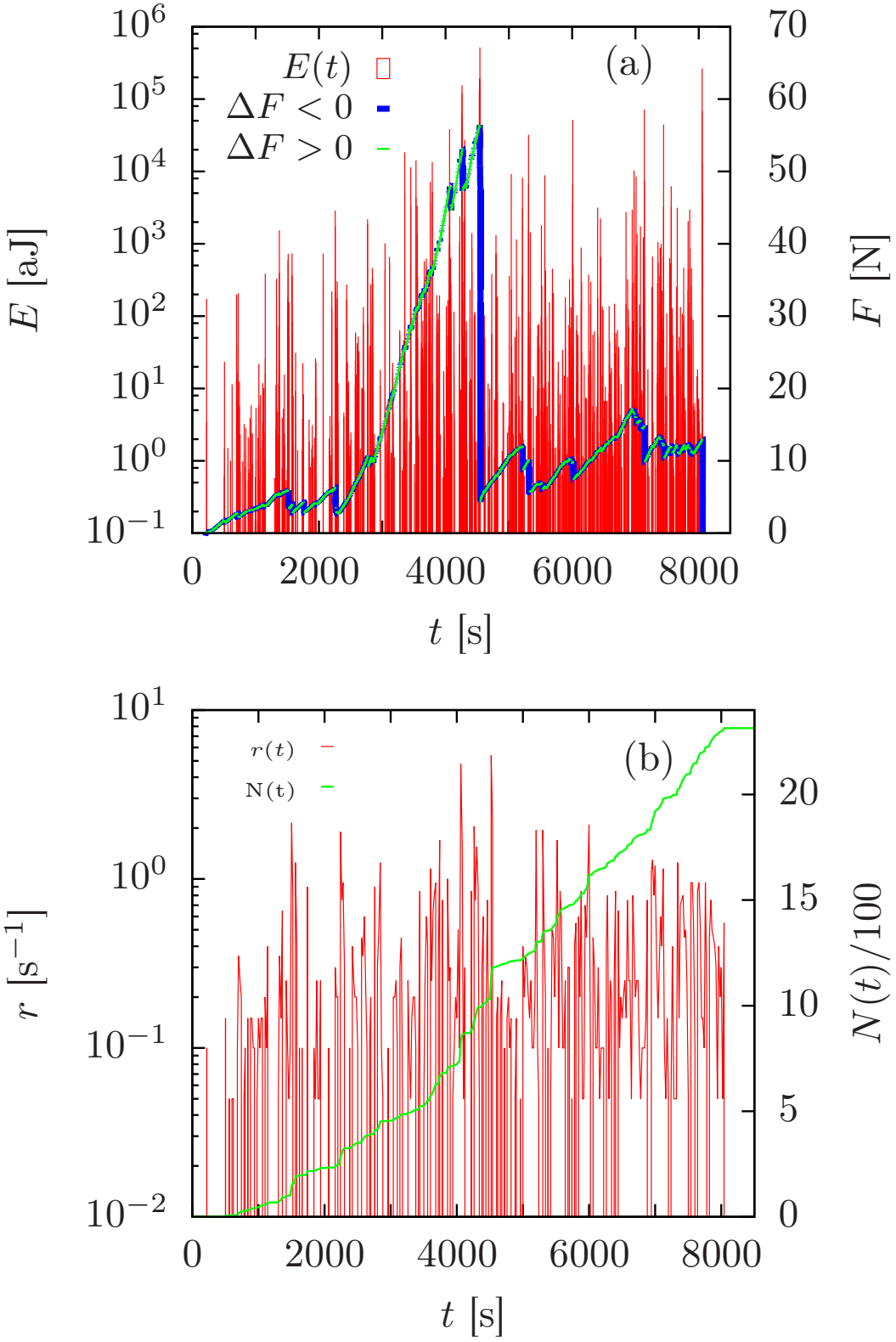

Fig. 6.1: Typical output for the sample V12. (a) Energy of the AE events as well as the measure of the force as a function of time. Green lines represent those time intervals ($\Delta t = 0.1$ s) in which the force increases whereas blue lines represent those for which the force decreases (force drops). (b) Activity rate of the experiment as well as cumulative number of AE events $N(t)$ as a function of time.

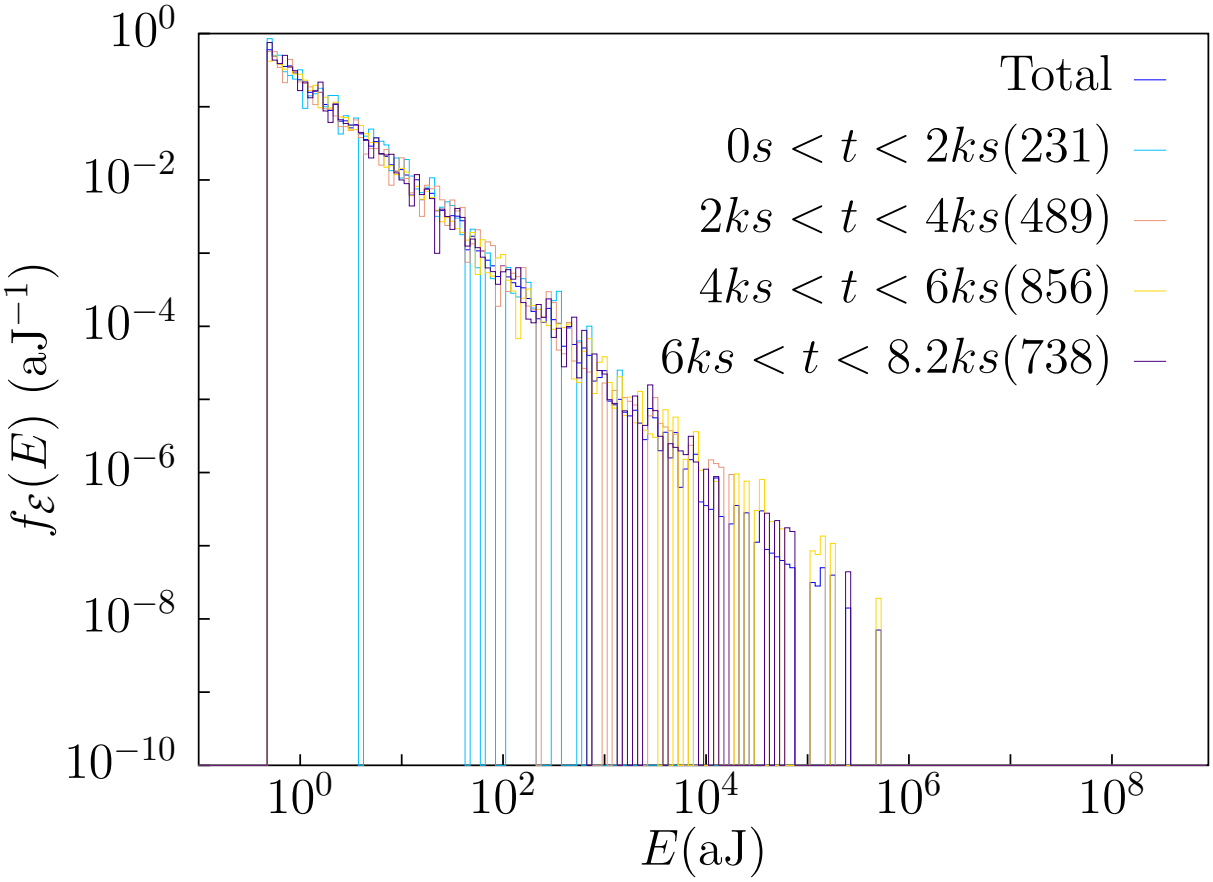

Fig. 6.2: Energy distribution for the sample V12 for different time windows (in kilo-seconds ks) as well as for the whole experiment. The numbers in parentheses account for the number of AE events in each time interval.

The different curves, corresponding to consecutive time windows of approximately 2000 seconds, reveal that the energy distribution looks stationary.

The procedure exposed in Sec. 5.5 is applied in order to guarantee statistical significance in the fit of the untruncated power-law distribution with exponent γ_E and lower threshold E_{min}. Considering as a null hypothesis that the energy distribution follows an untruncated power-law (see Eq. (6.1)), MLE for the exponent γ_E is computed for increasing values of the lower threshold E_{min} (see Figure 6.2). For each lower threshold and its corresponding exponent, the KS test explained in Sec. 5.4 is performed with a resulting p-value. The final values of the exponent and the threshold are chosen once the p-value has first overcome the significance level $p_c = 0.05$ and the power-law hypothesis cannot be rejected. The obtained values for every sample are shown in Figure 6.3 together with the standard deviation of the MLE. The horizontal line in panels (a) and (b) of Fig. 6.3 shows the average value and associated standard deviation of the exponents $\mu_{\gamma_E} = 1.35 \pm 0.05$. The values of the exponents show the

existence of some degree of sample to sample variability, and so, the concept of universality cannot be established. Despite the apparent change of the exponent with the cut-off E_{min} (see Fig. 6.3(a)), the fitting method does not rule out the presence of a power-law regime. Apparently, the value of the exponent follows a normal distribution with mean $\mu_{\gamma_E} = 1.35$ and standard deviation $\sigma_{\gamma_E} = 0.05$. Under the null hypothesis that the exponents are normally distributed with these ML estimators, the Lilliefors test for normality [Lilliefors, 1967] gives a p-value of 0.61, which implies that this null hypothesis cannot be rejected. Complementary information obtained from the fitting method is presented in Table 6.1. The average value of the fitted exponents ($\mu_{\gamma_E} = 1.35 \pm 0.05$) found for the present 13 displacement-driven experiments is compatible with the value found in force-driven measurements $\gamma_E = 1.39 \pm 0.05$ (3 experiments [Baró et al., 2013]). Not only the confidence intervals of these exponents overlap but also the test for comparing power-law exponents exposed in Sec. 5.9 gives a p-value of 0.61, thus finding not enough evidence to reject the hypothesis that both exponents are compatible. These statistical evidences elucidate that the difference between the exponent found for displacement-driven and force-driven compression experiments is not significant. Contrarily to what happens in martensitic transformations [Planes et al., 2013], it can be concluded that there are no clear evidences that the driving mechanism changes the value of the exponent of the AE energy distribution in compression experiments.

Sample	N	n	E_{min}(aJ)	E_{top} (aJ)	$\hat{\gamma}_E$	σ
V105	869	495	2.968	1.84×10^5	1.40	0.01
V11	1438	1358	0.635	2.69×10^6	1.310	0.008
V115	836	700	1.008	1.10×10^6	1.37	0.01
V12	2314	825	10.076	5.16×10^5	1.46	0.02
V125	1097	794	1.563	4.94×10^5	1.38	0.01
V205	4160	2429	3.075	9.97×10^6	1.358	0.007
V21	4170	1048	47.943	1.07×10^7	1.40	0.01
V22	3683	746	117.583	6.57×10^6	1.39	0.01
V23	1275	1146	0.645	7.16×10^6	1.284	0.008
V26	2071	2065	0.644	1.38×10^7	1.307	0.007
V28	974	973	0.503	2.82×10^6	1.288	0.009
V212	1646	1004	3.15	4.37×10^6	1.31	0.01
V24	2129	534	80.385	5.97×10^6	1.34	0.01

Table 6.1: Total number of AE events N, number of those which are power-law distributed n, value of the lower threshold E_{min}, maximum value E_{top}, exponent $\hat{\gamma}_E$ and standard deviation of the MLE σ.

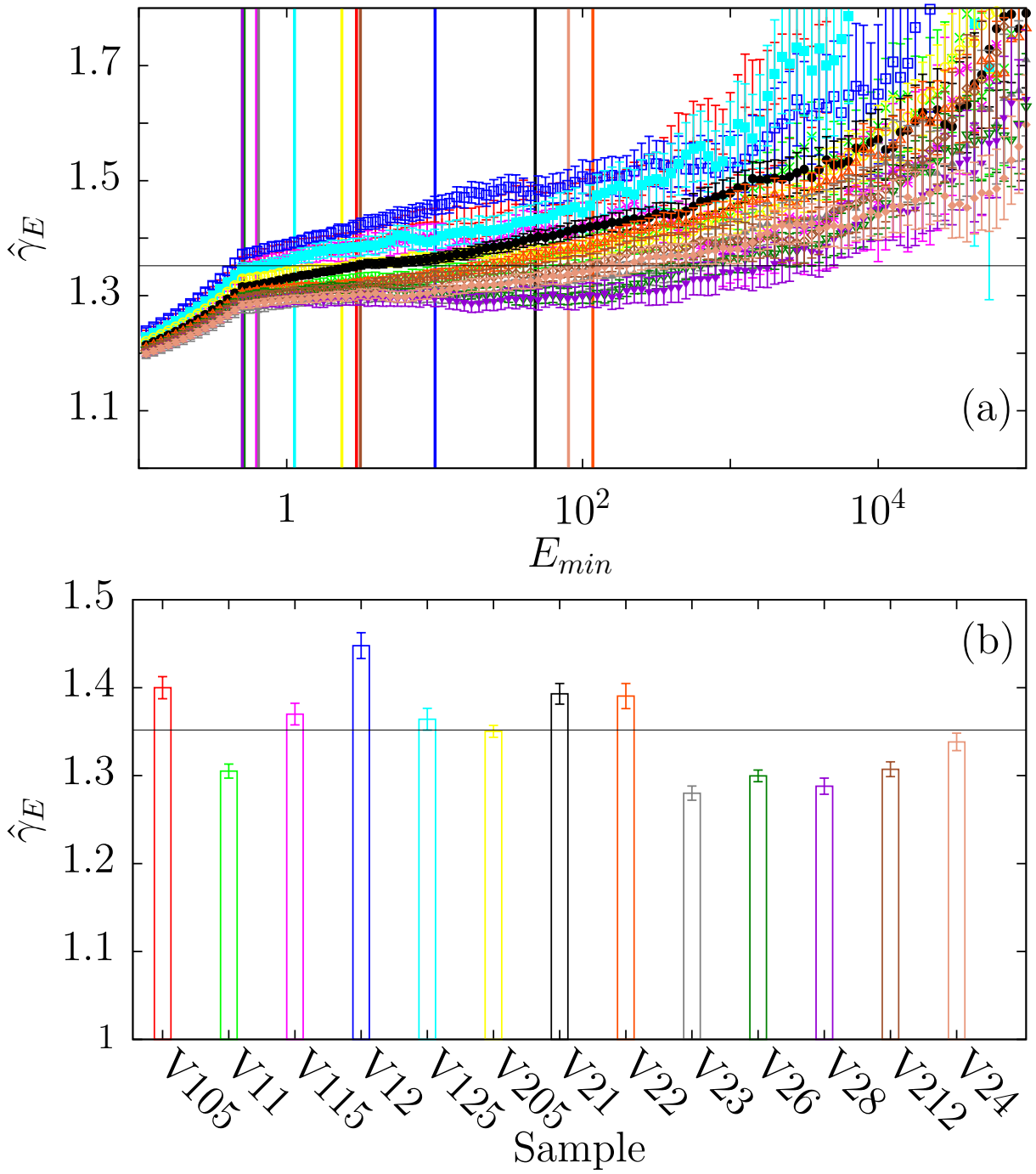

Fig. 6.3: (a) MLE of the exponent γ_E as a function of the lower threshold E_{min} for all the samples. Vertical lines correspond to the fitted values of E_{min} and $\hat{\gamma}_E$. (b) value of the exponent γ_E for each sample. The color code for each sample can be read from the color bars in (b). The dark horizontal line in both panels is the mean value of the exponent $\mu_{\gamma_E} = 1.35$.

6.2 Force drops and relationship between AE energy and force drops

When changing the driving mechanism from force driven to displacement driven, the first main macroscopic difference is that force fluctuates and shows drops that, as will be shown, correlate with AE energy. Recall that the force is measured and recorded according to the procedure explained in Sec. 4.1.4. Illa et al. [2015] have shown that the driving mechanism influences the nucleation process in martensitic transformations and these microscopic effects can lead to macroscopic changes in stress-strain curves in which force fluctuations appear. A power-law distribution with an exponential tail has been found for torque drops in shear experiments of granular matter [Lherminier et al., 2019]. Serrations or force drops have also been studied in metallic single crystals [Lebyodkin et al., 1995], metallic glasses [Antonaglia et al., 2014; Sun et al., 2010; Dalla Torre et al., 2010] and in high-entropy alloys [Carroll et al., 2015]. These studies are essentially focused on the presence of power-law distributions. Furthermore, Dalla Torre et al. [2010] studied the AE during the compression of metallic glasses and concluded that there exists a correlation between AE bursts and stress drops.

The evolution of the force as a function of time is shown in Figure 6.1(a). Force changes are defined as $\Delta F(t) = -\left(F(t+\Delta t) - F(t)\right)$, with a time step $\Delta t = 0.1$ s, so that force drops are positive. As can be observed in Fig. 6.4(a)-(c), the distribution of ΔF can exhibit several contributions. There is a possible Gaussian-like peak corresponding to negative ΔF that shifts to the left when increasing the compression rate. This peak is related to the average elastic behaviour of the porous material. The rest of contributions in the negative part of the histogram correspond to the different elastic regimes of the material as it experiences successive failures.

Statistical analysis is focused on the positive part of this distribution which corresponds to the force drops. The main question to be addressed is whether the force drops are power-law distributed or not. In Figure 6.5(a)-(c) the distribution of force drops ($\Delta F > 0$) corresponding to Figure 6.4 is shown in log-log scale. For completeness, CCDFs are also shown in Fig. 6.6 (a)-(c). The hypothesis is that force drops ΔF can be described as a random variable ϕ which follows an untruncated power-law PDF:

$$f_\phi(\Delta F) = \frac{\gamma_\phi - 1}{\Delta F_{min}^{1-\gamma_\phi}} \Delta F^{-\gamma_\phi} \, .$$

In order to determine from which value ΔF_{min} the power-law hypothesis holds, the fit of the right tail of the distribution of ΔF is performed following the same procedure as that followed for the energy distribution. In Figure 6.7(a) MLEs of the exponent $\hat{\gamma}_\phi$ as a function of the lower threshold ΔF_{min} for the samples compressed at different compression rates are shown. Three samples have been excluded due to wrong sampling of the measurement of the force [1].

[1]The wrong sampling of the force was due to a problem with the Test Expert software which stopped the recording of values of the force when the memory capacity was exceeded.

Sample	n_D	n	ΔF_{min}(N)	ΔF_{top}(N)	$\hat{\gamma}_\phi$	σ
V115	9960	174	1.73×10^{-2}	8.31	1.79	0.06
V12	32323	445	2.12×10^{-2}	27.61	1.95	0.05
V125	26104	208	1.93×10^{-2}	24.31	1.80	0.06
V205	5603	334	3.55×10^{-2}	853.36	1.53	0.03
V21	10787	149	0.16	977.78	1.72	0.06
V22	6609	133	9.05×10^{-2}	801.63	1.57	0.05
V23	9987	162	1.61×10^{-2}	593.19	1.46	0.04
V28	8881	113	1.72×10^{-2}	340.22	1.53	0.05
V212	9030	202	1.45×10^{-2}	247.56	1.55	0.04
V24	3742	53	5.80×10^{-2}	797.63	1.32	0.04

Table 6.2: Total number of force drops N and the resulting values of the number of those data which are power-law distributed n, values of the lower threshold ΔF_{min} and the value of the largest force drop ΔF_{top}, fitted exponent $\hat{\gamma}_\phi$ and standard deviation of the MLE.

Vertical lines of different colors represent the selected threshold ΔF_{min} for each sample. Note that, contrarily to what happens in the MLE of the energy exponent, for the lowest values of ΔF_{min} where the power-law hypothesis is not already valid, there is an overestimation of the exponent due to the presence of the Gaussian peak. The value of the exponents $\hat{\gamma}_\phi$ for the different samples is shown in Figure 6.7(b). The value of the exponent is higher for the slower compression rate and decreases for increasing compression rates. Moreover, power-law regime holds for three decades in the case of the slower compression rate and four decades for the higher ones. This difference is essentially due to the difference of sample surfaces. The larger the surface contact between the sample and the plate, the larger the force opposed by the material. Note that, in contrast to Fig. 6.4, the distribution of ΔF is conditioned to ΔF larger or equal than the lower threshold ΔF_{min} obtained from the fit. Additional parameters resulting from fits are shown in Table 6.2.

Dalla Torre et al. [2010] found that there exists a correlation between force drops and AE events but no evidence of correlation between the amplitude of these signals and the magnitude of the force drops was found. Nevertheless, the energy could show a certain correlation with force drops since not only the amplitude plays an important role in its calculation but also the duration of the AE events. As it can be appreciated in Figure 6.1(a), the largest force drops correspond with the highest energy of AE events. This could indicate that it actually exists a certain relationship with the values of the energy and force drops.

It is interesting to study this correlation for two reasons: on the one hand, it would set a relation between the energy of AE events, which is from microscopic nature (aJ), and force drops, which are at the macroscopic scale (N). On the other hand, force drops appear every time there is a micro-failure in the sample and thus they can be understood as releases of elastic energy. In the same way as Dalla Torre et al. [2010] a correspondence in time between the occurrence of force drops and the presence of AE events is found.

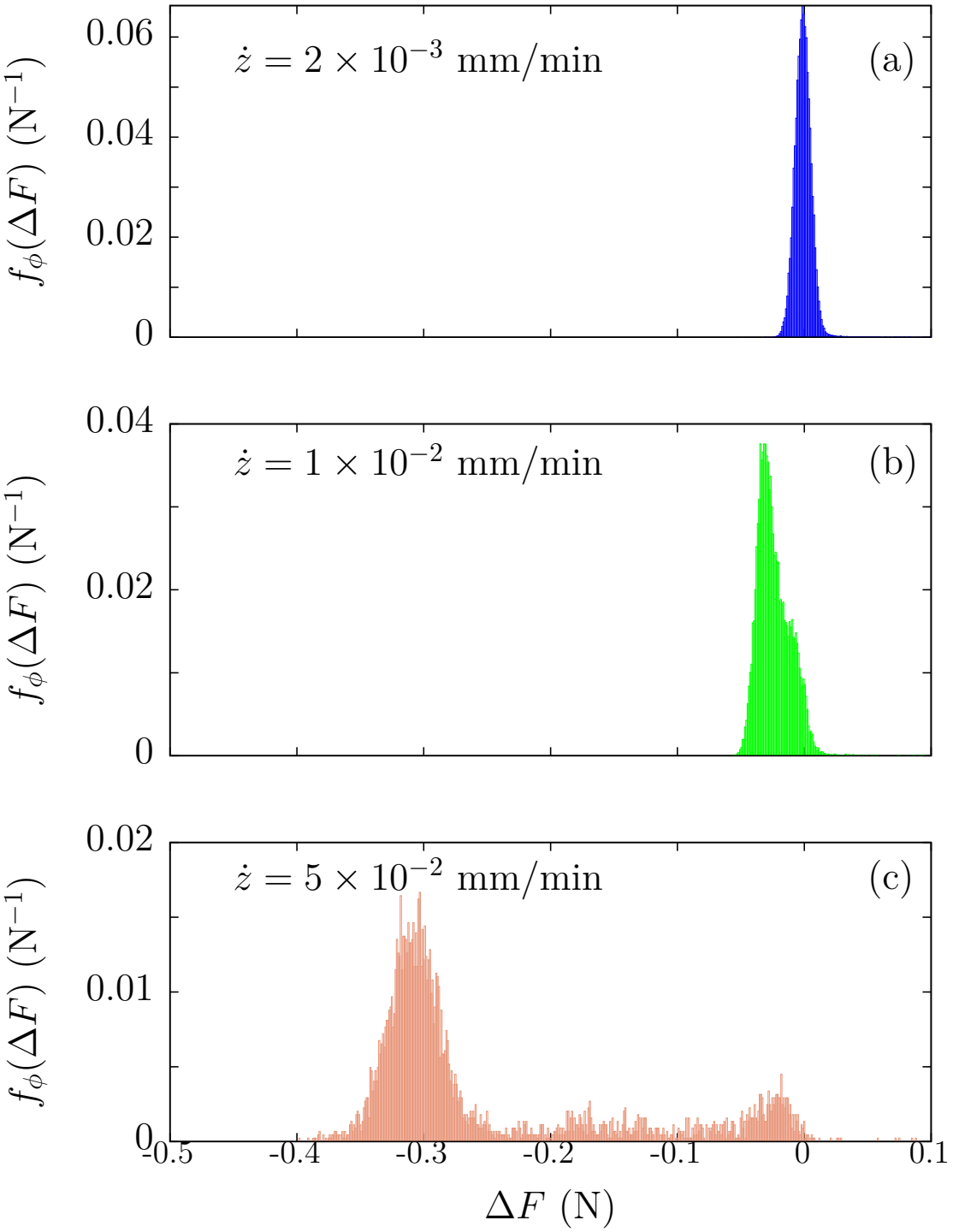

Fig. 6.4: PDFs of ΔF for three samples with different compression rates restricted to the values of ΔF close to zero. Sample V12 compressed at $\dot{z} = 2 \times 10^{-3}$mm/min is shown in (a), sample V212 compressed at $\dot{z} = 1 \times 10^{-2}$mm/min in (b) and sample V24 compressed $\dot{z} = 5 \times 10^{-2}$mm/min is presented in (c).

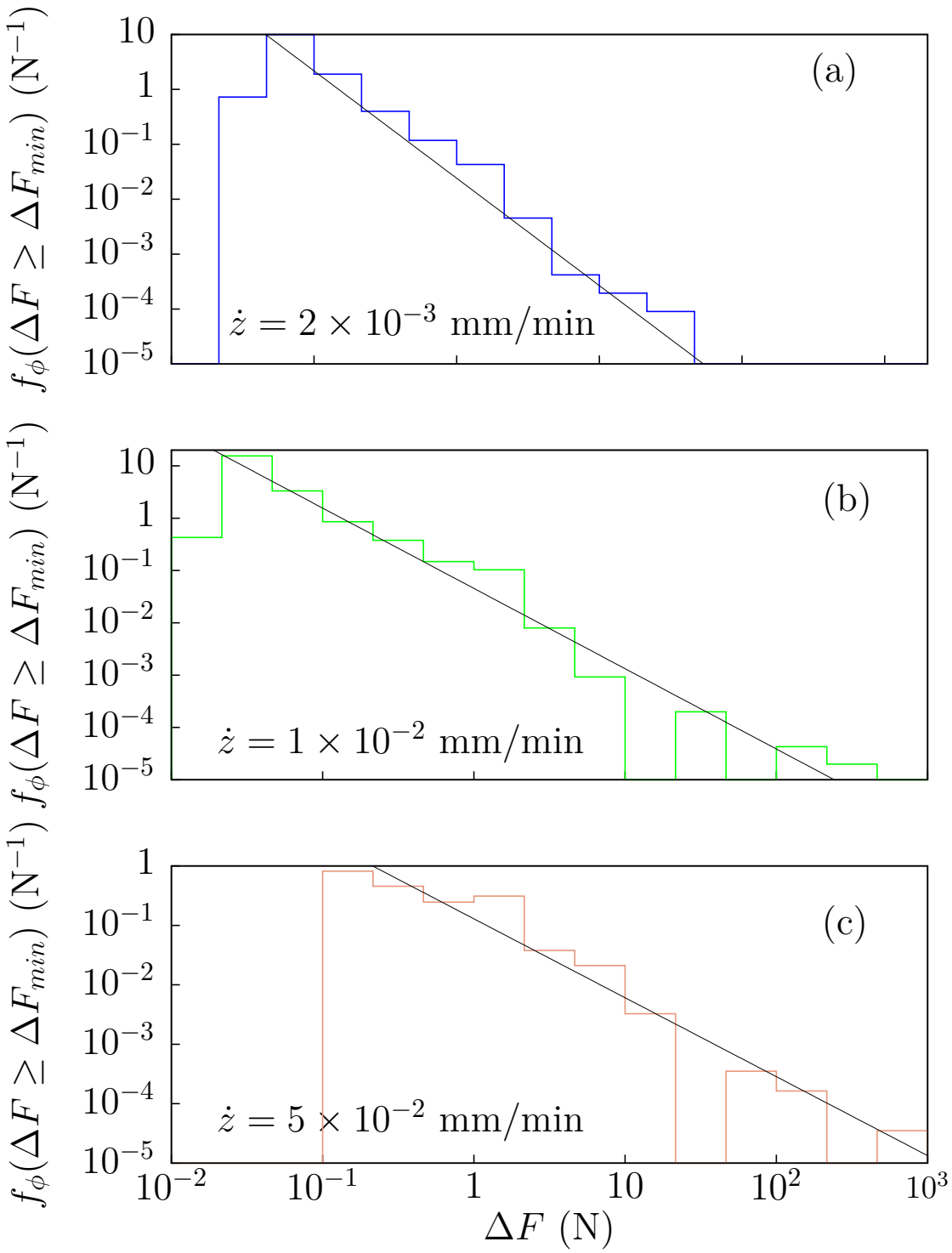

Fig. 6.5: PDFs of force drops ΔF and corresponding power-law fits for V12 (a), V212 (b) and V24 (c). Distributions are displayed and normalized for $\Delta F \geq \Delta F_{min}$.

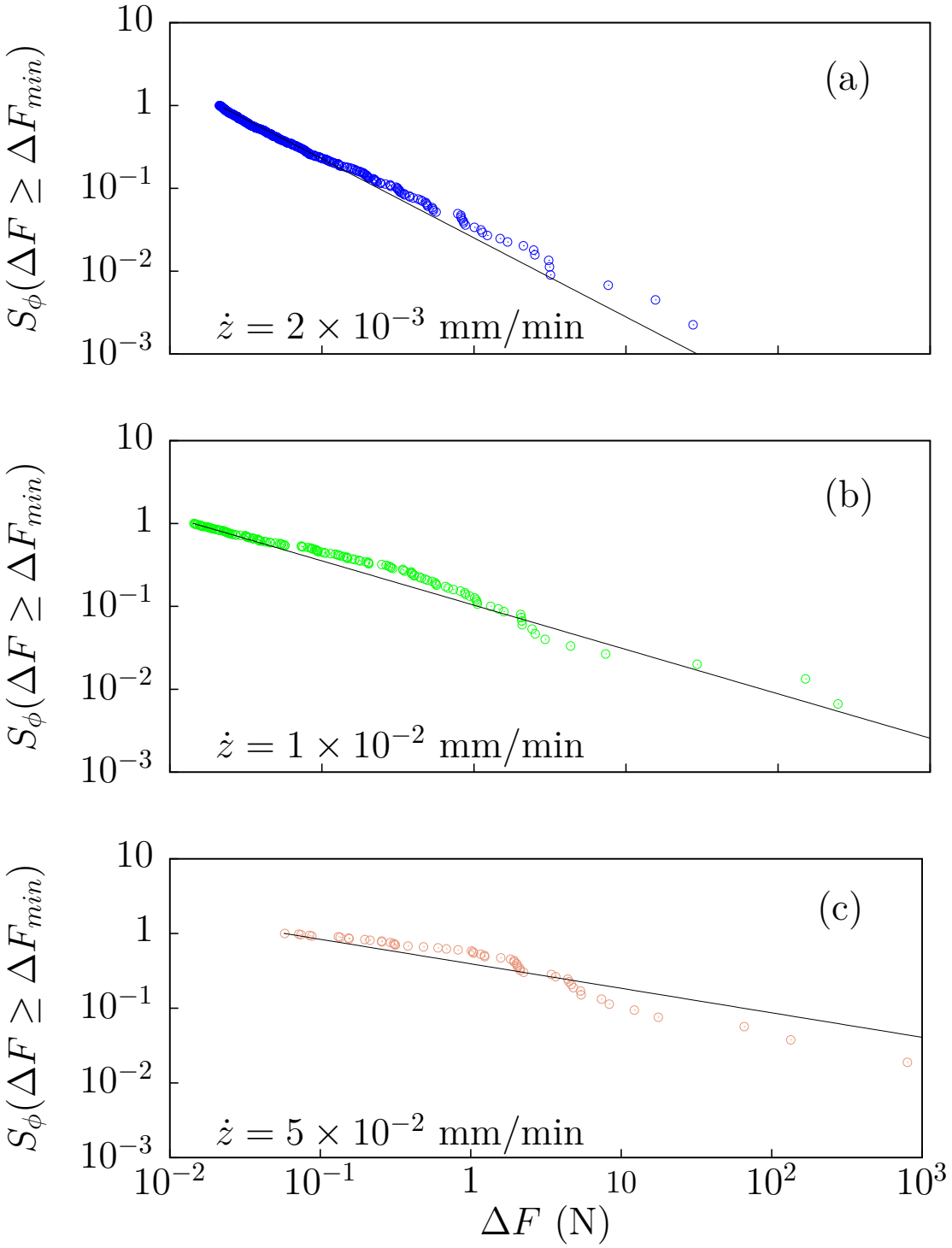

Fig. 6.6: CCDFs and corresponding power-law fits for V12 (a), V212 (b) and V24 (c). Survivor functions are displayed and normalized for $\Delta F \geq \Delta F_{min}$.

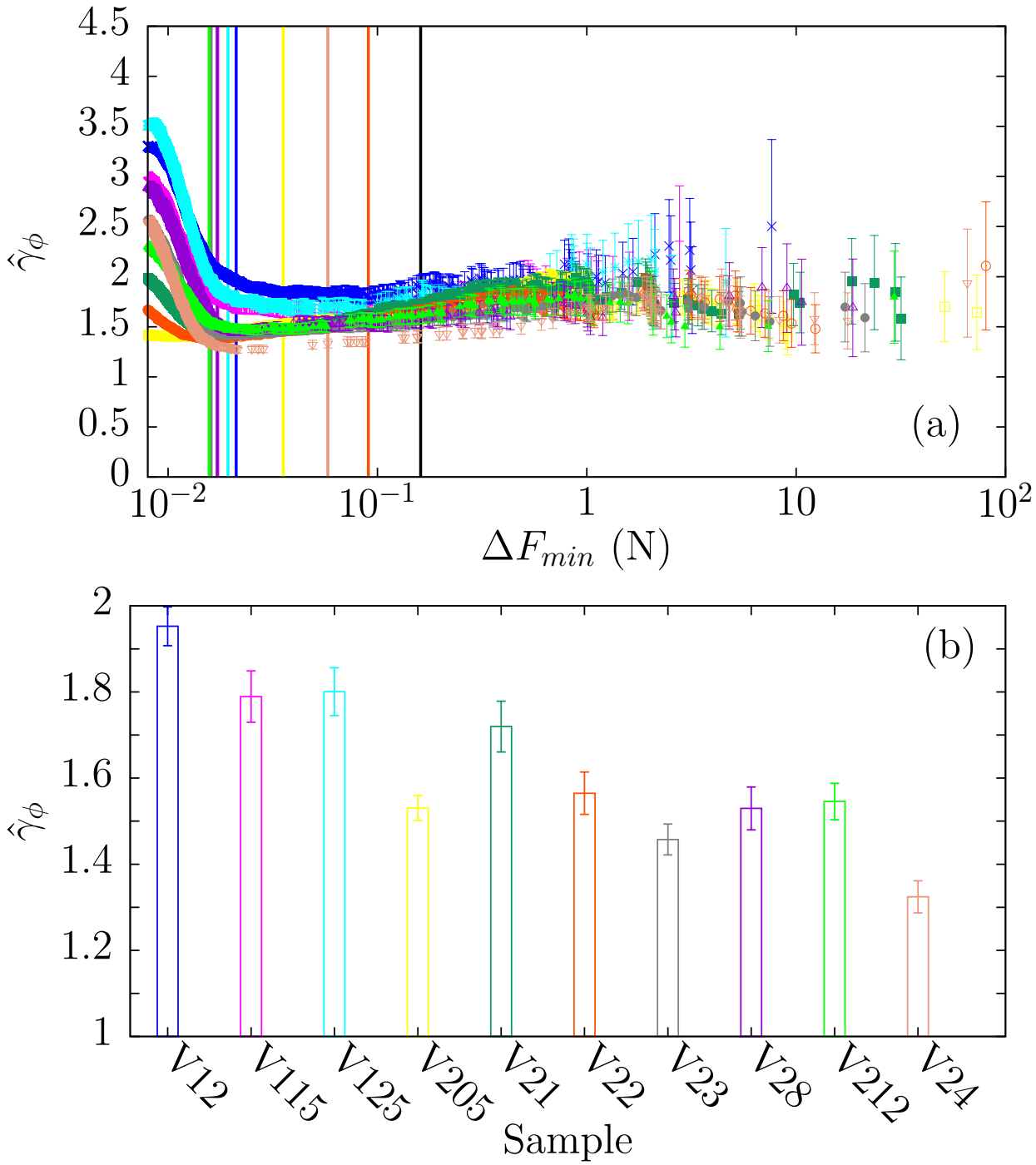

Fig. 6.7: (a) MLE of the exponent $\hat{\gamma}_\phi$ as a function of the lower threshold ΔF_{min} for all the samples. Vertical lines in each panel mark the threshold ΔF_{min} which is selected by the fitting and testing procedure. (b) Values of the exponent for each sample. The color code for each sample can be read from the color bars in (b).

Sample	n_U	n_D	n_U^{AE}	n_D^{AE}	n_{AE}	n_{AE}^U	n_{AE}^D	$\frac{n_{AE}^U}{n_U^{AE}}$	$\frac{n_{AE}^D}{n_D^{AE}}$
V115	20119	9960	119	217	836	191	645	1.61	2.97
V12	47663	32323	251	820	2314	345	1969	1.37	2.40
V125	37028	26104	141	313	1097	215	882	1.52	2.82
V205	26564	5603	1028	336	4160	2093	2067	2.04	6.15
V21	32380	10787	1324	223	4170	2572	1598	1.94	7.17
V22	24388	6609	930	177	3683	2066	1617	2.22	9.14
V23	24922	9987	423	70	1275	804	471	1.90	6.73
V28	25468	8881	359	83	974	638	336	1.78	4.05
V212	27367	9030	453	135	1646	882	764	1.95	5.66
V24	8114	3742	602	41	2129	1745	384	2.90	9.37

Table 6.3: Numbers which are involved in the construction of $W^{\Delta t}$. n_U and n_D are the total number of intervals where the force has raised up or dropped. n_U^{AE} and n_D^{AE} are the number of force rises and drops with AE events. n_{AE} is the total number of AE events, n_{AE}^U and n_{AE}^D are the number of AE events associated to rises and drops of the force, respectively. The last two columns correspond to the mean number of AE events encapsulated in force rises and drops respectively.

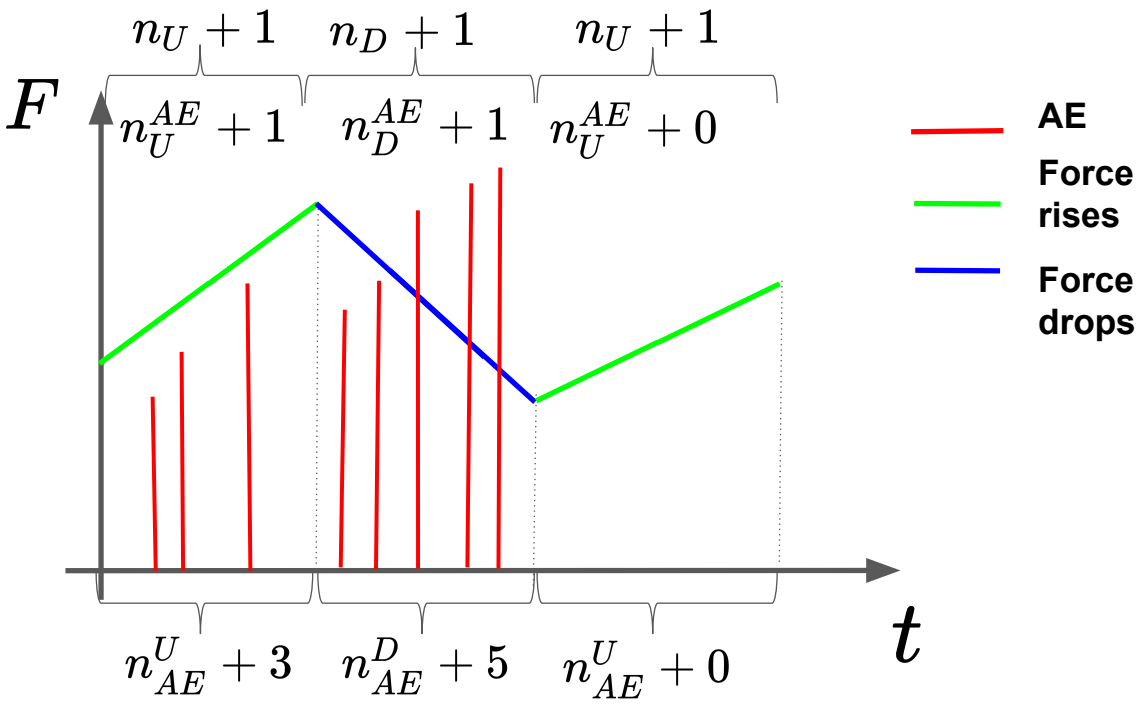

Fig. 6.8: Schematic plot with illustrating the numbers exposed in Table 6.3. Red lines correspond to AE events, green lines to force rises and blue lines to force drops.

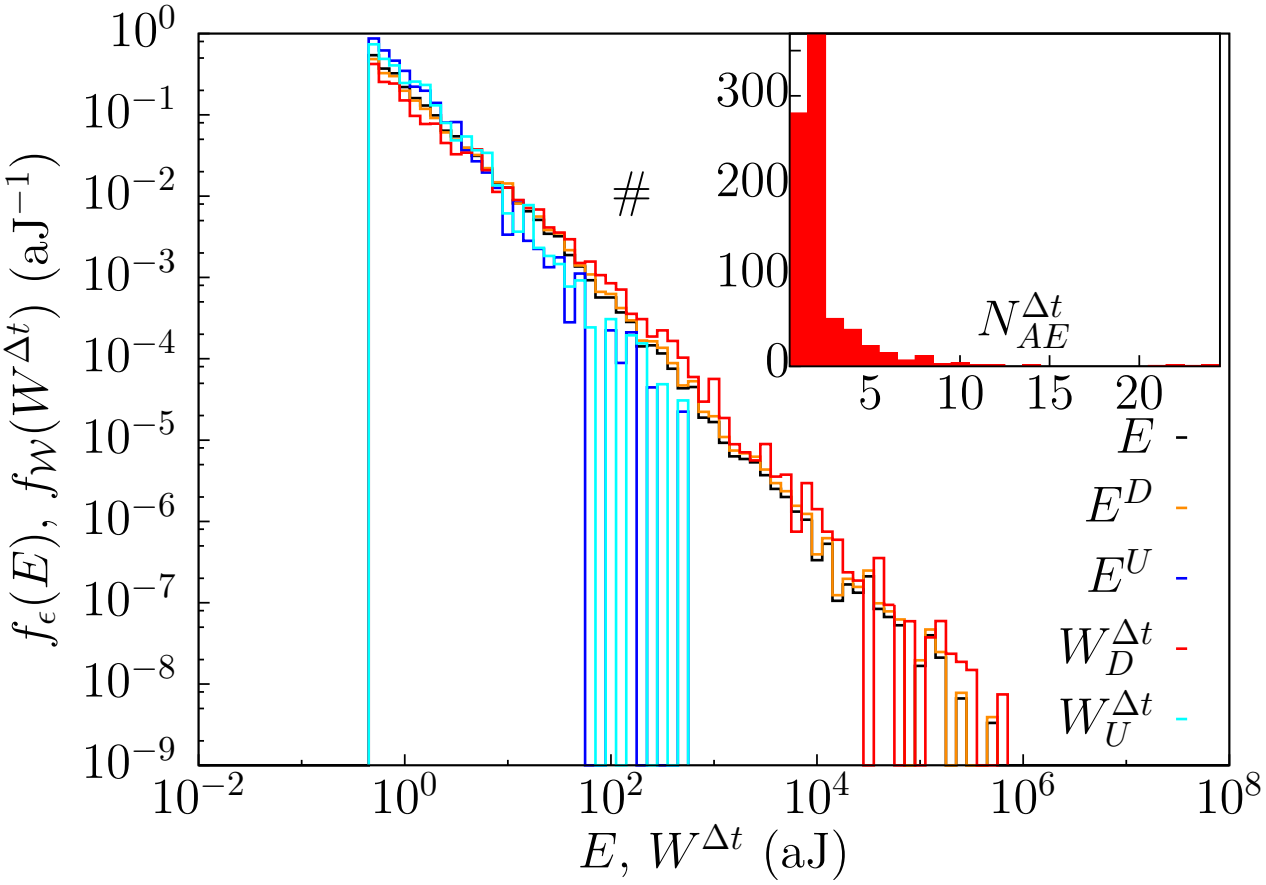

Fig. 6.9: Main panel shows the distribution of E, the distributions of energies E^D and E^U that appear when a force drop or a force rises occurs, and the distributions of $W_D^{\Delta t}$ and $W_U^{\Delta t}$, which refer to the sum of AE energies for a certain force drop or force rise. The inset represents the histogram of the number of AE events encapsulated in time intervals where force drops occur. All these distributions correspond to the sample V12.

In order to associate a certain energy to the i-th force drop, the following quantity can be defined:

$$W_{D,i}^{\Delta t} = \sum_{j=1}^{n_{AE}^i} E_j^D \,, \tag{6.2}$$

where n_{AE}^i is the number of AE events that occur within the time interval of duration $\Delta t = 0.1$ s where the i-th force drop appears, E_j^D is the energy of those AE events and D refers to "Drop". The same construction can be done for force rises by defining $W_U^{\Delta t}$ from E_j^U, where U refers to "Up". This construction is divided in two steps (see Fig. 6.8): the first one consists in splitting the time axis in intervals of duration Δt so that there is a correspondence between AE events and force rises or drops. The second step consists in applying Eq. (6.2) and its counterpart for $W_U^{\Delta t}$ for every interval with AE events. In Figure 6.9(a) the different

distributions involved in this construction for the sample V12 are presented. The inset shows the histogram of the number n_{AE} of AE events encapsulated in time intervals of Δt in which there are force drops for the sample V12. There are two random variables corresponding to the first step of the transformation: E^D corresponds to the energy of AE hits when a force drop appears whereas E^U corresponds to the energy of AE hits when force rises appear. In other words, the random variable $\mathcal{E}$ corresponding to the energy is divided in two subsets. The second step of the transformation is reflected in the quantities $W_D^{\Delta t}$ and $W_U^{\Delta t}$, which correspond to the sum of energies in every force drop and in every force rise respectively.

This kind of construction implies that the quantity $W^{\Delta t}$ (representing either $W_D^{\Delta t}$ or $W_U^{\Delta t}$) can be considered as a random variable $\mathcal{W}$ which is composed by the sum of n_{AE} independent and power-law distributed energies all of them sampled from the same power-law distribution:

$$\mathcal{W} = \mathcal{E}_1 + \mathcal{E}_2 + ... + \mathcal{E}_{n_{AE}} .$$

One could wonder whether the distribution of this new variable exhibits a power-law behaviour or not. One can try to estimate the distribution of the sum of n_{AE} power-law distributed random varialbes (energies in this case) in the Laplace space, where the convolution becomes the Laplace transform of a power-law distribution raised to the power of n_{AE}:

$$f_{\mathcal{W}}(s) = f_{\mathcal{E}}(s)^{n_{AE}} ,$$

where s is the variable in Laplace space.

Let us first try to unveil the functional expression of $f_{\mathcal{E}}(E)$ in Laplace space [Corral, 2015]:

$$\begin{aligned} f_{\mathcal{E}}(s) &= (\gamma_E - 1) E_{min}^{\gamma_E - 1} \int_{E_{min}}^{\infty} dE e^{-sE} E^{-\gamma_E} \\ &= (\gamma_E - 1) E_{min}^{\gamma_E - 1} \left[s^{\gamma_E - 1} \Gamma(1 - \gamma_E, sE_{min}) \right] . \end{aligned} \tag{6.3}$$

Note that Γ refers to the incomplete Gamma function $\Gamma(x, m) = \int_m^\infty t^{x-1} e^{-t} dt$. At this point, it is useful to consider the following equality (Eq. (6.5.29) in Abramowitz and Stegun [1965]):

$$\begin{aligned} &(sE_{min})^{\gamma_E - 1} \Gamma(1 - \gamma_E, sE_{min}) = \\ &= (sE_{min})^{\gamma_E - 1} \Gamma(1 - \gamma_E) - \sum_{k=0}^{\infty} \frac{(-sE_{min})^k}{(k + 1 - \gamma_E) k!} , \end{aligned}$$

where Γ refers to the Gamma function $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. The sum appearing in the last term can be approximated by:

$$\begin{aligned} &(sE_{min})^{\gamma_E - 1} \Gamma(1 - \gamma_E, sE_{min}) = \\ &= \frac{1}{\gamma_E - 1} \left[\tilde{C} s^{\gamma_E - 1} + \left(1 + E_{min} \frac{1 - \gamma_E}{2 - \gamma_E} s \right) + h.o.t(s) \right] , \end{aligned}$$

where $\tilde{C} = (\gamma_E - 1)\,\Gamma(1-\gamma_E)\,(E_{min})^{\gamma_E - 1}$ has been defined in order to group all the constants and simplify the expression. Note that this expression is valid as long as $\gamma_E \neq 1, 2$. If one is interested in the asymptotic behaviour at the tail of the distribution of the original variable $\mathcal{E}$, the interesting regime in Laplace space can be obtained by taking the limit $s \to 0$. Note that up to this point, the original distribution was a power-law distribution but, in this asymptotic analysis, the result can be extended to any distribution with a power-law tail. In this context, by taking the logarithms and expanding $\log(1+x) \simeq x$, one obtains:

$$\log\left[(sE_{min})^{\gamma_E - 1}\,\Gamma(1-\gamma_E, sE_{min})\right] = \\ -\log(\gamma_E - 1) + \tilde{C}s^{\gamma_E - 1} + E_{min}\frac{1-\gamma_E}{2-\gamma_E}s + h.o.t(s)\,.$$

If this expression is substituted into Eq. (6.3):

$$f_{\mathcal{E}}(s) = \exp\left(\tilde{C}s^{\gamma_E - 1} + E_{min}\frac{1-\gamma_E}{2-\gamma_E}s + h.o.t(s)\right)\,. \tag{6.4}$$

If $1 < \gamma_E < 2$, which is exactly the situation under study, the first term in the exponential is the one determining the functional behaviour:

$$f_{\mathcal{E}}(s) \simeq \exp\left(\tilde{C}s^{\gamma_E - 1}\right) \simeq 1 + \tilde{C}s^{\gamma_E - 1}\,. \tag{6.5}$$

Hence, the Laplace transform of the distribution of the variable $\mathcal{W}$:

$$f_{\mathcal{W}}(s) = f_{\mathcal{E}}(s)^{n_{AE}} \simeq \left(1 + \tilde{C}s^{\gamma_E - 1}\right)^{n_{AE}} \simeq 1 + n_{AE}\tilde{C}s^{\gamma_E - 1} \tag{6.6}$$

This last expression confirms that the Laplace transform at the tail of the distribution of the new variable $\mathcal{W}$ follows the same functional behaviour as the Laplace transform of the former variable $\mathcal{E}$. This means that the power-law exponent of the energy distribution is not altered when studying the tail of the distribution of the variable $\mathcal{W}$, which is indeed a sum of n_{AE} power-law distributed random variables. Note that n_{AE} is not fixed but a random variable. However, mixing a finite number of power-law distributed random variables will give a power-law tail with the same exponent. In other words, the probability distribution of the sum of a random number of power-law distributed random variables exhibits a power-law tail with a similar exponent as the original one.

In order to have numerical evidences of this mathematical reasoning, Fig. 6.10 shows the PDF corresponding to $n_W = 10^4$ values of the variable $W^{\Delta t}$ constructed from $n = 23983$ power-law simulated energies with exponent $\gamma_E = 1.35$ and $E_{min} = 1$. In this particular case, the number n is due to the choice of n_{AE}, which is sampled from the empirical distribution shown in the inset of Fig. 6.9. As it can be observed, the tail of both PDFs is power-law with the same exponent $\hat{\gamma}_W = \gamma_E$.

Given that the power-law exponent is kept in this construction, the key question is whether the range of values that the variable $\mathcal{W}$ can take is the same for force drops or force rises.

The main plot in Fig. 6.9 reinforces the importance of the relation between force drops and AE events since the distributions of E^U and $W_U^{\Delta t}$ are restricted to low values of the energy whereas the range of the distributions of E^D and $W_D^{\Delta t}$ is very similar to the original one. The maximum of the histogram in the inset is at $n_{AE} = 2$ and decreases up to the maximum encapsulation of $n_{AE} = 24$. The numbers involved in these constructions are shown for all the samples in Table 6.3. The fact that there are force rises associated to AE activity could be explained by the presence of force drops that have not been identified in a Δt interval where the force has globally increased. This prediction agrees with the assumption that the energy associated to force rises covers a small range corresponding to low energy values of the total energy distribution (see Fig. 6.9). It is important to remark that, despite the fraction of AE events associated to force drops decreases as the compression rate increases, the fraction that accounts for the average number of events encapsulated in a force drop (n_{AE}^D/n_D^{AE}) is

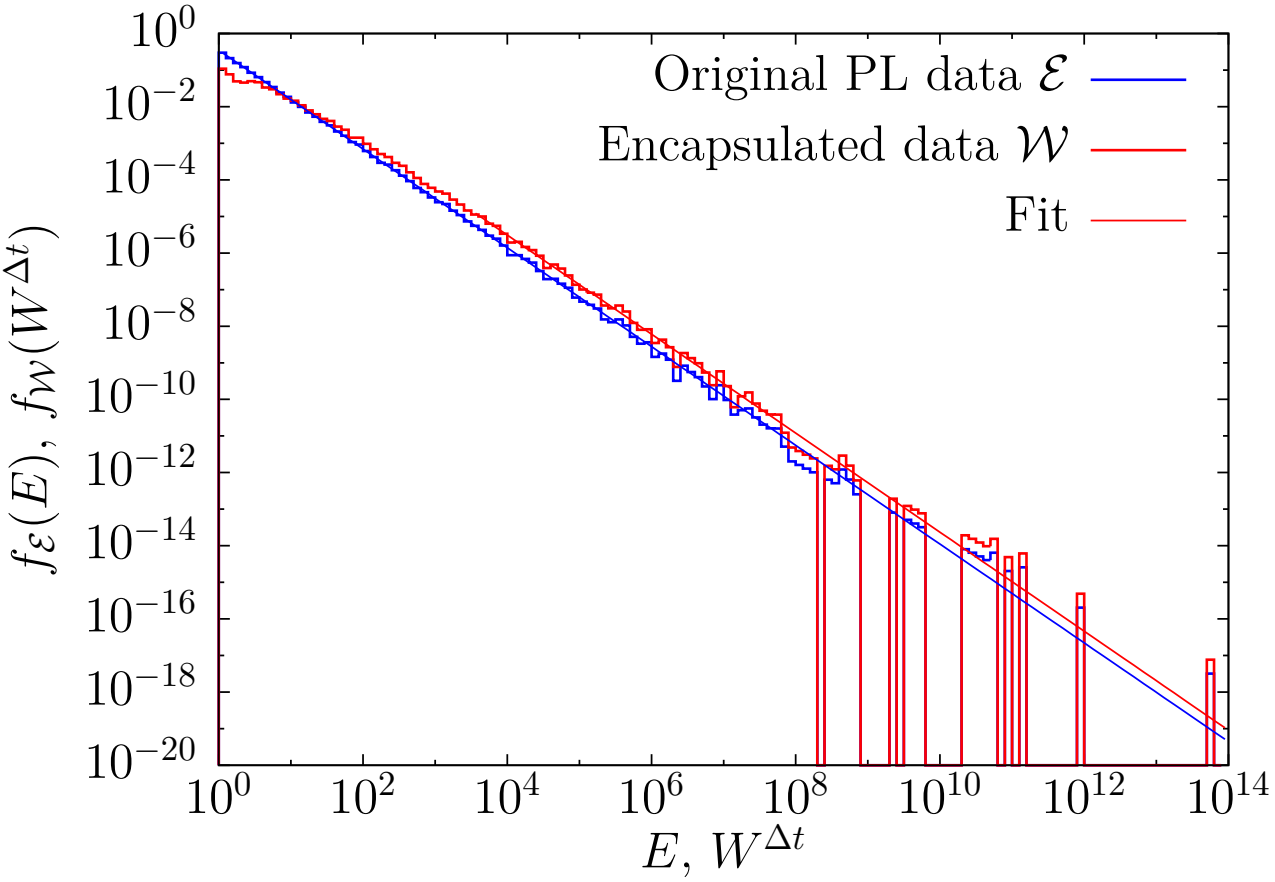

Fig. 6.10: In blue, PDF corresponding to $n = 23983$ power-law simulated energies with exponent $\gamma_E = 1.35$ and $E_{min} = 1$ together with the theoretical curve (in blue). In red, PDF of $n_W = 10^4$ values of the variable $W^{\Delta t}$ constructed by assuming that n_{AE} is sampled from the empirical distribution shown in the inset of Fig. 6.9. Red-lined fit presents the following parameters $\hat{\gamma}_W = 1.353$, $W_{min} = 2818.38$, $n = 1369$.

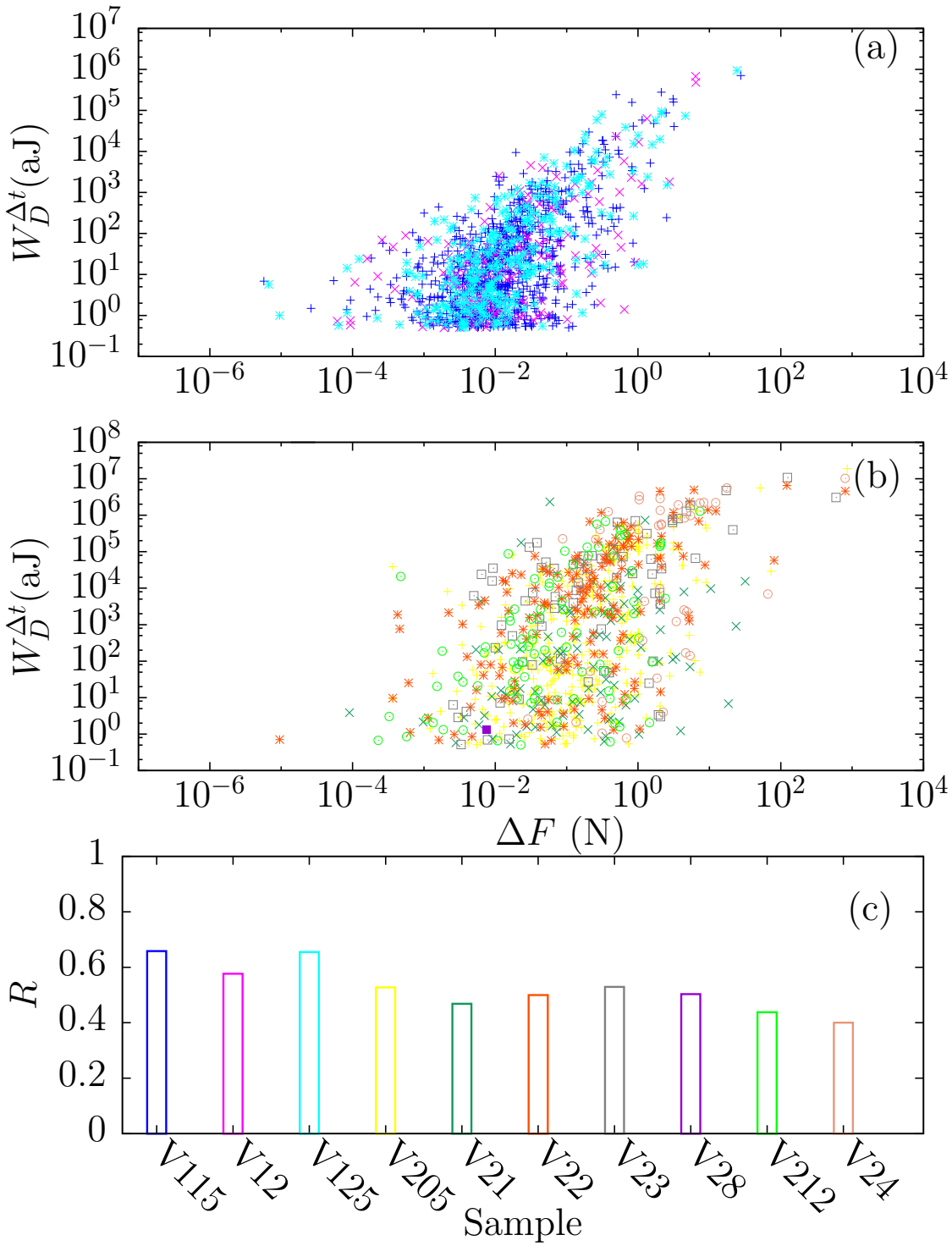

Fig. 6.11: Scatter plots of the energy released in each force drop versus the value of the force drop for all the samples compressed at $\dot{z} = 2 \times 10^{-3}$ mm/min in (a), at $\dot{z} = 10^{-2}$ mm/min and $\dot{z} = 5 \times 10^{-2}$ mm/min in (b). Panel in (c) shows the Pearson correlation coefficient for the logarithm of the variables for all the samples. Colour code provided for each sample is the same as in panel (c).

always larger than the average number of AE events encapsulated in intervals where the force is increasing (n_{AE}^{U}/n_{U}^{AE}) (See the last two columns in Table 6.3). Hence, increments of AE activity are essentially associated to drops in the force. The total duration of the experiment is given by $T = (n_U + n_D)\,\Delta t$ [2]. Note that, despite the big difference between the total number of force drops n_D and the number of force drops with AE activity n_D^{AE} (see Table 6.3), this second number is in the same order of magnitude as the number of power-law data in Table 6.2. This could indicate that power-law distributed force drops are those containing AE events.

In Figure 6.11 scatter plots for the different compression rates are presented. It must be noticed that the largest cumulative AE events are appreciated in those force drops which are power-law distributed ($\Delta F > 10^{-2}$ N, see values of ΔF_{min} in Table 6.2). The associated energy of the remaining force drops is relatively low compared with those with large values of ΔF.

Under these circumstances, the energy associated to force drops is studied with the aim of unveiling whether there exists any correlation between the magnitude of force drops and their corresponding AE energies. It must be mentioned that, as it has been already exposed in Sec. 4.1.4, the range of interest of force drops is restricted to those values which exceed 10^{-2} N. Below this value, the load cell exhibits noise and the values of the force are not reliable.

In Fig. 6.11(c) the Pearson correlation coefficient [Ross, 1998] of the logarithm of the variables $W_D^{\Delta t}$ and ΔF shown for each sample ($\Delta F > 0$). By considering the null hypothesis that the Pearson correlation coefficient is null: $\mathrm{H}_0 : R = 0$, one can perform a permutational test in order to state whether the correlations are significant (rejection of H_0). The Pearson correlation coefficient is computed for the reshuffled data R^* in order to compare it with the empirical value R. By repeating this procedure many times, a p-value for the test can be found [Good, 2011; Chave, 2017]. Table 6.4 shows the values of the Pearson correlation coefficient together with the p-values of the test. All the p-values are below a significance level of 0.05. Therefore, it can be concluded that the correlations can be considered to be statistically significant. The correlation is positive and it establishes a relation between AE events, which are of microscopic nature, with a magnitude of macroscopic character, the force drops.

[2] This equality is true for all the experiments presented in this work except for the sample V12, where three time intervals of $\Delta t = 0.1$ s have been found with a null value of Δ and no associated AE.

	V115	V12	V125	V205	V21	V22	V23	V28	V212	V24
R	0.66(10)	0.56(5)	0.66(8)	0.53(6)	0.47 (11)	0.50(8)	0.53(13)	0.48(15)	0.44(12)	0.40(16)
p_{value}	0	0	0	0	0.0168(4)	0	$1(1) \times 10^{-5}$	$4(2) \times 10^{-5}$	0.0028(2)	0.0101(3)

Table 6.4: Values of the Pearson correlation coefficient for the logarithm of the variables and the p-values of the permutational test for all the samples. p-values are computed with 10^5 realizations.

6.3 Energy distribution of charcoal labquakes

The aim of this section is to present the difference in the exponent of the energy distribution of AE events when labquakes are generated during the compression of a disordered material whose inner structure is substantially different from the one of Vycor. Whereas Vycor is a synthetic SiO_2-based material with an homogeneous structure composed by pores at micro and nanoscale, charcoal has been created through different natural processes and presents more heterogeneities. These differences in the internal structure might involve a different failure process [Xu et al., 2019]. The mean features characterizing the failure processes can be qualitatively appreciated in Fig. 6.12, where the force curves are presented together with the AE activity. Note that the activity is presented in logarithmic scale.

For both materials, the force curve starts with an adaptation with the compression plate (which is not relevant for the study due to the poor AE activity) and evolves in an elastic regime where the force increases monotonously. See Sec. 3.1 where the different regimes that a material may experience when it is under stress are presented. As it was exposed in Sec. 6.2, this elastic regime can be suddenly interrupted by small micro-fractures that might nucleate in the sample and are manifested by the presence of force drops and AE activity. However, the material is able to recover and the force can acquire higher values than the ones which previously caused the micro-fracture. For the case of Vycor, these small fractures do not alter the elastic behaviour and the force keeps monotonously increasing with a similar slope as before. This effect is known as hardening [Xu et al., 2019]. For the case of charcoal, the elastic regime tends to exhibit a flat or decreasing trajectory after a force drop (see Fig. 6.12 and Figs. 2 and 9 in Xu et al. [2019]). This kind of behaviour is very common for materials exhibiting plasticity and is known as weakening [Xu et al., 2019]. In the case of Vycor, the elastic regime is suddenly interrupted when the force reaches a maximum value and the curve exhibits a large force drop where the sample gets destroyed. Nevertheless, the force curve for charcoal does not exhibit such an extreme event. Although there are large force drops, these are not so sharp, and the sample gets destroyed by a continuously weakening process manifested in a sequence of smaller stress drops with a constant presence of AE activity. These qualitative differences indicate a different mechanical behaviour that might affect the statistics of labquakes for these materials [Xu et al., 2019]. In order to check whether there is difference in the statistics or not, the distribution of the energy of charcoal labquakes is studied by considering an untruncated power-law distribution.

N	n	E_{min} **(aJ)**	E_{top} **(aJ)**	$\hat{\gamma}_E$	p_{value}
101524	18448	5.012	1.86×10^7	1.657(5)	0.088(9)

Table 6.5: Fitting parameters for the enegy distribution of charcoal labquakes. An untruncated power-law with exponent $\hat{\gamma}_E$ with lower cut-off E_{min} gives p-value larger than $p_c = 0.05$. Numbers in parenthesis correspond to the standard deviations of the MLE and the p-value. N is the total number of data and n is the number of points entering into the fit.

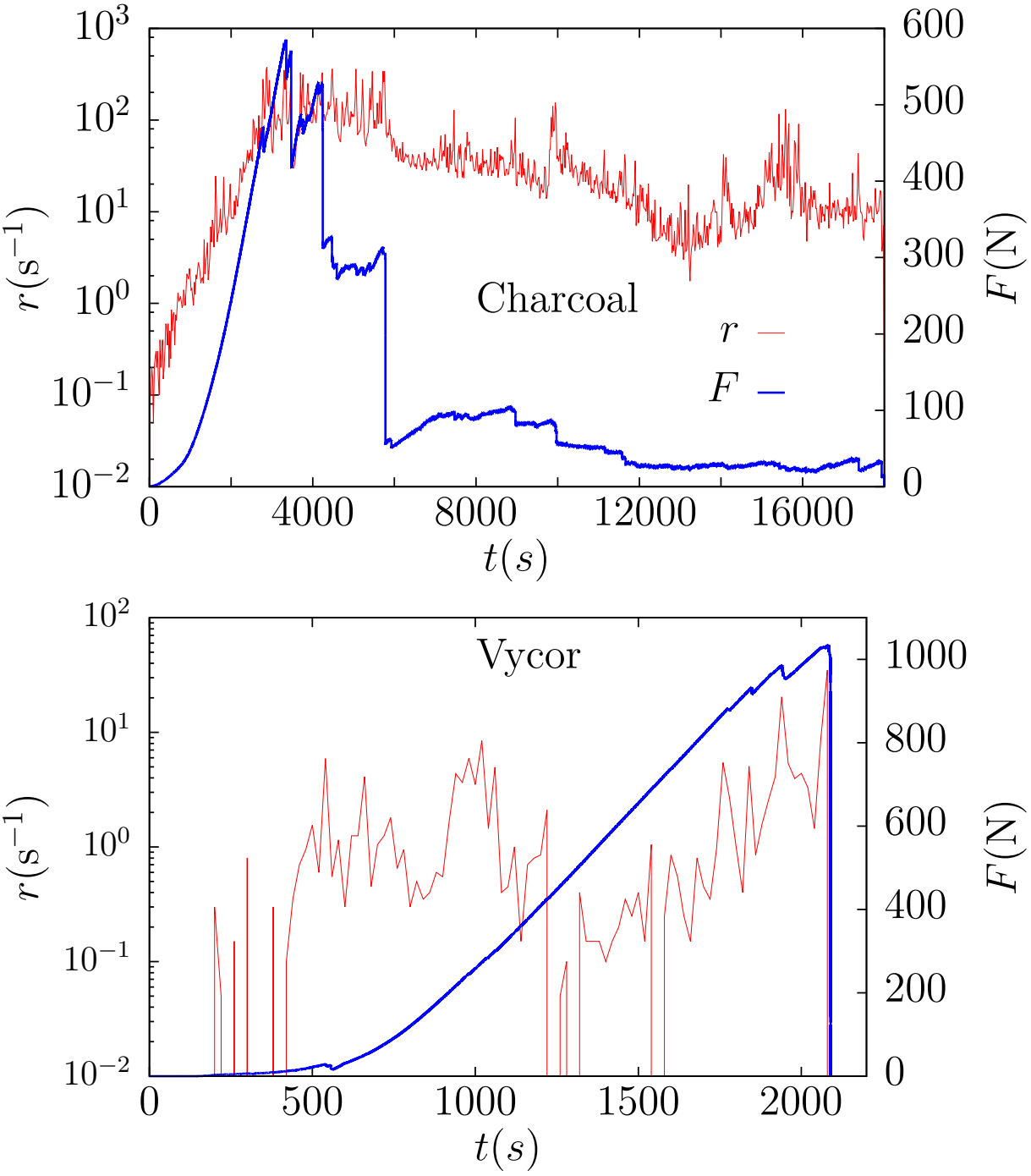

Fig. 6.12: Evolution of the AE activity (in red) and the force (in blue) as a function of the time for experiments performed with two different materials: charcoal (a) and Vycor (V22 sample in (b)).

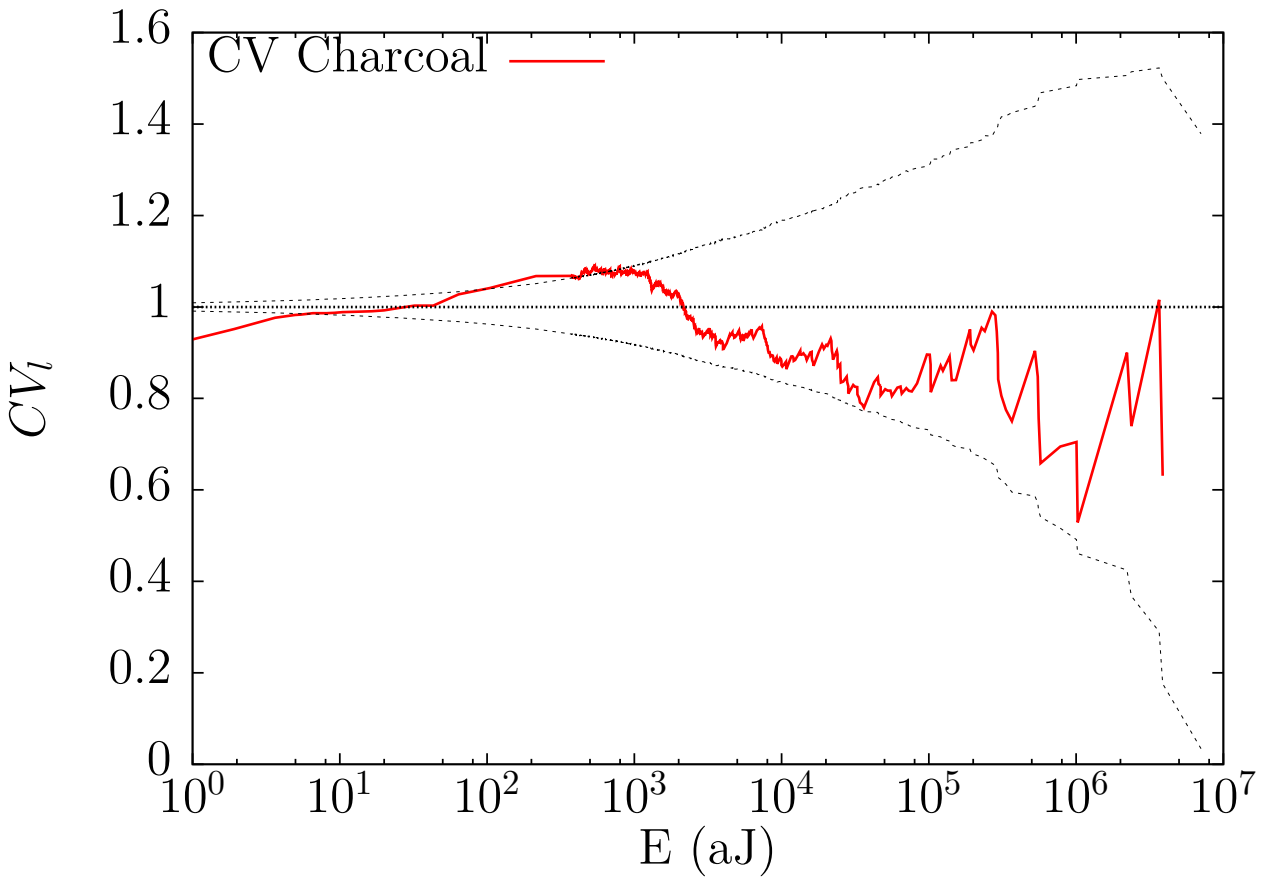

Fig. 6.13: In red, residual CV-plot for the logarithm of the energy of the charcoal sample. Confidence intervals of 95% are represented by black-dashed lines as well as an horizontal dashed-black line placed at the value 1.

This is supported by the CV plot shown in Fig. 6.13, where the hypothesis of a unique power-law tail cannot be rejected for $E > 10$ aJ. Therefore, there are not clear evidences of deviations from the power-law behaviour along all the distribution and no upper cut-offs are needed. Let us consider that the energy of charcoal labquakes is a random variable $\mathcal{E}$ which is distributed according to an untruncated power-law (see Eq. (6.1)) whose fitting parameters are computed according to the procedure explained in Sec. 5.5 and are shown in Table 6.5. A power-law exponent of $\gamma_E = 1.656 \pm 0.005$ holds for six orders of magnitude in energy (see Fig. 6.14).

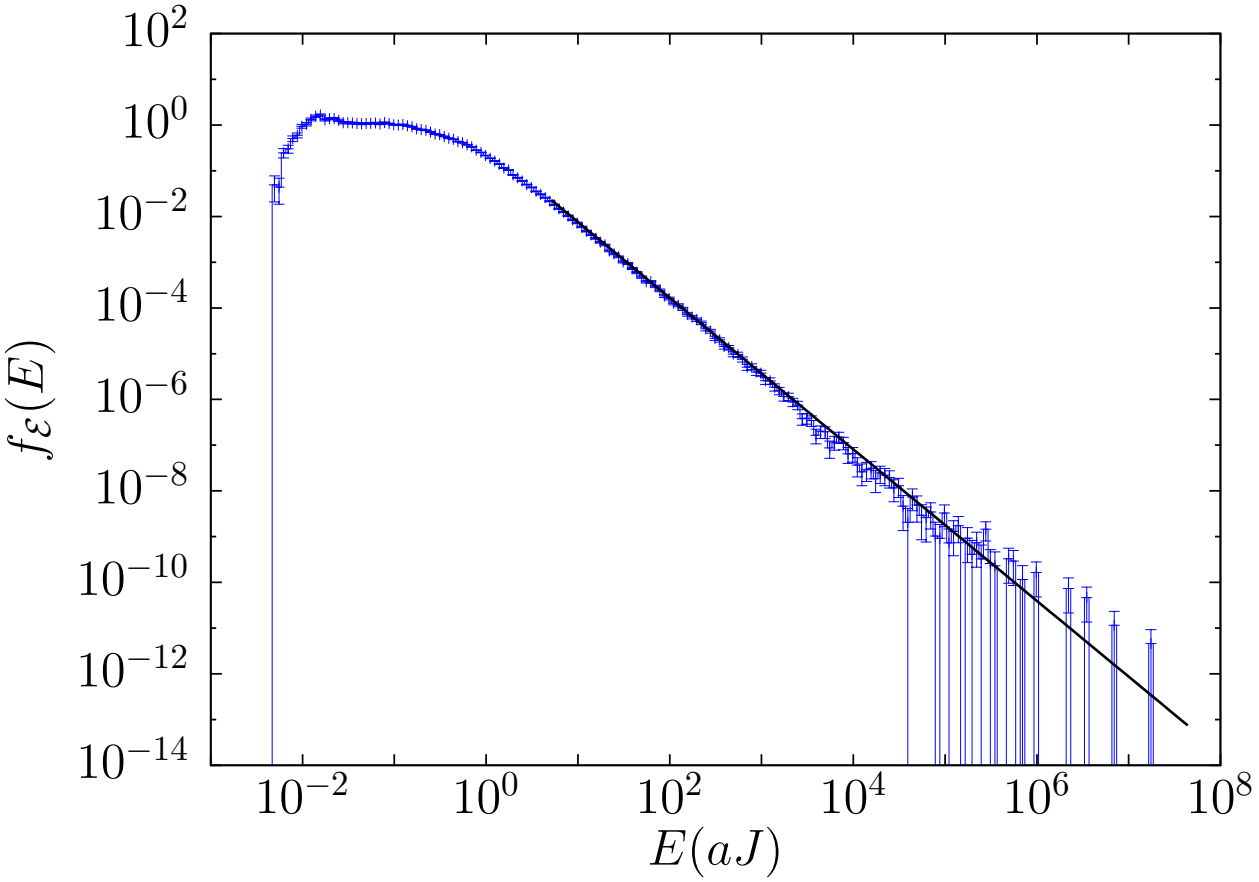

Fig. 6.14: In blue, estimated PDF of the energy E of charcoal labquakes. In black, fitted untruncated power-law distribution with the fitting parameters of Table 6.5.

As it can be appreciated in Fig. 6.15, the value of the exponent $\hat{\gamma}_E$ as a function of the cut-off E_{min} is always within 2σ. This exponent is clearly different from the ones found for Vycor samples. If one considers the results for Vycor exposed in Sec. 6.1 ($\mu_{\gamma_E} = 1.35 \pm 0.05$) and computes the z-statistic ($z = 3.108$) exposed in Sec. 5.9, the test rejects the compatibility of the exponents with a $p_{norm} = 0.002$. Therefore, the inner structure of the material can be a factor that affects not only the mechanical behaviour but also the statistics of labquakes which are generated during the compression experiment.

At this point it is important to stress the fact that, given the incompatibility of these power-law exponents, charcoal and Vycor labquakes cannot be considered into the same universality class.

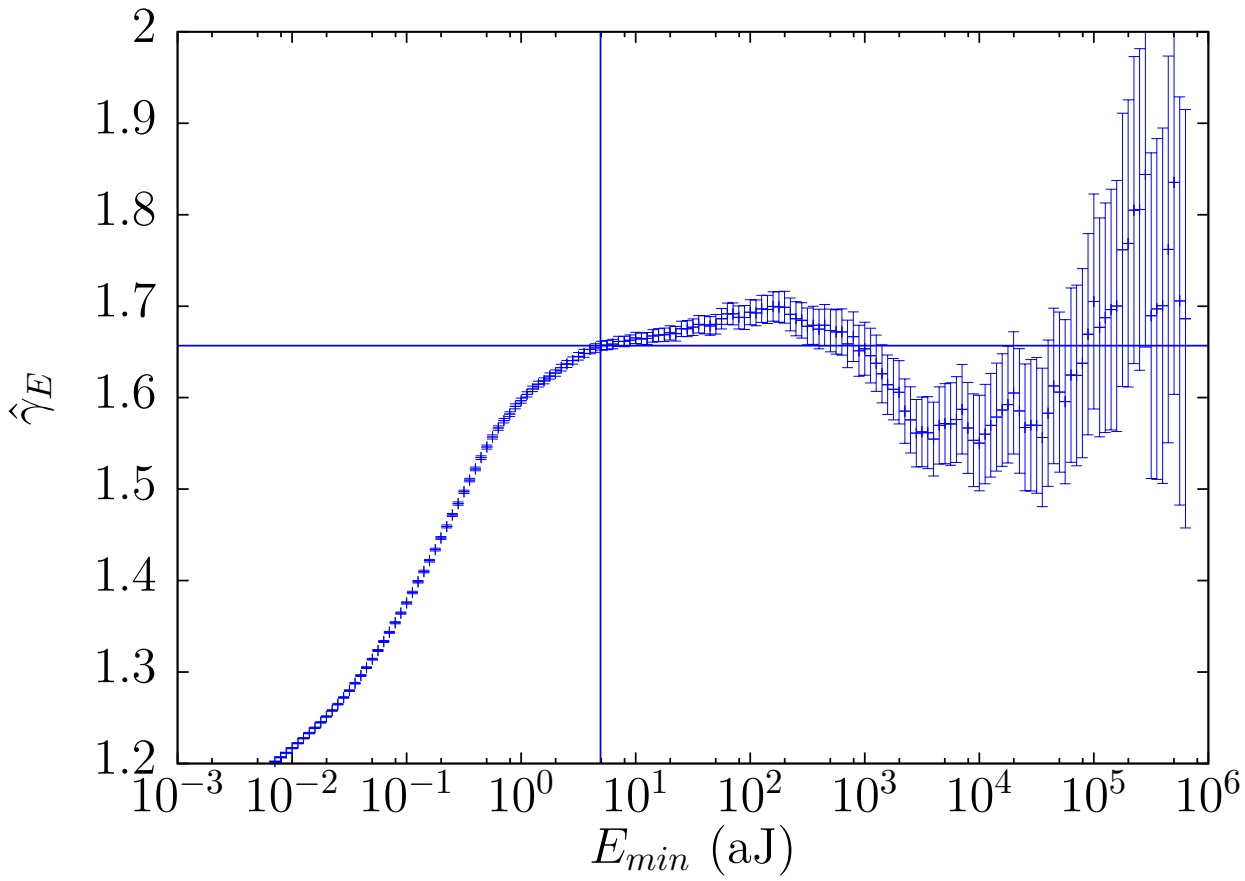

Fig. 6.15: Evolution of the value of the fitted exponent $\hat{\gamma}_E$ as a function of the cut-off E_{min}. The lines indicate the fit parameters E_{min} and γ_E exposed in Table 6.5.

Chapter 7

Universality of power-law exponents by means of maximum likelihood estimation

A statistical procedure to merge different datasets is first presented in this chapter with two different separate aims. First, obtaining a broader fitting range for the statistics of different experiments or observations of the same system. Second, establishing if two or more different systems may belong to the same universality class. By means of maximum likelihood estimation and goodness-of-fit test, this methodology provides rigorous statistical information to discern whether power-law exponents characterizing different datasets can be considered equal among them or not. Afterwards, some novel results found in the statistical study of labquakes are presented. Vycor labquake datasets are merged according to the methodology presented in Sec. 7.1, finding global energy and amplitude distributions ranging for more than nine and five orders of magnitude respectively. The methodology can also elucidate that, at least in statistical terms, labquakes generated from the compression of the mesoporous silica glass Vycor are different from those generated by the compression of charcoal. The methodology is also applied for merging earthquake catalogs finding a Gutenberg-Richter law with global exponent $\Gamma \simeq 1.67$ ($b \simeq 1$) holding for more than eight orders of magnitude. Finally, earthquake and labquake datasets are merged in order to check whether these phenomena are candidates to be classified into the same universality class or not. The results exposed in this chapter have been published in:

- Víctor Navas-Portella, Isabel Serra, Álvaro Corral, and Eduard Vives, Increasing power-law range in avalanche amplitude and energy distributions, Physical Review E 97, 022134 (2018) [Navas-Portella et al., 2018].
- Víctor Navas-Portella, Álvaro González, Isabel Serra, Álvaro Corral, and Eduard Vives, Universality of power-law exponents by means of maximum likelihood estimation, Physical Review E 100, 062106 (2019) [Navas-Portella et al., 2019].

7.1 Merging catalogues/datasets method

By considering $n_{\rm ds}$ datasets of N_i ($i = 1, ..., n_{\rm ds}$) observations each, one wants to fit a general power-law distribution with a unique global exponent for all of them. Let us assume that for the i-th dataset, the variable $\mathcal{X}$ (seismic moment if one works with the GR law for earthquakes or AE energy or amplitude if one works with labquakes) follows a power-law PDF $f_{\mathcal{X}}(x; \gamma_i, x_{min}^{(i)}, x_{max}^{(i)})$ given by Eq. (5.25) from a certain lower cut-off $x_{min}^{(i)}$ to an upper cutt-off $x_{max}^{(i)}$ with exponent γ_i and number of data n_i ($n_i \leq N_i$) in the range $\left[x_{min}^{(i)}, x_{max}^{(i)}\right]$. Note that one also can consider the untruncated power-law model for the i-th dataset if $x_{max}^{(i)} \to \infty$. By means of the methods explained in Sec. 5.5 (see also Deluca and Corral [2013]; Corral and González [2019], o alternatively, Clauset et al. [2009]), one can state that data from the i-th dataset does not lead to the rejection of the the power-law hypothesis for a certain range. Note that, in the i-th dataset, the variable $\mathcal{X}$ can acquire values in a range typically spanning several orders of magnitude.

Generally, the procedure of merging datasets is performed by selecting upper and lower cut-offs $x_{min}^{(i)}$ and $x_{max}^{(i)}$ ($x_{min}^{(i)} < x_{max}^{(i)}$) for each dataset. Data outside these ranges are not considered. All the possible combinations of cut-offs $\{x_{min}\}$ and $\{x_{max}\}$ are checked with a fixed resolution (see below). By assuming that a different power-law regime can be found at the tail of the distribution, the CV test can be used to fix some upper cut-offs, thus reducing the computational effort. For more details about the CV test, see Sec. 5.10.

Given a set of cut-offs, datasets can be merged by considering two different models:

- **Model One Exponent (OneExp)**: All datasets are merged by considering a unique global exponent Γ ($\gamma_i = \Gamma$ for all datasets).
- **Model Multi Exponent (MultiExp)**: All datasets are merged, but each one with its own exponent γ_i ($i = 1, ..., n_{\rm ds}$).

Note that model OneExp is nested in model MultiExp and the difference in the number of parameters characterizing these models is $n_{\mathcal{L}}^{(MultiExp)} - n_{\mathcal{L}}^{(OneExp)} = n_{ds} - 1$. Since one is interested in merging datasets with a unique global exponent (model OneExp), no statistical evidence that this MultiExp model is more suitable to fit the data should be found. Note that model OneExp can be a good model to fit if the particular exponents $\hat{\gamma}_i$ do not exhibit large differences among each other in relation to their uncertainty. For a given set of cut-offs, the fit is performed by means of the following protocol:

1. **Maximum Likelihood Estimation (MLE) of model OneExp**: The log-likelihood function of model OneExp can be written as:

$$\log \mathcal{L}_{OneExp} = \sum_{i=1}^{n_{\rm ds}} \sum_{j=1}^{n_i} \log f_{\mathcal{X}} \left(x_{ij}; \Gamma, x_{min}^{(i)}, x_{max}^{(i)} \right) , \tag{7.1}$$

where x_{ij} corresponds to the n_i values of the variable $\mathcal{X}$ that are in the range $x_{min}^{(i)} \leq x_{ij} \leq x_{max}^{(i)}$ in the i-th dataset, log is the natural logarithm and Γ is the global exponent. The definition of likelihood is consistent with the fact that likelihoods from different datasets can be combined in this way [Pawitan, 2013][p.27]. At this step, one has to find the value $\hat{\Gamma}$ of the global exponent Γ that maximizes the log-likelihood expression in Eq. (7.1). For the particular expressions corresponding to the truncated and untruncated power-law PDF, see Eqs. (5.25) and (5.16) in Sec. 5.2. If all the power-law distributions are untruncated, this exponent can be easily found analytically [Navas-Portella et al., 2018] as

$$\hat{\Gamma} = 1 + \frac{\sum_{i=1}^{n_{ds}} n_i}{\sum_{i=1}^{n_{ds}} \frac{n_i}{\hat{\gamma}_i - 1}} ,$$

where the hats denote the values of the exponents that maximize the log-likelihood of the particular power-law distribution (model MultiExp) and the general one in Eq. (7.1). See Sec. 7.1.1 for more details about the relationship between the global and the particular exponents. If truncated power-law distributions are considered, one has to use a numerical method in order to determine the exponent $\hat{\Gamma}$ that maximizes this expression [Baró and Vives, 2012; Deluca and Corral, 2013].

2. **MLE of model MultiExp**: The log-likelihood function of model MultiExp can be written as:

$$\log \mathcal{L}_{MultiExp} = \sum_{i=1}^{n_{\mathrm{ds}}} \sum_{j=1}^{n_i} \log f_{\mathcal{X}} \left(x_{ij}; \gamma_i, x_{min}^{(i)}, x_{max}^{(i)} \right) , \tag{7.2}$$

using the same notation as in Eq. (7.1). For the particular expressions corresponding to the truncated and untruncated power-law PDF, see Eqs. (5.25) and (5.16) in Sec. 5.2. The values of the exponents that maximize Eq. (7.2) are denoted as $\hat{\gamma}_i$.

3. **Likelihood-ratio test**: The likelihood-ratio test (LRT) for the models OneExp and MultiExp in order to check whether model OneExp is good enough to fit data or not in comparison with model MultiExp. For more details about the LRT see Sec. 5.6. If model OneExp "wins", the procedure goes to step (4). Otherwise, this fit is discarded and a different set of cut-offs $\{x_{min}\}$ and $\{x_{max}\}$ is chosen, and the procedure goes back to step (1).

4. **Goodness-of-fit test**: In order to check whether it is reasonable to consider model OneExp as a good candidate to fit data, the next null hypothesis is formulated H_0: the variable $\mathcal{X}$ is power-law distributed with the global exponent $\hat{\Gamma}$ for all the datasets. Two different statistics are used in order to carry out the goodness-of-fit tests: the Kolmogorov-Smirnov Distance of the Merged Datasets (KSDMD) and the Composite Kolmogorov-Smirnov Distance (CKSD). The KSDMD statistic can be used as long as datasets overlap each other whereas the CKSD statistic does not require this condition. For more details about how these statistics are defined and how the p-value of the test

is found, see Sec. 7.1.3. If the resulting p-value is greater than a threshold value p_c (in the present work the thresholds $p_c = 0.05$ and $p_c = 0.20$ are used) this is considered as a valid fit and it can be stated that the variable $\mathcal{X}$ is power-law distributed with exponent $\hat{\Gamma}$ along all the different datasets for the different ranges $\{x_{min}\}$ and $\{x_{max}\}$. Otherwise, this fit is not be considered as valid. A different set of cut-offs is chosen and the procedure goes back to step (1).

When all the combinations of cut-offs have been checked, one may have a list of valid fits. In order to determine which of them is considered the definitive fit, the following procedure is carried out:

1. The fit that covers the largest sum of orders of magnitude: $\max\left[\sum_{i=1}^{n_{ds}} \log_{10}\left(\frac{x_{max}^{(i)}}{x_{min}^{(i)}}\right)\right]$ is chosen. If the power-law fit is untruncated, $x_{max}^{(i)}$ can be substituted by the maximum observed value $x_{top}^{(i)}$. If there is a unique candidate with a maximum number of orders of magnitude, then this is considered as the definitive global fit. Otherwise the procedure goes to the next step.

2. The fit with the broadest global range: $\max\left[\frac{\max\left(x_{max}^{(i)}\right)}{\min\left(x_{min}^{(i)}\right)}\right]$ for $i = 1, ..., n_{ds}$ is chosen. If there is a unique candidate, this is considered as the definitive global fit. Otherwise the procedure goes to the next step.

3. The fit with the maximum number of data $\mathcal{N} = \sum_{i=1}^{n_{ds}} n_i$ is considered as the definitive global fit.

By means of these three steps, a unique fit has been found for all the datasets analysed in this work. Nevertheless, one could deal with datasets in which more conditions are needed in order to choose a definitive fit unambiguously. At the end of this procedure, if a solution is found, one is able to state that the datasets that conform the global fit correspond to phenomena that are candidates to be be classified into the same universality class, at least regarding the observable $\mathcal{X}$. If no combination of cut-offs is found to give a good fit, then it can be said that there exists at least one catalog that corresponds to a phenomenon that must be classified in a different universality class.

7.1.1 Global exponent and particular exponents

Let us consider $n_{\rm ds}$ catalogs with n_i events characterized by a variable $\mathcal{X}$. In order to simplify and have exact calculations, let us consider also that data in the i-th catalog follows an untruncated power-law distribution with exponent γ_i from a lower cut-off $x_{min}^{(i)}$:

$$f_{\mathcal{X}}(x;\gamma_i,x_{min}^{(i)}) = \frac{\gamma_i - 1}{x_{min}^{(i)1-\gamma_i}} x^{-\gamma_i} \,. \tag{7.3}$$

By finding the values of the particular exponents that maximize the log-likelihood expression in Eq. (5.27) one obtains [Deluca and Corral, 2013; Clauset et al., 2009]:

$$\hat{\gamma}_i = 1 + \frac{n_i}{\sum_{j=1}^{n_i} \log\left(x_j/x_{min}^{(i)}\right)} \,. \tag{7.4}$$

Let us consider that data in these $n_{\rm ds}$ catalogs follow PDFs with different cut-offs $x_{min}^{(i)}$ but sharing a global exponent Γ. If these PDFs are brought to the global log-likelihood fo Eq. (7.1), one obtains:

$$\log \mathcal{L} = \sum_{i=1}^{n_{\rm ds}} n_i \log\left(\Gamma - 1\right) + \sum_{i=1}^{n_{\rm ds}} n_i \left(\Gamma - 1\right) \log x_{min}^{(i)} - \Gamma \sum_{i=1}^{n_{\rm ds}} \sum_{j=1}^{n_i} \log x_{ij} \,. \tag{7.5}$$

By deriving this expression with respect Γ, the value of the global exponent that maximizes the log-likelihood is:

$$\hat{\Gamma} = 1 + \frac{\mathcal{N}}{\sum_{i=1}^{\rm n_{ds}} \sum_{j=1}^{n_i} \log\left(x_{ij}/x_{min}^{(i)}\right)} \,, \tag{7.6}$$

where $\mathcal{N} = \sum_{i=1}^{\rm n_{ds}} n_i$. This global exponent can be related with the particular exponents from Eq. (7.4) leading to:

$$\frac{\mathcal{N}}{\hat{\Gamma} - 1} = \sum_{i=1}^{\rm n_{ds}} \frac{n_i}{\hat{\gamma}_i - 1} \,, \tag{7.7}$$

which is indeed the same equation as the one exposed in Sec. 7.1. Hence, for the case of non-truncated power-law PDFs, the global exponent is related to the weighted harmonic mean of $\hat{\gamma} - 1$, being $\hat{\gamma}$ the exponent of the particular PDF of the i-th catalog.

7.1.2 Merging histograms: Algorithmic and numerical procedures

Histograms have been commonly used to graphically represent PDFs of empirical data. The most common situation is to have a dataset sampled from some probability distribution in a certain range and the histogram is simply built up by the usual techniques explained in Sec. 5.3. However, the situation gets more complicated if one tries to construct a single histogram by using different datasets that may span different ranges of values and different sampling "intensity" (number of data).

The problem can be addressed in the following terms: Let us consider n_{ds} samples from the same population with sizes n_i each one spanning $\left[x_{low}^{(i)}, x_{top}^{(i)}\right]$, where $i = 1, ..., n_{ds}$. According to this notation, it is important to remark that datasets are ordered according to $x_{low}^{(i)} \leq x_{low}^{(i+1)}$.

Let us define $x_{low}^{abs} = \min\left(x_{low}^{(1)}\right)$ and $x_{top}^{abs} = \max\left(x_{top}^{(i)}\right)$ and let us split the real axis by a set of w points such that $x_{[1]} = x_{low}^{abs} < x_{[2]} \leq ... < x_{top}^{abs} < x_{[w]}$. Note that the notation $x_{[j]}$ does not denote the j-th value in the sample but the value of the real axis of the variable $\mathcal{X}$ in which the j-th point $(j = 1, ..., w)$ is placed. This axis is thus divided in $n_{int} = w - 1$ consecutive semi-open intervals $\left[x_{[j]}, x_{[j+1]}\right)$ separating regions in which different datasets overlap each other or not. Overlapping intervals are those in which the intersection of two or more different datasets

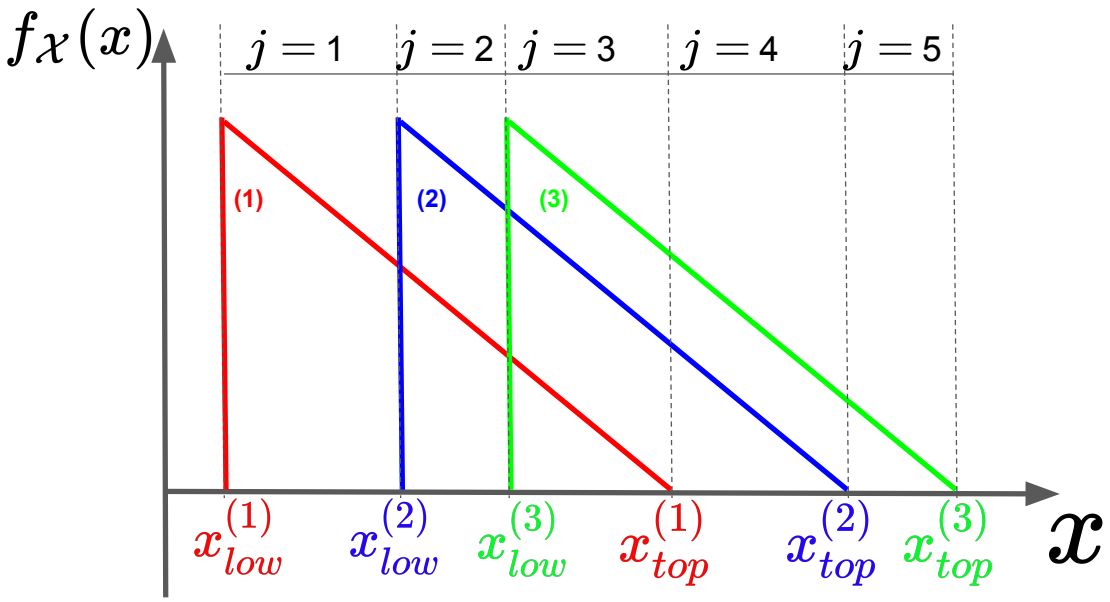

Fig. 7.1: Schematic representation of the PDFs of three different power-law datasets (1), (2) and (3) overlapping each other. According to the values of $x_{low}^{(i)}$ and $x_{top}^{(i)}$ $(i = 1, 2, 3)$ the real axis is splitted in five intervals $(j = 1, 2, 3, 4, 5.)$

is not null. Contrarily to the procedure exposed in Sec. 5.3, note that these semi-open intervals do not necessarily correspond to the bins. For instance, let us consider a simple case in which there are three samples (1), (2) and (3) for which $x_{top}^{(1)} \leq x_{top}^{(2)} \leq x_{top}^{(3)}$ (See Fig. 7.1). For this particular case, the configuration of the intervals would be:

- $j = 1$: $(1) \cap (2) = \emptyset$ and $(1) \cap (3) = \emptyset$. (Non-Overlapping interval)
- $j = 2$: $(1) \cap (2) \neq \emptyset$ and $(1) \cap (3) = \emptyset$. (Overlapping interval)
- $j = 3$: $(1) \cap (2) \cap (3) \neq \emptyset$. Overlapping interval)
- $j = 4$: $(1) \cap (2) = \emptyset$, $(1) \cap (3) = \emptyset$ and $(2) \cap (3) \neq \emptyset$. Overlapping interval)
- $j = 5$: $(1) \cap (2) = \emptyset$, $(1) \cap (3) = \emptyset$ and $(2) \cap (3) = \emptyset$. (Non-Overlapping interval)

Each particular problem will have a certain configuration of the intervals depending on the values of each dataset.

In this framework, the goal is to construct a single histogram composed by these samples. In order to deal with this problem two different approaches can be considered: the algorithmic procedure, which essentially consists in giving a certain weight to each data point, and the numerical procedure, which consists in minimizing a χ^2 statistic. The underlying philosophy of both methods is quite different. On the one hand, the algorithmic procedure considers that all the points in the sample should account differently according to the overlapping or non-overlapping region they belong to. This methodology can be used as long as datasets overlap each other. On the other hand, the numerical procedure can be thought as a non-parametric inference technique which can assign counts to regions where no observations are measured. In this sense, this method can be quite counter-intuitive when it is applied to merge datasets where there is a big difference in the scale of the observations. These methodologies are applied to power-law distributed synthetic data with the same value of the exponent but in different ranges. Both methodologies yield similar results.

7.1.2.1 Algorithmic procedure

The methodology is based on assembling recursively datasets by pairs and is independent on the shape of the histogram. Contrarily to the usual procedure for constructing histograms explained in Sec. 5.3, instead of considering one element as a count (as in Fig. 7.2 (a) and (c)) each individual contributes as the inverse of the sample intensity (number of data). For simplicity, let us first consider two uniformly distributed datasets (1) and (2) and let k_1 be the number of elements from (1) overlapping with those in (2) and k_2 the number of elements from (2) overlapping with those in (1).

The first step consists in considering that non-overlapping elements from (1) and (2) contribute $\frac{1}{k_1}$ and $\frac{1}{k_2}$ respectively whereas the overlapping elements contribute $\frac{1}{k_1+k_2}$ to the global histogram. The latter assumption comes from the fact that if one uses $\frac{1}{k_i}$ for the

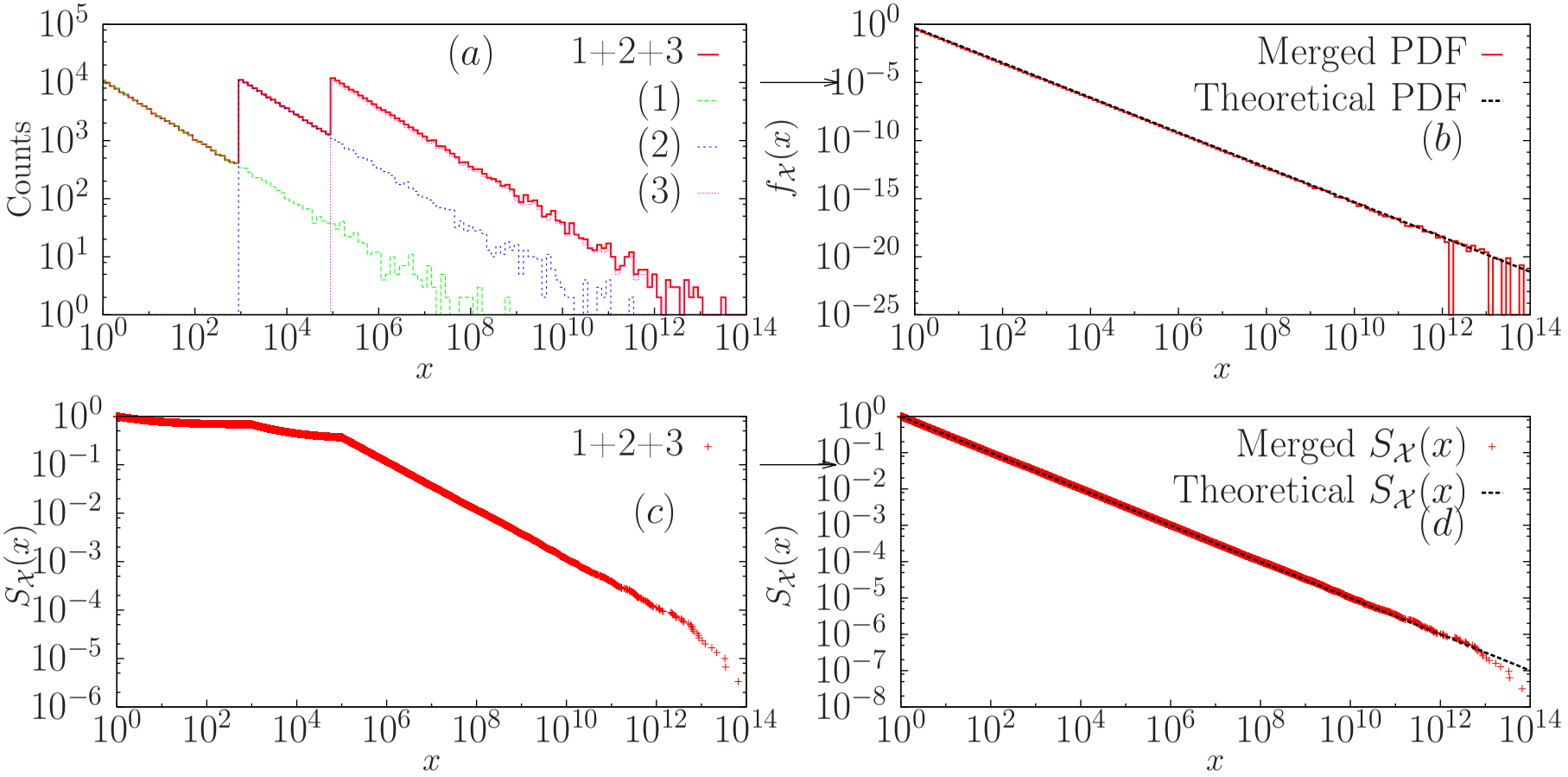

Fig. 7.2: (a) Original histogram of counts for three datasets of the same size $n = 10^5$ sampled from untruncated power-law distributions with the same exponent 1.5 but with different lower cut-offs: $x_{min}^{(1)} = 1$, $x_{min}^{(2)} = 10^3$, $x_{min}^{(3)} = 10^5$. (c) Original CCDF for the three datasets. Merged PDF (b) and CCDF (d) of the three datasets together with the theoretical curves.

overlapping data, an histogram twice as high as it should be (due precisely to the overlapping) would be obtained. Weighting each contribution by $\frac{k_i}{k1+k2}$ leads to the $\frac{1}{k_1+k_2}$ value. After this first rescaling, datasets (1) and (2) are merged in a single (smooth) histogram $(1+2)$.

In order to add a new dataset (3), the contribution of $((1)+(2))$ that overlaps with (3) has to be evaluated. Let us define k_{12} as the number of elements of (1) and (2) overlapping with (3). Each of these elements contribute with its respective weight to a global weight

$$\omega_{12} = \underbrace{\frac{1}{k_1} + \frac{1}{k_1} + ... + \frac{1}{(k_1+k_2)} + \frac{1}{(k_1+k_2)} + ... + \frac{1}{k_2} + ...}_{k_{12}\text{terms}}$$

where the three different terms in the sum come from the data in (1) that do not overlap with (2), the data in (1) that overlap with (2) or vice versa, and the data in (2) that do not overlap with (1), respectively (and with all of them overlapping with (3)).

In order to compare properly the histogram $(1)+(2)$ with (3), the former should come with a weight k_{12} (instead of ω_{12}), so the former is rescaled by the factor $\frac{k_{12}}{\omega_{12}}$. In this way, the histogram from $(1)+(2)$ turns out to have the different contributions $\frac{k_{12}}{k_1\omega_{12}}$, $\frac{k_{12}}{(k_1+k_2)\omega_{12}}$, and $\frac{k_{12}}{k_2\omega_{12}}$ (depending on which part of the domain each data point comes from).

The next step consists in applying the original procedure to $(1)+(2)$ and (3), thus rescaling in each case by $\frac{1}{k_{12}}$, $\frac{1}{k_{12}+k_3}$, or $\frac{1}{k_3}$, where k_3 refers to the number of data from (3) overlapping with $(1)+(2)$ (i.e., with (1) or (2)). To be precise, data points from the already assembled histogram $((1)+(2))$ not overlapping with (3) contribute introducing the extra factor $\frac{1}{k_{12}}$, yielding the weights Ω

$$\frac{1}{k_1\omega_{12}}, \frac{1}{(k_1+k_2)\omega_{12}}, \frac{1}{k_2\omega_{12}},$$

for each case; the overlapping data points contribute $\frac{1}{(k_{12}+k_3)}$, yielding the weights Ω

$$\frac{k_{12}}{k_1(k_{12}+k_3)\omega_{12}}, \frac{k_{12}}{(k_1+k_2)(k_{12}+k_3)\omega_{12}}, \frac{k_{12}}{k_2(k_{12}+k_3)\omega_{12}},$$

if the data point came from (1) or (2), and $\frac{1}{k_{12}+k_3}$ if they come from (3); and non-overlapping data points from the new dataset (3) contribute $\frac{1}{k_3}$.

One can iterate this procedure in order to add as many datasets as necessary. In any case, note that one can only add datasets that overlap, as the overlapping region is used as a way of "calibration". This method is independent of the shape of the histogram (a convenient change of variables can transform any distribution into a uniform one, without affecting the procedure). Once each point has a certain weight Ω_r ($r = 1, ..., \mathcal{N}$, where $\mathcal{N} = \sum_{i=1}^{n_{ds}} n_i$) assigned, one constructs the histogram in the same way as it was explained in Sec. 5.3. However, note that instead of contributing one unity, each point in the dataset contributes according to the weight Ω_r assigned by the algorithmic procedure. In order to have an estimation of a PDF, firstly, rescaled counts should be divided by the bin-width for

each bin and, secondly, the resulting histogram has to be divided by the effective number of counts ($n_{eff} = \sum_{r=1}^{\mathcal{N}} \Omega_r$) in order to fulfil normalization.

As already mentioned, depending on which overlapping or non-overlapping region a point from data comes from, each point has a certain weight Ω_r. By sorting data in ascending order, it is easy to construct the empirical Cumulative Distribution Function (CDF) for the merged datasets by using the following expression:

$$F_{\mathcal{X},e}(x_{(l)}) = \frac{\sum_{r=1}^{l} \Omega_r}{\sum_{r=1}^{\mathcal{N}} \Omega_r}, \tag{7.8}$$

with $l = 1, ..., \mathcal{N}$ and the sub-index e refers to the empirical CDF. Note that, for a standard situation in which datasets are not merged, all the weights would be $\Omega_r = 1 \quad \forall r$ [Deluca and Corral, 2013]. Note also that the Complementary Cumulative Distribution Function ($S_{\mathcal{X},e}$) is $S_{\mathcal{X},e}(x_{(k)}) = 1 - F_e(x_{(k)})$.

Figure 7.2 shows this procedure for three datasets (Datasets (1), (2) and (3)) of the same size $n = 10^5$ sampled from untruncated power-law distributions with the same exponent 1.5 but with different lower cut-offs: $x_{min}^{(1)} = 1$, $x_{min}^{(2)} = 10^3$, $x_{min}^{(3)} = 10^5$. The structure of these samples is similar to the schematic representation in Fig. 7.1. In left panels of Fig. 7.2, irregularities can be appreciated in the histogram of counts as well as in the CCDF for three datasets, where the algorithmic procedure has not been applied yet. Right panels show the merged PDF and CCDF after applying the procedure.

7.1.2.2 Numerical procedure

Let $\mathcal{C}_{ij}$ be the number of counts from the i-th sample in the j-th interval defined by $\left[x_{[j]}, x_{[j+1]}\right)$. The number of counts from the i-th sample in the j-th interval will be zero if none of the n_i observations is contained in $\left[x_{[j]}, x_{[j+1]}\right)$. However, this does not necessarily imply that the probability to find an observation in that interval is zero. If no observations have occurred, this might be due to experimental limitations in the acquisition of data, due to poor statistics, etc. This procedure is based on the assumption that the probability to find an observation in the j-th interval is proportional to the number of counts; even where there are no observations from the i-th sample, one would expect a finite value of $\mathcal{C}_{ij}$. In this context, it must be considered that the observed $\mathcal{C}_{ij}$ is unknown if the i-th sample is not contained in the j-th interval. Under this framework, the goal of building a single histogram is reduced to determine the unknown $\mathcal{C}_{ij}$.

If independence is both assumed among samples and among observations, then an arbitrary observation can be measured by two random variables $\mathcal{D}$ and $\mathcal{J}$: $\mathcal{D}$ corresponds to the sample index where the observation was collected and $\mathcal{J}$ is the label of the interval where it is contained. The variable $\mathcal{D}$ takes values $i = 1, ..., n_{ds}$ and $\mathcal{J}$ takes values $j = 1, ..., n_{int}$. The probability that an observation from the i-th sample lies in the j-th interval can be estimated

by:

$$\mathbb{P}\left(\mathcal{J}=j|\mathcal{D}=i\right)\simeq\frac{\mathcal{C}_{ij}}{\sum_{j=1}^{n_{int}}\mathcal{C}_{ij}}\,.$$

Under the assumption of independence among samples and among observations, one can compute an expected value of the ij-th entry:

$$e_{ij}=n_i\sum_{r=1}^{n_{ds}}\mathbb{P}\left(\mathcal{J}=j|\mathcal{D}=r\right)=\frac{\sum_{r=1}^{n_{ds}}\mathcal{C}_{rj}\sum_{l=1}^{n_{int}}\mathcal{C}_{il}}{\sum_{r=1}^{n_{ds}}\sum_{l=1}^{n_{int}}\mathcal{C}_{rl}}\,, \tag{7.9}$$

where r and l are dummy indexes. Then the unknowns c_{ij} can be determined by minimizing the Pearson's χ^2 statistic:

$$\chi^2=\sum_{i=1}^{n_{ds}}\sum_{j=1}^{n_{int}}\frac{\left(\mathcal{C}_{ij}-e_{ij}\right)^2}{e_{ij}}. \tag{7.10}$$

Therefore, under the assumption of independence, building up an histogram of merged datasets has been reduced to an optimization problem where one might use numerical methods to find the values of the unknonws that minimize the χ^2 statistic.

Let us apply this numerical procedure to the same samples that were merged in the previous section. In this particular case, there are three samples $i=1,2,3$ which can be classified into five different intervals $j=1,...,5$. According to this procedure, one can construct the table of counts shown in Table 7.1.

	Interval_j	1	2	3	4	5	$\sum_j$
Sample_i							
1		96774	2882	344	$\mathcal{C}_{14}=0$	$\mathcal{C}_{15}=0$	10^5
2		$\mathcal{C}_{21}=0$	90165	9762	73	$\mathcal{C}_{25}=0$	10^5
3		$\mathcal{C}_{31}=0$	$\mathcal{C}_{32}=0$	99356	636	8	10^5
$\sum_i$		96774	93047	109462	709	8	3×10^5

Table 7.1: Table of the initial observed (real) counts for each sample in each interval.

In this particular example, there are six unknowns $\mathcal{C}_{14}$, $\mathcal{C}_{15}$, $\mathcal{C}_{21}$, $\mathcal{C}_{25}$, $\mathcal{C}_{31}$, $\mathcal{C}_{32}$ which are initialized at zero. Once this table is known, it is easy to construct the table of the expected values according to Eq. (7.9) (see Table 7.2). Table 7.3 shows all the contributions to the chi-square statistic computed from Eq. (7.10). At this step, the main goal is to find the value of the six unknowns (which are assumed to be larger or equal than zero) that minimize the χ^2 statistic. In order to do that, the solver tool provided by the Excel software [Microsoft Corporation, 2018] as well as the function *"optimize"*, which is already implemented in R programming language [R Documentation, 2017; Brent, 2002], enable to optimize the value of the statistic for these unknowns. Once this numerical method is applied, it is found that the minimum value of the statistic in this case is $\chi^2=4.34$, and the solutions of the unknowns are exposed in Table 7.4.

	Interval_j	1	2	3	4	5	$\sum_j$
Sample_i							
1		32258	31015.67	36487.33	236.33	2.67	10^5
2		32258	31015.67	36487.33	236.33	2.67	10^5
3		32258	31015.67	36487.33	236.33	2.67	10^5
$\sum_i$		96774	93047	109462	709	8	3×10^5

Table 7.2: Table of the expected counts for each sample in each interval.

	Interval_j	1	2	3	4	5	$\sum_j$
Sample_i							
1		129032	25519.46	35802.58	236.33	2.67	190593.04
2		32258	112802.46	19575.11	112.88	2.67	164751.12
3		32258	31015.67	108324.42	675.88	10.67	172284.64
$\sum_i$		193548	169337.60	163702.10	1025.10	16	χ^2=527628.80

Table 7.3: Table with the different contributions to the χ^2 statistic. Initial value of the χ^2 statistic is shown.

Given that the value of the unknowns that minimize the statistic are determined, the number of counts of each sample is known for each interval and the PDF can be easily constructed just by adding all the contributions of the samples for each interval. Once each interval has a certain contribution from each sample, one builds the histogram in the same way as it was explained in Sec. 5.3. In Fig. 7.3 the merged histogram is obtained by means of the numerical procedure. As it can be seen, algorithmic and numerical approaches practically overlap. Tiny deviations can be associated to the numerical method.

Although this numerical method gives similar results to those obtained with the algorithmic procedure, the computational effort is higher and it requires complementary numerical methods whose convergence is not ensured or it might be difficult to reach. Note that, in the worst situation, datasets will overlap consecutively ($x_{top}^{(i)} > x_{low}^{(i+1)}$ and $x_{top}^{(i)} < x_{low}^{(i+2)} < x_{top}^{(i+1)}$, $\forall i$) in pairs thus maximizing the number of unknowns $n_{un} = 2\left(n_{ds} - 1\right)^2$. This implies that for three datasets one would have to find eight unknowns and for ten datasets, 162 unknowns. The number of unknowns increases dramatically with the number of merged datasets.

	Interval_j	1	2	3	4	5	$\sum_j$
Sample_i							
1		96774	2882	344	$\mathcal{C}_{14} = 2.07$	$\mathcal{C}_{15} = 0.03$	100002.08
2		$\mathcal{C}_{21} = 2.997 \times 10^6$	90165	9762	73	$\mathcal{C}_{25} = 0.79$	3.097×10^6
3		$\mathcal{C}_{31} = 3.037 \times 10^6$	$\mathcal{C}_{32} = 9.137 \times 10^5$	99356	636	8	3.139×10^7
$\sum_i$		3.347×10^7	1.007×10^6	109462	711.06	8.81	3.458×10^7

Table 7.4: Table with the solutions that minimize the χ^2 statistic.

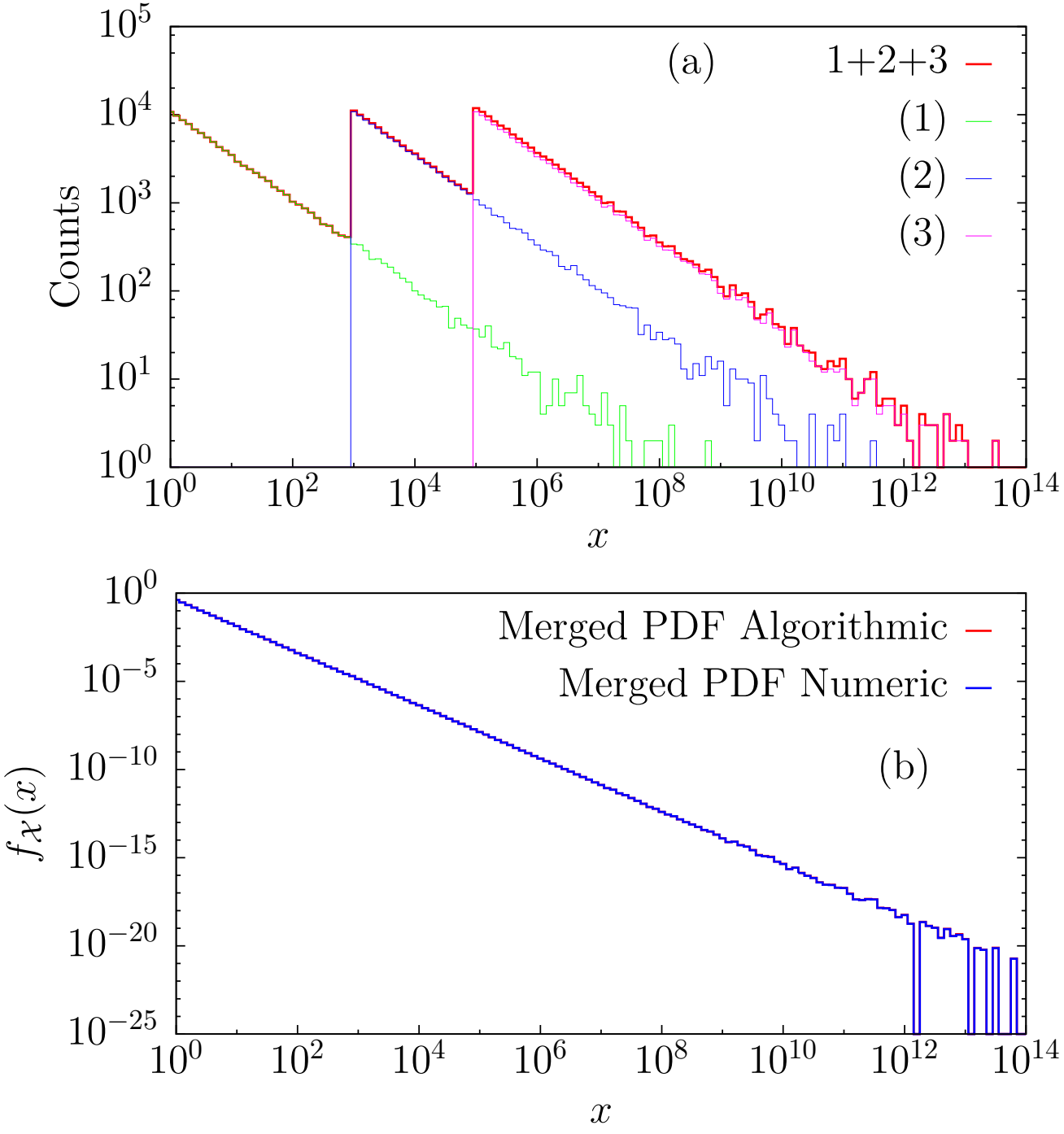

Fig. 7.3: (a) Original histogram of counts for three datasets of the same size $n = 10^5$ sampled from untruncated power-law distributions with the same exponent 1.5 but with different lower cut-offs: $x_{min}^{(1)} = 1$,$x_{min}^{(2)} = 10^3$, $x_{min}^{(3)} = 10^5$. (b) Merged PDF of the three datasets together by means of the numerical procedure and the algorithmic procedure.

This is the main reason why, except stated otherwise, the algorithmic method will be the preferred one when plotting PDFs or in goodness-of-fit tests.

7.1.3 Different goodness-of-fit test

Determining whether the null hypothesis of considering a global exponent Γ is compatible with the values of the particular fits shows some differences with the standard Kolmogorov-Smirnov (KS) goodness-of-fit test presented in Sec. 5.4. Two different statistics for the goodness-of-fit test that have been used for the global fit are presented: the first statistic is essentially an adaptation of the standard KS test for the merged case whereas the second method uses a statistic which is a composition of KS distances. When the value of the global exponent $\hat{\Gamma}$ has been found through MLE, one can understand that each dataset contributes to the global PDF with a global exponent $\hat{\Gamma}$ in their particular ranges $\left[x_{min}^{(i)}, x_{max}^{(i)}\right]$ $(i = 1, ..., n_{\rm ds})$. The global distribution can be understood as a global power-law PDF with exponent Γ ranging from $X_{min} = \min\{x_{min}^{(i)}\}$ to $X_{max} = \max\{x_{max}^{(i)}\}$. In this situation where several datasets are merged, it is necessary to redefine the KS distance.

- **Kolmogorov-Smirnov Distance of the Merged Datasets (KSDMD):** This statistic can be used as long as datasets overlap each other. Once the merged $F_{\mathcal{X},e}(x)$ has been constructed according to the procedure explained in Sec. 7.1.2, the KS distance can be easily computed for the truncated power-law case by:

$$d_e^{(M)} = \max\left| \left(\frac{x^{1-\hat{\Gamma}} - X_{min}^{1-\hat{\Gamma}}}{X_{max}^{1-\hat{\Gamma}} - X_{min}^{1-\hat{\Gamma}}} \right) - F_{\mathcal{X},e}(x) \right|, \tag{7.11}$$

  where the values of x are taken from the ranges $\left[x_{min}^{(i)}, x_{max}^{(i)}\right]$. In Fig. 7.4 the distribution of the KSDMD statistic is shown for three simulated power-law datasets of size $n = 10^3$ each, with exponent $\gamma = 1 + \frac{2}{3}$ for 10^4 realizations. The distribution approximately follows a KS distribution.

- **Composite KS Distance (CKSD):** This second statistic can be computed independently on whether datasets are overlapping each other or not. This statistic is constructed from the $n_{\rm ds}$ particular KS distances:

$$d_{e,i} = \max_{x_{min}^{(i)} \leq x \leq x_{max}^{(i)}} \left| \left(\frac{x^{1-\hat{\Gamma}} - x_{min}^{(i)1-\hat{\Gamma}}}{x_{max}^{(i)1-\hat{\Gamma}} - x_{min}^{(i)1-\hat{\Gamma}}} \right) - F_{\mathcal{X},e,i}\left(x; x_{min}^{(i)}, x_{max}^{(i)}\right) \right|, \tag{7.12}$$

  where the subindex i refers to the i-th catalog and $F_{\mathcal{X},e}$ is the empirical CDF computed according the methodology explained in Sec. 5.4. Note that these distances are computed for each individual dataset, therefore, the power-law cut-offs correspond to those of

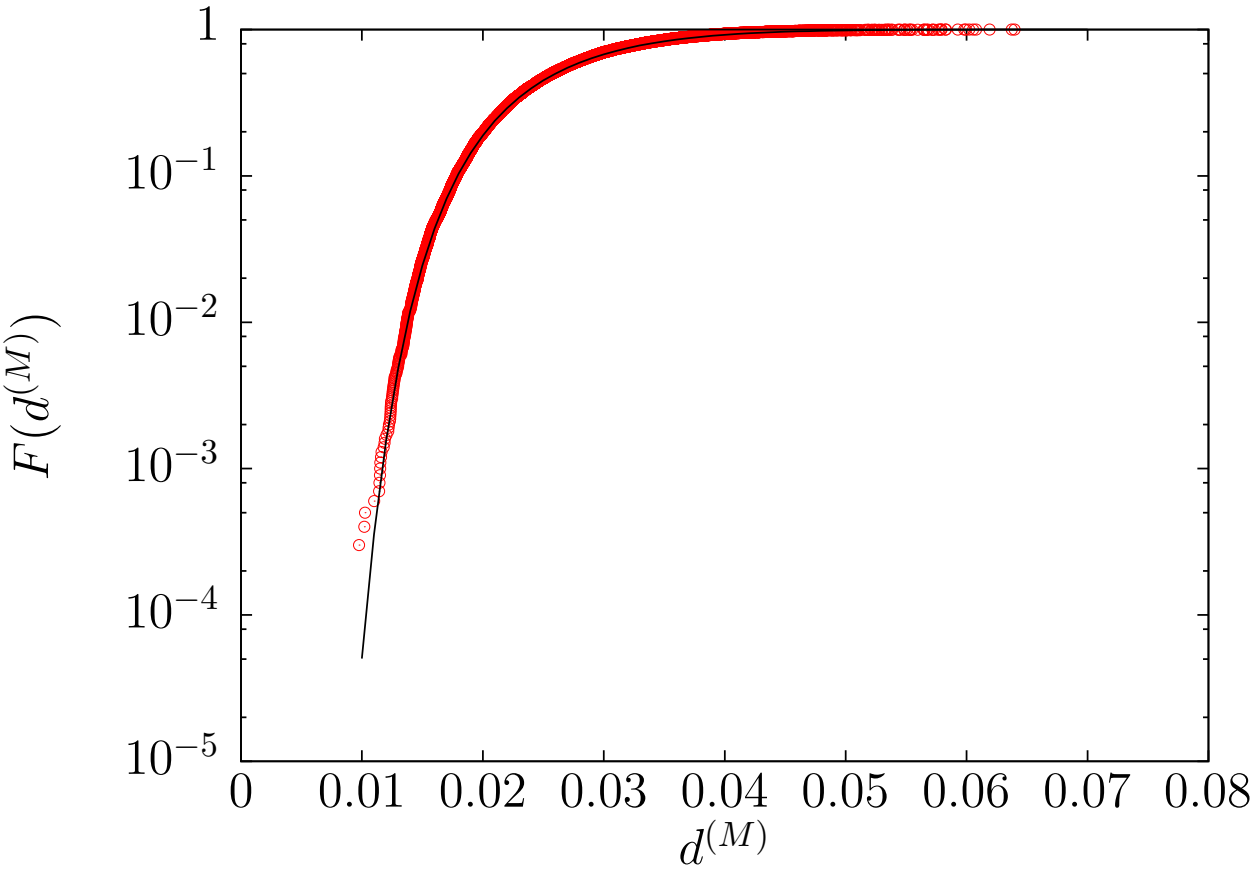

Fig. 7.4: Probability distribution of the KSDMD statistic for three simulated untruncated power-law datasets with $n = 10^3$ each, overlapped ranges and $\gamma = 1 + \frac{2}{3}$ for 10^4 realizations. Black solid line correspond to a Kolmogorov-Smirnov distribution.

each dataset $\left[x_{min}^{(i)}, x_{max}^{(i)}\right]$ In order to compute the statistic, the following summation is performed:

$$d_e^{(C)} = \sum_{i=1}^{\mathrm{n_{ds}}} \sqrt{n_i} D_{e,i}\,, \tag{7.13}$$

where the factors $\sqrt{n_i}$ are due to the scaling of the KS distance with the number of data [W. H. Press and Flannery., 1992].

In Fig. 7.5 the distribution of the CKSD statistic is shown for three simulated power-law datasets of size $n = 10^3$ each, with exponent $\gamma = 1 + \frac{2}{3}$ for 10^4 realizations.

Once the statistic is computed, one needs to determine whether it is big or small in relation to the one found for data sampled from a PDF with the same parameters $\{x_{min}^{(i)}\}$,$\{x_{max}^{(i)}\}$, Γ and $\{n_i\}$. Similarly to the goodness-of-fit test in Sec. 5.4, the p-value obtained from the theoretical distribution of the KSDMD statistic is biased when the KSDMD is computed with

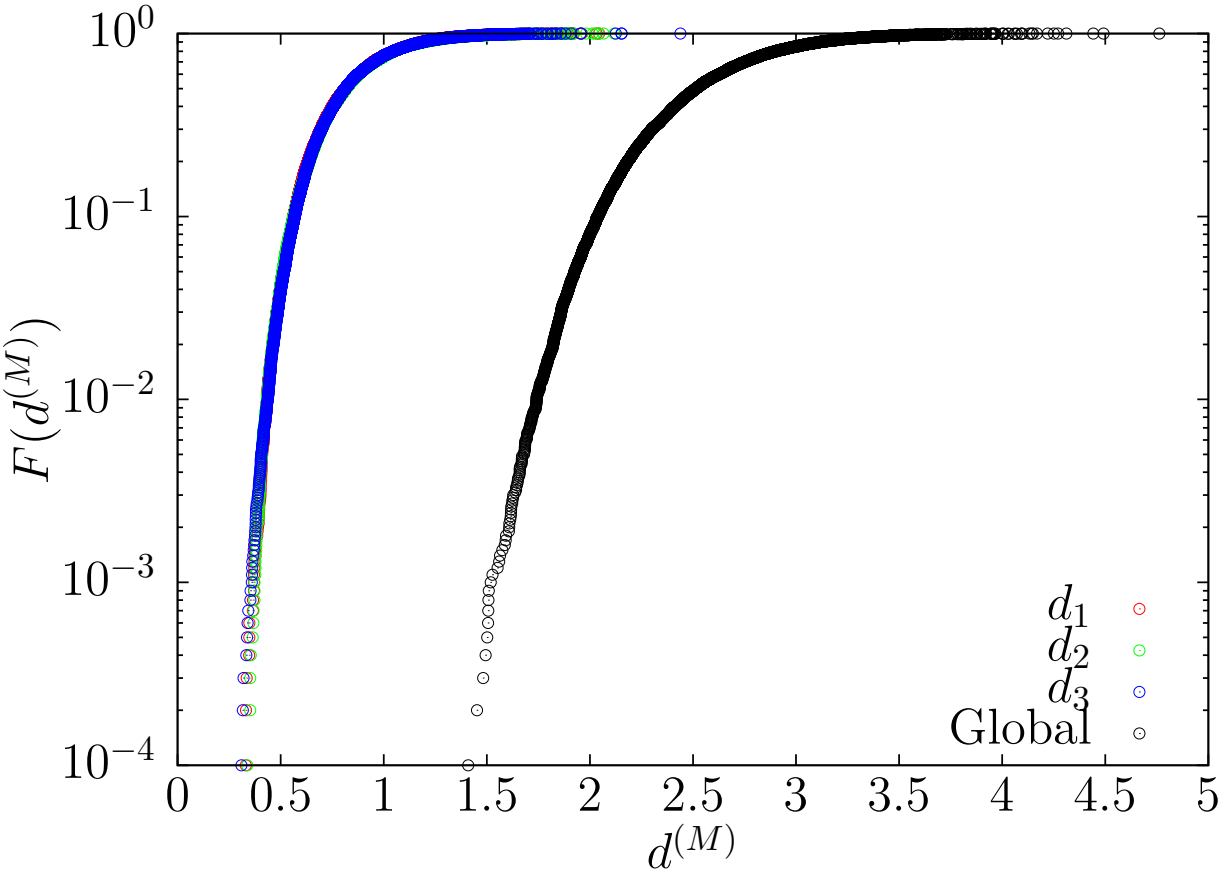

Fig. 7.5: Probability distribution of the CKSD Statistic for three simulated untruncated power-law datasets with $n = 10^3$ each, overlapped ranges and $\gamma = 1 + \frac{2}{3}$ for 10^4 realizations. Probability distributions of the particular datasets (1), (2) and (3) are also shown.

the estimated value of the exponent. Consequently, it is necessary to perform Monte Carlo simulations in order to obtain an unbiased p-value.

Simulated data is sampled from a power-law PDF in the range given by the i-th catalog with probability $q_i = n_i/\mathcal{N}$, where $\mathcal{N} = \sum_{i=1}^{\mathrm{n_{ds}}} n_i$. Note that the particular number in each simulated dataset is not necessarily the empirical one n_i but the total number of data $\mathcal{N}$ is maintained. Hence, one needs a first random number to choose the dataset i and therefore the range $\left[x_{min}^{(i)}, x_{max}^{(i)}\right]$ and a second one to generate the random truncated power-law number in that range with exponent $\hat{\Gamma}$ [Deluca and Corral, 2013]. When $\mathcal{N}$ events have been generated according to this procedure, one finds the global exponent $\hat{\Gamma}_{\mathrm{sim}}$ by maximizing the global log-likelihood and computes either $d_{sim}^{(M)}$ or $d_{sim}^{(C)}$ for the merged simulated datasets. By performing several realizations of the previous procedure, one can estimate the p-value by computing the fraction of simulated datasets where the simulated statistic is larger than the empirical one.

7.1.4 Performance over synthetic datasets

Once the methodology for merging datasets has been presented together with the different goodness-of-fit tests, it is important to check the performance of the method over synthetic data. In order to carry out this analysis, two untruncated power-law distributed datasets with exponents γ_1 and γ_2, lower cut-offs $x_{min}^{(1)}$ and $x_{min}^{(2)}$ and sizes n_1 and n_2 are generated. In order to simplify the analysis, the sizes of both datasets are considered to be equal $n_1 = n_2$. The global exponent $\hat{\Gamma}$ of the merged catalogs is estimated according to the methodology exposed in the main text and the LRT statistic $2\mathcal{R}_e$ is computed according to the methods exposed in Sec 5.6. Both KSDMD and CKSD statistics are computed together with their respective p-vales $p_{value}^{(M)}$ and $p_{value}^{(C)}$. The source code in C for this analysis is shown in Appendix 9.4.

As mentioned in Sec. 5.6, it must be checked whether the distribution of the statistic follows a χ^2 distribution with as many degrees of freedom as the difference of parameters between the more complex and the simpler model or not. Figure 7.6 shows the distribution of the statistic $2\mathcal{R}_e$ for 10^4 pairs of power-law datasets under the null hypothesis that data can be described by the simpler model (both exponents are the same $\gamma_1 = \gamma_2 = 1.5$). As it can be observed, the distribution is in agreement with a χ_1^2 distribution with one degree of freedom. This is in agreement with the results exposed by [Serra and Corral, 2017]. This result ensures using the critical value provided by the theoretical distribution of the statistic. This is important in order to avoid performing Monte-Carlo simulations to compute a p-value for the LRT, which would imply a higher computational effort. Given that the difference between between the number of parameters in model 1 and 2 in this case is 1, the critical value of the test with a significance level equal to 0.05 is $2\mathcal{R}_c = 3.84$. If the empirical LRT statistic is found to be larger than the critical one, one then rejects the null hypothesis that the simpler model 1 is good enough to describe data and, consequently, more parameters are needed. Once the global exponent and the likelihood ratio statistic are found, the two different goodness-of-fit test exposed in Sec 7.1.3 are performed.

The analysis is performed for different dataset sizes as well as different power-law exponents γ_1 and γ_2. The results of the method for synthetic datasets are shown in Table 7.5. Five groups are presented depending on the relative difference δ_γ between exponents. The p-values of the CKSD and the KSDMD goodness-of-fit tests have been computed with 10^4 Monte Carlo simulations. In order to compare our results with a test that checks whether two power-law exponents are significantly different or not, the p_{norm}-value of the z-test exposed in Sec. 5.9 is also shown. These p-values can be computed by assuming that, for a sufficiently large sample sizes, the z-statistic follows a normal distribution with zero mean and standard deviation equal to one. As one would expect, if both datasets have exactly the same power-law exponent, the null hypothesis that variable $\mathcal{X}$ is power-law distributed for all the datasets with the global exponent $\hat{\Gamma}$ is not rejected. When the CKSD goodness-of-fit test rejects the null hypothesis of a power-law distribution with a global exponent, the z-test also rejects the null hypothesis of considering $\gamma_1 = \gamma_2$. The same does not apply for the KSDMD goodness of fit test, where some fits yield to non-rejectable p-values whereas the z-test clearly rejects

n_1	γ_1	σ_1	n_2	γ_2	σ_2	δ_γ (%)	$\hat{\Gamma}$	$2\mathcal{R}_e$	$d_e^{(C)}$	$p_{value}^{(C)}$	$d_e^{(M)}$	$p_{value}^{(M)}$	z	p_{norm}-z
10^2	1.50	0.05	10^2	1.50	0.05	0	1.471	0.624	1.635	0.425	0.075	0.453	0.788	0.431
10^3	1.50	0.02	10^3	1.50	0.02	0	1.503	1.284	1.316	0.801	0.023	0.595	1.132	0.257
10^4	1.50	0.005	10^4	1.50	0.005	0	1.501	0.139	1.289	0.837	0.009	0.986	0.373	0.707
10^2	1.53	0.05	10^2	1.50	0.05	2	1.484	1.443	1.660	0.395	0.070	0.578	1.195	0.232
10^3	1.53	0.02	10^3	1.50	0.02	2	1.518	0.029	1.239	0.876	0.017	0.893	0.170	0.865
10^4	1.53	0.005	10^4	1.50	0.005	2	1.516	14.037	3.187	5×10^{-4}	0.013	0.024	3.745	1.81×10^{-4}
10^2	1.575	0.058	10^2	1.50	0.05	5	1.514	3.154	1.815	0.249	0.058	0.755	1.755	0.079
10^3	1.575	0.018	10^3	1.50	0.02	5	1.541	2.458	2.032	0.111	0.031	0.180	1.567	0.117
10^4	1.575	0.006	10^4	1.50	0.005	5	1.536	64.957	4.740	0	0.029	0	9.474	0
10^2	1.65	0.065	10^2	1.50	0.05	10	1.528	6.956	2.262	0.046	0.054	0.832	2.571	0.010
10^3	1.65	0.02	10^3	1.50	0.02	10	1.57	22.365	3.369	10^{-4}	0.044	0.015	4.690	2.73×10^{-6}
10^4	1.650	0.007	10^4	1.50	0.005	10	1.566	329.552	10.601	0	0.053	0	17.933	0
10^2	1.725	0.073	10^2	1.50	0.05	15	1.551	11.566	2.661	0.008	0.090	0.260	3.262	0.001
10^3	1.725	0.023	10^3	1.50	0.02	15	1.598	51.266	4.608	0	0.066	10^{-4}	7.026	2.12×10^{-12}
10^4	1.725	0.007	10^4	1.50	0.005	15	1.593	667.096	14.507	0	0.075	0	25.204	0

Table 7.5: Performance of the methodology for merging datasets explained in the main text for two untruncated power-laws with exponents γ_1 and γ_2, lower cut-offs $x_{min}^{(1)} = 10^{12}$ and $x_{min}^{(2)} = 10^{14}$ (arbitrary units) and sizes n_1 and n_2. Five different groups are presented depending on the relative difference of the power-law exponents δ_γ. For these datasets, the fitted global exponent $\hat{\Gamma}$ is estimated and the LRT statistic $2\mathcal{R}_e$ is computed. The statistics $D_e^{(C)}$ and $D_e^{(M)}$ together with their corresponding p-values $p_{value}^{(C)}$ and $p_{value}^{(M)}$ are presented for each combination of datasets. As a complement, the z-statistic and the p-value p_{norm}-z from to the z-test are also shown.

the null hypothesis. In this sense, one can consider that the KSDMD statistic is less powerful than the CKSD statistic. However, for the sample sizes which are involved in this work, both goodness-of-fit tests reject the null hypothesis for sufficiently large difference in the exponents. It can also be seen that, the null hypothesis is rejected independently on the goodness-of-fit test for those fits in which the Likelihood Ratio statistic exceeds the critical value $2\mathcal{R}_c = 3.84$. This fact justifies the decision of performing the LRT before the goodness-of-fit test.

Finally, let us remark that, based on this analysis, the goodness-of-fit test performed with the CKSD statistic is much better than the one performed with the KSDMD. Not only it is much powerful but also it can always be computed, independently on whether datasets overlap or not.

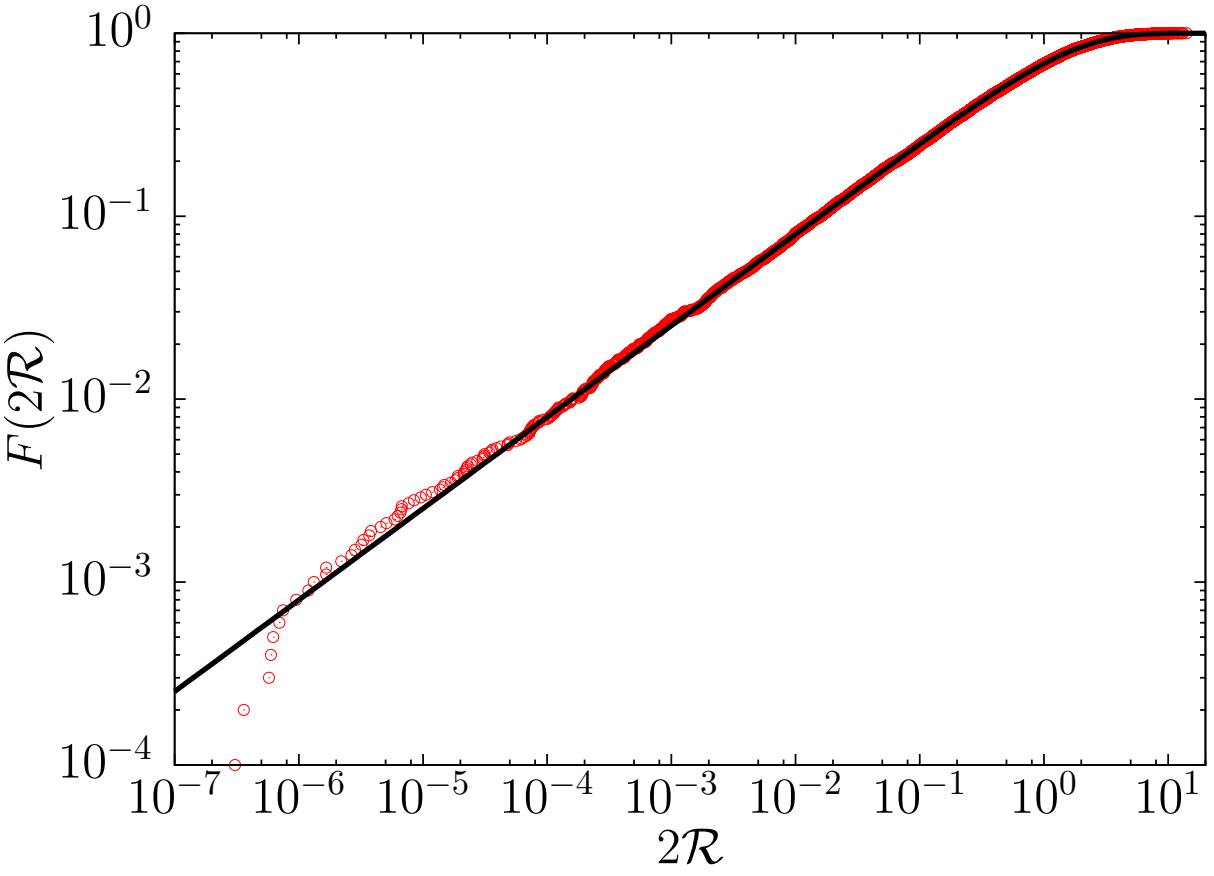

Fig. 7.6: Red points correspond to different values of the statistic $2\mathcal{R}$ computed with a simpler (one global fitted exponent $\hat{\Gamma}$) and a more complex model (assuming two different fitted exponents $\hat{\gamma}_1$, $\hat{\gamma}_2$) of two untruncated power-laws with exponents γ_1 and γ_2 (under the null hypothesis $\gamma_1 = \gamma_2$), sizes $n_1 = n_2 = 10^4$. Black line corresponds to a χ^2_1 distribution with one degree of freedom.

7.1.4.1 Effect of the magnitude resolution in earthquake catalogs

A different question that can be addressed is whether the resolution of magnitudes m in earthquake catalogs affects the fitting procedures when PDFs are considered as continuous. As explained in Sec. 4.2, some catalogs provide the magnitudes with a certain resolution Δm, leading to discretized values when they are transformed into seismic moments in order to fit a power-law distribution. In order to check the effect of the resolution Δm when fitting a power-law, n exponentially distributed magnitudes are generated according to the Gutenberg-Richter law from uniform random numbers u according to

$$m = m_{min} - \frac{\log_{10}(1-u)}{b}\,.$$

Once they are generated, these values are rounded according to the bin-width or resolution Δm and converted into seismic moment according to Eq. (3.11).

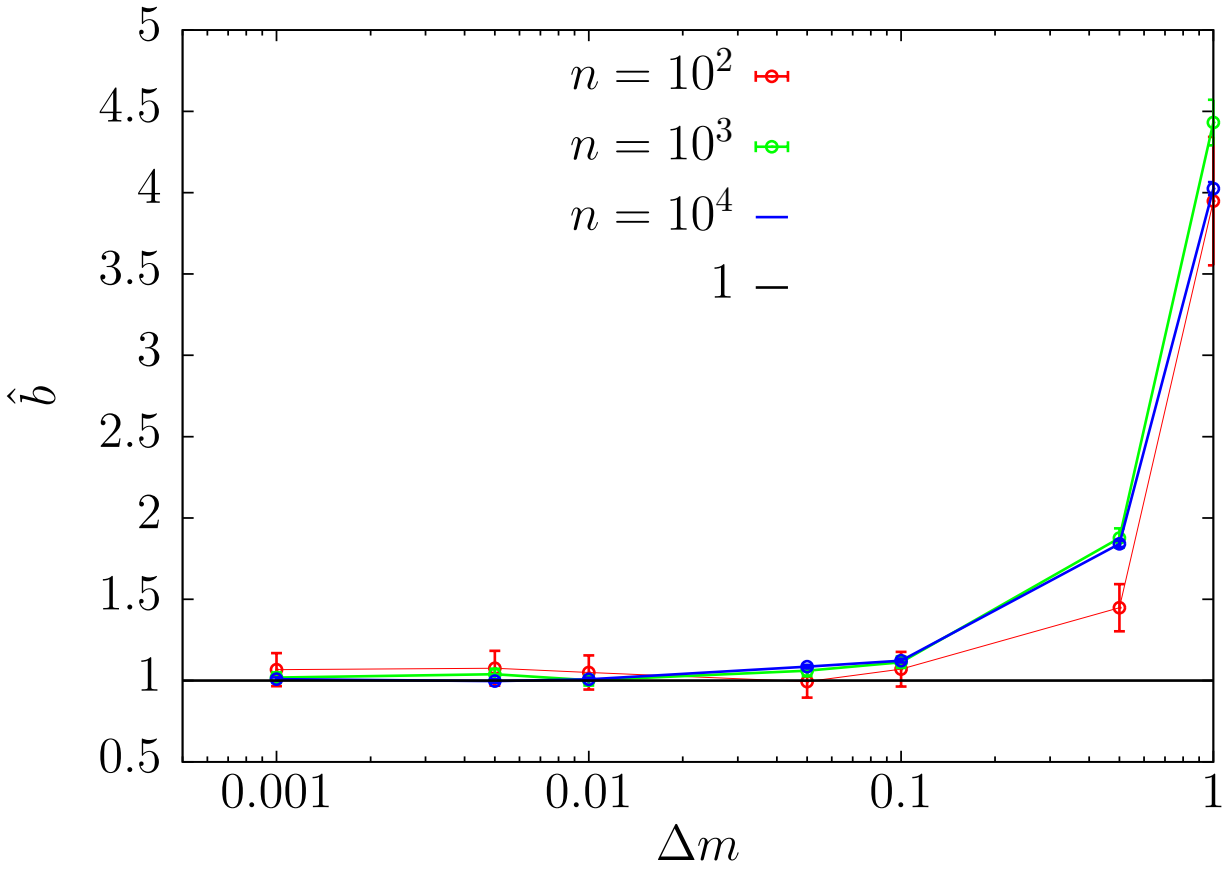

Fig. 7.7: ML estimator $\hat{b}$-value as a function of the resolution in magnitudes Δm for different sample sizes. Data are sampled from a Gutenberg-Richter law with $m_{min} = 3$ and b-value$= 1$. Horizontal black line corresponds to the real b-value.

The fitting procedure explained in Sec. 5.5 is then applied and the results for different sample sizes ($n = 10^2$, $n = 10^3$ and $n = 10^4$) of magnitudes sampled from a Gutenberg-Richter law (b-value= 1 and $m_{min} = 3$) and discretized according to different values of the resolution Δm are shown in Table 7.6.

In Fig. 7.7 the fitted $\hat{b}$-value as a function of the resolution Δm is shown for the three different sample sizes. As it can be observed, the value of the ML estimator converges to the expected b-value= 1 as the magnitude bin-width is decreased. The ML estimator is significantly biased for resolutions larger than 0.1 for the sample sizes analyzed and is acceptable, within the error bars, for smaller values of the resolution [Bender, 1983]. However, the results of the goodness-of-fit test lead to rejectable p-values for large samples even when the ML estimator is not significantly biased.

Δm	m_{min}	b-**value**	**n**	$\hat{b}$	σ	p_{value}
1	3	1	10^2	3.948	0.394	0
			10^3	4.4326	0.140	0
			10^4	4.024	0.040	0
0.5	3	1	10^2	1.448	0.145	0
			10^3	1.876	0.059	0
			10^4	1.840	0.018	0
0.1	3	1	10^2	1.070	0.106	0
			10^3	1.113	0.035	0
			10^3	1.122	0.011	0
0.05	3	1	10^2	0.994	0.099	0.022 ± 0.001
			10^3	1.060	0.033	0
			10^4	1.086	0.011	0
0.01	3	1	10^2	1.050	0.105	0.578 ± 0.004
			10^3	1.013	0.032	0.267 ± 0.004
			10^4	0.997	0.010	0
0.005	3	1	10^2	1.076	0.108	0.856 ± 0.004
			10^3	1.039	0.033	0.396 ± 0.005
			10^4	0.995	0.010	0.033 ± 0.002
0.001	3	1	10^2	1.067	0.102	0.588 ± 0.005
			10^3	1.018	0.0322	0.828 ± 0.004
			10^4	1.009	0.010	0.386 ± 0.005

Table 7.6: Results of simulating magnitudes sampled from a Gutenberg-Richter law with $m_{min} = 3$ and b-value= 1 for different sample sizes n and different resolutions Δm. The fitted $\hat{b}$-values and their standard deviations σ are computed by means of MLE and p-values are extracted from the KS goodness-of-fit test explained in Sec. 5.4.

A resolution of $\Delta m = 10^{-3}$ ensures that the largest sample of $n = 10^4$ yields non-rejectable p-values. Some results exposed in this chapter deal with catalogs whose resolution and number of events would result in rejectable fits. However, non-rejectable p-values are found instead leading to the conclusion that data in real catalogs are even more reliable to follow a power-law distribution.

7.2 Merging labquake catalogs

As explained in Sec. 4.1.6, experiments performed at different values of the pre-amplification lead to different observation ranges in the energy and amplitude distributions. In this section, the distributions of energy and amplitudes of labquakes generated during the displacement-driven compression experiments of charcoal and Vycor are studied by using techniques explained in Secs. 5.5 and 7.1.

7.2.1 Energy distribution of Vycor Labquakes

Let us consider four labquake catalogs generated during the displacement-driven compression of Vycor and recorded for different values of the pre-amplification: 60 dB, 40 dB, 20 dB and 0 dB. Let us consider that, for the i-th ($i = 1, ..., 4$) experiment, the random variable $\mathcal{E}$ corresponding to energy follows a continuous truncated power-law distribution,

$$f_{\mathcal{E}}(E)dE = \frac{1-\gamma_{E,(i)}}{E_{\max}^{1-\gamma_{E,(i)}} - E_{\min}^{1-\gamma_{E,(i)}}} E^{-\gamma_{E,(i)}} dE \,. \tag{7.14}$$

Given that each experiment covers a different observation window, the methodology explained in Sec. 7.1 can be applied in order to merge the four labquake catalogs with the aim of exploring the range of validity of the power-law behaviour. In this particular case, the log-likelihood of model OneExp is written as:

$$\log \mathcal{L}_{OneExp} = \sum_{i=1}^{n_{ds}=4} \sum_{j=1}^{n_i} \log f_{\mathcal{E}} \left(E_{ij}; \Gamma_E, E_{min}^{(i)}, E_{max}^{(i)} \right) . \tag{7.15}$$

Given that there exist limitations due to saturation effects for high energy AE events, it is convenient to check at which value of the energy it is necessary to establish an upper truncation for the power-law regime. A first exploration can be performed by plotting the variations of the MLE of the particular exponents $\hat{\gamma}_{E,(i)}$ as a function of the cut-offs (see the MLE maps shown in Fig. 7.9). The residual CV-plots are shown in Fig. 7.8 in order to illustrate the different behaviour at the tail of the distributions. The residual CV-test reveals the upper truncations of the experiments performed at 60, 40 and 0 dB but it does not provide conclusive results for the experiment at 20 dB. In order to explore fewer combinations of cut-offs, thus reducing the computation time, the upper truncations $E_{max}^{(1)}$, $E_{max}^{(2)}$ and $E_{max}^{(4)}$

(corresponding to vertical black lines in Fig. 7.8) are fixed whereas $E_{max}^{(3)}$ is considered as a free parameter. The labels (1), (2), (3) and (4) correspond to the experiments performed at 60, 40, 20 and 0 dB respectively.

Five intervals per decade equally spaced in logarithmic scale are sampled for each catalog and all the possible combinations of cut-offs $E_{min}^{(1)}$, $E_{min}^{(2)}$, $E_{min}^{(3)}$, $E_{max}^{(3)}$, $E_{min}^{(4)}$ are checked for fixed upper-truncations $E_{max}^{(1)}$, $E_{max}^{(2)}$ and $E_{max}^{(4)}$.

In Table 7.7 the results of the global fits for Vycor catalogs for the KSDMD and the CKSD statistics are presented. In both cases, the global exponents are very similar but the fit performed with the KSDMD statistic maximizes the sum of orders of magnitude $\sum_{i=1}^{n_{ds}} \log_{10}\left(\frac{x_{max}^{(i)}}{x_{min}^{(i)}}\right)$. The same combination of cut-offs would not be a good fit if the CKSD statistic were used instead. This fact can be justified by the performance of the method for synthetic data by applying the goodness-of-fit test with both statistics. The goodness-of-fit test for the KSDMD statistic tends to exhibit non-rejectable p-values for the same datasets for which the CKSD statistic clearly rejects the null hypothesis. See Sec. 7.1.4 for more details. Although the fit performed with the KSDMD covers a larger number of orders magnitude, the definitive fit is the one performed with the CKSD and $p_c = 0.20$ (see Table 7.7). Given that the ranges covered by the other non-rejectable fits lead to the rejection of the global power-law hypothesis when the CKSD is used with $p_c = 0.20$, this definitive fit is more restrictive than the other ones (see Sec. 7.1.4) and it is the 'safest' choice. The definitive fit covers about nine orders of magnitude in energy with a number of data $\mathcal{N} = 24496$ entering into the fit.

Note that the value of the global exponent is in agreement with the weighted harmonic mean

$$\hat{\Gamma}_E = 1.354 \pm 0.002 \simeq 1 + \frac{\mathcal{N}}{\sum_{i=1}^{n_{\rm ds}} \frac{n_i}{\hat{\gamma}_{E,(i)}-1}} = 1.355,$$

see Sec. 7.1.1 for a justification of this result. Due to the upper-truncations, the value of the global exponent is not exactly the same as the value of the harmonic mean of the particular exponents.

By considering the sample performed with zero pre-amplification and under the hypothesis that a power-law distribution with the same exponent can be extended for larger values of the energy, the probability $\mathcal{P}\left(E > E_{top}^{(0dB)}\right)$ turns out to be 0.0461. The number of observed events whose amplitude is larger than $E_{min}^{(0dB)}$ is 209. For these $N_i = 209$ trials, the probability of having no events in this range can be estimated by $(1 - 0.0461)^{209} = 5.22 \times 10^{-5}$. Based on these simple calculations, one could justify the existence of a corner value due to the finite size of the sample properties [Serra and Corral, 2017]. However, this corner value would be only visible for the experiment at zero amplification; the same calculation for the experiment at 20 dB gives a probability of having no events above E_{top} of 0.02, which is not a negligible value at all.

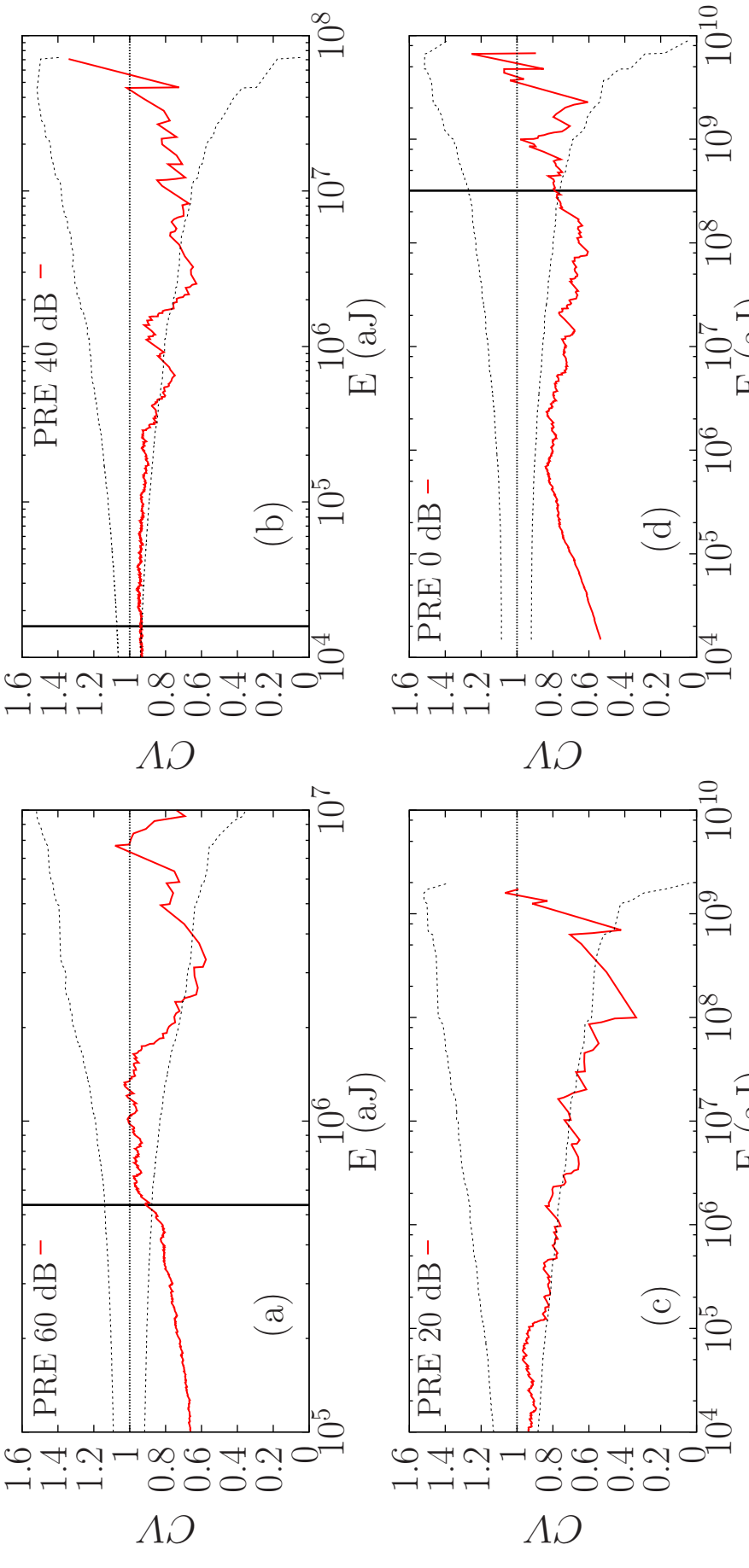

Fig. 7.8: In red, residual CV-plots for the logarithm of the energy from the four samples of Vycor performed at different pre-amplifications 60, 40, 20 and 0 dB ((a), (b), (c) and (d) panels respectively). Confidence intervals of 95% are represented by black-dashed lines as well as an horizontal dashed-black line placed at the value 1. Black vertical bars correspond to those values where a truncation x_{max} is performed and a different power-law behaviour at the tail of the distribution is detected.

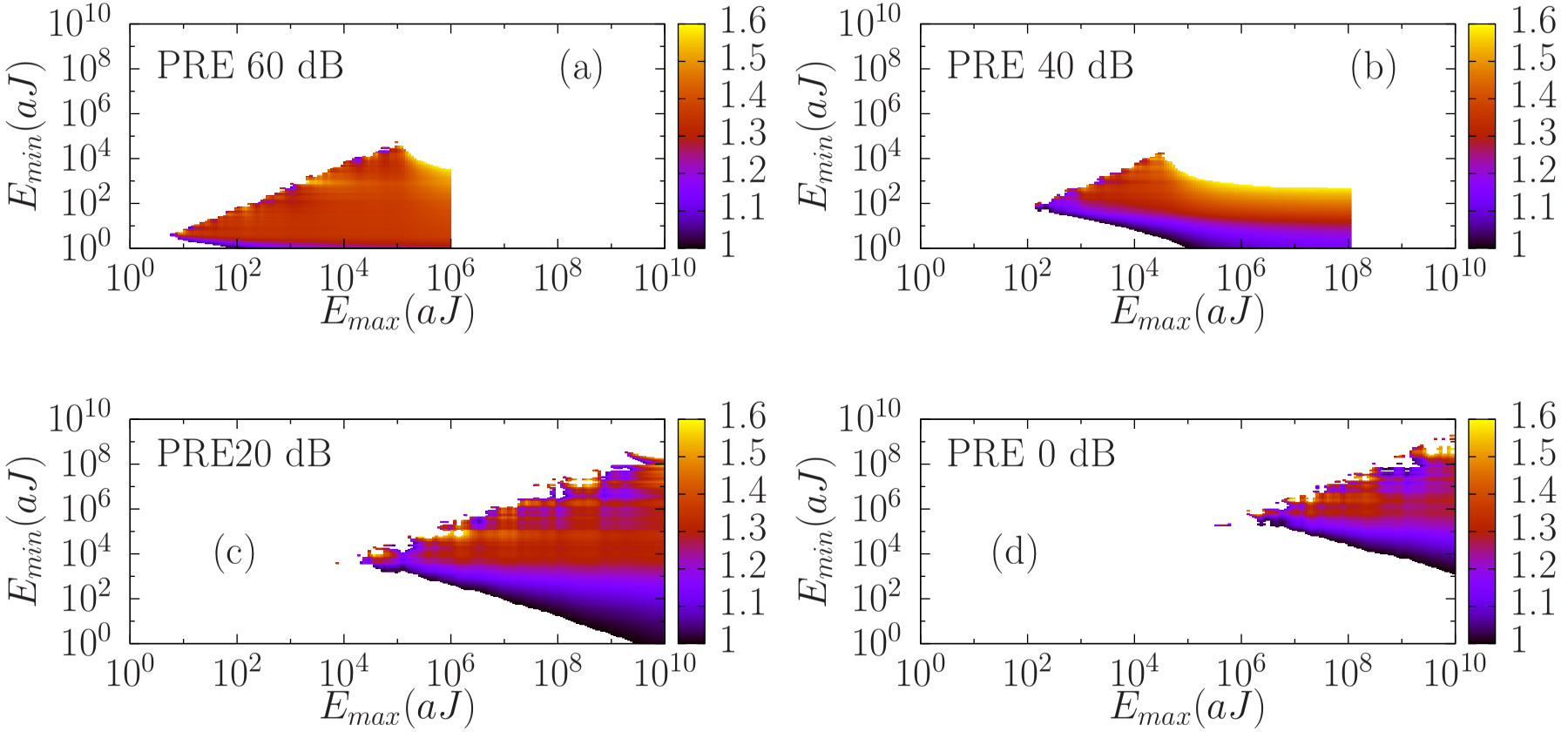

Fig. 7.9: MLE Exponents Maps for the exponent $\hat{\gamma}_E$ for the distribution of energies for the four samples of Vycor performed at different pre-amplifications 60, 40, 20 and 0 dB ((a), (b), (c) and (d) panels respectively).

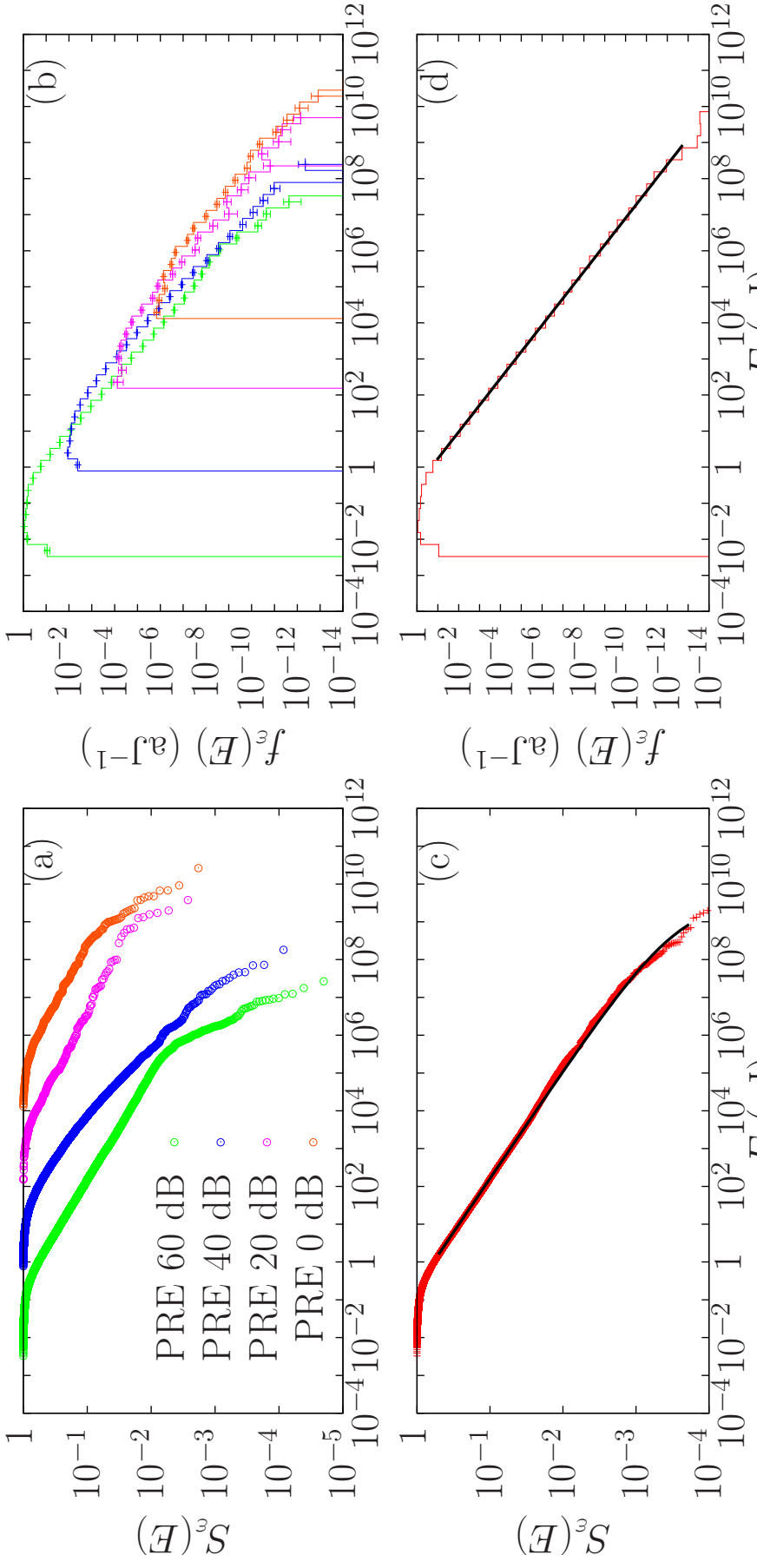

Fig. 7.10: Estimated Complementary Cumulative Distribution Functions ($S_\mathcal{X}$) of the energy distribution for each Vycor catalog (a) and for the merged Vycor catalogs (c). Estimated PDFs $f_\mathcal{E}(E)$ of the energy distribution for each Vycor catalog (b) and for the merged catalogs (d). Merged PDFs and CCDFs are plotted by following the procedure explained in Sec. 7.1.2. Events corresponding to the experiments performed at 40, 20, and 0 dB below their respective lower cut-offs as well as those corresponding to the experiments performed at 60, 40, and 20 dB above their respective upper cut-offs have been excluded. Power-law fits are represented by solid black lines.

$p_c = 0.20$	n	E_{min} **(aJ)**	E_{top} **(aJ)**	E_{max} **(aJ)**	**OM**	$\hat{\gamma}_E$	p_{value}
(1) PRE 60 dB	24338	1.58	2.66×10^7	5.37×10^5	5.53	1.351(2)	0.080(6)
(2) PRE 40 dB	5083	158.489	1.82×10^8	15860	2.00	1.34(1)	0.087(9)
(3) PRE 20 dB	263	6309.57	3.76×10^9	10^9	5.20	1.30(2)	0.67(1)
(4) PRE 0 dB	321	6.31×10^5	2.62×10^{10}	3.19×10^8	2.70	1.31(1)	0.25(1)
Model OneExp	$\mathcal{N}$				$\sum$ **OM**	$\hat{\Gamma}_E$	$p_{value}^{(M)}$
$2\mathcal{R}_e = 7.398$							
$d_e^{(M)} = 5.911 \times 10^{-3}$	30005	1.58	2.62×10^{10}	10^9	15.43	1.350(2)	0.20(1)

$p_c = 0.05$	n	E_{min} **(aJ)**	E_{top} **(aJ)**	E_{max} **(aJ)**	**OM**	$\hat{\gamma}_E$	p_{value}
(1) PRE 60 dB	24338	1.58	2.66×10^7	5.37×10^5	5.53	1.351(2)	0.080(6)
(2) PRE 40 dB	4179	251.189	1.82×10^8	15860	1.80	1.35(1)	0.19(1)
(3) PRE 20 dB	181	2.51×10^4	3.76×10^9	2.51×10^9	5	1.29(3)	0.18(1)
(4) PRE 0 dB	198	2.51×10^6	2.62×10^{10}	3.19×10^8	2.10	1.36(6)	0.36(2)
Model OneExp	$\mathcal{N}$				$\sum$ **OM**	$\hat{\Gamma}_E$	$p_{value}^{(C),0.05}$
$2\mathcal{R}_e = 4.722$							
$d_e^{(C)} = 3.363$	28896	1.58	2.62×10^{10}	2.51×10^9	14.43	1.351(2)	0.052(7)

$p_c = 0.20$	n	E_{min} **(aJ)**	E_{top} **(aJ)**	E_{max} **(aJ)**	**OM**	$\hat{\gamma}_E$	p_{value}
(1) PRE 60 dB	20778	2.51	2.66×10^7	5.37×10^5	5.33	1.354(3)	0.51(2)
(2) PRE 40 dB	3404	398.107	1.82×10^8	15860	1.60	1.36(2)	0.36(2)
(3) PRE 20 dB	155	3.98×10^4	3.76×10^9	10^9	4.40	1.32(4)	0.20(1)
(4) PRE 0 dB	159	3.98×10^6	2.62×10^{10}	3.19×10^8	1.90	1.34(7)	0.29(1)
Model OneExp	$\mathcal{N}$				$\sum$ **OM**	$\hat{\Gamma}$	$p_{value}^{(C),0.20}$
$2\mathcal{R}_e = 1.295$							
$d_e^{(C)} = 2.586$	24496	2.51	2.62×10^{10}	10^9	13.23	1.354(2)	0.20(1)

Table 7.7: Results of fitting models OneExp and MultiExp to the Vycor Labquake catalogs when doing the goodness-of-fit test with the KSDMD and the CKSD statistics and different values of p_c. $\hat{\gamma}_E$ is the estimated power-law exponent for a given experiment in the range $[E_{min}, E_{max}]$ including n data points. Values of the LRT statistic $2\mathcal{R}_e$ are provided in this case (recall that, the critical value in this case corresponds to the percentile 95 of a χ^2_3 distribution with 3 degrees of freedom $2\mathcal{R}_c = 7.81$). The p-value of the individual fits is given by p_{fit} and $OM = \log\left(x^{(i)}_{max}/x^{(i)}_{min}\right)$ corresponds to the number of orders of magnitude covered by the fit. The global power-law exponent $\hat{\Gamma}_E$ holds from $\min\{E^{(i)}_{min}\}$ $\max\{E^{(i)}_{max}\}$ and includes $\mathcal{N}$ points. The KSDMD statistic has a unique fit because the fit already exceed $p_c = 0.20$ and no valid fits are found with a smaller p-value.

The values of the particular exponents $\hat{\gamma}_{E,(i)}$ as well as the global exponent $\hat{\Gamma}_E$ of Vycor labquake catalogs differ remarkably from those found for earthquakes and charcoal labquakes [Navas-Portella et al., 2018]. As it was expected from the results obtained in Sec. 6.3, no combination of the cut-offs has lead to a good fit if Vycor labquakes are merged with charcoal labquakes or earthquakes. Therefore, Vycor labquakes can be considered outside this universality class.

7.2.2 Amplitude distribution of Vycor Labquakes

The same procedure for merging labquake catalogs in order to obtain a broader fitting range for the energy can also be performed for the amplitude of Vycor labquakes (See Fig. 7.11). Let us consider that, for the i-th experiment, the random variable $\mathcal{V}$ corresponding to amplitude (i.e., the peak voltage, whose values are denoted by V) follows a truncated continuous power-law distribution,

$$f_{\mathcal{V}}(V)dV = \frac{1-\gamma_{V,(i)}}{V_{\max}^{1-\gamma_{V,(i)}} - V_{\min}^{1-\gamma_{V,(i)}}} V^{-\gamma_{V,(i)}} dV \,. \tag{7.16}$$

However, the true value of $\mathcal{V}$ is not accessible from the experiments, and what one has instead is its discretized counterpart Ψ (the discretized peak voltage, whose values are denoted by ψ), which is concentrated in k discrete values (but not equispaced) of the amplitude A (see Sec. 4.1.3). In fact, the values V that the variable $\mathcal{V}$ can take are the real values of the voltages read by the AD card, but they are transformed into amplitudes A in dB, losing precision.

Under the assumption of a power-law distributed $\mathcal{V}$, one is allowed to state that the variable Ψ has probability mass function

$$f_{\Psi}(\psi)\, d\psi = \mathcal{P}(\Psi = \psi) = \mathcal{P}\left[g(A-\Delta) \leq \mathcal{V} < g(A+\Delta)\right] d\psi \,, \tag{7.17}$$

$$\begin{aligned} f_{\Psi}(\psi)\, d\psi &= \frac{\left[g(A+\Delta)\right]^{1-\gamma_{V,(i)}} - \left[g(A-\Delta)\right]^{1-\gamma_{V,(i)}}}{V_{\max}^{1-\gamma_{V,(i)}} - V_{\min}^{1-\gamma_{V,(i)}}} d\psi \\ &= \frac{2\sinh[2.30(1-\gamma_{V,(i)})\Delta/20]}{V_{\max}^{1-\gamma_{V,(i)}} - V_{\min}^{1-\gamma_{V,(i)}}} \frac{1}{\psi^{\gamma_{V,(i)}-1}} d\psi \,, \end{aligned} \tag{7.18}$$

where $\psi = g(A)$ is the transformation from amplitudes in voltages V to amplitudes in dB A defined in Eq. (4.3), $\Delta = 0.5$ dB is a vicinity around the values of A, $2.30 \simeq \log 10$. Note that, although in practice one can use the expression with the cumulatives in Eq. 7.17, $f_{\Psi}(\psi)$ is a power law but with exponent $\gamma_{V,(i)} - 1$. As a matter of notation, it must be mentioned that Δ (which was defined as the bin-width in a histogram in Sec. 5.3) has been maintained because it can also be understood as half-of the bin-width for the discretized or binned amplitudes.

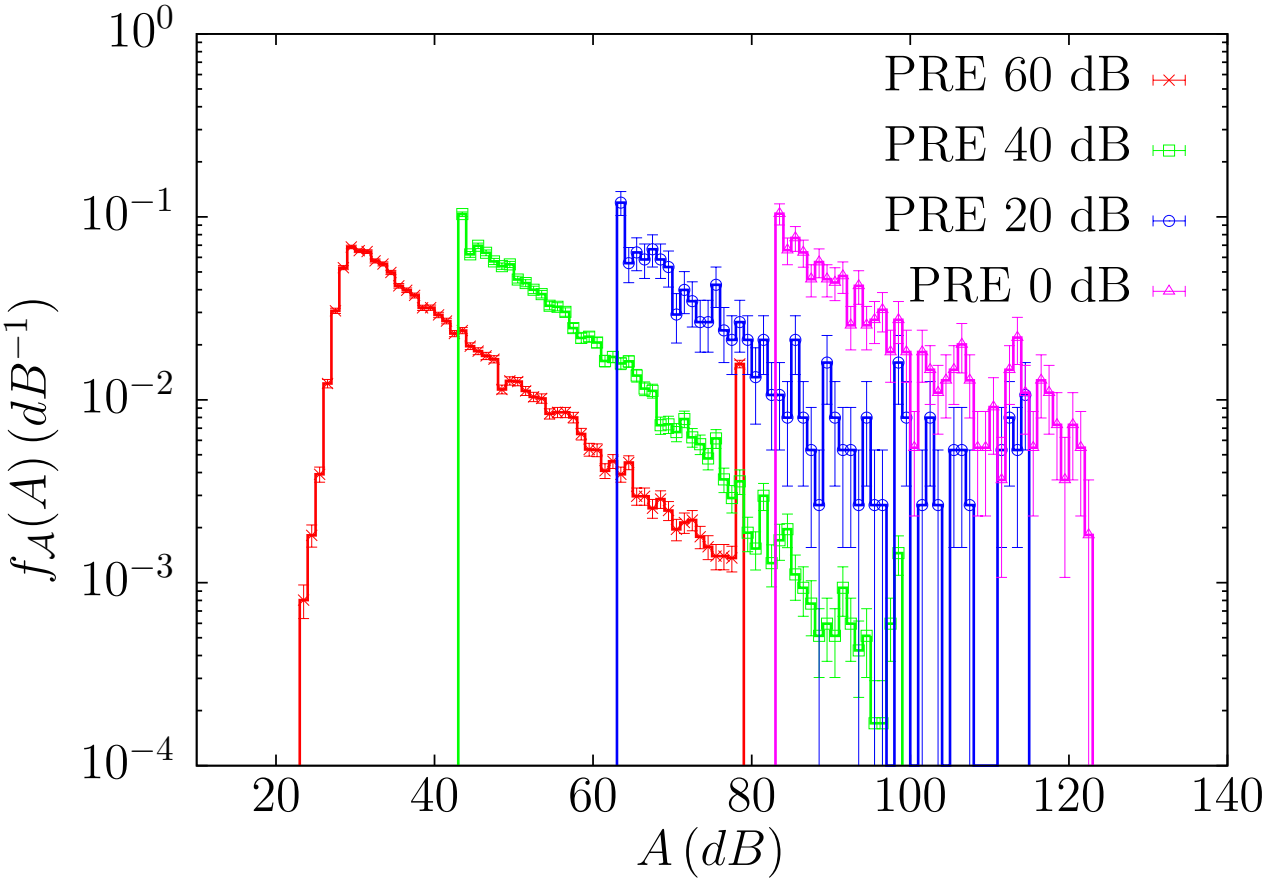

Fig. 7.11: Estimated probability mass functions of the amplitude in dB for the complete datasets of the different experiments performed at different pre-amplifications Preamp and Thresholds THR. Details about the experiments are exposed in Table 4.2. Error bars are estimated as the standard deviation for each bin.

The log-likelihood of the amplitudes for a particular experiment can be written as:

$$\begin{aligned}\log \mathcal{L} &= \sum_{j=A_{\min}}^{A_{\max}} \omega_j \log f_{\Psi}\left(g(j)\right) \\ &= \sum_{j=A_{\min}}^{A_{\max}} \omega_j \left[\log\left(\frac{2\sinh[2.30(1-\gamma_V)\Delta/20]}{V_{\max}^{1-\gamma_V} - V_{\min}^{1-\gamma_V}}\right) - (\gamma_V - 1)\log\psi_j\right],\end{aligned} \tag{7.19}$$

where γ_V is the power-law exponent, and $A_{\min}$ and $A_{\max}$ are the values of the amplitude in dB corresponding to the cutoffs imposed on the sample for the analysis. Note that the index j does not consider all the sample values one by one but the values of the random variable. The frequency ω_j is the number of events with discretized peak voltage $\psi_j = g(j)$. The value of the exponent γ_V that maximizes Eq. (7.19) is computed by means of the function *"optimize"*, which is already implemented in R programming language [R Documentation, 2017; Brent,

2002]. The values of the fitted exponents $\hat{\gamma}_V$ for different values of $A_{\min}$ and $A_{\max}$ are shown in Fig. 7.12 by using MLE exponent maps [Baró and Vives, 2012].

With the aim of merging labquake catalogs for the amplitude, let us write the log-likelihood function of model OneExp as:

$$\begin{aligned}\log \mathcal{L}_{OneExp} &= \sum_{i=1}^{n_{\mathrm{ds}}} \sum_{j=A_{\min}^{(i)}}^{\mathrm{A}_{\max}^{(\mathrm{i})}} \omega_{ij} \log f_{\Psi}\left(g(j)\right) \\ &= \sum_{i=1}^{n_{\mathrm{ds}}} \sum_{j=A_{\min}^{(i)}}^{\mathrm{A}_{\max}^{(\mathrm{i})}} \omega_{ij} \left[\log\left(\frac{2\sinh[2.30(1-\Gamma_V)\Delta/20]}{V_{\max,(\mathrm{i})}^{(i),1-\Gamma_V} - V_{\min,(\mathrm{i})}^{1-\Gamma_V}}\right) - (\Gamma_V - 1)\log\psi_{ij}\right],\end{aligned} \tag{7.20}$$

where Γ_V is the global power-law exponent, and $\mathrm{A}_{\min}^{(\mathrm{i})}$ and $\mathrm{A}_{\max}^{(\mathrm{i})}$ are the values of the amplitude in dB corresponding to the cutoffs imposed on the sample for the analysis. The frequency ω_{ij} is the number of events from the i-th experiment with discretized peak voltage $\psi_{ij} = g(j)$. As it was done for the energy, MLE maps are useful for a first exploratory analysis of the dependence of the power-law exponent with the different cut-offs that can be imposed in the samples. As it can be appreciated in Fig. 7.12, power-law exponents are quite stable along all the detected range of amplitudes except for the sample with pre-amplification of 40 dB. The stability of the power-law exponents as well as the saturation effects allow for the imposition of an upper cut-off that can be set to $A_{max} = A_{top} - 1$. By fixing some of the upper cut-offs, the number of combinations to explore is reduced. However, the only sample for which there is not a clear stability of the exponent is the one performed at 40 dB. The upper cut-off in this sample is considered as a free parameter.

After maximization of the log-likelihood function in Eq. (7.20), next step consists in determining whether the null hypothesis of considering a power-law distribution global exponent $\hat{\Gamma}_V$ is plausible or not. As it was explained in Sec. 7.1.3, two different goodness-of-fit test can be performed. The results of the global fits for the amplitude distributions of Vycor catalogs for the KSDMD and the CKSD statistics are presented in Table 7.8. In both cases, the ranges for which the largest number of orders of magnitude is achieved is the same, with very little difference in p-values. Furthermore, this fit overcomes both p-value thresholds $p_c = 0.05$ and $p_c = 0.20$ for both statistics. Consequently, there is a unique fit which is presented in Table 7.8. Note that each experiment detects labquakes within 2.8 decades in amplitude but all the experiments together would yield a total range of 5.8 decades. This range is broader than other ranges of AE amplitudes [Petri et al., 1994; Vives et al., 1995; Carrillo et al., 1998; Weiss et al., 2000; Koslowski et al., 2004; Richeton et al., 2005; Weiss et al., 2007].

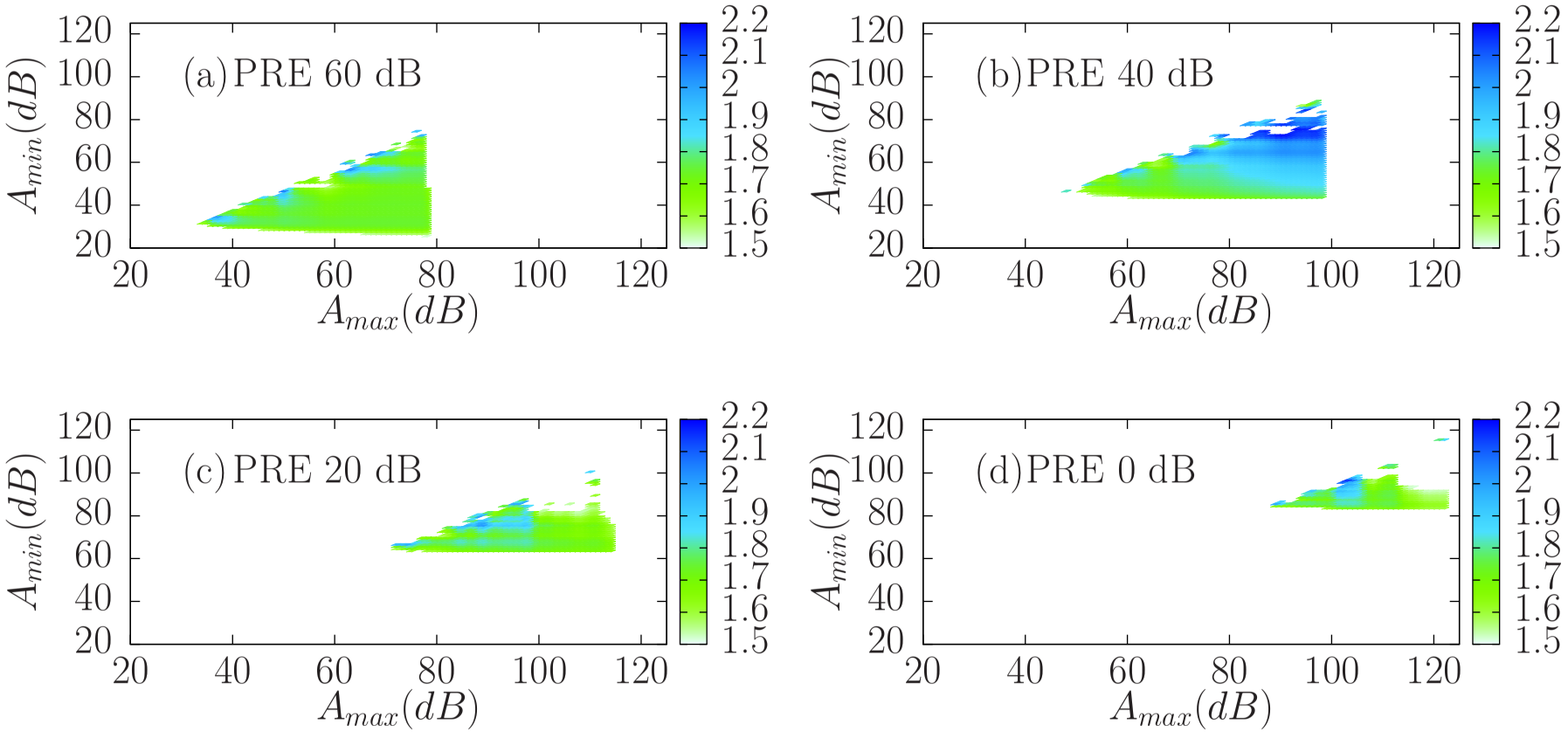

Fig. 7.12: MLE exponent maps for the exponent $\hat{\gamma}_V$ corresponding to the distribution of amplitudes in dB for the four experiments in Vycor performed at different pre-amplifications 60, 40, 20 and 0 dB ((a), (b), (c) and (d) panels respectively).

The global fit yields a global exponent $\hat{\Gamma}_V = 1.740 \pm 0.006$ with a $p_{value}^{(C)} = 0.36$ for $\mathcal{N} = \sum_{i=1}^{n_{\rm ds}} n_i = 31441$ events. Note that the value of the global exponent is in agreement with the weighted harmonic mean

$$\hat{\Gamma}_V = 1.740 \simeq 1 + \frac{\mathcal{N}}{\sum_{i=1}^{n_{\rm ds}} \frac{n_i}{\hat{\gamma}_{V,(i)}-1}} = 1.741,$$

see Sec. 7.1.1 for a justification of this result. Due to the upper-truncations, the value of the global exponent is not exactly the same as the value of the harmonic mean of the particular exponents.

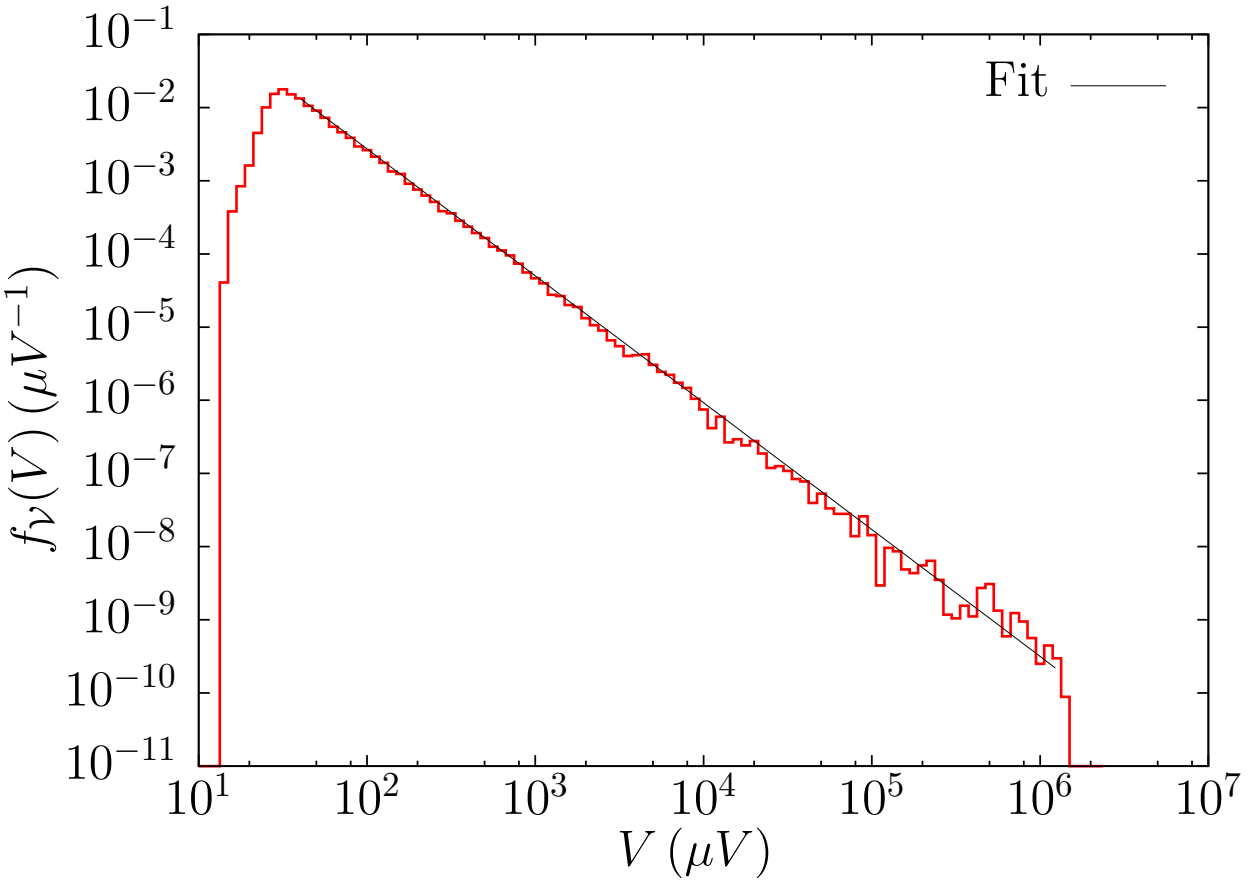

Fig. 7.13: Merged amplitude PDF of the global distribution with exponent $\alpha_g = 1.740 \pm 0.006$, p-value$= 0.36$, and number of fitted data $\mathcal{N} = 31441$. Events corresponding to the experiments performed at 40, 20, and 0 dB below their respective lower cut-offs as well as those corresponding to the experiments performed at 60, 40, and 20 dB above their respective upper cut-offs have been excluded. Black solid line shows the fit of a truncated power-law with exponent $\hat{\alpha}_g = 1.740$ and ranging almost 5 decades, from $g(32 - \Delta) = 37.58$ μV to $g(122 + \Delta) = 1.33 \times 10^6$ μV.

Similarly to the simulations performed in Sec. 7.1.4, four synthetic catalogs have been merged with the same parameters as in Table 7.8, yielding acceptable p-values. It must be mentioned that performing the particular fits by simply assuming that the discrete variable Ψ directly follows a truncated continuous power-law leads to the rejection of this hypothesis in the goodness-of-fit test.

Figure 7.13 shows the global PDF for the amplitudes and the global fit. Observe how the global exponent is valid along 4.5 orders of magnitude, giving an unprecedented broad fitting-range in amplitudes. The procedure to construct this aggregated histogram is explained in Sec. 7.1.2. As the estimation of the PDF is done using bins [Deluca and Corral, 2013], note that one can safely replace the unknown values of the random variable $\mathcal{V}$ by the known discretized values of Ψ. The only requirement is that the width of the bins is not smaller than the discretization of Ψ.

The same reasoning which was carried out for labquake energies can be also applied in this case in order to determine the existence of a corner value. According to the range of detection for the experiment performed at 0 dB, one could be able to observe events up to 139 dB. Nevertheless, the maximum in this sample corresponds to 123 dB. Under the hypothesis that a power-law distribution with the same exponent ($\hat{\gamma}_A = 1.61 \pm 0.04$) can be extended for larger values of the amplitude, the probability $\mathcal{P}\,(123\text{dB} < \text{A} < 140\text{dB})$ turns out to be 0.042. The number of observed events whose amplitude is larger than $A_{min} = 84$ dB is 529. For these $N_i = 529$ trials, the probability of having no events in this range can be estimated by $(1 - 0.042)^{529} = 1.4 \times 10^{-10}$. However, the same calculation for the experiment at 20 dB gives a probability of having no events above A_{top} of 0.11, which is not a negligible value at all. As it occurred for the energy, one could justify the existence of a corner value due to the finite size of the sample properties [Serra and Corral, 2017].

Once both amplitude and energy distributions have been fitted with global exponents holding for several decades, a relationship between amplitudes and energies can be established. Scatter plots for the complete datasets are shown in Fig. 7.14. Rectangular windows with the same color as the points of a single experiment correspond to the ranges in which both amplitudes and energies are power-law distributed. Saturation effects are clear and are left outside of these windows which conform global fits. Bak et al. [1988] proposed a relationship between exponents of power-law distributed avalanche observables.

Let us assume that the energy of labquakes is related to their amplitude by a relationship of this kind, $E\,(A) \sim A^{\varrho}$, where $\sim$ denotes proportionality and ϱ is an exponent establishing thus a scaling relationship between energy and amplitude. A rough estimation of this exponent can be easily found by applying:

$$\varrho = \frac{1 - \hat{\Gamma}_V}{1 - \hat{\Gamma}_E} = \frac{1 - 1.74}{1 - 1.35} = 2.11 \simeq 2\,. \tag{7.21}$$

PRE(dB)	$\hat{\gamma}_V$	$A_{\min}$ (dB)	A_{low} (dB)	$A_{\max}$ (dB)	A_{top}(dB)	OM	n	N	p_{value}
60	1.743(7)	32	23	78	79	2.3	21414	28614	0.921(6)
40	1.75(1)	46	43	72	99	1.3	9146	11717	0.20(1)
20	1.67(4)	64	63	114	115	2.5	353	376	0.50(2)
0	1.61(4)	84	83	122	123	1.9	528	548	0.64
Model OneExp	$\hat{\Gamma}_V$					$\sum OM$	$\mathcal{N}$		
$2\mathcal{R}_e = 5.728$	1.740(6)	32	23	122	123	8	31441	41255	p_{value}
							$d_e^{(C)}$	4.091	0.36
							$d_e^{(M)}$	4.72×10^{-3}	0.40

Table 7.8: Results of fitting models OneExp and MultiExp to the Vycor Labquake catalogs when doing the goodness-of-fit test with the KSDMD and the CKSD statistics. $\hat{\gamma}_V$ corresponds to the fitted exponent in the range $[A_{\min}, A_{\max}]$ whereas $\hat{\Gamma}_V$ corresponds to the fitted global exponent. Numbers in parentheses correspond to the errors in the estimations. The error of the exponent is computed as the standard deviation of the MLE. n is the number of data entering into the fit, N is the total number of events for a given dataset and $\mathcal{N}$ is the total number of data of the global fit. The number of orders of magnitude $OM = \log_{10}\left(\frac{A_{max}^{(i)}}{A_{min}^{(i)}}\right)$ covered by each fit as well as for model 1 is also shown. Value of the LRT statistic $2\mathcal{R}_e$ is provided in this case (recall that, the critical value in this case corresponds to the percentile 95 of a χ^2_3 distribution with 3 degrees of freedom $2\mathcal{R}_c = 7.81$). The same fits are obtained for $p_c = 0.05$ and $p_c = 0.20$.

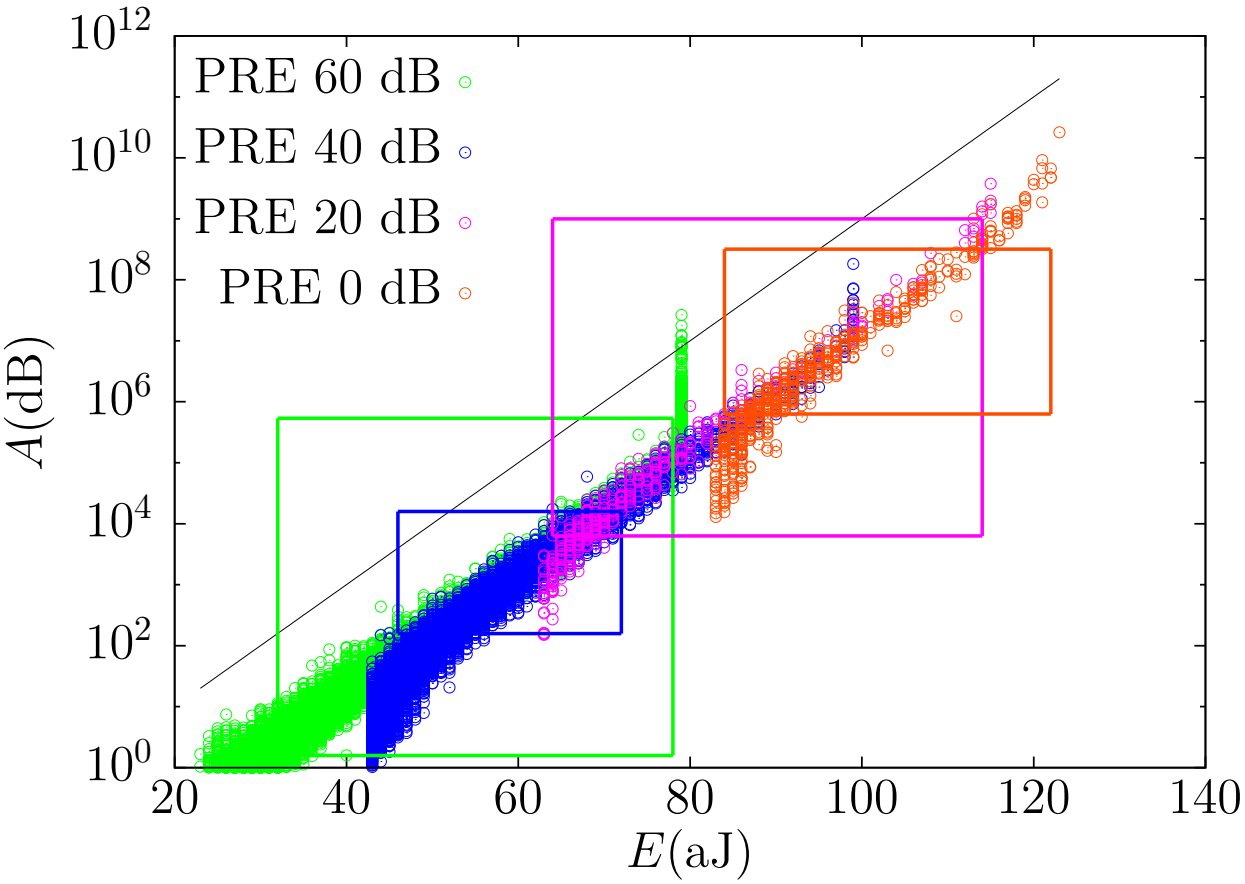

Fig. 7.14: Scatter plot for the complete datasets energies and amplitudes. Black solid line with slope 2 serves as a guide to eye to illustrate the relation $E \sim A^2$. Rectangular windows with the same color code as the point correspond to those ranges in which amplitude and energy are power-law distributed for that particular experiment and conform the global fit.

The relationship in Eq. 7.21 is illustrated in Fig.7.14 where a line with slope 2 is plotted as a guide to the eye. More accurate studies must be performed in order to give a rigorous mathematical sense to this kind of relationships.

7.3 Merging earthquake catalogs: Gutenberg-Richter law

As it was explained in Sec. 3.2.2.1, the GR law states that the seismic moment can be considered as a random variable $\mathcal{M}$ which is power-law distributed:

$$f_{\mathcal{M}}(M)dM = \frac{2}{3}\frac{b}{M_{min}}\left(\frac{M}{M_{min}}\right)^{-\left(1+\frac{2}{3}b\right)} dM = \frac{\gamma_M - 1}{M_{min}}\left(\frac{M}{M_{min}}\right)^{-\gamma_M} dM\,. \tag{7.22}$$

In its usual form, the GR law fits a power-law model that contemplates a unique power-law exponent. Nevertheless, several studies have elucidated the existence of a double-power-law behaviour in the GR law for global seismicity [Corral and González, 2019; Pacheco et al., 2012; Yoder et al., 2012]. Corral and González [2019] claim that a truncated power-law with exponent $\gamma \simeq 1.66$ cannot be rejected up to $m_{max} \simeq 7.4$ and a second power-law tail emerges from $m'_{min} = 7.67$ with an exponent $\gamma'_M = 2.1 \pm 0.1$. These results have been checked by using the residual CV-test (see Fig. 7.15). Furthermore, by fixing the upper truncation at $m_{max} = m'_{min} = 7.67$, the truncated power-law hypothesis cannot be rejected (see Table 7.9). Consequently, if one wants to fit a power-law PDF with a unique exponent, one has to exclude all those earthquakes with $m \geq m_{max} = 7.67$. For the CMT catalog, the upper cut-off is fixed at $M_{max} = 10^{\frac{3}{2}m_{max}+9.1}$ whereas the rest of catalogs exposed in Sec. 4.2.3 can be safely fitted by untruncated power-law PDFs because the CV-test does not reject the hypothesis of a unique power-law tail and the magnitudes which are studied are considerably smaller than those in the CMT catalog.

Therefore, two untruncated power-law distributions and a third one which is truncated for the CMT catalog are considered in order to merge catalogs. For each decade, 5 values of $M^{(i)}_{min}$ equally spaced in logarithmic scale are sampled, and all the possible combinations of cut-offs $M^{(1)}_{min}$, $M^{(2)}_{min}$, $M^{(3)}_{min}$ are checked for a fixed upper-truncation $M^{(3)}_{max}$. The labels (1), (2) and (3) correspond to the catalogs YSH B, YSH A and CMT respectively.

	N	n	m_{min}	M_{min} **(Nm)**	m_{top}	M_{top}**(Nm)**	$\hat{b}$**-value**	$\hat{\gamma}_M$	p_{fit}
(1) YHS B	26330	3412	1.93	10^{12}	5.39	1.53×10^{17}	0.99(3)	1.66(2)	0.072(8)
(2) YHS A	152924	4353	3.17	7.08×10^{13}	7.20	7.94×10^{19}	0.98(1)	1.65(1)	0.080(9)
(3) CMT	48637	22336	5.33	1.24×10^{17}	9.08	5.25×10^{22}	0.982(7)	1.655(5)	0.36(2)

Table 7.9: Results of fitting the GR law for each individual catalog. The total number of earthquakes in each catalog is given by N whereas the number of data entering into the fit is n. The value m_{top} corresponds to the maximum observed value for each catalog. The GR law is valid for each catalog from $[m_{min}, m_{max}]$ with a particular b-value. m_{max} has no upper-limit for any of the fits except for the CMT catalog, in which $m^{(3)}_{max} = 7.67$ ($M^{(3)}_{max} = 4.03 \times 10^{20}$ Nm). Numbers in parentheses correspond to the error bar estimated with one σ in the scale given by the last digit. The p-value of the fits has been computed with 10^3 simulations and $p_c = 0.05$. For completeness, values of cut-offs in seismic moment M and power-law exponents $\hat{\gamma}$ have been included.

In Table 7.10 the results of the global fit for models OneExp and MultiExp are shown. The values of the statistics as well as the resulting p-values are shown in Table 7.11. It can be observed that, for this particular set of cut-offs and as it happened with the synthetic datasets in Sec. 7.1.4, the CKSD test has a smaller p-value and can be considered to be a more powerful statistic than the KSDMD. A b-value very close to one holds for more than 8 orders of magnitude in seismic moment from $m_{min} = 1.93$ to $m_{max} = 7.67$. The value of the global exponent is approximately in agreement with the harmonic mean of the particular exponents of the GR-law for each catalog [Navas-Portella et al., 2018; Kamer and Hiemer, 2015]. Due to the upper-truncation, the value of the global exponent is not exactly the same as the value of the harmonic mean of the particular exponents. The results for the CKSD statistic do not show remarkable differences if the critical p-value is set to $p_c = 0.20$ (see Table 7.10). The definitive fit is the one whose goodness-of-fit test has been performed with a threshold value of $p_c = 0.20$.

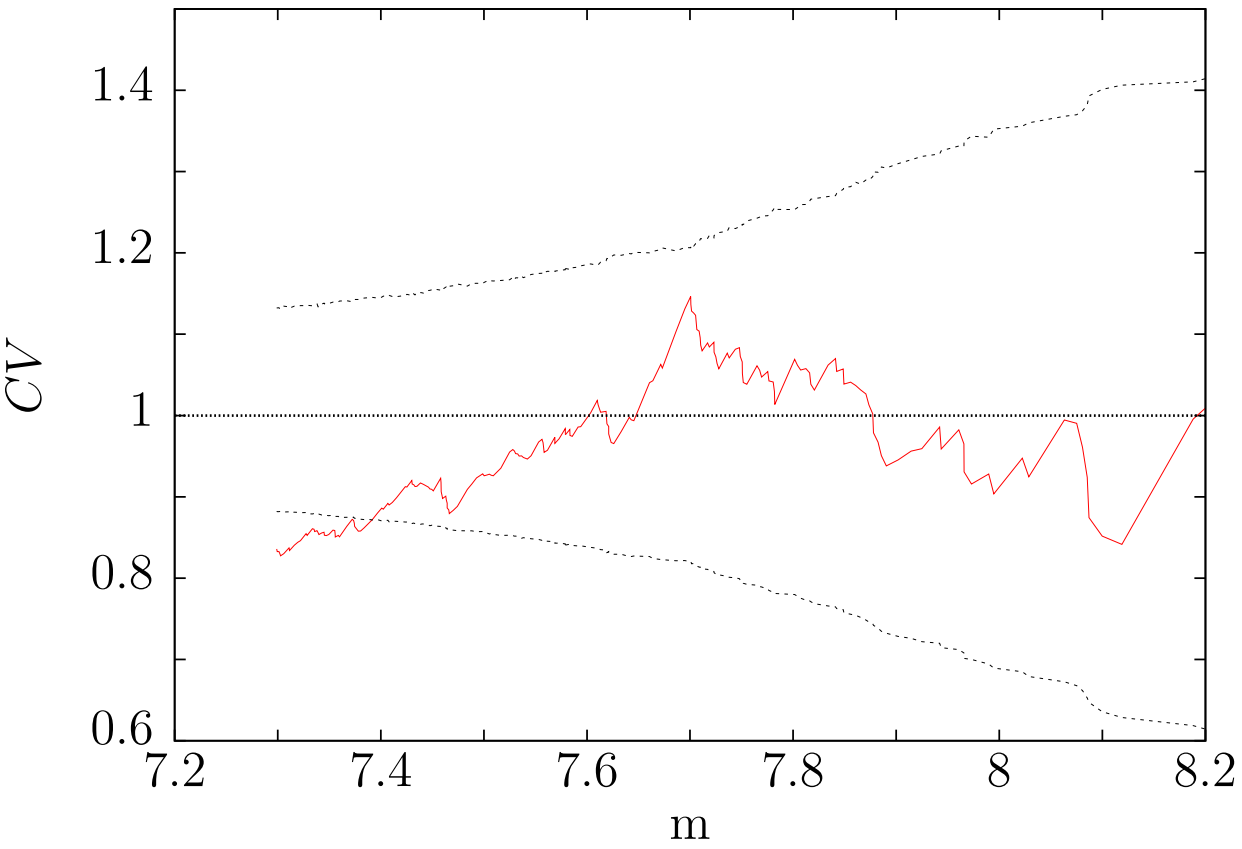

Fig. 7.15: Values of the residual coefficient of variation as a function of the moment magnitude m for the CMT catalog. Dotted lines correspond to 95% confidence intervals computed for 10^4 simulations.

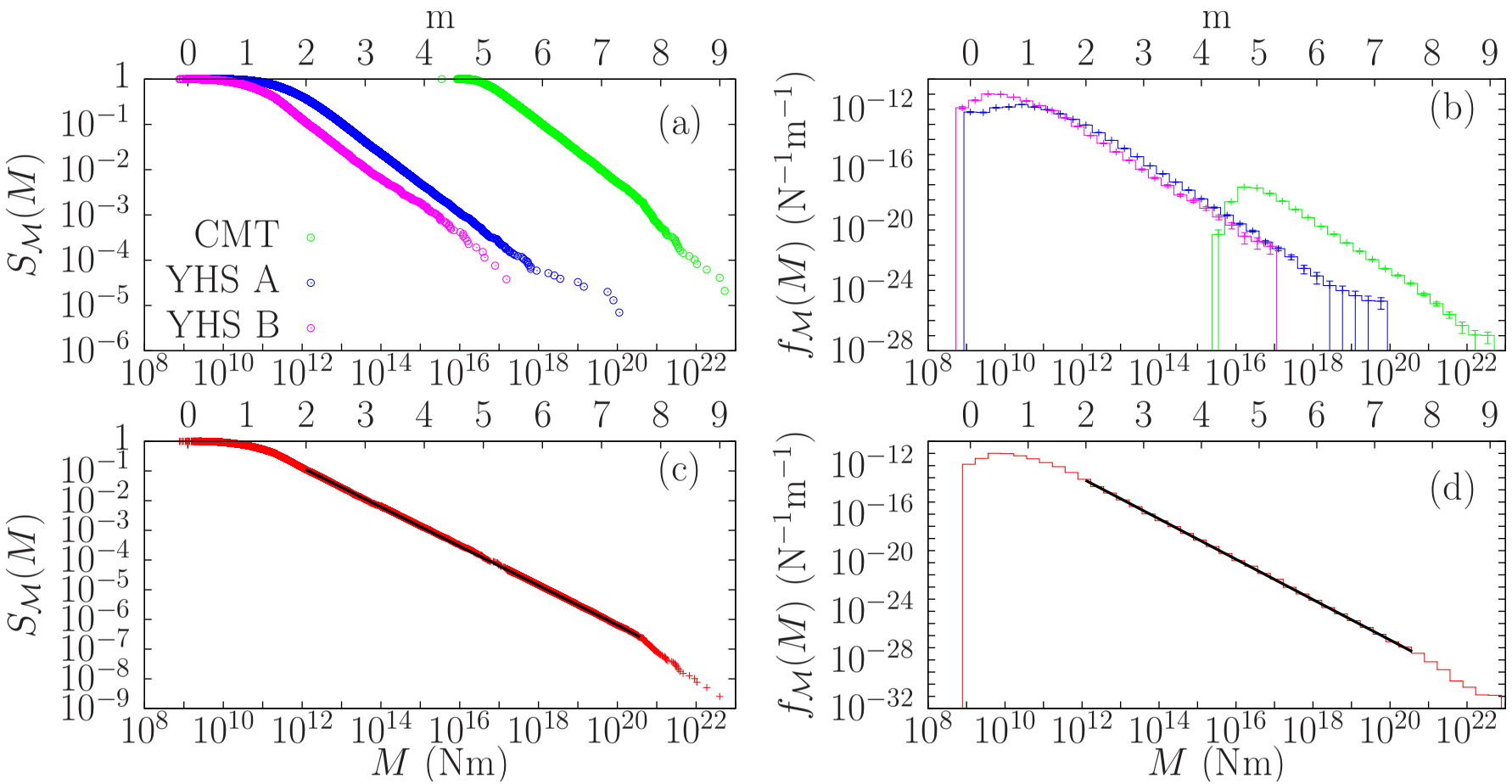

Fig. 7.16: Estimated Complementary Cumulative Distribution Functions ($S_{\mathcal{X}}$) of the Gutenberg-Richter law for each catalog (a) and for the merged catalogs (c). Estimated PDFs $f(M)$ of the Gutenberg-Richter law for each catalog (b) and for the merged catalogs (d). Merged histogram is plot by following the procedure explained in Ref. [Navas-Portella et al., 2018]. Fits are represented by solid black lines. Top axis represents the same scale in moment magnitude. . Note that earthquakes from the CMT and YSH A catalogs below their respective lower cut-offs $x_{min}^{(i)}$ have been excluded in the bottom plots.

$p_c = 0.05$	n	M_{min}	M_{min} **(Nm)**	**OM**	$\hat{b}$**-value**	$\hat{\gamma}_M$	p_{value}
(1) YHS B	3412	1.93	10^{12}	5.18	0.99(1)	1.66(2)	0.072(8)
(2) YHS A	3500	3.27	10^{14}	5.90	0.99(2)	1.66(1)	0.089(9)
(3) CMT	19003	5.40	1.58×10^{17}	3.40	0.98(8)	1.655(5)	0.26(1)
Model OneExp	$\mathcal{N}$			$\sum$**OM**	$\hat{b}_g$**-value**	$\hat{\Gamma}_M$	p_{value}
$2\mathcal{R}_e = 1.337$							
	25915	1.93	10^{12}	14.48	0.991(6)	1.661(4)	0.079(9)

$p_c = 0.20$	n	M_{min}	M_{min} **(Nm)**	**OM**	$\hat{b}$**-value**	$\hat{\gamma}_M$	$p_{value}^{(C),0.05}$
(1) YHS B	3412	1.93	10^{12}	5.18	0.99(1)	1.66(2)	0.072(8)
(2) YHS A	3500	3.27	10^{14}	5.90	0.99(2)	1.66(1)	0.089(9)
(3) CMT	10422	5.67	3.98×10^{17}	3	1.00(1)	1.663(7)	0.62(2)
Model OneExp	$\mathcal{N}$			$\sum$**OM**	$\hat{b}_g$**-value**	$\hat{\Gamma}_M$	$p_{value}^{(C),0.20}$
$2\mathcal{R}_e = 1.124$							
	17334	1.93	10^{12}	14.08	1.000(8)	1.667(5)	0.326(5)

Table 7.10: Results of fitting the models OneExp and MultiExp to the earthquake catalogs when doing the goodness-of-fit test with the CKSD statistis and different values of p_c. If the goodness-of-fit test is performed with the KSDMD statistic, the same set of cut-offs as the fit done by using $d_e^{(C),0.20}$ is found no matter the value of p_c (see Table 7.11). Same symbols as in Table 7.9. Values of the LRT statistic $2\mathcal{R}_e$ are provided in this case (recall that, the critical value in this case corresponds to the percentile 95 of a χ^2_2 distribution with 3 degrees of freedom $2\mathcal{R}_c = 5.99$). The value $M_{max}^{(i)}$ is replaced for $M_{top}^{(i)}$ for untruncated fits. Note that $7.67 - 1.93 = 5.74$ units in magnitude correspond to 8.6 orders of magnitude in seismic moment.

	d_e	p_{value}
$d_e^{(M),0.20}$	0.018487	0.20(1)
$d_e^{(C),0.05}$	3.25082	0.079(9)
$d_e^{(C),0.20}$	2.778202	0.326(5)

Table 7.11: Results of fitting model OneExp to all the catalogs when doing the goodness-of-fit test with the KSDMD and the CKSD statistics for $p_c = 0.05$ (same cut-offs for all the catalogs) and for the set of cut-offs for which the p-value is larger or equal than $p_c = 0.20$ for the CKSD statistic. KSDMD has a unique fit because the fit already exceed $p_c = 0.20$ and no valid fits are found with a smaller p-value.

As far as the author knows, this is the broadest fitting range that has been found for the Gutenberg-Richter law for earthquakes with a unique value of the exponent [Kagan, 2010]. There are catalogs of tiny mining-induced earthquakes which exhibit a much smaller completeness magnitude [Plenkers et al., 2011] than the ones of natural seismicity used in this work. They were not considered here because they are not currently public and show b-values significantly different [Kwiatek et al., 2010] from the one found here, which would result in non-acceptable fits when merging them with the rest of catalogs, possibly pointing to a different universality class.

7.4 Merging earthquakes and charcoal labquake catalogs

Motivated by the fact that the power-law exponents of earthquake catalogs and charcoal labquakes are very similar (see Tables 7.9 and 6.5), the methodology exposed of merging datasets explained in Sec. 7.1 is applied by adding also the catalog of charcoal labquakes.

As discussed in Sec. 3.2.2.1, it is important to stress the fact that the seismic moment does not correspond with the energy radiated E_r by the earthquake, which would be the reasonable energy to compare with the AE energy. For this study, the ratio of seismically radiated energy over the seismic moment is considered as constant so that the values of the seismic moment should just be multiplied by an unique factor. The value of this unique factor is $\frac{E_r}{M} = 10^{-4.6}$ [Bormann, 2015], where M is the seismic moment. In this case, E_r corresponds to the energy radiated in seismic waves by earthquakes and the AE energy. Let us remark that the measured AE energy is the part of the radiated energy which is collected by the transducer. In this sense, let us assume that the total energy radiated by a labquake is of the same order of magnitude as this detected energy.

It can be shown that, for both models OneExp and MultiExp, multiplying the variable by a constant factor only introduces a constant term in the log-likelihood that does not change neither its maximum nor the difference of the log-likelihoods.

Let us suppose n_{ds} datasets and that data in the i-th dataset is described, for simplicity and without losing generality, in terms of untruncated power-law PDFs. The log-likelihood function for both OneExp and MultiExp models are:

$$\begin{aligned} \log \mathcal{L}_{OneExp} &= \sum_{i=1}^{n_{ds}} \sum_{j=1}^{n_i} \log f_{\mathcal{X}}\left(x_{ij}\right) \\ &= \sum_{i=1}^{n_{ds}} \sum_{j=1}^{n_i} \log \left(\frac{\Gamma - 1}{x_{min}^{(i)1-\Gamma}} x_{ij}^{-\Gamma} \right) . \end{aligned} \tag{7.23}$$

$$\begin{aligned} \log \mathcal{L}_{MultiExp} &= \sum_{i=1}^{n_{ds}} \sum_{j=1}^{n_i} \log f_{\mathcal{X}}\left(x_{ij}\right) \\ &= \sum_{i=1}^{n_{ds}} \sum_{j=1}^{n_i} \log \left(\frac{\gamma_i - 1}{x_{min}^{(i)1-\gamma_i}} x_{ij}^{-\gamma_i} \right) . \end{aligned} \tag{7.24}$$

Let us now suppose that the values taken by the variable $\mathcal{X}$ in the i-th dataset are multiplied by a certain factor λ_i as well as the lower cut-off $x_{min}^{(i)}$:

$$x_{ij} \to \lambda_i x_{ij} \, ,$$

$$x_{min}^{(i)} \to \lambda_i x_{min}^{(i)} \, .$$

Note that this is not strictly speaking a scale transformation since not only the variable but some parameters characterizing the function are changed. The likelihoods of the transformed data $\mathcal{L}^T$ for both models are:

$$\begin{aligned}\log \mathcal{L}^T_{OneExp} &= \sum_{i=1}^{n_{ds}} \sum_{j=1}^{n_i} \log \left(\frac{\Gamma - 1}{\left(\lambda_i x^{(i)}_{min}\right)^{1-\Gamma}} \left(\lambda_i x_{ij}\right)^{-\Gamma} \right) \\ &= \log \mathcal{L}_{OneExp} - \sum_{i=1}^{n_{ds}} n_i \log \lambda_i \,,\end{aligned} \tag{7.25}$$

$$\begin{aligned}\log \mathcal{L}^T_{MultiExp} &= \sum_{i=1}^{n_{ds}} \sum_{j=1}^{n_i} \log \left(\frac{\gamma_i - 1}{\left(\lambda_i x^{(i)}_{min}\right)^{1-\gamma_i}} \left(\lambda_i x_{ij}\right)^{-\gamma_i} \right) \\ &= \log \mathcal{L}_{MultiExp} - \sum_{i=1}^{n_{ds}} n_i \log \lambda_i \,.\end{aligned} \tag{7.26}$$

The transformed likelihoods can be expressed in terms of the likelihood functions of the original data together with an additive constant term that does not depend neither on the exponent nor the cut-off. This means that maximum of the transformed log-likelihood is attained at exactly the same value as the maximum of the original one. Given that the likelihood-ratio statistic is a subtraction of the likelihood of both models, the relation between the original statistic $2\mathcal{R}$ and the one corresponding to transformed likelihood functions is:

$$\begin{aligned}2\mathcal{R}^T &= 2\left(\log \mathcal{L}^T_{MultiExp} - \log \mathcal{L}^T_{OneExp}\right) \\ &= 2\left(\log \mathcal{L}_{MultiExp} - \log \mathcal{L}_{OneExp}\right) = 2\mathcal{R} \,.\end{aligned} \tag{7.27}$$

Since the two additional terms in the transformed likelihoods are the same for both models, they cancel and the statistic for the likelihood-ratio test is invariant under this transformation. As the CKSD statistic is a weighted average of the particular KS distances of each dataset, it does not change neither. Therefore, the results shown in Table 7.9 as well as those resulting when merging catalogs would not change except for the values of the cut-offs.

For each decade, 5 different values of $x^{(i)}_{min}$ equally spaced in logarithmic scale, for a fixed upper-truncation $x^{(3)}_{max}$ are checked and all the possible combinations of cut-offs $x^{(0)}_{min}$, $x^{(1)}_{min}$, $x^{(2)}_{min}$, $x^{(3)}_{min}$ are checked for a fixed upper-truncation $x^{(3)}_{max}$. The labels (0), (1), (2) and (3) correspond to the catalogs of the charcoal experiment, YSH B, YSH A and CMT respectively.

The results of the global fit for the CKSD statistic are presented in Table 7.12. In this case, not all the catalogues overlap each other and the CKSD statistic is the only one that can be used for the goodness-of-fit test. The value of the global exponent is approximately in

agreement with the harmonic mean of the particular exponents of the GR-law for each catalog [Navas-Portella et al., 2018]. The results do not show remarkable differences if the critical p-value is set to $p_c = 0.20$ (see Table 7.12). The definitive fit is the one whose goodness-of-fit test has been performed with a threshold value of $p_c = 0.20$. These results are shown in Fig. 7.17 and are compatible with the ones obtained from synthetic catalogs shown in Sec. 7.1.4.

Given that there is no overlapping between the charcoal catalog and the earthquakes' ones, an alternative procedure must be applied in order to construct the histogram of the estimated global PDF. First of all, the three earthquake catalogs are merged according to the procedure exposed in Sec. 7.1.2. The earthquakes from the CMT catalog which are above the upper cut-off $x^{(3)}_{max}$ are also plotted. Earthquakes from the YSH B, YSH A and CMT catalogs which are below $x^{(1)}_{min}$, $x^{(2)}_{min}$ and $x^{(3)}_{min}$ are respectively exlcuded. This piece of the histogram, which is not normalized yet, is denoted as Ea, from Earthquakes. The charcoal catalog histogram is plotted by following the procedures exposed in Sec. 5.3 without normalizing by any number of events yet. The piece of histogram corresponding to charcoal is denoted as La, from Labquakes. Given that no normalization has been performed yet, at this point of the procedure, the pieces (Ea) and (La) are not aligned in logarithmic scale. Consequently, the

$p_c = 0.05$	n	m_{min}	$E_{r,min}$ **(Nm)**	**OM**	$\hat{b}$-value	$\hat{\gamma}_{E_r}$	p_{fit}
(0) Charcoal	15906	-	6.31×10^{-18}	6.47	0.988(8)	1.658(5)	0.15(1)
(1) YHS B	1353	2.33	9.99×10^{7}	4.58	0.98(3)	1.66(2)	0.10(1)
(2) YHS A	234	4.47	1.59×10^{11}	4.10	0.98(6)	1.65(4)	0.62(2)
(3) CMT	7689	5.80	1.59×10^{13}	2.80	1.00(1)	1.667(9)	0.393(5)
Model OneExp	$\mathcal{N}$			$\sum$**OM**	$\hat{b}_g$**-value**	$\hat{\Gamma}_{E_r}$	$p^{(C),0.05}_{value}$
$2\mathcal{R}_e = 4.794$							
$d^{(C)}_e = 3.590$	25182		6.3×10^{-18}	17.95	1.003(6)	1.669(4)	0.057(7)

$p_c = 0.20$	n	m_{min}	$E_{r,min}$ **(Nm)**	**OM**	$\hat{b}$-value	$\hat{\gamma}_{E_r}$	p_{fit}
(0) Charcoal	3555	-	6.31×10^{-17}	5.47	1.04(2)	1.69(1)	0.88(1)
(1) YHS B	1007	2.47	1.59×10^{8}	4.38	0.99(3)	1.66(2)	0.058(7)
(2) YHS A	234	4.47	1.59×10^{11}	4.10	0.98(6)	1.65(4)	0.62(2)
(3) CMT	3014	6.20	6.33×10^{13}	2.20	1.00(2)	1.67(2)	0.59(2)
Model OneExp	$\mathcal{N}$			$\sum$**OM**	$\hat{b}_g$**-value**	$\hat{\Gamma}_{E_r}$	$p^{(C),0.20}_{value}$
$2\mathcal{R}_e = 4.184$							
$d^{(C)}_e = 2.998$	7810		6.3×10^{-17}	16.15	1.03(1)	1.688(8)	0.21(1)

Table 7.12: Results of fitting models OneExp and MultiExp to the charcoal and earthquake datasets for two different values of p_c. Same notation as in previous tables. Values of the LRT statistic $2\mathcal{R}_e$ are provided in this case (recall that, the critical value in this case corresponds to the percentile 95 of a χ^2_3 distribution with 3 degrees of freedom $2\mathcal{R}_c = 7.81$). Note that E_r represents the radiated energy in seismic waves by earthquakes and the radiated AE energy.

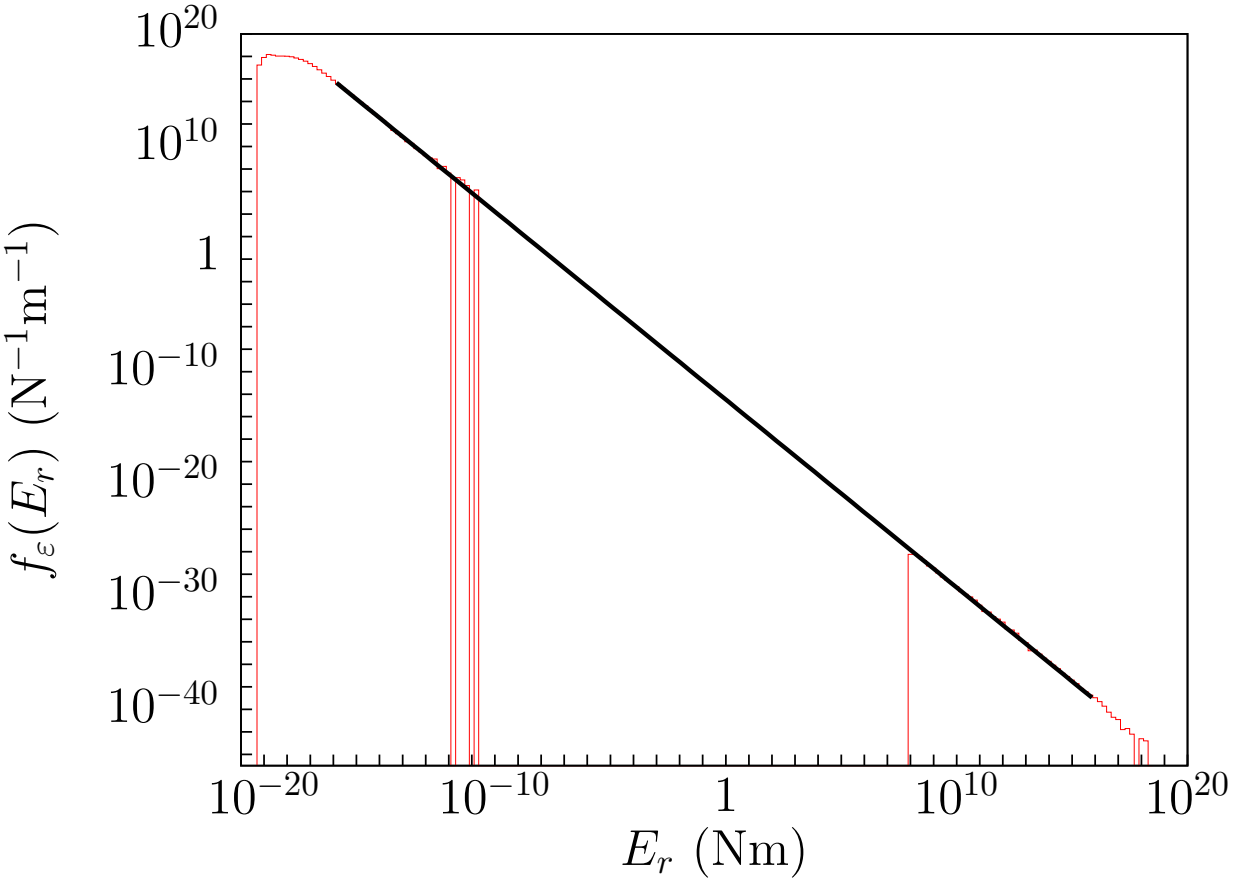

Fig. 7.17: Estimated PDF of the radiated energy E_r for the merged earthquake and charcoal labquake catalogs. The fit is represented by a solid black line. The methodology to construct the histogram with the three earthquake catalogs is the same as the one explained in Sec. 7.1.2, whereas the addition of the charcoal histogram to this fit has been done by conveniently rescaling both parts (those corresponding to charcoal and earthquakes respectively). Each part has been divided by an effective number of events by assuming that the probability in each part corresponds to that obtained from a global power-law exponent with exponent $\hat{\Gamma}$ from $x_{min}^{(0)}$ to $x_{max}^{(3)}$. Note that events from the CMT, YSH A and YSH B catalogs below their respective lower cut-offs x_{min} have been excluded in the plot.

fit of a global power-law with exponent $\hat{\Gamma}$ from $x_{min}^{(0)}$ to $x_{max}^{(3)}$ will not overlap with both pieces.

Note that one could be tempted to use the numerical procedure for merging histograms in Sec. 7.1.2 in this case where these pieces do not overlap each other. Nevertheless, the procedure would assign a certain number of counts to the big gap between the two pieces that conform the histogram thus blurring them and weakening the important message: the big difference in energy scales.

In order to align both pieces, one has to assume that data from $x_{min}^{(0)}$ to $x_{max}^{(3)}$ follows a power-law distribution with exponent $\hat{\Gamma}$. Given that this assumption is supported by statistical results that have been already found, it can be imposed that the probability represented by

each piece corresponds to this theoretical probability. In order to achieve this imposition, the total number of counts in each piece is divided by the effective number of events in that piece:

$$F_{\mathcal{X}}\left(x=x_{top}^{(0)};x_{min}^{(0)},x_{max}^{(3)},\hat{\Gamma}\right)-F_{\mathcal{X}}\left(x=x_{min}^{(0)};x_{min}^{(0)},x_{max}^{(3)},\hat{\Gamma}\right)=\frac{\mathcal{C}\left(x_{min}^{(0)}\leq x\leq x_{top}^{(0)}\right)}{n_{eff}^{La}}\,,$$

$$F_{\mathcal{X}}\left(x=x_{max}^{(3)};x_{min}^{(0)},x_{max}^{(3)},\hat{\Gamma}\right)-F_{\mathcal{X}}\left(x=x_{min}^{(1)};x_{min}^{(0)},x_{max}^{(3)},\hat{\Gamma}\right)=\frac{\mathcal{C}\left(x_{min}^{(1)}\leq x\leq x_{max}^{(3)}\right)}{n_{eff}^{Ea}}\,.$$

where $\mathcal{C}()$ corresponds to the number of counts in the interval considered inside the parenthesis, $F_{\mathcal{X}}$ corresponds to the CDF of a truncated power-law (see Eq. (5.26)) $F_{\mathcal{X}}\left(x=x_{min}^{(0)};x_{min}^{(0)},x_{max}^{(3)},\hat{\Gamma}\right)=0$ and $F_{\mathcal{X}}\left(x=x_{max}^{(3)};x_{min}^{(0)},x_{max}^{(3)},\hat{\Gamma}\right)=1$. Provided that $x_{min}^{(0)}$, $x_{top}^{(0)}$, $x_{min}^{(1)}$, $x_{top}^{(1)}$ as well as the observed number of counts in these ranges are known, the effective number of counts n_{eff}^{Ea} and n_{eff}^{La} can be easily isolated from the last expressions. Note that these numbers are integers for the (La) piece, but not for the (Ea) given that it has been constructed by merging several datasets and an event does not contribute necessarily one unity (see Sec. 7.1.2). By dividing the piece (Ea) by n_{eff}^{Ea} and the piece (La) by n_{eff}^{La}, the histograms are aligned and overlapping with the theoretical fit. However, given that data outside the fit is also considered in the construction of the histogram, the integral of both pieces (Ea) and (La) will not be one and the effective area A_{eff} at this point corresponds to:

$$A_{eff}=\frac{\mathcal{C}\left(x_{low}^{(0)}\leq x\leq x_{top}^{(0)}\right)}{n_{eff}^{(La)}}+\frac{\mathcal{C}\left(x_{min}^{(1)}\leq x\leq x_{top}^{(3)}\right)}{n_{eff}^{(Ea)}}\,.$$

In order to normalize the PDF, both pieces together with the theoretical fit must be divided by A_{eff}. Contrarily to the procedure to construct a merged histogram explained in Sec. 7.1.2, this procedure is performed by assuming that the theoretical PDF is known *a priori* in order to obtain a graphical representation.

Earthquake catalogs have been also merged with a charcoal labquake catalog with a global power-law exponent $\hat{\Gamma}_{E_r}=1.688$ suggesting that these different systems might be classified into the same universality class. Further investigations involving different observables, such as the distribution of waiting times, might be necessary in order to properly classify charcoal labquakes and real earthquakes into the same universality class. However, interesting conclusions can be extracted from this global statistical law. Recall that this truncated power-law has been fitted with cut-offs $x_{min}=6.31\times10^{-17}$ Nm, $x_{max}=1.01\times10^{16}$ Nm and the exponent $\hat{\Gamma}_{E_r}=1.688$. According to this statistical model, one could state that about 2×10^{21} events whose radiated energy is in the interval $10^{-16}<E_r<10^{-15}$ will occur for each earthquake in the interval $10^{15}<E_r<10^{16}$ Nm ($7<m<7.66$ in moment magnitude).

Chapter 8

Effect of the Coulomb Stress on the Gutenberg-Ricther law

The Mohr-Coulomb theory was exposed in Sec. 3.1.1.1 with the aim of contextualizing the motivation of the results explained in this chapter. This chapter pursues statistical evidence that the b-value is affected by the value of the Coulomb stress. First of all, the procedure through which Coulomb stresses are computed at the fault plane for each aftershock is explained in Sec. 8.1.1. The aftershock sequence after the Landers mainshock ($m = 7.3$, California, 1992) is chosen in order to compute the values of the Coulomb stress increase for each aftershock. The election of this mainshock is due to its importance (it is the largest earthquake in California at least since 1952), and because it is recorded in a catalog where the focal mechanisms of aftershocks (which are needed for the calculations of Coulomb stresses increase on the fault planes) are reported. This aftershock sequence has been broadly studied and different slip models have been proposed to characterize the slip at each point of the fault rupture caused by Landers. These models are briefly explained in Sec. 8.1.2.

After computing the increases of Coulomb stresses for Landers aftershocks, two different subsets of earthquakes can be distinguished: the ones which have received a positive increase in Coulomb stress and those which have received a negative increase in Coulomb stress. The b-values of the Gutenberg-Richter law are then computed for these subsets by using the statistical techniques exposed in Chpt. 5. These magnitude distributions and their corresponding b-values are compared in order to check whether they are sensible to the sign of the increase in Coulomb stress.

The results exposed in this chapter have been accepted for publication in:

- Víctor Navas-Portella, Abigail Jiménez, and Álvaro Corral, No Significant Effect of Coulomb Stress on the Gutenberg-Richter Law after the Landers Earthquake. Scientific Reports 10, 2901 (2020) [Navas-Portella et al., 2020]

8.1 Computation of the Coulomb stress

8.1.1 General procedure

Coulomb-stress theory has been used to forecast spatial patterns of aftershock rates, as well as assessing the likelihood of earthquake rupture sequences [Quigley et al., 2019; Parsons et al., 2000]. Although there exist instances where its predictive skills are arguable [Hardebeck et al., 1998; Marsan, 2003; Mallman and Zoback, 2007; Felzer and Brodsky, 2005], the monitoring of the changes in the stress field represents a valuable information for seismic and volcanic hazard forecasting and to proposing the adequate mitigation measures.

Whenever an earthquake occurs, an underlying mechanical process takes place and the faults or the systems of faults experience a change in geometry in the form of a rupture and a slip. In practice, it is generally difficult to know the configuration of faults due to the inaccessibility of this information or the complexity of the fault system. A useful approach to deal with this problem is by means of slip models, which try to geometrically characterize the faults and the displacement they have experienced. However, after a large earthquake (mainshock), not only the fault in which the earthquake has occurred experiences a rupture but also the surrounding region experiences a deformation. The deformation at each point of the surrounding (deformation field $\boldsymbol{u}$ or displacement) is computed by means of the dMODELS software [Battaglia et al., 2013]. The local coordinate system for dMODELS is east-north-up, ENU, which is explained in Sec. 3.2.1. Although there exist many programs that calculate deformation caused by earthquakes, this package has been thoroughly tested, and can introduce many different sources of deformation, which can be translated into stress changes in a straightforward way.

After introducing the corresponding slip model (also called source model) for the mainshock of interest (slip models for Landers mainshock are explained in the next Sec. 8.1.2), the projections in the ENU axes of the deformation field $\boldsymbol{u}$ caused by the mainshock at the position of each aftershock (and also at its neighborhood, in order to take derivatives) are obtained. The strain tensor associated to $\boldsymbol{u}$ can be obtained by calculating the (symmetrized) gradient of the deformation according to Eq. (3.1) (with a spatial step equal to 1 km). Afterwards, an isotropic and elastic material is assumed for calculating the stress tensor [Lautrup, 2011], or, more precisely, the contribution of the mainshock to the stress tensor, according to Eq. (3.2) with the Lamé elastic moduli given by $\mu = \lambda = 3 \times 10^4$ MPa [Kanamori and Brodsky, 2004] (Poisson ratio $\nu = \lambda\,(\lambda + \mu)^{-1}/2 = 0.25$). Note that in order to be realistic, the Coulomb-stress changes have to be calculated onto the planes of the actual faults so that one can actually evaluate if the new events could have been triggered by the induced stress or not. [McCloskey et al., 2003; Ishibe et al., 2017]. This contrasts with an approach in which Coulomb stresses are calculated onto the so-called optimally oriented planes [King et al., 1994], when the only information available is the regional stress. However, optimally oriented planes are imaginary planes that might not correspond to the actual geology.

Given the fault plane and slip vector of an aftershock in East-Nord-Up (ENU) coordinates, the changes in the normal σ_n and shear (or tangential) τ stresses are calculated in that orientation and position, as

$$\Delta\sigma_n = \sum_{kl} n_k s_{kl} n_l \text{ and } \Delta\tau = \sum_{kl} \ell_k s_{kl} n_l, \tag{8.1}$$

with n_i and ℓ_i the components of the normal and slip vectors, respectively (see Sec. 3.2.1).

The relationship between the focal mechanism (which is available from the YHS catalog, see Sec. 4.2.3) and the normal vector of the fault is given by

$$\hat{\boldsymbol{n}} = \begin{pmatrix} n_E \\ n_N \\ n_U \end{pmatrix} = \begin{pmatrix} \cos\Theta \sin\mathfrak{d} \\ -\sin\Theta \sin\mathfrak{d} \\ \cos\mathfrak{d} \end{pmatrix} \tag{8.2}$$

in the ENU coordinate system [Smith, 2006]. In the same way, the slip vector is obtained as

$$\hat{\boldsymbol{\ell}} = \begin{pmatrix} \ell_E \\ \ell_N \\ \ell_U \end{pmatrix} = \begin{pmatrix} \sin\Theta\cos\rho - \cos\Theta\cos\mathfrak{d}\sin\rho \\ \cos\Theta\cos\rho + \sin\Theta\cos\mathfrak{d}\sin\rho \\ \sin\mathfrak{d}\sin\rho \end{pmatrix}, \tag{8.3}$$

where $\mathfrak{d}$, Θ and ρ are the dip, strike and rake angles defined in Sec. 3.2.1.

As it has been exposed in Sec. 4.2, high quality catalogs incorporating focal-mechanism solutions are nowadays available. Given the distribution of acceptable mechanisms, the preferred solution is the most probable one [Hardebeck and Shearer, 2002]. The ambiguity of the actual fault plane is solved by considering that the preferred nodal planes are those associated with the preferred solution listed in the catalog [Yang et al., 2012]. The focal mechanism, in concrete, the rake angle, together with Landers stress field derived from the slip model, allows the calculation of the Coulomb-stress increases (positive or negative) induced by the mainshock on the actual orientations of the aftershock ruptures. Nevertheless, if one does not trust the preferred solutions reported in the catalog, an alternative procedure can be applied based on selecting the nodal plane in which the Coulomb stress increase is maximum [Strader and Jackson, 2014].

The change in Coulomb stress at the aftershock fault plane due to the mainshock will be

$$\Delta\mathfrak{C} = \Delta\tau + \mu'\Delta\sigma_n, \tag{8.4}$$

with $\Delta\tau$ and $\Delta\sigma_n$ coming from Eq. (8.1) (see Sec. 3.1.1.1). Thus, positive increases of the Coulomb stress bring the fault closer to failure, whereas negative increases distance it away from failure.

The complete procedure to obtain the Coulomb stress is summarized in Fig. 8.1.

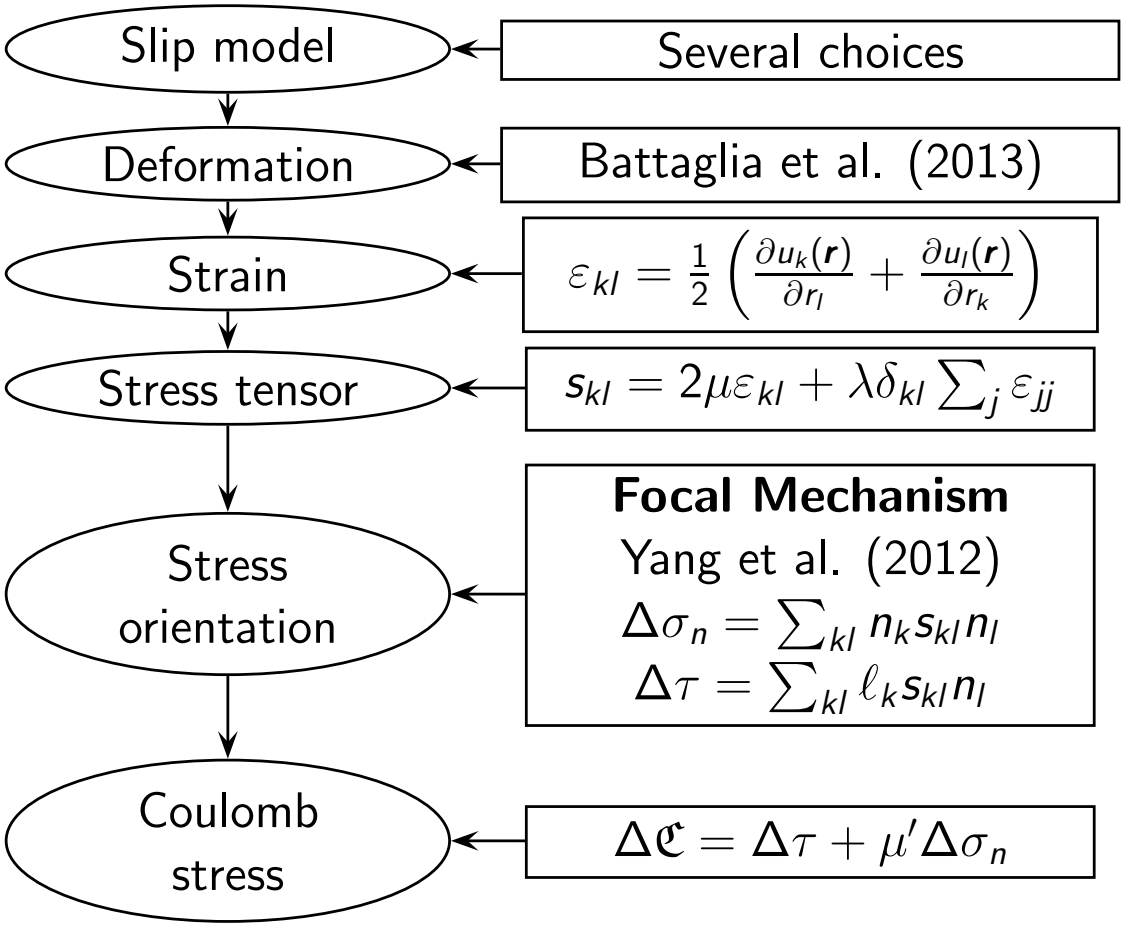

Fig. 8.1: Flowchart summarizing the procedure to obtain the Coulomb stress on each aftershock fault plane from the slip model and the focal-mechanism catalog.

8.1.2 Slip Models for Landers earthquake

Although describing the slip experienced by a fault or system of faults due to an earthquake is an extremely complex task, slip models try to simplify the fault geometry and associate a certain value of the slip to each part of the fault. The first simplification for modelling faults is based on the assumption that they can be considered as planar surfaces which are divided into equally-sized cells called fault patches. There exist several techniques, such as Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR), inversion techniques by using Global Position System (GPS) and Synthetic Aperture Radar interferometry (inSAR) (see for instance [Yeats et al., 1997] p.100-106), through which one can associate a certain slip, in meters, to each fault patch.

The June 28, 1992, Landers earthquake (34.22°N, 116.43°W at 11:57:35 Coordinated Universal Time (UTC)) occurred with a moment magnitude $m = 7.3$ and a rake angle $\rho = -177°$,

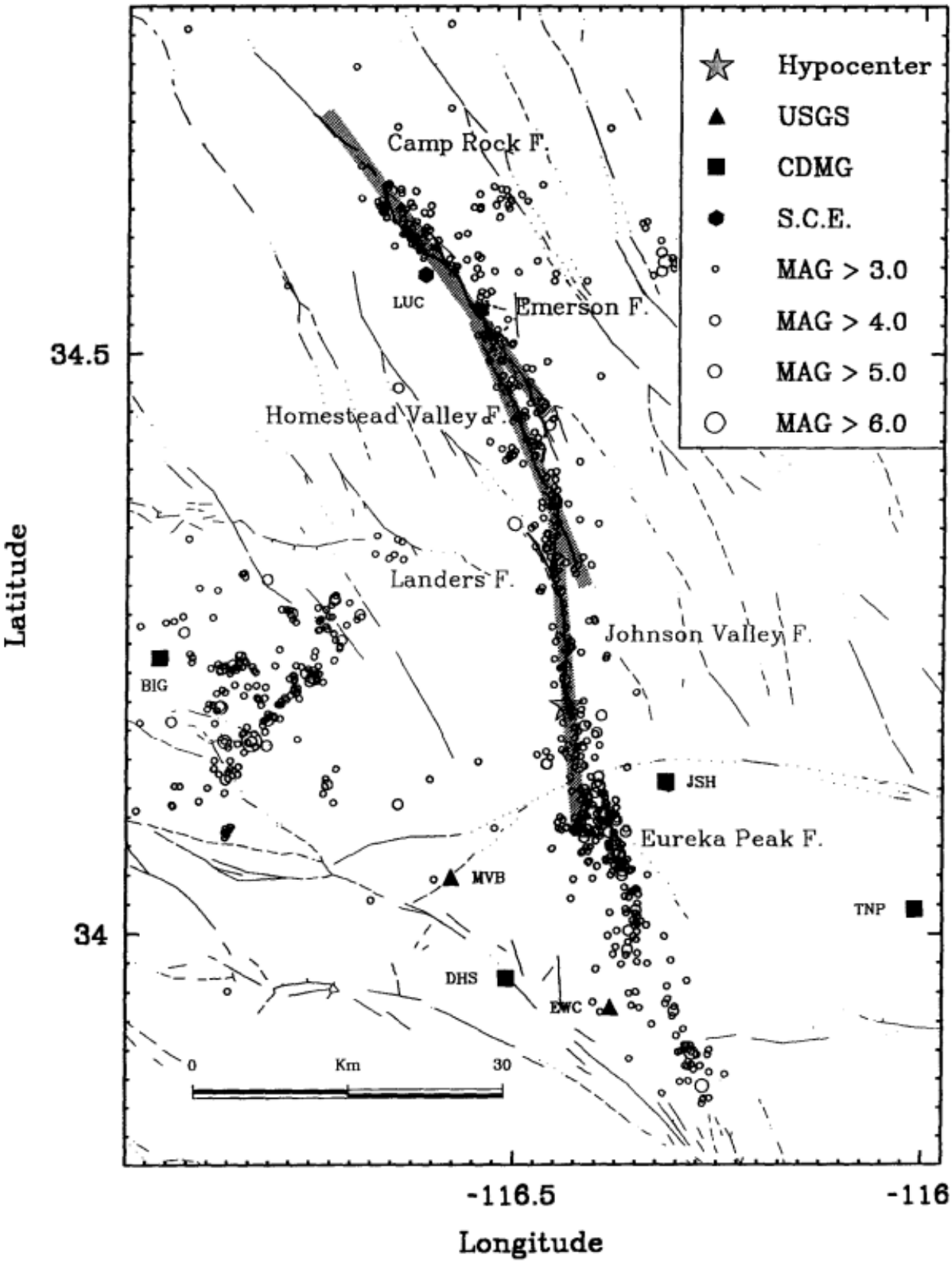

Fig. 8.2: Map of the surrounding region of Landers mainshock, which is represented by a star symbol ⋆ . Aftershocks are represented as circles with magnitudes given by the legend. Dark lines correspond to the fault surface and the hatchured lines are the projection of the three fault segments of the **wald** slip model. Figure extracted from Wald and Heaton [1994].

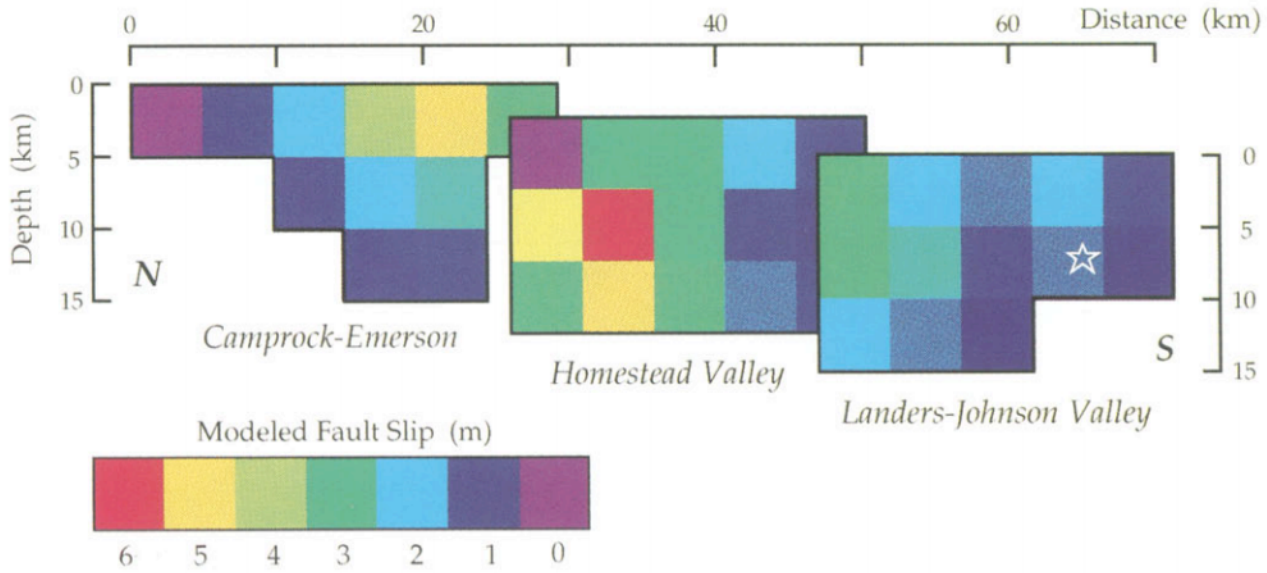

Fig. 8.3: Pictorical representation of the **wald** slip model for the Landers Earthquake [King et al., 1994]. White star corresponds with the hypocenter of the Landers earthquake. The color of every fault patch represents the experienced slip due to the Landers event in meters.

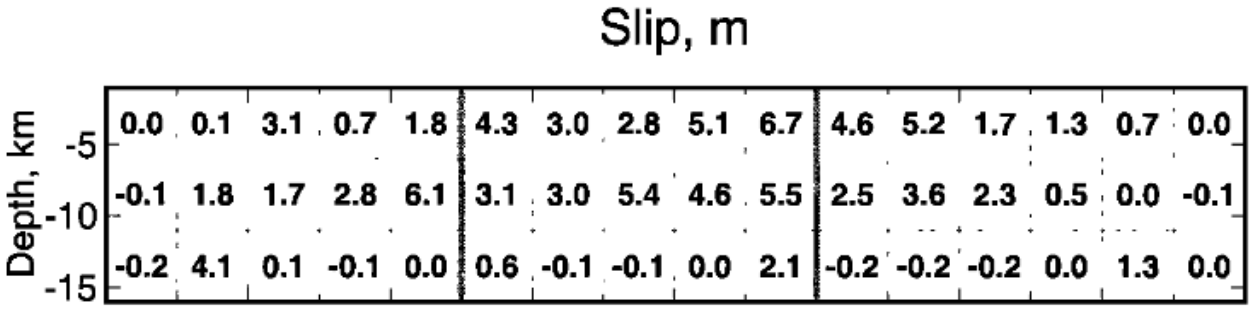

Fig. 8.4: Pictorical representation of the **hernandez** slip model for the Landers Earthquake [Hernandez et al., 1999]. The numbers in every fault patch represents the experienced slip due to the Landers event in meters. Positive numbers correspond to right-lateral (Landers' focal mechanism) whereas negative numbers correspond to left-lateral. Note that, contrarily to the orientation exposed in Fig. 8.3, the North is at right whereas South is at left.

corresponding to strike-slip focal mechanism, thus being the strongest one in Southern California at least since 1952. The earthquake and its subsequent aftershock sequence have been extensively studied [Hill et al., 1993; Hauksson et al., 1993; Gomberg et al., 2001], with a number of slip distributions that describe its rupture [Wald and Heaton, 1994; Spotila and Sieh, 1995; Hernandez et al., 1999; Steacy et al., 2004].

Four slip models are used here to calculate the strain:

- Wald and Heaton (referred here to as **wald**)[Wald and Heaton, 1994]: This slip model has been derived from joint inversion of strong motion, teleseismic, geodetic, and surface slip data. This model is composed by three planar surfaces, with 11, 15 and 13 cubic fault patches each with an edge of 5 km. The planar surfaces are normal to the plane conformed by the ground, where they span around 70 km and 15 km in depth. See Figs. 8.2 and 8.3.

- Hernandez et al. (**hernandez**) [Hernandez et al., 1999],: Slip model constructed by means of GPS and InSAR techniques. The model is geometrically defined by three planar surfaces with 15, 15 and 18 cubic patches each with an edge of 5 km. Note that this configuration is very similar to the one in **wald** and the spatial distribution of the planar surfaces can also be understood from Fig. 8.2. The planar surfaces are normal to the plane conformed by the ground, where they span around 70 km and 15 km in depth. Some fault patches change its sign from right-lateral (Landers' focal mechanism) to left-lateral. See Fig. 8.4.

- Landers Big-Bear California (**bbcal**) [Steacy et al., 2004]: The largest Landers' aftershock was the Big Bear earthquake $m = 6.3$ (34.29°N, 116.82°W which occurred the same day as Landers at 15:05:30 UTC (approximately three hours later than Landers) with a rake angle $\rho = -180°$ [Steacy et al., 2004]). This slip model is constructed with the same fault geometry as the **hernandez** slip model but, instead of Landers, the Big Bear earthquake is considered as a temporal reference. The slip corresponding to this earthquake is included in an additional fault patch.

- Landers Surface Rupture (**lsurfrup**) [Spotila and Sieh, 1995]: Slip model made of 11 irregular quadrilateral fault patches which are placed at the ground.

The terminology is the same as the one used by Steacy et al. [2004]. Distances from each event to the fault are computed as the minimum Euclidean distance from the aftershock hypocenter to the center of each fault patch as given by the slip model.

8.2 Results

8.2.1 Spatio-temporal window

In order to better detect the influence of the Landers stress change a time window of 100 days after the mainshock and a spatial window going from 10 to 150 km from the Landers rupture are taken. This spatio-temporal window defines Landers aftershocks for the purposes of the statistical analysis. The reason to exclude events closer than 10 km is the uncertainty of the deformation field near the edges of the subfaults [Okada, 1992], as the finite-fault approximation provides spurious values near the fault zone because of boundary effects. Landers

earthquake is taken as the temporal reference for all the slip models except for the **bbcal** whose temporal reference is the Big-Bear earthquake. The temporal window ensures that the activity rate has decreased after 100 days and the influence of Landers is rather small (see Fig. 8.5). Two different subsets with positive and negative increase in Coulomb stress can be distinguished for each slip model with a cut-off magnitude $m_{min} = 3$. Scatter plots illustrating the dependence of the absolute value of the increase of Coulomb stress $|\Delta\mathfrak{C}|$ as a function of the distance to the fault $d_{ev-fault}$ are shown in Fig. 8.6 for the four slip models. As it is implicit by the Coulomb theory, the value of the increase of Coulomb stress decays as the cube of the distance to the fault. Aftershocks with positive and negative increase of Coulomb stress placed in the window of study for the **hernandez** slip model and $\mu' = 0.4$ are presented in Figs. 8.7, 8.8 and 8.9. Given that the results exposed in this section do not show remarkably differences depending on the values of the friction coefficient, if not otherwise specified, this value is taken as $\mu' = 0.4$ (see Sec. 8.2.4 for more details).

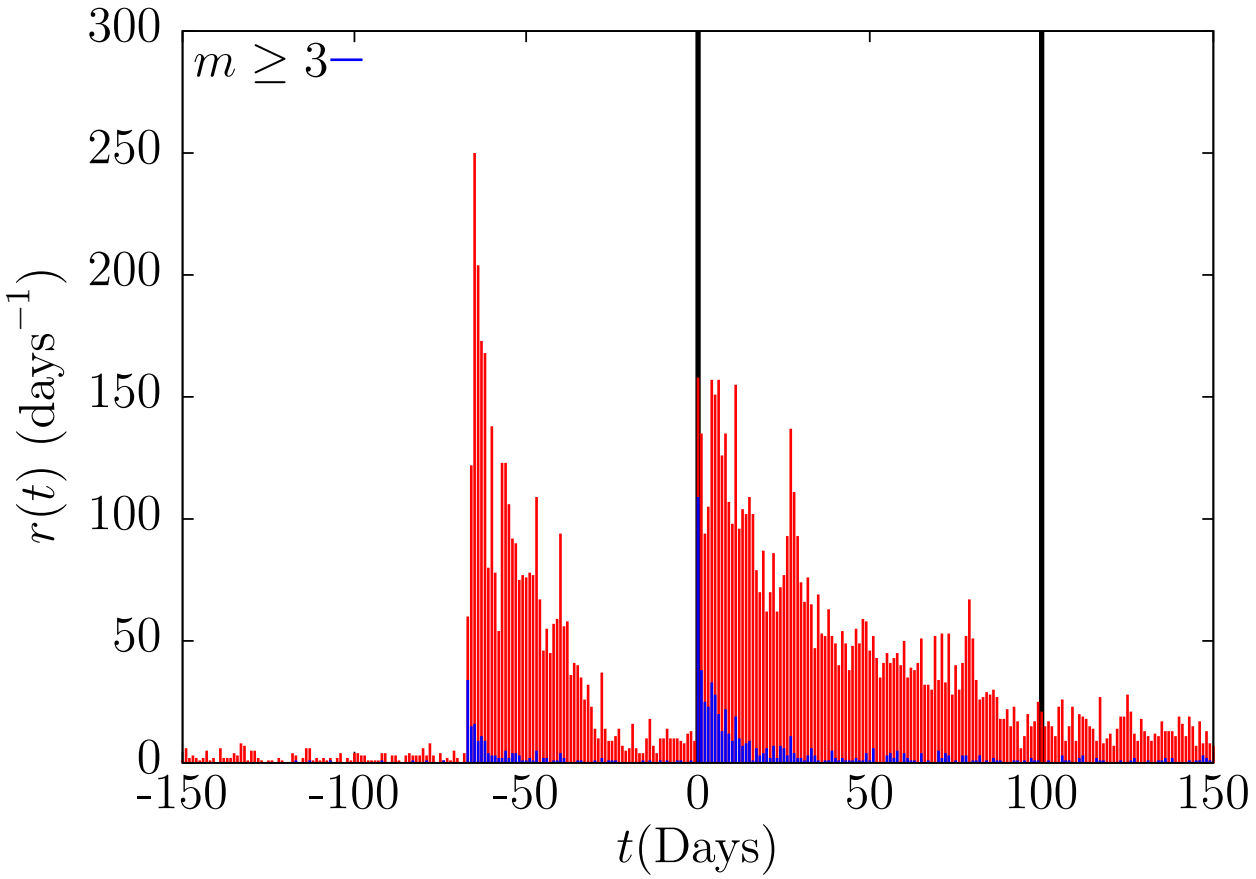

Fig. 8.5: Number of earthquakes per day of any magnitude (in red) and for $m \geq 3$ (in blue) before and after the Landers mainshock in the YHS catalog in the area selected for this study (10 km$\leq d_{ev-fault} \leq$ 150 km). Black lines show the temporal window chosen in this work.

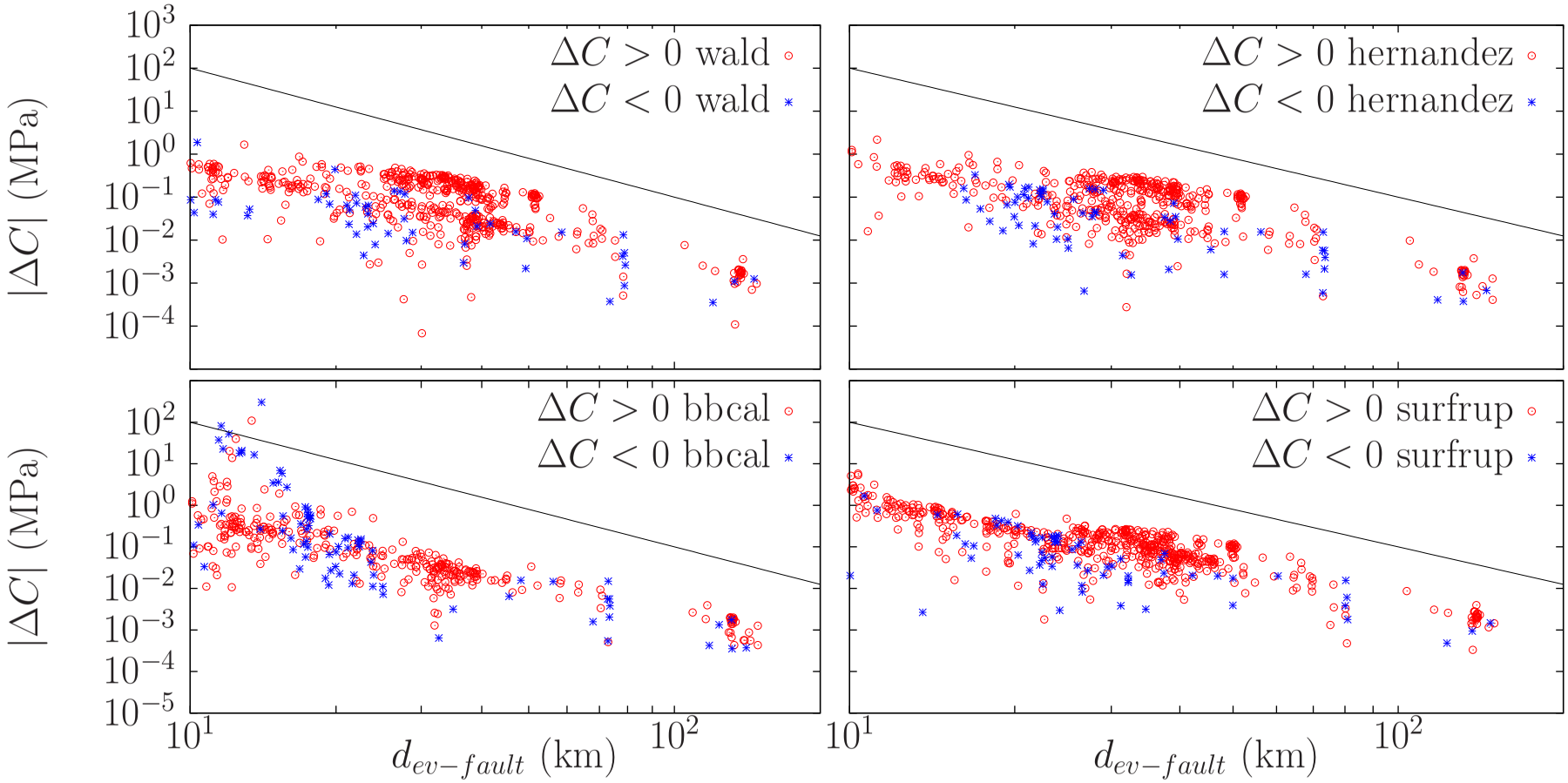

Fig. 8.6: Dependence of the absolute value of the change in the Coulomb stress $\Delta\mathfrak{C}$ as a function of the distance to the fault of the aftershocks for each slip model, $m_{min} = 3$ and $\mu' = 0.4$. Aftershocks correspond to the first 100 days after the Landers mainshock (except for the **bbcal** slip model, which has the Big Bear earthquake as a temporal reference). Distance of aftershocks to the Landers rupture is restricted to be between 10 and 150 km. Black dashed line with slope -3, as stated by Coulomb theory, is shown as a guide to the eye.

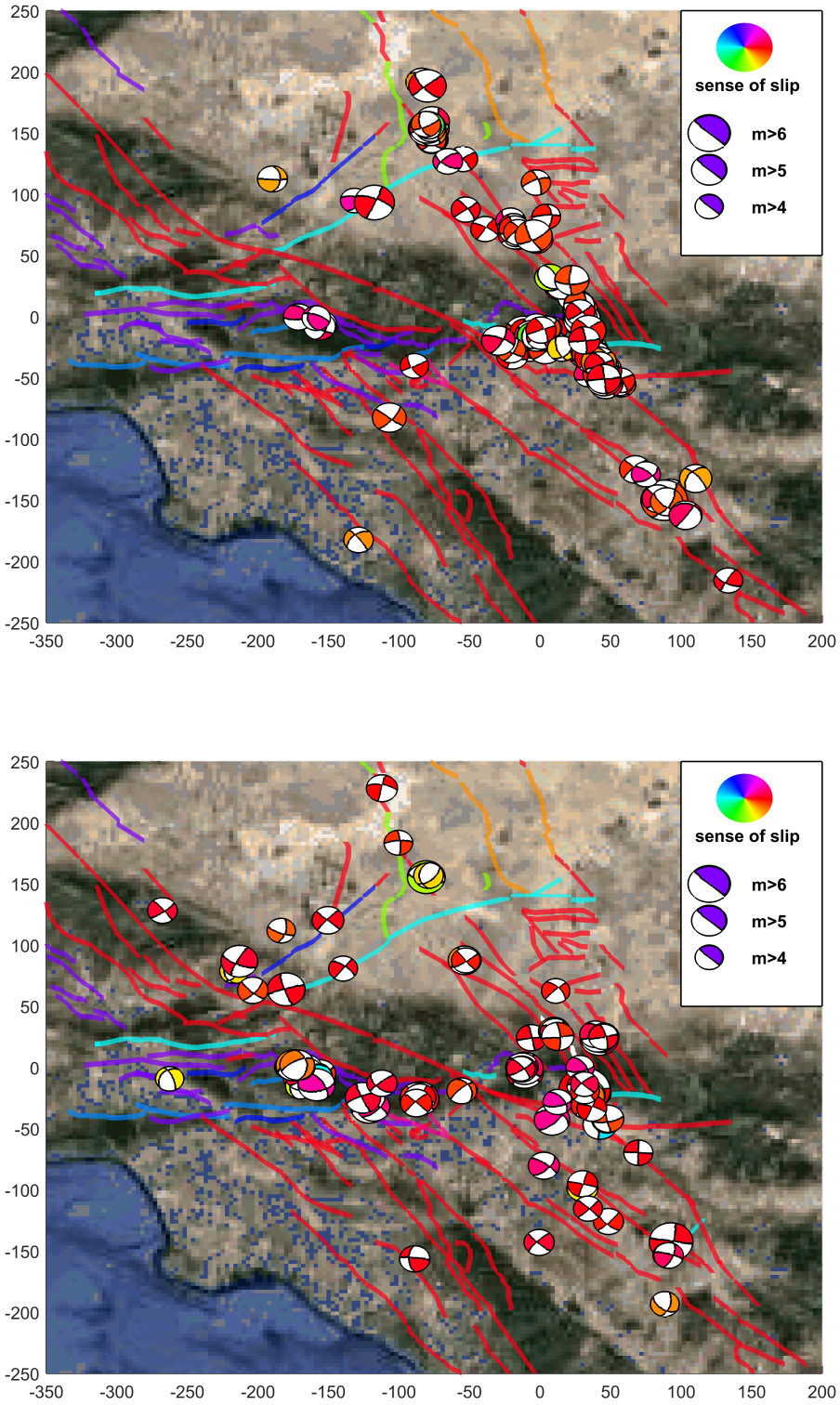

Fig. 8.7: Focal mechanism representation for strike-slip Landers aftershocks, separated in terms of $\Delta\mathfrak{C}$, as calculated from the hernandez slip model (time window of 100 days after the mainshock). Top: $\Delta\mathfrak{C} > 0$. Bottom: $\Delta\mathfrak{C} < 0$. Color scale represents the sense of slip (rake) and fault traces are also shown using the same color code: red for right-lateral (ρ close to $\pm 180°$), light blue for left-lateral (ρ close to $0°$), green for normal (ρ close to $-90°$) and dark blue or purple for thrust faulting (ρ close to $90°$). An area of 550×500 km is shown; aftershocks are restricted to $m \geq 4$ (for clarity sake). Aftershocks beyond the limit of 150 km are also shown. Both axes display distances with respect an arbitrary origin, in km.

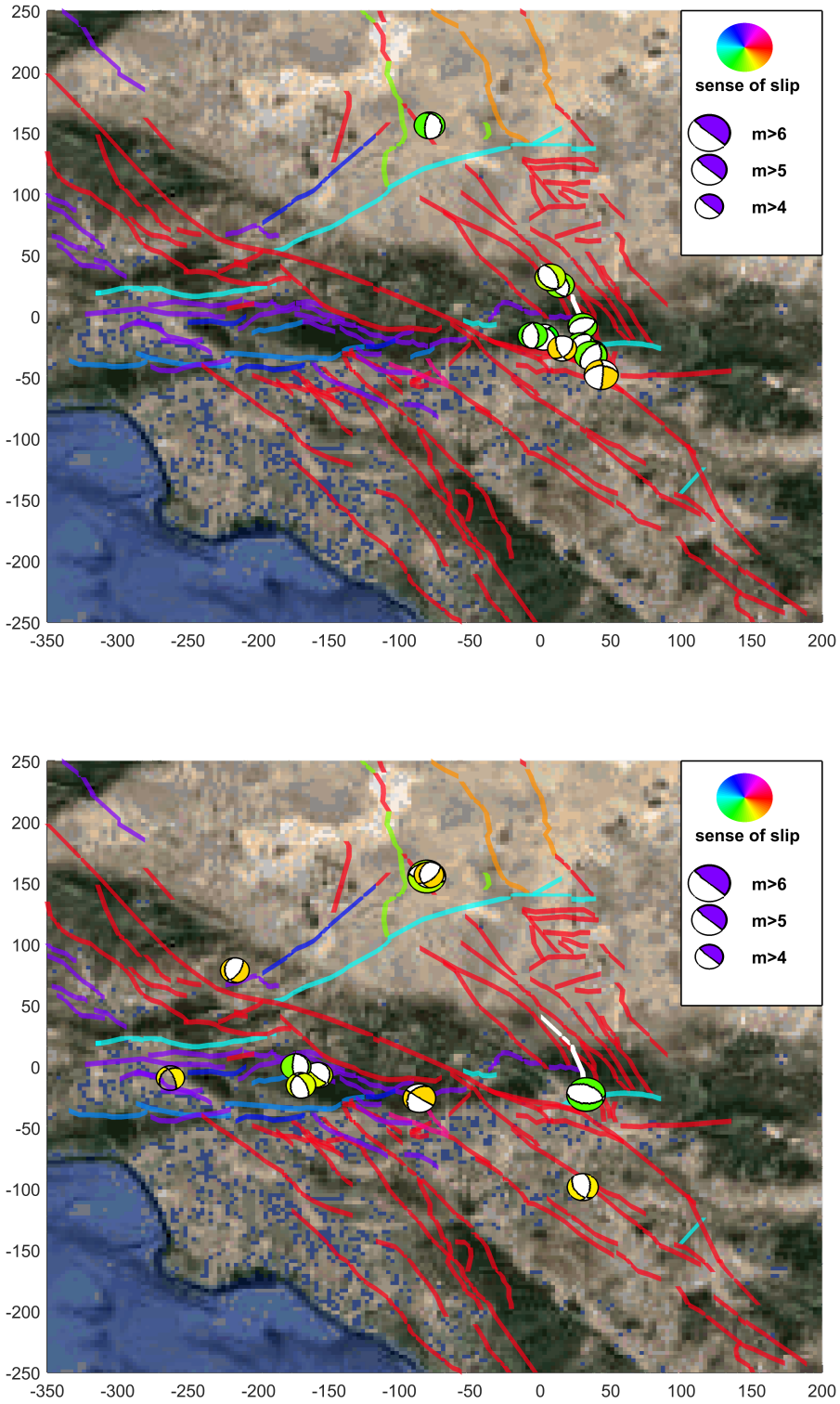

Fig. 8.8: Same as previous figure, for aftershocks with normal focal mechanism.

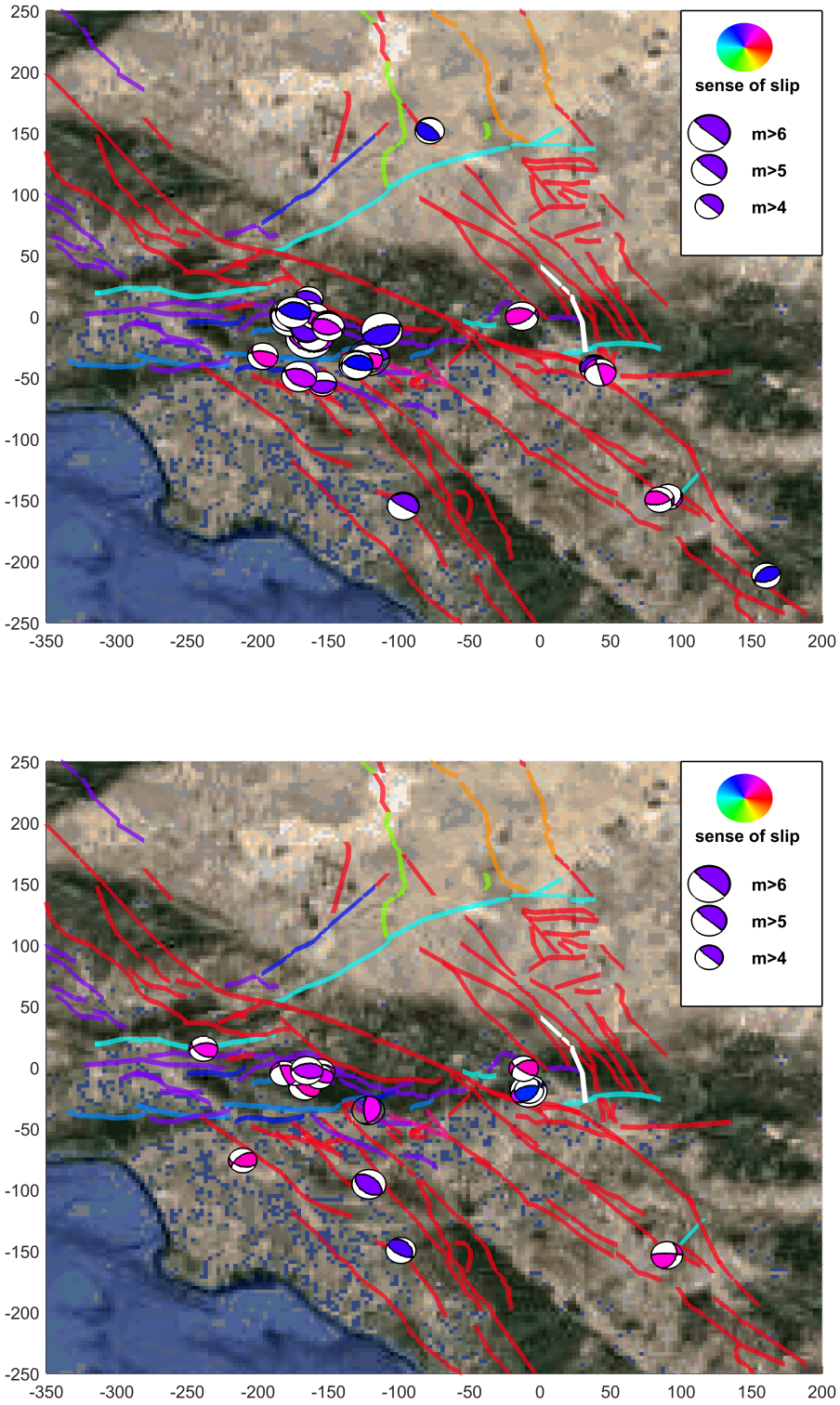

Fig. 8.9: Same as previous figure, for aftershocks with thrust focal mechanism.

8.2.2 Comparison of distributions

Table 8.1 shows the estimated b-values obtained from the application of the MLE and goodness-of-fit test explained in Sec. 5.2 to the different subcatalogs obtained from the Landers sequence. As it can be seen, in the overall case (when events are not separated in terms of Coulomb-stress change), the GR law is fulfilled with an average value $\hat{b}_{all} = 0.92$. Each slip model leads to a different value of $\hat{b}_{all}$ because the fault geometry is different for each of them, and events too close to the fault are discarded. This $b-$value for the Landers aftershocks is found, not surprisingly, to be close to the average for aftershocks in California, $\hat{b} \simeq 0.9$ [Reasenberg and Jones, 1989; Jones, 1994], and somewhat below the long-term value of Southern California (all events), $\hat{b} \simeq 1.0$ [Hutton et al., 2010] (although other works report $\hat{b} \simeq 1.0$ for Landers aftershocks, probably due to the consideration there of a much smaller magnitude of completeness [Shcherbakov et al., 2005]).

Slip Model		N	n	$\hat{b}-$**value**	σ	p_{value}
wald	$\Delta\mathfrak{C} > 0$	5213	509	$\hat{b}_{>} = 0.927$	0.041	0.313 ± 0.005
	$\Delta\mathfrak{C} < 0$	814	51	$\hat{b}_{<} = 0.766$	0.107	0.861 ± 0.003
	All	6027	560	$\hat{b}_{all} = 0.909$	0.038	0.243 ± 0.004
hernandez	$\Delta\mathfrak{C} > 0$	5027	465	$\hat{b}_{>} = 0.926$	0.043	0.505 ± 0.005
	$\Delta\mathfrak{C} < 0$	765	62	$\hat{b}_{<} = 0.866$	0.110	0.197 ± 0.004
	All	5792	527	$\hat{b}_{all} = 0.919$	0.040	0.231 ± 0.004
bbcal	$\Delta\mathfrak{C} > 0$	3641	309	$\hat{b}_{>} = 0.978$	0.056	0.232 ± 0.004
	$\Delta\mathfrak{C} < 0$	1191	82	$\hat{b}_{<} = 0.948$	0.105	0.327 ± 0.005
	All	4832	391	$\hat{b}_{all} = 0.971$	0.049	0.053 ± 0.002
surfrup	$\Delta\mathfrak{C} > 0$	5534	548	$\hat{b}_{>} = 0.890$	0.038	0.290 ± 0.005
	$\Delta\mathfrak{C} < 0$	774	68	$\hat{b}_{<} = 0.891$	0.108	0.555 ± 0.005
	All	6308	616	$\hat{b}_{all} = 0.890$	0.036	0.239 ± 0.004

Table 8.1: Results of fitting the Gutenberg-Richter law to the Landers aftershocks, separating positive and negative Coulomb-stress increases, for different Slip models, $\mu' = 0.4$ and $m_{min} = 3$. Aftershocks correspond to the first 100 days after the Landers mainshock. Distance of aftershocks to the Landers rupture is restricted to be between 10 and 150 km. The $p-$value of the goodness-of-fit test is computed with 10^4 simulations and is denoted by p_{value}. Its uncertainty corresponds to one standard deviation. In no case the Gutenberg-Richter law can be rejected.

After separating by the sign of the Coulomb-stress change, the first result that becomes apparent from the table is that the number of aftershocks with positive increases is much larger than the number for the negative case [King et al., 1994; Steacy et al., 2005], no matter the slip model (nor the value of μ') used to calculate $\Delta\mathfrak{C}$ (see next Sec. 8.2.4). All the quantities with sub-indices $>$ and $<$ correspond to those computed for aftershocks with $\Delta\mathfrak{C} > 0$ and $\Delta\mathfrak{C} < 0$ respectively. Regarding the $b-$values, although they depend on the slip model, they

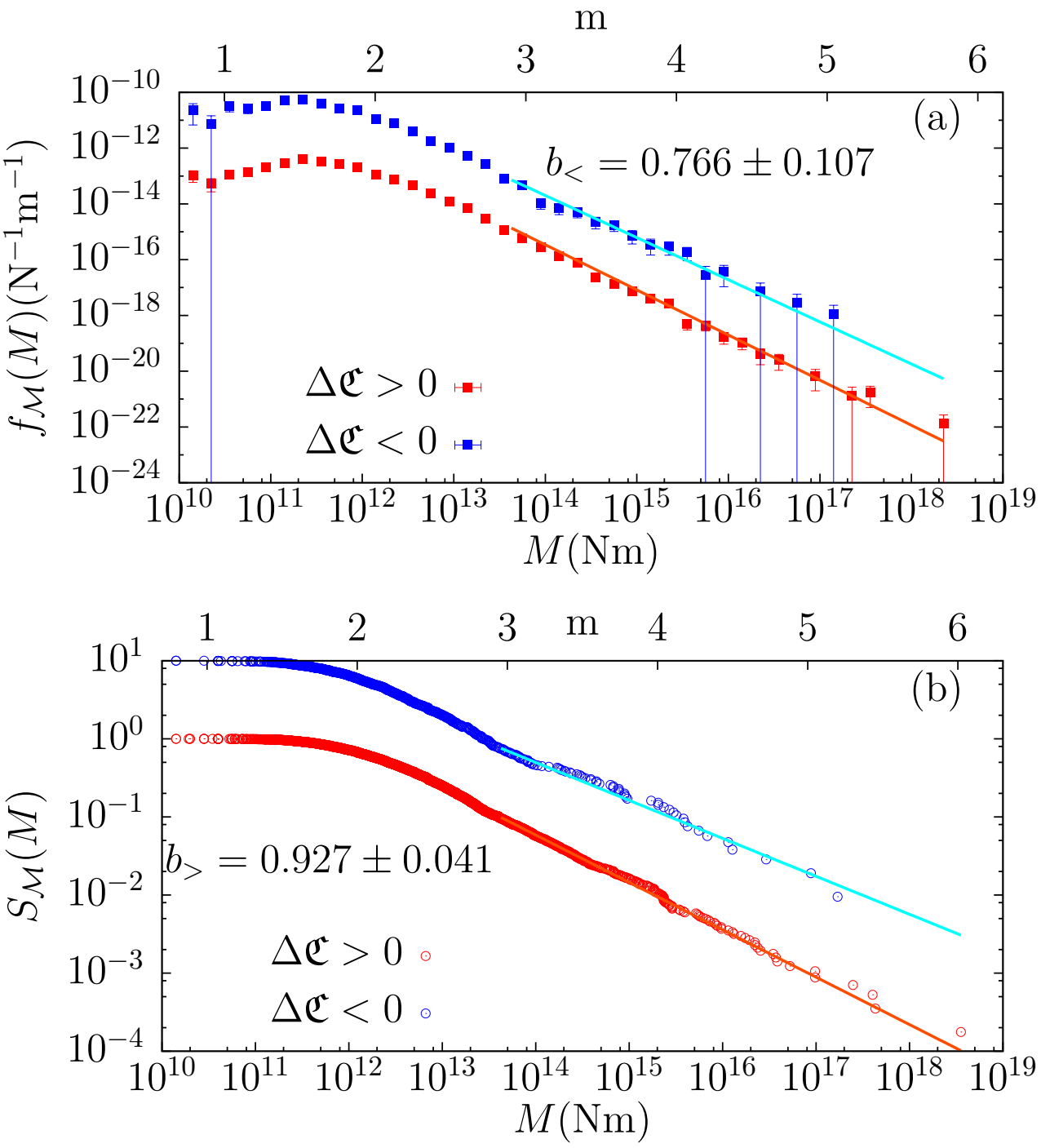

Fig. 8.10: Estimation of the PDFs (a) and of the complementary cumulative distribution functions (CCDF) (b) of seismic moment M for Landers aftershocks with $\Delta\mathfrak{C} > 0$ and $\Delta\mathfrak{C} < 0$ for the **wald** slip model using $\mu' = 0.4$ in the calculation of $\Delta\mathfrak{C}$. Curves corresponding to $\Delta\mathfrak{C} < 0$ have been conveniently multiplied by a factor 100 and 10 respectively for clarity sake. Error bars in (a) denote one standard deviation, and are symmetric, despite the appearance in log scale [Deluca and Corral, 2013].

can be summarized by taking the mean of the four models and taking $\mu' = 0.4$ as $\hat{b}_> \simeq 0.93$ and $\hat{b}_< \simeq 0.87$ with individual uncertainties around 0.04 and 0.11 respectively. Note that the magnitude distribution for the overall case is a mixture of the distributions corresponding to $\Delta\mathfrak{C} > 0$ and $\Delta\mathfrak{C} < 0$, and therefore, the value of $\hat{b}$ in the overall case turns out to be the harmonic mean of $\hat{b}_>$ and $\hat{b}_<$, i.e.,

$$\hat{b}_{all}^{-1} = \frac{n_> \hat{b}_>^{-1} + n_< \hat{b}_<^{-1}}{n_> + n_<}, \tag{8.5}$$

see Sec. 7.1.1. Despite the fact the values of $\hat{b}_>$ and $\hat{b}_<$ do not look much different between them, statistical testing becomes necessary in order to establish significance [Corral et al., 2015].

Slip Model	z	p_{norm}	p_{perm}	d_{2ks}	p_{2ks}	$\Delta\mathcal{A}$
wald	1.396	0.163	0.156 ± 0.004	0.139	0.311	0.234
hernandez	0.511	0.609	0.596 ± 0.005	0.095	0.690	1.748
bbcal	0.254	0.800	0.828 ± 0.004	0.063	0.952	1.936
surfrup	-0.010	0.992	0.994 ± 0.001	0.094	0.643	1.999

Table 8.2: Results of the statistical tests comparing the estimated $\hat{b}$-values and magnitude distributions for positive and negative Coulomb-stress changes, using different slip models and $\mu' = 0.4$. Values of the AIC $\Delta\mathcal{A} = \mathcal{A}_2 - \mathcal{A}_1$ are also included. Same data as previous table. Columns 2 to 4: testing the null hypothesis that there is no difference between the $b-$values (i.e., $b_> = b_<$). Both asymptotic normality of the z statistic and a permutation test are used for the calculation of the $p-$value (labeled as p_{norm} and p_{perm}, respectively). The number of permutations is 10^4, and the uncertainty of p_{perm} corresponds to one standard deviation. Columns 5 to 6: testing the null hypothesis using the 2-sample Kolmogorov-Smirnov test. d_{2ks} and p_{2ks} are the Kolmogorov-Smirnov statistic and its $p-$value.

Table 8.2 compares $\hat{b}_>$ and $\hat{b}_<$ for the different slip models taking $\mu' = 0.4$, and shows that the difference in the $\hat{b}-$values cannot be considered significantly different from zero with a confidence larger than 0.95; so, the null hypothesis $\hat{b}_> \simeq b_<$ can not be rejected. The statistical test to check the compatibility of power-law exponents exposed in Sec. 5.9 is first performed thus providing the value of the z-statistic as well as the p-values obtained through the permutational p_{perm} test and the one obtained from the normal distribution p_{norm}. Table 8.2 also shows the results of the two-sample KS test and the calculation of $\Delta\mathcal{A}$ explained in Secs. 5.8 and 5.7, leading in both cases to the result that no change in the distributions as a function of positive and negative $\Delta\mathfrak{C}$ can be established. In concrete, $\Delta\mathcal{A}$ is always greater than the critical value $\Delta\mathcal{A}_c = -1.84$ [Burnham and Anderson, 2002; Murtaugh, 2014] at significance level of 0.05. The **wald** slip model is the one for which both distributions (positive and negative) appear as more different; however, the difference is not significant. Figure 8.10 shows the PDFs as well as the CCDFs in this case.

8.2.3 Influence of the focal mechanism

As mentioned in the introduction, some authors have unveiled a direct dependence of the $b-$value on the focal mechanism of the events, which implies a dependence of b on the total stress (not the stress increase) [Schorlemmer et al., 2005]. The rake angle is associated to the focal mechanism in the following way: values of the rake around $-90°$ correspond to normal events (labelled as no), values around $0°$ or $\pm 180°$ to strike-slip events (ss), and values around $90°$ to thrust events (th).

After calculating and analyzing the b-values, no significant effect of the rake on the $b-$value is found (see Table 8.3), due to the low number of events in the normal and thrust regimes (which increases the uncertainty). But despite the large uncertainty, the values of b_{no} and b_{ss} are roughly in agreement with the results of Schorlemmer et al. [2005]; however, the value of b_{th} turns out to be rather large in comparison (but compatible, within the error bars).

It is also observed that in strike-slip and normal events the contribution from $\Delta\mathfrak{C} > 0$ is higher than in thrust events. This is verified by the ratios $n_{>ss}/n_{<ss}$ and $n_{>no}/n_{<no}$, which are larger than $n_{>th}/n_{<th}$ (see Table 8.3). Comparing with the number of earthquakes with each focal mechanism for the 5 years previous to Landers it can be concluded that it is indeed the low number of thrust aftershocks with positive $\Delta\mathfrak{C}$ which is anomalous (and not the relatively high number of them for negative $\Delta\mathfrak{C}$), due to an increase in the number of normal events and an even higher increase in strike-slip events triggered ($\Delta\mathfrak{C} > 0$) by the Landers mainshock. This difference in numbers becomes visually apparent in Figs. 8.7, 8.8 and 8.9.

fm	$n_{>fm}$	$n_{<fm}$	n_{fm}	n^{pre}_{fm}	$b_{>fm}$	$\hat{b}_{<fm}$	$\hat{b}_{fm}$	$\hat{b}^{pre}_{fm}$
wald								
No: $-135° \leq \rho \leq -45°$	39	5	$n_{no} = 44$	$n^{pre}_{no} = 77$	1.047	-	$b_{no} = 1.081$	$b^{pre}_{no} = 1.484$
Th: $45° \leq \rho \leq 135°$	9	8	$n_{th} = 17$	$n^{pre}_{th} = 77$	-	-	$b_{th} = 0.929$	$b^{pre}_{th} = 0.899$
SS: the rest	461	38	$n_{ss} = 499$	$n^{pre}_{ss} = 503$	0.914	0.726	$b_{ss} = 0.896$	$b^{pre}_{ss} = 0.960$
hernandez								
No: $-135° \leq \rho \leq -45°$	38	3	$n_{no} = 41$	$n^{pre}_{no} = 78$	0.995	-	$b_{no} = 1.011$	$b^{pre}_{no} = 1.502$
Th: $45° \leq \rho \leq 135°$	7	10	$n_{th} = 17$	$n^{pre}_{th} = 61$	-	-	$b_{th} = 0.929$	$b^{pre}_{th} = 0.899$
SS: the rest	420	49	$n_{ss} = 469$	$n^{pre}_{ss} = 506$	0.920	0.840	$b_{ss} = 0.911$	$b^{pre}_{ss} = 0.957$
bbcal								
No: $-135° \leq \rho \leq -45°$	22	4	$n_{no} = 26$	$n^{pre}_{no} = 80$	1.128	-	$b_{no} = 1.165$	$b^{pre}_{no} = 1.521$
Th: $45° \leq \rho \leq 135°$	7	5	$n_{th} = 12$	$n^{pre}_{th} = 69$	-	-	$b_{th} = 1.309$	$b^{pre}_{th} = 0.831$
SS: the rest	280	73	$n_{ss} = 353$	$n^{pre}_{ss} = 507$	0.970	0.886	$b_{ss} = 0.951$	$b^{pre}_{ss} = 0.958$
surfrup								
No: $-135° \leq \rho \leq -45°$	46	5	$n_{no} = 51$	$n^{pre}_{no} = 76$	0.939	-	$b_{no} = 0.966$	$b^{pre}_{no} = 1.470$
Th: $45° \leq \rho \leq 135°$	9	11	$n_{th} = 20$	$n^{pre}_{th} = 61$	-	1.010	$b_{th} = 0.886$	$b^{pre}_{th} = 0.899$
SS: the rest	493	52	$n_{ss} = 545$	$n^{pre}_{ss} = 504$	0.888	0.844	$b_{ss} = 0.884$	$b^{pre}_{ss} = 0.961$

Table 8.3: Number of events and $b-$values corresponding to Landers aftershocks with $m \geq 3$ separated by sign of the Coulomb-stress increase ($>$ and $<$) and by focal mechanism (fm) for each slip model. $fm = no$ (normal), ss (strike-slip), and th (thrust). Same data as in Tabs. 8.1 and 8.2. Values of b calculated with 10 or less events are not reported. Values for the 5 years previous to Landers are also included.

8.2.4 Robustness of the results

Given that there are many factors that might influence the results exposed in the last subsection 8.2.2, it is interesting to check the robustness of these results by selecting different spatial windows, a different criterion for the selection of the fault plane and different values of the effective friction coefficient.

First, Fig. 8.11 shows the values of $\hat{b}_>$ and $\hat{b}_<$ for other choices of the limits of the spatial window. Similar results as those presented in Table 8.1 are found and no significant differences in b-values are appreciated.

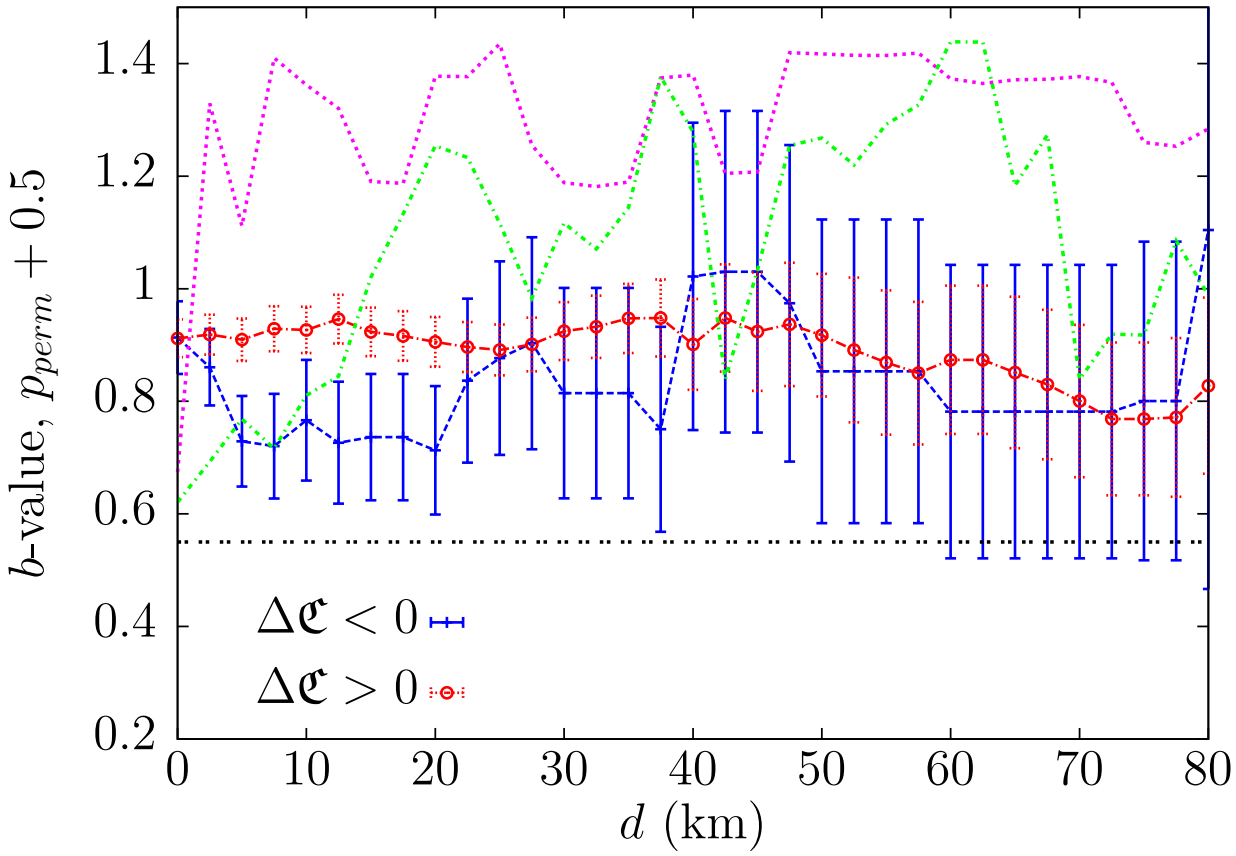

Fig. 8.11: Values of the exponents $b_>$ and $b_<$ (blue and red respectively) computed for those events placed at distances to fault larger than d and smaller than 150 km for the **wald** slip model, with $m_{min} = 3$ and $\mu' = 0.4$. Green and purple dashed lines correspond to the p-values of the goodness-of-fit test shifted 0.5 for convenience. Horizontal black dashed line corresponds to the threshold p_{perm}-value $p_c = 0.05$ shifted 0.5. Observe that in all cases $p_{perm} > 0.05$.

Second, instead of using the solutions provided by the YHS catalog one could also consider the focal mechanism in the complementary nodal plane. Given the focal mechanism in one plane $(\Theta, \mathfrak{d}, \rho)$, one can easily obtain the focal mechanism of the other plane $(\Theta^*, \mathfrak{d}^*, \rho^*)$ by imposing that the slip vector of one plane corresponds to the normal vector of the auxiliary plane and vice-versa $\hat{\ell} = \hat{n}^*$, $\hat{\ell}^* = \hat{n}$. By imposing these conditions and after applying some arithmetic and trigonometric calculations one can obtain the focal mechanism in the auxiliary plane as a function of the focal mechanism in the original plane:

$$\mathfrak{d}^* = \arccos\left(\sin\mathfrak{d}\sin\rho\right) \tag{8.6}$$

$$\rho^* = \arcsin\left(\frac{\cos\mathfrak{d}}{\sin\mathfrak{d}^*}\right) = \arcsin\left(\frac{\cos\mathfrak{d}}{\sin\left(\arccos\left(\sin\mathfrak{d}\sin\rho\right)\right)}\right) \tag{8.7}$$

$$\Theta^* = \Theta - \arcsin\left(\frac{\cos\rho}{\sin\mathfrak{d}^*}\right) = \Theta - \arcsin\left(\frac{\cos\rho}{\sin\left(\arccos\left(\sin\mathfrak{d}\sin\rho\right)\right)}\right) \tag{8.8}$$

Either $(\Theta, \mathfrak{d}, \rho)$ or $(\Theta^*, \mathfrak{d}^*, \rho^*)$ will correspond to the correct solution for the focal mechanism in the real fault plane. Care must be taken when finding the appropriate quadrants for the angles [Stein and Wysession, 2003]. Once the focal mechanisms are known for both nodal planes, the nodal plane in which the Coulomb stress increase is maximum is considered as the fault plane [Strader and Jackson, 2014]. The results obtained through this procedure are shown in Tables 8.4 and 8.5. The **wald** is the only slip model for which some significant difference between exponents is found. However, the results can be considered robust since just one from all the cases that have been analyzed exhibits significant differences.

Slip Model		N	n	$\hat{b}-$**value**	σ	p_{value}
wald	$\Delta\mathfrak{C} > 0$	5474	525	$\hat{b}_> = 0.936$	0.041	0.330 ± 0.005
	$\Delta\mathfrak{C} < 0$	553	35	$\hat{b}_< = 0.637$	0.108	0.909 ± 0.003
	All	6027	560	$\hat{b}_{all} = 0.909$	0.038	0.243 ± 0.004
hernandez	$\Delta\mathfrak{C} > 0$	5210	482	$\hat{b}_> = 0.938$	0.043	0.439 ± 0.005
	$\Delta\mathfrak{C} < 0$	582	45	$\hat{b}_< = 0.754$	0.112	0.255 ± 0.004
	All	5792	527	$\hat{b}_{all} = 0.919$	0.040	0.231 ± 0.004
bbcal	$\Delta\mathfrak{C} > 0$	4059	337	$\hat{b}_> = 0.993$	0.054	0.190 ± 0.004
	$\Delta\mathfrak{C} < 0$	773	54	$\hat{b}_< = 0.854$	0.116	0.357 ± 0.005
	All	4832	391	$\hat{b}_{all} = 0.971$	0.049	0.053 ± 0.002
surfrup	$\Delta\mathfrak{C} > 0$	5750	565	$\hat{b}_> = 0.902$	0.038	0.245 ± 0.004
	$\Delta\mathfrak{C} < 0$	558	51	$\hat{b}_< = 0.772$	0.108	0.864 ± 0.003
	All	6308	616	$\hat{b}_{all} = 0.890$	0.036	0.239 ± 0.004

Table 8.4: Same as Table 8.1 but computed at the nodal plane where $\Delta\mathfrak{C}$ is maximum.

Slip Model	z	p_{norm}	p_{perm}	$\Delta\mathcal{A}$
wald	2.600	0.009	0.006 ± 0.001	-3.460
hernandez	1.534	0.125	0.116 ± 0.003	-0.094
bbcal	1.081	0.280	0.312 ± 0.005	0.909
surfrup	1.136	0.256	0.256 ± 0.004	0.814

Table 8.5: Same as Table 8.2 but computed at the nodal plane where $\Delta\mathfrak{C}$ is maximum.

Finally, as the real value of the effective friction coefficient μ' is uncertain [Kanamori and Brodsky, 2004], the analysis is also done for different values of it [Hardebeck et al., 1998] for the **wald** slip model. These results are exposed in Tables 8.6 and 8.7.

	N	n	$\hat{b}-$**value**	σ	p_{value}
Overall	6027	560	$b_{all} = 0.909$	0.038	0.243 ± 0.004
$\mu' = 0.1$, $\Delta\mathfrak{C} > 0$	5083	488	$\hat{b}_{>} = 0.925$	0.042	0.451 ± 0.005
$\Delta\mathfrak{C} < 0$	944	72	$\hat{b}_{<} = 0.818$	0.096	0.162 ± 0.004
$\mu' = 0.2$, $\Delta\mathfrak{C} > 0$	5157	497	$\hat{b}_{>} = 0.929$	0.042	0.297 ± 0.006
$\Delta\mathfrak{C} < 0$	870	63	$\hat{b}_{<} = 0.783$	0.099	0.712 ± 0.006
$\mu' = 0.4$, $\Delta\mathfrak{C} > 0$	5213	509	$\hat{b}_{>} = 0.927$	0.041	0.313 ± 0.005
$\Delta\mathfrak{C} < 0$	814	51	$\hat{b}_{<} = 0.766$	0.107	0.861 ± 0.003
$\mu' = 0.6$, $\Delta\mathfrak{C} > 0$	5204	512	$\hat{b}_{>} = 0.934$	0.041	0.315 ± 0.005
$\Delta\mathfrak{C} < 0$	823	48	$\hat{b}_{<} = 0.710$	0.102	0.757 ± 0.004
$\mu' = 0.8$, $\Delta\mathfrak{C} > 0$	5138	509	$\hat{b}_{>} = 0.920$	0.041	0.315 ± 0.005
$\Delta\mathfrak{C} < 0$	889	51	$\hat{b}_{<} = 0.815$	0.114	0.632 ± 0.005

Table 8.6: Same as Table 8.1 for different values of the effective friction coefficient μ' in the **wald** slip model.

μ'	z	p_{norm}	p_{perm}	d_{2ks}	p_{2ks}
0.1	1.021	0.307	0.304 ± 0.002	0.158	0.047
0.2	1.949	0.051	0.0495 ± 0.002	0.156	0.041
0.4	2.029	0.042	0.042 ± 0.002	0.159	0.046
0.6	1.991	0.046	0.047 ± 0.002	0.170	0.030
0.8	1.834	0.067	0.069 ± 0.003	0.176	0.024

Table 8.7: Same as Table 8.2 for different values of the effective friction coefficient μ' in the **wald** slip model.

Discussion

The positive Coulomb-stress increase associated to the Landers mainshock triggered a very large number of strike-slip events and also a large number of normal events, but much less thrust events. Although this result seems easy to establish, as it can be obtained without the calculation of $\Delta\mathfrak{C}$ (due to the fact that most of the events have $\Delta\mathfrak{C} > 0$ and thus, this subset dominates the overall statistics) these events have been unambiguously associated to those with positive $\Delta\mathfrak{C}$. On the other side, the events in the opposite regime (with $\Delta\mathfrak{C} < 0$) keep a proportion between normal, strike-slip, and thrust events rather different to the $\Delta\mathfrak{C} > 0$ case, and close to that of the immediately previous record (1987-1992, up to Landers). These results are largely independent on the slip model used to calculate the change in Coulomb stress, the value of μ' and the selection of nodal plane. The b-values of the Gutenberg-Richter law for events with positive $\Delta\mathfrak{C}$ (for which $\hat{b}_{>} \simeq 0.93$) are in general larger than the $b-$values for the events with negative $\Delta\mathfrak{C}$ ($\hat{b}_{<} \simeq 0.87$); nevertheless, this difference is not statistically significant for any of the slip models used to compute the change in the Coulomb stress.

Finally, in a preliminary analysis it has been observed that there is no substantial difference in the fulfilment of the Omori law [Omori, 1894; Utsu, 1961; Utsu et al., 1995] in the two populations of events ($\Delta\mathfrak{C} > 0$ and < 0). Indeed, if one compares this for the two subsets the "characteristic" power-law Omori decay of the rate is found with very similar values of the Omori exponent. Note that this is in disagreement with the rate-and-state formulation [Dieterich, 1994], which does not predict Omori behavior in the case of negative $\Delta\mathfrak{C}$.

Certainly, more research using other mainshocks (for which detailed slip models were available) is necesary, in order to reduce the statistical uncertainty by means of aggregated distributions, which could lead to the detection of small significant differences in both populations of events.

Part IV

Conclusions and perspectives

Chapter 9

Conclusions and perspectives

9.1 Conclusions

The important results of this PhD thesis can be summarized as the development of statistical methods in order to establish avalanche criticality and universality. These methods have been applied to empirical datasets, mainly earthquakes and labquakes.

- Avalanche criticality and universality have been discussed in the introduction. The different limitations and contradictions associated to these terms have been exposed. The importance of rigorously determine power-law exponents in order to characterize criticality and provide a good classification of phenomena in different universality classes has been highlighted.
- Different displacement-driven compression experiments of porous materials have been performed in the Condensed Matter Physics Department of the Universitat de Barcelona. With the aim of determining whether the exponent of the energy distribution of labquakes is affected by the compressional driving mechanism, a first series of displacement-driven compression experiments of Vycor have been performed:
  - The Gutenberg-Richter law is found for the energy distribution of labquakes in the same way it was previously found for force-driven compression experiments [Baró et al., 2013]. Regarding the values of the exponents, it can be concluded that they do not seem to be affected by the driving mechanism in compression experiments. The independence with the driving mechanism has also been found in the measurement of slip events in microcrystals Maaß et al. [2015].
  - When the driving variable turns out to be the displacement, the release of elastic energy is not only expressed by means of AE, but it is also manifested as drops in the force which turn out to be power-law distributed with a compression-rate-dependent exponent. These drops can also be observed in computer simulations

near the big failure event [Kun et al., 2013, 2014]. Nevertheless, some tuning of the disorder should be arranged in simulations in order to replicate a situation with a similar level of heterogeneity as in the experiments in this PhD thesis. Furthermore, force drops and their associated energies have been found to be significantly correlated.

- A statistical procedure to merge different datasets in order to validate the existence of universal power-law exponents across different scales or phenomena has been developed. This methodology can be useful in the study of different complex systems in order to check whether the power-law exponents obtained via maximum likelihood estimation are statistically compatible among them or not. Therefore, this procedure provides a statistical tool that allows to establish whether different complex systems can be classified into the same universality class or not. In this work, the methodology has been applied to the Gutenberg-Ricther law for earthquakes and labquakes:
  - By merging earthquake catalogs, a global power law with a global exponent $\hat{\Gamma}_{\mathcal{M}} = 1.667$ holds for more than eight orders of magnitude in seismic moment (from $m_{min} = 1.9$ to $m_{max} = 7.7$ in moment magnitude). To our knowledge, this is the broadest fitting range that has been found for the Gutenberg-Richter law for earthquakes with a unique value of the exponent [Kagan, 1999]. There are catalogs of tiny mining-induced earthquakes which exhibit a much smaller completeness magnitude [Plenkers et al., 2011] than the ones of natural seismicity used in this work. They were not considered here because they are not currently public and show b values significantly different [Kwiatek et al., 2010] from the one found here, which would result in nonacceptable fits when merging them with the rest of catalogs, possibly pointing to a different universality class.
  - For the case of the amplitude probability distribution of labquakes, a global exponent $\hat{\Gamma}_V = 1.74$ has been found spanning 4.5 orders of magnitude. To our knowledge, this is the broadest fitting range that has been found for the amplitude distribution of AE hits.
  - For the case of the energy of Vycor labquakes, the catalogs (or datasets) have been merged and a Gutenberg-Richter law with exponent $\Gamma = 1.35$ for the energy acoustic emission hits has been found to hold for nine orders of magnitude. As it was proposed by Xu et al. [2019], the labquakes in this material have been found to belong to a different universality class than earthquakes and charcoal labquakes.
  - Earthquake catalogs have been also merged with a charcoal labquake catalog with a global power-law exponent for the energies $\hat{\Gamma} = 1.688$ suggesting that these different systems might be classified into the same universality class. Further investigations involving different observables, such as the distribution of durations, amplitudes and waiting times, might be necessary in order to properly classify charcoal labquakes and real earthquakes into the same universality class.

- A methodology to calculate the change in Coulomb stress induced by a mainshock in the fault plane for every aftershock has been presented. By using different slip models, this procedure has been applied to the Landers event (California, USA, 1992, moment magnitude 7.3). Earthquake models that relate earthquake rates and Coulomb stress after a main event, such as the rate-and-state model, assume that the magnitude distribution of earthquakes is not affected by the change in the Coulomb stress. Several statistical analyses are applied in order to test whether the distribution of magnitudes is sensitive to the sign of the Coulomb-stress increase.

  - Positive Coulomb-stress increase associated to the Landers mainshock triggered a very large number of strike-slip events and also a large number of normal events, but much less thrust events. Although this result seems easy to establish, as it can be obtained without the calculation of $\Delta\mathfrak{C}$ (due to the fact that most of the events have $\Delta\mathfrak{C} > 0$ and thus, this subset dominates the overall statistics), these events have been unambiguously associated to the positive $\Delta\mathfrak{C}$. On the other side, the events in the opposite regime (with $\Delta\mathfrak{C} < 0$) keep a proportion between normal, strike-slip, and thrust events rather different to the $\Delta\mathfrak{C} > 0$ case, and close to that of the immediately previous record (1987-1992, up to Landers). These results are largely independent on the slip model used to calculate the change in Coulomb stress.
  - The b-values of the Gutenberg-Richter law for which the Landers event yielded a positive $\Delta\mathfrak{C}$ (with $b_{>} \simeq 0.93$) are in general larger than the $b-$values for the events with negative $\Delta\mathfrak{C}$ ($b_{<} \simeq 0.87$); nevertheless, this difference is not statistically significant for any of the slip models used to compute the change in the Coulomb stress. A non-significant result teaches that the differences in $b-$values may be spurious and that certainly, more research is necessary in order to overcome statistical limitations.

9.2 Future works derived from this PhD Thesis

Different studies can be performed from the contents exposed in this PhD thesis.

- Different works can be done by using the datasets which have been used in this PhD thesis:
  - To develop rigorous statistical techniques in order to compare exponents in the Omori law both for earthquakes and labquakes [Baró i Urbea, 2015]. Note that, in principle, the Omori law is not a probability density function.
  - To apply rigorous statistical techniques to study the distribution of waiting-times δ both for earthquakes and labquakes. The double power-law distributions which are usually found [Corral, 2004a; Baró et al., 2013] could be studied with more statistical rigour in order to check whether they are compatible with those distributions found for earthquakes.
  - To study the distribution of labquake durations.
  - Inspired by the works trying to model the source of acosutic emission [Yu and Clapp, 1987], spectral analysis of acoustic emission waveforms could be applied in order to explore the existence of corner frequencies in labquakes. The existence of corner frequencies in labquakes could be a new way of measuring labquakes energies. It would be interesting to study how these energies computed from the corner frequency are related to the current definition of AE energy hit.
- Regarding the statistical methods applied to earthquake and labquake datasets, the works developed in this PhD thesis can be helpful in order to:
  - Incorporate new sources of uncertainty in the determination of the power-law exponents. The nature of the uncertainties derived from the different statistical techniques to check avalanche criticality or universality presented in this PhD thesis is purely statistical. However, there exist other uncertainty sources such as the experimental resolution of the experimental devices that should be also taken into account. Although there exists methods that take into account sample-to-sample variability for power-law distributed samples [Deluca et al., 2016], it would be interesting to develop a powerful method encompassing as many source of uncertainty as possible.
  - Study different earthquake catalogs in order to find a broader fitting range of the Gutenberg-Richter law and also to check whether different regions have compatible power-law exponents or not. Particular interest should be paid to those catalogs with small completeness magnitudes. This kind of studies would be of interest in order to statistically strengthen the geological arguments that justify the difference in the b values observed in some regions [Yoder et al., 2012].

  - The labquake energies obtained in the first series of displacement-driven compression experiments in Chapter 6 could also be merged according to the statistical procedure exposed in Chapter 7. If the the null hypothesis that all datasets can be described by a unique global exponent Γ_E is not rejected, this would provide an alternative comparison with the energy exponent derived from force-driven compression.
  - To deepen into the study of bivariate power-law distributions in order to establish relationships among the different avalanche observables.

- A number of extensions and improvements could be incorporated to the approach used in the study of the effects of the Coulomb stress in the seismicity after Landers:
  - Slip models with relatively low resolution in space have been used; so, it would be interesting to know if higher resolution slip models [Olsen et al., 1997; Peyrat et al., 2001] lead to somewhat different values of the strain and the stress, in particular close to the fault.
  - Some authors have argued that real faults should have rather low values of the μ' coefficient [Mulargia and Bizzarri, 2016]. Some check of this has been studied leading to the conclusion that μ' has little influence on the $b-$values. However, it would be interesting to obtain a spatial distribution of μ' based on geological information and check whether the results are robust under the spatial variations of this parameter.
  - Further, in the temporal window of 100 days, the effect of viscoelastic relaxation [Sabadini R., 2016] should be important; so, this would need to be incorporated into the calculation of the stress.
  - Perhaps more relevant but easier to implement would be the contribution to the stress of the triggered events [Felzer et al., 2003]. Knowing the focal mechanism of each of these events allows us to calculate the slip direction, which, together with an estimation of the slip value from the magnitude, can be considered an elementary slip model. From this, the event's contribution to the strain can be computed in the same way as for the mainshock.
  - The study of other aftershock sequences for which both knowledge of focal mechanisms as well as detailed slip models for the mainshock are available is crucial. It may happen, as for the Landers sequence, that the restrictions in space, time, magnitude, and Coulomb-stress change are too many to yield significant results (in our case, the restriction $m \geq 3$ is particularly strong, but necessary for the fulfilment of the Gutenberg-Richter law). In such a case of low statistics, aggregation (stacking) of sequences from different mainshocks could reduce statistical uncertainty and lead to significant results.

Part V

Appendixes

9.3 Source codes for fitting a truncated power-law

9.3.1 R Language

Source code in R which applies the procedure explained in Fig. 5.5 in Sec. 5.5 to a data array "dades". Definitions of all the variables, allocations, comments and printings are included.

```
library(poweRlaw)
library(sfsmisc)
library(stats)

dataoriginal<- # Array of data
sort(dataoriginal)
xmaxabsolute<-max(dataoriginal)
xminabsolute<-min(dataoriginal)
xmax<-max(dataoriginal)
xmin<-min(dataoriginal)
partitions<-5 # Number of partitions per decade
kmax<-1+partitions*log10(xmaxabsolute)  # Maximal index
kmin<-partitions*log10(minimabsolute)   # Minimal index

indexwrite=1 # Writing index
# Define arrays
pvalor<-c()
xminev<-c()
xmaxev<-c()
exponentev<-c()
errorexp<-c()
nev<-c()
pvalor[indexmin]=0

for(j in 1:partitions*log10(xmaxabsolute/xminabsolute)){
  corre<-0
  pvalorparticular<-0

  while(xmax>xmin & pvalorparticular<0.20){
    indexmin=indexmin+1
    corre=corre+1
    xmin<-10^((j+kmin)/parititions)
    xmax<-10^((kmax-corre)/parititions)
    data<-dataoriginal[dataoriginal>=xmin &  dataoriginal<=xmax]
    data<-sort(data)
    n<-length(data)
    loglikelihood<-function(exponent) n*log((1-exponent)/(xmax^(1-exponent)-xmin^(1-exponent)))
    -exponent*sum(log(data))
     exponent<-optimize(loglikelihood,interval=c(0.2,3),maximum =TRUE, tol=1E-3)$maximum
    dkse<-dksecomp(data,xmin,xmax,exponent)
```

```
     nsimulations<-1E4
     pexit<-0
     kssint<-c()
     exponentsint<-c()

     for(i in 1:nsimulations){
        datasint<-c()
        primer<-c()
        ps<-c()
        contadades<-0
        while(contadades<n){
          sint<-xmin/(1-(1-(xmin/xmax)^(exponent-1))*runif(1))^(1/(exponent-1))
          if(sint<=xmax){
            contadades=contadades+1
            datasim[contadades]<-sint
          }
        }
        loglikelihoodsim<-function(exponent) n*log((1-exponent)/
        (xmax^(1-exponent)-xmin^(1-exponent)))-exponent*sum(log(datasint))
        exponentsim<-optimize(loglikelihoodsim,interval=c(0.2,3),maximum =TRUE, tol=1E-3)$maximum
        exponentsint[i]<-exponentsim
        dkssint[i]<-dksecomp(datasint,xmin,xmax,exponentsim
        if(dkssint[i]>=dkse)  pexit=pexit+1
      }

          # Write the results of the fit
    xminev[indexmin]<-xmin
    xmaxev[indexmin]<-xmax
    nev[indexmin]<-n
    exponentev[indexmin]<-exponent
    pvalor[indexmin]<-pexit/nsimulations
    pvalorparticular<-pvalor[indexmin]
    errorexp[indexmin]<-sd(exponentsint)
  }

}


# Functions
#############################################################################################
# Function with the CDF of the truncated power-law
cdf<-function(x,xmin,xmax,alpha) (x^(1-alpha)-xmin^(1-alpha))/(xmax^(1-alpha)-xmin^(1-alpha))
# Function computing the KS distance
dksecomp<-function(dades,xmin,xmax,exponent){
  xs<-sort(dades)
  z<-ecdf(dades)
  n<-length(dades)
  dkse1<-c()
  dkse2<-c()
```

```
  dksemaxpartial<-c()
  for(i in 1:(n-1)){
    dkse+[i]<-abs(z(xs[i])-(cdf(xs[i],xmin,xmax,exponent)))
    dkse-[i]<-abs(z(xs[i-1])-(cdf(xs[i],xmin,xmax,exponent)))
    dksemaxpartial[i]<-max(dkse+[i],dkse-[i])
  }

  dkse<-max(dksemaxpartial)

}
```

9.3.2 C Language

Source code in C which applies the procedure explained in Fig. 5.5 in Sec. 5.5 to a data array "dades". Definitions of all the variables, allocations, comments and printings are included. This code also chooses the fit with largest number of data and the fit with the broadest range.

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <gsl/gsl_rng.h>
#include <gsl/gsl_sort.h>
#include <stdint.h>

int mle(double *dades, double *xmax,double *xmin, double *exponent, double *error,
double* pvalue,gsl_rng * rn, int size, int *n);
double func(double x, double xmin, double xmax, int n, double sumlog);
double rtflsp(double x1, double x2, double xacc, double xmin, double xmax, int n, double sumlog);
int main(){
int i, j;
// Set for random number generator
const gsl_rng_type * T;
gsl_rng * rn;
gsl_rng_env_setup();
T = gsl_rng_default;
rn = gsl_rng_alloc (T);
//Data
const int size=1E3; // Size of data
double *dades;  // Define array of data
dades=(double *)malloc(size*sizeof(double)); // allocate memory
gsl_sort(dades,1,size); // sort data
// Variables for MLE and allocation
double *xmin, *xmax, *exponent, *pvalue, *error;
int *n;
// Allocations
xmin=(double *)malloc(1*sizeof(double));
xmax=(double *)malloc(1*sizeof(double));
```

```
exponent=(double *)malloc(1*sizeof(double));
error=(double *)malloc(1*sizeof(double));
pvalue=(double *)malloc(1*sizeof(double));
n=(int *)malloc(1*sizeof(int));
// Fit parameters of maxim N fit
double xminmaxn, xmaxmaxn,exponentmaximn, errorexpmaximn,pvalormaximn  ;
//  Fit parameters of maxim Range fit
double xminmaxrange, xmaxmaxrange, exponentmaximrange, errorexpmaximrange, pvalormaximrange;
int nmaximrange; // number of data in fit with maximum range
double maximn=1; //initial value of maximum N
double maximrange=1; //initial value of maximum range

double maxim=dades[size-1]; // Determine maximum of your data
double minim=dades[0]; // Determine minimum of your data
double particions=5.0; // Number of subdivisions per decade
int imin=particions*log10(minim); // minimal index
int imax=1+particions*log10(maxim); // maximal index
//Variables for the loop
int indexcorremin=-1; // index for xmin in the loop
int indexcorremax; // index for xmax in the loop
int rejectmain=0; // rejection set to 0

// Loop for xmax
for(indexcorremax=0; indexcorremax< (particions*log10(maxim/minim)); indexcorremax++){

        xmax[0]=pow(10.0, (double)(imax-indexcorremax)/particions);
        indexcorremin=-1;

        do{
                indexcorremin=indexcorremin+1;
                xmin[0]=pow(10.0, (double)(indexcorremin+imin)/particions);

                if(xmax[0]>xmin[0]){
                        // MLE function
                        rejectmain=mle(dades,xmax,xmin,exponent,error,pvalue,rn,size,n);
                        // If the fit is not rejected
                        if(rejectmain==1){
                        // Change the fit with largest N
                                if(n[0]>maximn){
                                        maximn=n[0];
                                        xminmaxn=xmin[0];
                                        xmaxmaxn=xmax[0];
                                        exponentmaximn=exponent[0];
                                        errorexpmaximn=error[0];
                                        pvalormaximn=pvalue[0];
                                }
                                // Change the fit with largest range
                                if((xmax[0]/xmin[0])>maximrange){
                                        maximrange=xmax[0]/xmin[0];
```

```
                                        xminmaxrange=xmin[0];
                                        xmaxmaxrange=xmax[0];
                                        exponentmaximrange=exponent[0];
                                        errorexpmaximrange=error[0];
                                        nmaximrange=n[0];
                                        pvalormaximrange=pvalue[0];
                                }
                        }
                }else{
                        rejectmain=1; // if xmax<xmin, skip
                }
        }while(rejectmain == 0);
}

fprintf(stderr,"Fit largest N \n");
fprintf(stderr,"%lf \t %lf \t %lf \t %lf \t %d \t %lf \n",xminmaxn,xmaxmaxn, exponentmaximn,
errorexpmaximn,maximn, pvalormaximn);
fprintf(stderr,"Fit with largest range \n");
fprintf(stderr,"%lf \t %lf \t %lf \t %lf \t %d \t %lf \n",xminmaxrange,xmaxmaxrange,
exponentmaximrange, errorexpmaximrange,nmaximrange, pvalormaximn);

gsl_rng_free (rn);
return EXIT_SUCCESS;
}
////////////////////////////////////////////////////////////////////////////////////
//// FUNCTIONS
////////////////////////////////////////////////////////////////////////////////////
mle(double *dades, double *xmax,double *xmin, double *exponent, double *error, double *pvalue,
gsl_rng * rn,int size, int *n){
// Definition of variables for the function
int i, j, reject, pexit, sizeactual, conta;
double *dadesactuals, *dadesactualsaux,  xminim, xmaxim,*cumempirica, *cumteorica, sumlog,
distks, distksprima, dmaxempirica ; // Variables of the empirical data
double *vectoru, *cumsint, *cumteoricasint, dmaxsintetic, distkssintprima, distkssint,sumlogsim,
exposim,aux,meansim, meansim2; // Variables for synthetic data
const int nit=1E4; // number of simulations
double pc=0.20; // Threshol p-value
int contador=0;

// Allocation of auxiliar array with data
dadesactualsaux=(double *)malloc(size*sizeof(double));
// Reject is 0 at the beggining
reject=0;

//        Computing the PL exponent
n[0]=0;
sumlog=0;

        for(i=0; i<size; i++){
```

```
                if(dades[i]>=xmin[0] && dades[i]<=xmax[0]){
                        sumlog=sumlog+ log(dades[i]); // sum of logarithm of data
                        dadesactualsaux[contador]=dades[i];
                        contador=contador+1;
                        n[0]=n[0]+1;
                }
        }

sizeactual=n[0];

// Allocations of data, Empirical and theoretical CDFs
dadesactuals=(double *)malloc(sizeactual*sizeof(double));
cumempirica=(double *)malloc((sizeactual)*sizeof(double));
cumteorica=(double *)malloc((sizeactual)*sizeof(double));
        for(i=0; i<sizeactual; i++){
        // Current data after imposing cut-offs
                dadesactuals[i]=dadesactualsaux[i];
        }
free(dadesactualsaux); // Free auxiliar array¡
gsl_sort(dadesactuals,1,sizeactual); //Sorting Data
//xmin and xmax to compute the PL exponent
xminim=xmin[0];
xmaxim=xmax[0];
//Computing the exponent
exponent[0] = rtflsp(1.05, 3, 1E-3,  xminim,  xmaxim,  sizeactual,  sumlog);
//Construct the theoretical and empirical CDF and compute KS distance
dmaxempirica=0; // Initiallize KS distance at 0
        for(i=0; i<sizeactual; i++){
                cumempirica[i]=(double) i/ (double) (sizeactual-1);
                cumteorica[i]=(pow(dadesactuals[i],1.0-exponent[0])- pow(xmin[0],1.0-exponent[0])) /
                (pow(xmax[0],1.0-exponent[0])- pow(xmin[0],1.0-exponent[0]));
                distks=fabs(cumempirica[i]-cumteorica[i]); // KS+
                if(i>0){
                        distksprima=fabs(cumempirica[i-1]-cumteorica[i]);         // KS-
                        if(distksprima>distks){
                        distks=distksprima; // max(KS+,KS-)
                        }
                }
                if(distks>dmaxempirica){
                        dmaxempirica=distks; // Determine the KS distance
                }
        }

        // Allocations fot the variables with synthetic data
        vectoru=(double *)malloc(sizeactual*sizeof(double));
        cumsint=(double *)malloc(sizeactual*sizeof(double));
        cumteoricasint=(double *)malloc(sizeactual*sizeof(double));
        // Initiallize <exponentsim> and <exponenstims2> at 0
        meansim=0;
```

```
meansim2=0;
pexit=0; // number of times in which KSsim>KSemp
for(i=0; i<nit; i++){ // Loop for simulated Datasets
        conta=0;
        while(conta<sizeactual){
                // PL simulated data
                aux=xmin[0] /(pow( 1.0 - gsl_rng_uniform (rn)*(1.0-pow((xmin[0]/xmax[0])
                ,(exponent[0]-1.0)) ), (1.0/(exponent[0]-1.0)))) ;
                vectoru[conta]=aux;
                conta=conta+1;
        }
        gsl_sort(vectoru,1,sizeactual); // Sort the vector
        sumlogsim=0;
        for(j=0; j<sizeactual; j++){
                sumlogsim=sumlogsim+ log(vectoru[j]);
        }
        // Compute the simulated exponent
        exposim= rtflsp(1.05, 3, 1E-3,  xminim,  xmaxim,  sizeactual,  sumlogsim);
        meansim=meansim+exposim;
        meansim2=meansim2+ exposim*exposim;
        dmaxsintetic=0; // Initialize KS distance at 0
        // KS of simulated Data
        for(j=0; j<sizeactual; j++){
                cumsint[j]=(double) j/ (double) (sizeactual-1);
                cumteoricasint[j]=(pow(vectoru[j],1.0-exposim)-
                pow(xmin[0],1.0-exposim))/(pow(xmax[0],1.0-exposim)-
                pow(xmin[0],1.0-exposim));
                distkssint=fabs(cumsint[j]-cumteoricasint[j]);        // KS+ Sim
                if(j>0){
                        distkssintprima=fabs(cumsint[j-1]-cumteoricasint[j]); // KS- Sim
                        if(distkssintprima>distkssint){
                                distkssint=distkssintprima; // max(KS+Sim, KS-Sim)
                        }
                }
                if(distkssint>dmaxsintetic){
                        dmaxsintetic=distkssint; // KS for synthetic data
                }
        }
        if(dmaxsintetic>=dmaxempirica){
                pexit=pexit+1;
        }
}
// <exponentsim> and <exponenstims2>
meansim=meansim/ (double) nit;
meansim2=meansim2/(double) nit;
error[0]=sqrt(meansim2-meansim*meansim); // estimation of the error of the MLE
pvalue[0]=(double) pexit / (double) nit;         // Resulting p-value
// Decision
if(sizeactual<10 || pvalue[0]>=pc){
```

```
                reject=1;
        }
        // free arrays
        free(vectoru);
        free(cumsint);
        free(cumteorica);
        free(cumempirica);
        free(cumteoricasint);

return reject;

}

////////////////////////////////////////////////////////////////////////////
double func(double x, double xmin, double xmax, int n,double sumlog){
double value;
value = sumlog + (double) n/(1- x) -(double) n*( (pow(xmax,1- x)*log(xmax) -
pow(xmin,1- x)*log(xmin)) / (pow(xmax,1- x)-pow(xmin,1- x)) );
return value;
}

////////////////////////////////////////////////////////////////////////////

#include <math.h>
#define MAXIT 30
double rtflsp( double x1, double x2, double xacc, double xmin, double xmax,
int n, double sumlog){
void nrerror(char error_text[]);
int j;
double fl,fh,xl,xh,swap,dx,del,f,rtf;
fl= +sumlog + (double) n/(1- x1) - (double) n*( (pow(xmax,1- x1)*log(xmax) -
pow(xmin,1- x1)*log(xmin)) / (pow(xmax,1.- x1)-pow(xmin,1.- x1)) );
fh= +sumlog + (double) n/(1- x2) - (double) n*( (pow(xmax,1- x2)*log(xmax) -
pow(xmin,1- x2)*log(xmin)) / (pow(xmax,1.- x2)-pow(xmin,1.- x2)) );
if (fl*fh > 0.0);// fprintf(stderr,"Root must be bracketed in rtflsp \n ");
if (fl < 0.0) {
        xl=x1;
        xh=x2;
}else{
        xl=x2;
        xh=x1;
        swap=fl;
        fl=fh;
        fh=swap;
}
dx=xh-xl;
for (j=1;j<=MAXIT;j++) {
        rtf=xl+dx*fl/(fl-fh);
        f=+sumlog + (double) n/(1- rtf) - (double) n*( (pow(xmax,1- rtf)*log(xmax) -
```

```
        pow(xmin,1- rtf)*log(xmin)) / (pow(xmax,1- rtf)-pow(xmin,1- rtf)) );
        if (f < 0.0) {
                del=xl-rtf;
                xl=rtf;
                fl=f;
        } else {
                del=xh-rtf;
                xh=rtf;
                fh=f;
        }
        dx=xh-xl;
        if (fabs(del) < xacc || f == 0.0) return rtf;
};
return 0.0;
}
```

9.4 Source code for merging different datasets

Source code with the methodology for merging two datasets presented in Chpt. 7. By assuming that the global distributions follows an untruncated power-law, the different goodness-of-fit tests are also exposed. This code corresponds to the case where the cut-offs in both datased are fixed.

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <gsl/gsl_rng.h>
#include <gsl/gsl_sort.h>
#include <stdint.h>

double *agcdf(double *dades1, double *dades2, int n1, int n2); // cdf agregada
double ksd(double *dades, int n, double expo, double xmin); // KS distance
double pvalueksdmd(double *dades1, double *dades2, double xmin1, double xmin2, int n1, int n2,
double exponentglobal);
double pvaluecksd(double *dades1, double *dades2, double xmin1, double xmin2, int n1, int n2,
double exponentglobal);

int main(){
int i,j;
// Random number generator setup
const gsl_rng_type * T;
gsl_rng * rn;
gsl_rng_env_setup();
T = gsl_rng_default;
rn = gsl_rng_alloc (T);
```

```
double *data1, *data2, *dadesfinals; // datasets 1 and 2
int n1=1E4;
int n2=1E4;
int nt=n1+n2;
// allocations
dades1=(double *)malloc(n1*sizeof(double));
dades2=(double *)malloc(n2*sizeof(double));
dadesfinals=(double *)malloc((nt)*sizeof(double));
// Cut-offs which are proposed for the analysis
double xmin1=1;
double xmin2=1E3;
for(i=0; i<(nt); i++){
        if(i<n1){
                dadesfinals[i]=dades1[i];
        }
        if((i>=n1) & (i<(nt))){
                dadesfinals[i]=dades2[i-n1];
        }
}
//Sort data
gsl_sort(dades1,1,n1);
gsl_sort(dades2,1,n2);
gsl_sort(dadesfinals,1,nt);
double sumlog1=0;
double sumlog2=0;
for(i=0; i<n1;i++){
        sumlog1=sumlog1+log(dades1[i]);
}
for(i=0; i<n2;i++){
        sumlog2=sumlog2+log(dades2[i]);
}
// Empirical Exponents
double sumlog=sumlog1+sumlog2;
// In this case, the global exponent can be computed analitically
double exponentglobal=1+ (double)(nt)/(sumlog-n1*log(xmin1)-n2*log(xmin2));
double likelihoodglobal=(nt)*log(fabs(exponentglobal-1.0)) - (1.0-exponentglobal)*
(n1*log(xmin1)+n2*log(xmin2))-exponentglobal*sumlog;
double expo1= 1.0 + n1/ (sumlog1-n1*log(xmin1));
double expo2= 1.0 + n2/ (sumlog2-n2*log(xmin2));
double sigmadata1= (expo1-1.0)/sqrt(n1);
double sigmadata2=(expo2-1.0)/sqrt(n1);
// LR Statistic
double piece1=n1*log(expo1-1.0)-n1*(1.0-expo1)*log(xmin1)-expo1*sumlog1;
double piece2=n2*log(expo2-1.0)-n2*(1.0-expo2)*log(xmin2)-expo2*sumlog2;
double likelihoodcomplex= piece1 + piece2;
double likelihoodratio=2.0*fabs(likelihoodcomplex-likelihoodglobal);
double likelihoodratioc=3.84; // In this case

if(likelihoodratio<likelihoodratioc) { // LR filter
```

```
// At this point, one can select any of two statistics
// CKSD or KSDMD
////////////////////////////////////// KSDMD ///////////////////////////////////////
double *cumempirica, *cumteorica;
cumempirica=(double *)malloc((nt)*sizeof(double));
cumteorica=(double *)malloc((nt)*sizeof(double));
double distksdmd;
double distksdmdprima;
double dmaxempirica;
// Call the function to construct the empirical merged CDF
cumempirica=agcdf(dades1, dades2,  n1, n2);
// Now compute KSDMD
for(i=0; i<nt; i++){
        cumteorica[i]=pow((dadesfinals[i]/xmin1),1.0-exponentglobal);
        distksdmd=fabs(cumempirica[i]-cumteorica[i]);
        if(i>0){
                distksdmdprima=fabs(cumempirica[i-1]-cumteorica[i]);
                if(distksdmdprima>distksdmd){
                        distksdmd=distksdmdprima;
                }
        }
        if(distksdmd>dmaxempirica){
        dmaxempirica=distksdmd;
        }
}
// Definitive DKSDMD
double dksempiricadefksdmd;
dksempiricadefksdmd=dmaxempirica;
// Now the p-value of the KSDMDM
double pksdmd=pvalueksdmd(dades1,dades2,xmin1,xmin2, n1,n2,exponentglobal,dksempiricadefksdmd);

////////////////////////////////////// CKSD ///////////////////////////////////////
double cksd, ks1, ks2;
// Particular KS Distances
ks1= ksd(dades1,n1,exponentglobal, xmin1);
ks2= ksd(dades2,n2,exponentglobal, xmin2);
// CKSD
double dksempiricadefcksd;
dksempiricadefcksd= sqrt((double)n1)*ks1 + sqrt((double)n2)*ks2;
//Now the p-value of the CKSD
double pcksd=pvaluecksd(dades1,dades2,xmin1,xmin2, n1,n2,exponentglobal,dksempiricadefcksd);
}else{
fprintf(stderr,"LRT rejected \n");
}
gsl_rng_free (rn);
return EXIT_SUCCESS;
}
```

```
////////////////////////////////////////////////////////////////////////////////
/////////////// FUNCTION ///////////////////////////////////////////////////////
////////////////////////////////////////////////////////////////////////////////

// Function to compute the empirical KS distance for a dataset
double ksd(double *dades, int n, double expo, double xmin){ // KS distance
double distks,distksprima,dmaxempirica,*cumemprica,*cumteorica;
cumempirica=(double *)malloc(n*sizeof(double));
cumteorica=(double *)malloc(n*sizeof(double));

for(i=0; i<(n); i++){
cumempirica[i]=1.0-(double) (i) / (double) (n1-1);
cumteorica[i]=pow((dades[i]/xmin),1.0-expo);
distks=fabs(cumempirica[i]-cumteorica[i]); // KS+
        if(i>0){
                distksprima=fabs(cumempirica[i-1]-cumteorica[i]);//KS-
                if(distksprima>distks){
                        distks=distksprima;
                }
        }
        if(distks>dmaxempirica1){
        dmaxempirica=distks;
        }
}
return dmaxempirica;

}

////////////////////////////////////////////////////////////////////////////////
/////////////// FUNCTION ///////////////////////////////////////////////////////
////////////////////////////////////////////////////////////////////////////////

// function to compute the pvalue of the ksdmd statistic
double pvalueksdmd(double *dades1, double *dades2, double xmin1, double xmin2, int n1, int n2,
double exponentglobal, double distksdmd){
double *dadessim1, *dadessim2, *cumempiricasim, *cumteoricasim, *dadesfinalssim;
dadessim1=(double *)malloc((n1)*sizeof(double));
dadessim2=(double *)malloc((n2)*sizeof(double));
cumempiricasim=(double *)malloc((nt)*sizeof(double));
cumteoricasim=(double *)malloc((nt)*sizeof(double));
dadesfinalssim=(double *)malloc((nt)*sizeof(double));
int pexit=0;
int nsimu=1E4;

for(j=0; j<nsimu; j++){
        double sumlogsim1=0;
        double sumlogsim2=0;
        for(i=0; i<n1;i++){
                dadessim1[i]=xmin1*pow((1-gsl_rng_uniform (rn)),(-1.0/(exponentglobal-1.0)));
```

```
        sumlogsim1=sumlogsim1+log(dadessim1[i]);
}

for(i=0; i<n2;i++){
        dadessim2[i]=xmin2*pow((1-gsl_rng_uniform (rn)),(-1.0/(exponentglobal-1.0)));
        sumlogsim2=sumlogsim2+log(dadessim2[i]);
}
double sumlogsim=sumlogsim1+sumlogsim2;
double exponentglobalsim=1+ (nt)/(sumlogsim-n1*log(xmin1)-n2*log(xmin2));
double dmaxempirica1sim=0;
double dmaxempirica2sim=0;

gsl_sort(dadessim1,1,n1);
gsl_sort(dadessim2,1,n2);

for(i=0; i<(nt); i++){

        if(i<n1){
                dadesfinalssim[i]=dadessim1[i];
        }
        if( (i>=n1) & (i<(nt))){
                dadesfinalssim[i]=dadessim2[i-n1];
        }
}
gsl_sort(dadesfinalssim,1,nt);


double dmaxempiricasim=0;
// Construct the merged CDF
cumempiricasim=agcdf(dadessim1, dadessim2,  n1, n2);

for(i=0; i<nt; i++){
        cumteoricasim[i]=pow((dadesfinalssim[i]/xmin1),1.0-exponentglobalsim);
        distksdmd=fabs(cumteoricasim[i]-cumteoricasim[i]);
        if(i>0){
                distksdmdprima=fabs(cumempiricasim[i-1]-cumteoricasim[i]);
                if(distksdmdprima>distksdmd){
                        distksdmd=distksdmdprima;
                }
        }
        if(distksdmd>dmaxempiricasim){
        dmaxempiricasim=distksdmd;
        }
}

double dkssimadef= dmaxempiricasim;


if(dkssimadef>=dksempiricadef){
```

```
                pexit=pexit+1;
        }
}

double pvalueksdmd= (double)pexit/(double)nsimu;
return pvalueksdmd;
}

////////////////////////////////////////////////////////////////////////////
/////////////// FUNCTION ///////////////////////////////////////////////////
////////////////////////////////////////////////////////////////////////////

// function to compute the pvalue of the cksd statistic
double pvaluecksd(double *dades1, double *dades2, double xmin1, double xmin2, int n1, int n2,
double exponentglobal, double distksdmd){
double *dadessim1, *dadessim2, *cumempiricasim, *cumteoricasim, *dadesfinalssim;
dadessim1=(double *)malloc((n1)*sizeof(double));
dadessim2=(double *)malloc((n2)*sizeof(double));
cumempiricasim=(double *)malloc((nt)*sizeof(double));
cumteoricasim=(double *)malloc((nt)*sizeof(double));
dadesfinalssim=(double *)malloc((nt)*sizeof(double));
int pexit=0;
int nsimu=1E4;

for(j=0; j<nsimu; j++){
        double sumlogsim1=0;
        double sumlogsim2=0;
        for(i=0; i<n1;i++){
                dadessim1[i]=xmin1*pow((1-gsl_rng_uniform (rn)),(-1.0/(exponentglobal-1.0)));
                sumlogsim1=sumlogsim1+log(dadessim1[i]);
        }

        for(i=0; i<n2;i++){
                dadessim2[i]=xmin2*pow((1-gsl_rng_uniform (rn)),(-1.0/(exponentglobal-1.0)));
                sumlogsim2=sumlogsim2+log(dadessim2[i]);
        }
        double sumlogsim=sumlogsim1+sumlogsim2;
        double exponentglobalsim=1+ (nt)/(sumlogsim-n1*log(xmin1)-n2*log(xmin2));
        double dmaxempirica1sim=0;
        double dmaxempirica2sim=0;

        gsl_sort(dadessim1,1,n1);
        gsl_sort(dadessim2,1,n2);

        double dks1sim=ksd(dadessim1,n1,exponentglobalsim, xmin1);
        double dks2sim=ksd(dadessim2,n2,exponentglobalsim, xmin2);

        double dkssimadef= sqrt((double)n1)*dks1sim + sqrt((double)n2)*dks2sim;
```

```
        if(dkssimadef>=dksempiricadef){
                pexit=pexit+1;
        }
}

double pvaluecksd= (double)pexit/(double)nsimu;
return pvaluecksd;
}

////////////////////////////////////////////////////////////////////////////
/////////////// FUNCTION ///////////////////////////////////////////////////
////////////////////////////////////////////////////////////////////////////

// Function that constructs the merged CDF for the KSDMD
double *agcdf(double *data1, double *data2,  int cont1, int cont2){
int i, j;
gsl_sort(data1,1,cont1);
gsl_sort(data2,1,cont2);
double min1=data1[0];
double max1=data2[cont1-1];
double min2=data2[0];
double max2=data2[cont2-1];
double k1=0;
double k2=0;
double cdfaux=0; // auxiliar variable to construct the CDF
double sumweight=0; // sum of weights
int sizet=cont1+cont2;
double *weightdata, *dadesfinals, *cdf;

// Allocations
dadesfinals=(double *)malloc((sizet)*sizeof(double));
weightdata=(double *)malloc((sizet)*sizeof(double));
cdf=(double *)malloc((sizet)*sizeof(double));

for(i=0; i<cont1; i++){
        if((data1[i]>=min2) & (data1[i]<=max1)){
                        k1=k1+1;
        }
}

for(i=0; i<cont2; i++){
        if((data2[i]>=min2) & (data2[i]<=max1)){
                k2=k2+1;
        }
}

for(j=0; j<sizet; j++){
        if(j<cont1){
                datafinal[j]=data1[j];
```

```
        }
        if( (j>=cont1) & (j<(cont1+cont2))){
                datafinal[j]=data2[j-cont1];
        }
}

gsl_sort(datafinal,1,sizet);

for(j=0; j<sizet; j++){

//Datapoints from (1)
if((dadesfinals[j]>=minim1) & (dadesfinals[j]<minim2)){
        weightdata[j]=(1.0/(k1));
}
//Overlapping(1)(2)
if((dadesfinals[j]>=minim2)&(dadesfinals[j]<maxim1)){
        weightdata[j]=(1.0/((k1+k2)));
}
//Datapoints from (2)
if((dadesfinals[j]>=minim2)&(dadesfinals[j]>=maxim1)){
        weightdata[j]=(1.0/((k2)));
}
//CDF
cdf[j]=cdfaux+weightdata[j];
cdfaux=cdf[j];
}

for(i=0; i<sizet; i++){
// Calculo larea que cobreix l'histograma
        sumweight=sumweight+pesdada[i];
}
for(j=0; j<sizet; j++){
        cdf[j]=1.0-(cdf[j]/area);
}
return cdf;
// Free arrays
free(pesdada);
free(dadesfinals);
free(cdf);
}
```

Part VI

List of publications

Part of the original work presented in this PhD thesis is based on the following publications:

- Víctor Navas-Portella, Álvaro Corral, and Eduard Vives, Avalanches and force drops in displacement-driven compression of porous glasses Physical. Review E 94, 033005 (2016) [Navas-Portella et al., 2016].
- Víctor Navas-Portella, Isabel Serra, Álvaro Corral, and Eduard Vives, Increasing power-law range in avalanche amplitude and energy distributions. Physical Review E 97, 022134 (2018) [Navas-Portella et al., 2018]
- Víctor Navas-Portella, Álvaro González, Isabel Serra, Álvaro Corral, and Eduard Vives, Universality of power-law exponents by means of maximum likelihood estimation. Physical Review E 100, 062106 (2019) [Navas-Portella et al., 2019]
- Víctor Navas-Portella, Abigail Jiménez, and Álvaro Corral, No Significant Effect of Coulomb Stress on the Gutenberg-Richter Law after the Landers Earthquake. Scientific Reports 10, 2901 (2020) [Navas-Portella et al., 2020]

During this PhD thesis, the author has also contributed in the following publications:

- Víctor Navas-Portella, and Eduard Vives, Influence of the aspect ratio and boundary conditions on universal finite-size scaling functions in the athermal metastable two-dimensional random field Ising model. Physical Review E, 93,22129 (2016) [Navas-Portella and Vives, 2016]
- Djordje Spasojević, Svetislav Mijatović, Víctor Navas-Portella, and Eduard Vives, Crossover from three-dimensional to two-dimensional systems in the nonequilibrium zero-temperature random-field ising model. Physical Review E 97, 012109 (2018) [Spasojević et al., 2018]

Part VII

Nomenclature

Nomenclature

As a general fact, bold symbols represent vectors or matrices.

$2\mathcal{R}$	Likelihood ratio statistic
α	Significance level
α_p	Productivity exponent
Δ	Bin-width
δ	Waiting time between consecutive events
Γ	Generic exponent of the Global power-law distribution
γ	Generic power-law exponent
$\hat{\boldsymbol{\theta}}$	Array with the set of the estimated parameters characterizing the statistical model
λ	First Lamé elastic modulus
$\mathcal{A}$	Akaike information criterion
$\mathcal{C}(k)$	Number of counts inside the k-th bin of an histogram of counts
$\mathcal{H}$	Hamiltonian
$\mathcal{L}$	Likelihood
$\mathcal{P}$	Probability
$\mathcal{X}$	Generic random variable
$\mathfrak{C}$	Coulomb failure stress
$\mathfrak{d}$	Dip angle in fault geometry
μ	Lamé elastic shear modulus

μ'	Effective pore pressure
Ω	Weight given by the algorithmic procedure to each data point in a merged histogram
$\overline{\ell}$	Mean earthquake slip
π	Omori exponent
$\hat{\boldsymbol{\ell}}$	Slip vector
$\hat{\boldsymbol{n}}$	Normal vector to the fault
$\boldsymbol{\mathcal{I}}$	Expected information matrix
$\boldsymbol{\mathcal{S}}$	Score function
$\boldsymbol{\theta}$	Set of parameters characterizing a distribution F
$\boldsymbol{I}$	Matrix of the second derivatives of the log-likelihood function
$\boldsymbol{r}$	Position vector
$\boldsymbol{u}$	Deformation Field
ψ	Discretized values of the peak voltage
ρ	Rake angle in fault geometry
τ	Tangential or shear stress.
Θ	Strike angle in fault geometry
$\tilde{\boldsymbol{\theta}}$	Array with the set of the real parameters characterizing the statistical model that has generated a dataset
ε_{ij}	Components of the strain Tensor
ς	Source time function
A	Amplitude of the AE avalanche in dB
A_r	Rupture area of an earthquake
b	b-value: Exponent of the Gutenberg-Richter law
D	Avalanche duration.
d	Kolmogorov-Smirnov distance

$d^{(C)}$	Composite Kolmogorov-Smirnov Distance
$d^{(M)}$	Kolmogorov-Smirnov distance of the Merged Datasets
E	Avalanche energy
E_r	Radiated energy by either earthquakes or labquakes
$F_{\mathcal{X},e}$	Estimated empirical Cumulative Distribution Function (CDF) of the random variable $\mathcal{X}$
$f_{\mathcal{X},e}$	Estimated empirical Probability Density Function (PDF) of the random variable $\mathcal{X}$
$F_{\mathcal{X}}$	Probability Distribution or Cumulative Distribution Function (CDF) of the random variable $\mathcal{X}$
$f_{\mathcal{X}}$	Probability density function of the random variable $\mathcal{X}$
H	External driving field
H_0	Null hypothesis
H_1	Alternative hypothesis
h_i	Local quenched random fields in the Random Field Ising model
M	Seismic moment
m_{min}	Completeness magnitude
n	Number of observations in a sample
$n_{\mathcal{L}}$	Number of parameters characterizing the Likelihood Function $\mathcal{L}$
n_{bdec}	Number of bins per decade in a histogram of counts
n_{bins}	Number of bins in a histogram of counts
N_{sim}	Number of Monte Carlo simulations
nn	Nearest neighbours
OM	Orders of magnitude
p_{value}	p-value
R	Amount of disorder in Ising Models

r Activity rate of events (number of events per unit time)

$S_{\mathcal{X}}$ Complementary cumulative distribution function of the random variable $\mathcal{X}$

s_{ij} Components of the stress tensor

t Time

u Random uniform number

V Amplitude of the AE avalanche in V

x Values that the generic random variable $\mathcal{X}$ can take

x_{low} Minimum value in a sample

x_{max} Upper cut-off of a power-law distribution

x_{min} Lower cut-off of a power-law distribution

x_{top} Maximum value in a sample

z Generic statistic used in statistical hypothesis testing

AE Abbreviation for Acoustic emission

AIC Abbreviation for Akaike Information Criterion

CCDF Abbreviation for Complementary cumulative Distribution Function

CDF Abbreviation for Cumulative Distribution Function

CKSD Abbreviation for Composite Kolmogorov-Smirnov Distance

DIC Abbreviation for Disorder Induced Criticality

DIRP Abbreviation for Driven Interfaces in Random Potentials

FHHM Abbreviation for Failure of Highly Heterogeneous Materials

GR Abbreviation for Gutenberg-Richter

HDT Abbreviation for Hit Definition Time

HLT Abbreviation for Hit Lockout Time

i.i.d. Independent Identically Distributed Random Variables

KS Abbreviation for Kolmogorov-Smirnov

KSDMD Abbreviation for Kolmogorov-Smirnov Distance of the Merged Datasets

LRT Abbreviation for Likelihood ratio test

ML Abbreviature for Maximum Likelihood

MLE Abbreviation for Maximum Likelihood Estimation

RFIM Abbreviation for Random Field Ising Model

SOC Abbreviation for Self-Organized Criticality

SOI Abbreviation for Sweeping of an Instability

THR Abbreviation for Threshold

Part VIII

Bibliography

Bibliography

Aban, I. B., Meerschaert, M. M., and Panorska, A. K. (2012). Parameter Estimation for the Truncated Pareto Distribution. *J. Am. Stat. Assoc.*

Abramowitz, M. and Stegun, I. A., editors (1965). *Handbook of Mathematical Functions*. Dover, New York.

Ahern, T. K. (2003). 86 - The FDSN and IRIS Data Management System: Providing Easy Access to Terabytes of Information. *Int. Geophys.*, 81:1645–1655.

Aizenman, M. and Wehr, J. (1989). Rounding of first-order phase transitions in systems with quenched disorder. *Phys. Rev. Lett.*, 62:2503–2506.

Akaike, H. (1973). *Information Theory and an Extension of the Maximum Likelihood Principle*, pages 199–213. Springer New York, New York, NY.

Aki, K. and Richards, P. G. (2002). *Quantitative Seismology, 2nd Ed.*

Alava, M. J., Nukala, P. K. V. V., and Zapperi, S. (2006). Statistical models of fracture. *Advances in Physics*, 55(3-4):349–476.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97.

Alder, B. J. and Wainwright, T. E. (1959). Studies in Molecular Dynamics. I. General Method. *J. Chem. Phys.*, 31(2):459–466.

Alstrøm, P. (1988). Mean-field exponents for self-organized critical phenomena. *Phys. Rev. A*, 38:4905–4906.

Anderson, P. W. (1972). More Is Different. *Science*, 177(4047):393–396.

Angus, J. E. (1994). The Probability Integral Transform and Related Results. *SIAM Rev.*, 36:652–654.

Antonaglia, J., Xie, X., Schwarz, G., Wraith, M., Qiao, J., Zhang, Y., Liaw, P. K., Uhl, J. T., Dahmen, K. A., Peker, A., and Johnson (2014). Tuned Critical Avalanche Scaling in Bulk Metallic Glasses. *Sci. Rep.*, 4:2342.

Bak, P., Christensen, K., Danon, L., and Scanlon, T. (2002). Unified scaling law for earthquakes. *Phys. Rev. Lett.*, 88:178501.

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Phys. Rev. Lett.*, 59:381–384.

Bak, P., Tang, C., and Wiesenfeld, K. (1988). Self-organized criticality. *Phys. Rev. A*, 38:364–374.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Barés, J., Bonamy, D., and Rosso, A. (2019). Seismiclike organization of avalanches in a driven long-range elastic string as a paradigm of brittle cracks. *Phys. Rev. E*, 100:023001.

Barés, J., Wang, D., Wang, D., Bertrand, T., O'Hern, C. S., and Behringer, R. P. (2017). Local and global avalanches in a two-dimensional sheared granular medium. *Phys. Rev. E*, 96:052902.

Baró, J., Corral, A., Illa, X., Planes, A., Salje, E. K. H., Schranz, W., Soto-Parra, D. E., and Vives, E. (2013). Statistical similarity between the compression of a porous material and earthquakes. *Phys. Rev. Lett.*, 110(8):088702.

Baró, J., Shyu, P., Pang, S., Jasiuk, I. M., Vives, E., Salje, E. K. H., and Planes, A. (2016). Avalanche criticality during compression of porcine cortical bone of different ages. *Phys. Rev. E*, 93:053001.

Baró, J. and Vives, E. (2012). Analysis of power-law exponents by maximum-likelihood maps. *Phys. Rev. E*, 85(6):066121.

Baró i Urbea, J. (2015). Avalanches in Out of Equilibrium Systems: Statistical Analysis of Experiments and Simulations. *TDX (Tesis Dr. en Xarxa), Avalanches in Out of Equilibrium Systems: Statistical Analysis of Experiments and Simulations.*

Baró, J., Planes, A., Salje, E. K. H., and Vives, E. (2016). Fracking and labquakes. *Philosophical Magazine*, 96(35):3686–3696.

Battaglia, M., Cervelli, P. F., and Murray, J. R. (2013). dmodels: A matlab software package for modeling crustal deformation near active faults and volcanic centers. *J. Volcanol. Geotherm. Res.*, 254:1 – 4.

Bauke, H. (2007). Parameter estimation for power-law distributions. *Eur. Phys. J. B*, 58(2):167–173.

Beggs, J. M. and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J. Neurosci.*, 23(35):11167–11177.

Ben-Zion, Y. (2008). Collective behavior of earthquakes and faults: Continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Rev. Geophys.*, 46(4).

Bender, B. (1983). Maximum likelihood estimation of b values for magnitude grouped data. *Bull. Seismol. Soc. Am.*, 73(3):831–851.

Bender, C. M. and Orszag, S. A. (1999). *Advanced Mathematical Methods for Scientists and Engineers I.* Springer, New York, NY.

Bentz, D. P., Garboczi, E. J., and Quenard, D. A. (1998). Modelling drying shrinkage in reconstructed porous materials: application to porous Vycor glass. *Modell. Simul. Mater. Sci. Eng.*, 6(3):211–236.

Bernard, C. (1957). *An introduction to the study of experimental medicine*, volume 400. Courier Corporation.

Bernt Aadnøy, R. L. (2010). *Failure Criteria.* Elsevier, New York, NY.

Bertotti, G. and Mayergoyz, I. D. (2006). *The science of hysteresis: Hysteresis in materials*, volume 3. Gulf Professional Publishing.

Bilek, S. L., Lay, T., and Ruff, L. J. (2004). Radiated seismic energy and earthquake source duration variations from teleseismic source time functions for shallow subduction zone thrust earthquakes. *J. Geophys. Res.: Solid Earth*, 109(B9).

Bonamy, D., Santucci, S., and Ponson, L. (2008). Crackling dynamics in material failure as the signature of a self-organized dynamic phase transition. *Phys. Rev. Lett.*, 101:045501.

Bordeu, I., Amarteifio, S., Garcia-Millan, R., Walter, B., Wei, N., and Pruessner, G. (2019). Volume explored by a branching random walk on general graphs. *Sci. Rep.*, 9(15590):1–9.

Bormann, P. (2015). Are new data suggesting a revision of the current mw and me scaling formulas? *Journal of Seismology*, 19(4):989–1002.

Bramwell, S. T., Holdsworth, P. C. W., and Pinton, J.-F. (1998). Universality of rare fluctuations in turbulence and critical phenomena. *Nature*, 396(6711):552–554.

Brent, R. P. (2002). *Algorithms for minimization without derivatives.* Dover Publications.

Brodsky, E. E. and Kanamori, H. (2001). Elastohydrodynamic lubrication of faults. *J. Geophys. Res.: Solid Earth*, 106(B8):16357–16374.

Brune, J. N. (1968). Seismic moment, seismicity, and rate of slip along major fault zones. *J. Geophys. Res.*, 73(2):777–784.

Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference.* Springer-Verlag New York.

Cagniard De La Tour, C. (1822). Supercritical fluids. *Ann. Chim. Phys*, 21:127–132.

Caldarelli, G., Di Tolla, F. D., and Petri, A. (1996). Self-organization and annealed disorder in a fracturing process. *Phys. Rev. Lett.*, 77:2503–2506.

Callen, H. B. (1985). *Thermodynamics and an Introduction to Thermostatistics.* Wiley.

Carbone, G. and Giannoccaro, I. (2015). Model of human collective decision-making in complex environments. *Eur. Phys. J. B*, 88(12):1–10.

Carrillo, L., Mañosa, L., Ortín, J., Planes, A., and Vives, E. (1998). Experimental evidence for universality of acoustic emission avalanche distributions during structural transitions. *Phys. Rev. Lett.*, 81(9):1889.

Carroll, R., Lee, C., Tsai, C.-W., Yeh, J.-W., Antonaglia, J., Brinkman, B. A. W., LeBlanc, M., Xie, X., Chen, S., Liaw, P. K., and Dahmen, K. A. (2015). Experiments and Model for Serration Statistics in Low-Entropy, Medium-Entropy, and High-Entropy Alloys. *Sci. Rep.*, 5:16997.

Cattania, C., Hainzl, S., Wang, L., Enescu, B., and Roth, F. (2015). Aftershock triggering by postseismic stresses: A study based on Coulomb rate-and-state models. *J. Geophys. Res.*, 120:2388–2407.

Cattania, C., Hainzl, S., Wang, L., Roth, F., and Enescu, B. (2014). Propagation of Coulomb stress uncertainties in physics-based aftershock models. *J. Geophys. Res.*, 119:7846–7864.

Cesca, S., Grigoli, F., Heimann, S., González, A., Buforn, E., Maghsoudi, S., Blanch, E., and Dahm, T. (2014). The 2013 September–October seismic sequence offshore Spain: a case of seismicity triggered by gas injection? *Geophys. J. Int.*, 198(2):941–953.

Chan, C. H., Wu, Y. M., and Wang, J. P. (2012). Earthquake forecasting through a smoothing kernel and the rate-and-state friction law: Application to the Taiwan region. *Nat. Hazards Earth Syst. Sci.*, 12:1–13.

Chave, A. D. (2017). *Computational Statistics in the Earth Sciences: With Applications in MATLAB*. Cambridge University Press.

Cheng, J., Adamic, L., Dow, P. A., Kleinberg, J. M., and Leskovec, J. (2014). *Can cascades be predicted?* ACM.

Christensen, K. and Moloney, N. R. (2005). *Complexity and Criticality*. Imperial College Press, London.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM review*, 51(4):661–703.

Cocco, M. and Rice, J. R. (2002). Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions. *J. Geophys. Res.: Solid Earth*, 107(B2):ESE 2–1–ESE 2–17.

Concise (2008). *Kolmogorov–Smirnov Test*, pages 283–287. Springer New York.

Corning (2020). Corning | Materials Science Technology and Innovation. https://www.corning.com/worldwide/en.html.

Corominas-Murtra, B., Hanel, R., and Thurner, S. (2015). Understanding scaling through history-dependent processes with collapsing sample space. *Proc. Natl. Acad. Sci. U.S.A.*, 112(17):5348–5353.

Corral, A. (2003). Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Phys. Rev. E*, 68:035102.

Corral, A. (2004a). Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys. Rev. Lett.*, 92:108501.

Corral, A. (2004b). Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A*, 340(4):590 – 597. Complexity and Criticality: in memory of Per Bak (1947–2002).

Corral, A. (2005a). Mixing of rescaled data and bayesian inference for earthquake recurrence times. *Nonlinear Proc. Geoph.*, 12(1):89–100.

Corral, A. (2005b). Renormalization-group transformations and correlations of seismicity. *Phys. Rev. Lett.*, 95:028501.

Corral, A. (2015). Scaling in the timing of extreme events. *Chaos, Solitons & Fractals*, 74:99 – 112. Extreme Events and its Applications.

Corral, A., Boleda, G., and Ferrer-i-Cancho, R. (2015). Zipf's law for word frequencies: Word forms versus lemmas in long texts. *PLoS ONE*, 10(7):e0129031.

Corral, A. and Christensen, K. (2006). Comment on "earthquakes descaled: On waiting time distributions and scaling laws". *Phys. Rev. Lett.*, 96:109801.

Corral, A. and Font-Clos, F. (2013). *Self-Organized Criticality Systems*, chapter Branching processes, criticality, and self-organization: application to natural hazards. Aschwanden, M.

Corral, A., Garcia-Millan, R., and Font-Clos, F. (2016). Exact derivation of a finite-size scaling law and corrections to scaling in the geometric galton-watson process. *PloS one*, 11(9):e0161586.

Corral, A., Garcia-Millan, R., Moloney, N. R., and Font-Clos, F. (2018). Phase transition, scaling of moments, and order-parameter distributions in brownian particles and branching processes with finite-size effects. *Phys. Rev. E*, 97(6):062156.

Corral, A. and González, Á. (2019). Power Law Size Distributions in Geoscience Revisited. *Earth Space Sci.*, 6(5):673–697.

Corral, A. and Paczuski, M. (1999). Avalanche Merging and Continuous Flow in a Sandpile Model. *Phys. Rev. Lett.*, 83(3):572–575.

Corral, A., Telesca, L., and Lasaponara, R. (2008). Scaling and correlations in the dynamics of forest-fire occurrence. *Phys. Rev. E*, 77:016101.

Costa, L. S., Lenzi, E. K., Mendes, R. S., and Ribeiro, H. V. (2016). Extensive characterization of seismic laws in acoustic emissions of crumpled plastic sheets. *Europhys. Lett.*, 114(5):59002.

Cote, P. J. and Meisel, L. V. (1991). Self-organized criticality and the Barkhausen effect. *Phys. Rev. Lett.*, 67:1334–1337.

Cubero, R. J., Jo, J., Marsili, M., Roudi, Y., and Song, J. (2019). Statistical criticality arises in most informative representations. *J. Stat. Mech.: Theory Exp.*, 2019(6):063402.

Dahmen, K. and Sethna, J. P. (1993). Hysteresis loop critical exponents in 6-ϵ dimensions. *Phys. Rev. Lett.*, 71:3222–3225.

Dahmen, K. and Sethna, J. P. (1996). Hysteresis, avalanches, and disorder-induced critical scaling: A renormalization-group approach. *Phys. Rev. B*, 53(22):14872.

Dahmen, K. A., Ben-Zion, Y., and Uhl, J. T. (2009). Micromechanical model for deformation in solids with universal predictions for stress-strain curves and slip avalanches. *Phys. Rev. Lett.*, 102:175501.

Dahmen, K. A., Ben-Zion, Y., and Uhl, J. T. (2011). A simple analytic theory for the statistics of avalanches in sheared granular materials. *Nat. Phys.*, 7(7):554–557.

Daley, D. J. and Vere-Jones, D. (2008). *An introduction to the theory of point processes. Vol. II.* Probability and its Applications (New York). Springer, New York, second edition. General theory and structure.

Dalla Torre, F. H., Klaumünzer, D., Maaß, R., and Löffler, J. F. (2010). Stick–slip behavior of serrated flow during inhomogeneous deformation of bulk metallic glasses. *Acta Mater.*, 58(10):3742–3750.

Daniels, H. E. and Jeffreys, H. (1945). The statistical theory of the strength of bundles of threads. i. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 183(995):405–435.

Davidsen, J. and Goltz, C. (2004). Are seismic waiting time distributions universal? *Geophys. Res. Lett.*, 31(21).

Davidsen, J., Stanchits, S., and Dresen, G. (2007). Scaling and universality in rock fracture. *Phys. Rev. Lett.*, 98:125502.

De Haan, L. and Ferreira, A. (2007). *Extreme value theory: an introduction*. Springer Science & Business Media.

del Castillo, J. and Puig, P. (1999). The best test of exponentiality against singly truncated normal alternatives. *J. Am. Stat. Assoc.*, 94(446):529–532.

Deluca, A. and Corral, A. (2013). Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophys.*, 61(6):1351–1394.

Deluca, A., Moloney, N. R., and Corral, A. (2015). Data-driven prediction of thresholded time series of rainfall and self-organized criticality models. *Phys. Rev. E*, 91(5):052808.

Deluca, A., Puig, P., and Corral, A. (2016). Testing universality in critical exponents: The case of rainfall. *Phys. Rev. E*, 93(4):042301.

Dessai, S. and Walter, M. E. (2000). Self-organised criticality and the atmospheric sciences: selected review, new findings and future directions. *XE Extreme events: developing a research agenda for the 21st century*, pages 34–44.

Detcheverry, F., Rosinberg, M. L., and Tarjus, G. (2005). Metastable states and T=0 hysteresis in the random-field. *Eur. Phys. J. B*, 44(3):327–343.

Devroye, L. (1986). *Non-Uniform Random Variate Generation(originally published with.* Springer-Verlag.

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. .Geophys. Res.: Solid Earth*, 99(B2):2601–2618.

Dieterich, J., Cayol, V., and Okubo, P. (2000). The use of earthquake rate changes as a stressmeter at Kilauea volcano. *Nature*, 408:457–460.

Dobson, I. (2012). Estimating the propagation and extent of cascading line outages from utility data with a branching process. *IEEE Transactions on Power Systems*, 27(4):2146–2155.

Doornbos, D. (1982). Seismic source spectra and moment tensors. *Physics of the Earth and Planetary Interiors*, 30(2):214 – 227. Special Issue Earthquake Algorithms.

Dussauge, C., Grasso, J.-R., and Helmstetter, A. (2003). Statistical analysis of rockfall volume distributions: Implications for rockfall dynamics. *J. Geophys. Res. Solid Earth*, 108(B6).

Dziewonski, A. M., T.-A. C. and Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res*, 86:2825–2852.

Ekström, G., M. N. and Dziewonski, A. M. (2012). The global cmt project 2004-2010: Centroid-moment tensors for 13,017 earthquakes. *Phys. Earth Planet. Inter.*, 1.

Embrechts, P., Resnick, S. I., and Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41.

Falk, J., Jensen, M. H., and Sneppen, K. (1994). Intermittent dynamics and self-organized depinning in propagating fronts. *Phys. Rev. E*, 49:2804–2808.

Feller, W. (1948). On the kolmogorov-smirnov limit theorems for empirical distributions. *Ann. Math. Statist.*, 19(2):177–189.

Felzer, K. R., Abercrombie, R. E., and Ekström, G. (2003). Secondary Aftershocks and Their Importance for Aftershock Forecasting. *Bull. Seismol. Soc. Am.*, 93(4):1433–1448.

Felzer, K. R. and Brodsky, E. E. (2005). Testing the stress shadow hypothesis. *J. Geophys. Res.: Solid Earth*, 110(B5).

Font-Clos, F., Pruessner, G., Moloney, N. R., and Deluca, A. (2015). The perils of thresholding. *New Journal of Physics*, 17(4):043066.

Friedman, N., Ito, S., Brinkman, B. A. W., Shimono, M., DeVille, R. E. L., Dahmen, K. A., Beggs, J. M., and Butler, T. C. (2012). Universal critical dynamics in high resolution neuronal avalanche data. *Phys. Rev. Lett.*, 108(20):208102.

Frigg, R. (2003). Self-organised criticality—what it is and what it isn't. *Studies in History and Philosophy of Science Part A*, 34(3):613–632.

Frontera, C. and Vives, E. (1999). Numerical signs for a transition in the two-dimensional random field Ising model at $T = 0$. *Phys. Rev. E*, 59:R1295–R1298.

Garcia-Millan, R., Pausch, J., Walter, B., and Pruessner, G. (2018). Field-theoretic approach to the universality of branching processes. *Phys. Rev. E*, 98:062107.

Garcimartín, A., Mankoc, C., Janda, A., Arévalo, R., Pastor, J. M., Zuriguel, I., and Maza, D. (2009). Flow and Jamming of Granular Matter Through an Orifice. *SpringerLink*, pages 471–486.

Garcimartín, A., Zuriguel, I., Janda, A., and Maza, D. (2011). Fluctuations of grains inside a discharging two-dimensional silo. *Phys. Rev. E*, 84:031309.

Gardiner, C. (2009). *Stochastic Methods - A Handbook for the Natural and Social Sciences | Crispin Gardiner | Springer*. Springer-Verlag Berlin Heidelberg.

Gautschi, G. (2002). *Piezoelectric Sensorics*. Springer, Berlin, Heidelberg.

Gillespie, C. S. (2015). Fitting heavy tailed distributions: The poweRlaw package. *J. Stat. Software*, 64(2):1–16.

Ginzburg, V. and Landau, L. (1965). On the theory of superconductivity. In HAAR, D. T., editor, *Collected Papers of L.D. Landau*, pages 546 – 568. Pergamon.

Gleeson, J. P., Ward, J. A., O'Sullivan, K. P., and Lee, W. T. (2014). Competition-induced criticality in a model of meme popularity. *Phys. Rev. Lett.*, 112:048701.

Godano, C. and Pingue, F. (2000). Is the seismic moment–frequency relation universal? *Geophys. J. Int.*, 142(1):193–198.

Goel, S., Anderson, A., Hofman, J., and Watts, D. J. (2015). The structural virality of online diffusion. *Management Science*, 62(1):180–196.

Goff, J. A. and Holliger, K. (2003). *Heterogeneity in the Crust and Upper Mantle - Nature, Scaling, and Seismic Properties | John A. Goff | Springer*. Springer US.

Goldstein, H. (1980). *Classical Mechanics*. Addison-Wesley.

Goldstein, M. L., Morris, S. A., and Yen, G. G. (2004). Problems with fitting to the power-law distribution. *Eur. Phys. J. B*, 41(2):255–258.

Gomberg, J., Reasenberg, P. A., Bodin, P., and Harris, R. A. (2001). Earthquake triggering by seismic waves following the Landers and Hector Mine earthquakes. *Nature*, 411(6836):462–466.

Good, P. I. (2011). *Resampling Methods*. Birkhäuser, Boston, 3rd edition.

Goodfellow, S. D. and Young, R. P. (2014). A laboratory acoustic emission experiment under in situ conditions. *Geophys. Res. Lett.*, 41(10):3422–3430.

Grabowski, A. and Kosiński, R. A. (2006). Ising-based model of opinion formation in a complex network of interpersonal interactions. *Physica A*, 361(2):651–664.

Hamilton, J. D. (1994). *Time series analysis*, volume 2. Princeton university press Princeton, NJ.

Hanks, T. C. and Kanamori, H. (1979). A moment magnitude scale. *J. Geophys. Res.: Solid Earth*, 84(B5):2348–2350.

Hansen, A., Hemmer, P. C., and Pradhan, S. (2015). *The fiber bundle model: modeling failure in materials*. John Wiley & Sons.

Hardebeck, J. L., Nazareth, J. J., and Hauksson, E. (1998). The static stress change triggering model: Constraints from two Southern California aftershock sequences. *J. Geophys. Res.*, 103:24,427–24,437.

Hardebeck, J. L. and Shearer, P. M. (2002). A new method for determining first-motion focal mechanisms. *Bull. Seismol. Soc. Am.*, 92:2264–2276.

Hargitai, C., Shtrikman, S., and Wohlfarth, E. (1972). Range of validity of the landau theory of phase transitions for very weak itinerant ferromagnets. *Phys. Lett. A*, 39(2):87 – 88.

Harris, P. (1999). On charcoal. *Interdisciplinary Science Reviews*, 24(4):301–306.

Haskell, N. A. (1966). Total energy and energy spectral density of elastic wave radiation from propagating faults. Part II. A statistical source model. *Bull. Seismol. Soc. Am.*, 56(1):125–140.

Hauksson, E., Jones, L. M., Hutton, K., and Eberhart-Phillips, D. (1993). The 1992 Landers earthquake sequence: Seismological observations. *J. Geophys. Res.: Solid Earth*, 98(B11):19835–19858.

Helmstetter, A. (2003). Is earthquake triggering driven by small earthquakes? *Phys. Rev. Lett.*, 91:058501.

Hergarten, S. (2002). *Self-Organized Criticality in Earth Systems*. Springer, Berlin, Heidelberg.

Hernandez, B., Cotton, F., and Campillo, M. (1999). Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 landers earthquake. *J. Geophys. Res.: Solid Earth*, 104(B6):13083–13099.

Hidalgo, R. C., Kun, F., Kovács, K., and Pagonabarraga, I. (2009). Avalanche dynamics of fiber bundle models. *Phys. Rev. E*, 80:051108.

Hill, D. P., Reasenberg, P. A., Michael, A., Arabaz, W. J., Beroza, G., Brumbaugh, D., Brune, J. N., Castro, R., Davis, S., dePolo, D., Ellsworth, W. L., Gomberg, J., Harmsen, S., House, L., Jackson, S. M., Johnston, M. J. S., Jones, L., Keller, R., Malone, S., Munguia, L., Nava, S., Pechmann, J. C., Sanford, A., Simpson, R. W., Smith, R. B., Stark, M., Stickney, M., Vidal, A., Walter, S., Wong, V., and Zollweg, J. (1993). Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science*, 260(5114):1617–1623.

Hinrichsen, H. (2000). Non-equilibrium critical phenomena and phase transitions into absorbing states. *Adv. Phys.*, 49(7):815–958.

Hirata, T. (1987). Omori's power law aftershock sequences of microfracturing in rock fracture experiment. *J. Geophys. Res.: Solid Earth*, 92(B7):6215–6221.

Hohenberg, P. C. and Halperin, B. I. (1977). Theory of dynamic critical phenomena. *Rev. Mod. Phys.*, 49(3):435–479.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558.

Horgan, J. (1995). From complexity to perplexity. *Scientific American*, 272(6):104–109.

Huang, K. (2000). *Statistical Mechanics*. John Wiley and Sons.

Hutton, K., Woessner, J., and Hauksson, E. (2010). Earthquake monitoring in Southern California for seventy-seven years (1932-2008). 100(2):423.

Ide, S. and Beroza, G. C. (2001). Does apparent stress vary with earthquake size? *Geophys. Res. Lett.*, 28(17):3349–3352.

Illa, X., Rosinberg, M.-L., and Vives, E. (2006). Influence of the driving mechanism on the response of systems with athermal dynamics: The example of the random-field Ising model. *Phys. Rev. B*, 74:224403.

Illa, X., Winkelmayer, P., and Vives, E. (2015). Local strain variability and force fluctuations during the martensitic transition under different driving mechanisms. *Phys. Rev. B*, 92(18):184107.

Iori, G. (1999). Avalanche dynamics and trading friction effects on stock market returns. *Int. J. Mod. Phys. C*, 10(06):1149–1162.

Ishibe, T., Ogata, Y., Tsuruoka, H., and Satake, K. (2017). Testing the Coulomb stress triggering hypothesis for three recent megathrust earthquakes. *Geosci. Lett.*, 4.

Ishibe, T., Satake, K., Sakai, S., Shimazaki, K., Tsuruoka, H., Yokota, Y., Nakagawa, S., and Hirata, N. (2015). Correlation between coulomb stress imparted by the 2011 tohoku-oki earthquake and seismicity rate change in kanto, japan. *Geophys. J. Int.*, 201(1):112–134.

Ising, E. (1925). Beitrag zur Theorie des Ferromagnetismus. *Z. Angew. Phys.*, 31(1):253–258.

Izutani, Y. and Kanamori, H. (2001). Scale-dependence of seismic energy-to-moment ratio for strike-slip earthquakes in japan. *Geophys. Res. Lett.*, 28(20):4007–4010.

James, F. (2006). *Statistical methods in experimental physics*. World Scientific Publishing Company.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2014). *An Introduction to Statistical Learning: with Applications in R*. Springer Publishing Company, Incorporated.

Janićević, S., Jovković, D., Laurson, L., and Spasojević, D. (2018). Threshold-induced correlations in the Random Field Ising Model. *Sci. Rep.*, 8(2571):1–9.

Janićević, S., Laurson, L., Måløy, K. J., Santucci, S., and Alava, M. J. (2016). Interevent correlations from avalanches hiding below the detection threshold. *Phys. Rev. Lett.*, 117:230601.

Janićević, S., Mijatović, S., and Spasojević, D. (2017). Critical behavior of the two-dimensional nonequilibrium zero-temperature random field Ising model on a triangular lattice. *Phys. Rev. E*, 95:042131.

Jensen, H. J. (1998). *Self-organized criticality: emergent complex behavior in physical and biological systems*, volume 10. Cambridge university press.

Jones, L. M. (1994). Foreshocks, aftershocks, and earthquake probabilities: Accounting for the Landers earthquake. *Bull. Seismol. Soc. Am.*, 84(3):892.

Jones, R. A. L., Jones, R., Jones, R. A., et al. (2002). *Soft condensed matter*, volume 6. Oxford University Press.

Kadanoff, L. P. (1966). Scaling laws for Ising models near T_c. *Physics Physique Fizika*, 2:263–272.

Kadanoff, L. P., Götze, W., Hamblen, D., Hecht, R., Lewis, E. A. S., Palciauskas, V. V., Rayl, M., Swift, J., Aspnes, D., and Kane, J. (1967). Static phenomena near critical points: Theory and experiment. *Rev. Mod. Phys.*, 39:395–431.

Kadanoff, L. P., Nagel, S. R., Wu, L., and Zhou, S.-m. (1989). Scaling and universality in avalanches. *Phys. Rev. A*, 39:6524–6537.

Kagan, Y. Y. (1999). *Universality of the Seismic Moment-frequency Relation*, pages 537–573. Wyss M., Shimazaki K., Ito A. (eds) . Pageoph Topical Volumes. Birkhäuser, Basel.

Kagan, Y. Y. (2010). Earthquake size distribution: Power-law with exponent $\beta = 12$? *Tectonophysics*, 490(1):103 – 114.

Kagan, Y. Y. (2013). *Earthquakes: models, statistics, testable forecasts*. John Wiley & Sons.

Kamer, Y. and Hiemer, S. (2015). Data-driven spatial b value estimation with applications to california seismicity: To b or not to b. *J. Geophys. Res.: Solid Earth*, 120(7):5191–5214.

Kanamori, H. and Brodsky, E. E. (2004). The physics of earthquakes. *Rep. Prog. Phys.*, 67:1429–1496.

Kanamori, H. and Heaton, T. H. (2013). *Microscopic and Macroscopic Physics of Earthquakes*, pages 147–163. American Geophysical Union (AGU).

Kardar, M. (1998). Nonequilibrium dynamics of interfaces and lines. *Phys. Rep.*, 301(1):85–112.

King, G., Stein, R. S., and Lin, J. (1994). Static stress changes and the triggering of earthquakes. *Bull. Seismol. Soc. Am.*, 84:935–953.

Klafter, J. and Sokolov, I. M. (2011). *First steps in random walks: from tools to applications*. Oxford University Press.

Kleinbaum, D. G. and Klein, M. (2012). *Survival Analysis - A Self-Learning Text, Third Edition | David G. Kleinbaum | Springer*. Springer-Verlag New York.

Klemm, K., Eguíluz, V. M., Toral, R., and San Miguel, M. (2003). Nonequilibrium transitions in complex networks: A model of social interaction. *Phys. Rev. E*, 67:026120.

Knopoff, L. and Kagan, Y. Y. (1977). Analysis of the theory of extremes as applied to earthquake problems. *J. Geophys. Res.*, 82(36):5647–5657.

Kolmogorov, A. (1933). Sulla determinazione empirica di una lgge di distribuzione. *Inst. Ital. Attuari, Giorn.*, 4:83–91.

Kornei, K. (2019). Nuclear Bomb or Earthquake? Explosions Reveal the Differences - Eos. https://eos.org/articles/nuclear-bomb-or-earthquake-explosions-reveal-the-differences.

Koslowski, M., LeSar, R., and Thomson, R. (2004). Avalanches and Scaling in Plastic Deformation. *Phys. Rev. Lett.*, 93(12):125502.

Kun, F. and Herrmann, H. J. (1999). Transition from damage to fragmentation in collision of solids. *Phys. Rev. E*, 59:2623–2632.

Kun, F., Hidalgo, R. C., Raischel, F., and Herrmann, H. J. (2006). Extension of fibre bundle models for creep rupture and interface failure. *Int. J. Fracture*, 140(1):255.

Kun, F., Varga, I., Lennartz-Sassinek, S., and Main, I. G. (2013). Approach to failure in porous granular materials under compression. *Phys. Rev. E*, 88:062207.

Kun, F., Varga, I., Lennartz-Sassinek, S., and Main, I. G. (2014). Rupture cascades in a discrete element model of a porous sedimentary rock. *Phys. Rev. Lett.*, 112:065501.

Kuntz, M. C., Perkovic, O., Dahmen, K. A., Roberts, B. W., and Sethna, J. P. (1999). Hysteresis, avalanches, and noise. *Comput. Sci. Eng.*, 1(4):73–81.

Kusanagi, H., Kimura, H., and Sasaki, H. (2005). Acoustic emission characteristics during magnetization of ferromagnetic materials. *J. Acoust. Soc. Am.*, 64(S1):S175.

Kwiatek, G., Plenkers, K., Nakatani, M., Yabe, Y., Dresen, G., and JAGUARS-Group (2010). Frequency-Magnitude Characteristics Down to Magnitude -4.4 for Induced Seismicity Recorded at Mponeng Gold Mine, South Africa. *Bull. Seismol. Soc. Am.*, 100(3):1165–1173.

Lancia, C. and Scoppola, B. (2013). Equilibrium and non-equilibrium Ising models by means of pca. *J. Stat. Phys.*, 153:641–653.

Landau, L. (1936). The Theory of Phase Transitions. *Nature*, 138(3498):840–841.

Larson, K. M., Freymueller, J. T., and Philipsen, S. (1997). Global plate velocities from the Global Positioning System. *J. Geophys. Res. Solid Earth*, 102(B5):9961–9981.

Laurson, L., Illa, X., and Alava, M. J. (2009). The effect of thresholding on temporal avalanche statistics. *J. Stat. Mech. - Theory E.*, 2009(01):P01019.

Lautrup, B. (2011). *Physics of Continuous Matter*. CRC Press, 2nd edition.

Lay, T. and Wallace, T. (1995). *Modern Global Seismology*. ISSN. Elsevier Science.

LeBlanc, M., Angheluta, L., Dahmen, K., and Goldenfeld, N. (2012). Distribution of maximum velocities in avalanches near the depinning transition. *Phys. Rev. Lett.*, 109(10):105702.

Lebyodkin, M. A., Brechet, Y., Estrin, Y., and Kubin, L. P. (1995). Statistics of the catastrophic slip events in the Portevin-Le Châtelier effect. *Phys. Rev. Lett.*, 74(23):4758–4761.

Lebyodkin, M. A., Shashkov, I. V., Lebedkina, T. A., and Gornakov, V. S. (2017). Experimental investigation of the effect of thresholding on temporal statistics of avalanches. *Phys. Rev. E*, 95:032910.

Lebyodkin, M. A., Shashkov, I. V., Lebedkina, T. A., Mathis, K., Dobron, P., and Chmelik, F. (2013). Role of superposition of dislocation avalanches in the statistics of acoustic emission during plastic deformation. *Phys. Rev. E*, 88:042402.

Leschhorn, H. and Tang, L.-H. (1994). Avalanches and correlations in driven interface depinning. *Phys. Rev. E*, 49(2):1238–1245.

Lherminier, S., Planet, R., Vehel, V. L. d., Simon, G., Vanel, L., Måløy, K. J., and Ramos, O. (2019). Continuously sheared granular matter reproduces in detail seismicity laws. *Phys. Rev. Lett.*, 122:218501.

Lilliefors, H. W. (1967). On the kolmogorov-smirnov test for normality with mean and variance unknown. *J. Am. Stat. Assoc.*, 62(318):399–402.

Lilly, M. P., Finley, P. T., and Hallock, R. B. (1993). Memory, congruence, and avalanche events in hysteretic capillary condensation. *Phys. Rev. Lett.*, 71:4186–4189.

Lilly, M. P., Wootters, A. H., and Hallock, R. B. (1996). Spatially extended avalanches in a hysteretic capillary condensation system: Superfluid ^{4}He in nuclepore. *Phys. Rev. Lett.*, 77:4222–4225.

Lindman, M., Jonsdottir, K., Roberts, R., Lund, B., and Bödvarsson, R. (2005). Earthquakes descaled: On waiting time distributions and scaling laws. *Phys. Rev. Lett.*, 94:108501.

Lippiello, E., de Arcangelis, L., and Godano, C. (2010). Time-energy correlations in solar flare occurrence. *A&A*, 511:L2.

Liu, Y. and Dahmen, K. A. (2009). Random-field Ising model in and out of equilibrium. *Europhys. Lett.*, 86(5):56003.

Lowen, S. B. and Teich, M. C. (2005). *Fractal-based point processes*, volume 366. John Wiley & Sons.

Lu, E. T. and Hamilton, R. J. (1991). Avalanches and the Distribution of Solar Flares. *Astrophys. J.*, 380:L89.

Maaß, R., Wraith, M., Uhl, J. T., Greer, J. R., and Dahmen, K. A. (2015). Slip statistics of dislocation avalanches under different loading modes. *Phys. Rev. E*, 91(4):042403.

Madariaga, R. (1978). The dynamic field of Haskell's rectangular dislocation fault model. *Bull. Seismol. Soc. Am.*, 68(4):869–887.

Madariaga, R. (1989). *Seismic source: Theory*, pages 1129–1133. Springer US, Boston, MA.

Main, I. (1996). Statistical physics, seismogenesis, and seismic hazard. *Reviews of Geophysics*, 34(4):433–462.

Main, I. (2013). Viewpoint: Little Earthquakes in the Lab. *Physics*, 6.

Majewski, J., Li, H., and Ott, J. (2001). The Ising model in physics and statistical genetics. *Am. J. Hum. Genet.*, 69(4):853–862.

Mäkinen, T., Miksic, A., Ovaska, M., and Alava, M. J. (2015). Avalanches in wood compression. *Phys. Rev. Lett.*, 115(5):055501.

Malevergne, Y., Pisarenko, V., and Sornette, D. (2011). Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities. *Phys. Rev. E*, 83:36111.

Mallman, E. P. and Zoback, M. D. (2007). Assessing elastic Coulomb stress transfer models using seismicity rates in Southern California and southwestern Japan. *J. Geophys. Res.: Solid Earth*, 112(B3).

Marro, J. and Dickman, R. (2005). *Nonequilibrium Phase Transitions in Lattice Models*. Cambridge University Press.

Marsan, D. (2003). Triggering of seismicity at short timescales following Californian earthquakes. *J. Geophys Res.*, 108:1–14.

Maydeu-Olivares, A. and García-Forero, C. (2010). Goodness-of-Fit Testing. *International Encyclopedia of Education (Third Edition)*, pages 190–196.

McCloskey, J., Nalbant, S. S., Steacy, S., Nostro, C., Scotti, O., and Baumont, D. (2003). Structural constraints on the spatial distribution of aftershocks. *Geophys. Res. Lett.*, 30(12):1610–1613.

Mead, R. (2017). *Statistical methods in agriculture and experimental biology*. Chapman and Hall/CRC.

Menyhárd, N. (1970). Range of validity of the landau theory in the $h \neq 0$ case. *Solid State Commun.*, 8(17):1337 – 1339.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092.

Microsoft Corporation (2018). *Microsoft Excel*.

Mistras, H. G. (2004). *PCI-2 – PCI-Based Two-Channel AE Board & System, by Physical Acoustics*.

Mogi, K. (1967). Earthquakes and fractures. *Tectonophysics*, 5(1):35 – 55.

Molins, J. and Vives, E. (2016). Model risk on credit risk. *Risk and Decision Analysis*, 6(1):65–78.

Mulargia, F. and Bizzarri, A. (2016). Earthquake friction. *Physics of the Earth and Planetary Interiors*, 261:118 – 123. Microseismicity from all scales.

Murtaugh, P. A. (2014). In defense of p values. *Ecology*, 95(3):611–617.

Nandan, S., Ouillon, G., Woessner, J., Sornette, D., and Wiemer, S. (2016). Systematic assessment of the static stress triggering hypothesis using interearthquake time statistics. *J. Geophys. Res.: Solid Earth*, 121:1890–1909.

Narayan, O. (1996). Self-similar Barkhausen noise in magnetic domain wall motion. *Phys. Rev. Lett.*, 77:3855–3857.

Narayan, O. and Fisher, D. S. (1993). Threshold critical dynamics of driven interfaces in random media. *Phys. Rev. B*, 48(10):7030.

Narteau, C., Byrdina, S., Shebalin, P., and Schorlemmer, D. (2009). Common dependence on stress for the two fundamental laws of statistical seismology. *Nature*, 462:642–645.

Nataf, G. F., Castillo-Villa, P. O., Baró, J., Illa, X., Vives, E., Planes, A., and Salje, E. K. H. (2014). Avalanches in compressed porous SiO_2-based materials. *Phys. Rev. E*, 90(2):022405.

Navas-Portella, V., Corral, A., and E., V. (2016). Avalanches and force drops in displacement-driven compression of porous glasses. *Phys. Rev. E*, 94(3):033005.

Navas-Portella, V., González, Á., Serra, I., Vives, E., and Corral, A. (2019). Universality of power-law exponents by means of maximum-likelihood estimation. *Phys. Rev. E*, 100(6):062106.

Navas-Portella, V., Jiménez, A., and Corral, Á. (2020). No Significant Effect of Coulomb Stress on the Gutenberg-Richter Law after the Landers Earthquake. *Sci. Rep.*, 10(2901):1–13.

Navas-Portella, V., Serra, I., Corral, A., and Vives, E. (2018). Increasing power-law range in avalanche amplitude and energy distributions. *Phys. Rev. E*, 97:022134.

Navas-Portella, V. and Vives, E. (2016). Influence of the aspect ratio and boundary conditions on universal finite-size scaling functions in the athermal metastable two-dimensional random field Ising model. *Phys. Rev. E*, 93:022129.

Neyman, J., Pearson, E. S., and Pearson, K. (1933). Ix. on the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231(694-706):289–337.

Nitram (2020). Nitram Académie Fusains HB.

Ódor, G. (2004). Universality classes in nonequilibrium lattice systems. *Rev. Mod. Phys.*, 76(3):663.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *JASA*, 83:9–27.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2):379–402.

Okada, Y. (1992). Internal deformation due to shear and tensile faults in a half-space. 82(2):1018–1040.

Olami, Z., Feder, H. J. S., and Christensen, K. (1992). Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Phys. Rev. Lett.*, 68(8):1244.

Olsen, K. B., Madariaga, R., and Archuleta, R. J. (1997). Three-dimensional dynamic simulation of the 1992 landers earthquake. *Science*, 278(5339):834–838.

Omori, F. (1894). On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo*, 7:111–200.

O'Neill, M. and Summers, E. (2015). Collins English dictionary.

Pacheco, J. E., Scholz, C. H., and Sykes, L. R. (2012). Changes in frequency–size relationship from small to large earthquakes. *Nature*, 355:71–73.

Paczuski, M., Boettcher, S., and Baiesi, M. (2005). Interoccurrence times in the bak-tang-wiesenfeld sandpile model: A comparison with the observed statistics of solar flares. *Phys. Rev. Lett.*, 95(18):181102.

Papanikolaou, S., Bohn, F., Sommer, R. L., Durin, G., Zapperi, S., and Sethna, J. P. (2011). Universality beyond power laws and the average avalanche shape. *Nat. Phys.*, 7(4):316–320.

Park, E.-S., Shin, H.-S., Cheon, D.-S., and Jung, Y.-B. (2013). Management and concept of the monitoring system considering the characteristics of subsea tunnels. *J. Kor. Tunn. Und. Space Assoc.*, 15(5):523–536.

Parsons, T., Toda, S., Stein, R. S., Barka, A., and Dietrich, J. H. (2000). Heightened odds of large earthquakes near Istanbul: An interaction-based probability calculation. *Science*, 288:661–665.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979.

Pathria, R. K. and Beale, P. D. (2011). *Statistical Mechanics*. Elsevier Science.

Pawitan, Y. (2013). *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. OUP Oxford.

Pérez-Benitez, J. A., Capó-Sánchez, J., and Padovese, L. R. (2009). Simulation of the Barkhausen Noise using random field Ising model with long-range interaction. *Comput. Mater. Sci.*, 44(3):850–857.

Peierls, R. (1936). On Ising's model of ferromagnetism. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(3):477–481.

Pérez-Reche, F.-J., Tadić, B., Mañosa, L., Planes, A., and Vives, E. (2004). Driving rate effects in avalanche-mediated first-order phase transitions. *Phys. Rev. Lett.*, 93:195701.

Pérez-Reche, F.-J., Truskinovsky, L., and Zanzotto, G. (2008). Driving-Induced Crossover: From Classical Criticality to Self-Organized Criticality. *Phys. Rev. Lett.*, 101(23):230601.

Pérez-Reche, F. J., Vives, E., Mañosa, L., and Planes, A. (2001). Athermal character of structural phase transitions. *Phys. Rev. Lett.*, 87:195701.

Perković, O., Dahmen, K., and Sethna, J. P. (1995). Avalanches, Barkhausen noise, and plain old criticality. *Phys. Rev. Lett.*, 75(24):4528.

Peters, O., Deluca, A., Corral, A., Neelin, J. D., and Holloway, C. E. (2010). Universality of rain event size distributions. *J. Stat. Mech.: Theory Exp.*, 2010(11):P11030.

Petri, A., Paparo, G., Vespignani, A., Alippi, A., and M., C. (1994). Experimental evidence for critical dynamics in microfracturing processes. *Phys. Rev. Lett.*, 73(25):3423–3426.

Peyrat, S., Olsen, K., and Madariaga, R. (2001). Dynamic modeling of the 1992 landers earthquake. *J. Geophys. Res.: Solid Earth*, 106(B11):26467–26482.

Planes, A., Mañosa, L., and Vives, E. (2013). Acoustic emission in martensitic transformations. *J. Alloys Compd.*, 577:S699–S704.

Planet, R., Santucci, S., and Ortín, J. (2009). Avalanches and non-gaussian fluctuations of the global velocity of imbibition fronts. *Phys. Rev. Lett.*, 102:094502.

Plenkers, K., Schorlemmer, D., Kwiatek, G., and Group, J. R. (2011). On the probability of detecting picoseismicity. *Bull. Seismol. Soc. Am.*, 101(6):2579–2591.

Pruessner, G. (2012). *Self-organised criticality: theory, models and characterisation.* Cambridge University Press.

Quigley, M. C., Jiménez, A., Duffy, B., and King, T. R. (2019). Physical and statistical behavior of multifault earthquakes: Darfield earthquake case study, New Zealand. *J. Geophys. Res.: Solid Earth*, 124(5):4788–4810.

R Documentation (2017). *R: One Dimensional Optimization.*

Raischel, F., Kun, F., and Herrmann, H. J. (2005). Simple beam model for the shear failure of interfaces. *Phys. Rev. E*, 72:046126.

Reasenberg, P. A. and Jones, L. M. (1989). Earthquake hazard after a mainshock in California. *Science*, 243(4895):1173–1176.

Regev, O. (2006). Chaos and Complexity in Astrophysics. *Cambridge University Press.*

Ribeiro, H. V., Costa, L. S., Alves, L. G. A., Santoro, P. A., Picoli, S., Lenzi, E. K., and R.S., M. (2015). Analogies between the cracking noise of ethanol-dampened charcoal and earthquakes. *Phys. Rev. Lett.*, 115(2):025503.

Richeton, T., Weiss, J., and Louchet, F. (2005). Breakdown of avalanche critical behaviour in polycrystalline plasticity.

Rodney, D., Tanguy, A., and Vandembroucq, D. (2011). Modeling the mechanics of amorphous solids at different length scale and time scale. *Modelling and Simulation in Materials Science and Engineering*, 19(8):083001.

Roe, B. P. (2012). *Probability and statistics in experimental physics.* Springer Science & Business Media.

Rosinberg, M.-L. and Vives, E. (2004). Recent topics on metastability, hysteresis, avalanches, and acoustic emission associated to martensitic transitions in functional materials. Technical report.

Ross, S. M. (1998). *A First Course in Probability.* Prentice Hall, Upper Saddle River, N.J., fifth edition.

Rowlinson, J. S. (1969). Thomas Andrews and the Critical Point. *Nature*, 224(5219):541–543.

Rundle, J. B., Turcotte, D. L., Shcherbakov, R., Klein, W., and Sammis, C. (2003). Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev. Geophys.*, 41(4).

Sabadini R., V. B. (2016). *Global Dynamics of the Earth.* 2nd Edition, Springer (2016).

Sabhapandit, S., Shukla, P., and Dhar, D. (2000). Distribution of avalanche sizes in the hysteretic response of the random-field Ising model on a bethe lattice at zero temperature. *J. Stat. Phys.*, 98(1-2):103–129.

Salje, E. K. and Dahmen, K. A. (2014). Crackling noise in disordered materials. *Annu. Rev. Conden. Ma. P.*, 5(1):233–254.

Salje, E. K. H., Soto-Parra, D. E., Planes, A., E.Vives, M.Reinecker, and Schranz, W. (2011). Failure mechanism in porous materials under compression: crackling noise in mesoporous SiO_2. *Phil. Mag. Lett.*, 91:554–560.

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bull. Seismol. Soc. Am.*, 58(1):399–415.

Scholz, C. H. (2019). *The Mechanics of Earthquakes and Faulting*. Cambridge University Press, 3 edition.

Schorlemmer, D., Wiemer, S., and Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, 437:539–542.

Schorlemmer, D. and Woessner, J. (2008). Probability of detecting an earthquakeprobability of detecting an earthquake. *Bull. Seismol. Soc. Am.*, 98(2):2103–2117.

Schubert, G. (2015). *Treatise on Geophysics*. Elsevier Science.

Schwab, D. J., Nemenman, I., and Mehta, P. (2014). Zipf's Law and Criticality in Multivariate Data without Fine-Tuning. *Phys. Rev. Lett.*, 113(6):068102.

Scruby, C. B. (1987). An introduction to acoustic emission. *J. Phys. E: Sci. Instrum.*, 20(8):946–953.

Seeber, L. and Armbruster, J. G. (2000). Earthquakes as beacons of stress change. *Nature*, 407:69–72.

Segall, P., Llenos, A. L., Yun, S.-H., Bradley, A. M., and Syracuse, E. M. (2013). Time-dependent dike propagation from joint inversion of seismicity and deformation data. *J. Geophys. Res.: Solid Earth*, 118(11):5785–5804.

Serra, I. and Corral, A. (2017). Deviation from power law of the global seismic moment distribution. *Sci. Rep.*, 7:40045.

Sethna, J. (2006). *Statistical mechanics: entropy, order parameters, and complexity*, volume 14. Oxford University Press.

Sethna, J., Dahmen, K., and Perkovic, O. (2006). *Random-field Ising models of hysteresis*, volume 2.

Sethna, J. P., Dahmen, K., Kartha, S., Krumhansl, J. A., Roberts, B. W., and Shore, J. D. (1993). Hysteresis and hierarchies: Dynamics of disorder-driven first-order phase transformations. *Phys. Rev. Lett.*, 70:3347–3350.

Sethna, J. P., Dahmen, K. A., and Myers, C. R. (2001). Crackling noise. *Nature*, 410:242–250.

Shcherbakov, R., Turcotte, D. L., and Rundle, J. B. (2005). Aftershock statistics. *Pure appl. geophys.*, 162(6):1051–1076.

Sherrington, D. (2010). Physics and complexity. *Philos. T. R. Soc. A*, 368(1914):1175–1189.

Shirzaei, M., Ellsworth, W. L., Tiampo, K. F., González, P. J., and Manga, M. (2016). Surface uplift and time-dependent seismic hazard due to fluid injection in Eastern Texas. *Science*, 353(6306):1416–1419.

Shoji, Y., Tanii, K., and Kamiyama, M. (2005). A study on the duration and amplitude characteristics of earthquake ground motions. *Soil Dyn. Earthquake Eng.*, 25(7):505–512.

Shukla, P. and Thongjaomayum, D. (2017). Criteria for infinite avalanches in the zero-temperature nonequilibrium random-field Ising model on a bethe lattice. *Phys. Rev. E*, 95:042109.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *The annals of mathematical statistics*, 19(2):279–281.

Smith, D. E. (2006). *A new paradigm for interpreting stress inversions from focal mechanisms: how 3D stress heterogeneity biases the inversions toward the stress rate*. PhD thesis, California Institute of Technology.

Sornette, D. (1994). Sweeping of an instability: an alternative to self-organized criticality to get powerlaws without parameter tuning. *Journal de Physique I*, 4(2):209–221.

Sornette, D. (2009). Why Stock Markets Crash: Critical Events in Complex Financial Systems. *Princeton University Press*.

Sornette, D. and Werner, M. J. (2011). Seismicity, statistical physics approaches to. *Extreme Environmental Events: Complexity in Forecasting and Early Warning*, pages 825–843.

Spasojević, D., Bukvić, S., Milošević, S., and Stanley, H. E. (1996). Barkhausen noise: Elementary signals, power laws, and scaling relations. *Phys. Rev. E*, 54:2531–2546.

Spasojević, D., Janićević, S., and Knežević, M. (2011a). Avalanche distributions in the two-dimensional nonequilibrium zero-temperature random field Ising model. *Phys. Rev. E*, 84:051119.

Spasojević, D., Janićević, S., and Knežević, M. (2011b). Numerical evidence for critical behavior of the two-dimensional nonequilibrium zero-temperature random field Ising model. *Phys. Rev. Lett.*, 106:175701.

Spasojević, D., Mijatović, S., Navas-Portella, V., and Vives, E. (2018). Crossover from three-dimensional to two-dimensional systems in the nonequilibrium zero-temperature random-field Ising model. *Phys. Rev. E*, 97:012109.

Spotila, J. A. and Sieh, K. (1995). Geologic investigations of a "slip gap" in the surficial ruptures of the 1992 landers earthquake, Southern California. *J. Geophys. Res.: Solid Earth*, 100(B1):543–559.

Stanley, H., Amaral, L., Gopikrishnan, P., Ivanov, P., Keitt, T., and Plerou, V. (2000). Scale invariance and universality: organizing principles in complex systems. *Physica A*, 281(1):60 – 68.

Stanley, H. E. (1999). Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Rev. Mod. Phys.*, 71:S358–S366.

Stauffer, D. (2008). Social applications of two-dimensional Ising models. *Am. J. Phys.*, 76(4):470.

Steacy, S., Gomberg, J., and Cocco, M. (2005). Introduction to special section: stress transfer, earthquake triggering, and time-dependent seismic hazard. *J. Geophys. Res.*, 110:B05S01.

Steacy, S., Marsan, D., Nalbant, S. S., and McCloskey, J. (2004). Sensitivity of static stress calculations to the earthquake slip distribution. *J. Geophys. Res.: Solid Earth*, 109(B4).

Stehly, L., Campillo, M., Froment, B., and Weaver, R. L. (2008). Reconstructing Green's function by correlation of the coda of the correlation (C3) of ambient seismic noise. *J. Geophys. Res. Solid Earth*, 113(B11).

Stein, R., King, G., and Lin, J. (1992). Change in failure stress on the Southern San Andreas fault system caused by the 1992 magnitude=7.4 Landers earthquake. *Science*, 258:1328–1332.

Stein, S. and Wysession, M. (2003). *An Introduction to Seismology, Earthquakes, and Earth Structure*. Blackwell Pub.

Steven, H. S. (1994). *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press.

Stork, A. L. and Verdon, J. P. (2014). The robustness of seismic moment and magnitudes estimated using spectral analysis. *Geophysical Prospecting*, 62(4):862–878.

Strader, A. and Jackson, D. D. (2014). Near-prospective test of coulomb stress triggering. *J. Geophys. Res.: Solid Earth*, 119(4):3064–3075.

Sun, B. A., Yu, H. B., Jiao, W., Bai, H. Y., Zhao, D. Q., and Wang, W. H. (2010). Plasticity of Ductile Metallic Glasses: A Self-Organized Critical State. *Phys. Rev. Lett.*, 105(3):035501.

Thornton, S. T. and Marion, J. B. (2004). *Classical Dynamics of Particles and Systems*. Brooks/Cole.

Toda, S., Stein, R., Richards-Dinger, K., and Bozkurt, S. (2005). Forecasting the evolution of seismicity in Southern California: Animations built on earthquake stress transfer. *J. Geophys. Res.*, 110:B05S16.

Touati, S., Naylor, M., and Main, I. G. (2009). Origin and nonuniversality of the earthquake interevent time distribution. *Phys. Rev. Lett.*, 102:168501.

Tsekenis, G., Uhl, J. T., Goldenfeld, N., and Dahmen, K. A. (2013). Determination of the universality class of crystal plasticity. *Europhys. Lett.*, 101(3):36003.

Turcotte, D. L. (1997). Fractals and Chaos in Geology and Geophysics. *Cambridge Core*.

Uhl, J. T., Pathak, S., Schorlemmer, D., Liu, X., Swindeman, R., Brinkman, B. A. W., LeBlanc, M., Tsekenis, G., Friedman, N., Behringer, R., Denisov, D., Schall, P., Gu, X., Wright, W. J., Hufnagel, T., Jennings, A., Greer, J. R., Liaw, P. K., Becker, T., Dresen, G., and Dahmen, K. A. (2015). Universal Quake Statistics: From Compressed Nanocrystals to Earthquakes. *Sci. Rep.*, 5(16493):1–10.

Urbach, J. S., Madison, R. C., and Markert, J. T. (1995). Interface depinning, self-organized criticality, and the Barkhausen effect. *Phys. Rev. Lett.*, 75:276–279.

Utsu, T. (1961). A statistical study of the occurrence of aftershocks. *Geophysical Magazine*, 30:521–605.

Utsu, T. (1999). Representation and Analysis of the Earthquake Size Distribution: A Historical Review and Some New Approaches. *Pure Appl. Geophys.*, 155(2-4):509–535.

Utsu, T., Ogata, Y., S, R., and Matsu'ura (1995). The centenary of the omori formula for a decay law of aftershock activity. *J. Phys. Earth*, 43(1):1–33.

Van der Laan, M. J. and Rose, S. (2011). *Targeted learning: causal inference for observational and experimental data.* Springer Science & Business Media.

Van Herwijnen, A., Heck, M., and Schweizer, J. (2016). Forecasting snow avalanches using avalanche activity data obtained through seismic monitoring. *Cold Regions Science and Technology*, 132:68–80.

Vassiliou, M. and Kanamori, H. (1982). The energy release in earthquakes. *Bull. Seismol. Soc. Am.*, 72(2):371–387.

Vavryčuk, V. (2015). *Earthquake Mechanisms and Stress Field.* Springer, Berlin, Heidelberg.

Villegas, P., di Santo, S., Burioni, R., and Muñoz, M. A. (2019). Time-series thresholding and the definition of avalanche size. *Phys. Rev. E*, 100:012133.

Vives, E., Ortín, J., Mañosa, L., Ràfols, I., Pérez-Magrané, R., and A., P. (1994). Distributions of avalanches in martensitic transformations. *Phys. Rev. Lett.*, 72(11):1694–1697.

Vives, E., Ràfols, I., Mañosa, L., Ortín, J., and Planes, A. (1995). Statistics of avalanches in martensitic transformations. i. acoustic emission experiments. *Phys. Rev. B*, 52(17):12644.

Vives, E., Soto-Parra, D., Mañosa, L., Romero, R., and Planes, A. (2009). Driving-induced crossover in the avalanche criticality of martensitic transitions. *Phys. Rev. B*, 80(18):180101.

Von Neumann, J., Burks, A. W., et al. (1966). Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, 5(1):3–14.

W. H. Press, S. A. Teukolsky, W. T. V. and Flannery., B. P. (1992). Numerical recipes in fortran.

Wald, D. J. and Heaton, T. H. (1994). Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bull. Seismol. Soc. Am.*, 84(3):668–691.

Wang, J.-H. (2015). The energy-magnitude scaling law for $m \leq 5.5$ earthquakes. *J. Seismol.*, 19(2):647 – 652.

Wang, W.-H., Dong, C., and Shek, C. (2004). Bulk metallic glasses. *Materials Science and Engineering: R: Reports*, 44(2-3):45–89.

Weidlich, W. (2006). *Sociodynamics: A systematic approach to mathematical modelling in the social sciences*. Courier Corporation.

Weiss, J., Lahaie, F., and Grasso, J. R. (2000). Statistical analysis of dislocation dynamics during viscoplastic deformation from acoustic emission. *J. Geophys. Res. Solid Earth*, 105(B1):433–442.

Weiss, J. and Miguel, M. C. (2004). Dislocation avalanche correlations. *Materials Science and Engineering: A*, 387-389:292 – 296.

Weiss, J., Richeton, T., Louchet, F., Chmelik, F., Dobron, P., Entemeyer, D., Lebyodkin, M., Lebedkina, T., Fressengeas, C., and McDonald, R. J. (2007). Evidence for universal intermittent crystal plasticity from acoustic emission and high-resolution extensometry experiments. *Phys. Rev. B*, 76(22):224110.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W., and Tian, D. (2019). The generic mapping tools version 6. *Geochem. Geophy. Geosy.*, 20.

White, R. A. and Dahmen, K. A. (2003). Driving rate effects on crackling noise. *Phys. Rev. Lett.*, 91:085702.

Wielandt, E. (2002). Seismic sensors and their calibration. *New Manual of Seismological Observatory Practices*, 1(46).

Wilson, K. G. (1971). Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. *Phys. Rev. B*, 4:3174–3183.

Wilson, K. G. (1975). The renormalization group: Critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47:773–840.

Wilson, K. G. (1983). The renormalization group and critical phenomena. *Rev. Mod. Phys.*, 55:583–600.

Wolfram, S. (1983). Statistical mechanics of cellular automata. *Rev. Mod. Phys.*, 55:601–644.

Wood, S. N. (2015). *Core Statistics*. Institute of Mathematical Statistics Textbooks. Cambridge University Press.

Wu, W. and Adams, P. W. (1995). Avalanches and slow relaxation: Dynamics of ultrathin granular superconducting films in a parallel magnetic field. *Phys. Rev. Lett.*, 74:610–613.

Wyss, M. and Wiemer, S. (2000). Change in the probability for earthquakes in Southern California due to the landers magnitude 7.3 earthquake. *Science*, 290(5495):1334–1338.

Xu, Y., Borrego, A. G., Planes, A., Ding, X., and Vives, E. (2019). Criticality in failure under compression: Acoustic emission study of coal and charcoal with different microstructures. *Phys. Rev. E*, 99:033001.

Yaghoubi, M., de Graaf, T., Orlandi, J. G., Girotto, F., Colicos, M. A., and Davidsen, J. (2018). Neuronal avalanche dynamics indicates different universality classes in neuronal cultures. *Sci. Rep.*, 8(1):3417.

Yang, W., Hauksson, E., and Shearer, P. M. (2012). Computing a large refined catalog of focal mechanisms for Southern California (1981-2010): Temporal stability of the style of faulting. *Bull. Seismol. Soc. Am.*, 102:1179–1194.

Yeats, R., Yeats, R., Yeats, R., Sieh, K., Sieh, K., Allen, C., and Allen, P. (1997). *The Geology of Earthquakes*. Oxford University Press.

Yoder, M. R., Holliday, J. R., Turcotte, D. L., and Rundle, J. B. (2012). A geometric frequency–magnitude scaling transition: Measuring b=1.5 for large earthquakes. *Tectonophysics*, 532-535:167 – 174.

Yu, Z.-z. and Clapp, P. C. (1987). Quantitative analysis of acoustic emission signals. *J. Appl. Phys.*, 62(6):2212–2220.

Zaliapin, I. and Ben-Zion, Y. (2013). Earthquake clusters in Southern California i: Identification and stability. *J. Geophys. Res.: Solid Earth*, 118(6):2847–2864.

Zapperi, S., Cizeau, P., Durin, G., and Stanley, H. E. (1998). Dynamics of a ferromagnetic domain wall: Avalanches, depinning transition, and the Barkhausen effect. *Phys. Rev. B*, 58:6353–6366.

Zapperi, S., Vespignani, A., and Stanley, H. E. (1997). Plasticity and avalanche behaviour in microfracturing phenomena. *Nature*, 388(6643):658–660.

Zemansky, M. and Dittman, R. (1981). *Heat and Thermodynamics: An Intermediate Textbook*. McGraw-Hill international editions. Physics series. McGraw-Hill.

Zollo, A., Orefice, A., and Convertito, V. (2014). Source parameter scaling and radiation efficiency of microearthquakes along the irpinia fault zone in Southern Apennines, Italy. *J. Geophys. Res.: Solid Earth*, 119(4):3256–3275.

Zreihan, N., Faran, E., Vives, E., Planes, A., and Shilo, D. (2018). Coexistence of a well-determined kinetic law and a scale-invariant power law during the same physical process. *Phys. Rev. B*, 97(1):014103.

Zwick-Roell (2019). *testXpert II Testing Software–Intelligent and Reliable*. https://www.zwickroell.com/en/testing-software/testxpert-ii.